UNDERSTANDING INTERACTION MECHANICS IN TOUCHLESS TARGET SELECTION

Debaleena Chattopadhyay

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the School of Informatics and Computing
Indiana University

September 2016

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Davide Bolchini, Ph.D., Chair

Karl F. MacDorman, Ph.D.

Doctoral Committee

July 28, 2016

Stephen Voida, Ph.D.

Erik Stolterman, Ph.D.

© 2016
Debaleena Chattopadhyay

Dedication

To my parents

Acknowledgements

This dissertation was made possible by the perseverance, guidance, and encouragement of my mentors. In what follows, I thank those who were crucial to my doctoral training. However, a doctorate is a terminal degree, culminating one's education in a particular field of study. So I first take this opportunity to thank all my teachers who came before and equipped me to undertake this training effectively.

I would like to thank my advisor, Davide, for his unwavering support toward my research career. He continues to invigorate my half-baked ideas—even when I am half-heartedly pursuing them. As goes the adage, recognizing a good idea is as important as having a good idea—if not more. He gave me the utmost freedom to pursue my research goals while investing countless hours in training me to become a skilled researcher. He believed in my vision when I did not have the skills to undertake my research and helped me in acquiring those skills. Enumerating Davide's role in my doctoral training would befit a longer story and this acknowledgment, by no means, measures up to that. Nevertheless, Davide is a great teacher, both in and out of the classroom—and among those rare ones, who can inspire greatness. I doubt if a doctoral student could ask for an advisor with any more positivity.

In 2012, I had earned an A+ in my graduate research design course; the course instructor's feedback read: "Your significance section is inadequate. Significance means, if you are lucky, and your experiment goes as planned, and you collect the results anticipated and publish them, how does *that* change the world?" Earlier in that course, I had got my first B in a class assignment—an annotated bibliography. In research (or otherwise), Karl never settles for anything less than perfect; his *obsession* with high standards has significantly contributed to the groundwork of my doctoral training. Karl continues to be my touchstone for excellence in research. I would like to thank him for all the different hats he had put on during my Ph.D. life—a teacher, collaborator, mentor, dissertation committee member, and above all, a ruthless critic. Without his unnerving demand for excellence, I wouldn't be the researcher that I am today. Thank you for everything, and a day.

Sincere thanks go out to my committee members, Stephen Voida and Erik Stolterman, who advised this dissertation from proposal to defense—ensuring my plans are feasible and diligently curbing my underestimation of time and effort required for a successful completion. Steve's course on ubiquitous computing—that I took in the Spring of 2014—was instrumental in bringing the different pieces of my research together, under the umbrella of embodied interaction. I owe Steve another round of thanks for introducing me to Heidegger, and to Karl (rather, some of his old papers that he doesn't cite any more) for introducing me to Gibson's affordances.

A doctorate may be a terminal degree, but it also resides at the lowest rung of academia or a career in research. Indeed, a doctorate is at best the license to conduct independent research. My dreams of furthering that coveted research career continue to be supported by mentors outside the school, and Kenton, with whom I spent the summer of 2015 at Microsoft Research Cambridge (MSRC), deserves special mention. I would like to thank him for giving me an eye for simplicity and a taste of how sociotechnical systems can (and should) dissolve into the fabric of a familiar social milieu. My internship at the Human Experience and Design (HxD) lab at MSRC will always be a cherished memory—professionally and personally.

I am grateful for receiving research support during my Ph.D. from Indiana University-Purdue University Indianapolis (IUPUI) Graduate Office, IUPUI Office of the Vice Chancellor for Research, IUPUI School of Informatics and Computing, Xerox Foundation, Microsoft, Association for Computing Machinery (ACM), and National Science Foundation (NSF). I would like to thank Chauncey Frend, Jeff Rogers, and Michael Boyles of the Advanced Visualization Lab (AVL) for their generous support toward conducting my research studies. Many thanks go out to Elizabeth Bunge and Nancy Barker for their administrative assistance throughout my doctoral studies.

Doctoral research (or any research) is a poignantly lonesome pursuit. I would like to thank my friends and family, who put up with my perennial absence—yet being there at times I needed them. Particularly, heartfelt thanks go out to Manisha Deogharia (for listening to my computing research problems amid her HeLa cell cultures) and Raquel Perales (for always being open to trading her doctorate in Mathematics with mine).

Finally, my *most* sincere gratitude to my parents: During their college lives, my father wanted to inspire an auditorium full of students, and my mother wanted to *live* in a lab around pipettes and burners; I owe my love of knowledge to them.

Debaleena Chattopadhyay

UNDERSTANDING INTERACTION MECHANICS IN TOUCHLESS TARGET SELECTION

We use gestures frequently in daily life—to interact with people, pets, or objects. But interacting with computers using mid-air gestures continues to challenge the design of touchless systems. Traditional approaches to touchless interaction focus on exploring gesture inputs and evaluating user interfaces. I shift the focus from gesture elicitation and interface evaluation to touchless interaction mechanics.

I argue for a novel approach to generate design guidelines for touchless systems: to use fundamental interaction principles, instead of a reactive adaptation to the sensing technology. In five sets of experiments, I explore visual and pseudo-haptic feedback, motor intuitiveness, handedness, and perceptual Gestalt effects. Particularly, I study the interaction mechanics in touchless target selection. To that end, I introduce two novel interaction techniques: touchless circular menus that allow command selection using directional strokes and interface topographies that use pseudo-haptic feedback to guide steering–targeting tasks.

Results illuminate different facets of touchless interaction mechanics. For example, motor-intuitive touchless interactions explain how our sensorimotor abilities inform touchless interface affordances: we often make a holistic oblique gesture instead of several orthogonal hand gestures while reaching toward a distant display. Following the Gestalt theory of visual perception, we found similarity between user interface (UI) components decreased user accuracy while good continuity made users faster. Other findings include hemispheric asymmetry affecting transfer of training between dominant and nondominant hands and pseudo-haptic feedback improving touchless accuracy.

The results of this dissertation contribute design guidelines for future touchless systems. Practical applications of this work include the use of touchless interaction techniques in various domains, such as entertainment, consumer appliances, surgery, patient-centric health settings, smart cities, interactive visualization, and collaboration.

Davide Bolchini, Ph.D., Chair

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

Interacting with computers besides using mouse and keyboard has been a significant leap for human-computer interaction (Jacob et al., 2008). Over the last decade, with smartphones, tablets, and tabletops, interactive flat surfaces and touch interactions have pervaded our everyday life. With the current boom in sensing technologies, interactive computing has leaped further—from surfaces to spaces, from touch to *touchless* (Wigdor & Wixon, 2011).

Touchless is an interaction modality—enabling users to interact with computers using mid-air gestures, either with a bare hand (Bailly, Walter, Müller, Ning, & Lecolinet, 2011) or while wearing specialized hand gloves (Ni, McMahan, & Bowman, 2008). Unlike mouse, pen, keyboard, or touch, touchless gestures permit users to interact from a distance and untethered from a surface (Hespanhol, Tomitsch, Grace, Collins, & Kay, 2012). Touchless affords fluidity in physical navigation, along with the absence of an intermediate input device. Touchless interfaces are often deemed as natural user interfaces (NUIs). NUIs promise to offer an intuitive interface that does not require developing special skills for interacting with computers but allows people to use their natural abilities (Macaranas, Antle, & Riecke, 2015). For example, ad slogans, such as 'you are the controller' attained great popularity among consumers when Microsoft launched Kinect™ in 2010 (Nansen et al., 2014). Touchless interactions promise to turn our everyday gestures into meaningful commands to operate computer systems—from laptops to smart televisions to microwaves to large displays (Garzotto & Valoriani, 2012; Guimbretière & Nguyen, 2012; Vatavu & Zaiti, 2014).

Labeling touchless as *natural* raises a crucial question: What is natural (or intuitive or like real-world) for users? The emergence of NUIs has spurred interest in critically examining the concept of natural or intuitive, fueling many ongoing debates (Aigner et al., 2012; Grandhi, Joue, & Mittelberg, 2011; Hansen & Dalsgaard, 2015; Hespanhol et al., 2012; Lee, 2010; Malizia & Bellucci, 2012; Morris, 2012; O'Hara, Harper, Mentis, Sellen, & Taylor, 2013; Vatavu & Zaiti, 2014; Wigdor & Wixon, 2011). This dissertation sidesteps from generically labeling touchless as natural; instead, I explore the core mechanics of touchless interaction, such as feedback, affordances, abilities, or handedness. The premise here is that the naturalness of an interface is not an axiomatic truth, but achieved through sufficient feedback, effective feedforward, and perceived affordances (Norman, 2010; Wigdor, 2010).

Figure 1.1. Unlike mouse, pen, or touch, touchless interaction is device-less—enabling people to interact with computers with bare hands and facilitating physical navigation.

I study touchless interaction from an embodied cognition perspective—drawing on different theories of cognitive science and motor behavior. The embodied cognition perspective argues that our perceptions and motor actions depend on how the body experiences the world through our various sensorimotor abilities (Dourish, 2004; Kirsh, 2013). Using this theoretical lens, I deconstruct the sensorimotor relations in touchless—explore how interface affordances and ability play a role in the intuitiveness of touchless interactions and use theories of visual perception and motor action to inform the design of touchless interfaces.

Particularly, this dissertation focusses on the device-less property of touchless (Figure 1.1). Different touchless interfaces use different kind of sensing technologies, ranging from infra-red (IR) body markers (Zhou & Hu, 2008), IR-enabled handheld remote controllers (Kamuro, Minamizawa, Kawakami, & Tachi, 2009), hand gloves (Ni et al., 2008) to depth-based, markerless sensing (Bailly et al., 2011, Figure 1.2). Marker-based technologies—where individuals wear a set of IR markers on their bodies—are commonly used to study motion-tracking, but are intrusive and cumbersome for interacting with systems (Zhou & Hu, 2008). Infra-red handhelds enable interacting with computers using mid-air gestures but involve an input device (Kamuro et al., 2009). It is the introduction of markerless sensing of whole-body movements—without any intermediate device—that propelled the emerging research on touchless interfaces in a variety of domains, such as entertainment (Morris, 2012; Nebeling, Huber, Ott, & Norrie, 2014; Rovelo Ruiz, Vanacken, Luyten, Abad, & Camahort, 2014; Vatavu & Zaiti, 2014), surgery (Mentis, O'Hara, Sellen, & Trivedi, 2012; O'Hara et al., 2014; Ruppert et al., 2012; Schwarz, Bigdelou, & Navab, 2011), patient-centric health settings (Dsouza et al., 2014; Johnson, O'Hara, Sellen, Cousins, & Criminisi, 2011; Morrison et al., 2016;

Mullaney, Yttergren, & Stolterman, 2014; Rosa & Elizondo, 2014; Tan, Chao, Zawaideh, Roberts, & Kinney, 2013), interactive visualization (Dostal, Hinrichs, Kristensson, & Quigley, 2014; Kister, Reipschläger, Matulic, & Dachselt, 2015), and collaboration (Bragdon, DeLine, Hinckley, & Morris, 2011).

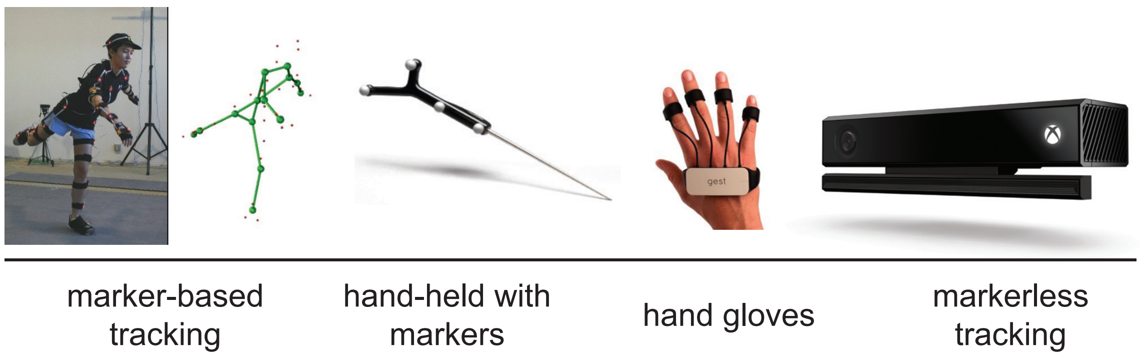

Figure 1.2. Touchless interactions use different kinds of sensing technologies, ranging from infra-red (IR) body markers to depth-based, markerless sensing.

Most recently, the naturalness of device-less touchless interaction was studied from an interactional perspective, focusing on Merleau-Ponty's lived-body view of individual experiences and Wittgenstein's socially organized view of the action (O'Hara et al., 2013). This dissertation looks into device-less touchless (hereafter touchless) from a different perspective: the implications of no embodied conception of a tool—no transition of an input device from *present-at-hand*, an object of activity, to *ready-to-hand*, absorbed in the fabric of the activity (Dourish, 2004; Heidegger, 1988). I argue that this device-less property of touchless creates unique interaction mechanics, different from the mouse, keyboard, pen, or touch.

Although touchless may involve interaction mechanics different than other more traditional modalities, such as touch or pen, touchless input remains strikingly similar to our everyday use of mid-air gestures. This similarity is the focus of some current approaches toward designing touchless interaction techniques—a method called gesture elicitation. Gesture elicitation aims to design intuitive interfaces by involving users in the process (Wobbrock, Morris, & Wilson, 2009). Gesture vocabularies are identified by typically showing the outcome of user interface actions or commands, and asking individual users to propose gestures that would trigger those actions. By the end of the process, a set of interaction commands emerges (Aigner et al., 2012; Grandhi et al., 2011; Morris, 2012; Nebeling et al., 2014; Vatavu & Zaiti, 2014; Vatavu & Wobbrock, 2015). Another approach to designing touchless systems is expert design: proposing new or emulating successful interaction techniques from other interaction modalities, such as pen or mouse, and then iterating and evaluating them with users (Bailly et al., 2011; Guimbretière & Nguyen, 2012; Hespanhol et al., 2012; Ni et al., 2008). In this dissertation, I shift the focus from gesture elicitation and interface evaluation to touchless interaction mechanics, present empirical results, and lay out several design and research implications.

Touchless techniques have been explored in various setups, from small to large interactive surfaces, from near to far-away interactions (Garzotto & Valoriani, 2012; O'Hara et al., 2013). This dissertation explores touchless interactions with distant, two-dimensional (2D), large displays (Figure 1.3). Touchless becomes relevant for interacting with large, distant displays when an interaction device is not at hand (e.g., in public spaces), when touching a device is not acceptable (e.g., in a sterile environment), or during sporadic browsing of multimedia information (e.g., in interactive TVs). Though in some of these scenarios users can use hand-held devices, such as smartphones or tablets, device-free interaction relieves users from the burden of searching, learning, connecting, and attending to an additional "medium" between the user and the display.

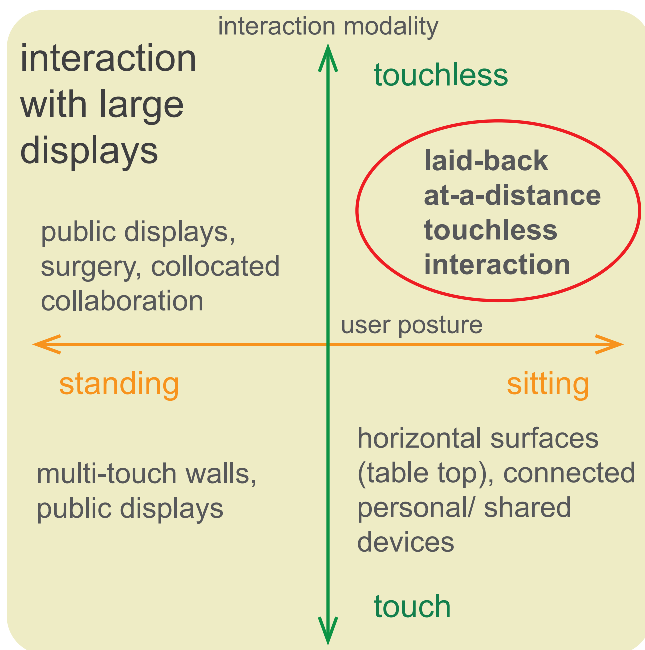

Figure 1.3. Touchless interaction with large displays while sitting at a distance and engaged in laid-back or high-bandwidth, sporadic tasks (Chattopadhyay & Bolchini, 2013).

However, it is important to note that touchless may not be suitable for all types of large-display interactions, primarily because of the lack of precision inherent in the interaction modality (Nancel, Wagner, Pietriga, Chapuis, & Mackay, 2011; Norman, 2010). Although researchers are exploring mid-air interaction techniques for fine-grained tasks such as text-entry (Markussen, Jakobsen, & Hornbæk, 2014; Sridhar, Feit, Theobalt, & Oulasvirta, 2015), it is highly unlikely that touchless would replace keyboard or pen for such precise interactions, such as typing or drawing illustrations (Fung, Lank, Terry, & Latulipe, 2008). Touchless is more suited for sporadic, high-bandwidth tasks, such as pointing-and-selecting, opening, moving, or lightly annotating (Beaudouin-Lafon et al., 2012; Nancel et al., 2011). To that end, I study *touchless target selection* in distant, large displays—a fundamental piece of interaction for any touchless interface.

Current research on touchless target selection techniques follows either of the two prevalent design approaches—gesture elicitation or expert design. Elicitation studies aim to understand user preference in touchless input gestures in different interaction contexts, such as in a living room with a large, flat screen television (Morris, 2012; Vatavu & Zaiti, 2014) or multiple collocated users viewing omnidirectional videos (Rovelo Ruiz et al., 2014). In the expert design approach, target selection techniques are introduced and evaluated. For example, pushing or dwelling (Hespanhol et al., 2012), making three-dimensional strokes (Guimbretière & Nguyen, 2012), posing a certain combination of fingers (Bailly et al., 2011; Kulshreshth & LaViola Jr, 2014), rolling the wrist and pinching (Ni et al., 2008) or crossing a delimiter (Ren & O'Neill, 2012) to select a target. Both elicitation and expert design approaches seek intuitive touchless techniques, but user studies have found certain interactions—that were earlier described as suitable or are effectively supported by the system—turn out difficult to perform during evaluation (Nebeling et al., 2014; Ren & O'Neill, 2012). For example, target selection by moving an open palm normal to the display (push-to-select) caused frequent false positives and false negatives while interacting with large displays (Hespanhol et al., 2012). When interacting with touchless marking menus on a distant display, researchers reported that most users had difficulties constraining their gestures in a two-dimensional (2D) plane (Bailly et al., 2011). Recent research on 3D marking menus also reported users' limitations in making precise hand trajectories in 3D space (Guimbretière & Nguyen, 2012; Ren & O'Neill, 2012). We often encounter such observations from evaluation studies about user limitations or failure of certain touchless gestures without any proper explanation. The current approaches are either treating human abilities as a 'black box', assuming that our ability to interact with the physical world directly translates into our ability to perform exact gestures in space or simply reacting to technological capabilities. I argue that the problem herein is twofold: neither of the existing approaches operationalizes the concept of intuitiveness nor seeks to understand the principles determining the fundamental interaction mechanics of touchless.

This dissertation sets out to understand touchless interaction mechanics in target selection on distant, large, vertical 2D displays. For example, what is intuitive in touchless? How can we design intuitive touchless interaction primitives, the basic units that constitute an interface control? How can we mitigate the lack of precision in touchless input? How can we design feedback languages to improve the touchless user experience? Feedback in touchless systems is exclusively visual and proprioceptive. How can theories of visual perception inform the design of input, feedback, or interface languages for touchless? This theoretical investigation is a crucial stepping stone toward unearthing fundamental knowledge about the potential and limitations of touchless as an interaction modality. Knowledge resulting from this inquiry will drive the design of next-generation touchless systems based on fundamental interaction principles—instead of a reactive adaptation to the sensing technologies.

I present use-driven basic research (Stokes, 1997). The emergence of touchless in different application domains motivated the research questions in this dissertation. Three types of outcomes are produced: (1) knowledge about how sensorimotor relations affect touchless performance, (2) interaction design guidelines for future touchless systems, and (3) a set of touchless interaction techniques for large displays. Overall, my contribution to human-computer interaction research is empirically understanding the interaction mechanics in touchless and using that knowledge to put forth interaction design guidelines for future touchless systems.

In sum, this dissertation explores touchless interaction mechanics through five sets of experiments (Chapters 4, 5, 7, 8, and 9) and two interaction techniques (Chapter 6). I begin with reviewing the literature on touchless interaction, defining the scope, and elaborating the significance of this research (Chapter 2). Chapter 2 also discusses the embodied cognition theory and the Gestalt theories of visual perception and motor action. Chapter 3 delves deeper into touchless interaction mechanics, focusing on target selection techniques. Here, I discuss the ‘crossing’ interaction primitive, less common in traditional input modalities, such as pen or mouse. Chapter 4 discusses empirical results from studies on visual feedback. Chapter 5 focusses on affordances and ability in touchless interfaces, operationalizes intuitiveness, and introduces motor-intuitive touchless interaction primitives. Armed with the results of these experiments, I then introduce two interaction techniques (Chapter 6). First, I present Touchless Circular

Menus (TCM), a command selection technique for large displays using directional strokes. Second, I present interface topographies, a targeting-steering technique using pseudo-haptic feedback. Chapter 7 discusses empirical results from studies on pseudo-haptic feedback in touchless target selection and steering. Chapter 8 discusses experiments on motor control and how hemispheric asymmetry, along with the lack of haptic feedback, affects touchless performance. Chapter 9 presents results from the experiments studying effects of perceptual Gestalt on touchless performance. I then discuss how the empirical results from different studies fit together to understand better touchless interaction mechanics (Chapter 10) and finally conclude the dissertation with open problems and future work (Chapter 11).

Chapter 2. Background, scope, and significance

In this chapter, I first review the emergence of touchless systems across a variety of application domains, such as public spaces, health, or information visualization, and discuss their different interaction patterns, user expectations and domain characteristics. Then, I define interaction mechanics, which encompass interface affordances and people's abilities. I further review the embodied cognition perspective—focusing on the device-less property of touchless and discussing the Gestalt theories of perception and motor action. Review of these theories is crucial as they inform the interaction design solutions of the emerging problems in current touchless research—as discussed in the later chapters. This Chapter concludes with the scope and significance of the dissertation.

2.1. The use of touchless systems across different domains

Current touchless systems can broadly be classified in terms of the size of their interfaces (e.g., large vs. small) or interaction proxemics (e.g., near vs. far-away interactions). Interaction proxemics is a property of an interactive system: the proxemic consequences of the interface and interaction mechanics (Mentis, O'Hara, Sellen, & Trivedi, 2012; O'Hara, Kjeldskov, & Paay, 2011). For example, pen- or touch-based interaction entails a proximal or near-the-display relation exclusively, while touchless interaction supports either distal or a mix of near-and-far interactions—based on the kind of sensors at play (Figure 2.1).

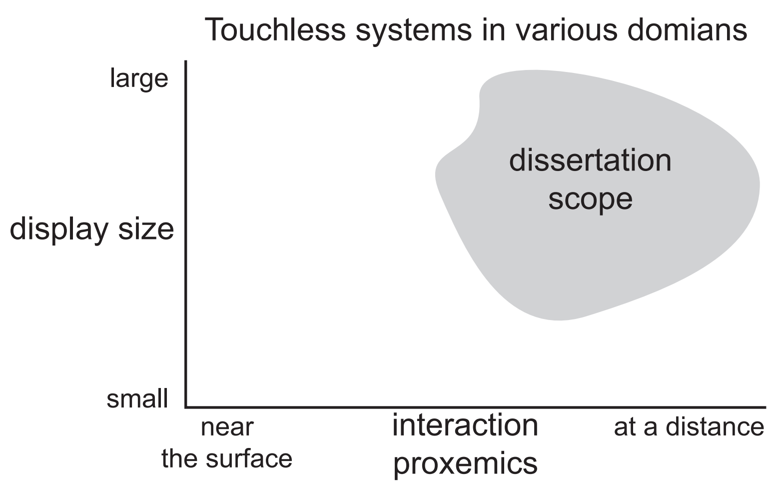

Figure 2.1. Current touchless systems can broadly be classified in terms of their interface size (e.g., large vs. small) or interaction proxemics (e.g., near vs. far-away interactions). The scope of this dissertation is touchless interactions with large displays from a distance.

Human factor studies have found touchless interactions lacking fine-grained precision (Nancel et al., 2011), making them more suitable for pointing, browsing, and lightly manipulating tasks. But in what contexts would it be useful to have gesture controls? While gesture control is exciting and frees users from learning a new input device, with the lack of accuracy and associated fatigue, users often wonder why "Should you care about putting your hands in the air?" (Jennings, 2014).

Like touch-enabled laptops, touchless control was recently introduced in laptops (Jennings, 2014). The ability to control games and applications with mid-air gestures, howver, did not receive many favorable reviews:

> "The lack of accuracy available put paid to all the games we tried and, even when Leap Motion worked as intended, keyboards and gamepads are still far more reliable and satisfying." –Jennings, 2014

Other touchless systems with small displays and near-the-surface interactions include facilitating bimanual interactions with desktop or laptop computers (Guimbretière & Nguyen, 2012) and interactions with household appliances like digital ovens (because during cooking physical contact with an interface is infeasible due to soiled hands or wearing gloves, Garzotto & Valoriani, 2012).

Current research on touchless systems mostly focuses on large displays and interacting from a distance. Because, it provides a context where other interaction modalities such as a mouse, keyboard, pen, or touch crucially limit the interaction proxemics (users are tethered to their input devices or required to be near the interfaces, O'Hara et al., 2013). While using handhelds can provide mobility in such scenarios (Liu, Chapuis, Beaudouin-Lafon, Lecolinet, & Mackay, 2014; Nancel, Chapuis, Pietriga, Yang, Irani, & Beaudouin-Lafon, 2013), touchless relieves the need of acquiring and carrying along an input device (Bailly et al., 2011). In what follows, I briefly discuss prior approaches to large display interaction and then identify specific contexts where researchers are exploring touchless interaction with large displays.

2.1.1. Touchless interaction with large displays

Large display research began with the conception of ubiquitous computing (Weiser, 1993). Historically, yard-scale whiteboards drove the vision of large displays (Czerwinski, Smith, Regan, Meyers, Robertson, & Starkweather, 2003; Swaminathan & Sato, 1997). For example, Liveboard (Elrod et al., 1992), MERBoard (Huang, Mynatt, & Trimble, 2006), or Tivoli (Pedersen, McCall, Moran, & Halasz, 1993) were some of the early works. But as large displays were being extensively built, deployed, and evaluated in Human-Computer Interaction (HCI) settings (Ni, Schmidt, Staadt, Livingston, Ball, &

May, 2006), their size started ranging from three to four standard desktop monitors to a whole wall (4 m x 1.5 m or larger). With the dropping cost of building large displays and the growing need to visualize large volumes of data, large display research in HCI took two important directions—understanding their advantages and innovating effective large-display interaction techniques.

Exploratory works investigated the efficacy of large displays in war-rooms (Jagodic, Renambot, Johnson, Leigh, & Deshpande, 2011; Jagodic, 2011), meeting rooms (Bragdon, DeLine, Hinckley, & Morris, 2011), and design studios (Oehlberg, Simm, Jones, Agogino, & Hartmann, 2012); with single users (Czerwinski, Tan, & Robertson, 2002) and multi-users (Jagodic et al., 2011); with collocated and remote users (Beaudouin-Lafon et al., 2012). Particularly, researchers found large displays improve task productivity (Czerwinski et al., 2003), spatial performance (Tan, Gergle, Scupelli, & Pausch, 2003; Tan, Gergle, Scupelli, & Pausch, 2006), collaborative sensemaking (Andrews, Endert, & North, 2010), difficult data manipulation (Liu et al., 2014), collocated brainstorming (Bragdon et al., 2011), and collaborative visualization (Dostal et al., 2014).

Early research on large display interaction explored traditional point-and-click techniques—mouse, pen-based stylus or single-touch input (Baudisch, Good, & Stewart, 2001; Baudisch et al., 2003; Baudisch, Cutrell, Hinckley, & Gruen, 2004; Bezerianos & Balakrishnan, 2004). Some of the research challenges were how to access remote content on the display, how to optimally manage content layout, or how to enhance display space organization (Bezerianos & Balakrishnan, 2004). To solve those challenges, techniques such as vacuum (Bezerianos & Balakrishnan, 2005), drag-and-pop, drag-and-pick (Baudisch et al., 2003), or tiling (Jagodic, 2011) were proposed. Later research focused on post-WIMP interaction techniques (windows, icons, menus, pointer), such as whole-body movements (Shoemaker, Tang, & Booth, 2007), ray casting (Jota, Nacenta, Jorge, Carpendale, & Greenberg, 2010), or gestures (Bragdon & Ko, 2011). For example, pen-based rectilinear gestures were found significantly efficient than direct selection of far-away targets on large displays (Bragdon & Ko, 2011). Apart from interaction techniques, interaction metaphors were studied to understand how the distance between the display and the user affects users' interaction experience (Jota, Pereira, & Jorge, 2009). Researchers also continue to explore large display experience for different tasks and interaction modalities, such as difficult data manipulation with handhelds (Liu et al., 2014), information visualization with tangible controllers (Jansen, Dragicevic, & Fekete, 2012), or up-close interaction during collocated collaboration (Jakobsen & Hornbæk, 2014).

Figure 2.2. Some example contexts where touchless systems are being increasingly explored. Touchless becomes relevant when interactions are sporadic and acquiring input devices either infeasible or effort some.

An alternative to up-close interaction is distal interaction with large displays. To that aim, people can use device-based (e.g., Gyro mouse or Wii remote) or device-less (e.g., touchless) interaction techniques (Bellucci, Malizia, Diaz, & Aedo, 2010; Nancel et al., 2011). Touchless becomes relevant when interactions are sporadic and acquiring input devices either infeasible or effort some. Other than gaming consoles, following are some primary areas, where touchless systems are being increasingly explored (Figure 2.2):

- *Public spaces*: People in public spaces, such as airports, shopping malls, or smart cities, interact with large displays for a brief amount of time. Hence, they may not spend the time and effort required to connect to an intermediate device to begin interaction (Valkanova, Walter, Vande Moere, & Müller, 2014; Walter, Bailly, & Müller, 2013; Walter, Bailly, Valkanova, & Müller, 2014). While touch displays are now commonly found in such contexts, touchless interaction allows interacting with displays that are out of hand's reach or from a distance (to get a bird's eye view of the display content). Example applications include interactive systems in shopping malls (Walter et al., 2013), street games (O'Hara et al., 2013), and installations for civic participation (Valkanova et al., 2014).

- *Sterile operating rooms*: In sterile environments, at times surgeons need to browse and manipulate images without physical contact to maintain asepsis. Touchless interfaces provide them direct control without the assistance of an intermediary nurse (O'Hara et al., 2013; O'Hara et al., 2014; Ruppert et al., 2012; Schwarz et al., 2011).
- *Patient-centric health settings*: With the increasing urge to make health interventions patient-centric, touchless systems with full-body tracking provide patients more control in managing clinical tools, such as positioning during radiotherapy treatments (Dsouza et al., 2014; Johnson et al., 2011; Morrison et al., 2016; Mullaney et al., 2014; Rosa & Elizondo, 2014; Tan et al., 2013).
- *Consumer electronics*: Touchless interactions can support the sporadic browsing of multimedia in interactive televisions or facilitate interaction with omnidirectional videos (Morris, 2012; Rovelo Ruiz et al., 2014; Vatavu & Zaiti, 2014).
- *Beyond-the-desktop visualization*: Visualizing large data sets have moved from desktops to large displays (Roberts, Ritsos, Badam, Brodbeck, Kennedy, & Elmqvist, 2014); touchless techniques allow multiple users to engage in both proximal and distal interactions with these visualizations (Death of the Desktop, 2014; Dostal et al., 2014; Isenberg, 2014).
- *Collocated collaboration*: Current computing devices vary widely in shapes, sizes, and affordances, ranging from smartphones to centrally shared displays (Bragdon et al., 2011). In such contexts of differently-abled devices, touchless techniques can facilitate distal interaction with shared displays during collocated collaboration and brainstorming (Oehlberg et al., 2012).

2.2. Interaction mechanics

In HCI, novel interaction techniques are frequently proposed, from desktop-based systems to touch to touchless. Such point designs are, however, insufficient toward wider dissemination of research as well as adoption by designers (Beaudouin-Lafon, 2004). In terms of research, it is difficult to build on a gamut of different interaction techniques, without an underlying theory of interaction design, a set of rules and principles that explain their advantages and guide ways to combine and choose between those techniques. Furthermore, to transfer novel interaction techniques into commercial applications, developers need models, methods, and tools that can identify the immediate benefits of shifting to a new interaction paradigm. In sum, the premise of designing interactions over interfaces argue for a theoretical foundation that combines both an understanding of the broader context of use and the sensory-motor details of the interaction. Studying the details of interaction as a sensory-motor phenomenon is as essential as devising new computer algorithms. Such explorations can provide a scientific basis to evaluate the interaction performance and inform new interaction models and design principles.

One of the crucial propellers of post-WIMP computing are the emerging input technologies, and their growing similarity to the devices we use in our everyday world (e.g., pen) or interactional ways of the daily life (e.g., surface or mid-air gestures). These input technologies bring along their distinct set of potential and limitations, thus requiring fundamental interaction design guidelines for different user interface designs (Wigdor, 2010). Wigdor (2010) argues that when architecting these next-generation user interfaces, it is crucial to adapt to human motor, cognitive, and social abilities, which can produce easy-to-learn interfaces and enable interaction scenarios that current mouse-based user interfaces do not. To that end, the five key areas requiring exploration are sensing and processing, input, affordance, and feedback languages, and applications. Sensors and effective processing of the accumulated data are critical to capture users' actions and surroundings. For example, the innovation of multi-touch trackpads enabled rich user interfaces by sensing multiple points of contact simultaneously, compared with the earlier generation of interactive systems with single touch point detection capability. Input languages constitute of a vocabulary of interface commands that are designed by combining interaction primitives—a narrow subset of system-recognized actions that are mapped to system responses (Wigdor & Wixon, 2011). For example, single click, double click, drag, or tap are interaction primitives; *point and click* is a kind of input language. Affordance languages complement input languages. Their influence on interface design is two-fold. First, they identify how sensor capabilities can draw on user abilities and inform the design of interaction primitives and interface commands. Second, they may also provide feedforward mechanisms, in concert with input languages, to guide user input. Feedback languages assist users in understanding a system's reactions to their actions. For example, the auditory feedback when emptying a recycling bin in a Windows system. Finally, investigating applications for emerging technologies can provide a holistic view of interactive systems in the context of use.

Interacting with mid-air gestures or touchless is a novel input technology. The advent of markerless sensors has fueled its popularity and speculated its use in different contexts. While novel touchless systems are being widely explored, an underlying theory of interaction design or a fundamental set of rules and principles are lacking. Because touchless interactions are so markedly different than traditional mouse and keyboard input, it is crucial to invest in theoretical foundations and draw on them toward developing input, affordance, and feedback languages.

But as we dig deeper into the interaction mechanics and study touchless interactions as a sensory-motor phenomenon, it is equally important not to abandon the holistic view of these interactions in the context of use. O'Hara et al. (2013) studied the "naturalness" of touchless from an interactional perspective in different contexts of use. This work was based on Merleau-Ponty's lived-body view of individual experiences and Wittgenstein's socially organized view of action. Not to lose a broader context of use, this dissertation is positioned around interaction with large displays; but it departs from earlier works on social organization of action around touchless systems (Mentis et al., 2012; Morrison et al., 2016; O'Hara et al., 2014a; O'Hara et al., 2014b) to study sensory-motor details of touchless interaction. To that aim, I look at touchless from the embodied interaction perspective (Gibson, 1979; Dourish, 2004) and explore what does the absence of an input device entails—like reckoning on visual perception and proprioception as primary ways of feedback. In what follows, I discuss the overarching theories that informed the experiments in this dissertation (chapters 4, 5, 7, 8, and 9).

2.3. An embodied interaction perspective: The tool, or lack thereof

So far, this chapter tried to convince the need of exploring touchless interaction mechanics for designing future interactive systems. To pursue this investigation, I adopt the embodied interaction perspective (reviewed in Dourish, 2004). Detail discussion of the rich history of embodiment is beyond the scope of this dissertation; the keen reader may read chapters four and five of Paul Dourish's *Where the action is* (2004). Embodied interaction and its antecedent phenomenology, of course, has lent a theoretical lens to many HCI investigations. That is not new. The goal here, instead, is to introduce embodied interaction, explain its relevance in studying touchless, and set up the stage to design empirically testable interaction design theories.

Dourish (2004) defines *embodiment* as "the property of our engagement with the world that allows us to make it meaningful." Embodied phenomena are the ones where we encounter the real world—not the abstract—and find meaning in it through exploration. For example, imagine making a conversation; simple actions like turn taking, turn allocation or repairing organization (resolving problems in speaking, hearing, or understanding) are conversational rules that humans become familiar with while engaged in the activity. However, when designing conversational computer systems, such rules need to be pre-specified to produce natural interaction with users. Thus, a conversation is an embodied phenomenon in our everyday world. Other similar examples are grasping an object, walking down the stairs, using and interpreting nonverbal cues, or understanding social stigmas. Indeed, all such phenomena draw on a sense of *familiarity* with our everyday surroundings—the physical objects, the laws of physics, and the socially constructed world.

Embodiment is central to all our daily experiences with the everyday world. But then, what is particular about *embodied interaction*? What is *not* embodied interaction? To answer this question, it is crucial to understand that embodied interaction is not a kind of interaction per se. It is an approach to design and analyze interactions in HCI—that capitalizes on our physical skills, abilities, familiarity with real-world objects or the relation between social actions and where it is situated (Suchman, 1987). O'hara et al. (2013) studied the role of embodiment in touchless in the social milieu (what are the social implications of touchless in different settings or from a social computing perspective). This dissertation draws on embodiment to study touchless interactions from a different aspect—as a sensory-motor phenomenon (Beaudouin-Lafon, 2004). The goal here is, however, not to focus on abstract cognitive processes, but the phenomenal world we experience daily. To that aim, I will later discuss Gibson's (1979) ecological psychology and how it informs this dissertation's study of touchless interaction mechanics. But before that, I introduce the phenomenological backdrop of embodiment. Instead of the view of Cartesian dualism between the mind and body, I adhere to Heidegger's *hermeneutic phenomenology*—that the meaningfulness of the world lies in how we encounter it practically (1988).

Heidegger's view of phenomenology argues for a fundamental intertwining of thinking and being. A central premise of his work is the concept of *Dasein*—the essence of being in the world, inhabiting as a human being. While inhabiting the world, we act upon it; however, the world is not merely the object of our actions, but also a medium through which we find ways to accomplish our goals. For instance, part of the world—like some physical objects in it—turns into tools or equipment for some task. This view of the world as both an object and medium is how Heidegger couples *intentionality* with being in the world.

In HCI, Heidegger's phenomenology has inspired the analysis of computational theories of cognition (Winograd & Flores, 1986). Particularly, of crucial importance is how he distinguishes the roles of a tool—interplaying between an *object* of experience and the *means* of experience (Miller, 2011). For example, Dourish (2004) provides the example of a mouse while interacting with a GUI: When the mouse is connected to the computer, it is an extension of the user's hand, and the user is acting through the mouse—in Heidegger's terms, *ready-to-hand*. If the mouse reaches the edge of the mousepad, requiring the user to lift and reposition it, the user becomes aware of the mouse as an object of her activity, mediating her action—in Heidegger's terms, *present-at-hand*. Other examples include eyeglasses ceasing to be a separate object of experience and becoming part of the user's experience of seeing the world (Ihde, 1990), the craftsman perceiving the hardness and position of a nail through the hammer which has become an extended limb of his body (Heidegger, 1988) a blind man's cane allowing him to experience the world when in contact with the pavement (Merleau-Ponty, 1962), or a driver's mastery of the steering, through which he achieves the experience of driving (Richardson, 2007). A common theme across these examples is the presence of a device, transitioning from being an object to being absorbed into the fabric of an activity—as the means of experiencing the world or interacting with a computer.

Departing from its antecedents, the traditional input modalities like mouse, keyboard, pen, or touch, touchless features *device-less* interaction with a computer—where the body plays both the *tool* and the *medium*. This presents a unique opportunity for interaction design theorists to inform future designs from a deep theoretical underpinning—exploiting the implications of a lack of tool in interactive computing. Touchless has transformed human-computer interactions on a par with our everyday interactions with the physical world; however, the physical world is governed by the laws of physics, while the computing interface is synthetic (Beaudouin-Lafon, 2004). What does this mismatch entail for touchless systems?

As I discussed before, interaction as a sensory-motor phenomenon includes users' execution of goals, the system's reaction and feedback to their action, and users' assessment of that feedback to continue the interaction. In what follows, I explore implications of the lack of tool in touchless, while focusing on interactions with large, distant, vertical, two-dimensional (2D) displays.

In touchless interaction, we use mid-air gestures to interact with computers—our body acting both as the *tool* and *medium*. In the history of interactive computing, such duality in the role of the human body is new. Although touch-enabled systems are similar to touchless—with no requirement of acquiring a tool—they include a touch surface, which embodies the concept of a tool, later transitioning to a medium. This absence of tool in touchless is often celebrated as a "natural" mode of interaction. Natural, because in daily life, we use our body to interact with the everyday physical world; we grab a book, open a door, lift a box, throw a ball, wave a friend goodbye, wipe the whiteboard, gesture to a direction, and so on and so forth. In our everyday interactions, a tool is not always necessary; we use our body (such as body parts, arms, and legs) to both accomplish a task and experience it (and find meaning in it). But gesturing with a distant 2D display is quite unlike gesturing with a three-dimensional (3D) physical world—thus invalidating the premise that touchless gestures are natural simply because we are *familiar* using them in our surrounding physical world.

Then what are the differences between gesturing in our familiar environment and touchless interaction with distant, large, 2D displays? This stands as the pressing question now. To find an answer, I build on Gibson's exploration of the relation between *seeing* and *acing*—a classic problem in visual perception (1979). Visual perception deals with how living beings can see, recognize the seen, and act on it. To study this phenomenon, Gibson introduced the concept of *ecological psychology*, which encompasses the central construct of *affordances*. Ecological psychology acknowledges the significance of our physical embodiment by positioning *cognition* within the environment, as a concept involving the organism, action, and its environment. Gibson defines affordance as a construct relating the ability of an entity, action, and the environment (Gibson, 1971; discussed in Dourish, 2004). For example, a chair affords sitting to a human, but not to an entity inappropriate for sitting (e.g., a fish). Water affords breathing to fish with its gills, it does not afford breathing to human beings, because we are not appropriately equipped. The concept of affordance has been extensively studied in HCI and extended in different, such as perceived affordances (Norman, 1988), technology affordances (Gaver, 1991), and social affordances (Gaver, 1996). In this dissertation, I use Gibson's affordances to study touchless interaction mechanics.

To study touchless, it is important to understand its affordance and users' abilities in the interaction context. This investigation is crucial to identifying the differences between gesturing in our everyday environment and touchless interaction.

Touchless input is three dimensional. In the absence of a device and its constraints, our whole body and the complete set of physical abilities become available toward realizing affordances of a touchless system. So while interacting with a distant 2D display, we can use our hands or fingers to push, pull, roll, or make directional strokes in mid-air as interaction commands—similar to mid-air gestures we use in our everyday (3D) world. However, the response of our input is available on a 2D, distant display that lacks the 3D worldview of the everyday world (Gibson, 1979). There lies the mismatch—the availability of all physical abilities we use in a 3D world, but to act on, a 2D user interface (UI) without any haptic feedback. Because of the lack of haptic feedback, touchless interaction primarily depends on visual perception and proprioception. Thus, I draw on psychological principles of visual perception (Koffka, 1922) and motor control (Klapp & Jagacinski, 2011) and theories of motor behavior (Sigrist, Rauter, Riener, & Wolf, 2013) to explore touchless interaction mechanics. Each of the chapters 4, 5, 7, 8, and 9 will discuss the pertinent theories and how they inform the subsequent empirical studies. Before that, I look at the emerging problems in touchless and explain the scope and significance of this dissertation.

2.4. Emerging problems

Up to now, touchless systems have been explored largely as a practical exercise—with a variety of prototypes developed opportunistically (Bailly et al., 2011), driven much by the innovation in body tracking sensor technology and the emergence of new algorithms than by a reasoned understanding of the role of physicality (our body) in such interactions. Interaction design theories that govern the traditional input modalities have limited applicability to this new domain. But there is no theory of touchless interaction. How does touchless capitalize users' physical abilities, skills, and everyday familiarity? Which features of touchless are important, which are merely convenient in certain contexts, and which are simply infeasible with average human abilities? This dissertation is about developing answers to some of these questions.

The previous section explained why gesturing in a 3D physical world is unlike using mid-air gestures to interact with distant, 2D displays—theoretically. In practice, researchers studying user performance with touchless prototypes have reported several breakdowns too. What is lacking is an explanation of these observations—a theory. This dissertation attempts to address this limitation. Although I do not set out to identify and explain an exhaustive set of breakdowns in touchless systems—observed till date, I provide an overarching theoretical perspective to study them (Figure 2.3). Under the umbrella of embodied interaction, I illustrate the use of specific theories (from traditional fields like visual perception and motor behavior) to explain certain interaction breakdowns; and then go on to generate new theories and interaction design principles. These principles inform new touchless interaction techniques (Chapter 5), which in turn facilitates studying further aspects of touchless interaction mechanics (Figure 2.3).

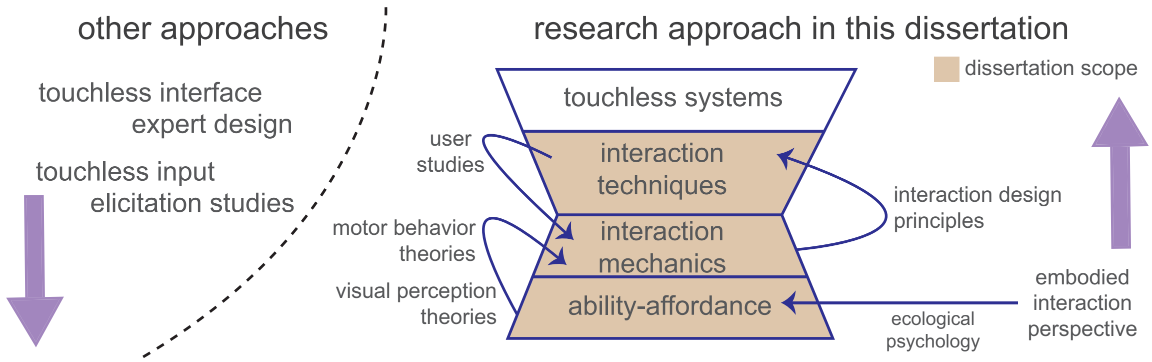

Figure 2.3. Up to now, touchless systems have been explored through building prototypes or eliciting gesture input from users—driven much by the innovation in sensor technology and new algorithms than by a reasoned understanding of the role of our body in such interactions. Instead, this dissertation follows a bottom-up approach: understanding the sensory-motor relations in touchless interactions and then using that knowledge to drive design guidelines.

The growing popularity of touchless stems from its expectation as something *natural* to use, something already familiar to us. But such an attribution has led to many debates; mainly because current studies often adopt a vernacular definition of 'natural' or 'intuitive' as instinctive or spontaneous—thereby lacking an operationalization. How do we define natural? As effective, accurate, a feeling of familiarity, easy to learn, easy to remember, or fun to use? For example, empirical studies have shown that due to the lack of haptic guidance, touchless gestures are less efficient and more fatiguing than device-based gestures (Nancel et al., 2011). Does that make touchless less natural? Or just less efficient than touch?

Investigating touchless as a sensory-motor phenomenon can address this question. However, current research is either exploring the naturalness of touchless input through elicitation studies (Aigner et al., 2012; Grandhi et al., 2011; Vatavu & Zaiti, 2014), or designing touchless interface languages motivated by mouse (Hespanhol et al., 2012; Jota et al., 2009), pen (Guimbretière & Nguyen, 2012) or touch-based interfaces (Bailly et al., 2011; Ren & O'Neill, 2012). In elicitation studies, users suggest gestural inputs based on the outcome shown on the user interface (UI). This method aims to identify an input language that is based on everyday metaphors (Lakoff & Johnson, 1980). For example, Grandhi et al. (2011) reported user preference toward bimanual gestures and intuitiveness of dynamic gestures (iconic representation of the motion required for the manipulation) over static iconic hand poses. For example, users would prefer a “wiping” hand movement over a static hand sign to trigger a “delete” action.

On the other hand, expert design studies first iteratively design touchless interface languages, such as target selection, pan-and-zoom techniques or menus, and then evaluate their user performance. These proposed interaction techniques are either motivated by our everyday metaphors, similar to elicitation studies (e.g., pushing to select, like pushing to open a door, Hespanhol et al., 2012), or other traditional UI languages (e.g., marking menu for pen-based interaction, Guimbretière & Nguyen, 2012; finger-count menu for touch interaction, Bailly et al., 2011; or linear menu for WIMP interaction, Bailly et al., 2011).

The challenges with this approach to touchless research are two-fold. First, uncoupling the touchless input and interface leaves no space to explore the mechanics of touchless interactions (Beaudouin-Lafon, 2004). Second, in the absence of any knowledge of touchless interaction mechanics, when designing touchless UIs, designers resort to WIMP, pen or touch-based interaction principles. As a result, when touchless evaluation studies report certain interaction techniques to be intuitive, they fail to explain why other techniques were unintuitive or ineffective.

For example, a number of recent studies have studied touchless target selection, using static poses (a fist, Bailly et al., 2011), finger count (Bailly et al., 2011; Kulshreshth & LaViola Jr, 2014), crossing a delimiter (Ren & O'Neill, 2012), 3D angular strokes (Guimbretière & Nguyen, 2012), push (Hespanhol et al., 2012), dwell (Hespanhol et al., 2012), multi-finger pinch (Guimbretière & Nguyen, 2012) or roll-and-pinch (Ni et al., 2008). In touchless target selection, researchers noted a number of limitations. Guimbretière and Nguyen (2012) report the unreliability of a three-dimensional marking menu because users failed to gauge a 3D angle for the mark gesture. Ren and O’Neill (2012) report similar findings for their stroke technique. For push-to-select gesture, Hespanhol et al. (2012) report a translation-action ambiguity problem. A touchless gesture suffers from translation-action ambiguity when users frequently trigger actions while repositioning their body in space. They also report accidental invocation problems with *dwell* or holding gesture. Bailly et al. (2011) found users faced difficulty in constraining their hand movements in a 2D plane, thus often triggering inadvertent commands. Markussen et al. (2014) found their proposed mid-air word-gesture keyboard slower than touch—in spite of the increased fluidity in touchless movements during target selection. Some of the possible reasons that authors discuss are the incompatibility between the stimulus and response, gestures in the motor space compared with the keyboard and feedback on the display and the heavy reliance on visual feedback.

This lack of theory and principles to explain observations encountered during touchless studies can only be mitigated using a bottom-up approach: understanding the sensory-motor relations in touchless interactions and then using that knowledge to drive design guidelines.

2.5 Scope of the work

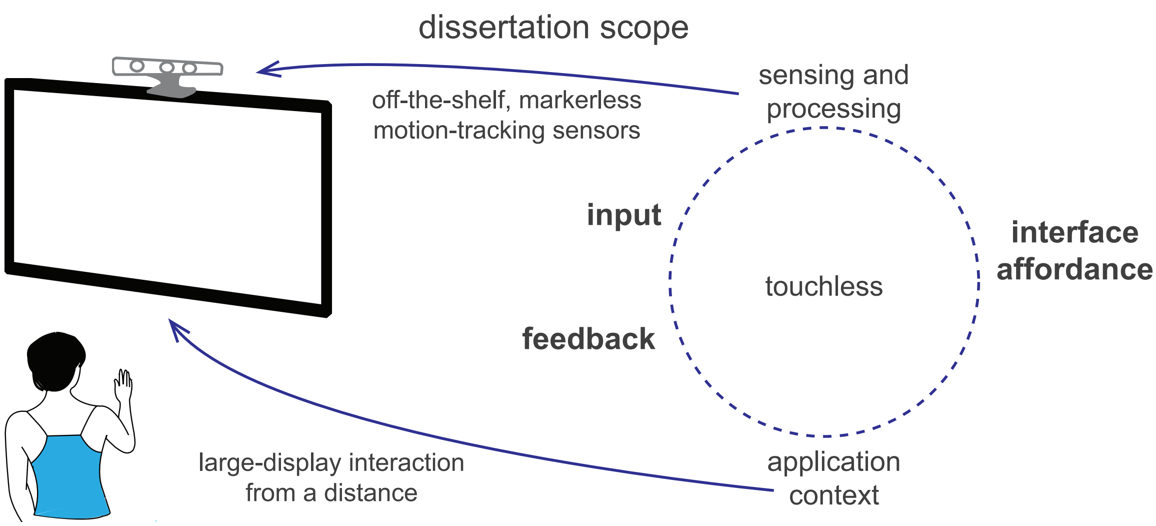

Figure 2.4. Around touchless target selection, this dissertation studies input, feedback, and affordances—using off-the-shelf, markerless motion-tracking sensors (Kinect™), in the application context of large-display interaction.

Within touchless interaction with large, 2D, distant displays, this dissertation focuses on target selection. Target selection is the most fundamental task in interactive computing with a variety of ways to accomplish it—from the command line argument *cat* to a voice command *open*. Around touchless target selection, I study input, feedback, and affordances (Figure 2.4). For touchless input, I operationalize naturalness or intuitiveness, introduce motor-intuitive interaction primitives (Chapter 5), and study motor control (Chapter 8). Chapters 4 and 7 discuss experiments on visual and pseudo-haptic feedback respectively. Toward studying interface affordances, I design interaction techniques (Chapter 6) and explore effects of Gestalt principles in touchless interactions

(Chapter 9). Among the five key aspects requiring innovation for architecting next-generation interfaces (Wigdor, 2010), this dissertation does not delve into sensing or applications of touchless. All experiments in this dissertation use off-the-shelf, markerless motion-tracking sensors (Kinect™) and emulates the application context of large-display interaction.

Methodologically speaking, this dissertation is primarily a number of controlled, quantitative studies. To study touchless as a sensory-motor phenomenon, I use theories from more traditional fields, like cognitive psychology and motor behavior. Thus, their style of empirical investigation is borrowed. Such rigorous hypothesis testing approach has led to many important advances in HCI because it provides a scientific basis for users' performance evaluation (Newell & Card, 1985). However, controlled experiments provide internal validity at the cost of ecological validity. So it is important to stress that the results in this dissertation should not be overgeneralized. Laboratory studies allow measuring user performance without any extraneous factors at play, which may differ significantly within different application contexts. It is out of the scope of this dissertation to make such generalizable claims, and future work must associate the findings here with the holistic level of the interaction context in use.

2.6. Significance of this research

The significance of this research is to address the crucial need for understanding the fundamental interaction principles of touchless—instead of a reactive adaptation to the advancements in motion-sensing technology. The overarching research aim here is to generate a set of theories explaining the sensory-motor relations in touchless interaction. Whereas prior HCI approaches to designing touchless systems have been either building prototypes or eliciting a gesture vocabulary, this dissertation sets out to generate fundamental knowledge that can inform touchless interaction design principles. Designing interactions grounded in interaction theory has long been argued for (Beaudouin-Lafon, 2004). I employ that design philosophy to provide a theory of touchless interaction—in terms of quantifiable results testing the sensory-motor properties of touchless and design principles informed by those empirical results.

Chapter 3. Understanding touchless interaction mechanics

The backdrop is now complete. Chapter 2 detailed the theoretical outlook of this dissertation, introduced its scope and explained its significance. Up to this point, I have discussed touchless interactions in general. The goal of this chapter is to identify the key aspects of touchless interaction mechanics and serve as the necessary introduction to the empirical studies in the next six chapters. It also delves deeper into the current approaches of touchless target selection. In a sense, this chapter bridges the broad theoretical abstractions of Chapter 2 with the particular functional aspects of touchless interactions that are investigated hereafter (chapters 4 to 9).

3.1. Related work

This dissertation looks into three aspects of touchless interactions mechanics—input, feedback, and interface affordances—using off-the-shelf sensing technology. Although in the experiments discussed later, users interact with large displays while sitting at a distance, users' body posture is not a topic of interest here. It served as a convenience to participants during the study (often around two hours) and increased the ecological validity of the empirical results (for laid-back settings, where users may remain seated during the interaction). However, the interaction mechanics, explored here, exclusively deal with hand gestures—not arm, other body parts, or full-body gestures. This is crucial to note because different body parts imply different movement and control abilities, thus affecting touchless interactions differently.

Touchless performance, such as efficiency, accuracy, and levels of fatigue, has been explored before—but not toward generating touchless interaction design guidelines per se. For example, while investigating mid-air pan-and-zoom techniques for very large, wall-sized displays, Nancel et al. (2011) showed that due to the lack of haptic guidance, touchless gestures are less efficient and more fatiguing than device-based gestures (e.g., a mouse wheel or touchpad). However, they found touchless gestures causing significantly fewer overshoots (task errors) than 2D surface-based gestures (e.g., touchpad). Within touchless, linear gestures were faster than circular gestures.

To measure upper-arm fatigue (a condition often called gorilla arm), Hincapié-Ramos et al. (2014) proposed a novel quantitative metric drawing on the biomechanical structure of the arm—consumed endurance. Fatigue in HCI is usually measured with self-reporting scales, where researchers ask users to rate their perceived physical effort, such as the NASA TLX or the Borg CR10 scale. During validations studies, the consumed endurance metric correlated strongly with the Borg CR10 scale, a standard measurement instrument of perceived exertion (Hincapié-Ramos et al, 2014). Authors also provided a set of guidelines to design less-fatiguing touchless interfaces. For example, they suggested that having arms bent and the interaction plane center to the body (see Figure 4, Hincapié-Ramos et al., 2014) is least tiring when selecting targets on a 2D plane and the SEATO keyboard layout (see Figure 7, Hincapié-Ramos et al, 2014) best balances efficiency with effort for touchless text entry.

Kajastilan et al. (2012) studied touchless gestures to accomplish a secondary task (such as tuning a radio) while attending to a primary task (such as driving). When comparing control gestures (touch vs. touchless, both circular) for visual and auditory interfaces, they found that user accuracy of the auditory interface was at par with the visual when using touchless gestures (see Figure 4, Kajastilan et al., 2012). However, overall, with visual and auditory feedback, the touchless interface was slower than the touchscreen.

A how-to-guide for designing touchless interactions with Microsoft Kinect is also available for developers (Microsoft, 2016), which provides pointers on how to optimize sensor performance and design appropriate interfaces and feedback languages for different application domains.

In sum, examining interaction mechanics of touchless has been the byproduct of several research endeavors, and they have identified efficiency, accuracy, and fatigue among the important outcome measures. This dissertation brings touchless interaction mechanics to the primary focus.

3.2. Interaction mechanics

3.2.1. Sensing

Microsoft's Kinect is a camera-based solution for full-body tracking. This technology, enabling markerless motion capture using a camera system, was first introduced as a commercial videogame console in 2010 (Figure 3.1). Kinect uses a range camera technology from PrimeSense™ that understands a 3D scene in two steps. First, it emits a continuously-projected infrared structured light in the environment. Then, it uses its depth sensors (infrared laser projector combined with a monochrome CMOS sensor) to record video data in 3D under any ambient light conditions. The computation of depth map broadly uses two classic computer vision techniques for 3D scene reconstruction, depth from focus and depth from stereo. When a live scene is processed by the Kinect, two versions of the scene is recorded, the color map (using the RGB camera) and depth map (using the depth sensors). Once a live scene is captured, machine learning algorithms are used to discover the 3D skeleton of a human body—if present at the scene. It also provides an estimate of robustness of the tracking output.

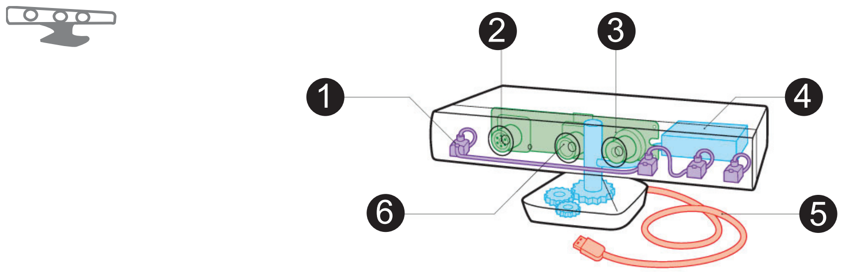

Figure 3.1. Experiments in this dissertation use off-the-shelf tracking sensors, Microsoft's Kinect. It is a camera-based solution for full-body tracking, enabling markerless motion capture using a camera system, and was first introduced as a commercial videogame console in 2010.

Human skeleton detection works as following. For each 3D scene, the Kinect evaluates how well each pixel matches the typical features of an example template. For example, does the pixel looks similar to one at the top of the body, or at the bottom? Each of the pixels is then scored accordingly. This evaluation uses a randomized decision forest search algorithm (Shotton et al., 2013). Broadly speaking, the randomized decision forest search is a collection of decisions, each of which asks whether a pixel (with a certain set of features) of the scene is a candidate for a particular body part. This evaluation algorithm is already trained with a collection of motion capture data (around 500,000 frames). Once the candidacy of each pixel to a particular body part is decided, the likely location of the skeletal joints is computed based on biomechanical constraints, and a 3D skeleton is built. Microsoft Xbox computes this algorithm 200 times per second—way faster than prior skeletal recognition algorithms. Due to its speed and robustness, these sensors are being used not only in games, but also in many computer vision tasks (Han, Shao, Xu, & Shotton, 2013), such as interaction recognition (Chattopadhyay, 2011; Yun, Honorio, Chattopadhyay, Berg, & Samaras, 2012) or activity recognition (Sung, Ponce, Selman, & Saxena, 2012).

3.2.2. Input, feedback, and affordances

Touchless input can be bare hand (Hespanhol et al., 2012) or require wearing hand gloves (Vogel & Balakrishnan, 2005). With markerless camera-based sensors, like Kinect, users can interact with a bare hand. What kind of hand gestures would be suitable for touchless interaction is an emerging area of HCI research—gesture elicitation. Elicitation studies (Grandhi et al., 2011), and its variants (Nebeling et al., 2014) have explored different gesture inputs drawing on gestures used in our daily world (Morris, 2014). Instead of asking users to report intuitive gestures at the macro level (e.g., how would you like to indicate an undo action after deleting a folder inadvertently?), my approach to touchless input is deconstructing its intuitiveness from the perspective of human abilities and interface affordances (e.g., Is it intuitive for us to accurately make orthogonal hand movements facing a 2D interface? More importantly, what is intuitive in touchless?). I explore this research question by drawing on the differences between the physical world and touchless, and emphasizing the role of embodiment in touchless interaction (chapters 5 and 8).

Touchless feedback is exclusively visual and proprioceptive—with no haptic guidance. Prior research has shown its telling effects on touchless performance—slow and tiring. Touchless interaction using hand gloves or other wearables have studied workarounds this problem, like vibrotactile feedback (Foehrenbach, König, Gerken, & Reiterer, 2009; Freeman, Brewster, & Lantz, 2014; Lehtinen, Oulasvirta, Salovaara, & Nurmi, 2012; Pasquero, Stobbe, & Stonehouse, 2011; Richter, Loehmann, Weinhart, & Butz, 2012). For example, Freeman et al. (2014) found no significant effect of tactile feedback on selection time, but on reducing task workload.

Other systems have also explored non-contact tactile feedback, such as AIREAL (Sodhi, Poupyrev, Glisson, & Israr, 2013), Ultrahaptics (Carter, Seah, Long, Drinkwater, & Subramanian, 2013), HaptoMime, (Monnai, Hasegawa, Fujiwara, Yoshino, Inoue, & Shinoda, 2014), or ultrasound transducers (Hoshi, Takahashi, Iwamoto, & Shinoda, 2010), and auditory feedback (Kajastila & Lokki, 2013; Vogel & Balakrishnan, 2005). This dissertation looks into visual feedback (Chapter 4) and pseudo-haptic feedback (Chapter 7) in touchless target selection.

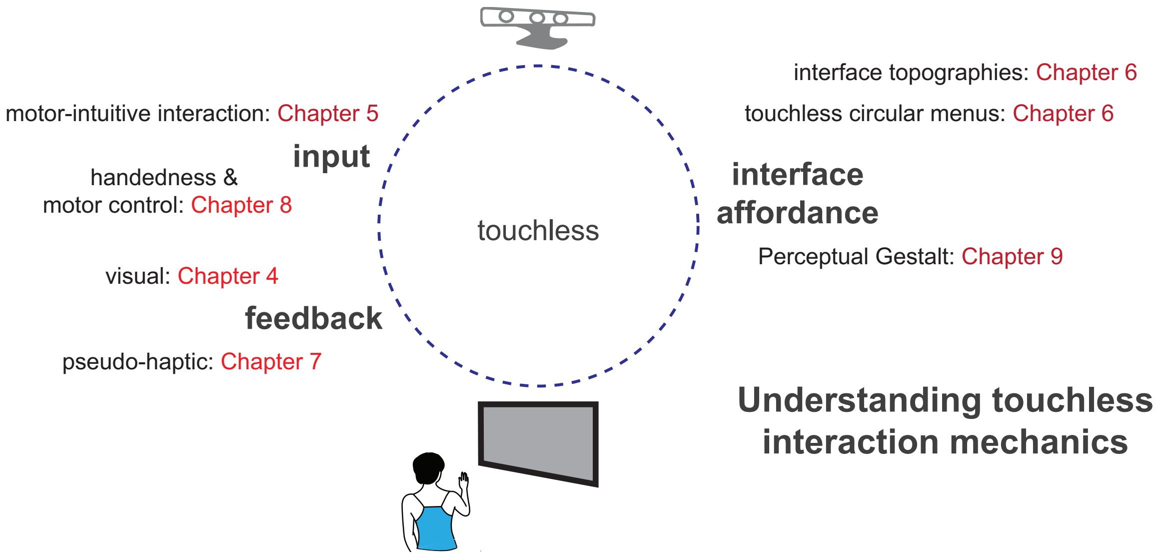

Figure 3.2. Although the dissertation chapters focus on a primary area of exploration, either input, feedback, or affordances, each of them is also inclusive of the other aspects—altogether studying touchless interaction mechanics.

I study affordance in touchless interaction mechanics in two ways. First, human abilities and interface affordances are explored to study touchless input—touchless interaction primitives that make up the building blocks of a touchless interface (chapters 5 and 9). Second, I propose touchless interaction techniques capitalizing interface affordances and evaluate them with users (Chapter 6).

Although this section tried to decouple the dissertation chapters and inject them into the three areas of touchless interaction mechanics, they are inherently intertwined (Figure 3.2). Each Chapter, thus has a primary area of exploration, either input, feedback, or affordances, but are also inclusive of the other aspects—altogether studying touchless interaction mechanics.

3.3. Target selection

This dissertation studies interaction mechanics around touchless target selection. Touchless target selection techniques were briefly discussed in Chapter 2. In this section, I provide a detailed review of the current approaches (Table 1), discuss their performance, and identify the emerging problems. I do not claim this review to be exhaustive—rather it is representative of recent research. Furthermore, although point and select interactions are often studied together in HCI, pointing performance in touchless interfaces is beyond this dissertation's scope; the focus is exclusively on target selection.

Table 3.1. Current touchless target selection techniques in different contexts of use.

Target selection method	Technology type	Related works
wrist rotation	wearable glove	Ni et al., 2008; Ni et al., 2011
tap (angle between palm and another finger)	wearable glove	Markussen, Jakobsen, & Hornbæk, 2013
pinch, thumb to another finger	wearable glove/ IR markers	Banerjee, Burstyn, Girouard, & Vertegaal, 2011; Markussen, Jakobsen, & Hornbæk, 2014; Ni et al., 2008; Ni et al., 2011; Vogel & Balakrishnan, 2005
	bare hand	Guimbretière & Nguyen, 2012
push, orthogonal to a 2D display	bare hand	Hespanhol et al., 2012; Kajastila & Lokki, 2013; Pyryeskin, Hancock, & Hoey, 2012
push, orthogonal to a 2D display, with a velocity threshold	bare hand	Seixas, Cardoso, & Dias, 2015
dwell, for a time window	bare hand	Freeman, Brewster, & Lantz, 2014; Hespanhol et al., 2012; Microsoft Kinect®; Pyryeskin, et al., 2012
directional stroke	bare hand	Bailly et al., 2011; Ren & O'Neill, 2012
crossing	bare hand	Ren & O'Neill, 2012; Schwaller, Brunner, & Lalanne, 2013
freehand movement	wearable glove	Markussen, et al., 2014
grab/fist/closed palm	bare hand	Bailly et al., 2011; Hespanhol et al., 2012; Pyryeskin, et al., 2012; Seixaset al., 2015; Song, Goh, Hutama, Fu, & Liu, 2012

finger combination, static pose	bare hand	Bailly et al., 2011; Freeman, et al., 2014; Kulshreshth & LaViola Jr, 2014; Sridhar, Feit, Theobalt, & Oulasvirta, 2015;
lassoing	bare hand	Hespanhol et al., 2012
enclosing with two hands	bare hand	Hespanhol et al., 2012

Touchless target selection is not a brand new area of research. As evident in Table 1, target selection methods are being explored for more than a decade. However, the continual emergence of advanced sensing technologies and a shift toward intuitive interactions (rather than designer-driven techniques) makes this research both timely and relevant.

To enable precise gestures like in-air tap (making an angle between a finger(s) and the palm), tilting of the wrist, or pinching using multiple fingers, researchers have used wearable gloves with IR markers (Ni et al., 2008; Ni et al., 2011; Markussen, et al., 2014; Vogel & Balakrishnan, 2005). Such gestures have been studied for command selection from menus (Ni et al., 2008) or mid-air text entry using posture-letter mapping (Sridhar, et al., 2015). In bare-hand interactions, with off-the-shelf camera-based tracking solutions, the common gestures are push (Hespanhol et al., 2012), dwell (Pyryeskin, et al., 2012), grab (Bailly et al., 2011), and 3D directional strokes (Ren & O'Neill, 2012). Some of these target selection methods were studied with horizontal surfaces, to enable mid-air interaction just above a multi-touch surface (Banerjee, et al., 2011), while others with vertical displays (Bailly et al., 2011). Another broad classification of the selection methods reviewed here is the temporal aspect—a static gesture or a dynamic gesture. For example, making a certain combination or arrangement of fingers to select a menu option (Kulshreshth & LaViola Jr, 2014) or entering a particular alphabet (Sridhar, et al., 2015) is a static gesture. While pushing orthogonal to a display to indicate a selection (Hespanhol et al., 2012) or moving over a series of alphabets on a keyboard to type in a word is a dynamic gesture. Gestures can also be a mix of two, such as roll-and-pinch, where users make a pinch to start a selection, then tilt their wrist toward the circular menu option of their choice, and finally release the pinch to indicate their intention of target selection (Ni et al., 2008).

Each of the selection techniques listed in Table 1 has different performance benefits and limitations. Rather than enumerating all of them in details, it is interesting to note some common trends. For example, grab gestures are reported more accurate than push (Seixas et al., 2015). Although built upon the success of marking menus (Kurtenbach, Buxton, 1994; Lepinski, Grossman, & Fitzmaurice, 2010) that interpret directional strokes as target selection commands, studies report users' limitations in making accurate 3D strokes (Guimbretière & Nguyen, 2012; Ren & O'Neill, 2012). Both dwell and push suffer from a limitation of accidental invocations; selections are invoked inadvertently when repositioning the body in space, a problem in distinguishing between the translation and action movements (Hespanhol et al., 2012).

Now that the stage is set for a deeper exploration of touchless interaction mechanics in target selection, we move on to the next chapters, where I will present detailed empirical studies testing a set of hypothesis on feedback, input, and interface affordances.

Chapter 4. Visual feedback

In the absence of any haptic feedback, touchless primarily rely on visual cues, but properties of visual feedback remain little explored. This Chapter systematically investigates how large-display touchless interactions are affected by (1) types of visual feedback—discrete, partial, and continuous; (2) alternative forms of touchless cursors; (3) approaches to visualize target-selection; and (4) persistent visual cues to support out-of-range and drag-and-drop gestures.

4.1. Feedback or lack thereof?

In spite of the abundant enthusiasm about more "natural" forms of interaction, the lack of feedback in touchless scenarios raises important usability concerns (Nancel et al., 2011; Norman, 2010). In fact, unlike mouse or touch-based interactions, touchless synthesizes input and output from physically disconnected motor and display spaces, and without any haptic feedback. This lack of haptic guidance reduces users' efficiency and accuracy, because users are excessively dependent on other forms of sensory feedback, such as visual, auditory, or proprioception (Markussen et al., 2014; Nancel et al., 2011). Researchers have tried to compensate this lack of haptic feedback using visual and auditory feedback (Kajastila & Lokki, 2013; Vogel & Balakrishnan, 2005), or tactile feedback (Gupta, Morris, Patel, & Tan, 2013; Sodhi, Poupyrev, Glisson, & Israr, 2013). Specifically, visual feedback has been used to improve the learnability of touchless gestures (Walter, Bailly, & Müller, 2013), to identify multiple users (O'Hara et al., 2014), to communicate gesture ambiguity (Vogel & Balakrishnan, 2005), and to represent clicking and swiping gestures (Markussen et al., 2014; Vogel & Balakrishnan, 2005). Although visual feedback is being actively used in touchless interaction, a systematic exploration of its properties is lacking.

Visual feedback in touchless interactions should guide users' movement effectively. It should also be salient among an array of artifacts on a large display. The role of visual feedback in acquiring and learning movements has been extensively studied in human motor science (Saunders & Knill, 2004; Sigrist, Rauter, Riener, & Wolf, 2013). Similarly, attributes of display artifacts have been widely explored in the visual search literature (Wolfe, 1998; Wolfe & Horowitz, 2004). But these findings have not been significantly adopted to guide the design of visual feedback in touchless interactions. Designers simply consider representing users' position and their actions: "where the user is" (e.g., with an open hand) and "what the user is doing" (e.g., a grab posture). To help users learn, retain, and perform touchless gestures effectively, we are faced with the challenge of designing visual feedback as a salient yet non-distracting aide.

The main contribution of this chapter is to explore visual feedback in large-display touchless interactions—using six controlled experiments—along four aspects: (1) types of visual feedback; (2) alternative forms of touchless cursors; (3) alternative approaches to visualize target-selection; and (4) persistent visual feedback for two common user actions: *drag-and-drop* and when users *land out of the display range*. Our approach to explore visual feedback is informed by the motor science and the visual perception literature. A successful design of visual feedback have the potential to augment users' proprioception, and somewhat compensate the lack of haptic feedback in touchless interactions. Our work makes the following contribution:

- We discuss related work about visual feedback from the motor science and the visual perception literature—such as timing, attributes, and semantics—that can inform future research on designing appropriate visual feedback for different touchless interactions (section 4.2).
- We provide empirical results from six controlled experiments that explore types of visual feedback, shape, size, color and opacity of touchless cursors, different approaches to visualize target selection, and persistent visual feedback in touchless interactions (sections 4.3 – 4.9).
- Grounded in our empirical results, we provide practical guidelines for designing visual feedback in large-display touchless environments (section 4.11). Finally, we illustrate our guidelines by designing a visual feedback routine for drag-and-drop operations across a touchless system's three interaction states—idle, active, and engaged.

How visual perception regulates attention and controls movement is complex and being extensively studied. Still, our work is a first step toward adopting some existing results and rethinking the design of visual feedback in touchless interactions. Our findings can facilitate the development of a visual feedback language for large-display touchless interfaces.

4.2. Background

Visual feedback in motor responses

Visual feedback plays a twofold role in motor responses: motor control and motor learning. Hence, the impact of visual feedback on movement is widely studied in rehabilitation, sports training, and minimally invasive surgery. Two aspects that mediate the role of visual feedback in motor responses are task complexity and feedback visualization.

*Motor control***.** While proprioception estimates the initial body posture and selects a motor command, pointing movements are continually corrected by the visual feedback of the hand (Scheidt, Conditt, Secco, & Mussa-Ivaldi, 2005). Processing of visual feedback while pointing movements can be quite short (e.g., 100 ms, Zelaznik et al., 1983), and thus facilitate the accuracy of rapid movements. In dynamic environments, where closed-loop control (sensory feedback of the users' action) is possible, visual feedback informing motion pattern and position coordinates significantly affect hand movements—in both early and later stages of the movement (Saunders & Knill, 2004).

Motor learning. In any desktop environment, transfer functions (or gain factors) define how amplitudes of hand and cursor movements relate to each other; these are a type of visuomotor transformation that we can easily master due to our sensorimotor abilities (Verwey & Heuer, 2007). In general, when users need to retain mastery of visuomotor transformations, the type of visual feedback during the practice plays a key role: While terminal visual feedback (at the end of the movement) facilitates simple tasks, such as aiming movements using a mouse, continuous visual feedback helps complex tasks, such as inter-limb coordination skills (Sigrist et al., 2013; Sülzenbrück, 2012). Even the frequency of visual feedback—when decreased with decreasing task complexity—further facilitates motor learning. Touchless interactions in large-display environments range from bimanual gestures for data manipulation to static gestures for mode switching. Visual feedback, if appropriately used, can augment learnability of such visuomotor transformations.

Visualization. Visual feedback designs are effective when they enable parallel processing of the visual and the kinesthetic information about the ongoing movement (Sigrist et al., 2013). In motor learning, they range from abstract (lines, bars, curves, Lissajous figures) to natural visualizations (virtual avatars, 3D animations). Studies indicate that it is very important to provide feedback about only the relevant key features of the task (Huegel, Celik, Israr, & O'Malley, 2009). While it is common to provide user information in large-display touchless interactions using a skeleton representation, rethinking our visual feedback designs may facilitate user performance.

Visual attributes guiding attention

Design-dimensions of display artifacts have been widely explored in visual search literature (Smith and Thomas, 1964; Wolfe, 1998; Wolfe & Horowitz, 2004). But these findings have not been significantly adopted to guide the design of visual feedback in touchless interactions. For example, research suggests that color coding leads to efficient visual search (Carter, 1982), but in a dense display efficiency is retained only if the distractors and the targets are widely separated in color space (D'Zmura, 1991). Although debatable, the topological property of a "hole" or the number of line terminators are often considered as features that guide attention in visual search (Wolfe & Horowitz, 2004). The relative size of a target item and how densely packed it is (spatial density) compared to other display artifacts also plays a role in guiding attention (Wolfe, 1998). Empirical studies suggest that attention can be efficiently guided to opaque targets among transparent objects, but it is more difficult to find one transparent item among all opaque items. Interestingly, the effect of opacity is explained by the human tendency to combine multiple cues—namely motion, luminance and structural features (Wolfe, Birnkrant, Kunar, & Horowitz, 2005).

With the absence of haptic feedback in touchless interactions, we are faced with the challenge of designing visual feedback that can help users control and learn touchless gestures effectively. Inspired by the role of visual perception in motor responses and visual search, our work is a first step to investigate the effects of visual feedback in large-display touchless interactions.

We conducted six within-subject experiments to understand how the following four aspects of visual feedback affect large-display touchless interactions: (1) types of visual feedback (discrete, partial, and continuous); (2) alternative forms of touchless cursors; (3) alternative approaches to visualize *target-selection*; and (4) persistent visual feedback for *drag-and-drop* operation and when users *land out of the display range* (Figure 4.1). Findings from these empirical studies can facilitate the development of a visual feedback language for future touchless interfaces.

4.3. General method

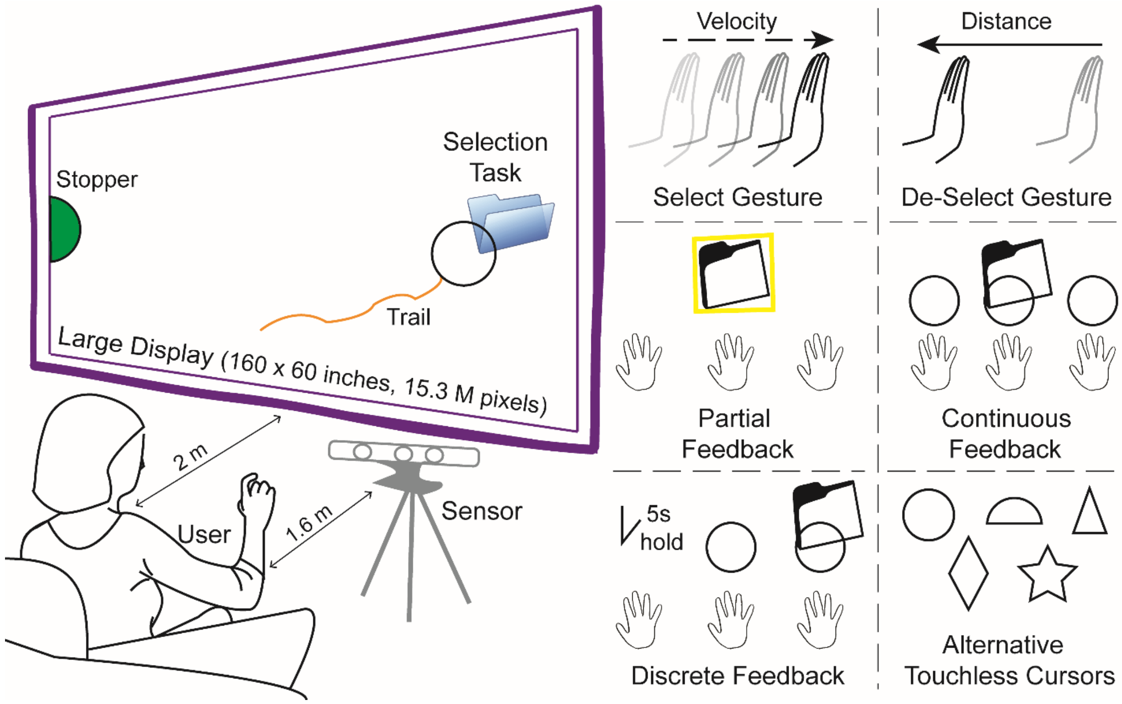

Figure 4.1. We conducted six controlled experiments to understand how visual feedback affects user experience in large-display touchless interactions. (Left) In our experiment, participants used touchless gestures to select display objects while sitting away from a large display. (Right) They used a velocity-based *select* and a distance-based *de-select* gesture. We evaluated three types of visual feedback (partial, continuous, and discrete) and alternative touchless cursors. (Left) We also designed and evaluated *Stoppers*—semantic visual feedback informing users when they are out of the display range, and *Trail*— persistent visual feedback echoing the path of movement during drag-and-drop operations.

Apparatus

Our experiments were conducted using a high-resolution large display that comprises of eight 50–inch projection cubes laid out in a 4 x 2 matrix. The large display is driven by a single computer. Each of these cubes has a 1600 x 1200 pixel resolution, resulting in a 160-inches wide and 60-inches high display with over 15.3 million pixels (Figure 4.2). For motion tracking, we used a Kinect™ (for Windows) sensor. All experiments were written in C# running on Windows 7, and were implemented with OpenNI 1.4 SDK and PrimeSense's NiTE 1.5 middleware.

Participants

A total of 37 right-handed participants with no color-blindness were recruited from an urban university campus; experiments were conducted in two rounds (December 2012 and August 2013). 18 participants (9 females, 13 familiar with touchless gestures) took part in the first five experiments (first round), and 19 participants (8 females, 11 familiar with touchless gestures) took part in the sixth experiment (second round). 15/18 and 15/19 participants were below 30 years of age. Participants were randomly recruited by sending out emails using the university's mailing list. The study was approved by the university's Office of Research Administration (IRB Study No. 1210009814 and 1303010855), and participants were compensated with a $20 gift card for two hours of participation.

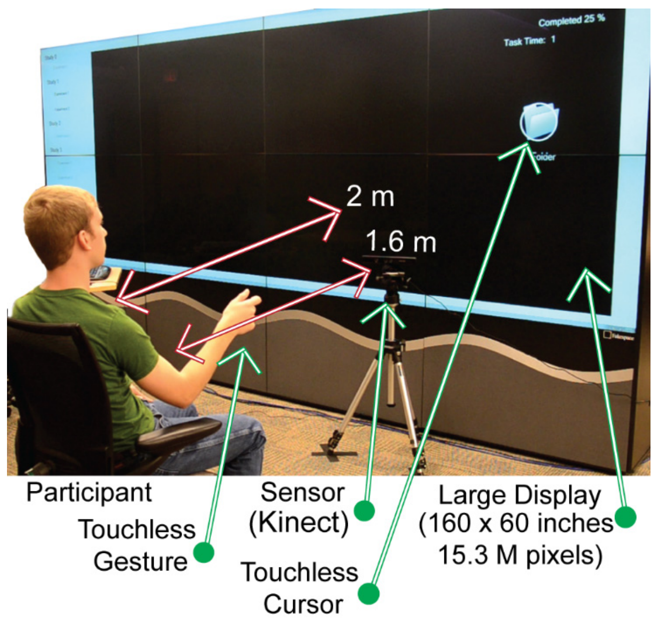

Figure 4.2. Our experiments were conducted using a 160 x 60 inches large display with a resolution of 15.3 M pixels. We used Microsoft's Kinect sensor for motion tracking, and across all six experiments, participants sat in a chair 2 meters away from the large display.

Gesture primitives

To explore visual feedback in large-display touchless interactions, we designed two gesture primitives: *select* and *de-select*. A select gesture was defined as a forward movement of the hand with a certain velocity (350 mm/s), and a de-select gesture was defined as a backward movement of the hand by a certain distance (100 mm, Figure 4.1). Using these two gestures, participants performed two basic actions: (1) *point-and-*

select— point to an object, select and de-select, and (2) *drag-and-drop*— point to an object, select it, drag it to a specified location, and de-select.

Procedure

Across all five experiments, participants sat in a chair 2 m away from the large display and took about two hours to complete all trials. They were situated 1.6 m away from the sensor, and the chair-seat was 53 cm high. The sensor was 89 cm from the ground with a horizontal field of view of 57 degrees, and a vertical field of view of 43 degrees. (In the second round, for experiment 6, participants sat in a couch 2.25 m away from the display and 1 m away from the sensor, and the couch-seat was 44 cm high.) In the XY plane (parallel to the display), hand movements were mapped from real space to display space as 1: 2.4 (when a participant moved 1 cm in real space, the cursor moved 2.4 cm in the display space). Before the experiment, all participants spent about 10 – 15 minutes practicing *select* and *de-select* gestures while solving a picture puzzle on the large display (see Figure A1 in Appendix A.1). Each block of an experiment began by selecting a 'Start' circle. Each trial began with a blue folder appearing on the display with a black background (Figure 4.2). To successfully complete a trial, participants either performed a point-and-select or a drag-and-drop operation on the folder. Participants were required to take at least a 10–second break in between each block. (For experiments 1 – 4, 20 trials constituted a block.) Trials were recorded using a video camera capturing users' gestures and the display. In the first round, across the five experiments, randomized partial counterbalancing was used to control order effects.

Measures

User experience was operationalized as efficiency (performance time), effectiveness (selection and de-selection error rates), and user satisfaction (users' ranking of experimental conditions and qualitative comments). We also logged the location where selection and de-selection errors occurred. Time was measured from when a folder (target) appeared on the display to when users successfully selected the folder or moved the folder to a specified location. To ensure that participants do not spend too long on any particular trial, and could complete the entire experiment, point-and-select trials were skipped after 20 seconds and drag-and-drop trials were skipped after 40 seconds. Data were analyzed only for successfully completed trials.

4.4. Experiment 1: different types of visual feedback

In WIMP-based interfaces, the mouse pointer provides visual feedback for two input states—tracking and engaged. In direct-touch paradigm, visual feedback is usually available for the engaged state (e.g., user tapping on an icon, or pinching to zoom). Touchless systems are typically one-state input devices, where users are always being tracked (Wigdor and Wixon, 2011). What kind of visual feedback should be available for touchless interactions? In this experiment, we studied three different types of visual feedback—discrete, partial, and continuous (Figure 4.1). Discrete feedback required users' explicit invocation by holding their hand stationary for 5 seconds in front of the sensor. Once discrete feedback was invoked, the touchless cursor was continually visible on the display. It would disappear after a certain period of user's inactivity. Partial feedback only visualized the target's response to user input but did not provide any visual feedback otherwise (This condition was inspired by terminal feedback in motor learning). For example, when users' hand hovered over a folder, the folder got highlighted. Though user's hand was continually tracked, no visual feedback was available at any other time. Continuous feedback did not require any explicit invocation. A touchless cursor was always visible as long as the user's hand was within the display range. Overall, continuous feedback operated similar to the mouse pointer; partial feedback operated similar to tapping an on-screen object in touch-based systems, and discrete feedback provided strict user control on the system's behavior.

Method

The experimental target-selection task was adapted from Fitts' 1D reciprocal task (Fitts, 1954). For each consecutive trial, a folder appeared at a certain amplitude (1100 pixels in display space, 29 cm in control space) left or right of the previous trial position. Experimental conditions were randomly counterbalanced. The size of the white-bordered touchless cursor was equal to the size of the target folder (256 pixels, or 163 mm). In summary, the study design was as followed: 3 types of feedback (condition) x 4 trials x 18 participants = 216 trials

For discrete feedback, participants needed to invoke the touchless cursor for each trial. The invocation time was not considered as part of their performance time. We did not evaluate dismissal of discrete feedback. The time threshold for discrete feedback was informed by our pilot studies. When previous work used lower time-out thresholds (e.g., 1 second) for selection by dwelling (Hespanhol, Tomitsch, Grace, Collins, & Kay, 2012), authors reported that users found it too sensitive, and even after considerable training users could not avoid unintentional triggering. However, we do not argue that our time-out threshold is an optimal choice. We simply wanted to measure the user experience, when participants perceived an explicit invocation of visual feedback.

Results and discussion

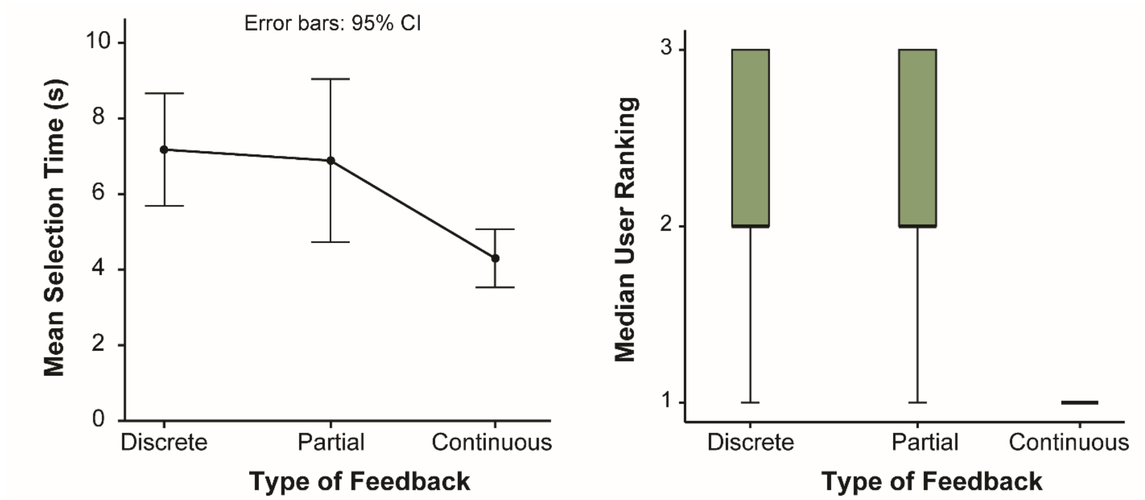

Figure 4.3. Types of feedback (discrete, partial, and continuous) significantly affected selection time and user preference. Continuous feedback was most efficient and most preferred by users.

Shaprio-Wilk test of normality showed that performance time was normally distributed, but error rates were not. A repeated measures ANOVA found that performance time was significantly affected by the type of feedback, $N = 72$, $F(2, 12) = 5.09$, $p < .05$, $\eta^2 = .46$ (Figure 4.3, left). Only successful selections were considered for data analysis; participants were unsuccessful with 51% of the trials in discrete, 75% in partial, and 21% in continuous feedback condition. Unsuccessful trials were treated as data missing completely at random (MCAR). Planned contrasts showed that participants were significantly efficient with continuous feedback ($M = 4.30$ s, $SD = .83$) compared with discrete feedback ($M = 7.17$ s, $SD = 1.61$), $p < .01$, $d = 2.24$.

A Friedman test showed significant effects of the type of feedback on error rates, $\chi^2(2, n = 19) = 16.00$, $p < .001$. Follow-up pairwise comparisons were conducted using a Wilcoxon test and Type I error was controlled using *Bonferroni-Holm* correction. Error rates were significantly more in partial feedback condition (*Mdn* = 0%, *IQR* = 50) than both in continuous feedback (*no errors*), $Z = 2.83$, $r = .65$, and discrete feedback condition (*no errors*), $Z = 2.83$, $r = .65$, ps < .01.

Each participant was asked to rank the three types of feedback according to their order of preference. A Friedman's ANOVA showed a significant effect of the type of feedback on user preference, $\chi^2(2, N = 18) = 17.56$, $p < .001$ (Figure 4.3, right). Follow-up Wilcoxon tests showed that users significantly preferred continuous feedback over discrete feedback, $Z = 3.23$, $r = .76$, and partial feedback, $Z = 2.99$, $r = .70$, ps < .01.

Among the three conditions, continuous feedback provided the best user experience, thus confirming the critical role of visual feedback in controlling touchless interactions. Although discrete feedback differed from continuous feedback only in invocation, users were less efficient with the former. Holding their hand stationary not only made users dislike discrete feedback, but also affected their efficiency. This suggests that simply holding the hand stationary may not be an ideal candidate for mode switching. However, in a touchless system, this effect would only be articulated in the first task following the mode switching. For partial feedback 7 out of 18 participants mentioned that they guessed where to point, which explains the significant decrease in their efficiency and effectiveness. This suggests that in device-less touchless interactions, point-and-select tasks on a large display cannot be guided sufficiently with proprioception.

4.5. Experiment 2: alternative shapes, sizes, and colors of the touchless cursor

A mouse pointer is an icon from a semiotic perspective (Pierce, 1931-58). By default, it resembles an arrow and signifies the concept of pointing. It may also take up other forms, such as an hour clock (to signify that the user needs to *wait* for a computer response) or a blinking vertical line (to signify the possibility of *text* input). The mouse pointer provides visual feedback for point-and-click interactions. Similarly, in touchless systems, the touchless cursor could change its form (e.g., shape, size) to provide necessary visual feedback on the ongoing status of the interaction. In this experiment, we studied three different properties of the touchless cursor—shape, size and color. But why can't we simply replicate the existing representations of the mouse pointer? Because the lack of kinesthetic feedback in touchless interactions and the inherent ambiguity with hand-gesture input requires unobtrusive yet effective visual feedback at many instances—unwarranted in point-and-click interactions (e.g., see Vogel & Balakrishnan, 2005). This makes our investigation of visual feedback in large-display touchless interactions pertinent.

We studied five shapes: circle, semi-circle, triangle, diamond, and star; three sizes: small, medium, and large; and five outline colors: green, blue, white, red and yellow. Searching the mouse pointer on a traditional desktop screen is not a pressing problem, but it is often reported that users lose track of the cursor in very large displays and multi-monitor configurations (e.g., see Baudisch, Cutrell, & Robertson, 2003). On the other hand, large displays are suited for visualizing and manipulating large datasets (Beaudouin-Lafon, 2012). Hence, it is crucial that a touchless cursor can easily be searched while interacting with information-dense displays. Our shape and color coding dimensions were inspired by a class counting study (most common visual search task) by Smith and Thomas (1964). The shapes used in this experiment are geometric forms with vertices ranging from 0-5. We conducted a pilot study to confirm the user perception of the five Munsell colors (Fig.1, p. 139, Smith and Thomas, 1964) when converted to RGB space (see Appendix A.2 for conversion details). Seven observers classified each color on the large display. Fleiss' kappa was used to measure the reliability of their agreement. All observers substantially agreed on all colors ($\kappa > .75$) except white ($\kappa = .30$). Following the analysis, we changed the white color to be described by its hex color code, FFFFFF. Small-sized cursors were bounded by a square of 128 pixels (81 mm), medium by 256 pixels (163 mm), and large by 512 pixels (325 mm). Overall, the cursors were 50%, 100% or 200% of the display object (256 x 256 pixels) that was required to be selected during the point-and-select task.

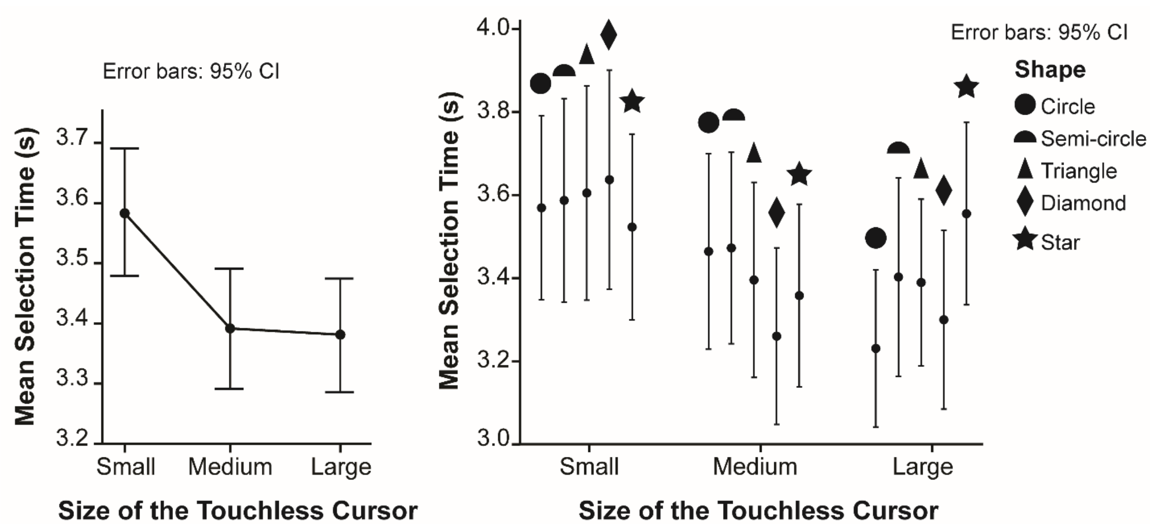

Figure 4.4. (Left) Selection time was significantly correlated with the size of the touchless cursor, $r = -.10$, $p < .01$. (Right) We found an interaction effect of shape x size on selection time. The increase in number of corners did not increase efficiency across all sizes of touchless cursors.

Method

For this experiment, we used the same target-selection task as experiment 1. Visual feedback was continuously present. The touchless cursor was not filled with any solid color. All experimental conditions were randomized across trials. In summary, the study design was as followed: 5 shapes x 3 sizes x 5 colors x 4 trials x 18 participants = 5400 trials.

Results and discussion

Among the three independent variables (shape, size, and color), we only found a significant correlation between the size of the touchless cursor and performance time, $r = -.10$, $p < .01$ (Figure 4.4, left). No main effect of shape, size, or color was found on participants' efficiency or effectiveness. We only found an interaction effect of shape x size, $F(8, 184) = 2.15$, $p < .05$, $\eta^2 = .09$. Increase in the number of corners did not increase efficiency across all sizes, which explains the interaction effect (Figure 4.4, right). No significant performance benefit of the large-sized cursor was found over the medium-sized cursor, but 10/18 participants reported preference for large-sized cursors. Nine out of 18 participants preferred circular cursors. No color preference was reported.

Our results suggest that a touchless cursor of size equivalent to display objects (equal bounding areas) provides an optimal user experience, and an increase in cursor size do not improve user performance. We did not find any significant effect of shape or color coding of the touchless cursors. Overall, participants reported their preference for symmetrical shapes. A limitation of this study was the simplicity of the selection task, and a non-distracting background. Future research on the effects of shape and color of touchless cursors should investigate complex scenarios, where the display already contains artifacts of different shapes and colors.

4.6. Experiment 3: alternative levels of transparency of the touchless cursor

Researchers have found that different levels of transparency of user interface elements, such as a tool palette, affect users' selection time (Harrison, Kurtenbach, & Vicente, 1995). In this experiment, we investigated user experience for different levels of transparency (100%, 50%, 25%, and 0%) of the touchless cursor. The level of transparency affected the fill of the touchless cursor, but not its outline.

Method

We used the selection task from experiment 2. The touchless cursor always had a white outline, and was equal to the size of the target folder (256 pixels, or 163 mm). Different transparency levels with the base color white were randomized across trials. In summary, the study design was as followed: 4 transparency levels x 4 trials x 18 participants = 288 trials.

Results and discussion

Performance time or error rates were not significantly affected by levels of transparency, *p*s > .05; but user preference was significantly affected (Figure 4.5). Each participant was asked to rank the four types of touchless cursors according to their order of preference. A Friedman's ANOVA showed a significant effect of transparency on user preference, $\chi^2(3, N = 18) = 18.17$, $p < .001$. Follow-up Wilcoxon tests showed that users significantly preferred medium transparency (50%) over low-transparent (25%), $Z = 3.56$, $r = .84$ and opaque touchless cursors, $Z = 3.06$, $r = .76$, ps < .01.

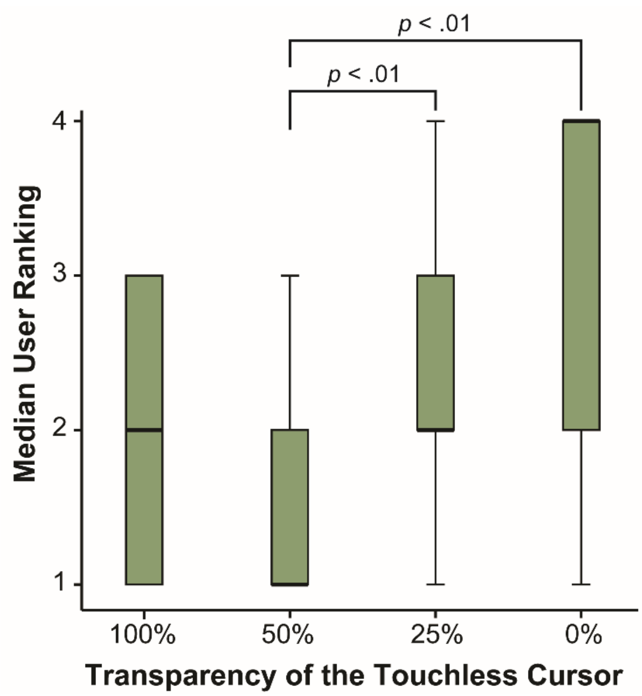

Figure 4.5. User preference of touchless cursors was significantly affected by their level of transparency. Participants significantly preferred medium transparency (50%), both over low transparency (25%) and opaque touchless cursors.

Participants mentioned that they disliked the opaque touchless cursor because it obstructed the view of the display object, but a 50% transparent touchless cursor was equally preferred as a completely transparent touchless cursor (with only an outline). This is an important finding since we are used to an opaque mouse pointer in desktop environments, but the mouse pointer is significantly smaller than the icons, thus not producing the obstruction problem that participants faced in this experiment. As we found in experiment 2, having a touchless cursor smaller than the display icon reduces user's selection efficiency.

4.7. Experiment 4: alternative approaches to represent selection

The touchless cursor should not only inform users where they are on the display, but also what they are doing. How can we represent operations (e.g., selection, de-selection) using the touchless cursor as a 'sign vehicle'? This is particularly important because of the absence of any kinesthetic feedback in touchless interactions that is conveniently available with a mouse or on a touch surface. In this experiment, we investigated different approaches to represent target-selection: change in the cursor's shape (circle to semi-circle, semi-circle to triangle, triangle to diamond, and diamond to star), change in depth (sphere to circle, and circle to sphere), and change in transparency (0% to 100%, 100% to 0%, 50% to 25%, and 25% to 50%). For example, when hovering over a folder, a user would see a circular touchless cursor, a successful select gesture would transform the cursor into a semi-circle, and a successful de-select gesture would convert the cursor back to a circle.

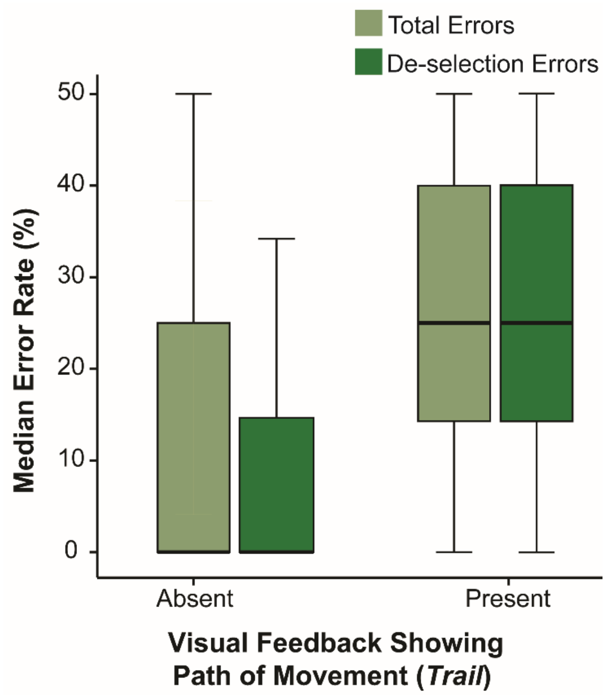

Figure 4.6. Participants made significantly more errors when Trail was present compared with no Trail condition, $p < .05$, $r = .50$.

Method

We used the selection task from experiment 2. The touchless cursor always had a white outline (except for depth changes, where the cursor was filled white), and was equal to the size of the target folder (256 pixels, or 163 mm). In summary, the study design was as followed: 10 cursor transitions x 4 trials x 18 participants = 720 trials.

Results and discussion

Performance time or error rates were not significantly affected by different cursor transitions, *p*s > .05. Although most participants could not report clear ranking preferences for the 10 cursor transitions, overall they reported that a change of opacity was more informative and less distracting than a change in shape or depth. Ten out of

18 participants liked cursor transitions to represent target-selection. One participant commented, *"I felt I am accomplishing something. It made me feel good."*

4.8. Experiment 5: persistent visual feedback for *drag-and-drop* operations

All interactive systems are affected by some amount of lag: a delay between users' input and the visualized response. Working with multitouch systems, Wigdor et al. (2009) reported that such lag reduces users' perception of reactivity of the system, and designed a *trail* visualization that renders behind a finger as its contact point moves from one position to another. Large-display touchless interactions are device-less. With no surface friction of any device, the user moves faster, and with a larger screen the delayed reactivity of the system becomes a significant problem. Moreover, without any tactile feedback, the user solely depends on proprioception to perceive their path of movement. Since continuous visual feedback controls motor responses (see section 2.3.1), this lack of immediate visual feedback can affect operations where users are dragging an object on the display. In this experiment, we evaluated *trail* – persistent visual feedback that echoes the immediate history of user's hand positions (for a pre-defined time window). A trail was visualized as a Bézier spline (using cubic Bézier curves) along 10 previously tracked hand positions.

Method

The experimental task was a drag-and-drop operation. For each trial, participants moved a folder across the display (2000 pixels in display space, 53 cm in control space) to the left or to the right. The white-bordered touchless cursor (equal to the size of the target folder, 256 pixels) was filled with solid white, when a successful select gesture was interpreted; and the trail was visualized as a yellow line (Figure 4.1). In summary, the study design was as followed: 2 directions x 3 blocks of repetitions x 18 participants = 108 trials.

Before this experiment, participants practiced drag-and-drop operations in 8 compass directions (1100 pixels in display space, 29 cm in control space) for 3 blocks of repetition (Figure 4.9 shows the de-selection errors during those practice sessions).

Results and Discussion

Shaprio-Wilk test of normality showed that neither performance time, nor error rates were normally distributed. The presence of trail did not significantly affect performance time; but error rates were significantly more with trail present (*Mdn* = 25%, *IQR* = .28) than without trail (*Mdn* = 0%, *IQR* = .29), $n = 17$, $Z = —2.08$, $p < .05$, $r = .50$ (Figure 4.6). Specifically, trail did not affect error rates for selection, but de-selection errors were more with trail present (*Mdn* = 25%, *IQR* = .33) than without trail (*Mdn* = 0%, *IQR* = .14), $n = 17$, $Z = —2.20$, $p < .05$, $r = .53$. Participants commented that the continuous updating of the trail was distracting and exacerbated the natural tremor in hand motions.

Unlike in device-based interactions (such as with touch), hand movements in mid-air are rarely smooth—they frequently create a convoluted trail, thus distracting rather than supporting the user's task at hand. Moreover, the echo feedback provided information not entirely relevant to users' task at hand. Our results suggest that a trail significantly affected participants' effectiveness, mainly while dropping objects on the display (de-selection errors). Why selection was not equally affected by trail may be explained by the inherent difficulty of the de-select gesture (for details see additional observations, Figure 4.9). Based on participants' comments, video recordings, and logged data, we re-designed trail: A straight line joining the initially selected position to the user's current hand position (Figure 4.10, bottom-left).

4.9. Experiment 6: persistent visual feedback for out-of-range events

In large-display touchless interactions, when the sensor's tracking range does not match with the system's display range, a gap is created between the system's behavior and the user's mental model. This happens when users perform a gesture that erroneously steps out of the display range. During our pilot studies in the first round of experiments, we observed that when participants' gestures go off the display and the touchless cursor becomes unavailable, participants stop and get disoriented. They do not further attempt to move their hands and return within the display range. In the absence of any visual feedback, users fail to perceive that they are still being tracked by the sensors. From our observations, we hypothesized that participants halted because they perceived a lack of feedback as an error, and their reaction to an error was to slow down, a well-known phenomenon called post-error slowing (Notebaert et al., 2009).

Based on our hypothesis, we iteratively developed and tested *Stoppers* (Figure. 4.1), a type of semantic feedback (p. 83, Wigdor & Wixon, 2011) that uses the metaphor of stoppers (or plugs) to inform users that the system is still tracking their gesture, thus giving them the opportunity to instantly step back within the display's range. Stoppers support this action by providing visual feedforward (direction to move) and visual feedback (user's current position). When users gesture within a display range, a touchless cursor (such as a circle) is available. When users go off the display range, a semi-circle appears at the last-recorded within-display position of their gesture. In our current visualization of *Stoppers*, the change in feedback from a circle to a semi-circle subtly informs users that they are out of the display range and need to retrace their way back (see Figures A2 and A3 in Appendix A.3 for a detailed visualization). Stoppers disappear as soon as the user is back within the display range. During pilot studies in the first round of experiments, users found *Stoppers* intuitive and helpful (Chattopadhyay, Pan & Bolchini, 2013). In the second round, we systematically investigated the effect of stoppers on user's efficiency in returning within the display range.

Method

For this experiment, participants pointed to a target object (a text label or a display icon of size 256 pixels) appearing randomly at certain positions at the top, left or bottom border of the display (see Figure A4 in Appendix A.3 for a description of the experimental task). Because of the difficulties of our de-select gesture in the previous round of experiments, we decided to use a pointing task. To successfully complete a trial, participants pointed to the target object with a white-bordered touchless cursor (sized equal to the target). In summary, the study design was as followed: 14 target positions x 5 blocks x 19 participants = 1330 trials

Results and Discussion

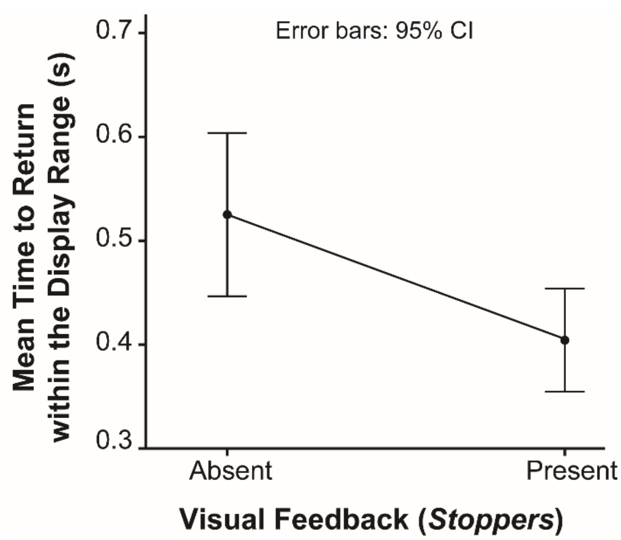

Figure 4.7. Participants were significantly faster in returning within the display range with Stoppers present than without Stoppers, $p < .01$, $d = .87$.

Participants were significantly faster in returning within the display range with stoppers present (M = 411 ms, SD = 104) than without stoppers (M = 533 ms, SD =169), $t(18) = 2.97$, $p < .01$, $d = .87$ (Figure 4.7). Participants also reported stoppers as a non-distracting, helpful guide to keep them within the display's range and to help them retrace their steps back.

Our results from experiments 5 and 6 confirm that the type of visualization plays a key role in visual feedback: relevant and semantic visual feedback seems to be more effective than echo feedback in large-display touchless interactions.

4.10. Additional findings

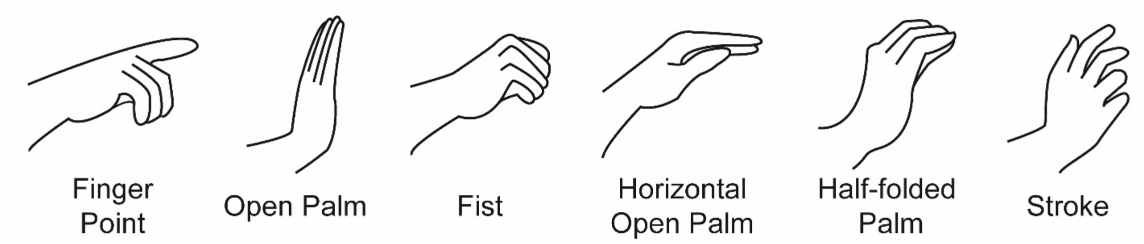

Figure 4.8. While using the select gesture, participants spontaneously created and used a rich range of hand poses.

Apart from our six controlled studies we made two interesting observations: one throughout the first round of the experiment, and another during the drag-and-drop practice sessions. Since our gesture primitives and hand tracking algorithm was agnostic of participants' hand poses, we encouraged participants to user their preferred hand pose. Across all experiments, we observed a rich paradigm of spontaneous gesture variations that participants created to perform touchless selection (Figure 4.8).

Throughout our first five experiments, we used two gesture primitives: *select* and *de-select*. While a select gesture was defined as a forward movement of the hand with a certain velocity, a de-select gesture was defined as a backward movement of the hand by a certain distance (Figure 4.1). During the drag-and-drop practice sessions (prior to experiment 5), participants performed de-selection at 8 different locations of the display. We observed an interesting phenomenon: While participants intended to move backward from the sensor (in Z-direction), they actually moved *down* vertically (during de-selecting objects in northern regions, such as NW, N, or NE) or moved *up* vertically (during de-selecting objects in southern regions, such as SW, S, or SE) (Figure 4.9). Overall, there was a strong trend among participants to bring their hand closest to the center of their torso, probably for energy conservation. An inverse, but related phenomenon was reported by researchers while using push-to-select gestures on large displays (Hespanhol et al., 2012): While translating from one position on the display to another (parallel to the display), users often moved their hands forward (orthogonal to the display), and accidentally invoked the select gesture.

4.11. General discussion

We conducted six controlled experiments to explore four different aspects of visual feedback in large-display touchless interactions. Specifically, we investigated: types of feedback, alternative forms of touchless cursors, alternative approaches to visualize target-selection, and persistent visual feedback for drag-and-drop operations and out-of-range events. Although we studied visual feedback using a point-and-select task, our findings are applicable beyond our experimental tasks. In the following sections, we discuss how our findings can be extended to inform the design of visual feedback for touchless interactions with large displays. To frame our discussion properly, it is important to note two different kinds of large-display touchless interactions: An interaction that happens in the context of a display object (e.g., using a marking menu to operate on an icon, Bailly et al., 2011), and an interaction that is object-agnostic (e.g., making a *teapot* gesture to create an avatar; Walter et al., 2013). Our findings and design guidelines are relevant to object-oriented touchless interactions that require users to point to a display object prior to any gesture invocation.

Design Implications

First, our findings suggest that *continuous* visualization of users' current position on the display—independent of an application's response to user input—is crucial for touchless interactions. The designer may choose to represent tracking information corresponding to one or more body parts depending upon the interaction vocabulary in use. For example, a touchless system allowing two-hand manipulations would require visual feedback for both hands; a system allowing foot interactions should further represent tracking information of users' feet. Visual feedback of an application's response does not provide enough feedback to users before any successful gesture registration or during gesture relaxation (Wu, Shen, Ryall, Forlines, & Balakrishnan, 2006). For example, an application allowing users to rotate 3D images bimanually in a sterile environment should show the hand positions in addition to the rotation of the object as a result of users' hand movements (similar to Rosa & Elizondo, 2014).

Second, a touchless cursor can be efficiently used as a 'sign vehicle' to represent many critical aspects of touchless interactions, such as when a user is engaged in an on-going interaction or when multiple users are collaborating synchronously. Our results suggest that shape or color coding of touchless cursors do not significantly affect user experience in large-display touchless interactions. Yet, users informally commented on their preference toward symmetrical shapes. Hence, colors may be used to distinguish multiple users interacting at a time, while shapes may be used to represent different interaction states (e.g., when the user is clutching instead of interacting).

We found that a touchless cursor of size equivalent to a display object is significantly more efficient than a smaller cursor (50% of the display object), but not significantly less efficient than a larger cursor (200% of the display object). While using a cursor equivalent to the size of a display object, users disliked an opaque cursor, but significantly preferred a slightly transparent touchless cursor (50% opacity). The applicability of our results on the size of the touchless cursor may be limited by our gesture primitives. Nevertheless, similar to shape coding, our results on transparency can be applied to represent a touchless cursor during an interaction. For example, multiple users reported envisioning a scenario where during touchless selection the cursor would transform from an outline to a transparent fill to represent a successful select gesture, and revert to its default outline when deselected. Although we explored different transitions of the touchless cursor to represent touchless selection (experiment 4), no particular condition emerged as significantly more efficient or effective. Still users reported a preference for transparency changes and mentioned that shape transitions were distracting.

Most current systems use the icon of an open hand as a touchless cursor, and transform the icon to a closed hand or corresponding poses (such as finger counts) on successful pose recognition (Microsoft, 2013). This visual feedback technique may not be scalable for a collaborating environment. Our results can be used to augment the visual feedback along with pose information in collaborative touchless environments. For example, let us imagine a collaborative touchless environment that uses both hands and feet toward performing gestures. Multiple users may be color coded. Hands and feet may be distinguished using shape coding (or iconic images). The touchless cursors can appear as outlines while users are being tracked, but are not engaged. On successful gesture recognition, a touchless cursor may simply be filled with a certain level of transparency, or an iconic image of the pose can be transparently overlaid on the cursor.

Third, persistent visual feedback can benefit touchless operations that are affected by users' fast and large movements. When users erroneously gestured out of the display range, *Stoppers* significantly increased their efficiency in returning within the display range (experiment 6). However, trail—persistent visual feedback that echoed users' path of movement during drag-and-drop operations decreased users' efficiency (experiment 5). Users reported them as distracting and redundant. While stoppers provided users with *semantic feedback* (a meaningful representation of the system's knowhow about the user), trail provided *echo feedback* (an echo of minimally processed sensor data; p. 83, Wigdor and Wixon, 2011). Although further research is required to make a more general claim, semantic feedback seems to be more effective than echo feedback in large-display touchless interactions. Our findings suggest that persistent visual feedback in large-display touchless interactions should be: (1) visually unobtrusive, (2) salient, and (3) communicate only relevant information for the ongoing interaction. Based on these guidelines, we redesigned trail from a cubic Bezier curve to a simple straight path connecting the initial selection position during a drag-and-drop operation and the current position of the user's hand.

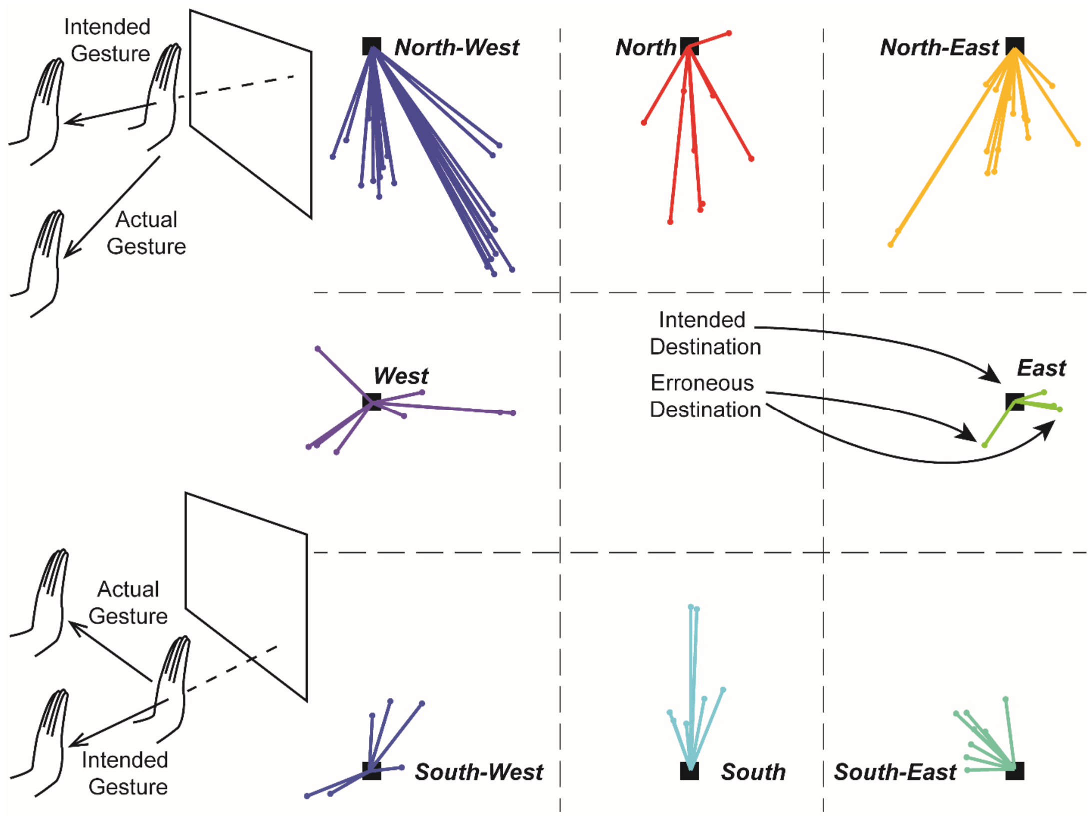

Figure 4.9. During drag-and-drop practice sessions, participants moved display objects in 8 directions (N, S, W, E, SW, SE, NE, and NW). We found an interesting pattern in the de-selection errors across different positions of the display: While moving backward from the sensor (in Z-direction), participants often moved down vertically (during de-selecting objects in northern regions, such as NW, N, or NE) or moved up vertically (during de-selecting objects in southern regions, such as SW, S, or SE). Overall, there was a strong trend among participants to bring their hand closest to the center of their torso, probably for energy conservation.

Additionally, we discovered a caveat about touchless gesture primitives that parametrize orthogonal movements. Our video recordings and logged data of users' de-selection errors showed that users always tend to follow the shortest path toward the center of their torso, rather than orthogonal movements (Figure 4.9). While performing de-select gestures, users frequently moved vertically downwards (or vertically upwards) while intending to move only orthogonal to the large display. This observation well aligns with the minimum energy cost model of human movement planning (Alexander, 1997); it states that while reaching an object, among infinitely many paths, we choose the one path that minimizes our metabolic energy cost. This phenomenon is most relevant for *large-display* touchless interactions, where to interact with display objects users stretch their hands beyond the space directly in front of their torso—up, down, left or right.

Overall, our findings suggest that given the large size of the display, and the lack of haptic feedback in touchless interactions, effective visual feedback plays a key role in improving the touchless user experience with large display interfaces. When proprioception is the only feedback for an interaction modality, visual cues can somewhat compensate the lack of haptic feedback. This work provides the first step toward building a visual feedback language for touchless interactions.

Finally, to crystallize in a coherent view the lessons learned across our six experiments, we propose a visual feedback routine for a simple interaction scenario: moving a folder using a drag-and-drop operation (Figure 4.10). We envision the large-display touchless system in three interaction states: idle, active, and engaged. In the *idle* state, though users are being tracked by the motion sensor, they cannot interact with the system; for example the user may be out of the display range, or clutching. In *active* state, users are interacting with the system (e.g., pointing), but *not* performing any action, such as selecting, dragging, or resizing. In *engaged* state, users either make a gesture to initiate an operation, or continue an ongoing operation; the system in this state would register a gesture, allow the user to continue a gesture, or recognize gesture termination. In our visualization instance, we provide *stoppers* to represent when users are out of the display range (Figure 4.10.a); a circular, unfilled touchless cursor to show users' position on the display (Figure 4.10.b); a partially filled (50%) touchless cursor to indicate that selection has been registered (Figure 4.10.c); and a trail to provide semantic context of the ongoing drag-and-drop operation (Figure 4.10.d). When users complete the drag-and-drop operation, the touchless cursor would change back to its default state, and indicate that de-selection has been registered. This simple *idle-active-*

engaged model provides a preliminary framework to conceptualize interactions and their corresponding visual feedback routine in a touchless system.

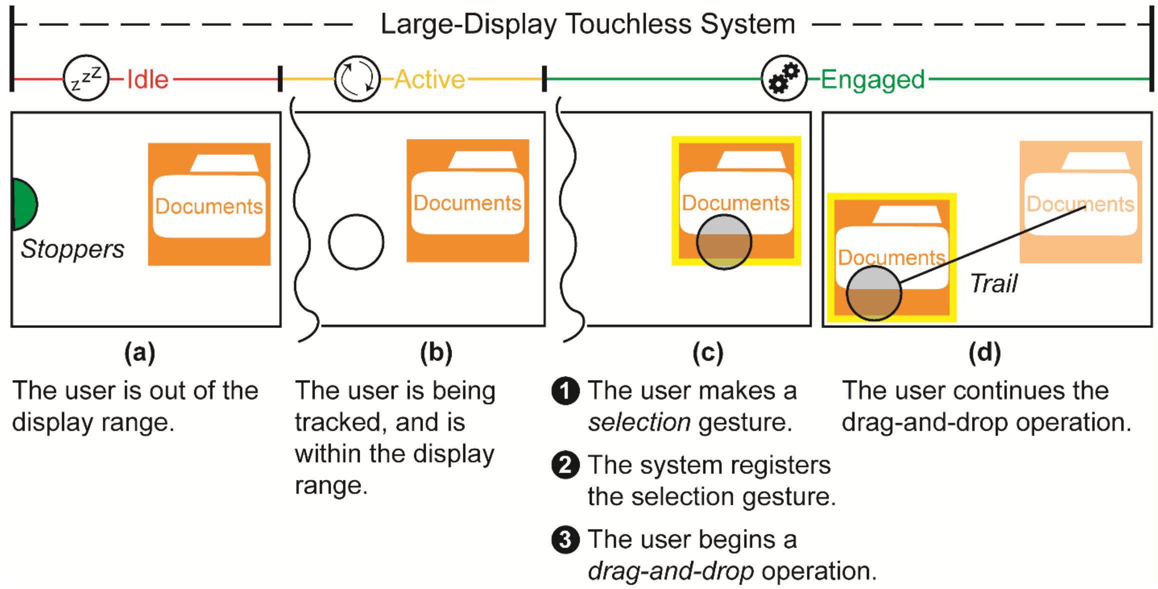

Figure 4.10. Demonstrating visual feedback for the three interaction states—idle, active, and engaged—during a drag-and-drop operation: (a) Stoppers represent when users are out of the display range; (b) a circular, unfilled touchless cursor shows users' position on the display; (c) a partially filled (50%) touchless cursor indicates that selection has been registered; and (d) a trail provides semantic context of the ongoing drag-and-drop operation.

Limitations

The capability of our motion tracking sensor limits our findings. It operated with a maximum refresh rate of 30 fps: Users perceived a lag of about 33 ms between their movements and screen update. In our experimental setup, participants sat in a comfortable chair. This may have affected their ability to make certain gestures; but neither did we observe any ergonomic constraints, nor was reported by the participants. Moreover, our participants were right-handed. Although we do not think that this would affect our findings on visual feedback, we cannot claim a generalization of our findings across left-handed users.

We investigated visual feedback using only *select* and *de-select* gestures. Our performance measures may be biased by the gesture primitives we used in the experiment, and further research is necessary to tailor visual feedback to any particular interaction vocabulary. Our experimental system received a mean SUS score of 66 that suggests an average usability; but we did not record any subjective ratings for intuitiveness. Informally, users did not report any significant physical strain after the experiment. Based on current research, future studies should record user fatigue using objective measurements, such as consumed endurance (Hincapié-Ramos, Guo, Moghadasian, & Irani, 2014). Users' difficulty in performing the de-select gesture (Figure 4.9) was obvious during the practice trials; but that may not have significantly affected the experimental trials (in experiment 5) because participants only performed select and deselect gesture at their chest-level (when seated).

Our experiment used a simple point-and-select task, and a solid black background. Most real world tasks are complicated, and the display background is populated with other artifacts. Future research investigating visual feedback in large-display touchless interactions should use the display density of the background as an experimental factor. More complex tasks, such as matching, sorting or grouping of display objects may be used.

Though we provide some guidelines on how to design visual feedback for multiple users interacting simultaneously, future experiments—controlled or in-the-wild—are required to identify their role in collaborative touchless environments. Moreover, we did not investigate the aspect of clutching in touchless interactions. It is important to investigate how visual feedback can intuitively allow users to reposition their body parts without affecting the screen output.

External validity

Our results are generalizable for large-display touchless interactions. Specifically, our findings about different types of visual feedback (experiment 1) and observations about de-select gestures (Figure 4.9) may not apply in gaming scenarios where users interact with standard television screens, such as 50" HDTVs, from a 7-9 feet distance. This is because in such scenarios the operating region of user's motor space (also known as user's control space) is much smaller compared with while interacting with larger displays. (Shrinking the motor space in large-display interactions—using a very high control display gain—would lead to quantization errors.) Although users were seated in our experiments, we expect our findings to stay valid in a standing posture. Visibility depends on the distance from the display. Our experiments were conducted at a fixed distance from the large display. Though distance from the display may affect the task efficiency of the users (since display objects get smaller), it is unlikely to affect our general findings on visual feedback. Finally, our design guidelines are agnostic of the control-display gain of the system, or how the control space is mapped to the display space. For our study, we used an off-the-shelf sensor inside a room with normal levels of fluorescent lighting. Outdoor lighting may affect the tracking noise, the screen glare, and the perception of color coding.

4.12. Conclusions

Touchless interactions lack haptic feedback, but effectively designed visual feedback can guide users to control their movements and still perform operations efficiently. Because large displays are often densely populated with artifacts, visual feedback in large-display touchless interactions should be easily perceivable. Motor science research suggests that visual feedback can improve motor control and learning; studies on visual perception present attributes that can be used to facilitate users' attention in visual search. Inspired by the potential of visual feedback in related fields, we systematically investigated types of feedback, alternative forms of touchless cursors, approaches to visualize target selection, and persistent visual feedback during drag-and-drop operation and out-of-range event.

Our findings suggest that continuous visual feedback is significantly effective than partial feedback; users' efficiency did not increase with their cursors increasing beyond the size of the display objects (200%); and users preferred slightly transparent (50%) cursors over completely opaque ones. We also found that semantic feedback located at the border of the display (Figures 4.1, A3 and A4) informing users when they were out of the display range helped users to return efficiently; but echo feedback showing the path of users' movement made users inefficient during drag-and-drop operations. We additionally observed users making a wide range of hand postures during touchless selection. We also found that orthogonal movements as interaction primitives are limited: users obviously take the shortest path toward their torso, thus misfiring touchless gestures.

This work aligns with the research on imaginary interfaces that show users can reliably perform spatial interaction using bare-hand movements without any visual feedback (Gustafson, Bierwirth, & Baudisch, 2010), or eyes-free distal pointing (Cockburn, Quinn, Gutwin, Ramos, & Looser, 2011). Instead, our work puts forth the importance of visual feedback in effectively controlling touchless interactions with large displays—where the display space is entirely decoupled from the motor space. The overarching contribution of our work is to confirm the key role of visual feedback in touchless interactions, and providing some early pointers on how the design of visual feedback can somewhat compensate the lack of haptic feedback. Future research on visual feedback needs to mine specific requirements in different interaction scenarios, such as swiping-to-type on a keyboard, crossing-to-select a menu, or making finger poses to trigger commands. These requirements related to motor control, motor learning, and visual attention can then guide the design of a visual feedback language for those interaction scenarios. Another direction of research is—given our dependency on visual perception for triggering motor responses in touchless interactions—what other phenomena that affect visual perception (e.g., Gestalt principles) also affects touchless user experience. This is explored in Chapter 9.

Chapter 5. Affordance and ability

Elicitation and evaluation studies explore intuitive touchless gestures but do not operationalize intuitiveness. For example, studies found that users fail to make accurate 3D strokes as interaction commands. But this phenomenon remains unexplained. In this chapter, we first explain how making accurate 3D strokes is generally unintuitive because it exceeds our sensorimotor knowledge. We then introduce motor-intuitive, touchless interaction that uses sensorimotor knowledge by relying on image schemas. Specifically, we propose an interaction primitive—mid-air, directional strokes—based on space schemas, up-down and left-right. Finally, we present results from a controlled study, where users interact with large displays using directional strokes.

In sum, this chapter operationalizes intuitive touchless interaction and demonstrate how user performance of a motor-intuitive, touchless primitive based on sensorimotor knowledge (image schemas) is affected by biomechanical factors.

5.1. Operationalizing intuitiveness in touchless interactions

To explore *intuitiveness* (or naturalness) in touchless interactions, researchers mostly follow either of these two approaches: gesture elicitation (e.g., Aigner et al., 2012, Vatavu & Zaiti, 2014) or gesture evaluation (e.g., Ren & O'Neill, 2012). For example, a gesture elicitation study reported that users would prefer dynamic “wiping” hand movements over a static hand posture (e.g., a certain combination of fingers) to trigger a “delete” action (Grandhi, et al., 2011). In a gesture evaluation study, researchers found that users evaluated “dwelling” as the most intuitive gesture to select a target (Hespanhol, et al, 2012). Neither of these existing approaches to investigate intuitiveness of touchless interactions operationalizes the concept of intuitiveness. Therefore, we often encounter observations from evaluation studies about the poor performance of certain gestures without any proper explanation. For example, a common touchless interaction primitive to indicate “selection” uses dynamic gestures, where meaning is assigned to particular translations (i.e., hand movements) in space. Recent works examining this interaction primitive (Guimbretière, & Nguyen, 2012; Ren, & O'Neill, 2012) report users’ limitations in *making precise hand trajectories in 3D space.* Despite repeated observations of this phenomenon, we still lack a causal explanation.

We argue that to explain the potential and limitations of current touchless primitives, we need to consider the level of knowledge that is being used in such interaction contexts. The level of knowledge at play while interacting with computers is classified into a *continuum of knowledge* by the intuitive interaction framework (Blackler

& Hurtienne, 2007). In this continuum, the level of intuitiveness of the interaction grammar is inversely proportional to the artificiality of the knowledge that a user relies on to interact with. Intuitive interaction is thus characterized as the extent to which users' unconscious application of prior knowledge leads to effective interaction (Hurtienne & Israel, 2007). In the case of touchless interactions, designers often treat human abilities as a "black box", assuming that our ability to interact with the physical world *directly translates* into our ability to perform exact gestures in space. Yet, intuitive interaction does not work in this way. To unleash intuitive user experiences, designers need to examine the relationship between a given level of knowledge and the corresponding interaction primitives that align well with that knowledge.

The main contribution of this chapter is to introduce the concept of motor-intuitive, touchless interactions. Specifically, we propose and evaluate a novel, motor-intuitive, touchless interaction primitive—mid-air, directional strokes—based on space schemas: *up-down* and *left-right.* To investigate how other factors, such as biomechanical properties of the human body, affect the performance of our proposed motor-intuitive touchless primitive, we conducted a controlled experiment. As per the intuitive interaction framework, motor-intuitive interactions have the potential to establish a new touchless interaction grammar that is based on what users can accomplish without further cultural or advanced expertise. Our work makes the following contributions:

- We provide a theoretical explanation of human limitations in making accurate 3D trajectories (section 5.3) by drawing an analogy between 'reaching for an object' and freehand gesturing toward a display. This explanation is based on the consideration of the sensorimotor level in the continuum of knowledge that is at play during such interactions. We further discuss how the lack of feedback in touchless interactions can also explain such motor limitations.
- We introduce *motor-intuitive,* touchless interactions based on image schemas. Specifically, we propose a touchless interaction primitive that draws on the sensorimotor level of knowledge—the two space schemas, *up-down* and *left-right* (section 5.4).
- Finally, we investigate how biomechanical factors affect user performance of our proposed interaction primitive. Grounded in our empirical results, we provide practical design guidelines for intuitive touchless interactions and large-display touchless interactions (section 5.7). These include pointers on designing dynamic

touchless gestures, characterization of right-handed users' control space based on user performance, and implications for designing UI elements for large displays (e.g., touchless menus).

Our work is a first step toward applying the continuum of knowledge in intuitive interaction to define touchless interaction primitives. Our findings can inform fundamental design decisions to align touchless user interfaces with human sensorimotor abilities, thus making them intuitive to use.

5.2. Background

While designing gesture primitives for touchless interfaces—often referred as a kind of Natural User Interface (NUI)—existing studies associate the same meaning to 'natural' and 'intuitive' (Aigner et al., 2012; Hespanhol et al., 2012; Grandhi et al., 2011; Lee, 2010; Morris, 2012; O'Hara et al., 2013; Vatavu & Zaiti, 2014; Wigdor & Wixon, 2011). The meaning of 'natural' or 'intuitive' (these terms are used interchangeably in this dissertation) that is adopted by these studies does not go beyond the vernacular definition of *instinctive* or *spontaneous*. Our work is an attempt to operationalize 'intuitive' in touchless interactions and builds upon the crossroads of two research areas: intuitive interaction and natural user interfaces.

Intuitive interaction

The intuitive interaction framework defines intuitive interaction (or intuitivity) as the extent to which users' unconscious application of prior knowledge leads to effective interaction (Blackler & Hurtienne, 2007). While a similar framework, reality-based interaction (Jacob et al., 2008), identifies core themes (such as naïve physics or body awareness and skills) to *scope* what can be called real (or natural), intuitive interaction framework provides a continuum of knowledge to *classify* intuitivity (Hurtienne & Israel, 2007). This bottom-up continuum of knowledge classifies intuitive interaction according to four different levels of prior knowledge: innate, sensorimotor, culture, and expertise. According to this continuum, the higher an interface requires specialization of knowledge the lower is the expected speed of knowledge retrieval, and hence less intuitive to use. Although this continuum of knowledge has been used to propose tangible interaction primitives (Hurtienne & Israel, 2007), use of this continuum in touchless interaction remains largely unexplored. According to this continuum of knowledge, touchless primitives drawing on the sensorimotor level of knowledge would be far more intuitive to use than primitives based on the expertise level.

Natural user interface

Many ongoing debates stem from the term *natural* in natural user interfaces (NUIs) (Norman, 2010; O'Hara, 2013; Wigdor & Wixon, 2011). NUIs promise to offer an intuitive interface modality, one that does not require users to develop special skills for communicating with computers, but allows users to use their natural abilities. But what is natural (or intuitive or like real-world) for users? Norman (2011) discussed that the notion of naturalness in a user interface is not an axiomatic truth, but achieved through sufficient feedback, effective feedforward, and perceived affordances. O'Hara et al. (2013) discuss how naturalness of an interaction modality, such as touchless, is derived from the actions it enables in different communities of practice and settings (the interactional perspective). According to Wigdor & Wixon (2011, p. 9), natural is a design philosophy that enables an iterative product-creation process, rather than a mimicry of the real world. Overall, there is an urgent need to understand what is natural for users, and then leverage it toward building NUIs.

In touchless interaction, elicitation and evaluation studies on hand gestures continue to inform the naturalness of interaction primitives. For example, empirical studies have shown that unguided mid-air gestures—especially circular in design—are generally less efficient and more fatiguing than linear gestures (Nancel, et al., 2011). Grandhi et al. (2011) reported user preference toward bimanual gestures and intuitiveness of dynamic gestures (iconic representation of the motion required for the manipulation) over static iconic hand poses. Different kinds of hand gestures have also been evaluated as command selection techniques, such as push (Hespanhol et al., 2012), grab, finger-count (Bailly, Walter, Müller, Ning, & Lecolinet, 2011), mark (Guimbretière & Nguyen, 2012; Ren, & O'Neill, 2012), or roll-and-pinch (Ni, McMahan, & Bowman, 2008). While these studies report certain gestures to be intuitive compared with others, they do not classify their intuitivity or provide an explanation about why other gestures failed to be intuitive (performed poorly). We argue that the continuum of knowledge in intuitive interaction can operationalize the intuitiveness of touchless interfaces by informing the design of touchless interaction primitives, which are the building blocks of any interaction language (Wigdor & Wixon, 2011, p. 116).

5.3. Touchless interaction primitives and our limitation to perform accurate 3d trajectories

Human gesturing has been used in different application domains of HCI for over 50 years. In 2005, Karam and Schraefel provided a high-level classification of human gestures according to gesture styles, input technologies, output technologies, and application domains. Since 2010, with recent advancements in markerless tracking, mid-air gestures are being increasingly used as interaction primitives in touchless interaction. To classify the physical mechanics of these gestures, we build upon the taxonomy proposed by Vatavu & Pentiuc (2008) (Figure 5.1). Vatavu and Pentiuc classified hand gestures into four categories: static simple, static generalized, dynamic simple and dynamic generalized gestures. Static simple gestures are gestures that only involve the use of a single posture over a certain period of time (e.g., a closed hand, Bailly et al., 2011). Static generalized gestures are gestures that involve a series of consecutive postures over certain periods of time (e.g., rolling the wrist and pinching, Ni et al., 2008; or finger movements, Vogel & Balakrishnan, 2005). Dynamic simple gestures are gestures that use information about the underlying motion trajectory but not the posture information (e.g., drawing shapes or characters in mid-air, Gustafson, Bierwirth, & Baudisch, 2010; or performing accurate 3D strokes to invoke commands in a 3D marking menu, Ren & O'Neill, 2012). Dynamic generalized gestures are gestures that use the information about both the motion trajectory and the posture (e.g., select by moving an open palm *normal* to the display, Hespanhol et al., 2012; or pinch and 3D stroke, Guimbretière & Nguyen, 2012). Each of these four categories of gestures is defined as a function of time. Hence, we call this a temporal classification.

Mid-air gestures as interaction primitives can also be classified from a spatial perspective—describing the relationship between the position of the gesture in the input space and the UI (user interface) elements in the display space. Spatially, a gesture can be referential or non-referential. Referential gestures are gestures that use the spatial information along with posture and/or motion trajectory. For example, to select an icon with a *reach* gesture users need to move across the icon's boundary (Ren & O'Neill, 2012); or to select using a *dwell* gesture users need to point to an object and hold their open palm (Hespanhol et al., 2012). Non-referential gestures are gestures that do not use any spatial information but only the posture and/or motion trajectory (e.g., touching the hip in StrikeAPose, Walter, Bailly, & Müller, 2013; or making a posture for entering a letter, Sridhar, et al., 2015).

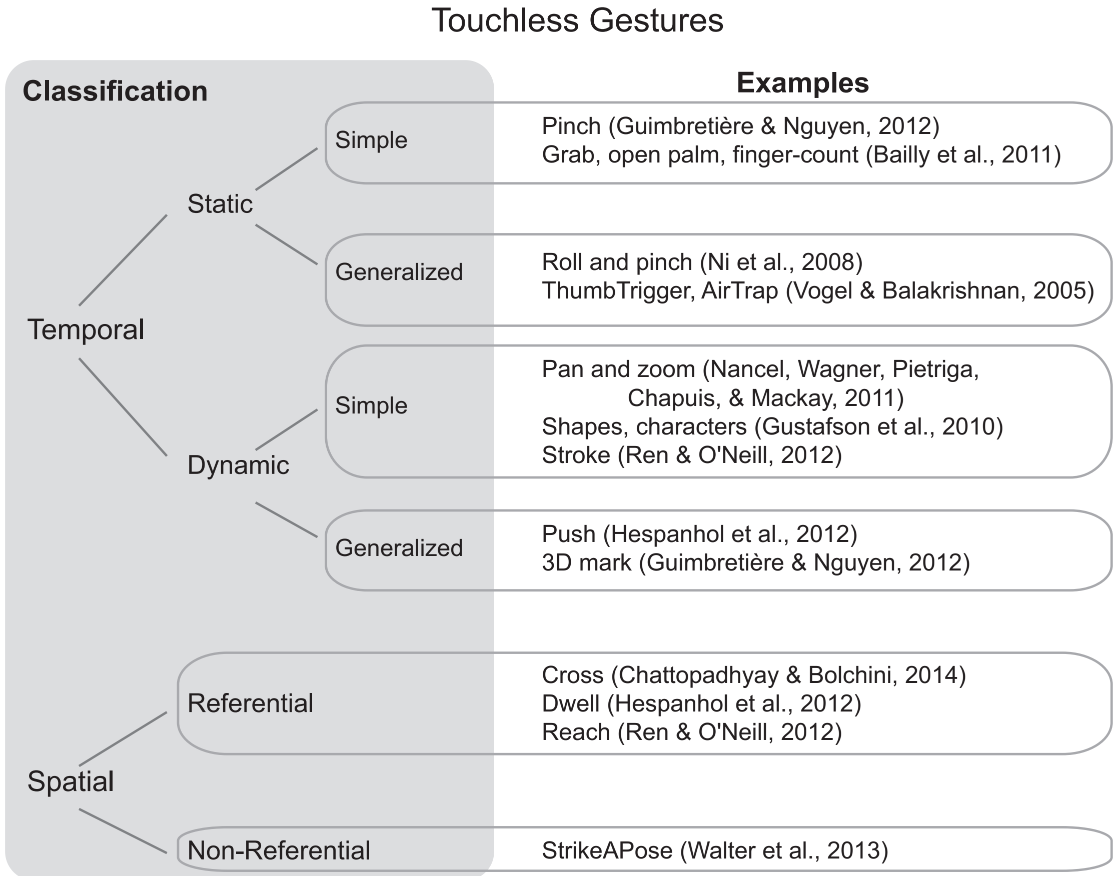

Figure 5.1. We present a taxonomy to classify the physical mechanics of device-free, mid-air gestures. We generalize the taxonomy proposed by Vatavu & Pentiuc (2008) as temporal, and further provide a spatial classification.

Touchless interaction is limited by the absence of haptic feedback, and the decoupling between the display space (containing the goal of the interaction) and the input space (containing the motor action) (O'Hara, et al., 2013). Specifically, dynamic touchless gestures (simple or generalized) suffer from human limitations to make accurate three-dimensional movements in mid-air (such as making accurate 3D strokes, or constraining hand movements in a 2D plane). Previous research that evaluated touchless gestures has reported this phenomenon (Bailly et al., 2011; Guimbretière & Nguyen, 2012; Hespanhol et al., 2012; Ren & O'Neill, 2012). Guimbretière and Nguyen (2012) report the unreliability of a three-dimensional marking menu because users failed to gauge a 3D angle for the *mark* gesture. Ren and O'Neill (2012) report similar findings for their *stroke* technique. For *push-to-select* gesture, Hespanhol et al. (2012) report a translation-action ambiguity problem. A touchless gesture suffers from translation-action ambiguity when users frequently trigger actions while repositioning their body in space

(Figure 5.2). Although the literature widely reports human limitations to make precise 3D trajectories, we still lack a causal explanation.

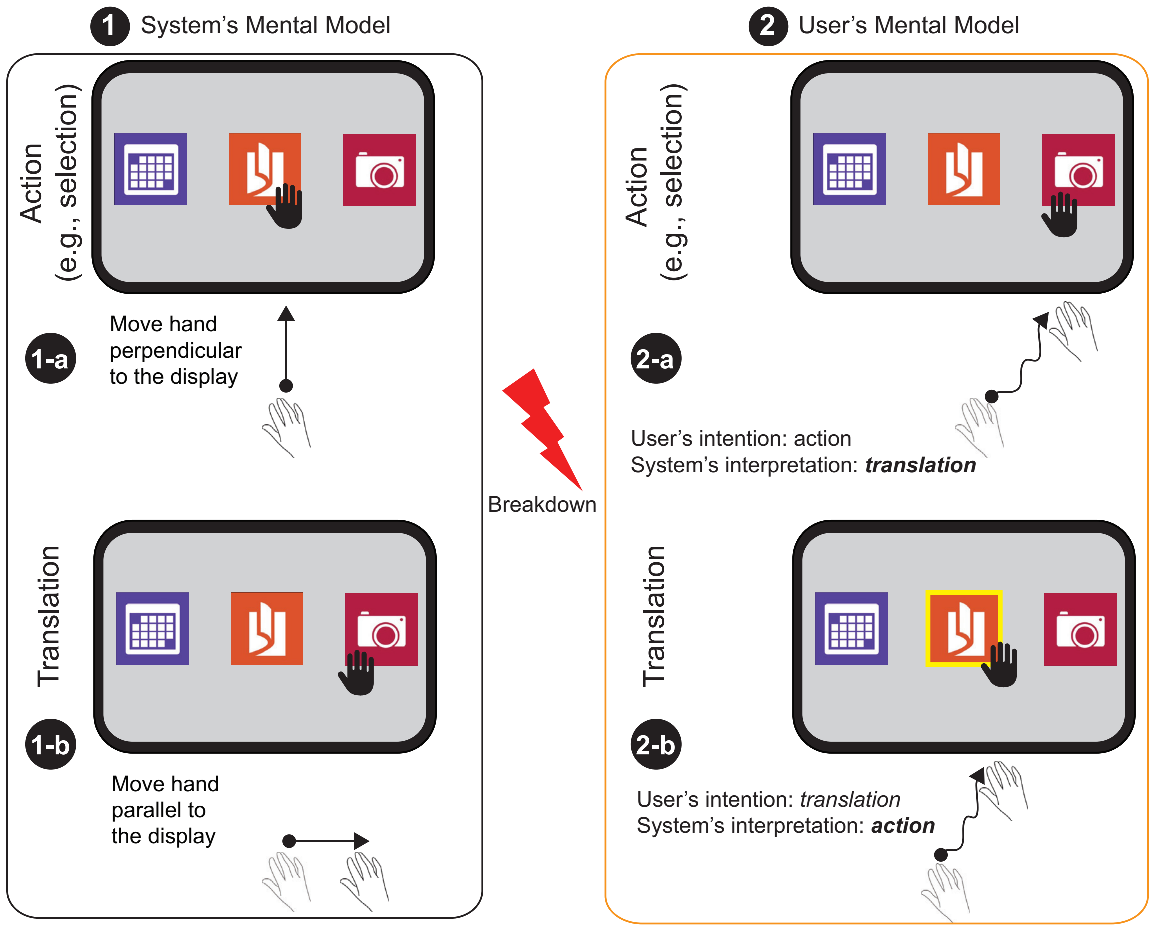

Figure 5.2. Some of the current technological systems (1) expect users to discriminate between action-gestures (1-a) and translation-gestures (1-b) by making orthogonal hand-movements. However, in daily life, we are continually moving our hands in an unconstrained, three-dimensional space. This tension between our familiar movements (2-a, 2-b) and technological expectations (1-a, 1-b) poses a translation-action ambiguity in touchless interactions.

We explain human limitations in making accurate 3D trajectories by drawing an analogy between 'reaching for an object' (a sensorimotor level of knowledge) and freehand gesturing toward a display. In daily life, we mostly move our hands in an unconstrained, three-dimensional space. To reach for an object, among infinitely many trajectories, we choose the one that minimizes our metabolic energy costs (Alexander, 1997). Hence, we are not familiar with planning movements that force us to calculate accurate 3D trajectories, or follow a combination of orthogonal paths. Based on this minimum energy cost model, we argue that users fail to perform accurate 3D strokes in mid-air as they cannot leverage their familiar mental model of movement planning. Since making accurate 3D strokes exceeds our sensorimotor level of knowledge, according to the continuum of knowledge in intuitive interaction, this would be classified as an expertise level of knowledge (Hurtienne & Israel, 2007).

Furthermore, the lack of accuracy in making 3D trajectories can be explained by the limited feedback in touchless interactions. To perform touchless interactions we rely exclusively on visual feedback and proprioception (our sense of position and orientation of the body, Mine, Brooks, & Sequin, 1997) because current touchless systems only provide visual cues on the display and no haptic feedback. Visual feedback—provided on a two-dimensional display—and proprioception cannot sufficiently guide users to make accurate 3D trajectories. Whether manipulating visual feedback (e.g., laser rays in mid-air, or 3D visualization) or adding vibrotactile feedback (e.g., airwave, Gupta, Morris, Patel, & Tan, 2013) can assist users to make accurate 3D trajectories is yet to be explored.

5.4. Motor-intuitive interactions: designing touchless primitives based on image schemas

Our explanation for the lack of accuracy in making 3D trajectories is based on the sensorimotor level of knowledge in the continuum of intuitive interaction: users fail to make 3D trajectories because they cannot apply their prior knowledge that they learned while interacting with the physical world. Hence, we argue that the potential and limitations of touchless primitives can be explained using the continuum of knowledge in intuitive interaction (Hurtienne & Israel, 2007). To illustrate our argument, we introduce motor-intuitive, touchless interactions based on image schemas that draw on our sensorimotor level of knowledge.

Motor-intuitive touchless interactions

Motor-intuitive touchless interactions are interactions where users can apply their pre-existing sensorimotor knowledge unconsciously. Specifically, they do not need to learn new motor planning or execution skills. Since childhood, we perform basic motor movements, such as pushing, pulling, grasping, or moving up and down. These motor intuitions are closely related to image schemas, such as *up-down*, *near-far*, or *left-right*. Image schemas are a schematic representation of our daily sensorimotor experiences—an abstraction of the different patterns by which our body interacts with the physical world (Johnson, 1987; Lakoff & Johnson, 1980). Hurtienne & Israel (2007) classified image schemas in eight different groups: basic, space, containment, identity, multiplicity, process, force, and attribute (Table 1, p. 130, Hurtienne & Israel, 2007). Motor-intuitive interaction primitives are based on space schemas: schemas that represent our everyday motor-actions in navigating 3D space such as up-down, left-right, near-far, front-back, center-periphery, straight-curved, contact, path, scale, or location. Intuitiveness of a motor-intuitive interaction cannot be determined solely by its performance measures (efficiency and accuracy), but depends on the level of knowledge at play during the interaction. With practice, users may perform certain motor actions accurately (expertise level), but motor-intuitive interactions are based on image schemas that act beyond our conscious awareness (sensorimotor level). Hence, motor-intuitive interactions would be easy-to-perform, learn, and remember.

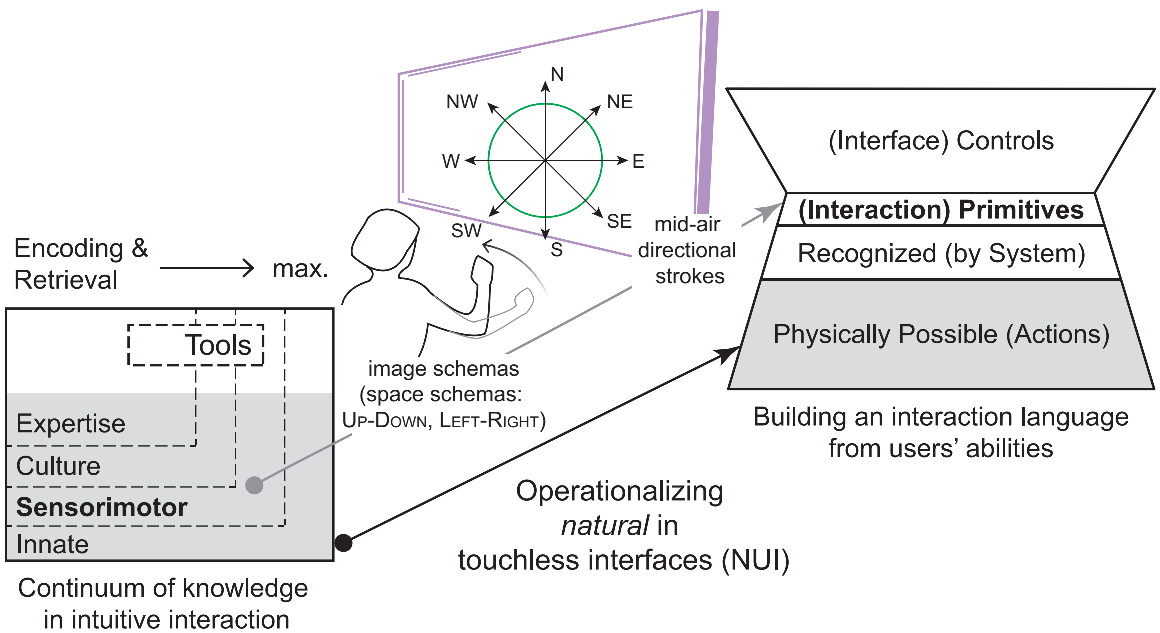

Figure 5.3. We argue that the continuum of knowledge in intuitive interaction (left, Hurtienne & Israel, 2007) can classify mid-air gestures into different levels of intuitiveness, and thereby operationalize the intuitiveness of touchless interfaces (right, Wigdor & Wixon, 2011, p. 116). Our work illustrates this argument by designing and evaluating a touchless interaction primitive (mid-air, directional strokes) that draws on our sensorimotor level of knowledge (image schemas, more specifically the up-down and the left-right space schema). To evaluate our proposed interaction primitive, we investigated user performance when making directional strokes in eight compass directions.

Because motor-intuitive interactions are based on image schemas that act beyond our conscious awareness, they are unlikely to be self-reported in traditional gesture elicitation studies. Gesture elicitation studies aim at gathering gesture primitives as suggestions from end users for any particular interaction (e.g., moving hand upward to increase the volume of a TV, Vatavu & Zaiti, 2014). As expected, participants of these studies use their previous knowledge and acquired skills to suggest touchless interaction primitives. They certainly use metaphors to map the gestures to their meaning (Lakoff & Johnson, 1980), such as the motion of cutting with an imaginary knife to mean a *slice* gesture (Grandhi, et al., 2011). However, with respect to the continuum of knowledge (Figure 5.3, left), these metaphors mostly reside at the levels of tool, expertise, or culture. Thus, it is not surprising that researchers report limitations of elicitation studies due to expertise bias (previously acquired gesture interaction models, such as the mouse, Morris, et al., 2014; Vatavu & Zaiti, 2014) or cultural bias. As an alternative to gesture elicitation, in our approach toward designing intuitive touchless interaction primitives, we shifted to the sensorimotor level of the continuum of knowledge and introduced motor-intuitive, touchless interactions. To exemplify our concept, we propose a novel, motor-intuitive, touchless primitive: mid-air directional strokes.

5.5. Mid-air directional strokes: a motor-intuitive touchless primitive based on image schemas

We propose a motor-intuitive, touchless interaction primitive: mid-air strokes dynamically mapping the *up-down* and the *left-right* schema. Using these two space schemas, users can make any two-dimensional directional movements, such as north, south or southwest. (Making accurate 3D movements would require the use of an additional *front-back* schema. While physical tokens allow tangible interactions to use the *front-back* schema, the absence of haptic feedback in touchless interactions limits the use of that space schema.) To leverage the up-down and the left-right space schemas, a touchless system would provide visual cues on a 2D display and use an orthographic projection to interpret users' 3D hand movements as 2D trajectories. This design proposal opens up a number of questions. Most importantly, given that the sensorimotor knowledge is constant across different directions, *what other factors could affect such mid-air movements*? How will different directions affect users' performance? Will users be more effective with smaller strokes?

Effect of biomechanical factors on mid-air directional strokes

Our proposed motor-intuitive, touchless interaction primitive is based on space schemas that use the sensorimotor level of knowledge. In touchless interactions, user performance depends on both the level of knowledge at play and biomechanical properties of the human body. To investigate how biomechanical properties can affect a motor-intuitive, touchless primitive, we designed a controlled experiment. Theoretically, users can make any two-dimensional directional movements using the two space schemas left-right and up-down. For our controlled experiment, participants performed mid-air strokes in eight compass directions while sitting away and interacting with a large display. The directions of movement were represented visually on the display to leverage users' sensorimotor skills (image schemas; for details see the Tasks and Procedure section). In our study, we were specifically interested to understand how directions of movement and stroke lengths affect user performance of mid-air strokes. Our experiment did not investigate intuitiveness in touchless interactions (as studied by Aigner et al., 2012, Grandhi et al., 2011, or Hespanhol et al., 2012), but explored how the same motor-intuitive, interaction primitive can cause different user performance (operationalized as accuracy and efficiency). We did not measure users' self-reported satisfaction because during the pilot studies most users reported equal preferences for all directions of movement and stroke lengths.

5.6. Evaluating user performance of mid-air directional strokes

When we move our arms in mid-air, biomechanical properties of the human body (such as the position of the forearm relative to the upper body) affect how accurately and quickly we can make arm movements (Werner, Armstrong, Bir, & Aylard, 1997). Although empirical studies suggest that hand pointing at shoulder level requires more effort than pointing at center level, no significant effects of arm-configuration or arm-extension on performance time (efficiency) or accuracy has been reported (Hincapié-Ramos et al., 2014). Because of the required effort, we argue that arm postures will affect the efficiency and accuracy of hand movements.

Hypothesis 1 (H1): Direction of movement will affect the efficiency of mid-air directional strokes.

Hypothesis 2 (H2): Direction of movement will affect the accuracy of mid-air directional strokes.

Pointing and target acquisition has been widely studied in device-based input modalities (Fitts, 1954; Grossman & Balakrishnan, 2004; MacKenzie & Buxton, 1992;

Shoemaker, Tsukitani, Kitamura, & Booth, 2012). It is well established in the literature that time taken to complete a movement is directly proportional to the amplitude of the movement. Moreover, Nancel et al. (2011) found that unguided mid-air gestures are more tiring than device-based mid-air gestures, which suggests that users would be more precise with a smaller amplitude of movements.

Hypothesis 3 (H3): Increase in stroke length will decrease the efficiency of mid-air directional strokes.

Hypothesis 4 (H4): Increase in a stroke length will decrease the accuracy of mid-air directional strokes.

Method

We conducted a within-subject experiment to understand how well participants can perform mid-air strokes in different directions. Specifically, we wanted to test the effect of direction and stroke length on the efficiency and accuracy of mid-air strokes. Furthermore, we wanted to compare the paths that participants took across different directions and stroke lengths. This important data can inform future research on designing touchless interfaces that draw on dynamic gestures.

Participants

We recruited 17 right-handed participants (7 females) from an urban university campus. Ten participants had prior familiarity with touchless gestures. Twelve participants were below 30 years of age. Participants were randomly recruited by sending out emails using the university's mailing list. The study was approved by the Indiana University Institutional Review Board (Protocol# 1303010855), and participants were compensated with a $20 gift card for an hour of participation.

Apparatus

We used a high-resolution large display integrated by Fakespace Systems that comprises of eight 1.27 m projection cubes laid out in a 4 x 2 matrix. It is driven by a single computer. Each cube has a resolution of 1600 x 1200 pixels, resulting in a 4.06 m wide by 1.52 m high display with over 15.3 million pixels. We used a Kinect™ for Windows to track users' hand position. The experiments were written in C# running on Windows 7, and were implemented with OpenNI 1.4 SDK and PrimeSense's NITE 1.5.

Tasks and procedure

To test our hypotheses, we designed an experimental task (Figure 5.4, right) inspired by a previous study (Lepinski, Grossman, & Fitzmaurice, 2010). On a large interactive display (Figure 5.4, left), participants were presented with a direction (at random) and a target line in that direction. The (640-pixel long) target line informed users of the minimum travel length and appeared at 500, 800 or 1100 pixels. Participants were situated 1 m away from the sensor and were asked to make a hand movement in the provided direction as accurately as possible. The motion-tracking sensor had a horizontal field of view of 57 degrees and a vertical field of view of 43 degrees. Participants' movements were mapped from real space to display space as 1: 3.7 (when a participant moved 1 cm in real space the cursor moved 3.7 cm in the display space). Trajectory lengths in real space were 86 mm, 137 mm, and 189 mm. We chose smaller movements because a survey on social acceptability of touchless gestures (Bragdon et al., 2011) found that 80% of respondents felt comfortable performing smaller hand motions over larger body motions, such as sweeping their arms well across their body. Eight different directions were presented at random: 0, 45, 90, 135, 180, 225, 270 and 360 degrees.

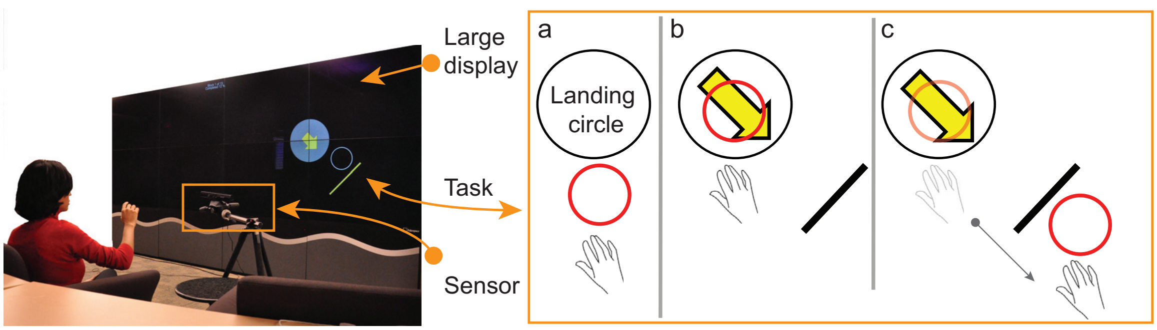

Figure 5.4. (Left) In our experiment, participants used touchless gestures to interact with a large display, while sitting away from it. (Right) The experimental task began with a landing circle appearing on the display (a). As participants reached the landing circle, the direction of movement and the target line appeared (b). Participants completed the task by making a directional stroke with a minimum travel distance as informed by the target line (c).

Participants sat on a comfortable couch at 2.25 m away from the large display and took about 20-30 minutes to complete all trials. Existing studies on touchless interaction with large displays have mostly investigated settings where users are standing in front of the display. However, a sitting posture may limit users' fluidity of hand-movements more than a standing posture. We chose a sitting position for our experiments to avoid standing fatigue and uncover any limitations posed by a sitting posture. Trials were recorded using a video camera capturing users' gestures and the display. Before the actual experiment, all participants completed three blocks of practice trials. Participants were required to take at least a 10-second break in between each block. Trials were randomized within subjects. In summary, the study design was as followed: 8 directions (trials) x 3 trajectory lengths x 5 blocks x 17 participants = 2040 trials.

Participants hovered over a 'Start' circle to begin a block. Each trial began with a *landing circle* appearing on the display, which participants landed on to begin the trial. The landing circle was horizontally aligned with the participants' body midline, and 142 cm from the ground. The sensor was 84 cm from the ground, and the couch-seat was 44 cm high. As soon as participants reached the landing circle, two things would appear: an arrow representing one out of eight directions and a *target line* at one of the three stroke lengths (Figure 5.4). For a trial to be considered successful, participants were required to move past the target line with an angular error less than 45 degrees. Participants' hand movements in the 3D space were measured as their orthographic projections on the 2D display.

Measures

We recorded performance time, error rate, angular error, and trajectory paths. Time was measured from when participants left the landing circle to when they moved past the target line. We measured the stroke angle using the last point recorded inside the landing circle and the first point recorded after crossing the target line (hence the target line, though 20-pixel wide, did not influence the calculation of angular error). The angular error was calculated as the absolute difference between this stroke angle and the required angle for the trial. An error was recorded when the angular error was more than 45 degrees. In the case of an error, the trial was repeated until participants successfully completed it. We measured the efficiency as time to complete a trial and accuracy as error rates and angular error.

Results

Performance data was analyzed using nonparametric tests for within-subject experimental design because Shapiro-Wilk tests were significant, $p < .001$, and Q-Q plots were non-linear. In our experimental setup, participants sat in a couch away from the large display (Figure 5.4). We observed that some participants ran into considerable ergonomic constraints in making movements in the south direction (270 degrees) because of the sitting posture. Their arm movements got hindered by their knees or the armrest of the couch (more in Limitations). This effect is obvious in all of our following results. To ensure that this experimental artifact would not affect the conclusions we draw from our results, we also tested our data without considering the S-direction as one of the levels of the direction variable. When these tests showed major differences in terms of the significance level, we reported the test statistic and the level of significance.

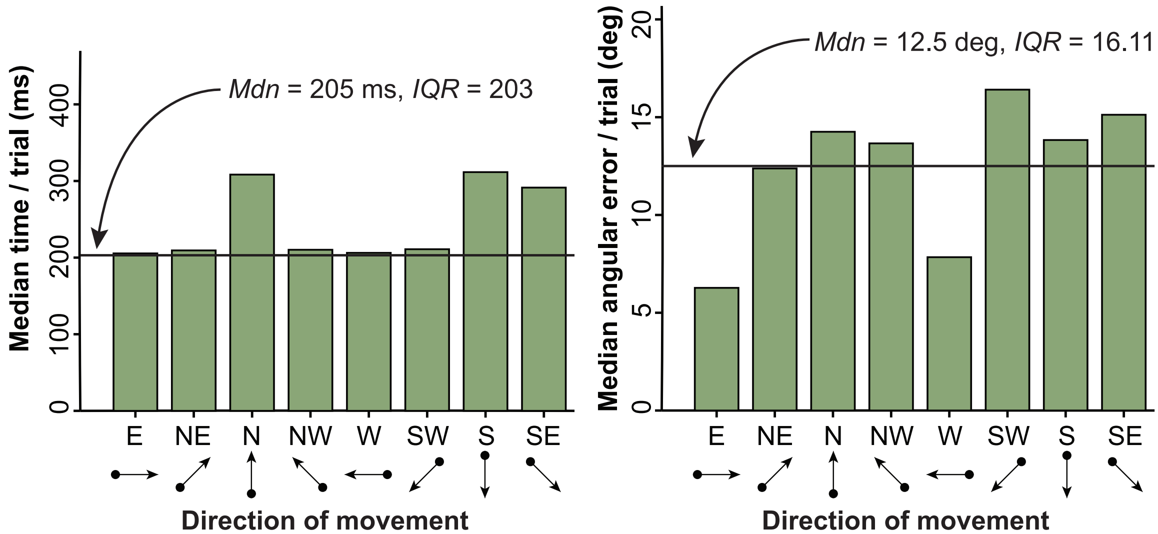

Figure 5.5. The direction of movement significantly affected performance time and the angular error of mid-air strokes, p < .001. Participants made significantly less angular error (p < .001) in E and W direction compared with all other directions (NE, N, NW, SW, S, and SE).

Direction of movement affects efficiency and accuracy of mid-air strokes

Direction of movement significantly affected performance time (*Mdn* = 205 ms, *IQR* = 203), $\chi^2(7) = 146.93$, $p < .001$ (Figure 5.5). We conducted 13 pairwise comparisons: N vs. rest of the directions, and S vs. rest of the directions. Post-hoc Wilcoxon Signed-rank tests (with *Bonferroni* correction, significance level .0038) revealed that participants took significantly more time making strokes in N-direction than E, W, NE, or NW, $p < .001$, with a medium effect, $.33 < r < .46$. We found a significant learning effect across blocks, $p < .01$. Participants were about 66 ms faster in the last block than in the first block. *H1* was supported.

A trial was considered erroneous, when participants made an angular error more than 45 degrees in clockwise or counter-clockwise direction. Direction of movement significantly affected error rate (*Mdn* = 4.76%, *IQR* = 7.08), χ2 (7) = 28.82, $p < .001$ (without S-direction, $\chi^2(6) = 20.7$, $p < .01$).

Direction of movement significantly affected angular error (*Mdn* = 12.5 degrees, *IQR* = 16.11), $\chi^2(7) = 159.14$, $p < .001$. We conducted 13 pairwise comparisons: E vs. rest of the directions, and W vs. rest of the directions (with *Bonferroni* correction, significance level .0038). Post-hoc Wilcoxon Signed-rank tests revealed that angular error in directions N, NE, NW, S, SE and SW were significantly more than angular error in E direction, $p < .001$, with a medium effect, $.37 < r < .50$; and in W direction, $p <.001$, with a small to medium effect, $.28 < r < .44$. Angular error was more in W direction (*Mdn* = 7.00 deg.) than in E direction (*Mdn* = 8.60 deg.), $Z = 2.08$, but not significant, $p = .04$. *H2* was supported.

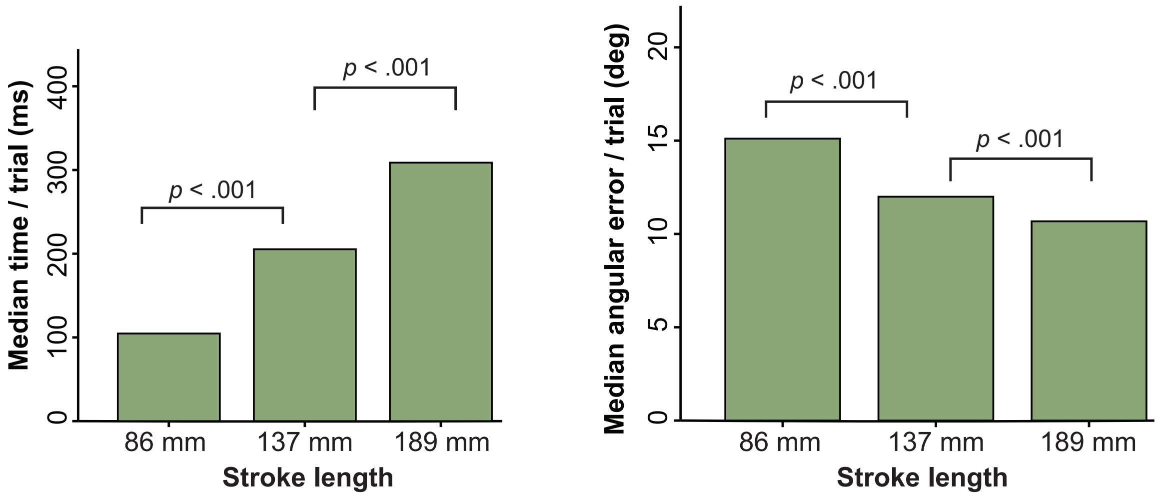

Figure 5.6. Stroke length significantly affected performance time and angular error of mid-air strokes, p < .001. Interestingly, participants made significantly less angular error with increase in stroke length, p < .001.

Stroke length affects efficiency and accuracy of mid-air strokes

Stroke length significantly affected performance time, $\chi^2(2) = 385.39$, $p < .001$ (Figure 5.6). Post-hoc Wilcoxon Signed-rank tests (with *Bonferroni* correction, significance level .016) revealed that performance time was significantly different between each pair of distances, $p < .001$. Small stroke length (*Mdn* = 105 ms, *IQR* = 28.7) was significantly faster than medium stroke length (*Mdn* = 205 ms, *IQR* = 182.12) with a medium effect, $Z = 13.48$, $p <.001$, $r = .47$; and medium stroke length was significantly faster than large stroke length (*Mdn* = 309 ms, *IQR* = 230.63) with a medium effect, $Z = 12.44$, $p <.001$, $r = .44$. *H3* was supported.

Stroke length did not significantly affect error rate, but significantly affected angular error, $\chi^2(2) = 42.19$, $p < .001$ (without the S-direction: $\chi^2(2) = 34.66$, $p < .001$). Moreover, post-hoc tests revealed that angular error significantly decreased with increase in stroke lengths. Angular error for small strokes (*Mdn* = 15.1 degrees, *IQR* = 17.61) was significantly more than angular error for medium strokes (*Mdn* = 12 degrees, *IQR* = 15.74) with a small effect, $Z = 4.44$, $p < .001$, $r = .13$; angular error for medium strokes was significantly more than angular error for large strokes (*Mdn* = 10.68, *IQR* = 12.4) with a small effect, $Z = 4.04$, $p < .001$, $r = .11$. *H4* was not supported.

Trajectory patterns indicate asymmetric ability in touchless interactions

We recorded the paths participants took to move in different directions across different stroke lengths (Figure 5.7). Participants were asked to make directional strokes as accurately as possible. We recorded paths only for successful trials, and a trial was successful if a participant's angular error was less than 45 degrees. From the visualization of these paths, a number of patterns emerged. First, participants' trajectories were longer on their dominant side compared with their non-dominant side. Second, confirming previous findings, their angular error decreased as stroke length increased. Third, we observed a trend in participants' hand movements toward the eastern hemisphere (dominant side) and the northern hemisphere. For example, in both N and S direction of movement, participants' strokes tended toward the eastern hemisphere; in E and W direction, their strokes tended toward the northern hemisphere.

In the following section, we discuss the lessons learned from our experiments, and the implications suggested by our findings. Specifically, we discuss how our findings can inform the design of intuitive touchless interactions and UI elements for large-display touchless interactions (such as menus, or toolbars).

Discussion

In this paper, we introduced motor-intuitive, touchless interactions based on image schemas that draw on our sensorimotor level of knowledge. To illustrate our concept, we proposed a motor-intuitive, touchless interaction primitive: mid-air directional strokes mapping the *up-down* and the *left-right* space schemas. We then argued that in touchless interactions, a motor-intuitive primitive is affected by the biomechanical properties of the human body. To that aim, we explored how the same motor-intuitive, interaction primitive can result in different user performance across different directions of movement and different stroke lengths.

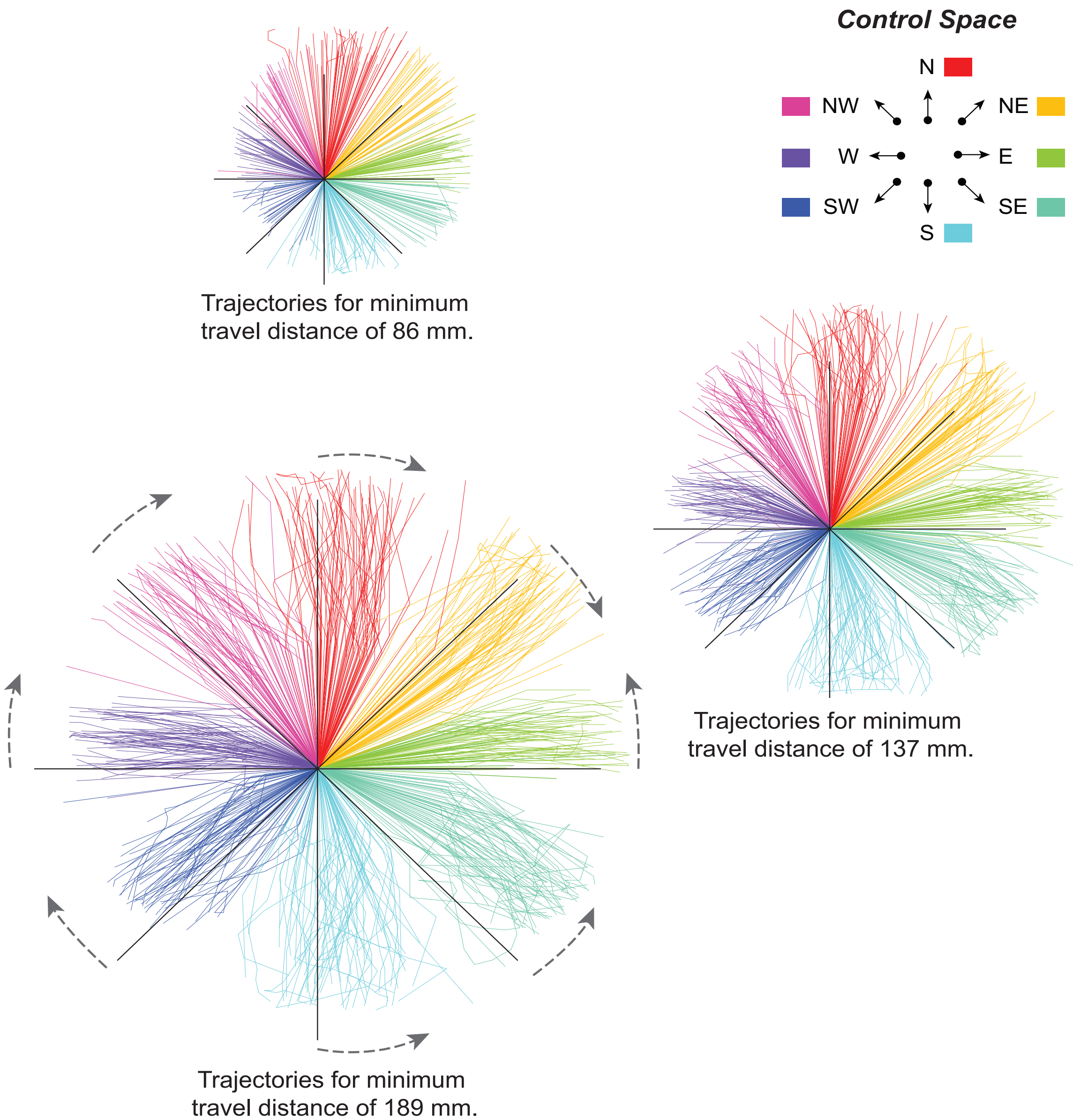

Figure 5.7. We recorded trajectories (across 8 directions, and 3 distances) from 17 right-handed participants as they performed directional strokes in mid-air (see Figure 5.4). In right-handed users' control space, we observed the following: (a) participants performed longer trajectories while operating on their dominant side than in their non-dominant side; (b) participants' angular error decreased with an increase in the stroke length (similar to Figure 5.6); and (c) participants' hand movements tended toward the eastern hemisphere and the northern hemisphere (illustrated by dashed arrows).

Lessons learned

In a controlled experiment (N = 17), we investigated efficiency and accuracy of mid-air strokes. We learned the following from our study. First, direction of movement significantly affected efficiency and angular error of mid-air strokes. On average, participants were very efficient and took only 0.2 seconds (median performance time) to make a directional stroke. However, their median angular error was *12.5 degrees*, which is slightly more than twice compared with a previous study on multitouch strokes (5.6 degrees, Lepinski et al., 2010). Increase in angular error from multitouch to mid-air strokes contradicts a previous finding, where 2D-surface gestures were more erroneous than 3D-free gestures (Nancel et al., 2011). However, such a comparison is limited, because these studies used different experimental tasks and settings. Previous studies that explored 3D strokes as interaction commands (Guimbretière & Nguyen, 2012; Ren & O'Neill, 2012) do not report any performance measures because users were extremely inaccurate. Instead, the gesture primitives were either redefined or reported as infeasible. Unlike accurate 3D strokes (based on the expertise level of knowledge), we found 2D directional strokes (based on the image schemas, which is a sensorimotor level of knowledge) generally effective and efficient. This supports our premise that the intuitiveness of touchless interactions can be operationalized using the continuum of knowledge in intuitive interaction (Hurtienne & Israel, 2007): the higher the level of knowledge used in an interaction primitive, the lower would be the expected speed of knowledge retrieval, and the lesser would be the primitive's intuitiveness to general population.

Second, an increase in stroke length increased performance time. This is an expected result that aligns well with previous findings for other input modalities, where movement time increased with movement amplitude (Fitts, 1954; MacKenzie & Buxton, 1992). The increase in stroke length also decreased angular error. This is an unexpected finding that suggests that we tend to over-correct our movements based on forward planning (Shadmehr, Smith, & Krakauer, 2010). This finding advises against designing touchless gestures that require users to make directional strokes with a very short trajectory length (more in the Design Implications).

Third, we found an effect of cross-lateral inhibition on user's ability to make mid-air strokes. Cross-lateral inhibition occurs when users' hand crosses the body midline and operates away from their dominant side (Figure 5.8): Crossing the 'body midline' offers more resistance than operations limited to the same side of the dominant hand

(Schofield, 1976). In line with this biomechanical property, we observed that across all stroke lengths, right-handed participants made longer strokes on their dominant side (Figure 5.7). However, we did not find any significant effect of cross-lateral inhibition on users' efficiency or accuracy. This effect of cross-lateral inhibition indicates how handedness—an innate level of knowledge—affected an interaction primitive that used the sensorimotor level of knowledge. This observation follows the inherent dimensionality of the continuum of knowledge in intuitive interaction: the lower the level of knowledge the higher the frequency of encoding and retrieval of knowledge. Hence, interaction primitives designed to use any particular level of knowledge in the continuum would still be affected by the levels of knowledge residing below (in varied amounts based on prior use and training).

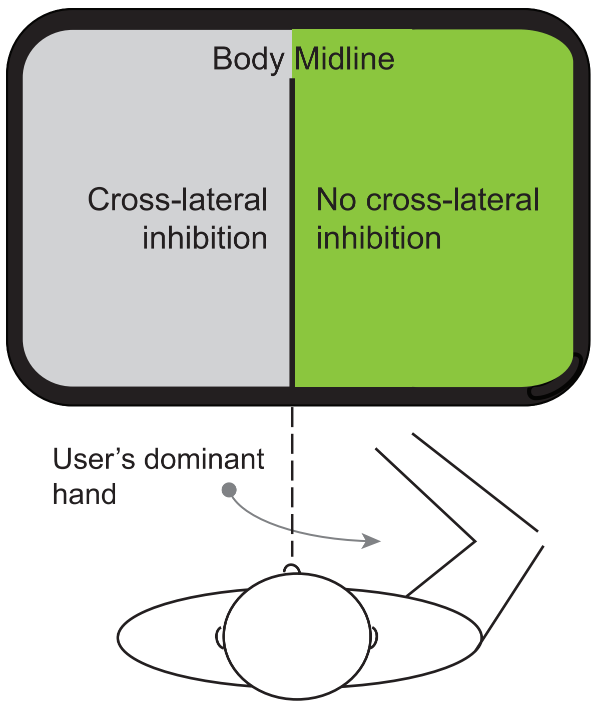

Figure 5.8. Cross-lateral inhibition occurs when users' hand crosses the body midline and operates away from their dominant side (e.g., left side for right-handed participants).

Overall, our findings suggest that in intuitive touchless interactions, user performance of a motor-intuitive, touchless primitive is significantly affected by the biomechanical properties of the human body. Based on efficiency, angular errors, and the trajectory-patterns that participants took to make directional strokes in mid-air, we identified three regions that are characterized by decreasing performance and increasing effort: top-right, top-left, and top-middle (Figure 5.9). Our findings align with previous results where researchers found that users' physical effort was significantly more for interactions in the shoulder plane (similar to our top-middle) compared with interactions in the center plane (similar to our top-left and top-right) (Hincapié-Ramos et al., 2014). We do not comment on users' relative performance in the southern hemisphere because we observed that our experimental setting constrained some users' southward movements. Hence, the relatively inferior user performance may be an artifact of our experiment. In the following paragraphs, we use our findings to inform some design implications for both intuitive touchless interactions and large-display touchless interactions.

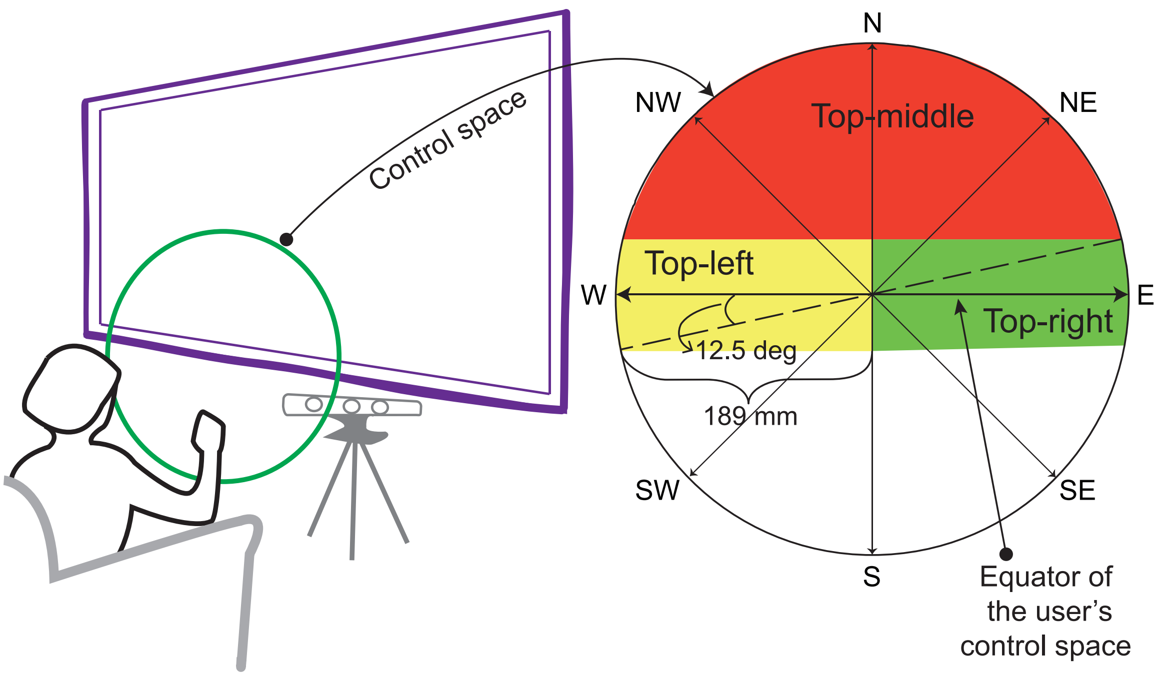

Figure 5.9. In a right-handed user's control space, while sitting away and interacting with a large display, our study on mid-air directional strokes identified three regions that are characterized by decreasing performance and increasing effort: top-right, top-left, and top-middle.

5.7. Design implications

Design implications for intuitive touchless interactions.

Compared with previous reports on users' failure to perform 3D trajectories as interaction commands, we found that mid-air directional movements based on image schemas (*up-down* and *left-right*) were efficient (median 0.2 s) and effective (median angular error of 12.5 degrees). We did not record users' self-reported satisfaction because during our pilot studies most users reported equal preferences across all experimental conditions. Such equal preferences can be explained by the same level of knowledge at play (sensorimotor level) during interacting in different directions and different stroke lengths. To compare a motor-intuitive touchless primitive with another gesture primitive, in Chapter 6, we introduce a command-selection technique based on mid-air directional strokes (Touchless circular menus). In a controlled study, we found that touchless circular menus were twice more efficient than linear menus that used “grab” gestures; users also perceived less workload while using the touchless circular menus.

Our empirical results suggest that in touchless interaction, *intuitive interaction* depends on both the sensorimotor level of knowledge and the biomechanical properties of the body. We showed that even when touchless interactions with mid-air strokes draw on the sensorimotor level of knowledge, other factors such as directions of movement and stroke lengths significantly affect the user performance. Specifically, we present two design guidelines for intuitive touchless interaction. First, scale-dependent, directional gestures should not be of very small length (e.g., 86 mm in our experiment) because though such gestures take less time to complete, they seem to produce more angular error. Instead, designers should consider selecting a stroke length that involves more than just flipping the hand, such as moving the forearm (portion of users’ arm between the elbow and the wrist), because longer strokes are more accurate. However, it must be noted that large hand-movements sweeping across the body have been reported as fatiguing and socially unacceptable (Bragdon et al., 2011). Hence, large hand-movements requiring users to move their arm (not just forearm) might be more effective but would be less efficient and less acceptable to users in a specific context.

Second, dynamic touchless gestures should only require users to make 2D directional strokes, rather than accurate angular movements in 3D. For example, a menu option should be accessible by making a stroke in any compass direction (such as NE or SW) that is based on space schemas (up-down or left-right), rather than a three-dimensional angle in freespace (such as vectors in 3D space, cf. Figure 5 in Guimbretière & Nguyen, 2012). Continuous visual feedback to guide users in making such directional gestures will be helpful. For example, rather than depending on users’ proprioception to execute mid-air strokes (Guimbretière & Nguyen, 2012), a dynamic illustration (e.g., a visual trace) of how users’ hand is moving could be shown (section 4.11, Chapter 4). In general, visual feedback is a major factor affecting the intuitiveness of touchless interactions.

In the absence of any haptic feedback, visual feedback plays a major role to ensure that while interacting with motor-intuitive touchless primitives, users are actually drawing upon their sensorimotor level of knowledge. For example, it is crucial to provide a proper visual representation of an image schema on the display (e.g., Figure 5.4) and adopt an effective frame of reference (egocentric or allocentric, Klatzky, 1998). Our experiments used an allocentric (viewer's) frame of reference and traditional GUI-type visual feedback because the alternative—egocentric frame of reference and full-body avatar visualization—may not be suitable in certain scenarios, such as visualization or collaborative work (Bragdon et al., 2011; Dostal, Hinrichs, Kristensson, & Quigley, 2014). Because egocentric frame of reference and avatars are used in immersive full-body games, touchless interactions in those games need not be grounded on image schemas to feel intuitive. But in more traditional settings, touchless primitives based on image schemas will be more intuitive than primitives based on expertise level of knowledge. While we do not discuss the role of visual feedback in touchless interactions in this paper, the effect of visual feedback on acquiring, learning, and retaining motor actions is well studied (e.g., see Sigrist, Rauter, Riener, & Wolf, 2013).

Design implications for large-display touchless interactions.

Our findings can also be leveraged to design interface elements for large-display touchless UIs. First, directional strokes to trigger frequently used commands should be in the top-right or the top-left of the user because users' performance suffers as they operate in the top-middle of their control space (see Figure 5.9). Moreover, user effort increases as the dominant hand suffers from cross-lateral inhibition when it crosses the body midline (Figure 5.8). Designers can leverage this characterization of users' control-space to define rarely used gestures. For example, to operate a media player, users could make a stroke in E-direction to play/pause, and a stroke in N-direction to quit the media player. Similarly, crucial interface widgets such as toolbars should be around the equator of the users' control space (see Figure 5.9) because we found that users were most effective and efficient in executing mid-air strokes around the equator.

Second, the average angular error of 12.5 degrees in mid-air directional strokes suggests that pie-based touchless menus can offer about 25 command-selection options that can still be accurately selected with mid-air strokes. Future controlled experiments can be informed by this range of touchless menu options to further determine the precise cardinality of menu items. Apart from menus, mid-air directional strokes will also play an important role as an interaction primitive in touchless user interfaces for sketching (Taele & Hammond, 2014).

Finally, our experiment with a large display explored intuitiveness of touchless primitives that leverage the entire control space available for hand gestures. Hand gestures, when used with current game consoles, only involve a small control space (i.e., users' hand movements are very small and directly in front of the sensor) because users are situated about 2 – 3 m away from a 1.27 m HDTV. In comparison, large-display interaction opens up the potential of a larger control space. In our work, we showed how using a larger control space poses new limitations to touchless interactions, such as biomechanical factors, even when a gesture primitive is based on our sensorimotor level of knowledge (motor-intuitive). Furthermore, we argue that motor-intuitive, touchless interactions will outperform expertise-based interactions because touchless interactions are sporadic, spontaneous, and short-lived: They are often used for exploratory tasks (e.g., browsing images, opening and closing files, or using media controls) rather than fine-grained, repetitious tasks (e.g., editing) (Chattopadhyay & Bolchini, 2013).

Limitations

Our experiment was limited by the capabilities of our motion-tracking sensor, which operated at a refresh rate of 30 frames per second. We could not record some of the trajectory paths (Figure 5.7) because some tracking points were lost when participants moved their hands very fast. We also placed our sensor in such a way that the execution of the longest stroke was within the sensor's optimal tracking range. Our experimental setup also limits our findings. Specifically, we observed that some of our participants faced considerable ergonomic constraints while performing southward movements. We chose a 'sitting' position for our experiments to avoid users' standing fatigue. We did not anticipate that users would face ergonomic constraints in this position, but users often moved their hands backward (toward the center of their body) instead of southward, thus causing the arm-rest to restrain their movements.

For small-length strokes, some users completed an entire experimental trial in the E and the W direction while resting their hands on the arm rest. This was not possible for trials in any other direction or with medium or large strokes. While those few trials may have increased the efficiency and accuracy of mid-air strokes in E and W direction, they do not confound our general conclusion that biomechanical factors affect motor-intuitive, touchless interactions. Furthermore, other experiments using a standing posture and without any armrest has also shown that touchless interactions in the center plane (e.g., E and W) requires significantly less effort than interactions in shoulder plane (e.g., N, S, or NW). In addition, all our participants were right-handed. Hence, we cannot claim a generalization of our findings across left-handed users.

We did not investigate the effect of control-display gain or pointer acceleration on the execution of mid-air strokes. This would be necessary to design the required length of mid-air strokes in a touchless interface. We anticipate an effect of pointer acceleration on user performance of mid-air strokes. Furthermore, we did not record any subjective ratings for user fatigue or intuitiveness. Informally, users did not report any physical strain after 30 minutes of execution of mid-air strokes.

We need to further consider the role of visual feedback in guiding users to make mid-air strokes. In our study, the direction of movement was presented as a static image. Users mentioned that a dynamic illustration of their hand movement would be helpful in making accurate strokes. We think that adequate visual feedback will somewhat mitigate the absence of haptic feedback, and also improve users' learnability. However, this needs to be further explored.

Though we mention that the median angular error for mid-air, directional strokes—12.5 degrees—can inform the design of touchless pie-menus, future experiments are required to identify the precise cardinality of such menus. Moreover, in our experiments, we used a landing circle to mark the beginning of a mid-air stroke. It is necessary to investigate specific invocation techniques when such dynamic gestures are applied to touchless interfaces.

External validity

Our findings can be generalized to touchless interaction settings, where users are sitting away from a large display, facing the display, and within the sensor's tracking range. Though our study used a couch with an arm-rest (see Figure 5.4) our findings can be extended to other furniture setups. However, it must be noted that an arm-rest in such scenarios plays a two-fold role: (a) it can help reduce user fatigue by allowing the elbow to rest during hand movements; (b) it can also constrain southward movements. Since sitting posture already constrained users' hand movements to a certain extent, we expect our general findings to stay valid in a standing posture. For example, 2D strokes would be more intuitive than 3D strokes without prior expertise and directions of strokes and stroke-length would still affect the user performance of mid-air strokes. However, in a standing posture, users would be more efficient in utilizing the southern hemisphere of their control space than while sitting. Finally, our design guidelines are agnostic of the control-display gain of the system, or how the control space is mapped to the display space. We provided insights into how human sensorimotor abilities (in the control space) can inform the design of intuitive touchless interfaces (in the display space).

5.8. Conclusions

How intuitively users perform a mid-air hand gesture can inform what subset of physically possible actions should constitute intuitive touchless interactions. For example, in this paper, we contrasted between two touchless gesture primitives—making accurate 3D strokes that draw on the expertise level of knowledge and making 2D directional strokes that draw on the sensorimotor level of knowledge. The fact that making accurate 3D strokes is less intuitive for the general population than making 2D strokes can be explained by the intuitive interaction framework where the expertise level of knowledge resides above the sensorimotor level. Hence, we argued that the *continuum of knowledge in intuitive interaction* can operationalize the intuitiveness of touchless interfaces because it informs the design of touchless primitives by considering the level of knowledge that is at play during their execution. Specifically, we introduced motor-intuitive, touchless interactions based on image schemas that draw on our sensorimotor level of knowledge. To illustrate motor-intuitive interactions, we proposed a touchless primitive—mid-air, directional strokes—based on space schemas *up-down* and *left-right.* We then investigated how our proposed touchless primitive is affected by the biomechanical properties of the human body.

Our findings suggest that mid-air (2D) directional strokes are efficient (median time of 0.2 seconds) and effective (median angular error of 12.5 degrees). From our results, we discovered that directions of movement (2D) and stroke length affect users' performance of mid-air directional strokes. Interestingly, users made significantly accurate strokes while traveling longer trajectories. While sitting away and interacting with a large display, our results identified three regions in a right-handed user's control space that can be characterized by decreasing accuracy and increasing effort: top-right, top-left, and top-middle. Finally, grounded in our findings, we provided practical guidelines on designing intuitive touchless interaction and UI elements for large displays.

This is but a first step in understanding how the continuum of knowledge in intuitive interaction can inform the design of motor-intuitive, touchless interaction primitives. Our findings can inform fundamental design decisions to align touchless user interfaces with human sensorimotor abilities, thus making them intuitive to use. An important result from this study is how asymmetric motor abilities—due to biomechanical factors—affect user performance of motor-intuitive, touchless interactions. This research opens up an immediate line of inquiry—a need to explore the proposed motor-intuitive interaction primitive, 2D directional strokes, as part of an interaction technique. We explore this in Chapter 6.

Chapter 6. Interaction techniques

This Chapter focuses on interface affordances and exclusively serves two purposes. First, it builds upon the motor-intuitive interaction primitive introduced in Chapter 5, mid-air directional strokes, and introduces a touchless interaction technique. The interaction technique is then evaluated in a controlled study. Empirical results from this user study, then prompts the proposal of the second interaction technique—discussed later—and the experiments in chapters 7 and 8.

6.1. Touchless circular menus

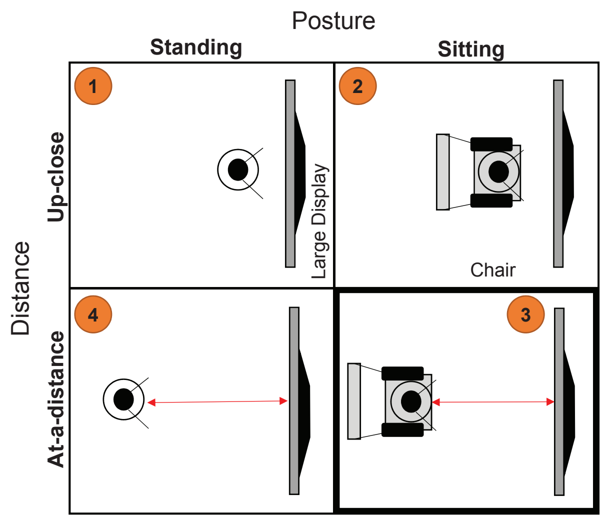

Figure 6.1. Large display interaction space across two dimensions: user posture and distance from the display. Scenario 3 represents our experimental setting.

To support touchless interactions with large displays, we still need a fundamental set of interface conventions for frequent user-operations, such as pointing, text-entry, or command-selection (Figure 6.1). This area is a largely uncharted territory. Specifically, whereas an extensive body of works investigated optimal menu designs for mouse-and-keyboards, pen-input, or multitouch surfaces, few have explored touchless command-selection techniques for large displays. Recent solutions that have appeared in product platforms (e.g., Samsung Smart TV) or research venues require users to comply strictly with system-defined poses, such as closing the hand, pinching with fingers, or making different finger combinations. These approaches are problematic because they are analogous to command-line interfaces: users need to remember an interaction vocabulary and input a pre-defined symbol (via gesture or command). Not only have expert reviews commented on such products' low user-acceptance (CNET reviews, 2013), but in-lab user-studies have also reported high mental and physical demand (Bailly et al., 2011).

Our approach to address this problem is based on our prior work on motor-intuitive touchless interactions (Chapter 5)—drawing on users' prior knowledge, such as sensorimotor abilities, which is acquired since childhood while continuously interacting with the physical world. We propose Touchless circular menus (TCM) – a contextual circular menu, through which users can select commands by making directional strokes and *crossing* menu options (Figure 6.2). TCM utilize our sensorimotor ability to make directional strokes in mid-air. Therefore, it relieves users from both recalling a vocabulary of precise postures and complying with those pre-defined poses.

In a two-part, controlled experiment, we first investigated how different triggering locations of TCM affect user performance. Then, we compared between TCM and contextual linear menus with *grab* gestures. Our work contributes the following:

- A command-selection technique that solely builds upon human sensorimotor abilities. Although the menu structure, the menu-triggering mechanism, and the menu-selection delimiter already exist in practice, a combination of these to harness our motor abilities is a novel approach toward designing touchless menu systems.
- We further provide important empirical evidence applicable to the design of touchless user interfaces for large displays.

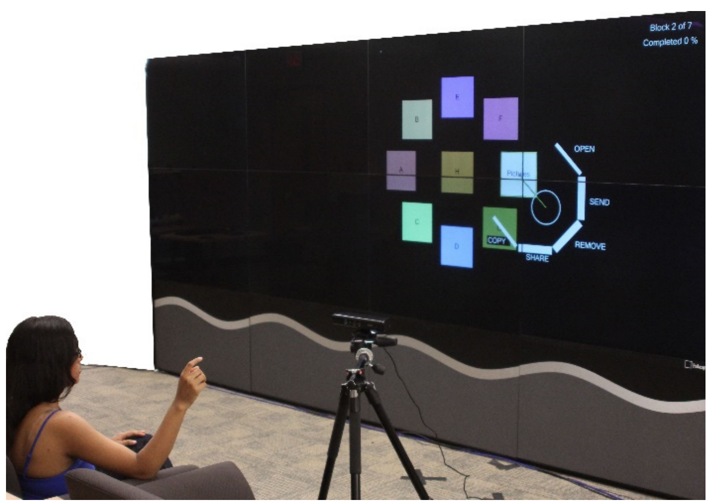

Figure 6.2. Touchless circular menus (TCM) relieve users from the need to comply strictly with system-defined postures and supports command selection by movement in mid-air.

Our results show that the performance of TCM depends significantly on their triggering locations on the visual display, suggesting an effect of our *asymmetric motor abilities* on touchless interactions with large displays. Our experiments also suggest that

TCM is more efficient and causes less workload than command-selection techniques using strict postures, such as *grab*.

6.1.1. Background

Freehand input techniques are increasingly becoming popular due to the recent advances in markerless motion tracking and improved gesture-recognition techniques. The growing popularity of touchless interactions stems from its expectation as something *natural* to use. While critics have repeatedly refuted such a claim of inherent naturalness to this modality, researchers have explained that naturalness of touchless modality lies in the actions enabled and settings (or communities of practice) that give meaning to such actions (O'hara et al., 2013). In a similar line, designers have been encouraged to find naturalness in users, rather than in interaction techniques or interface components (Wigdor & Wixon, 2011).

Another research domain, which investigates how to design interfaces that are intuitive to use, has proposed the intuitive interaction model (Blackler & Hurtienne, 2007). Their model explains how different levels of prior knowledge—from innate abilities to expertise—and their unconscious application define an interface's intuitiveness. For example, any interface that uses *motion to attract attention* (e.g., inertial scrolling) taps into our innate abilities to respond toward movement; while advanced software features often require a certain level of expertise. Until now, all touchless interaction techniques have been proposed as an extension of what has proven efficient for mouse-based, pen-based, or multitouch interfaces (Bailly et al., 2011; Lenman, Bretzner, & Thuresson, 2002). Our design approach for touchless interactions uses human abilities to inform interface components.

6.1.2 Command-selection techniques

Command-selection techniques have been studied for decades (Tables 6.1.a and 6.1.b). Different menu techniques have been proposed for point-and-click (Callahan, Hopkins, Weiser, & Shneiderman, 1988; Kabbash, Buxton, & Sellen, 1994; Kurtenbach & Buxton, 1994; Pook, Lecolinet, Vaysseix, & Barillot, 2000) and multitouch systems (Lepinski, Grossman, & Fitzmaurice, 2010). The major difference between other interactive systems and touchless systems is the device-free nature of the later. Due to the absence of a device, freehand interaction lacks control and precision (Lepinski et al., 2010). Hence, it becomes important to consider the strength and limitations of human motor abilities while extending any device-based menu-techniques to touchless systems.

Table 6.1.a. Different features of some device-based menu techniques that have been widely studied.

	Traditional linear menu (Kabbash et al., 1994)	Pie menu (Callahan et al., 1988)	Marking menus (Kurtenbach & Buxton, 1994)
Uni/ Bimanual	one-handed	one-handed	one-handed
Shape	vertical	radial	radial
Menu triggering mechanism	not applicable	press and hold mouse	press and hold mouse/stylus
Menu selection delimiter	release the mouse button	release the mouse button	release the mouse button/stylus
Gesture semantics	none	none	scale-invariant directional strokes
Menu breadth	4 (later studies suggest 8)	8	12 (later studies suggest 8)
Expert mode	No	No	Yes

Table 6.1.b. Different features of some device-based menu techniques that have been widely studied.

	Control menus (Pook et al., 2000)	FlowMenu (Guimbretiére, & Winograd, 2000)	Toolglass (Bier, Stone, Pier, Buxton, & DeRose, 1993)
Uni/Bimanual	one-handed	one-handed	two-handed
Shape	radial	radial	radial
Menu triggering mechanism	press and hold mouse/stylus	press mouse/stylus	non-dominant hand positions the widget
Menu selection delimiter	moving a threshold distance from the menu-center (no crossing of any interface element)	re-entering the menu-center	mouse click with the dominant hand
Gesture semantics	none	none	none
Menu breadth	8	8	~70 x 70 pixel Toolglass sheet
Expert mode	Yes	No	No

Prior research (Bailly et al., 2011; Guimbretière & Nguyen, 2012; Lenman et al., 2002) proposed touchless menus by extending successful device-based menus (Tables 6.2a and 6.2b). From their evaluation, researchers report interesting findings on how our motor abilities limit touchless interactions. In an informal testing (Guimbretière & Nguyen, 2012), researchers found that a 3D marking menu was most efficient when users were not required to make accurate 3D marks: Users found it difficult to gauge a 3D angle. Bailly, et al (2011) reported that most users had difficulties constraining their gestures in a 2D plane. These observations suggest human limitations to perform 3D movements accurately in mid-air. Most importantly, this emphasizes our premise that designing touchless menus require more than a mere extension of device-based menus.

Table 6.2a. Different features of touchless menus for distant and near-surface interactions

	[1]Linear menu (Bailly et al., 2011)	[1]Marking menu (Bailly et al., 2011)	[1]Finger-Count menu (Bailly et al., 2011)
Uni/Bimanual	one-handed	one-handed	two-handed
Shape	vertical	radial	vertical/ radial
Triggering mechanism	opening the hand toward the display	none	none
Menu selection delimiter	closing the hand	closing the hand	closing both hands at the same time
Gesture semantics	opening and closing hand	strokes, and closing hand	finger combinations with both hands, and closing hand
Menu breadth	8	8	5
Expert Mode	No	Yes	Yes

[1]Interactions from a distance; [2]Interactions near surface

It is also important to identify the features of touchless menu techniques that require different considerations than device-based techniques. For example, with pen-input or multitouch surfaces, triggering a menu is straightforward: Users put the pen down or touch the surface with fingers. Similarly, command-selection is delimited by breaking contact with the interface. In device-based paradigms, both linear and radial menus are common. Now without the guidance of a device, we are faced with the obvious questions: What would be an efficient triggering mechanism or a menu selection delimiter? Can we accurately make directional movements in mid-air to operate a radial menu?

All existing touchless menu techniques (Tables 6.2.a and 6.2.b) employ hand-postures (e.g., grab, finger-count, or pinch) for menu-invocation and menu-selection. Only Guimbretière & Nguyen (2012) investigated scale-invariant marks as an alternative menu-selection delimiter but reports its limitations due to 3D angular movements. Bailly, et al (2011) reported no significant difference in accuracy for linear and marking menus. Alternative to these existing techniques, we propose a touchless menu system that relieves users from both recalling a precise vocabulary of hand postures and strictly complying with them.

Table 6.2b. Different features of touchless menus for distant and near-surface interactions

	[1]Roll-and-pinch menu (Ni et al., 2008)	[2]Bimanual marking menu (Guimbretière & Nguyen, 2012)	[1]Touchless Circular Menu
Uni/Bimanual	one-handed	one-handed	one-handed
Shape	radial	radial	radial
Triggering mechanism	thumb-to-forefinger pinch gesture	middle or index-finger pinch (non-dominant hand)	reaching the ROI of a target
Menu selection delimiter	releasing the pinch	releasing the pinch	moving passed the boundary of any interface element (crossing)
Gesture semantics	rolling the wrist, and pinching with fingers	pinch, and 3D directional strokes (non-dominant hand)	**none**
Menu breadth	8/12	26/ 48/ 52	5
Expert Mode	No	No	No

[1]Interactions from a distance; [2]Interactions near surface

6.1.3. Designing touchless circular menus (TCM)

During our qualitative exploration phase, we looked for human capabilities that could relieve users from the burden of complying with pre-defined hand-postures. It would save users from recalling a fixed vocabulary of gestures and from maintaining positions optimal for the pose-recognizer. We found that users can reliably make directional gestures in mid-air, a sensorimotor ability that we frequently use in our everyday lives, such as during conversations or to give directions. Since such everyday movements happen unconstrained in 3D space, we observed the same problem as reported earlier (Bailly et al., 2011; Guimbretière & Nguyen, 2012; Hespanhol et al., 2012): users' obvious difficulty in gauging 3D angles accurately. We mitigated this problem by shifting the burden of users' input to the interface—interpreting users' 3D translation by its orthographic projection on the 2D display. Based on our ability to make directional strokes in mid-air, and informed by some of the successful features of device-based menus, we designed iteratively a contextual menu system for large displays: Touchless circular menus (Figure 6.3).

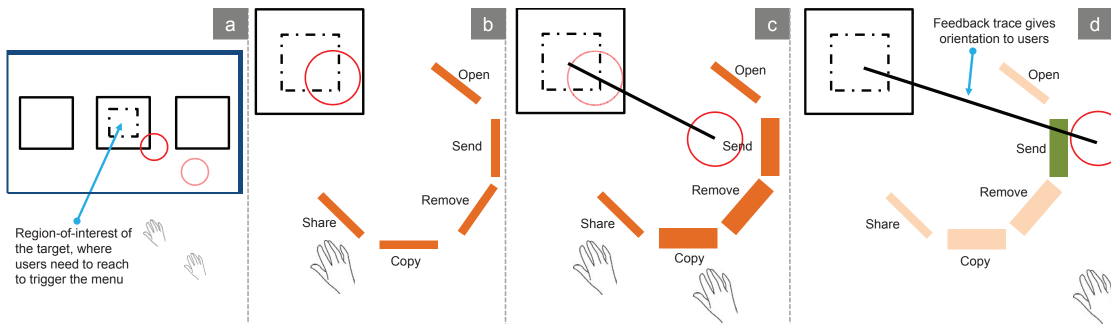

Figure 6.3. Touchless circular menus: (a) a user approaches a target, (b) and reaches the ROI of the target. TCM appear against the user's direction of approach. (c) The user makes a directional stroke towards TCM, (d), and selects a command by crossing it. The selected menu option changes color to indicate a successful command-selection.

Menu invocation

To trigger the contextual menu, a user must cross the region-of-interest (ROI) of a display object. The ROI can be of any symmetrical shape around the center of the target, with its size directly proportional to the technique's sensitivity. To support rapid exploration without accidental invocation of the menu, the menu appeared against the users' direction of movement. So if users would reach the ROI of a display object from the top, or left, the menu would appear against their direction of approach: at the top-left corner of the target. Users can then make a directional stroke toward the command (see Figure 6.3) and select it by crossing; but if they continue in their direction of movement the touchless circular menu would disappear.

Command selection by crossing

To select a command after triggering the menu, users cross it using a stroke in the command's direction. Device-based marking menus are scale invariant, and a mark's angle is interpreted to select commands. However, when extended to touchless techniques (Bailly et al., 2011; Guimbretière & Nguyen, 2012), this implies the need of a posture-based menu-invocation, and a posture-based menu-selection delimiter, which negates the typical advantages of marking menus. Hence, it is not surprising that Bailly et al. (2011) reported similar accuracy for touchless implementations of linear and marking menus.

Until the crossing happens, users can cancel TCM by moving in any direction away from the triggered menu. To allow easy escape routes, we designed the structure of TCM as a semi-circular array of options appearing at the top-left or the bottom-right corner of the target. As users approach the menu, to give them orientation, a trace is drawn connecting the target and the users' hand position. Based on Fitts' law (Fitts, 1954), to improve users' pointing performance, we designed the menu options to increase in amplitude as users approached them. To provide further feedback, menu options changed color when selected by crossing.

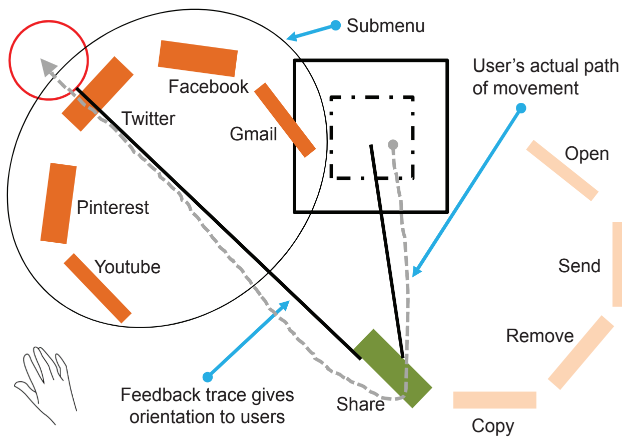

Figure 6.4. The second-level menu in TCM (dashed path represents the user's actual movement).

Accessing submenus

Currently, our menu design scales up to two levels (5 x 5), with users performing continuous strokes (Figure 6.4). When users cross a command in the root menu, a submenu appears opposite to it, pivoted around the center of the selected command. To operate the submenu, users then change their track and cross another command. In device-based hierarchical menus, submenus appear in the same direction of the root menu. Due to the lack of precision and control of freehand movements, TCM require users to make inflections in their continuing trajectories, and thereby avoid accidental command-selections. Users can dismiss a submenu by continuing in their direction of movement after selecting a command from the root menu. In the following sections, we discuss our experiments with single-level TCM.

6.1.4. Experiment 1: Evaluating touchless circular menus

TCM are contextual menus for large displays, and ideally they are expected to perform optimally across the entire display canvas. Hence, we conducted a controlled experiment to investigate how effectiveness and efficiency of TCM is affected by their triggering locations on the visual interface.

Hypotheses

Our menu design was motivated by our abilities to make directional strokes in mid-air. When we move our arms in mid-air, biomechanical properties of the human body (such as the position of the forearm relative to the upper body) affect how accurately and quickly we can make arm-movements. Certain arm-postures result in a more static equilibrium of the body and hence are more comfortable than others. The absence of any guidance device, such as a remote (Nancel et al., 2011) or a wand (Cao & Balakrishnan, 2003), further aggravates the control and the precision of such mid-air movements (Nancel et al., 2011). Based on these theories, we made the following hypotheses:

H1: Triggering location will affect the efficiency of TCM.

H2: Triggering location will affect the effectiveness of TCM.

Furthermore, in our experimental setup, based on our sensor's tracking specifications and pilot testing, we ensured that the tracking performance was optimal across all triggering locations.

Apparatus

The high-resolution large display (Figure 6.2) integrated by Fakespace Systems comprises of eight 50" projection cubes laid out in a 4 x 2 matrix. It is driven by a single computer. Each cube has a resolution of 1600 x1200 pixels, resulting in a 160" wide by 60" high display with over 15.3 million pixels. Our goal was to evaluate TCM as a potential user interface component using off- the-shelf motion-capture sensors. We used a Kinect for Windows to track users' hand position and recognize gestures. Though this system is limited from a technological perspective, we wanted to evaluate user performance with a commodity-range camera. The experiments were written in C# running on Windows 7, and were implemented with OpenNI 1.4 SDK and PrimeSense's NITE 1.5.

Task

To test our hypotheses, we designed a menu selection task informed by the ISO standard 9241-9 (ISO, 2002). On a large interactive display (Figure 6.2), participants were shown a circular arrangement (594-pixel diameter) of 9 equally sized (320-pixel) squares, aligned to the horizontal and the vertical center of the background (Center, N, NW, NE, S, SW, W and E). Participants' task was to invoke TCM for a (randomly generated) white square and select the 'Remove' command by crossing (Figure 6.2). The ROI was set to 256 pixels, and TCM's diameter was set to 400 pixels.

Procedure

We recruited 15 right-handed participants (4 females) from a university campus, with 8 participants having prior familiarity with touchless gestures, and 11 participants below 30 years of age.

Participants sat on a comfortable couch (Figure 6.2) at 2.25 m away from the display (~1 m away from the sensor), and took 20-30 minutes to complete all trials. Prior to the experiment, all participants completed 3 blocks of practice trials. Throughout the experiment, participants were required to take at least a 10-second break in between each block. Trials were randomized within subjects. In summary, the study design was as followed: 9 triggering locations (trials) x 7 blocks x 15 participants = 945 total trials.

Participants hovered over a 'Start' circle, to begin a block. Trials were defined as a successful selection of the 'Remove' command. We recorded performance time, command-selection errors, and encouraged participants to make comments about the menu. Time was measured from the target's appearance to a successful command selection. A command-selection error was recorded, when participants selected a wrong command from the triggered menu. When a command-selection error occurred, 'error' was flashed on the display, and the trial restarted. Participants received a $20 gift card for 2 hours of participation.

Successful Trigger Rate. TCM are contextual menus. They are triggered when users reach the ROI of a target and are dismissed if users move away from the triggered menu. During selecting commands, users may inadvertently dismiss the menu before selecting any command and re-trigger it again. To understand how unwanted menu dismissals affect users' efficiency, we defined successful trigger rate as successful triggers / (successful + unsuccessful triggers). Successful triggers: When users trigger a menu, and continue to select a command from the triggered menu. Unsuccessful triggers: When users trigger a menu, but the menu is dismissed before any command is selected. Obviously, a high successful trigger rate would increase a menu's efficiency as users would not have to re-trigger it.

Results

Performance time was normally distributed, but error rate and successful trigger rate were not. We used repeated measures ANOVA (and its nonparametric version) for data analysis.

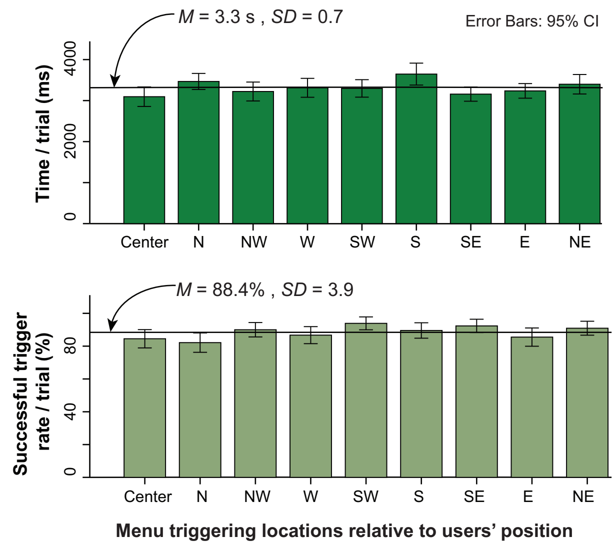

Figure 6.5. The triggering location of TCM significantly affected selection time and successful trigger rate.

Triggering location {Center, N, NW, NE, S, SW, W and E} had a significant effect on task time, $F(6.9, 718.12) = 4.74$, $p < .001$ (Figure 6.5). Planned contrasts revealed that both north (3466 ms) and south (3646 ms) locations took significantly more time than the center location (3095 ms), $p < .001$. We found a significant learning effect across blocks, $p < .01$. Participants were about half a second faster in the last block than the first block. Menu triggering location also significantly affected successful trigger rates, $\chi^2(8) = 18.83$, $p < .05$. Across all triggering locations, the average successful trigger rate per trial was 88.4%. H1 was supported.

Triggering location did not significantly affect error rate. The average error rate (participants selecting a wrong command from the menu) across all triggering locations was 2.7%. During 88.5% of the command-selection errors, users chose the nearest neighbor options ('Send' or 'Share', Figure 6.2). H2 was not supported.

Apart from the initial novelty effect that excited the participants, they appreciated the use of fewer muscles in the crossing gesture. However, participants also commented on the lack of control: "*I felt I had to rush to select the menu option*" and precision: "*It was sometimes difficult to be precise.*" Overall, users liked the feedback language of the menu: "*It feels like the menu is a bow, and I am aiming an arrow to select one of the* *options.*" Finally, some users were excited about their performance: "*I was surprised that I could do so well.*"

Discussion

From our user study, we learned the following about TCM:

- Depending upon the menu's triggering location on the display users' control on their hand movements varied significantly.
- A visual comparison of successful trigger rates and time spent in command-selection across all triggering locations (Figure 6.5) reveals that unsuccessful triggers were not the sole reason behind the variability in efficiency of TCM. For example, at certain triggering locations (such as, N and S), users did not lose the triggered menu more than the average but spent more than average time in command-selection. One possible explanation is that participants had to put more physical effort, thereby spending more time at certain triggering locations.

Overall, our results suggest that touchless interaction with large displays is significantly affected by the *asymmetric nature of human motor abilities* (control and precision).

The average efficiency of TCM was 3.3s and accuracy 97.3%. Bailly et al. (2011) reported performance measures for a linear menu as 6.6s (94.2%), marking menu as 7.2s (95.3%) and finger-count menu as 8.5s (93.4%). Our results cannot be directly compared to Bailly et al. (2011) because we used different experimental tasks and menu hierarchy (details in Table 6.4). However, this is an encouraging result. Although such performance time is higher than the menu-selection time in typical Xbox games, it is important to note that Xbox gamers are continually (visually) guided to position themselves in an optimal space—in front of the sensor (2–3 meters)—so that the sensor can track users' entire body (Microsoft, 2014). TCM was implemented using hand tracking algorithms that did not require whole body tracking.

Limitations. Due to sensor limitations, when participants moved their arms very fast, tracking points were lost, thereby causing unwanted menu dismissals. This may have decreased the successful trigger rate for TCM. As TCM do not require any static poses, their invocation and selection suffer from certain limitations. To provide users escape routes, the breadth of TCM is limited to 5. Moreover, menu invocation is not tolerant to target overshooting (when hand movements trail the eye gaze), and may cause accidental invocations if users decide to change the direction of movement for target acquisition. One possible approach to mitigate these limitations is using explicit dynamic gestures (e.g., lassoing or *pigtails*, Hinckley, Baudisch, Ramos, & Guimbretiere, 2005) as a menu-selection delimiter. As a delimiter, dynamic gestures would be more efficient than static poses as users would not have to halt-and-execute a pose, but fluidly end the selection. Furthermore, we do not foresee a large number of commands in large-display touchless interfaces, as they are not fitted for intense editing but suited for exploratory data browsing. As the location of menu options in TCM depends on users' direction of movement, users cannot exploit spatial memory to locate them. However, TCM appear at either the NW- or the SE-corner of a target in a symmetric layout (as mirror images of one another). Further research is required to understand if users can exploit this symmetry to locate menu options in TCM.

External Validity. Our findings can be generalized to settings, where users are sitting away from a large display, facing the display, and within the sensor's tracking range. Since sitting posture already constrains our arm movements to a certain extent (e.g., when leaning back or resting the elbow), we expect similar or better user performance of TCM in a standing posture.

Experiment 1 suggested an encouraging performance of TCM. However, it was unclear how this performance would compare with menu systems that employ static postures, especially in similar settings.

6.1.5. Experiment 2: Touchless circular menus vs. linear menus

Experiment 1 focused on investigating the performance of TCM across different triggering locations. In experiment 2, we investigated how the overall user experience of TCM compares with contextual linear menus using *grab* gestures.

Contextual Linear Menus. With linear menus, participants could point-and-select a display object by doing a *grab* gesture. They would do a *grab* gesture by making a fist, and opening their hand again (Figure 6.6). To trigger the linear menu, users would do a *grab* gesture on a target, and the menu would appear to its right. Then users would select a command by doing another *grab*. In this technique, *gesture registration* happens with the first *grab*; then *gesture relaxation* follows, where users point to a command, and then *grab* gesture is *reused* to select that command (Wu, Shen, Ryall, Forlines, & Balakrishnan, 2006)

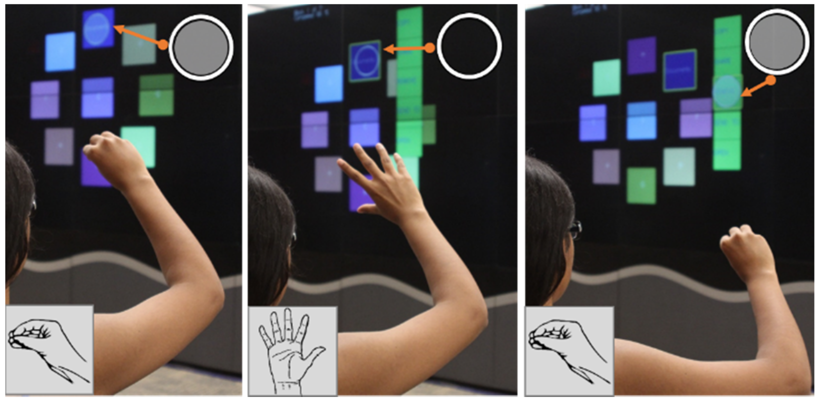

Figure 6.6. To trigger linear menus, users made a grab gesture on the target by closing (left) and opening their hand (center). A command was then selected by another grab gesture (right).

Hypotheses

Based on previous research and our pilot studies, we made the following two hypotheses:

H3: Compared with TCM, the linear menu design uses more muscle groups (Werner et al., 1997) and involves *reuse* of gesture primitives (Wu et al., 2006). We predicted TCM would be more efficient than linear menus.

H4: We hypothesized that TCM would be easier to use than linear menus because of the use of more muscle groups (Werner et al., 1997) in *grab* pose than in a *crossing* gesture.

Task and procedure

In experiment 2, we compared the user experience of TCM with that of linear menus. Thus, it used the same experimental task, procedure and evaluation metrics as experiment 1. However, due to sensor limitations, we designed the command-selection task for linear menus only at six different locations (Center, N, NW, NE, W and E). A successful *grab* gesture on the target triggered the linear menu 200 pixels right and 700 pixels top from the top-right corner of the target. The menu consisted of five equally sized (256-pixel) squares (Figure 6.6), and the participants' task was to select the 'Remove' command by a *grab* gesture. Users would dismiss the linear menu if they performed a *grab* gesture anywhere outside the menu. We recorded menu dismissals to calculate the menu's successful trigger rate. Self-reported system usability scores were recorded using SUS, and perceived workload using NASA-TLX. Participants responded to SUS (Brooke, 1996) after using each menu (except questions 1, 2 and 6). After using both the menus, they completed the NASA-TLX scale (Hart & Staveland, 1988). Since we conducted both parts of our experiment on the same day, and with the same participants, the menu condition was counter-balanced. Participants took a break of about 10 minutes in between sessions. Trials were randomized within subjects.

Apparatus. The linear menu experiment was written in C running on Windows 7 and was implemented with OpenNI 2.2 SDK, NITE 2.2 and Windows SDK 1.7. For the *grab* gesture recognition, we used PrimeSense's Grab detector library (PrimeSense Labs, 2013).

Results

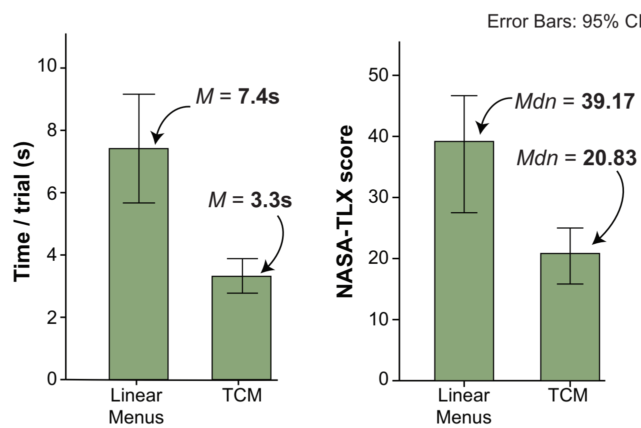

Figure 6.7. Compared with linear menus, users were more efficient with TCM, and perceived lower overall workload.

TCM are More Efficient than Linear Menus. TCM (M = 3.3s, SD = .7) were more than twice as fast as the linear menus (M = 7.4s, SD = 2), $t(14) = 7.43$, $p < .001$, r = 0.89. *H3* was supported. However, there was no significant difference in successful trigger rates between TCM (*Mdn* = 89%, *IQR* = 9.55) and linear menus (*Mdn* = 92%, *IQR* = 7.62).

TCM are Less Effective than Linear Menus. TCM (*Mdn* = 1, *IQR* = 3) were significantly less effective than linear menus, Z = 2.68, p < .01, r = .69. With TCM, on an average, users made about 3 errors per 100 trials. Given the lack of precision and control associated with freehand movements, 97.3% accuracy is an encouraging result. Leaving out the outliers, users made no command-selection errors with linear menus.

TCM elicit Less Workload than Linear Menus. System usability scores were not significantly different between TCM (M = 82.86, SD = 13.58) and linear menus (M = 72.62, SD = 19.96). However, overall workload was significantly higher for linear menus (*Mdn* = 39.17, *IQR* = 19.17) than TCM (*Mdn* = 20.83, *IQR* = 9.17), Z = 2.89, $p < .01$, r = .75. When the NASA-TLX scale was analyzed separately, we found significant differences between the menus regarding physical demand, temporal demand, and effort. *H4* was partially supported.

User Comments. Compared with TCM, linear menus received mixed user reactions. A male participant younger than thirty was enthusiastic: "*This is how I envision using touchless gestures*." A female participant over fifty said: "*It was a lot of effort*." She pointed out that Arthritis patients would find it difficult to do *grab* gestures.

Discussion

In experiment 2, we compared the overall user experience of TCM with linear menus (Table 6.3). TCM utilize our sensorimotor abilities to make directional strokes in mid-air, while the linear menu was designed to emulate the current *status quo*: contextual menu using *grab* gestures.

Table 6.3. Contrasting characteristics of touchless circular menus vs. contextual linear menus.

	Touchless circular menus	Contextual linear menus
Menu selection delimiter	crossing the boundary of an interface element	*grab* gesture
Triggering mechanism	reaching a pre-defined ROI of a display object	*grab* gesture
Gesture types	(dynamic) stroke	(static) grab
Technology	hand tracking	hand pose recognition
Shape	radial	linear

Surprisingly, the linear menu had an accuracy of 100%, which means participants did not select any wrong command from the triggered menu. Nevertheless, participants lost the triggered menu in 8% of the trials. Our videos revealed that while *grabbing* a menu command, participants often moved their hands horizontally away from a specific command (right or left); thereby dismissing the menu. As they did not move their hands up or down, and the linear menu options were stacked vertically (Figure 6.6), command-selection errors did not occur. Compared with linear menus, TCM had an accuracy of 97.3%. This maybe because:

- The options in linear menus were 256-pixel squares and more than eight times wider than the options in TCM (306 pixels in length, 30 pixels in breadth).

- In linear menus, users triggered the menu with a *grab* gesture. They also selected a command using another *grab* gesture. Between these two gesture registrations, users could move their hands freely around the display. However, for TCM, after the menu is triggered, users could inadvertently move their hand and select a wrong command. Unlike linear menus, TCM required users to constrain strictly their freehand movements after triggering the menu.

6.1.6. Conclusion

Overall, we learned the following from part II of our study:

- In our experimental settings, TCM were more efficient but less effective than linear menus. TCM elicited significantly less workload than linear menus.
- Compared with linear menus, participants were more than two times faster with TCM, but there was no significant difference in successful trigger rates between them. This suggests that menu-triggering by reaching the ROI (88% accuracy) performed on par with menu triggering by *grab* (92%). Moreover, participants seemed to spend more time with linear menus due to more effort required in performing a *grab* gesture than a *crossing* gesture.

Limitations. Capabilities of our tracking sensor limit our results. An ideal gesture recognition algorithm may have made the linear menus more efficient than TCM. In this work we proposed a touchless menu system that does not employ any pose-recognition techniques, but performed on par with current available menu techniques (Table 6.4). Furthermore, with future improvements in tracking capabilities, we expect that TCM will outperform linear menus because it builds on users' previously learned skills of making in-air directional gestures. Aimed at a preliminary understanding of a touchless menu system that does not employ any pose-recognition techniques, both our visual interface and task were simple (always selecting the 'Remove' command from a single-level menu). Future research is required to assess the user experience of TCM in more realistic usage scenarios.

External Validity. Large displays are becoming popular in consumer electronics (e.g., interactive TVs), healthcare settings and public spaces. Touchless gestures offer a promising interaction modality for these novel devices. Our proposed touchless menu system uses dynamic gestures for selecting commands on large displays while interacting from a distance.

Table 6.4. Performance measures across touchless menus.

Touchless menu system	**Menu options**	**Average time**	**Average accuracy**
[1]Linear menu (Bailly et al., 2011)	8 x 8	6.6s	94.2%
[1]Marking menu (Bailly et al., 2011)	8 x 8	7.2s	95.3%
[1]Finger-count menu (Bailly et al., 2011)	5 x 5	8.5s	93.4%
[2]Contextual linear menu	5	7.4s	100%
[2]TCM	5	3.3s	97.3%

[1]Participants standing; [2]Participants sitting.

Prior work on touchless interaction with large displays contributed interaction techniques that require users to comply with pre-defined postures. Our research suggests that dynamic gestures—such as simple crossing—when coupled with human sensorimotor abilities—such as making directional strokes—is more efficient than posture-based techniques. Specifically, whereas existing touchless menu systems for selecting commands from a distance are posture-based (Bailly et al., 2011; Lenman et al., 2002), we introduced a novel touchless menu system (TCM) for large displays, which solely uses our ability to make directional strokes in mid-air and relieves users from recalling a vocabulary of gestures.

Our comparative study suggests that TCM are more than two times efficient than contextual linear menus using grab gestures. Users also perceived less workload with TCM. However, TCM caused 3% more errors than linear menus. This may happen because, unlike linear menus, TCM required users to constrain strictly their freehand movements after triggering the menu. Touchless input is inherently imprecise, which is exacerbated by the lack of haptic feedback. To improve the accuracy of touchless selections, we now explore *pseudo-haptic feedback* (Lécuyer, Burkhardt, & Etienne, 2004). To that end, we first introduce an interaction technique, interface topographies, in the next section, and then report empirical studies in Chapter 7.

In evaluating TCM, we also found that the asymmetric nature of human motor capabilities significantly affected the efficiency of our proposed touchless circular menus. We expect this effect to be pervasive in touchless interactions with large displays, which requires further investigation. The design of future touchless interfaces can be informed by identifying these asymmetric motor abilities. In Chapter 8, we revisit motor control and study handedness in touchless interactions.

6.2. Interface topographies

The lack of haptic feedback in touchless interactions causes users' gestures difficult to control and to move off interface elements unintentionally. This lack of control increases users' effort to perform accurate actions, such as steering and targeting. To mitigate this problem, we introduce *interface topographies*: pseudo-haptic textures that modify cursor movements to guide touchless interactions along the contours of interface content (e.g., data visualizations) or components (e.g., widgets). We designed and implemented three topography primitives—*holes*, *valleys,* and *pits*.

Designing appropriate touchless interfaces is still in its infancy (Guimbretière & Nguyen, 2012; Dostal et al., 2014). We still need basic interface standards to design touchless widgets and to support frequent user tasks, such as searching, targeting, or steering. In the last section, we introduced and evaluated a touchless command selection technique. We found that target acquisition was imprecise due to a lack of control in steering toward the menu option (see section 6.1.5). In this section, we explore interface affordances in steering-targeting tasks—tasks requiring trajectory-based movements before target acquisition.

The lack of haptic guidance reduces touchless precision because users are exclusively dependent on other forms of sensory feedback, such as visual, auditory, or proprioception (Nancel et al., 2011; Chapter 5). Hence touchless gestures add abundant fluency to an interaction scenario, but fail to provide fine-grained, pixel-level, motor guidance for accurate interaction tasks. To compensate this lack of haptic feedback, researchers have explored visual, auditory and tactile feedback in touchless target-acquisition tasks (Lehtinen et al., 2012; Van Mensvoort, 2002). But as touchless interfaces mature to provide traditional controls (e.g., menus or scrollbars) and its contents call for trajectory-based tasks (e.g., interacting with data visualizations, such as heat maps or bubble charts; or drawing), understanding how to improve the precision of trajectory-based interactions becomes essential.

To improve the precision of touchless interactions, specifically trajectory-based interactions on large displays, our paper makes the following contributions:

- We introduce *Interface Topographies*: pseudo-haptic textures (e.g., holes, valleys, and pits) that can virtually overlay on interface controls (menu or scrollbar) or interface contents (e.g., data visualization structures, such as nodes, lines, or regions) and constrain the touchless cursor's imprecise movements during navigation—conveniently along the structure of the interface content or control.

- We implement three topography primitives, *holes*, *valleys*, and *pits* (Figure 6.8), and introduce two techniques to augment their effectiveness, *adaptive* and *additive* topographies. Adaptive topographies dynamically adapt their shapes to constrain imprecise cursor movements. Additive topographies combine multiple primitives to suit a specific trajectory-based interaction.

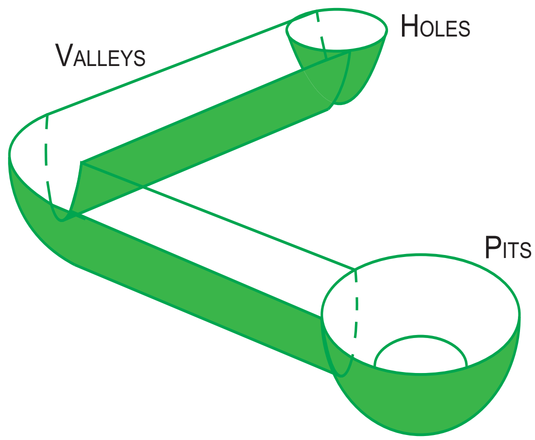

Figure 6.8. Topography primitives (e.g., holes, valleys, or pits) operate as virtual surfaces that overlay on an interface and modify cursor movements to improve the precision of touchless interactions.

6.2.1. Background

Touchless interactions suffer from lower accuracy than device-based interactions—due to the absence of haptic feedback. To improve user experience, haptic feedback is explored across different input modalities for over two decades. Under this umbrella term of *haptic feedback*, however, lays important distinctions (cf. Table 1, Oakley, McGee, Brewster, & Gray, 2000), such as force feedback, tactile feedback, or pseudo-haptic feedback. *Force* feedback relates to the mechanical production of sensations perceived by the human kinesthetic system (muscles, tendons, and joints). *Tactile* feedback pertains to the cutaneous sense of pressure perceived by the skin surface. *Pseudo-haptic* feedback is proprioceptive and can simulate haptic effects (e.g., slopes or friction) with a passive input device (Lécuyer et al., 2004). It is generated by purposely violating the isometric mapping between the motor space and the display space. For example, while crossing a bump on an interface, the cursor is artificially slowed down, forcing the user to move the device more, thereby creating an illusion of force feedback (Lécuyer et al., 2004).

In the last two decades, researchers explored force feedback to haptically augment desktop interfaces. For example, a PHANTOM haptic device (SensAble Technologies Inc., now part of Geomagic) was used to create a haptically enhanced XWindows Desktop (now commercially available as Geomagic Touch™ X). When such haptically-enhanced GUIs were evaluated on target-acquisition tasks, researchers found that haptic effects reduced errors and overall workload, but did not affect task completion times (Oakley et al., 2000). Furthermore, basic research looking into the interplay of perceptual processes in haptic feedback showed that between the vertical and the lateral force information, the lateral forces dominate other perceptual cues (Robles-De-La-Torre & Hayward, 2001).

Lately, integrating tactile feedback in user interfaces has garnered increased attention. To generate a broad range of different tactile sensations, researchers proposed using electrovibration—controlled electrostatic friction—in touch interfaces (Bau, Poupyrev, Israr, & Harrison, 2010). In touchless interfaces, approaches to generate tactile feedback followed two broad categories—wearable sensor gloves and feedback projected on users' unadorned hands. For example, to augment visual search, wearable gloves with vibrotactile actuators was proposed for dynamic tactile cueing (Lehtinen et al., 2012). Approaches to project tactile feedback included the use of air voxels in AIREAL (Sodhi et al., 2013), and ultrasonic waves in UltraHaptics (Carter et al., 2013) and HaptoMime (Monnai et al., 2014). Evaluations of touchless systems with continuous tactile feedback for target-acquisition did not report significant performance benefits (Foehrenbach et al., 2009). But touchless gestures with dynamic tactile feedback in search-and-select tasks significantly improved task completion times, when visual complexity was high (Lehtinen et al., 2012). Rest of the touchless systems with tactile feedback focused on psychophysical experiments, but not empirical evaluations (Carter et al., 2013; Sodhi et al., 2013).

Pseudo-haptic feedback does not require additional hardware, and is explored in a number of applications (surveyed in Lécuyer, 2009). Approaches to implement pseudo-haptic feedback include adding tiny displacements to the cursor (ActiveCursor, Van Mensvoort, 2002), using Flash-based animation templates (PowerCursor), or varying the cursor's motion with a transfer function (pseudo-haptic textures; Lécuyer, A., et al., 2004). For example, a transfer function adjusts the Control/Display (C/D) ratio of an input device to simulate textures and generate the illusion of lateral forces—same forces that dominate the perception of textures in force-feedback devices (Robles-De-La-Torre et al., 2001). Empirical studies suggest that users can feel pseudo-haptic textures, such as *bumps* and *holes* (Lécuyer et al., 2004). Apart from guiding target acquisition in GUI, pseudo-haptics have also been used for content creation tasks in digital drawing. For example, Kinematic Templates amplify or dampen the cursor's speed to guide users' strokes into drawing circles, parallel lines, or soft edges (Fung et al., 2008).

In the context of touchless systems past work focused on tactile feedback, but not pseudo-haptic feedback, and evaluated mostly target-acquisition tasks. Tactile feedback has the benefit of improving touchless experience, but with the exception of empirical evaluation of visual search tasks, past research did not evaluate its benefits on user performance. On the other hand, in GUI, force feedback improved user performance in target-acquisition (Oakley et al., 2000) and steering-targeting tasks (Dennerlein, Martin, & Hasser, 2000). Building upon prior work, we address this research gap: We introduce a touchless interaction technique with pseudo-haptic feedback and evaluate its performance in steering-targeting tasks (Accot & Zhai, 1997).

6.2.2. Designing interface topographies

Interface Topographies are pseudo-haptic textures overlaid on interface contents or controls that manipulate the touchless cursor's motion to improve interaction precision. The cursor is manipulated by adjusting the C/D ratio with a transfer function. This transfer function is formulated according to the geometrical structure of the interface content or the interface control. For example, a visualization presented on a touchless interface is morphed into a virtual topography. Thus, by overlaying a virtual terrain atop the visualization that reflects the visualization's structure (e.g., rows, columns, or regions), interface topographies can constrain users' imprecise touchless gestures during steering and targeting tasks.

We implemented topography using *height maps*. Height maps vary the cursor's speed to conjure up a feeling of traveling over uneven topographical surfaces (Lécuyer et al., 2004). To simulate different topographies, we first propose topography primitives and discuss their design parameters. We then introduce two techniques to augment the effectiveness of interface topography: *adaptive* and *additive* topographies. Finally, we describe a visual feedback routine that augments the pseudo-haptic feedback as users exit a topographical structure.

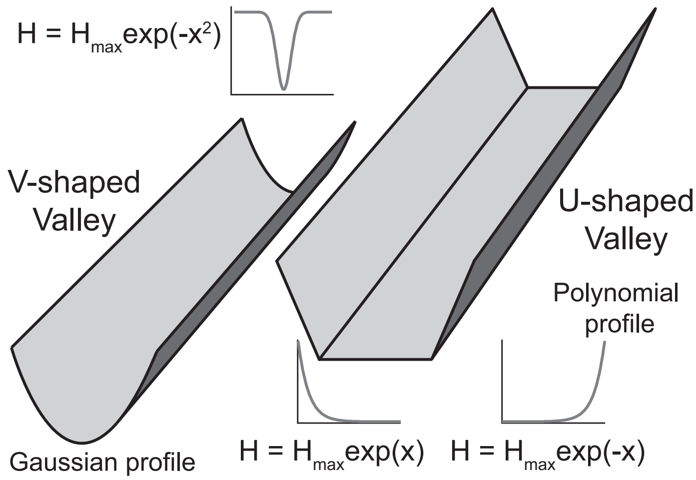

Figure 6.9. Two different types of valleys: V-shaped and U-shaped. (H = current height, H_{max} = maximum height)

Height Maps: Simulating a Topography

Topography of a surface is a function of different heights—a *height map*. Topography can be simulated on a user interface by maintaining a height value associated with each pixel of the screen. A slope is simulated using either a Gaussian profile or a Polynomial profile (e.g., Figure 6.9).

Our algorithm to implement topographies (Figure 6.10) is adapted from Lécuyer et al. (2004): Users' movement in the control space is mapped to the touchless cursor's movement in the display space as a function of the height map of the topography. The *cost of displacement* between two consecutive pixels is determined by their difference in height. When this difference in height is negative, the cost is greater than 1 (i.e., user has to move more in the control space than usual, or *ascend*) and when the height difference is positive, the cost is less than 1 (i.e., user moves less in the control space than usual, or *descends*). Until users' movement in control space exceeds the cost of displacement, the touchless cursor is constrained at its prior position, thus simulating movement over a virtual terrain.

```
AmPx ← Amount of pixels moved in control space
PrevPos ← Previous position in display space
CurrPos ← Current position in display space
T ← Topography constant

ApplyTopography (PrevPos, CurrPos, AmPx)
DO
        NextPixel ← CalcNextPx (PrevPos, CurrPos)
        DiffHeight ← CalcDh (PrevPos, NextPixel)

        IF DiffHeight > 0
              CostOfMovement ← 1 + T × |DiffHeight|
        ELSE
              CostOfMovement ← 1 − T × |DiffHeight|
        ENDIF
        IF AmPx > CostOfMovement
              PrevPos ← NextPixel
              AmPx  ← AmPx − CostOfMovement
        ELSE
              CurrPos = PrevPos
        ENDIF
```

Figure 6.10. Algorithm for traveling height maps (based on Lécuyer et al., 2004).

6.2.3. Topography primitives: holes, valleys, and pits

To match some common geometrical structures, such as points, lines, and circles, we propose using *height maps* (Figure 6.11) to simulate *holes*, *valleys,* and *pits* (Figure 6.8).

Holes: A hole is a narrow, circular depression from a baseline plane that is simulated using mathematical profiles, such as a Gaussian, a polynomial, or a linear profile (Lécuyer et al., 2004). For example, a vertical cross-section of a Gaussian Hole can be computed as: $H = H_{max} \times \exp(-x)^2$, where H is the height of the pixel *x*, and H_{max} is the height of the baseline plane.

Valleys: A valley is a linear depression from a baseline plane (Figure 6.9). The vertical cross section of a valley is either similar to a hole (a V-shaped valley, Gaussian profile) or to a pit (a U-shaped valley, polynomial profile).

Pits: A pit is a wide, circular depression from a baseline plane whose left slope is simulated with an exponential decay function, $H = H_{max} \times \exp(-x)$, and right slope with an exponential growth function, $H = H_{max} \times \exp(x)$. To simulate valleys and pits, we chose a polynomial profile:

FOR each pixel P along the length of the wall

$H = H_{max} \times \exp(-\ Slope \times P)$

ENDFOR

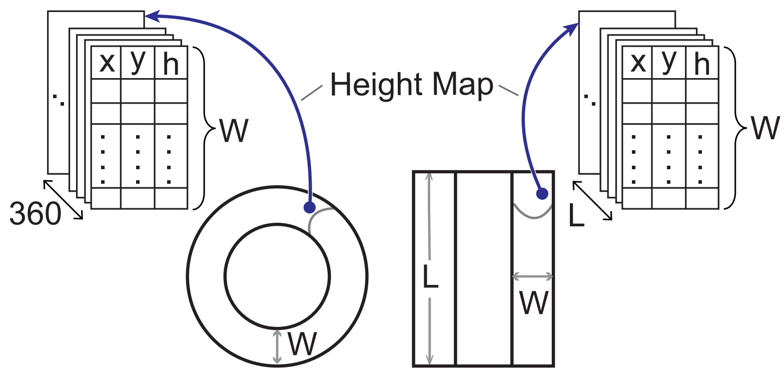

Figure 6.11. A vertical cross-section of a pit or a valley is stored as a height map, with $h = H_{max} \times f(step)$, $\forall$ step: step $\in$ W.

Invoking topography

On-demand invocation: To navigate *interface content* (e.g., a heatmap where a valley is overlaid on a row), topographies are invoked on demand. Such on-demand invocation ensures that topographies assist steering-targeting only after users have determined the navigation task (e.g. which particular row to traverse), and avoid accidental distractions *en route* to a target. On-demand dismissals for interface content topographies should be allowed for densely packed contents, because that allows easy, short movements into adjacent regions without much displacement in the motor space to exit the topography.

Automatic invocation and on-demand dismissal. To operate *interface controls* (e.g., menus or scrollbars), topographies are auto-invoked when users land on a controls' operation zone. For example, a valley is activated along a scrollbar to facilitate precise steering. Since landing may be accidental, and the topographic effect distracting, a "reserved" gesture (e.g., a closed fist or a non-dominant hand raise) allows emergency exit and promptly dismisses the topography.

Topography parameters

The proposed topography primitives are based on four parameters: *Wall Length*, *Slope*, H_{max}, and *T*. A higher value of *T* amplifies the slope of a topography and increases or decreases the cost of displacement (see Figure 6.10). Through iterative tuning, we identified an optimum range of *T* as [100, 500]. A combination of the parameters *Wall Length* and H_{max} play the same role as the parameter *Slope* in simulating a steep or a gradual ascent/descent. For user evaluation, we used H_{max} = 10, *Wall Length* = 5, *Slope* = 0.1, and *T* = 400 for valleys and T = 200 for pits. In the parameter tuning phase, we encountered the following phenomenon: When users moved obliquely to the wall of any topography, they took a longer path to exit; resulting in weaker constraints than when moving orthogonally to the wall. To mitigate this and effectively constrain imprecise touchless interactions, we introduced *Adaptive Topography*.

6.2.4. Adaptive topographies

The slope of topography allows a gradual descent into a hole, a pit, or a valley. However, when exiting, the ascent—the feature that ultimately constrains users—provides different resistance depending on how users move along the wall of the topography. Orthogonal movements provide the intended resistance, but oblique movements provide weak constraints due to small differences in height crossed while traversing the wall (similar to taking the ramp instead of a huge step). Our solution is *adaptive topographies* (Figure 6.12): after users enter a pit or a valley, the inclined walls become vertical, thereby eliminating the possibility for users to make oblique movements that would allow them to unwittingly leave the topography during data browsing. Holes, however, do not require any adaptation, because they map to points that do not require detailed interaction within the topography—instead, holes play the role of transitional stops along interconnected lines.

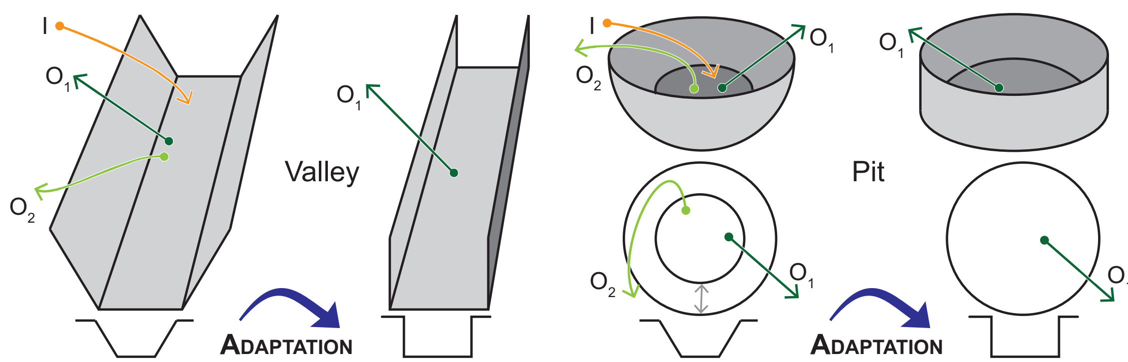

Figure 6.12. Slope of a valley (or a pit) allows users to move gradually into the topography (I). To get out, users can move orthogonal (O1) or oblique (O2) to the wall of the valley (or the pit). Due to small differences in height, however, a long oblique movement along the wall (O2) fails to sufficiently constrain users' touchless gestures. To mitigate this, we introduce Adaptive Topography: after users enter a valley (or a pit), its walls become vertical, thus requiring a higher cost of displacement to move out of the topography, and thereby appropriately constraining users' touchless interactions to the region.

6.2.5. Additive topographies

Similar to interaction primitives constituting interaction controls, topography primitives can be combined to match non-trivial interface content, such as graphical data visualizations. To that aim, we introduce *Additive Topographies*. Because the complex nature of additive topographies may over-constrain users' touchless interaction, we suggest dynamic invocation of primitives in these kinds of scenarios, thereby fostering a more seamless data browsing experience. For example, a graph (with nodes and edges) can be morphed into a set of holes and valleys. At first, the graph would only contain holes to provide a variety of flexible starting points for data exploration. As users enter a node/hole, valleys would be invoked on its connecting edges to guide graph traversal. As a particular edge (with valley) is being traversed, its endpoints would be overlaid with "destination" holes.

Additional visual feedback & offset recovery

Interface topographies provide pseudo-haptic feedback during steering-targeting tasks. This pseudo-haptic effect is generated by purposely violating the isometric mapping of the cursor between the motor space and the display space. For example, the touchless cursor—while ascending out of the topography—ceases to move until sufficient displacement occurs in the control space (see Figure 6.10). Traditionally, prior work on pseudo-haptic feedback exclusively used C/D ratio modification to elicit a sensorial experience—but only in device-based interactions (Lécuyer et al., 2004; Lécuyer, 2009). Our pilot studies explored pseudo-haptic feedback for touchless interfaces and found it as a double-edged sword. Modifying C/D ratio along interface content/controls improved accuracy, and users reported perceiving a "wall" constraining their interactions. Yet, decoupling motor and display spaces disoriented users as the topographic effect prolonged; rather than continuing to move in the control space to fully experience the pseudo-haptic effect, users often confused the "frozen cursor" as a tracking error and halted. This perceived post-error slowing was perhaps exacerbated by strong user expectations of interaction fluidity—common in human-human gestural interactions (Notebaert et al., 2009).

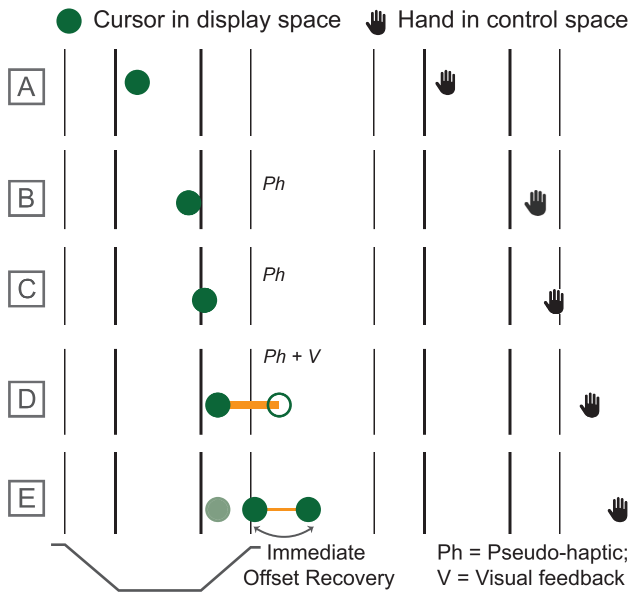

Figure 6.13. Although primarily designed for pseudo-haptic feedback (B, C), interface topographies also provide visual feedback (D) as users exit a topography: When the cursor is halfway ascending out of a topography (D), a secondary cursor shows users' position in the control space and a trail connects the two cursors. On a successful exit from the topography, the two cursors immediately merge to represent users' position in control space (E), thus recovering the control-display offset.

To mitigate this problem, we introduced an additional visual feedback routine for interface topographies (Figure 6.13). As the touchless cursor is halfway ascending out of the topography, a secondary cursor shows users' position in the control space and a trail connects the two cursors. The width of this trail represents the current *cost of displacement* to exit the topography (see Figure 6.10). On a successful exit from the topography, the two cursors immediately merge to represent users' position in control space, thus recovering the *control-display offset*. This immediate recovery of the offset—due to the C/D ratio manipulation—eventually generates a "no man's land". The display space—following the end of a topographic ascent—that corresponds to the excess space traversed in the control space is rendered unusable while exiting the topography (Figure 6.13). Thus, to navigate adjacent regions in densely-packed interface contents, users should employ on-demand dismissal of the topography.

Interface topographies, pseudo-haptic feedback in touchless interaction, are evaluated in a controlled study in Chapter 7. We will revisit touchless input in Chapter 8 and interface affordances in Chapter 9.

Chapter 7. Experiments on pseudo-haptic feedback

This Chapter discusses an empirical study on the effects of pseudo-haptic feedback in touchless steering-and-targeting. We evaluate interface topographies (introduced in Chapter 6) with 17 participants performing two steering-targeting tasks at two levels of difficulty.

As of yet, the most-frequently used evaluation task in touchless is target acquisition (Guimbretière & Nguyen, 2012; Van Mensvoort, 2002). However, other actions, such as navigation or drawing, often require users to perform trajectory-based tasks such as steering. And as trajectory length increases, trajectory-based tasks require more control—even more so in touchless interactions due to the lack of tactile feedback and an input device. Such expected precision in an interaction task is particularly suitable to explore the efficacy of topographies. Thus, in this chapter, we evaluate adaptive interface topographies (see section 6.2.4) in touchless steering-targeting tasks. While prior research evaluated visual (Vogel & Balakrishnan, 2005), auditory (Vogel & Balakrishnan, 2005), and tactile feedback (Lehtinen et al., 2012) in touchless interfaces, this study evaluates a touchless feedback language that includes pseudo-haptic feedback.

Furthermore, we wanted to assess whether using a physical token—not digitally connected to the interface—provides users an advantage of tactile feedback (similar to the use of token in Ballendat, Marquardt, & Greenberg, 2010). Hence, in total, we explored four types of interfaces: *Flat* (no topography), *Token* (with an unconnected physical device working identic to bare hands), *Topography* (with topography primitives, see Chapter 6), and *Topography & Token* (together).

7.1. Hypothesis

Prior empirical studies have found that force feedback in device-based interactions reduces errors and workload, but task completion times remain unaffected (Oakley et al., 2000). Because pseudo-haptic feedback mimics the lateral effects of force feedback (Robles-De-La-Torre, 2001), we expected similar results. We hypothesized the following:

H1: Topography will not affect touchless efficiency.

H2: Topography will increase touchless accuracy.

H3: Topography will increase accuracy more in a difficult than simple task (in terms of the required precision).

H4: Topography will reduce overall workload of touchless interactions.

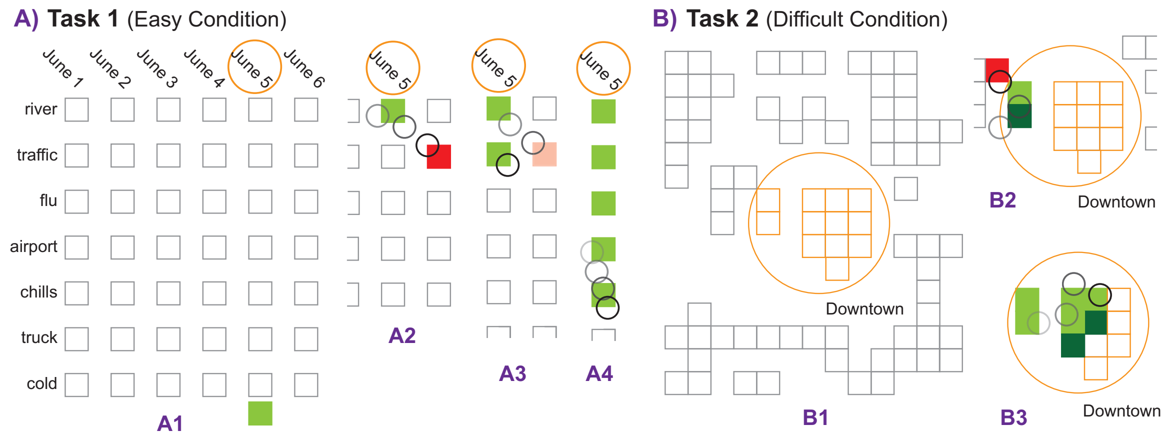

Figure 7.1. Participants performed two steering-targeting tasks on a large display, at two conditions of difficulty. For example, a vertical steering-targeting task on a low-density, contiguous grid in the easy condition (A), and a circular task on a high-density, contiguous non-grid in the difficult condition (B). The target column (A1) or region (B1) was labeled at the beginning of the task. When participants traversed a cell outside the target, it flashed red (A2, A3, B2). Target cells were selected either once (task 1, A4) or twice (task2, B3), and turned green on selection (light green, A4; dark, then light green, B3).

7.2. Method

Participants. We recruited 17 right-handed participants (M_{age} = 24.31, SE_{age} = 1.51, 7 females). Fourteen of them were familiar with Kinect, Wii, or Leap Motion. This study was approved by Indiana University IRB (1411698641) and participants were compensated $20 for their time and effort.

Apparatus. We used a 4.06 m wide and 1.52 m high display with over 15.3 million pixels. The display, integrated by Fakespace Systems, is composed of eight 1.27 m projection cubes (each with a resolution of 1600 x 1200 pixels) laid out in a 4 x 2 matrix, and is driven by a single computer. Instead of using submillimeter-accurate sensors, we evaluated interface topography using off-the-shelf hardware—a Kinect™ for Windows—reflecting more likely real-world configurations. All experiments were written in C#/WPF running on Windows 7, and were implemented with Windows Kinect SDK 1.8.

Tasks and procedure. To test our hypothesis, we designed two abstract steering-targeting tasks (Figure 7.1). Task 1 emulated a vertical steering-targeting task on a contiguous grid structure, similar to steering along a column in a heat map and selecting each cell (Figures 7.1-A1 and 7.2). Task 2 emulated a circular steering-targeting task on a contiguous non-grid structure, similar to steering a circular region of interest and selecting the cells within (Figure 7.1-B1 and 7.3). Both tasks broadly represented steering-targeting tasks in touchless interfaces, where topographies could be overlaid on interface content (data visualization). We did not, however, test the on-demand invocation of topographies on interface content, but only the user experience during the steering-targeting task (H1–4).

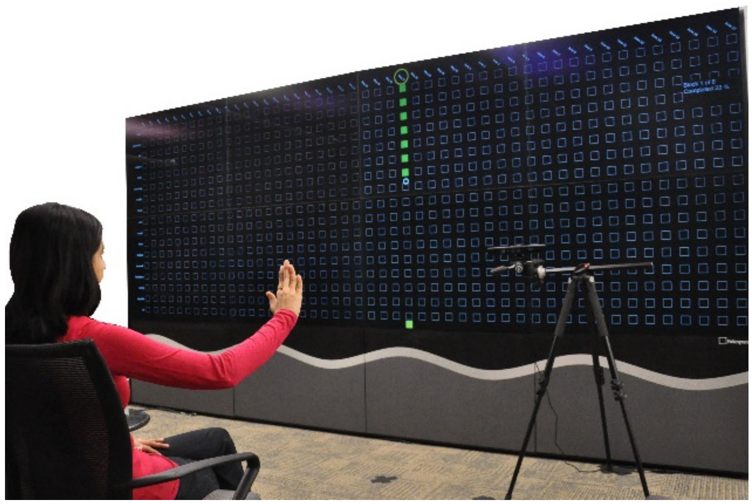

Figure 7.2. The vertical steering-targeting task (contiguous grid) on a large display while sitting at a distance at a low level of density.

In task 1, participants traversed the target column from the topmost to the bottommost cell (64px-square; Figure 7.1-A4). In task 2, participants passed over each cell within the target circle (576px diameter circle) at least twice, which first turned yellow and then green (Figure 7.1-B3). Passing a cell twice represented the typical repetitive interaction during processing information from a visualization. Across both tasks, when a target cell was traversed, it turned green (Figures 7.1-A4 & 7.1-B3); if a non-target cell was traversed, the cell flashed red to indicate an error (Figures 7.1-A2, 7.1-A3, & 7.1-B2). Overall, task 1 required more interaction precision than task 2 due to its implicit spatial complexity.

Figure 7.3. The circular steering-targeting task (contiguous non-grid) on a large display while sitting at a distance at a high level of density.

Moreover, each task comprised of two levels of difficulty—operationalized as spatial density. An easy task was half as densely populated as a difficult task. Overshoots occurred when participants moved out of the target region. To complete each trial, participants selected all cells within a target region (a trial continued until completed successfully). In a repeated-measures within-subject experiment, we measured task completion time (for efficiency) and the number of overshoots (for accuracy) for each of 816 trials: 2 tasks × 4 interfaces × 2 difficulty × 3 ROIs × 17 participants. Trials and tasks were completely randomized within subjects and across subjects.

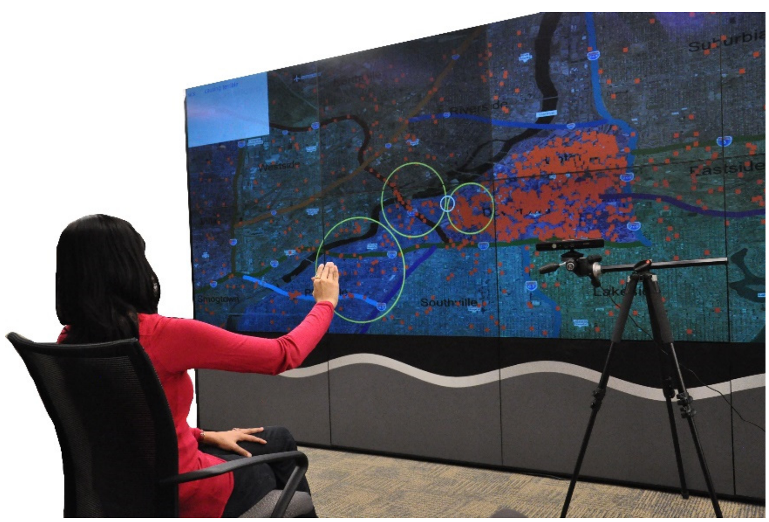

Figure 7.4. An open-ended, exploratory prototype, based on the VAST 2011 Epidemic Spread dataset (Grinstein et al., 2011).

At the conclusion of the controlled tasks, participants interacted with an ecologically-oriented InfoVis task (Figure 7.4), to provide their interaction preference. *Epidemic Spread* was designed using the dataset from the VAST 2011 challenge (a city map, text messages, and metadata, Grinstein et al., 2011). The points of origin of posts with at least one keyword related to the epidemic (or events leading to the epidemic) were shown on the map. As participants browsed over the map, a word cloud displayed all the keywords shared from the positions underlying the participants' cursor (128-pixel squares in all interaction conditions); the font size of a keyword indicated its frequency. We defined *three* ROIs on the map. Using the word cloud, participants tried to identify one major event that occurred in each of those ROIs. During this task (~ 20 minutes), they used *Mouse*, *Flat*, *Token*, and *Topography* techniques, in no particular order.

Participants sat in a ~0.5m high chair, situated 2 m away from the large display (~ 1.5 m from the sensor) and took about an hour to complete the study (Figure 7.2). Participants' movements were mapped from the control space to the display space with 1: 3.75 (baseline C/D ratio). Trials were video recorded. Prior to each task, all participants practiced three trials at each ROI with topography. Participants rested at least 10s after every 3 trials and 10 minutes after completing all trials in a task. They used a whiteboard marker as the token (Expo Original). After completing each task (all trials), participants self-reported their perceived workload using the NASA-TLX instrument. To prevent over-exposing participants to the instrument (16 times), we only measured workload for Flat, Token, and Topography across easy and difficult conditions (6 times). We also logged task completion times, the number of overshoots out of the target region, and the trajectory paths for each trial. Time, overshoots, and paths were measured from the first time participants landed on the target region until trial completion. Task completion times included the time spent overshooting target boundaries and subsequent recovering. Overshoots more than 500 pixels from the target boundaries were discarded as system (sensing) errors. Participants shared overall comments at the conclusion of the tasks.

7.3. Results

For all 17 participants, we analyzed task completion times, the number of overshoots, and overall workload. As expected, task completion times were positively skewed; thus, replications of unique experimental conditions were represented by their median. Our analysis used GLMM with standard repeated measures REML technique. Participants were handled as a random factor. We report F-statistic using type III

ANOVA with Satterthwaite approximation, and pairwise comparisons (using pooled variance) with Holm-Bonferroni correction. We found a learning effect across blocks: For the difficult task, participants performed about 3.7s more slowly in the first block of task 1 than the last block (2.8s more slowly for task 2). As the factors, interface and difficulty, were counter-balanced, this did not adversely affect our analysis.

Task 1: vertical steering in a contiguous grid

Efficiency. We found significant main effects of difficulty, $F(1, 112) = 23.27$, $p < .001$, and interface, $F(3, 112) = 2.87$, $p = .039$, but no significant interaction effect (Figure 9A). Participants took significantly more time to complete the difficult task (M = 20s, $SD = 5.13$) than the easy task (M = 18s, $SD = 4.84$), $p < .001$, $r = 0.50$, which confirmed our manipulations of task difficulty. Pairwise comparisons did not find any significant effect of topography on efficiency (Flat vs. Topography or Token vs. Topography & Token). *H1* was supported.

Accuracy. We found significant main effects of difficulty, $F(1, 112) = 199$, $p < .001$ ($M_{high} = 23.77$, $SD_{high} = 8.47$; $M_{low} = 10.04$, $SD_{low} = 3.15$), and interface, $F(3, 112) = 7.20$, $p < .001$, and an interaction effect of interface × difficulty, $F(3, 112) = 4.48$, $p = .005$ (Figure 7.5-B). Pairwise comparisons indicated tha t in the difficult task, participants made significantly fewer overshoots with topography ($M = 20.38$, $SD = 3.90$) than a flat interface ($M = 26.06$, $SD = 9.30$), $p = .031$, $r = 0.51$, and significantly fewer overshoots with topography & Token ($M = 19.47$, $SD = 4.40$) than Token ($M = 29.18$, $SD = 10.53$), $p < .001$, $r = 0.72$. No significant results were found for the easy task. Post hoc Tukey-tests did not find significant differences between Flat and Token for either easy or difficult task. *H2* was partially supported—only for the difficult task. *H3* was supported.

Workload. We found no significant effect of the interface on workload, $p = .132$ (Figure 7.5-C). However, interface significantly affected perceived effort, $p = .025$, but not perceived performance, $p = .793$. *H4* was not supported

Task 2: circular steering in a contiguous non-grid

Efficiency. We found a main effect of difficulty, $F(1, 112) = 459$, $p < .001$, but no significant effect of either interface or the interface × difficulty interaction (Figure 7.5-D). Participants took significantly more time for the difficult task (M = 29s, $SD = 6.55$) than the easy task (M = 16s, $SD = 2.82$), $p < .001$, $r = 0.92$. *H1* was supported.

Accuracy. Only difficulty had a significant effect on the number of overshoots, $F(1, 112) = 132$, $p < .001$ ($M_{high} = 30.43$, $SD_{high} = 17.62$; $M_{low} = 5.51$, $SD_{low} = 3.27$) (Figure 7.5-E). In the difficult task, participants made more overshoots in Token (M =

33.12, *SD* = 14.64) than Topography & Token (*M* = 24.59, *SD*=19.59), with results approaching significance, *p* = .056. Similar to task 1, post hoc tests did not find significant differences between Flat and Token for either easy or difficult task. Neither *H2* nor *H3* was supported.

Workload. Interface did not significantly reduce participant's overall workload, *p* = .292 (Figure 7.5-F). Interface neither significantly affected participants' perceived effort, *p* = .708, nor perceived performance, *p* = .902. *H4* was not supported. Figures 7.5-D & 7.5-H exemplifies how topographies constrained participant's interactions.

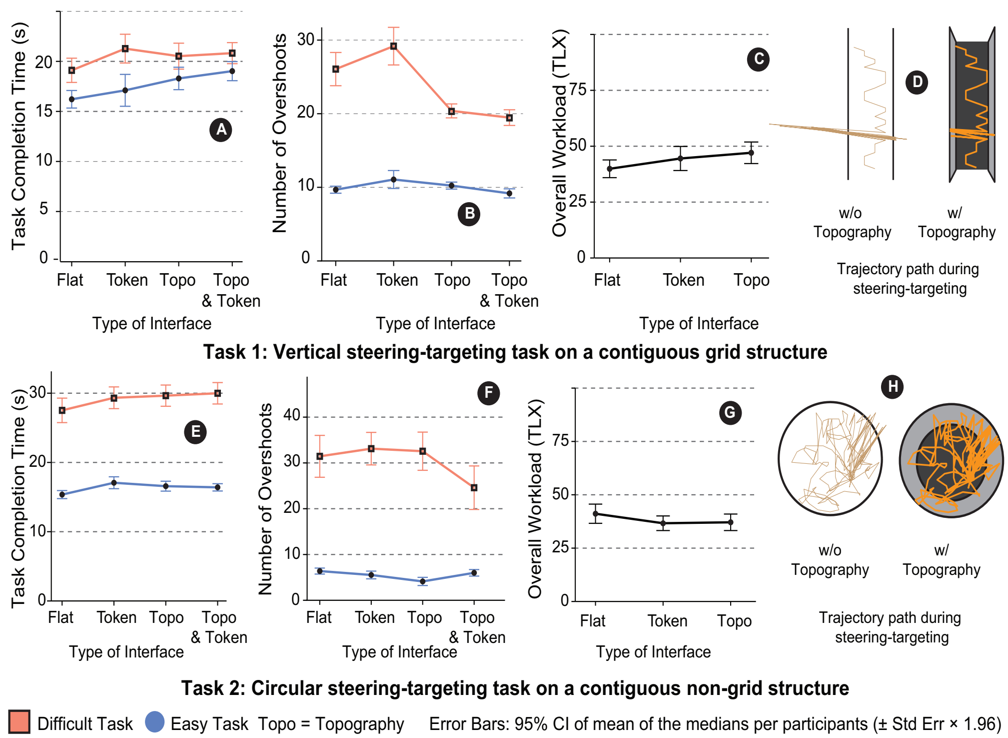

Figure 7.5. In task 1, interface topography significantly reduced the number of overshoots, but not overall workload, thus improving participant's interaction precision. In task 2, interface topography did not significantly affect efficiency or overall workload; but participants made fewer overshoots with topography & token than token alone, with results approaching significance, p = .056.

Performance across task 1 and task 2

To explore the effects of interface and difficulty on performance across tasks, we fitted a hierarchical mixed-effects model with participant as a random factor, task as a random factor, and task difficulty nested within the task factor: time/overshoot ~ interface × difficulty + rand(participant) + rand(task/difficulty).

Our fitted model found a main effect of interface on task completion time, $F(3, 246) = 3.80$, $p = .010$, but no interaction effect. Pairwise comparisons did not find any significant effect of Topography on efficiency. *H1* was supported. Number of overshoots was significantly affected by interface, $F(3, 262) = 6.74$, $p < .001$, and interface × difficulty, $F(3, 262) = 4.92$, $p = .002$. Pairwise comparisons indicated that across tasks, in the *difficult* condition, participants made significantly fewer overshoots with Topography & Token ($M = 14.71$, $SD = 7.68$) than with Token ($M = 21.53$, $SD = 11.42$), $p < .001$, $r = .62$; and significantly fewer overshoots with Topography ($M = 16.66$, $SD = 6.31$) than Flat ($M = 20.25$, $SD = 10.13$), $p = .024$, $r = .38$. No significant results were found for the *easy* task. *H2* was partially supported. *H3* was supported. Across the two tasks, post-hoc tests did not find any significant differences between Flat and Token for either *easy* or *difficult* task.

Based on participants' open-ended summary comments, they preferred a traditional interface (Flat) for completing task 1, which required a strict vertical steering-targeting. In this context, the lack of constraints "Helped me to move easily in the complex [more dense] matrix" [P13], provided "more freedom to move around" [P3], and felt both "free and smooth" [P4] and "faster and less constrained" [P9].

However, this freedom came with a perceived cost. Participants reported that completing the task without any constraints felt slower and required more effort due to the lack of precision in their interactions: "I'd get really far away [from the target region]" [P10]; Flat interface was "harder to control; I needed to concentrate more" [P6]; "[I] wasted a lot of time as I was moving away" [P7]. These responses are notable since we found no significant quantitative differences in task completion time between the flat and topography conditions; however, they do resonate with the significantly fewer overshoots in the topographic interface.

Six participants preferred using the flat touchless interface than using a physical token, because then they had "no physical objects to use" [P14], a perceived advantage in regards to system simplicity. Although the token was not digitally connected to the interface, participants reported that a flat touchless interface "was not as accurate as token, but required less focus." When using the token, five participants felt guided. For example, they reported that the token "helped me focus" [P2], "felt like painting" [P13], and "gave me more support to move along the track" [P3]. For these participants, the token increased their confidence in interaction: "I felt like I had more control" [P6] and "I felt I could more accurately focus my interactions" [P16]). However, four participants reported increased fatigue when using a token and perceived the system's response to this style of input as being less precise than a flat interface. For example, participants reported that "I didn't like token. It felt less precise, less accurate" [P4]; "I was more tired with the token; my arm felt rigid" [P16]; "I expected to be more precise using the token compared with my hand. But it was not happening, so there was a break of expectations" [P17]; and "It felt more natural with hand, but more straight with token" [P5, referring to the linear, vertical gesture needed to complete task 1].

The shortcomings of the traditional touchless interface that participants disliked were somewhat mitigated when using the interface topography. Participants perceived the cursor constraints that topography was designed to provide: "The subtle corrections were making me efficient" [P16]; "It was smoothing my… movements and keeping me in line" [P12]; "I did not have to continually focus on my hands afraid of getting out. It was easy to learn and helped me to be precise" [P17]. However, two participants found the guidance to be too constraining, especially for task 1: "It was too much constraining in the vertical movement" [P17]; "I didn't like the fact that I was not in control" [P4]; One participant [P6] found the trail feedback to be distracting, while some participants reported it to be useful: "[It] lets you know that you're out of the region. You can see if you are going out with your peripheral vision" [P4]; "I knew when I was out of line" [P6]. Overall, in task 1, the participants' responses on whether topography was useful or too constraining were mixed. But almost all found topographies to be helpful in task 2. These qualitative findings suggest an interesting nuance between the user perceptions and the user performance that emerged from our quantitative results. In task 1, participants were significantly more accurate with the topographic interface compared with the traditional touchless interface; in task 2, however, the improvement in accuracy with topographies only approached significance.

7.4. Discussion

Overall, results partially supported our main hypotheses. Interface topography improved the precision of touchless interactions (H2) in task 1, but did not significantly affect user efficiency (H1). Overall workload, however, was not reduced using topography. For task 1 (Figure 7.5, top row), which required greater overall precision than task 2, topography significantly increased interaction accuracy for difficult trials. However, for difficult trials of task 2 (Figure 7.5, bottom row), increase in accuracy because of topography only approached statistical significance. Moreover, across tasks 1 and 2, accuracy of easy trials was unaffected by topographies. Thus, H3 was supported, H2 was partially supported, but H4 was not supported. In what follows, we discuss some relevant implications of our findings.

Implications for touchless interaction research

Topographies improve Touchless Interaction Precision. Past studies provided empirical evidence that users can successfully identify macroscopic pseudo-haptic textures, such as bumps and holes, when simulated by modifying the C/D ratio of the mouse cursor (Lécuyer et al., 2004). Building upon pseudo-haptic textures, we (1) introduced topography primitives, (2) demonstrated adaptive techniques to morph interface content using topography primitives, and (3) provided empirical evidence that topographies do increase the accuracy of touchless interactions. Practical implications of our findings include (1) providing dynamic guidance in natural user interactions where typical tactile feedback is lacking or insufficient and (2) building virtual affordances for natural user interface components such as touchless menus or widgets. Notably, we found evidence that virtual constraints such as interface topography are effective only when an interaction is sufficiently difficult, i.e., for operations requiring high interaction precision (H3).

Fluency vs. Control in Touchless Interfaces. Our findings suggested a dichotomy between users' perceived performance (as self-reported) and observed performance (as captured in log). While topographies improved users' accuracy in both task 1 (significantly) and task 2 (approaching significance), most users reported topographies as being helpful in task 2, but often too constraining in task 1 (see Figure 7.5). This tension between perceived interaction fluency and input control signifies a familiar but crucial tradeoff in the evolution of interaction techniques. For example, the mouse allows more input control because of its characteristic resistance that its movement across a surface provides; but pen, touch, and touchless gestures provide more interaction fluency—a hallmark of natural user interfaces (NUI). Thus, attempts to provide more input control in natural user interfaces, such as touchless, implies immediately compromising some of its 'naturalness' or interaction fluency. As per our findings, the balance between perceived interaction fluency and input control is not absolute, but is highly situated. It depends on the nature and the difficulty of the task (i.e., required precision), the optimization of user feedback, and on the contingent break of user's expectations that occurs when novel systems substantially augment the ability of users but do not behave as smoothly as expected. We showed that pseudo-haptic feedback improves touchless accuracy. But, further research is required to understand how to optimally tradeoff between the perceived interaction fluency and input control in touchless interactions—with input control mediated by feedback and user abilities.

Implications for touchless interaction design

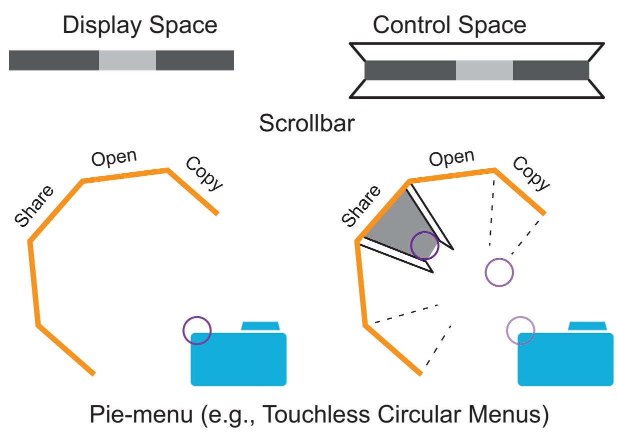

Figure 7.6. Designing widgets for touchless interaction that improves users' steering-targeting precision: A valley overlaid on a scrollbar (above) and valleys adaptively invoked along menu options of a pie-menu (below).

Apart from the key conceptual implications stemming from our findings, important design implications include offering static or adaptive virtual constraints in common interface controls—such as scrollbars or touchless menus (Figure 7.6). For example, to improve users' steering accuracy, a valley can be overlaid on top of a scrollbar or along a menu option of the touchless circular menus (TCM, see Chapter 6). TCM were found to be significantly more efficient—but less accurate—than linear menus, because users had to constrain their freehand movements between triggering a menu and steering toward a menu option.

Limitations. Our study's findings are limited by the capability of our off-the-shelf tracking sensors, which were intended to reflect current, widely available technologies.

We evaluated topographies using simple, abstract tasks in a controlled lab setting. Further research is required to assess their benefits during repeated invocation and dismissal of interface topographies.

7.5. Conclusion

In sum, we designed, implemented, and evaluated interface topographies—pseudo-haptic textures that increase the accuracy of touchless interactions in difficult steering-targeting tasks. During these tasks, users made fewer overshoots with topographies than with a traditional touchless interface. Specifically, our contribution is threefold. First, we implemented three topography primitives—holes, valleys, and pits—that map to common geometrical primitives, points, lines, and regions (Chapter 6). Second, with adaptive and additive topography, we demonstrated how these primitives can be combined to morph non-trivial interface content into topographies. Third, we provided empirical evidence suggesting that interface topography improves accuracy of touchless interactions, but do not affect users' overall workload or efficiency.

Until now, we have explored solely users' dominant hand. But bimanual touchless interactions can further complement that vocabulary of touchless gestures (Grandhi et al., 2011; Guimbretière & Nguyen, 2012; Nancel et al., 2011). To that end, in the next Chapter we study handedness in touchless input.

Chapter 8. Motor control: handedness and hemispheric asymmetry

In this chapter, we shift back to touchless input from feedback. We had studied touchless input before; Chapter 5 looked into human capabilities and introduced motor-intuitive interactions based on image schemas and sensorimotor abilities. In the previous Chapter, we evaluated pseudo-haptic feedback on touchless accuracy in steering-targeting tasks. Until now, we have explored single-handed manipulation—with users' dominant hand. But in user input, that is half the story. In interactive computing, bimanual techniques involving users' nondominant (or non-preferred) hand has been extensively studied and particularly found useful as a mid-air input in performing 3D object manipulation (Hinckley, Pausch, Goble, & Kassell, 1994). Because of the significance of the nondominant hand in interaction techniques, this chapter explores handedness and transfer of skill between dominant and nondominant hands. Broadly speaking, we study motor control in touchless. This research is grounded in the more traditional literature (nearly a century of research; Adams, 1987; Magill & Anderson, 2007) on motor behavior (e.g., motor control and learning, Todor & Doane, 1978).

In interactive computing, the research on bimanual methods follows two primary directions, understanding the performance constraints of the nondominant hand and evaluating user experience of bimanual interaction techniques. For example, in a seminal work, Guiard (1987) introduced the Kinematic chain model to explain why most human skilled manual activities involve two hands, and how they play different roles in the division of labor. He pointed out that the two manual motors representing the two hands work as if assembled in a serial fashion. This hierarchical division of role between the two hands results in the manipulative efficiency of bimanual gestures. Much work in HCI has been built upon Guiard's model to propose efficient bimanual interaction techniques; more recently in device-based mid-air and touchless (Hespanhol et al., 2012; Nancel et al., 2011; Pyryeskin et al., 2012). Bimanual methods require the use of the nondominant hand. The nondominant hand can also be used in single-handed interactions, when more precise tasks demand the use of the dominant hand and may involve a different interaction modality (e.g., a tablet or a pen; Guimbretière & Nguyen, 2012). We do not look into touchless bimanual interactions. Instead, this chapter focusses on exploring the performance constraints of the nondominant hand in touchless steering and targeting tasks. Most recently, Jude, Poor, & Guinness (2014) found that touchless pointing performance improved more than mouse and touchpad, and had the lowest degradation between hands. In this chapter, first, we review the relevant literature on motor behavior, then present a set of hypotheses, and finally discuss the results from a two-stage empirical study.

8.1. Background

The work in this chapter builds upon two important concepts in motor behavior—transfer of learning and motor control. Earlier, we had briefly touched upon motor control when discussing visual feedback in Chapter 4. Visual feedback plays an important role in motor control and assists in motor learning and retention (Sigrist et al., 2013). This Chapter delves deeper into the properties of touchless input and the lack of haptic feedback. Here, we aim to understand two aspects of touchless input: how insufficient feedback affects the performance of nondominant hand (motor control) and how prior training with dominant hand impacts the nondominant hand's performance (bilateral transfer of learning).

Motor control

Fitts's law (1954) is arguably the most frequently used theoretical premise in HCI (Wright & Lee, 2013). This classic finding represents movement time in relatively long movements as a function of the distance to the target and the size of the target. A less studied, but equally important, finding is that these parameters, distance and target amplitude, do not influence the *choice reaction time*—the time interval between the appearance of a signal and the beginning of the response (Ells, 1973; Fitts & Peterson, 1964). Choice reaction time reflects the time to program a response—that is the preparation time among a set of alternate options.

When studying choice reaction time in aimed movements, Klapp found (1975) that the time is influenced by the required precision of the movement only for shorter amplitudes, but not in longer ones (increased time for higher requirement of precision). Following this findings, Klapp (1975) concluded that long aimed movements are under feedback control, while very short movements are pre-programmed and simply ballistic; and that Fitts's law do not hold for very short movements, but for long movements that comprise of a fast initial movement, a pause, and then a slow final movement (ballistic and corrective movements, Casiez et al., 2008). Thus, information processing while control of aimed movements involves either *feedback control* or *preprogrammed motor plans*—although their roles may not always be mutually exclusive.

Motor behavior research contends that human cerebral hemispheric specialization influences motor performance (Cohen, 1973; Durnford & Kimura, 1971). For example, Todor and Doane (1978) reported data partially supporting that performance capabilities of the hands mirrors the dominant processing type of their contralateral hemisphere: in right-handed individuals, the left hemisphere and the right hand is dominant for sequential information processing or feedback-controlled motor actions, whereas the right hemisphere and the left hand is dominant for parallel information processing or preprogrammed motor plans. They found that the left hand (in right-handed individuals) fared superior in aimed movements that required greater preprogramming of motor plans. However, the right hand was not found superior in movements requiring the greatest demand for feedback control.

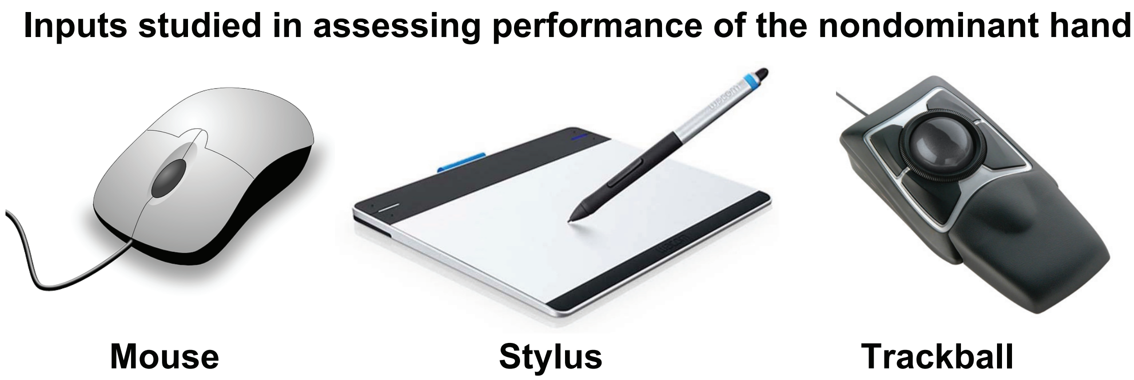

Figure 8.1. Kabbash, MacKenzie, & Buxton (1993) built upon Todor and Doane’s (1978) work and studied the user performance of right-handed individuals with mouse, stylus, and trackball.

In HCI, Kabbash, MacKenzie, & Buxton (1993) built upon Todor and Doane’s (1978) work; they studied the user performance of right-handed individuals with mouse, stylus, and trackball. As expected, mean movement times in pointing and dragging tasks were significantly greater with left than right hand. However, interestingly, they found the accuracy of left hand in trackball-dragging was superior to the right hand, in contrast with the opposite finding in mouse and stylus. Kabbash et al. (1993) explained this finding of a left-hand advantage to the finger-thumb independence requirement in the trackball-dragging task and the superiority of the right-hand to perform paired finger flexions (Kimura & Vanderwolf, 1970). However, it is also interesting to note that compared with mouse and stylus, trackball has an increased degrees of freedom for dragging and fewer feedback constraints (Figure 8.1). The left hand advantage, thus, may also be because of the greater role of preprogrammed motor plans in operating the trackball over the mouse and the stylus.

Because of the lack of haptic feedback, we argue that touchless relies more on preprogrammed motor plans than feedback control. Thus, in right-handed individuals, the left-hand performance will be superior in steering-targeting tasks. We hypothesized the following:

H1: Nondominant hand's accuracy will be significantly greater than the dominant hand.

Transfer of skill

The role of cerebral hemispheres in motor control and learning is also evidenced in the transfer of motor skill learned in one hand to the other hand (Criscimagna-Hemminger, Donchin, Gazzaniga, & Shadmehr, 2003). Such bilateral transfer of motor control means better speed and accuracy with one hand, when the particular skill was practiced with another hand. Researchers have suggested that such inter-arm generalization from dominant to nondominant hand is caused by neural elements within a cerebral hemisphere tuned to both the right and left hands (Criscimagna-Hemminger et al., 2003). However, such bilateral transfer of learning is asymmetric in nature (Malfait & Ostry, 2004; Teixeira, 2000). For example, when acquisition of motor skills has a strong perceptual component (e.g., timing) transfer between dominant (or preferred) and non-dominant hand is symmetric (Teixeira, 2000). But when motor skill is strongly effector-dependent, such as exerting force control, transfer of learning is asymmetric—only from dominant to nondominant hand (point-to-point reaching movements, cursor launching, etc.). Motor skills generalize from dominant to nondominant hand in a variety of tasks, such as reaching movements, pointing, rhythmic tapping, or wrist-flexion movement (Teixeira, 2000).

Touchless targeting and steering have a strong effector component. Thus, we expected a transfer of learning from dominant to nondominant hand. Furthermore, because the dominant hand is superior in feedback processing, we hypothesized that additional pseudo-haptic feedback (similar to force control) will augment learning in the dominant hand (Todor & Doane, 1978), and that will further increase the skill transferred to the nondominant hand. We hypothesized the following:

H2: Prior training with right hand will improve the nondominant hand's accuracy than without training.

H3: Prior training with right hand and additional feedback control will improve nondominant hand's accuracy than without training.

Building upon prior results, we hypothesize effects on task accuracy, not task completion times (Kabbash et al., 1993; Teixeira, 2000).

8.2. Method

*Participants***.** We study handedness and transfer of learning in touchless in two consecutive experiments. Experiment 1 was conducted along with the experiments evaluating pseudo-haptic feedback (Chapter 7). In Chapter 7, we report user performance of interface topographies—an interaction technique drawing on pseudo-haptic feedback—with right-handed participants using their dominant hand (right hand). Following that study, each participant further completed another session on using topographies with their non-dominant hand (left hand). Among the 17 right-handed participants (M_{age} =24.31, SE_{age} = 1.51, 7 females) taking part in this study, fourteen of them were familiar with Kinect, Wii, or Leap Motion. We recorded the user performance of right hand ($RH_{control}$) and left hand following training with right hand and pseudo-haptic feedback (LH_{rhf}). Experiments were conducted in December 2014.

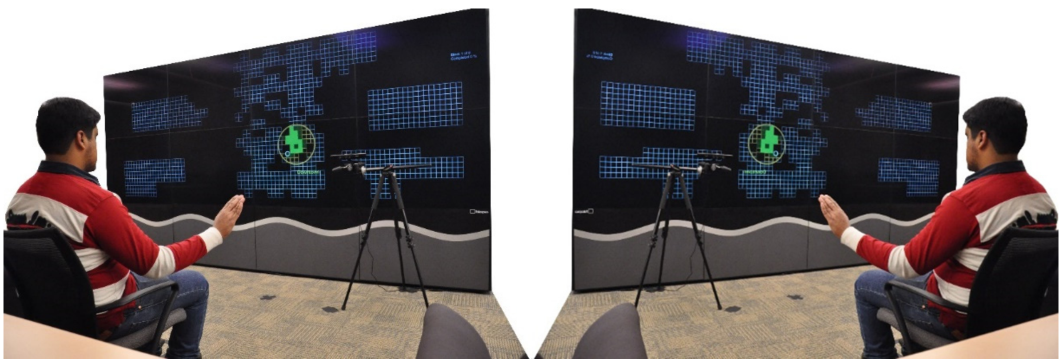

Figure 8.2. We studied handedness in touchless interactions with a circular steering-targeting task (same task used in the experiments evaluating interface topographies, Chapter 7). Right-handed users completed the task at a high level of density (high difficulty) on the large display while sitting at a distance.

We conducted experiment 2 in two sessions. In the first session (about 30 minutes) 16 right-handed participants (different than those recruited for experiment 1, M_{age} = 30.44, SE_{age} = 2.28, 9 females, two familiar with Kinect) completed tasks using their left hand. They revisited the lab (at least three days apart) to participate in the second session. In session 2 (about an hour), participants completed experimental tasks first using their dominant hand (right hand) and then their left hand. In this session, pseudo-haptic feedback was not available when using the right hand. The study was approved by Indiana University IRB (1411698641) and participants were compensated $15 for their time and effort. Experiments were conducted in May, 2015.

Apparatus. Study setup was the same as the experiments evaluating pseudo-haptic feedback (Chapter 7). We used the 4.06 m wide and 1.52 m high display with over 15.3 million pixels and a Kinect™ for Windows; our experiments were written in C#/WPF running on Windows 7, and were implemented with Windows Kinect SDK 1.8.

Tasks and procedure. In Chapter 7, we found that pseudo-haptic feedback improved task accuracy for difficult tasks. So experiment 2 only included difficult tasks or steering-targeting in a high-density condition. To ensure user performance was not affected by boredom and excessive fatigue, half of the participants (n = 8) completed task 1 and another half task 2 (see section 7.2 for details). Task 1 was a vertical steering-targeting task on a contiguous grid structure, similar to steering along a column in a heat map and selecting each cell and, and task 2 a circular steering-targeting task on a contiguous non-grid structure, similar to steering along a circular region of interest and selecting the cells within. Like experiment 1 (similar to Chapter 7), experiment 2 was repeated-measures and within-subject. Total number of trials in experiment 1 used toward this study is 816: 1 hand (left) x 2 tasks × 4 interface repetitions × 1 difficulty × 3 ROIs × 17 participants. In experiment 1, each participant completed both task 1 and task 2. Since the interface repetitions (completely randomly balanced) were a mix of with and without pseudo-haptic feedback, data from experiment 1 contributed to the condition LH_{rhf}, left hand following training with right hand and pseudo-haptic feedback. Total number of trials in experiment 2 was 576: 3 hands (left in session 1, right then left in session two) x 1 task x 4 interface repetitions × 1 difficulty × 3 ROIs × 16 participants. In experiment 2, each participant either completed task 1 or task 2. The interface types for session two with right hand (in experiment 2) was all without pseudo-haptic feedback; they were repeated to ensure the same number of trials prior to left-hand usage.

Participants sat on a ~0.5m high chair, situated 2 m away from the large display (~ 1.5 m from the sensor) and took about an hour to complete the study (Figure 8.2). Participants' movements were mapped from the control space to the display space with 1: 3.75 (baseline C/D ratio). Trials were video recorded. Prior to each task, all participants practiced one trial at one random ROI without pseudo-haptic feedback. Participants rested at least 10s after every 3 trials and 10 minutes after completing all trials in a task.

Measures. Task completion times included the time spent overshooting target boundaries and subsequent recovering. Task accuracy was operationalized as the number of errors, overshooting target boundaries. Overshoots more than 500 pixels from the target boundaries were discarded as system (sensing) errors.

8.3. Results

For all participants, we analyzed task completion times and a number of overshoots. For both experiments, data violated parametric assumptions ($p < .01$, Shapiro–Wilk test). Thus, we use Wilcoxon Signed Rank Test to compare dominant and nondominant hand's performance and report Pearson's *r* for effect size (with continuity correction). We first report results of experiment 1 that evaluated the user performance of right hand ($RH_{control}$) and left hand following training with right hand and pseudo-haptic feedback (LH_{rhf}). In experiment 2, we report results from the user performance of right hand ($LH_{control}$) and left hand following training with right hand and no additional feedback (LH_{rh}). For analysis within both experiments, dependent two-group Wilcoxon Signed Rank Test is used. When comparing performance across the two experiments, we use independent two-group Wilcoxon Rank Sum Test (with continuity correction). All missing data (owing to random sensing lapses) were treated as missing completely at random (MCAR). Data analysis was done in R version 3.1.1.

Preliminary analysis.

One participant's second session's data was lost due to system malfunction. We analyze data for task 2 with 7 participants. Data analysis on task 1, the vertical steering-targeting task did not lead to any significant results. They are, hence, not reported. In what follows, only user performance for task 2 is reported

8.3.1. Experiment 1

Using a one-tailed test, no significant differences were found between user performances of right hand ($RH_{control}$, $Mdn = 36.00$, $IQR = 43.5$) and left hand following training with right hand and pseudo-haptic feedback (LH_{rhf}, $Mdn = 9$, $IQR = 45$), $n = 199$, $p = 0.74$ (Figure 8.3). As expected, task times were positively skewed, but similar, $RH_{control}$, $M = 29.89$s, $SD = 7.55$, and LH_{rhf}, $M = 31.71$s, $SD = 7.73$.

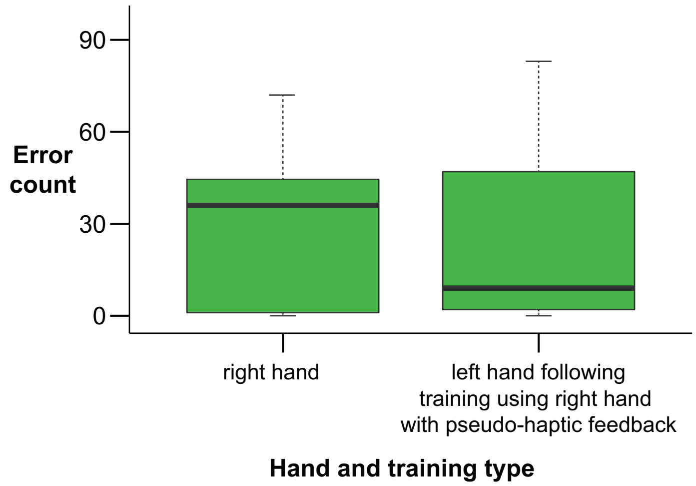

Figure 8.3. No significant differences were found between user performances of right hand and left hand following training with right hand and pseudo-haptic feedback.

8.3.2. Experiment 2

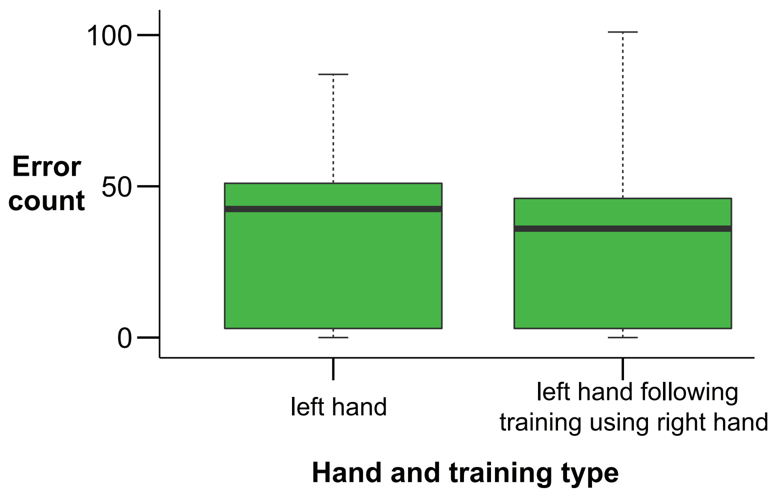

Figure 8.4. No significant differences were found between user performances of left hand and left hand following training with right hand.

Using a one-tailed test, no significant differences were found between user performances of left hand ($LH_{control}$, *Mdn* = 42.5, *IQR* = 48) and left hand following training with right hand (LH_{rh}, *Mdn* = 36, *IQR* = 43), *n* = 84, *p* = .21 (Figure 8.4). *H2* was not supported. Similar to experiment 1, task times were neither significantly different, $LH_{control}$, *M* = 35.47s, *SD* = 8.53, and LH_{rh}, *M* = 30.53s, *SD* = 9.63.

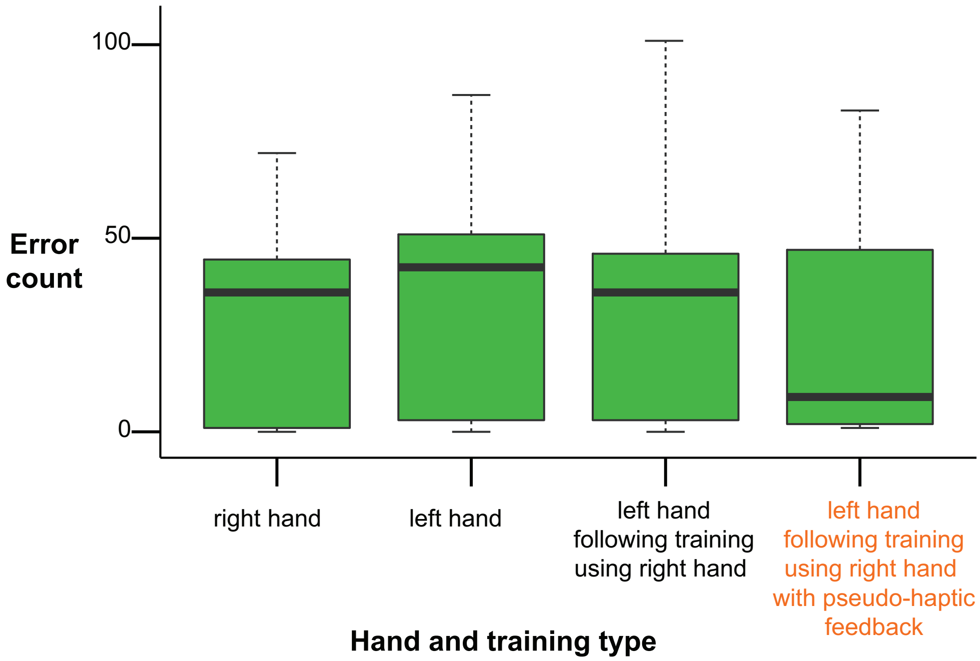

Figure 8.5. User performances of left hand following training with right hand and pseudo-haptic feedback was significantly more accurate than left hand without any prior training.

When comparing across experiments, we found that left hand following training with right hand and pseudo-haptic feedback (LH_{rhf}) was significantly more accurate than left hand without any training ($LH_{control}$) with a small-to-medium effect size, U = 9944.5, Z = 2.53, p = .006, r = 0.2 (Figure 8.5). H3 was supported.

Right hand ($RH_{control}$) was significantly more accurate than left hand ($LH_{control}$) with a small-to-medium effect size, U = 6565.5, Z = 2.86, p = .002, r = 0.2, (Figure 8.5). *H1* was not supported. No other performance differences were significant.

8.4. Discussion

This Chapter investigated handedness and transfer of training in a touchless circular steering-targeting task. Data failed to support the hypothesis that left-hand accuracy is superior to right hand (H1). Instead, the user performance of right hand ($RH_{control}$) was found significantly more accurate than left hand ($LH_{control}$) with a small-to-medium effect size. This result may be limited because of two reasons. First, the circular steering-targeting task with high-density arrangement may have required greater demands of feedback control, thus drawing on the strengths of users' dominant hand—and neutralizing the property of touchless depending on preprogrammed motor plans. This points to the fact that nondominant hand performance does not simply depend on the input modality, but a combination of the task-at-hand and the input modality. Second, in the $RH_{control}$ condition, users performed steering-targeting without any pseudo-haptic feedback. However, because we were not studying bilateral transfer of learning from nondominant to dominant hand, in the $LH_{control}$ condition, users performed steering-targeting both with and without pseudo-haptic feedback (randomly balanced). Thus, user performance of the nondominant hand may have been aggravated for trials that needed additional feedback control. Future experiments need to consider tasks with much less requirements of feedback control to evaluate if there is a nondominant hand advantage for simpler tasks.

We found a significant transfer of learning from dominant to nondominant hand. H3, but not H2, was supported. Transfer of learning was significant when users performed tasks with their right hand and used pseudo-haptic feedback in some of the trials. When users were not exposed to the additional feedback in the right-hand condition, the transfer of learning was not significant; left hand's performance was not significantly better than without any prior learning ($LH_{control}$). The additional feedback in right-hand condition must have augmented the motor skill learning in touchless circular steering-targeting, which is later transferred to the nondominant hand.

Although more systematic explorations are required to understand the role of nondominant hand in touchless, we were able to show a significant transfer of training from dominant to nondominant hand in touchless. Touchless interaction techniques can be designed to support this type of inter-limb transition for bimanual methods, thus supporting a novice to expert changeover.

It is important to note here that hemispheric asymmetry also affected user performance in a prior experiment in this dissertation—in Chapter 5. In Chapter 5, we had found that accuracy of mid-air directional strokes (within dominant hand) significantly increased as movements became longer (see Figure 5.6). This maybe because short movements using preprogrammed motor plans (Todor & Doane, 1978) was inhibited by the use of dominant hand compared with the feedback control required in longer movements.

Two research implications follow from these findings. First, touchless modality facilities tasks requiring pre-programmed motor plans over feedback control. Second, if tasks require greater feedback control, training the dominant hand with additional feedback can significantly improve the nondominant hand's performance due to bilateral transfer of learning.

As of now, we have studied either perceptual factors (visual feedback, image schema) or motor factors (pseudo-haptic feedback, transfer of learning in motor skill). The final set of experiments, in the next Chapter, investigates the confluence of perceptual and motor factors in touchless. How visual theories of perception may influence the motor action in touchless methods? We look for this answer in Chapter 9.

Chapter 9. Gestalt in touchless

This Chapter presents the final set of experiments of this dissertation: We investigate how Gestalt principles affecting visual perception influences motor control in touchless. While deconstructing intuitiveness in touchless—in chapters 2 and 5—we discussed wherein the mismatch lies between the physical and touchless world, in spite of the immediate resemblance of the gestural input. The mismatch lies in the availability of all physical abilities we use in a 3D world in touchless interactions, but to act on, a 2D user interface (UI) without any haptic feedback. Because of the lack of haptic feedback, touchless interaction exclusively depends on visual perception and proprioception. Thus, in this last study, we draw on the Gestalt principles of visual perception (particularly principles of similarity and continuity, Koffka, 1922) and motor control (Klapp & Jagacinski, 2011) to explore touchless interaction mechanics.

9.1. Gestalt psychology

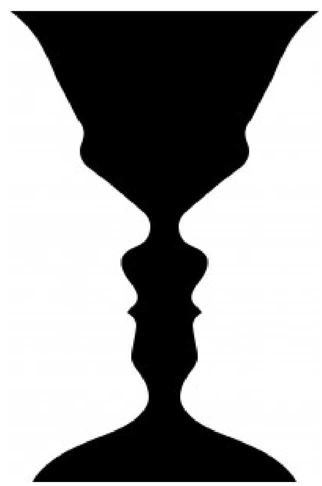

Figure 9.1. Rubin's face-vase is an example of visual illusion illustrating Gestalt principles of figure-ground organization (Rubin, 1915)

In spite of its criticisms—for over a century—Gestalt thinking has continued to influence the discoveries of psychological principles that explained visual perception (Wertheimer & Riezler, 1944). With Max Wertheimer's historical 1912 paper on phi motion, Gestalt theory emerged as an explanation of perception in terms of structured wholes or Gestalten, rather than an assimilation of more primitive percepts (Wertheimer, 1912). Decades later, true Gestalt phenomena again became relevant in visual perception with the importance of hierarchical structure in perceptual representations (Palmer, 1977). Ecological (Gibson, 1971) and computational (Marr, 1982) approaches to visual perception also acknowledged the influence of Gestalt thinking (Koffka, 1922;

Wertheimer & Riezler, 1944). Overall, Gestalt psychology—a popular proponent of holism—has played an illustrious role in providing theoretical foundations toward visual perception (Wagemans et al., 2012a; Wagemans et al., 2012b), and very recently motor action (Klapp & Jagacinski, 2011).

Gestalt psychologists have argued that perceptual experiences and motor actions are inherently holistic, rather than a composite of unrelated structural units. The most symbolic image of Gestalt is arguably the Rubin's vase (Figure 9.1, Rubin, 1915)—illustrating the figure-ground principle (Wagemans et al., 2012b): When two adjoining regions share a border and create a mosaic percept, the occluding region is perceived as the figure with the adjacent region not imparting a shape. This figure is said "to own the borderline". The border-ownership is switched when figure-ground reversals occur (e.g., when observers perceive two faces in Figure 9.1 instead of a vase). To find out why such a switching occurs, readers are directed to Wagemans et al., 2012b (section 3), and to find out the factors determining what is perceived as a figure to Wagemans et al., 2012a (section 5).

The centennial review on Gestalt research showed how different methodological shortcomings of this research program have somewhat been addressed (Wagemans et al., 2012a; Wagemans et al., 2012b). Specifically, the Gestalt principles of perceptual grouping in vision, such as proximity, similarity, or continuity, have been quantified (Wagemans et al., 2012a). Another recent review analyzed reaction-time results from previous studies and argued that four fundamental Gestalt principles in perception also apply to the control of motor action—holism, constancy, mutual exclusivity, and grouping in apparent motion (Klapp & Jagacinski, 2011). For example, certain motor actions, such as articulating a syllable during speech or making quick taps indicate the presence of motor Gestalts (chunks). However, neither perceptual nor motor Gestalt has been investigated in the context of touchless interactions.

The focus of this chapter is perceptual Gestalt. We argue that Gestalt principles can inform how visual perception influences touchless interactions, because visual perception plays a crucial role in terms of feedback, feedforward, or understanding ecological affordances (Gibson, 1979) in touchless systems. We also discuss later how one of our prior results on motor-intuitiveness (from Chapter 5) can be explained using motor Gestalt theories. In sum, this chapter's principal contribution is to introduce Gestalt thinking into touchless.

Specifically, I study the role of perceptual Gestalt in touchless target selection—with the *directional stroke* primitive (Chapter 5, or a *crossing* gesture; Accot & Zhai, 2002; Apitz, Guimbretière, & Zhai, 2010). Prior works have extensively investigated crossing in pen-based interfaces and suggested it as a promising interaction primitive for three-dimensional environments (Apitz et al., 2010). In what follows, we propose two experiments, their findings, and the significance of the results.

In Chapter 2, while setting up the background for this dissertation, we had emphasized the concept of embodiment in touchless, and argued that ecological affordances (Gibson, 1971) would be a suitable lens to explore touchless interface affordances—and more broadly touchless interaction mechanics. Gibson's work on affordances, like Marr (1982), is an approach to visual perception alternative to the more standard cognitive psychology and information-processing approaches. In addition, both these approaches show an explicit influence of Gestalt thinking (Wagemans et al., 2012a).

The oldest, most cited, and most studied aspect of Gestalt thinking in visual perception is perceptual grouping in simple 2D displays (Figure 9.2). Historically, Wertheimer (1923) proposed the first problem in perceptual grouping by exploring factors that determine the perceptual grouping of discrete elements (Wagemans et al., 2012a). Perceptual grouping is a kind of perceptual organization, which is a broader field, often studied by Gestalt psychologists. Another kind of perceptual organization is figure-ground organization. Their difference is important to note. Grouping establishes the qualitative elements of perception, such as similarity or continuity, while figure-ground determines how these elements are interpreted in terms of shape, relative location, or frame of reference in a 3D world (Wagemans et al., 2012a). Since we study 2D touchless interfaces, our focus is perceptual grouping. Two particular grouping principles are studied: by similarity of shape (Figure 9.2-E) and continuity (Figure 9.2-I).

Similarity of shape. With all other conditions being equal, the most similar visual elements in shape tend to be grouped together (Wagemans et al., 2012a).

Continuity. With all other conditions being equal, elements tend to be grouped together when they are aligned with each other (Wagemans et al., 2012a).

We chose the above two Gestalt principles based on prior research and our preliminary investigation of perceptual Gestalt in expert users (see Appendix B).

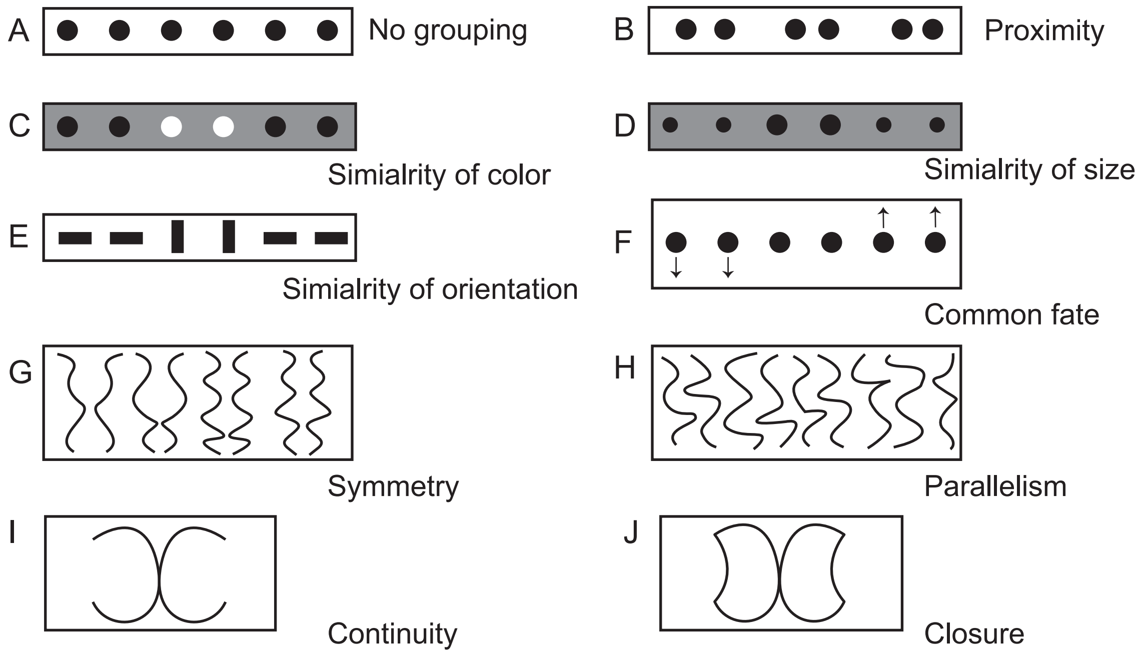

Figure 9.2. Some Gestalt principles of perceptual grouping (adapted from Wagemans et al., 2012a): Equally spaced dots do not group together (A), but when some are placed closed together, they group together strongly in pairs (B). All else being equal, the most similar elements will tend to be grouped together (by color, C; size, D; and orientation, E). Other examples include common fate (elements moving in the same direction, F), symmetry (G) and parallelism (H) of curves, continuity of lines (I), and closure (all else being equal, elements forming a closed figure will tend to form a group).

9.2. Research questions and hypothesis

Beyond visual perception, Gestalt principles of grouping was recently studied in motor action (Klapp & Jagacinski, 2011) and tactile perception (Gallace & Spence, 2011). Moreover, Gestalt principles continue to inform the design of traditional graphical user interfaces (interaction design book). Building upon prior work, this dissertation lays the foundations for designing Gestalt-informed touchless user interfaces. Our overarching research question is:

How perceptual Gestalt affects a crossing-based touchless user interface?

Within this dissertation crossing-based interface was first studied in Chapter 5 to understand the motor-intuitiveness of directional strokes, and then in Chapter 6 to design a touchless command-selection technique.

Grouping by similarity

The Gestalt principle of similarity states that, “all else being equal, the most similar elements (in color, size, and orientation) tend to be grouped together” (p. 9, Wagemans et al., 2012a). We hypothesize the following:

H1: User interface (UI) components representing similarity will *decrease* the efficiency of touchless target selection by crossing.

H2: UI components representing similarity will *decrease* the accuracy of touchless target selection by crossing.

The rationale of this hypothesis is that the perceptual similarity between different UI components will tend to group strongly into a perceptual whole. Such a perceived grouping would inhibit the action of crossing if one of those UI components represents action while the other a signifier of the action (e.g., a part of a widget and a cursor, Figure 9.3). Our hypothesis is informed by our preliminary findings where expert users were faster when crossing-to-select a rectangular menu option with a circular cursor than a circular menu option with a circular cursor (see Appendix B).

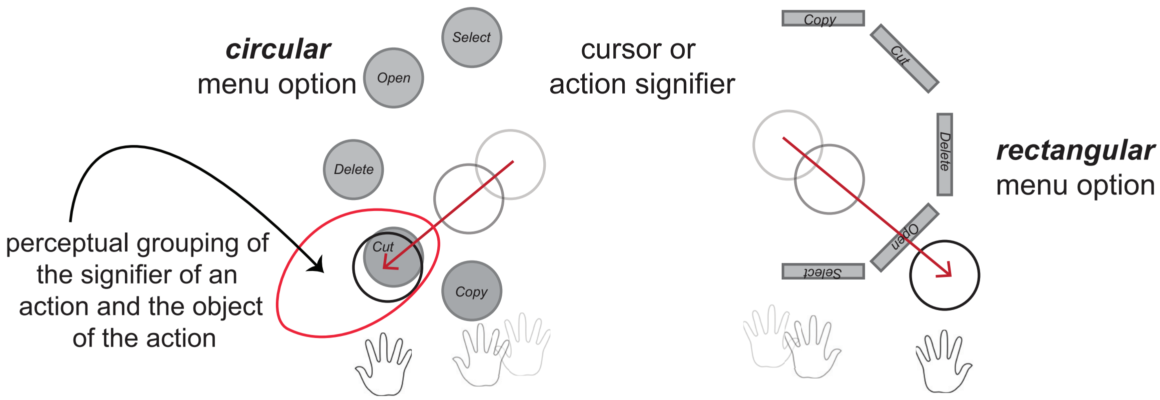

Figure 9.3. Strong tendency of a perceptual grouping would inhibit the action of crossing if one of those UI components represents action while the other a signifier of the action—due to Gestalt principle of similarity.

Grouping by continuity

The Gestalt principle of continuity or good continuation states that elements tend to be grouped together as a single uninterrupted object when they follow an established direction (Wagemans et al., 2012a). We hypothesize the following:

H3: UI components representing structural continuity will *increase* the efficiency of touchless target selection by crossing.

H4: UI components representing structural continuity will *decrease* the accuracy of touchless target selection by crossing.

The rationale of this hypothesis is that the continuity of UI components (e.g., a menu with multiple options) increases the effective target width, because users tend to group the UI components into a perceptual whole (Figure 9.4, left). But in the absence of good continuation, the target width is decreased (Figure 9.4, center) and different parts of the UI component act as distractors to the intended target (Figure 9.4, right). With the increase in effective width, users will be faster, but more prone to make angular errors. Our hypotheses are also informed by prior studies that found that efficiency in crossing-based interfaces is inversely related to the target width (Apitz et al., 2010).

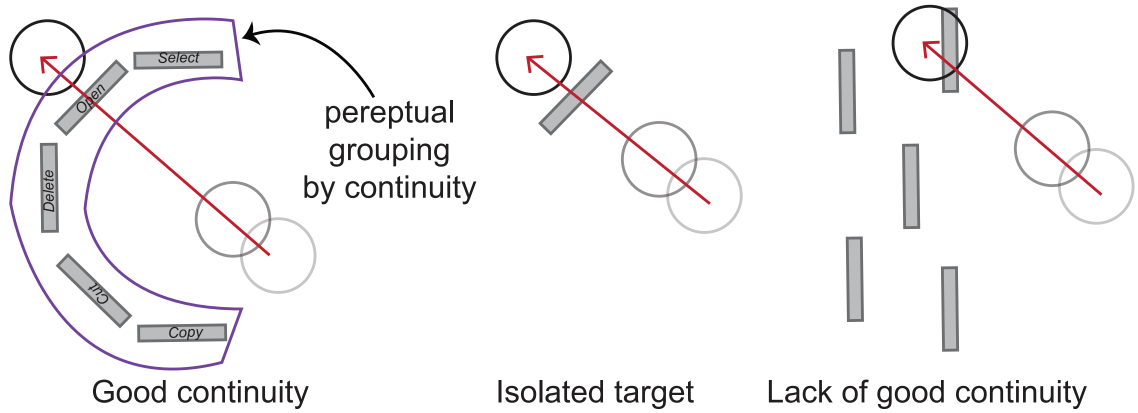

Figure 9.4. Good continuity of UI components (e.g., a menu with multiple options) increases the effective target width, because users tend to group the UI components into a perceptual whole (left). However, in the absence of good continuation, the target width is decreased (center) or the different parts of the UI component act as distractors to the intended target (right).

Method

Participant. We recruited 18 right-handed participants (M_{age} = 25.61, SE_{age} = 1.98, 8 females). Fifteen of them were familiar with Kinect, Wii, or Leap Motion. This study was approved by Indiana University IRB (1601477955) and participants were compensated $15 for their time and effort.

Apparatus. We used a 1.34 m wide and 0.79 m high LG TV with a resolution of 1920 x 1080 pixels and driven by a single computer (Figure 9.5). For motion tracking, we used off-the-shelf hardware—a Kinect™ for Windows. The experiments were written in C# running on Windows 7, and were implemented with OpenNI 1.4 SDK and PrimeSense's NITE 1.5. During the study, participants sat in a 56 cm high chair, situated 1.5 m away from the large display (1.54 m from the sensor) and took about an hour to complete the study (Figure 9.5). The sensor was 83 cm from the floor and aligned to the user's body midline horizontally. The armrest of the chair was 73 cm high. The motion-tracking sensor had a horizontal field of view of 57 degrees and a vertical field of view of 43 degrees. Participants' movements were mapped from real space to display space as 1: 3.7 (when a participant moved 1 cm in real space the cursor moved 3.7 cm in the display space, baseline C/D ratio). Trials were video recorded.

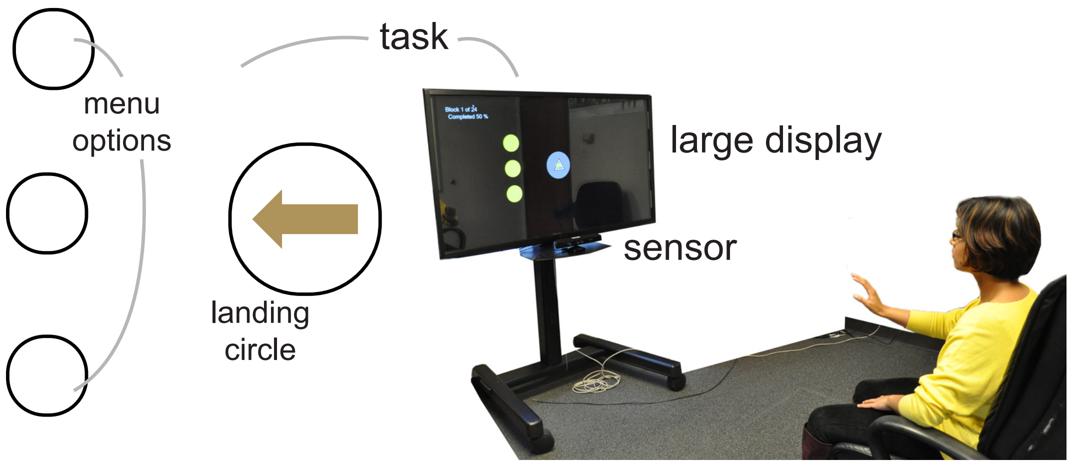

Figure 9.5. (Right) In our experiment, participants used touchless gestures to interact with a large display, while sitting away from it. (Left) The experimental task began with a landing circle appearing on the display. As participants reached the landing circle, the target appeared and participants completed the task by crossing-to-select the target.

Prior to each task, all participants practiced three blocks trials. Participants rested at least 5 s after every 3 trials and 5 minutes after completing all trials in a task. We logged task completion times, the number of errors, and the trajectory paths for each trial.

Participants hovered over a 'Start' circle to begin a block. Each trial began with a *landing circle* appearing on the display, which participants landed on to begin the trial. The landing circle was horizontally aligned with the participants' body midline. As soon as participants reached the landing circle, two things would appear: an arrow representing one out of four directions (0, 45, 135, and 180) and a *target* (Figures 9.7 and 9.9). Participants' hand movements in the 3D space were measured as their orthographic projections on the 2D display.

We recorded performance time, error rate, angular error, and trajectory paths. Time was measured from when participants left the landing circle to when they moved past the target. We measured the angle of crossing using the last point recorded inside the landing circle and the first point recorded after crossing the target (hence the width of the target did not influence the calculation of angular error). Angular error was calculated as the absolute difference between this crossing angle and the required angle for the trial. For a trial to be considered successful, participants were required to move past the target with an angular error less than ± 22.5°. If users land beyond the target without crossing or make an angular error greater than ± 22.5°, it is considered as an error (a target miss) In the case of an error, the trial was repeated until participants successfully completed it. We operationalize efficiency as time to complete a trial and accuracy the angular error.

9.3. Experiments on Gestalt similarity

Tasks and procedure

This experimental task emulated different cursors and widgets currently used in existing touchless systems (Figure 9.6). A linear menu (Callahan et al., 1988) was presented, and the shape of the touchless cursor and the shape of the menu options were systematically manipulated (three structures: circle—circle, triangle—triangle, and circle—triangle, Figure 9.7).

Figure 9.6. An example of a linear menu in a current touchless application (Xbox Kinect game, Dance Central 2)

Because past studies showed that certain angles between the target centerline and the horizontal line affect user performance of crossing-based interfaces (Accot & Zhai, 2002; chapters 5 and 6), this study randomized trials at the following four angles with similar levels of difficulty: 0°, 45°, 135°, and 180° (Figure 9.7). The total number of trials for this experiment was: 18 (participants) x 4 (angles) x 3 (structures) x 8 (blocks or repetitions) = 1728.

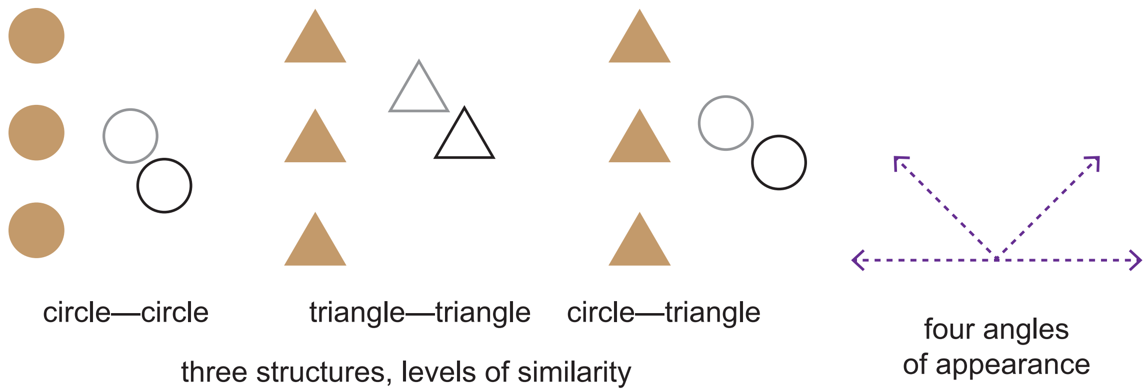

Figure 9.7. To test Gestalt similarity in touchless, a linear menu was presented at four different angles, and the shape of the touchless cursor and the shape of the menu options were systematically manipulated (circle—circle, triangle—triangle, and circle—triangle).

Results

For all 18 participants, for both the experiments, we analyzed performance time (for efficiency) and angular error (for accuracy). Trajectory paths are not reported here; for an analysis of paths during crossing-to-select targets, see Chapter 5. Because of the simplicity of the experimental task, overall workload was not measured, and we chose the continuous dependent variable angular error over the discrete error count to operationalize task accuracy. Across the experiments, participant reported their levels of fatigue as very low (on a 10-point scale, *Mdn* = 3, *IQR* = 2.75).

Recorded data were positively skewed; thus, replications of unique experimental conditions were represented by their median. Our analysis used GLMM with standard repeated measures REML technique. Participants were handled as a random factor. We report F-statistic using type III ANOVA with Satterthwaite approximation, and pairwise comparisons (using pooled variance) with Holm-Bonferroni correction. Effect sizes are reported using Cohen's *d*, and interpreted as: 0.2 or greater as small, 0.5 as medium, and 0.8 large (Cohen, 1992).

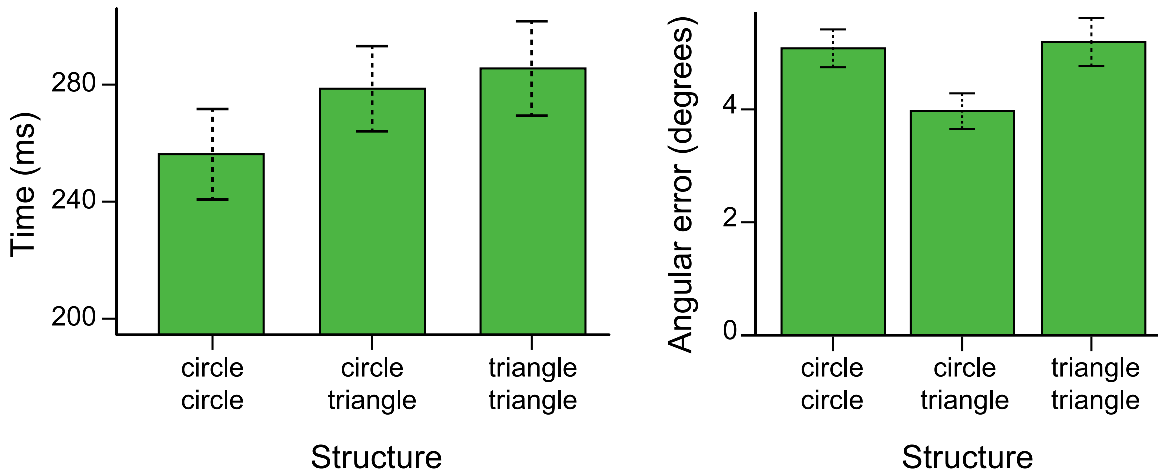

Figure 9.8. Similarity of shape in UI components did not significantly affect performance times, but it affected accuracy. Participants made significantly smaller angular error in the dissimilar condition (circle—triangle).

Efficiency. We found significant effects of structure $F(2, 186.97) = 8.68$, $p < .001$, and angle, $F(3, 186.97) = 9.03$, $p < .001$, but no significant interaction effect (Figure 9.8, left). However, participants took significantly less time in the circle—circle condition (M = 256 ms, SD = 131) than triangle—triangle (M = 286, SD = 137), $p < .001$, d = .22, and circle—triangle (M = 279, SD = 123), p =.002, r = 0.18. Similarity did not significantly decrease the efficiency of touchless target selection by crossing. *H1* was not supported.

Accuracy. We found significant effects of structure $F(2, 187) = 8.31$, $p < .001$, angle, $F(3, 187) = 65.69$, $p < .001$, and structure x angle interaction, $F(6, 187) = 5.08$, $p < .001$ (Figure 9.8, right). Participants made significantly smaller angular error in the circle—triangle condition (M = 3.97, SD = 2.68) than triangle—triangle (M = 5.19, SD = 3.62), $p < .001$, d = .38, and circle—circle (M = 5.08, SD = 2.85), $p < .001$, d = .40. Similarity significantly decrease the accuracy of touchless target selection by crossing. *H2* was supported.

The effects of the direction of movement (angle) on user performance was similar to findings previously reported in Chapter 5 and 6.

9.4. Experiments on Gestalt continuity

Tasks and procedure

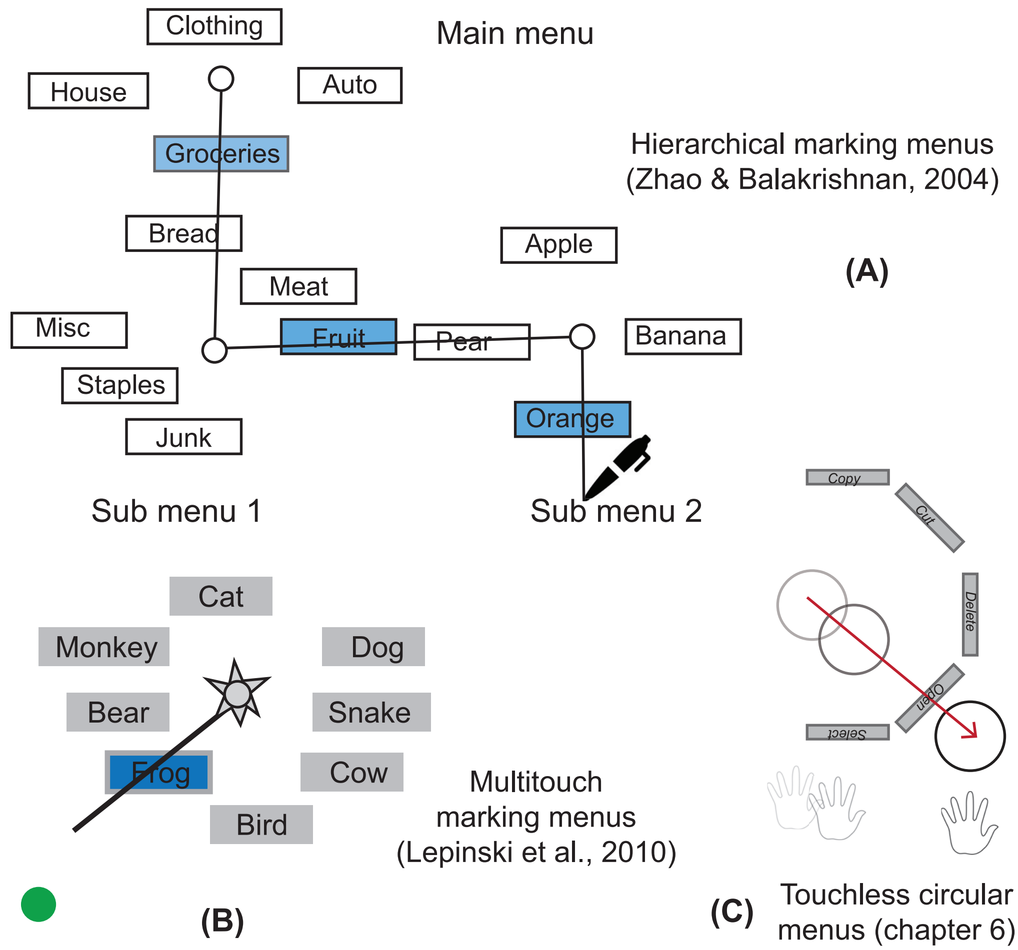

Figure 9.9. Examples of menu structures with no good continuation (A, Zhao, & Balakrishnan, 2004 © 2004 Association for Computing Machinery, Inc. Reprinted by permission; B, Lepinski et al., 2010 © 2010 Association for Computing Machinery, Inc. Reprinted by permission) and good continuation that can be organized as a perceptual whole (e.g., a semi-circle) (C).

This experimental task was inspired by two existing menu structures that employ the crossing interaction primitive—pen and touch-based marking menus (Zhao, & Balakrishnan, 2004; Lepinski et al., 2010; Figure 9.9) and touchless circular menus (Chapter 6, Figure 9.9). However, it is important to note that marking menus employ directional strokes that are delimiter-independent and do not require explicit crossing for target selection in expert mode (Kurtenbach & Buxton, 1994). Compared with pen or mouse (where pen up or coming off the screen means delimiting an action), touchless do not provide an easy way to indicate the end of a selection (action delimiter). Thus, touchless approximation of marking menus has used explicit delimiters, such as a closed fist (Bailly et al., 2011) or crossing (Ren & O'Neill, 2012).

The continuity experiment was designed same as the similarity experiment, except the task included targets of different shape and orientation. Three levels of the independent variable, continuity was tested: good continuity, no continuity, and distractors (Figure 9.4). Measures were same as in the previous experiment. Trials were randomized within subjects. The total number of trials for this study was: 18 (participants) x 4 (angles) x 3 (structures—good continuity, no continuity, no continuity with distractors) x 8 (blocks or repetitions) = 1728.

Results

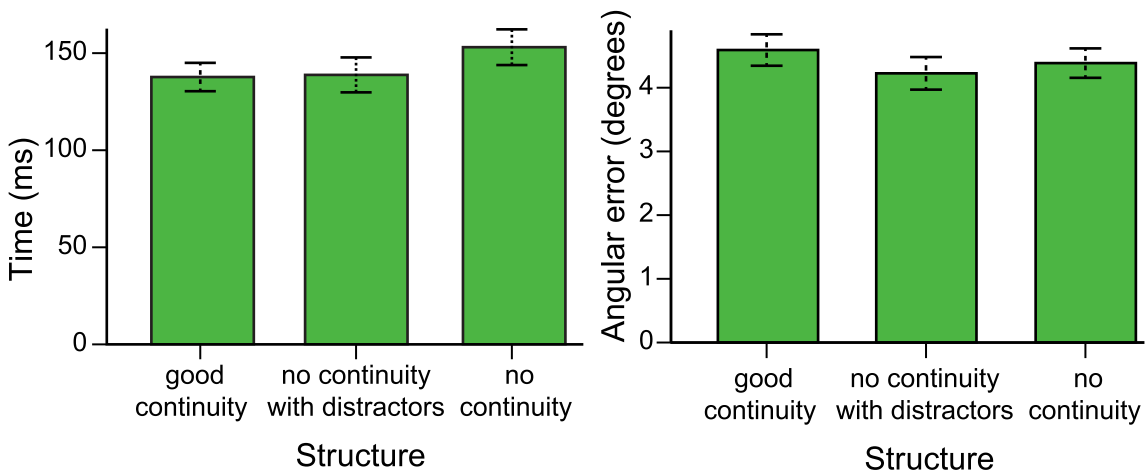

Figure 9.10. Good continuity in UI components significantly affected performance times, but not accuracy. Participants were significantly faster with good continuity than no continuity.

Efficiency. We found significant effects of structure $F(2, 187.01) = 8.87$, $p < .001$, and angle, $F(3, 187.01) = 5.35$, $p = .001$, but no significant interaction effect (Figure 9.10, left). Participants took significantly less time in the good continuity condition (M = 138 ms, SD = 62) than no continuity (M = 153, SD = 78), $p < .001$, $d = .21$. Participants were also significantly faster with the distractor condition (M = 139, SD = 76) than the no continuity condition, $p = .001$, $r = .19$. Continuity significantly increases the efficiency of touchless target selection by crossing. *H3* was supported.

Accuracy. We only found significant effects of angle, $F(3, 187) = 15.46$, $p < .001$ (Figure 9.10, right). Continuity did not significantly decrease the accuracy of touchless target selection by crossing. *H4* was not supported.

9.5. Discussion

Our results showed that Gestalt similarity of shape and Gestalt continuity in user interface (UI) components significantly affects user performance. However, the influence of similarity and continuity differed in terms of performance times and accuracy. Similarity made users less accurate, but did not affect performance times (Figure 9.8). H2 was supported. H1 was not supported. Good continuity made users faster, but did not affect task accuracy (Figure 9.10). H3 was supported. H4 was not supported.

Gestalt principles of visual perception are not mutually exclusive. They often act together, sometimes trumping one effect for another (e.g., proximity for similarity or similarity for continuity, Wagemans et al., 2012a). In this chapter, we studied the Gestalt effects of similarity and continuity on touchless motor action. Before discussing the results in detail and generalizing the design implications, it is important to note some limitations of the study.

Limitations. For motion tracking, we used an off-the-shelf tracking sensor Kinect™. It tracks users at 30 frames per second and suffers from occasional jitter that makes it suitable as a gaming console but not at par with the sub-millimeter accurate, marker-based tracking systems, such as VICON or OptiTrack. This tracking noise may have affected our findings.

Research and design implications

Gestalt similarity. In the similarity experiment, we found no significant improvement in efficiency for the dissimilar condition. We hypothesized that perceptual similarity between the signifier of an action (a cursor) and the object of an action would inhibit a crossing-to-select action because the UI components will tend to group strongly into a perceptual whole. Our findings, here, was different than our findings with expert users (n = 3), who were faster when crossing-to-select a rectangular menu option with a circular cursor than a circular menu option with a circular cursor (see Appendix B). A direct comparison between these two findings is infeasible because of the difference in participant types, task, and task parameters. Furthermore, our objective in this study was not to replicate prior findings, but conduct a more systematic, internally valid, exploration. However, it is interesting that between the two similar conditions, circle—circle condition was significantly faster than both circle—triangle (dissimilar) and triangle—triangle (similar) condition. This may be explained by *legacy bias* (Morris et al., 2014)—prior, extensive familiarity in a similar interaction context (circular cursors in current touchless applications), but needs further research. As expected, accuracy and task times were not significantly correlated. Our findings may also be explained by the typicality of the triangular shape over a vertex-less circle or less-angular straight line (like used in touchless circular menus, Chapter 6). The deviation of this study's findings from our preliminary results is not a limitation of this work; rather it opens up new research questions about touchless Gestalt: how familiarity and shape parameters mediate the effects of Gestalt principles on touchless. Such mediating effects are not new in HCI research. Pertinently, research shows the effect of familiarity on the use of image schema and metaphors in interaction (Blackler, Popovic, & Mahar, 2010) and shape on visual search (Smith & Thomas, 1964; Wolfe, 1998).

Gestalt continuity. In the continuity experiment, we found that good continuity made users faster than no continuity, which provides support to the premise that continuity created a perceptual grouping, thereby increasing the effective target width. However, there was no significant effect of continuity on accuracy (angular error). That accuracy was not significantly affected by continuity suggests that an increase in effective target width did not decrease users' targeting precision. This is an interesting finding and merits further exploration and explanations.

Furthermore, UI with no continuity and distractors also made users faster than the no continuity condition. This may be explained by the symmetrical structure used in the experimental task. If the increase in efficiency is caused by the increase in effective target width, the symmetry in the menu structure (Figure 9.4, right) may have caused it to appear as a perceptual whole (principle of Gestalt symmetry, Figure 9.2-G) instead of distractors. Whether symmetry prevailed over continuity in this occasion needs further exploration. In sum, in this very first study on touchless Gestalt, we showed some effects of Gestalt principles of perceptual grouping on touchless interactions.

9.6. Motor Gestalt in touchless

Apart from perceptual Gestalt, this chapter also reviewed recent findings of Gestalt principles affecting motor action (Klapp & Jagacinski, 2011). Although we did not empirically study motor Gestalt in touchless, we found that findings from a prior experiment (Chapter 4) could be explained using the lens of motor Gestalt.

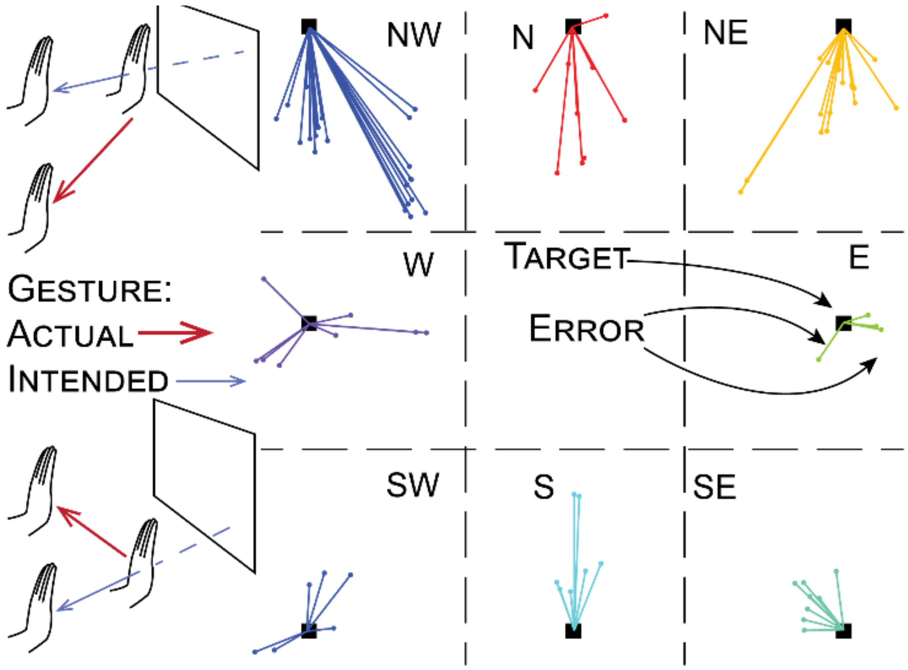

Figure 9.11. A holistic motor Gestalt exemplifying the Law of Prägnanz affected gesture intuitiveness: When intending to make mid-air movements perpendicular to a vertical display, such as a pull gesture, users repeatedly made oblique motions toward the center of their torso—to optimally reach static equilibrium, thus minimizing their body's energy expenditure.

Prägnanz affects intuitiveness of Gesture primitives

In a study on visual feedback, we explored push-to-select and pull-to-deselect gestures (Chapter 4). In a drag-and-drop task on a large display, we observed users often trying to select targets by following the shortest, oblique path, instead of a set of orthogonal paths. Similarly, when pulling to deselect, instead of a decoupled set of orthogonal movements (parallel to the display for translation and perpendicular for action), users intuitively made oblique motions toward the center of their torso (Figure 9.11). Such tendencies exemplify motor planning routines that seek to minimize our metabolic energy costs (Alexander, 1997). Overall, the holistic nature of this oblique motion can be explained by the concept of physical Gestalten or motor Gestalt (Klapp & Jagacinski, 2011), which is derived from a more general Law of Prägnanz: all physical systems, when left alone, tend to achieve a state of maximum equilibrium with minimum energy expenditure (Wagemans et al., 2012a).

Chapter 10. General discussion

10.1. Discussion

This dissertation studied the interaction mechanics of device-less, touchless target selection—on large, 2D, vertical, distant displays. Currently, predominant approaches to touchless interaction design are either exploring intuitive touchless gestures—using gesture elicitation studies—or introducing interaction techniques through expert design. Few researchers have also explored types of feedback or motor skills involved in touchless interactions. In this dissertation, I shift away from existing approaches toward understanding touchless as a sensorimotor phenomenon. I look at the device-less property of touchless from an embodied interaction perspective—what does the lack of tool entail? What is intuitive in touchless? How do the extreme reliance on proprioception and visual perception affect user performance? What theories and frameworks in visual perception and motor training and control can inform any future design of touchless techniques? This dissertation evolves from a theoretical stance (Chapter 2) to a pragmatic operationalization of intuitiveness (Chapter 5) to investigation of both interaction mechanics (e.g., feedback in chapters 4 and 7, motor control in Chapter 8, or interface affordances in Chapter 9) and interaction technique proposals (Chapter 6). As I claimed in the introduction, this work's theoretical investigation is a crucial stepping stone to generating fundamental knowledge about the potential and limitations of touchless as an interaction modality. Knowledge resulting from this work can drive the design of next-generation touchless systems based on fundamental interaction principles—instead of a reactive adaptation to the sensing technologies. In what follows, I highlight the results from the different chapters, discuss common patterns emerging from them, and indicate areas that merit further research.

In sum, the major findings of this work can be classified into the two classical aspects of any sensorimotor phenomenon: *visual perception* and *motor action*. This discussion will not attempt to generalize these findings into design implications. Some of the chapters already enumerated possible design directions. Rather, in this chapter, I will discuss research implications. Furthermore, it is important to note that the interaction techniques proposed in this work (Chapter 6) primarily contributed to the study of interaction mechanics—they served more as an apparatus than as a final proposal for future designers. So what is the takeaway for future designers of touchless systems from this dissertation? I address this question in Table 10.1. I also revisit each area of exploration and explain the emerging patterns.

Table 10.1. Dissertation findings on interaction mechanics for touchless target selection.

Primary category	Theme	Finding	Chapters
Visual perception	Gestalt principles of visual perception	Similarity of UI components made users less accurate but did not affect performance times.	9
		Good continuity made users faster but did not affect task accuracy.	
		Symmetry trumped the lack of good continuity and increased accuracy.	
Visual perception	semantic visual feedback	Persistent visual feedback increased users' efficiency to return to the display range when gestures went off accidentally.	4, 6
		Visual feedback assisted users, when they confused pseudo-haptic resistance as motion-sensing malfunction.	
Motor action	Hemispheric asymmetry	Right-hand accuracy was superior to left hand.	8
		Angular accuracy of mid-air directional strokes (within dominant hand) significantly increased as strokes became longer.	
Motor action	Bilateral transfer of motor skill	Transfer of learning was significant when users performed tasks with their right hand and used pseudo-haptic feedback in some of the trials.	8
Motor action	motor-intuitive interaction	We classified intuitiveness in touchless according to the continuum of knowledge in intuitive interaction. Motor-intuitive interaction draws on sensorimotor level of knowledge.	5

Motor action	2D directional stroke interaction primitive	Average angular accuracy of 2D mid-air directional strokes was 12 degrees compared with the infeasibility of 3D angular strokes.	5, 6
		Although 2D directional strokes are motor-intuitive and draw on the sensorimotor level of knowledge, we showed that bio-mechanical factors affect the performance of the interaction primitive (e.g., bilateral inhibition).	
		Touchless circular menus, build upon the 2D directional stroke primitive, was faster in target selection than contextual linear menus using grab gestures.	
Motor action	pseudo-haptic feedback	Pseudo-haptic feedback increased users' accuracy in difficult steering-targeting tasks.	6, 7
		Due to user expectation of high fluidity in touchless, pseudo-haptic resistance was often perceived as sensor malfunctions.	

The core research areas that this dissertation set out to explore were input, feedback, and interface affordances in touchless target selection (see Figure 3.2., Chapter 3). However, the findings can be better categorized as informing visual perception and motor action.

Visual perception

I explored visual perception primarily through visual feedback (Chapter 4) and Gestalt theories (Chapter 9)—both in 2D. Understanding visual perception was also significant to represent image schemas while designing motor-intuitive touchless interaction primitives (directional strokes, Chapter 5) and augmenting pseudo-haptic feedback in touchless steering-targeting (interface topographies, section 6.2). Notably, all these findings cannot be generalized to three-dimensional systems, such as gestural interaction in augmented or mixed reality or with head-mounted displays.

We found continuous visual feedback significantly improving user performance over object-oriented partial feedback (e.g., continually showing user's position as a touchless cursor on the display rather than highlighting a folder when the user is over it); 50% transparent cursors were preferred over opaque ones; semantic feedback and feedforwarding assisted users to reorient themselves when their gestures were out of the display range, but feedback echoing users' trajectory degraded performance during drag-and-drop tasks. Results suggest that, although touchless heavily relies on visual feedback, suitable feedback is not about providing maximum information, but sufficient information to allow users build a mental model of the interaction at hand.

While designing touchless interaction primitives (e.g., directional strokes), visual cues played an important role in representing image schemas and ensuring that users draw from the sensorimotor level of knowledge. We used an allocentric frame of reference instead of an egocentric one commonly found in immersive, avatar-based games. The space schemas in mid-air directional strokes were represented as straight lines in eight compass directions (e.g., north, south, or north-east). Although, user performance of crossing in different directions was affected by bio-mechanical constraints (see Chapter 5), average accuracy was around 12 degrees—compared with multiple earlier findings, where users' 3D strokes were inaccurate, to the extent of being infeasible.

Visual perception also became important when designing pseudo-haptic feedback for touchless steering-targeting. Manipulating visual feedback (in terms of control-display gain of the cursor, isometric to the users' motor movements) is a crucial ingredient in pseudo-haptic feedback (see chapters 6 and 7). But we augmented the traditional method with additional visual feedback to address users' confusion between haptic resistance and technological failure. Pseudo-haptic feedback generates the illusion of lateral forces when perusing an interface, similar to force-feedback devices. In touchless, this kind of feedback can address the lack of any haptic guidance. However, because of no input device and extreme reliance on proprioception, pseudo-haptic feedback in touchless was often confused as malfunctioning motion tracking—in a sense that any resistance to move freely was perceived as a *constraint* than a guidance. The expectation of abundant fluidity in touchless was violated by traditional rendering of pseudo-haptic feedback. Thus we augmented pseudo-haptic feedback with semantic visual feedback: a visualization of how much tensor force is currently at play and how much more is required to free the resistance (see Figure 6.13). Our design heuristics illustrate how visual feedback can bridge the gap between user expectation of fluidity in touchless and user requirement of guidance for improvements in task accuracy.

Finally, visual perception was explored through Gestalt theories of similarity and continuity. Similarity between user interface (UI) components—the action signifier and the object of action—decreased user accuracy, but did not affect performance time. Good continuity in UI components made users faster, but did not affect accuracy. Results also suggested that symmetry in UI—another Gestalt principle in visual perception (see Figure 9.2)—may have trumped a lack of good continuity in increasing user efficiency.

Overall, our findings indicate the crucial role of visual perception in touchless interactions, informing future designs of feedback and feedforward routines and the design of UI components.

Motor action

Motor action in touchless target selection was primarily explored through the work on touchless input and nondominant hand performance. Across the dissertation research, we studied three kinds of target selection methods. Mostly, we studied target selection by crossing. Chapter 4 also reports target selection by the dynamic gesture *push*—making orthogonal movements toward the display, and Chapter 7 discusses target selection by steering along a vertical or circular path.

Chapter 6 detailed the bio-mechanical factors affecting touchless target selection by crossing. We found user performance excels in their dominant sub-hemisphere, degrading at the poles and at the nondominant sub-hemisphere due to bilateral inhibition. We also found, quite interestingly, that accuracy of directional strokes (within dominant hand) significantly increased as movements became longer (see Figure 5.6). We think that this finding can be attributed to Todor & Doane's (1978) theory of motor behavior and its relation to hemispheric asymmetry: short movements using preprogrammed motor plans was inhibited by the use of dominant hand compared with more feedback control required in longer movements.

In Chapter 5, we explained why making orthogonal movements toward the display—in a push or hover gesture—can create ambiguity between translation and intended action. When reaching for targets on a distant display, users tend to choose a motion trajectory based on minimum energy cost (similar to what we do in our everyday environment while reaching for a physical object)—not a set of orthogonal movements.

We think that this tendency can also be explained by motor Gestalt—human preference toward making holistic gestures or chunking motor actions.

Results suggested that the untrained nondominant hand does not perform better than the dominant hand in touchless steering-targeting. However, the experimental task we used may have limited the generalizability of the results: the circular steering task (with or without pseudo-haptic feedback) may have required greater demands of feedback control, thus neutralizing the advantage of preprogrammed motor plans in touchless. When users first performed the task with their dominant hand with pseudo-haptic feedback in some of the trials, their left hand performance was significantly better than no prior learning. These findings suggest that the additional feedback in dominant-hand condition must have augmented the motor skill learning in touchless circular steering-targeting, which was later transferred to the nondominant hand.

Through this dissertation, I uncovered fundamental knowledge about touchless interaction mechanics in target selection. Broadly speaking, we found support that (1) Gestalt principles of visual perception affect touchless performance; (2) hemispheric asymmetry plays a role in bilateral transfer of motor skill and touchless performance with the dominant hand improves when pre-programmed motor planning is less involved than feedback control; (3) semantic visual feedback is more advantageous than echo feedback; (4) motor-intuitive interaction primitives must draw on users' sensorimotor level of knowledge; (5) 2D directional strokes based on space schemas are motor-intuitive; and (6) pseudo-haptic feedback can improve accuracy of steering-targeting tasks requiring greater overall precision.

These findings are, however, limited to our controlled lab settings. Future work needs to build touchless systems informed by these results and test them in different contexts of use. In the last Chapter, I conclude with overall contributions of this dissertation, and discuss some future directions for touchless interaction research.

10.2. Reflections

In this section, I reflect on the evolution of this dissertation as a scientific inquiry and discuss my methodological stance in HCI research. I also comment on the concept of natural or intuitive, particularly how this dissertation research shaped my understanding of intuitiveness as a property of an interaction modality. By reflecting on this research holistically, this is an attempt to capture a personal trajectory of this work.

10.2.1. Is naturalness a legacy bias?

So far, the biggest propeller of touchless interactions has been the often-used adjective ‘natural’. Because computerized systems, such as the TV, oven, or the car dashboard involves different interaction modalities and techniques, a majority of which needs extensive learning, natural user interfaces imply a panacea—no prerequisite to familiarize.

The alternative claim that natural user interfaces are not natural served as the starting point for this dissertation. That touchless gestures do not equate to everyday gestures was strongly argued by Norman in 2010. Indeed, touchless interactions resemble day-to-day gestures, but that does not deem them to be natural or intuitive. Over the last few years, many researchers have urged not to make that precipitous conclusion. I agree with them. My work began with operationalizing intuitiveness using the concept of the continuum of knowledge in intuitive interaction (Chapter 5). Similar to the intuitive interaction continuum where the higher up the framework the specialized the knowledge, what kind of user abilities interactions draw on would determine its naturalness. For example, since image schema is a sensorimotor knowledge, interaction primitives based on image schema would be natural; rather more natural than using a random combination of fingers as commands, because finger combination to touchless-command mapping would be an expertise. This is a positivistic approach toward naturalness—determining what is natural vs. what is not. Alternatively, phenomenological approaches, particularly the situated approach would look at how touchless experience blend into communities of practice. If the blending entails a natural experience, then the interaction modality in use is natural or intuitive in that context. Nevertheless, these two approaches are complementary.

Although the dissertation always remained at the level of interaction mechanics, i.e., exploring sensorimotor relations, the later chapters focused on uncovering affordances in touchless interactions primarily. For instance, consider the work on perceptual Gestalt (chapter 9). In the phenomenology of perception, Merleau-Ponty deems perception as a dialectical relation between the body and the world and acknowledges that the concept of Gestalt is essential to understand the basis of perception. To that end, instead of attempting to represent intuitiveness, that chapter uncovered perceptual grouping effects in touchless interactions, thus contributing to understating the role of physicality in embodied sensemaking. Although the primary role of Gestalt grouping principles is helping individuals perceive the world, my findings indicated that perceptual grouping on visual displays could affect motor actions, particularly in interaction modalities relying heavily on visual perception. Studying touchless interaction with 2D displays provided the unique opportunity to discover how perceptual grouping effects extend beyond visual sensemaking to motor responses.

While I argued for operationalizing intuitiveness using the extent to which prior knowledge can be unconsciously applied, the legacy bias concept considers prior experience a hindrance in capitalizing the full potential of novel interaction modalities. Legacy bias has been observed in touchless gesture elicitation studies: Users draw heavily on the skills of traditional interactions, such as the mouse or keyboard. Researchers report an explicit desire of users to transfer knowledge from legacy modalities, thus limiting the elicitation of interaction possibilities. When asked to propose different interaction possibilities in terms of touchless gestures, individuals frequently thought within the frame of reference of the known interaction techniques, primarily to minimize the physical and mental exertion. For example, a mid-air gesture resembling the mouse click, but without the mouse and in a vertical stance, is often described by users as a natural way for target selection. It is important to note that this kind of 'natural' designation is biased by legacy interaction techniques, which is different from the operationalization of naturalness or intuitiveness provided in chapter 5.

In sum, my research inquiry focused on investigating intuitiveness in touchless interactions through uncovering affordances. I argue that this approach can inform future explorations of intuitiveness in touchless interactions or other novel interaction modalities, which are susceptive to legacy bias and whose technological possibilities are still evolving.

10.2.2. Why study interaction mechanics?

As I discussed in the introduction, a majority of the current touchless research focuses on input and interface design separately. For example, elicitation studies determine which mid-air gestures users would find natural as touchless commands and interface design studies user performance for interaction techniques. What remains little explored are the interaction mechanics or understanding touchless interaction as a sensorimotor phenomenon. Investigating interaction mechanics give researchers a unique opportunity to study an interaction modality in a bottom-up approach: to take a theoretical stance and understand what user abilities can inform interaction primitives or interface commands. Furthermore, it paves the way for a systematic exploration of how factors affecting those user abilities influence the interaction performance. For example, because touchless interactions rely heavily on visual perception, the Gestalt principles of perceptual grouping were studied in chapter 9.

Studying interaction mechanics can play a significant role in designing techniques for interaction modalities that are like real, but not a replica of the physical world. In line with the forecast of ubiquitous computing, computers are increasingly getting blended into our daily life. Moreover, interaction modalities that draw on people's perceptual senses, such as touch, voice, or even smell, and day-to-day actions, such as gesture, speech, or applying force, are becoming popular. As this trend continues, it is essential that we determine the similarities of these modalities with the real-word—and their differences. Looking at interactions as a sensorimotor phenomenon helps in studying input, interface, and feedback altogether, which can inform the design of future interfaces.

Similarities of interaction modalities to day-to-day actions can lead to new challenges, such as mode switching. For instance, when are people using gestures for communicating with a friend, or a pet, compared with gestures for interacting with their TV? Or when is someone forcing on a device for getting a grip, compared to force touch as a command? When is speech a voice command? Both research studies and commercial solutions continue to explore mode switching in such 'natural' user interaction modalities, like voice commands. The more interaction modalities resemble people's daily actions, the more important it is to provide effective ways for mode switches.

10.2.3. HCI as problem-solving

This dissertation presented use-driven basic research, which entails conducting basic research inspired by use cases. Touchless interaction with distant, vertical, 2D, large displays was selected owing to its increasing exploration in a variety of domains (reviewed in chapter 2). Although controlled studies cannot cater to the contextual understanding of interactions, they play an important role in exploring well-defined phenomena. For instance, in graphical user interfaces, pointer acceleration or adaptive control-display gain was first investigated in controlled settings with proven benefits, and then adapted widely in both Windows and Macintosh operating systems in practice. A similar research trajectory was followed in other interaction designs, such as the Microsoft Office ribbon or semantic pointing. Because in-the-wild studies suffer from limited control, it is often challenging to uncover what design decisions ultimately went on to benefit the overall user performance.

But then what aspects of interaction design should be tested in controlled studies? I argue for grounding empirical investigation in pertinent theories. For example, experiments in this dissertation tested hypothesis generated from theories in more traditional fields, cognitive science and motor behavior. A theoretical grounding can provide testable hypothesis when inquiring HCI design principles. This approach is complementary to implementing prototypes and studying them in practice. While the situated approach possesses higher ecological validity and generalizability, controlled experiments provide high internal validity. Furthermore, it is important to note that the research approach must be justified by the research question at hand. Technology adoption of robotic nurses is not suited to be studied in controlled laboratory settings. Similarly, whether UI components exhibiting the Gestalt principle of symmetry decreases user accuracy is unlikely to be figured out through field deployments of a mid-air keyboard. I chose the controlled laboratory setting because of this dissertation's research questions, which were well-defined and well-grounded in theory and susceptible to external confounds such as the context of use (gaming vs. work or sitting vs. standing). However, I believe the next step to controlled studies is in-the-wild follow-ups in different contexts. Contextual studies integrating the results of controlled studies are not suitable for hypothesis testing, but for exploring generalizability and ecological validity. Maybe one finding from a laboratory study would not improve user experience substantially, as perceived within a context of use, but a set of findings would sum up to a better user experience.

Chapter 11. Conclusion and open problems

In this final Chapter of the dissertation, I revisit the premise of this work, provide a summary of contributions, and discuss two future directions for exploring touchless interaction mechanics.

11.1. Conclusion

The premise of this work was to understand touchless as a sensorimotor phenomenon and present the generated knowledge to inform future touchless interaction design. Instead of a reactive adaptation to the ever-evolving sensing technology, I urged for an exploration of touchless interaction mechanics. To that end, I focused on the device-less property of touchless, looking at it from an embodied interaction perspective: what does a 'lack of a tool' entail in touchless interaction mechanics? I particularly studied target selection in touchless—a key user interaction, and found significant effects of several facets of visual perception and motor action, such as a good Gestalt continuity in user interface (UI) components made users faster and a hemispheric asymmetry improved the dominant hand performance when pre-programmed motor planning was less involved than feedback control.

This dissertation presented basic findings that can inform touchless interaction design and also introduced two novel interaction techniques: touchless circular menus and interface topographies. Touchless circular menus demonstrated target selection using 2D directional strokes—a motor-intuitive touchless interaction primitive. They were found to be more efficient but less accurate than grab-gesture based linear menus—and further affected by biomechanical factors. To improve touchless accuracy, interface topographies employed pseudo-haptic feedback; but accuracy improved only for difficult steering-targeting tasks.

11.2. Contribution to human-computer interaction

This work contributes to human-computer interaction by informing future touchless interaction designs. Within target selection, I presented several findings on touchless interaction mechanics—broadly classified into aspects of visual perception and motor action. In what follows, I highlight the most significant contributions of this work:

- **Motor-intuitive touchless interactions.** This work operationalized *intuitiveness* in touchless interactions using the continuum of knowledge in intuitive interaction. I further defined *motor-intuitive* as a property of a touchless interaction primitive, where the interaction mechanics draw from users' sensorimotor level of knowledge. We demonstrated this with an example primitive, two-dimensional directional strokes that draw on space schemas, and also evaluated their performance in controlled settings.

- **Touchless interaction primitives.** Instead of eliciting gestures from users or directly emulating interaction primitives of more traditional modalities (e.g., mouse- or pen-based interfaces), my work urged for examining touchless interaction primitives through the lens of *affordance and ability*: how interaction techniques can realize interface affordances and user abilities suitably. For example, I identified the strength and limitations of several interaction primitives that make up target-selection controls, such as the *translation-action ambiguity* in a push gesture or the *gesture-relaxation* problem in a crossing gesture. My stance is not that touchless systems should always be designed bottom-up, from interaction primitives to interface controls. Rather, my work can complement the top-down touchless research that elicits user-preferred gestural interactions by identifying the primitives involved and gauging their effectiveness.
- **Touchless user interface.** Findings of this dissertation also contribute to the design of user interfaces (UI) for touchless systems. For example, we found that Gestalt theories of visual perception and biomechanical factors affected user performance. Drawing on these findings, we proposed design guidelines for UI components as well as implications for further research (e.g., making UI components represent good continuity to make users faster or providing less frequently used commands on the non-dominant hemisphere in a circular menu).
- **Touchless circular menus (TCM).** For interacting with large displays, we introduced a touchless command selection technique using 2D directional strokes. TCM is an alternative to posture-based selection techniques, such as finger menus or grab. In our user evaluations, TCM was faster but less accurate than a grab-based linear menu.
- **Interface topographies.** Touchless interactions afford ample fluidity due to the absence of an input device constraining free movements. Such fluidity, however, makes touchless input imprecise, difficult to control, and frequently tiring. We introduced interface topographies to provide guidance around UI components using pseudo-haptic feedback. While some techniques like air voxels and tactile feedback have been previously explored to provide haptic feedback in touchless, they use dedicated setups or wearable hand gloves. In contrast, our proposed

method only involved manipulating the control-display ratio of the touchless interface.

- **Touchless transfer of training.** Finally, we provided insights on bilateral transfer of training in touchless interactions. In circular steering-targeting tasks, we found dominant hand excels in performance over the nondominant hand; but the nondominant hand performs significantly better, if the dominant hand is trained with pseudo-haptic feedback before. However, further research is required to investigate transfer of training in other touchless tasks.

Overall, this work took a basic science approach toward understanding touchless interaction and presented design insights. Future work needs to explore their relevance in actual systems, such as in large display interaction or mixed-device ecologies. Other than adapting the proposed design insights into building touchless systems in different domains, two other important directions for future work on touchless interaction mechanics are motor Gestalt and touchless pointing.

11.3. Motor Gestalt

Recent advances in motor science have found the effect of Gestalt theories in motor action (Klapp & Jagacinski, 2011). A recent review analyzed reaction-time results from previous studies and argued that four fundamental Gestalt principles in perception also apply to the control of motor action—holism, constancy, mutual exclusivity, and grouping in apparent motion. For example, certain motor actions, such as articulating a syllable during speech or making quick taps indicate the presence of motor Gestalts (chunks). In other works, the effect of Gestalt principles has been found in tactile perception (Gallace & Spence, 2011). This dissertation showed that touchless user interface (UI) designs can also be informed by Gestalt theories. But our focus was the visual design of the UI. An obvious trajectory of this research is to investigate motor Gestalt in touchless: How touchless gesture design can draw on Gestalt principles in the control of motor action? This research is both timely and significant, as it complements the ongoing research on mid-air text entry (Markussen et al., 2014), mid-air drawing (Taele & Hammond, 2014), and the general pursuit of an intuitive touchless gesture vocabulary.

Recent research has already begun to look at aspects similar to motor Gestalt, such as rhythmic patterns in touchless gestures (Carter, Velloso, Downs, Sellen, O'Hara, & Vetere, 2016) or the unique stimulus-response incompatibility due to the decoupling between visual and motor space (Markussen et al., 2014). It will be interesting to explore what features of touchless gestures are favorable to make the interactions intuitive.

11.4. Touchless pointing

This dissertation did not explore touchless pointing. Pointing in mid-air, however, is a significant area of research. Although pointing is extensively studied in desktop interfaces (Casiez et al., 2008), touchless pointing involves several new challenges, some of which are specific to application domains. For example, while interacting with large displays, pointing in the mid-air involves large gains in the control-display ratio and clutching issues. Furthermore, there may be pointer acceleration and variable control-display gains. Ongoing research is exploring these issues, such as designing sub-space gestures, where users can dynamically design a personal interaction space for effective clutching (Rateau, Grisoni, & De Araujo, 2014), or using the tow-part Welford's model to capture mid-air pointing in large-display interaction (Shoemaker et al., 2012).

Shoemaker et al. (2012) found that Fitts's law does not appropriately model mid-air pointing on very large displays; instead, they showed that the two-part Welford's model is a better fit for constant control-display gains in large-display pointing. Previously, Fitts's law has been adapted in several ways, like for semantic pointing (Blanch, Guiard, & Beaudouin-Lafon, 2004) or virtual worlds (Balakrishnan, 2004). The overarching theory supporting the two-part model argues that pointing involves two distinct sensorimotor processes, one causing the initial ballistic movement and another the corrective movement, and while the ballistic movement depends on the amplitude of a target, the corrective movement depends on the target width. Thus, if these two processes occur at different rates, they would require different coefficients.

Future research could explore two directions related to touchless pointing: (1) model pointer acceleration and variable gain in touchless pointing on large displays or (2) investigate visual and pseudo-haptic feedback toward improving pointing accuracy. Effective clutching techniques in touchless pointing also remains an open problem.

Appendices

Appendix A. Visual Feedback

A.1.Training

During the training session in the first round of the study (experiments 1 - 5), participants practiced select and de-select gestures by solving a picture puzzle (Figure A1). They rearranged a puzzle using drag-and-drop operations. Each participant completed three picture puzzles, and on average took 10 – 15 minutes to complete all three of them.

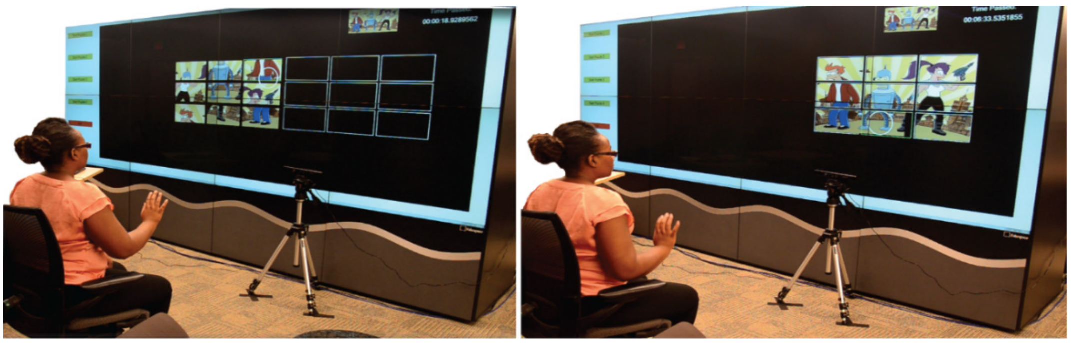

Figure A1. During the training session in the first round of the study, participants practiced select and de-select gestures by solving a picture puzzle.

A.2. Color Conversion from Munsell Notation to RGB

Five Munsell colors (Fig.1, p. 139, Smith and Thomas, 1964)—red, green, blue, yellow and white was used in experiment 2. Munsell notation was converted to RGB hex values using an R script. An example of the conversion code for color green (2.5G 5/8) is given below:

```
library(aqp)
library(colorspace)
rgbVal <- expand.grid(hue='2.5G', value=5,chroma=8)
rgbVal.rgb <- with(rgbVal, munsell2rgb(hue, value, chroma,return_triplets=TRUE))
newRgb = rgb(rgbVal.rgb$r, rgbVal.rgb$g, rgbVal.rgb$b)
```

After conversion, each color corresponded to a hex color code: green (2.5G 5/8) to #238C57; blue (5BG 4/5) to #156D69; white (5Y 8/4) to #D9CA93; red (5R 4/9) to #A34143; and yellow (10YR 6/10) to #C68A13.

A.3. Stoppers—Semantic Feedback for Out-of-Range Gestures

Figure A2. A user points in mid-air to a target folder on a large display (left); Stoppers provide visual feedback as the user's gesture goes out of the display range (center) and guide her back within the display range (right).

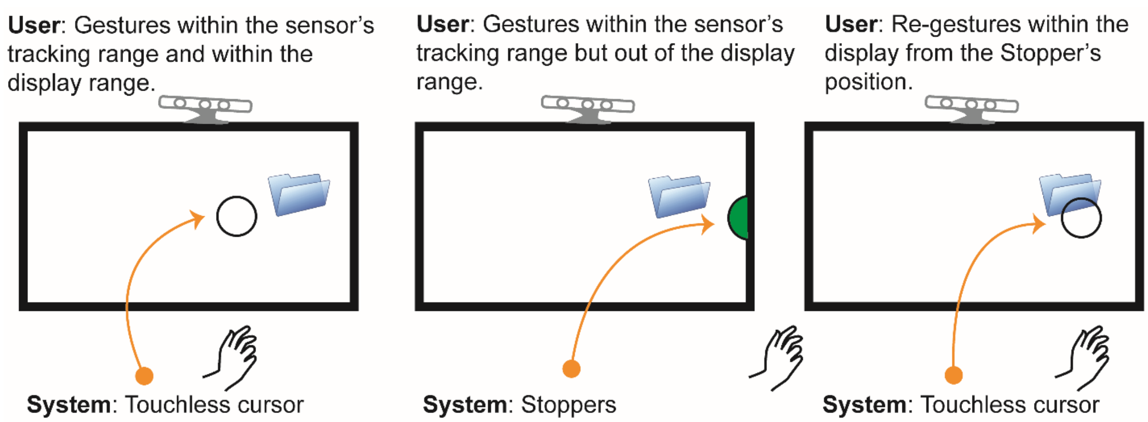

Figure A3. By introducing persistent visual feedback as users move out of the display range (center), Stoppers decrease users' disorientation and facilitate the recovery of touchless gestures within the display range (right).

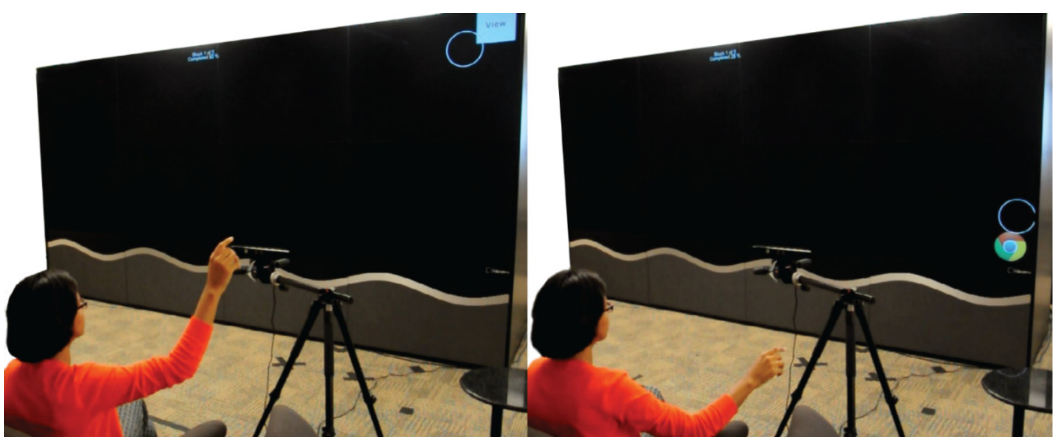

Figure A4. In the second round, participants performed a pointing task with targets (256 pixels x 256 pixels) randomly appearing at the top, left or right border of the large display.

Appendix B. Preliminary work on touchless Gestalt

While conducting controlled experiments to study command-selection techniques (Chapter 6) for large-display touchless interactions, we observed certain performance trends that were unnoticed by users. Visual elements of the UI affected users' effectiveness (e.g., aiming a menu for command-selection). Interestingly, our findings could be explained by Gestalt principles of perceptual grouping: similarity of shape.

Across our experiments, users sat about 1.5 – 2.5m away from a large display (4 x 1.5m) and were tracked by Kinect sensors.

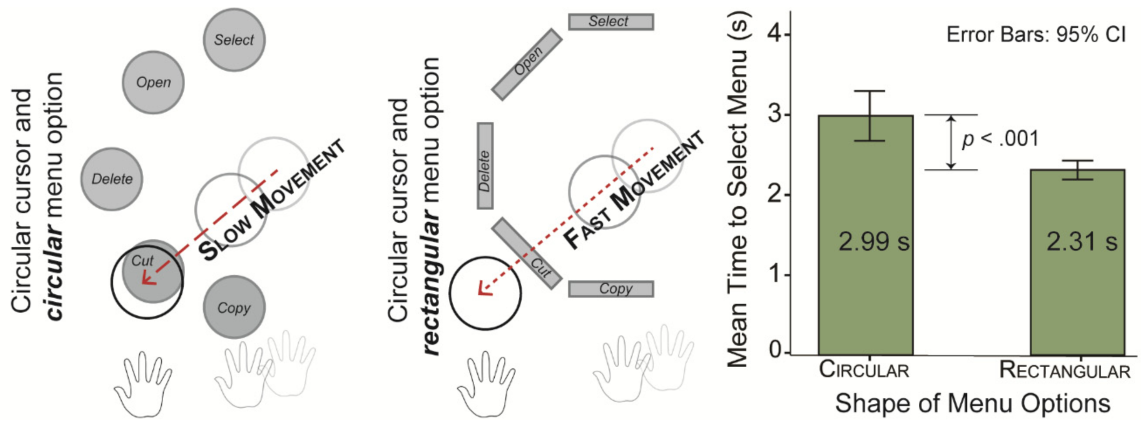

Figure B1. Perceptual grouping by Similarity of shape principle affected the efficiency of touchless interaction: Expert users were faster when crossing-to-select a rectangular menu option than a circular menu option with a circular cursor.

Similarity of shape decreases efficiency

To relieve users from strictly complying with system-defined postures as interaction commands, we introduced a command-selection technique using mid-air strokes—Touchless Circular Menus (Chapter 6). To trigger the contextual TCM, users would land on the target folder, and to select a command, users would simply cross the menu option (Figure B2). In our early iterations, the menu options (230px) were circular, isomorphic to the cursor (256px). During pilot testing with *three expert users*, we found them slowing down while crossing the menu-option. When the cursor was over the menu-option, users would tend to slow down as if they were placing the cursor over the menu-option, rather than crossing it (in spite of prior instructions and practice trials). The menu options were about 800 pixels away from the folder (13.7cm in control space). When the menu options were modified to rectangles (at the same distance), users became significantly faster. This occurred in spite of users essentially traversing the same distance: For circular options, users had to move across half the menu-option, and for rectangles cross the entire menu-option. Shape of the menu options (circular, M = 2.99s, SD = 2.0; rectangular, M = 2.31s, SD = 0.76) significantly affected efficiency (Log10 reaction time) with a small effect size, $n = 161$, $t(160) = 4.19$, $p < .001$, $d = .33$.

This finding can be explained using the Gestalt principle of perceptual grouping by similarity of shape: all else being equal, the most similar visual elements in shape tend to be grouped together (Wagemans et al., 2012a). Our results suggested that users must have perceived the circular cursor and the circular menu-option as a group—at least momentarily—and thus slowed their motor action to discriminate between the object of action (the circular menu option) and the symbolic referent of their action (the circular cursor).

Limitations

Our findings are posteriori arguments and are limited by our tracking sensors. Other limitations include not explicitly controlling for the index of difficulty in the crossing-based trials, expert users, and a small sample size.

Bibliography

Accot, J., & Zhai, S. (1997). Beyond Fitts' law: models for trajectory-based HCI tasks. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 295–302, ACM. doi:10.1145/258549.258760

Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological Bulletin*, *101*(1), 41–74. doi:10.1037/0033-2909.101.1.41

Aigner, R., Wigdor, D., Benko, H., Haller, M., Lindlbauer, D., Ion, A., … Redmond, W. A. (2012). Understanding mid-air hand gestures: A study of human preferences in usage of gesture types for HCI. (Microsoft Research Tech Report MSR-TR-2012-111). Redmond, WA: Microsoft. http://131.107.65.14/pubs/175454/GesturesTR-20121107-RoA.pdf (accessed January 8, 2015).

Alexander, R. M. (1997). A minimum energy cost hypothesis for human arm trajectoires. *Biological Cybernetics*, *76*(2), 97–105. doi:10.1007/s004220050324

Andrews, C., Endert, A., & North, C. (2010). Space to think: large high-resolution displays for sensemaking. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 55–64, ACM. doi:10.1145/1753326.1753336

Apitz, G., Guimbretière, F., & Zhai, S. (2010). Foundations for designing and evaluating user interfaces based on the crossing paradigm. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *17*(2), 9. doi:10.1145/1746259.1746263

Bailly, G., Walter, R., Müller, J., Ning, T., & Lecolinet, E. (2011). Comparing free hand menu techniques for distant displays using linear, marking and finger-count menus. In Proceedings of the INTERACT Conference, 248–262, Springer. doi:10.1007/978-3-642-23771-3_19

Balakrishnan, R. (2004). “Beating” Fitts’ law: virtual enhancements for pointing facilitation. *International Journal of Human-Computer Studies*, *61*(6), 857–874. doi:10.1016/j.ijhcs.2004.09.002

Ballendat, T., Marquardt, N., et al. (2010). Proxemic interaction: designing for a proximity and orientation-aware environment. In Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces, 121–130, ACM. doi:10.1145/1936652.1936676

Banerjee, A., Burstyn, J., Girouard, A., & Vertegaal, R. (2011). Pointable: an in-air pointing technique to manipulate out-of-reach targets on tabletops.

In Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces, 11–20, ACM. doi:10.1145/2076354.2076357

Bates, D., and M. Maechler. (2010). Package ‘lme4.’ http://lme4.r-forge.r-project.org/

Bau, O., Poupyrev, I., Israr, A., & Harrison, C. (2010). TeslaTouch: electrovibration for touch surfaces. In Proceedings of the Annual ACM Symposium on User Interface Software and Technology, 283–292, ACM. doi:10.1145/1866029.1866074

Baudisch, P., Good, N., & Stewart, P. (2001). Focus plus context screens: combining display technology with visualization techniques. In Proceedings of the Annual ACM Symposium on User Interface Software and Technology, 31–40, ACM. doi:10.1145/502348.502354

Baudisch, P., Cutrell, E., Robbins, D., Czerwinski, M., Tandler, P., Bederson, B., & Zierlinger, A. (2003). Drag-and-pop and drag-and-pick: Techniques for accessing remote screen content on touch-and pen-operated systems. In Proceedings of INTERACT, Vol. 3, 57–64.

Baudisch, P., Cutrell, E., & Robertson, G. (2003). High-density cursor: a visualization technique that helps users keep track of fast-moving mouse cursors. In Proceedings of INTERACT, 236–243, Springer.

Baudisch, P., Cutrell, E., Hinckley, K., & Gruen, R. (2004). Mouse ether: accelerating the acquisition of targets across multi-monitor displays. Extended Abstracts, In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1379–1382, ACM. doi:10.1145/985921.986069

Beaudouin-Lafon, M. (2004, May). Designing interaction, not interfaces. In Proceedings of the Working Conference on Advanced Visual Interfaces, 15–22, ACM. doi:10.1145/989863.989865

Beaudouin-Lafon, M., Huot, S., Nancel, M., Mackay, W., Pietriga, E., Primet, R., ... & Klokmose, C. (2012). Multisurface interaction in the WILD room. *Computer*, *45*(4), 48–56. doi:10.1109/MC.2012.110

Bellucci, A., Malizia, A., Diaz, P., & Aedo, I. (2010). Don't touch me: multi-user annotations on a map in large display environments. In Proceedings of the International Conference on Advanced Visual Interfaces, 391–392, ACM. doi:10.1145/1842993.1843072

Bezerianos, A., & Balakrishnan, R. (2004). Interaction and visualization techniques for very large scale high resolution displays. University of Toronto Technical Report

DGP-TR-2004-002. http://www.dgp.utoronto.ca/techreports/dgp-tr-2004-002.pdf (accessed on February 1, 2015).

Bezerianos, A., & Balakrishnan, R. (2005). The vacuum: facilitating the manipulation of distant objects. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 361–370, ACM. doi:10.1145/1054972.1055023

Bier, E. A., Stone, M. C., et al. (1993). Toolglass and magic lenses: the see-through interface. In Proceedings of the 20th annual conference on Computer graphics and interactive techniques, 73–80, ACM. doi:10.1145/166117.166126

Blackler, A. L., & Hurtienne, J. (2007). Towards a unified view of intuitive interaction: definitions, models and tools across the world. *MMI-Interaktiv*, *13*(2007), 36–54. Retrieved from http://eprints.qut.edu.au/19116/ on May 6, 2016.

Blackler, A., Popovic, V., & Mahar, D. (2010). Investigating users' intuitive interaction with complex artefacts. *Applied ergonomics*, *41*(1), 72–92. doi:10.1016/j.apergo.2009.04.010

Blanch, R., Guiard, Y., & Beaudouin-Lafon, M. (2004). Semantic pointing: improving target acquisition with control-display ratio adaptation. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 519–526, ACM. doi:10.1145/985692.985758

Bragdon, A., DeLine, R., Hinckley, K., & Morris, M. R. (2011). Code space: touch+ air gesture hybrid interactions for supporting developer meetings. In Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces, 212–221, ACM. doi:10.1145/2076354.2076393

Bragdon, A., & Ko, H. S. (2011). Gesture select: acquiring remote targets on large displays without pointing. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 187–196, ACM. doi:10.1145/1978942.1978970

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, *189*(194) 4–7.

Callahan, J., Hopkins, D., Weiser, M., et al. (1988). An empirical comparison of pie vs. linear menus. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 95–100, ACM. doi:10.1145/57167.57182

Cao, X., & Balakrishnan, R. (2003). VisionWand: Interaction techniques for large displays using a passive wand tracked in 3D. In Proceedings of the Annual ACM Symposium on User Interface Software and Technology, 173–182, ACM. doi:10.1145/964696.964716

Carter, R. C. (1982). Visual search with color. *Journal of Experimental Psychology: Human Perception and Performance*, *8*(1), 127. doi:10.1037/0096-1523.8.1.127

Carter, T., Seah, S. A., Long, B., et al. (2013). Ultrahaptics: multi-point mid-air haptic feedback for touch surfaces. In Proceedings of the Annual ACM Symposium on User Interface Software and Technology, 505–514, ACM. doi:10.1145/2501988.2502018

Carter, M., Velloso, E., Downs, J., Sellen, A., O'Hara, K., & Vetere, F. (2016). PathSync: Multi-User Gestural Interaction with Touchless Rhythmic Path Mimicry. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 3415–3427, ACM. doi:10.1145/2858036.2858284

Casiez, G., Vogel, D., Balakrishnan, R., & Cockburn, A. (2008). The impact of control-display gain on user performance in pointing tasks. *Human–Computer Interaction*, *23*(3), 215–250. doi:10.1080/07370020802278163

Chattopadhyay, D. and Bolchini, D. (2013) Laid-back, touchless collaboration around wall-size displays: visual feedback and affordances. In Proceedings of the SIGCHI POWERWALL Workshop. Retrieved from http://hdl.handle.net/1805/4526 on October 30, 2014.

Criscimagna-Hemminger, S. E., Donchin, O., Gazzaniga, M. S., & Shadmehr, R. (2003). Learned dynamics of reaching movements generalize from dominant to nondominant arm. *Journal of neurophysiology*, *89*(1), 168–176. doi:10.1152/jn.00622.2002

Collomb, M., Hascoët, M., Baudisch, P., & Lee, B. (2005). Improving drag-and-drop on wall-size displays. In Proceedings of Graphics Interface, 25–32, Canadian Human-Computer Communications Society. Retrieved from http://dl.acm.org/citation.cfm?id=1089514 on May, 6, 2016.

Cockburn, A., Quinn, P., Gutwin, C., Ramos, G., & Looser, J. (2011). Air pointing: Design and evaluation of spatial target acquisition with and without visual feedback. *International Journal of Human-Computer Studies*, *69*(6), 401–414. doi:10.1016/j.ijhcs.2011.02.005

Cohen, G. (1973). Hemispheric differences in serial versus parallel processing. *Journal of Experimental Psychology*, *97*(3), 349. doi:10.1037/h0034099

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155. doi:10.1037/0033-2909.112.1.155

Czerwinski, M., Tan, D. S., & Robertson, G. G. (2002). Women take a wider view. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 195–202, ACM. doi:10.1145/503376.503412

Czerwinski, M., Smith, G., Regan, T., Meyers, B., Robertson, G., & Starkweather, G. (2003). Toward characterizing the productivity benefits of very large displays. In Proceedings of INTERACT, Vol. 3, 9–16.

Death of the Desktop (2014). Envisioning Visualization without Desktop Computing. Workshop collocated with IEEE VIS 2014, Paris. Retrieved from http://beyond.wallviz.dk/ on January 8, 2015.

Dennerlein, J. T., et al. (2000). Force-feedback improves performance for steering and combined steering-targeting tasks. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 423–429, ACM. doi:10.1145/332040.332469

Dourish, P. (2004). *Where the action is: The foundations of embodied interaction*. Cambridge, MA: MIT press.

Dostal, J., Hinrichs, U., Kristensson, P. O., & Quigley, A. (2014). SpiderEyes: designing attention-and proximity-aware collaborative interfaces for wall-sized displays. In Proceedings of the International Conference on Intelligent User Interfaces, 143–152, ACM. doi:10.1145/2557500.2557541

Dsouza, M., Burggraaf. J., Kamm, C., Tewarie, P., Kontschieder, P., Dorn, J., …, Kappos, L. (2014). Infrared depth sensor based automated classification of motor dysfunction in multiple sclerosis - a proof-of-concept study, In European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS). Retrieved from http://research.microsoft.com/apps/pubs/default.aspx?id=231970 on January 8, 2015.

Durnford, M., & Kimura, D. (1971). Right hemisphere specialization for depth perception reflected in visual field differences. *Nature*, *231*(5302), 394–395. doi:10.1038/231394a0

D'Zmura, M. (1991). Color in visual search. *Vision research*, *31*(6), 951–966. doi:10.1016/0042-6989(91)90203-H

Ells, J. G. (1973). Analysis of temporal and attentional aspects of movement control. *Journal of Experimental Psychology*, *99*(1), 10. doi:10.1037/h0034740

Elrod, S., Bruce, R., Gold, R., Goldberg, D., Halasz, F., Janssen, W., ... & Welch, B. (1992). Liveboard: a large interactive display supporting group meetings, presentations, and remote collaboration. In Proceedings of the SIGCHI Conference on Human factors in Computing Systems, 599–607, ACM. doi:10.1145/142750.143052

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*(6), 381. doi:10.1037/h0055392

Fitts, P. M., & Peterson, J. R. (1964). Information capacity of discrete motor responses. *Journal of Experimental Psychology*, *67*(2), 103. doi:10.1037/h0045689

Foehrenbach, S., König, W. A., Gerken, J., & Reiterer, H. (2009). Tactile feedback enhanced hand gesture interaction at large, high-resolution displays. *Journal of Visual Languages & Computing*, *20*(5), 341–351. doi:10.1016/j.jvlc.2009.07.005

Freeman, E., Brewster, S., & Lantz, V. (2014). Tactile feedback for above-device gesture interfaces: Adding touch to touchless interactions. In Proceedings of the 16th International Conference on Multimodal Interaction, 419–426, ACM. doi:10.1145/2663204.2663280

Fung, R., Lank, E., Terry, M., & Latulipe, C. (2008). Kinematic templates: end-user tools for content-relative cursor manipulations. In Proceedings of the ACM annual symposium on User Interface Software and Technology, 47–56, ACM. doi:10.1145/1449715.1449725

Gallace, A., & Spence, C. (2011). To what extent do Gestalt grouping principles influence tactile perception? *Psychological Bulletin*, *137*(4), 538–561. doi:10.1037/a0022335

Garzotto, F., & Valoriani, M. (2012). Don't touch the oven: motion-based touchless interaction with household appliances. In Proceedings of the International Working Conference on Advanced Visual Interfaces, 721–724, ACM. doi:10.1145/2254556.2254693

Gaver, W. W. (1991). Technology affordances. In Proceedings of the SIGCHI Conference on Human factors in Computing Systems, 79–84, ACM. doi:10.1145/108844.108856

Gaver, W. W. (1996). Situating action II: Affordances for interaction: The social is material for design. *Ecological Psychology*, *8*(2), 111–129. doi:10.1207/s15326969eco0802_2

Gibson, J. J. (1971). The legacies of Koffka's Principles. *Journal of the History of the Behavioral Sciences.* 7, 3–9. doi:10.1002/1520-6696(197101)7:1<3::AID-JHBS2300070102>3.0.CO;2-1

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Houghton Mifflin.

Grandhi, S. A., Joue, G., & Mittelberg, I. (2011). Understanding naturalness and intuitiveness in gesture production: insights for touchless gestural interfaces. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 821–824, ACM. doi:10.1145/1978942.1979061

Grinstein, G., Cook, K., Havig, P., Liggett, K., Nebesh, B., Whiting, M., ... & Knoecni, S. (2011). VAST 2011 Challenge: Cyber security and epidemic. IEEE VAST'11, 299–301. Retrieved from https://www.semanticscholar.org/paper/VAST-2011-Challenge-Cyber-Security-and-Epidemic-Grinstein-Cook/17d93bc0d4c79edf4ac37f657db12256310573b0/pdf on May, 23, 2016.

Grossman, T., & Balakrishnan, R. (2004). Pointing at trivariate targets in 3D environments. In Proceedings of the SIGCHI conference on Human factors in computing systems, 447–454, ACM. doi:10.1145/985692.985749

Guiard, Y. (1987). Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *Journal of Motor behavior*, *19*(4), 486–517. doi:10.1080/00222895.1987.10735426

Guimbretiére, F., & Winograd, T. (2000). FlowMenu: combining command, text, and data entry. In Proceedings of the ACM Annual Symposium on User Interface Software and Technology, 213–216, ACM. doi:10.1145/354401.354778

Guimbretière, F., & Nguyen, C. (2012). Bimanual marking menu for near surface interactions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 825–828, ACM. doi:10.1145/2207676.2208521

Gupta, S., Morris, D., Patel, S. N., & Tan, D. (2013). Airwave: Non-contact haptic feedback using air vortex rings. In Proceedings of the ACM international joint conference on Pervasive and Ubiquitous Computing, 419–428, ACM. doi:10.1145/2493432.2493463

Gustafson, S., Bierwirth, D., & Baudisch, P. (2010). Imaginary interfaces: spatial interaction with empty hands and without visual feedback. In Proceedings of the ACM Annual Symposium on User Interface Software and Technology, 3–12, ACM. doi:10.1145/1866029.1866033

Han, J., Shao, L., Xu, D., & Shotton, J. (2013). Enhanced computer vision with Microsoft Kinect sensor: A review. *IEEE Transactions on Cybernetics*, *43*(5), 1318–1334. doi:10.1109/TCYB.2013.2265378

Hansen, L. K., & Dalsgaard, P. (2015). Note to Self: Stop Calling Interfaces "Natural". *Aarhus Series on Human Centered Computing*, *1*(1), 4. doi:10.7146/aahcc.v1i1.21316

Harrison, B. L., Kurtenbach, G., & Vicente, K. J. (1995). An experimental evaluation of transparent user interface tools and information content. In Proceedings of the ACM Annual Symposium on User Interface Software and Technology, 81–90, ACM. doi:10.1145/215585.215669

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human mental workload*, *1*(3), 139–183. doi:10.1016/S0166-4115(08)62386-9

Hayward, V., Astley, O. R., Cruz-Hernandez, M., Grant, D., & Robles-De-La-Torre, G. (2004). Haptic interfaces and devices. *Sensor Review*, *24*(1), 16–29. doi:10.1108/02602280410515770

Heidegger, M. (1988). *The basic problems of phenomenology*. Indiana University Press.

Hespanhol, L., Tomitsch, M., Grace, K., Collins, A., & Kay, J. (2012). Investigating intuitiveness and effectiveness of gestures for free spatial interaction with large displays. In Proceedings of the International Symposium on Pervasive Displays, 6, ACM. doi:10.1145/2307798.2307804

Hincapié-Ramos, J. D., Guo, X., Moghadasian, P., & Irani, P. (2014). Consumed Endurance: A metric to quantify arm fatigue of mid-air interactions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1063–1072, ACM. doi:10.1145/2556288.2557130

Hinckley, K., Pausch, R., Goble, J. C., & Kassell, N. F. (1994). A survey of design issues in spatial input. In Proceedings of the 7th Annual ACM Symposium on User Interface Software and Technology, 213–222, ACM. doi:10.1145/192426.192501

Hinckley, K., Baudisch, P., Ramos, G., & Guimbretiere, F. (2005). Design and analysis of delimiters for selection-action pen gesture phrases in scriboli. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 451–460, ACM. doi:10.1145/1054972.1055035

Hoshi, T., Takahashi, M., Iwamoto, T., and Shinoda, H. Noncontact Tactile Display Based on Radiation Pressure of Airborne Ultrasound. *IEEE Trans. on Haptics*, *3*(3), 155–165. doi:10.1109/TOH.2010.4

Huang, E. M., Mynatt, E. D., & Trimble, J. P. (2006). Displays in the wild: understanding the dynamics and evolution of a display ecology. Pervasive Computing, 321–336, Springer Berlin Heidelberg. doi:10.1007/11748625_20

Huegel, J. C., Celik, O., Israr, A., & O'Malley, M. K. (2009). Expertise-based performance measures in a virtual training environment. *Presence: Teleoperators and Virtual Environments*, *18*(6), 449–467. doi:10.1162/pres.18.6.449

Hurst, A., Mankoff, J., Dey, A. K., & Hudson, S. E. (2007). Dirty desktops: using a patina of magnetic mouse dust to make common interactor targets easier to select. In Proceedings of the ACM Annual Symposium on User Interface Software and Technology, 183–186, ACM. doi:10.1145/1294211.1294242

Hurtienne, J., & Israel, J. H. (2007). Image schemas and their metaphorical extensions: intuitive patterns for tangible interaction. In Proceedings of the International Conference on Tangible and Embedded Interaction, 127–134, ACM. doi:10.1145/1226969.1226996

Hutchins, E. L., Hollan, J. D., & Norman, D. A. (1985). Direct manipulation interfaces. *Human–Computer Interaction*, *1*(4), 311–338. doi:10.1207/s15327051hci0104_2

ISO, 2002. Ergonomic requirements for office work with visual display terminals (VDTs). Requirements for non-keyboard input devices (ISO 9241-9). Retrieved from http://www.iso.org/iso/iso_catalogue/catalogue_ics/catalogue_detail_ics.htm?csnumber=31169 on May 6, 2016.

Ihde, D. (1990). *Technology and the lifeworld: From garden to earth* (No. 560). Indiana University Press.

Jacob, R. J., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T., & Zigelbaum, J. (2008). Reality-based interaction: a framework for post-WIMP interfaces. In Proceedings of the SIGCHI conference on Human factors in computing systems, 201–210, ACM. doi:10.1145/1357054.1357089

Jagodic, R., Renambot, L., Johnson, A., Leigh, J., & Deshpande, S. (2011). Enabling multi-user interaction in large high-resolution distributed environments. *Future Generation Computer Systems, 27*(7), 914–923. doi:10.1016/j.future.2010.11.018

Jagodic, R. (2011). *Collaborative Interaction and Display Space Organization in Large High-Resolution Environments*. Ph.D. Dissertation. University of Illinois at Chicago, Chicago, IL, USA. Retrieved from http://search.proquest.com/docview/1285215124 on February 1, 2015.

Jakobsen, M. R., & Hornbæk, K. (2014). Up close and personal: Collaborative work on a high-resolution multitouch wall display. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *21*(2), 11. doi:10.1145/2576099

Jansen, Y., Dragicevic, P., & Fekete, J. D. (2012). Tangible remote controllers for wall-size displays. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2865–2874, ACM. doi:10.1145/2207676.2208691

Jennings, Mike. HP Envy 17 Leap Motion SE review. 2014. Retrieved from http://www.trustedreviews.com/hp-envy-17-leap-motion-se-review on May 6, 2016.

Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination, and reason*. The University of Chicago Press, Chicago & London.

Johnson, R., O'Hara, K., Sellen, A., Cousins, C., & Criminisi, A. (2011). Exploring the potential for touchless interaction in image-guided interventional radiology. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 3323–3332, ACM. doi:10.1145/1978942.1979436

Jota, R., Pereira, J. M., & Jorge, J. A. (2009). A comparative study of interaction metaphors for large-scale displays. Extended Abstracts, In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 4135–4140, ACM. doi:10.1145/1520340.1520629

Jota, R., Nacenta, M. A., Jorge, J. A., Carpendale, S., & Greenberg, S. (2010). A comparison of ray pointing techniques for very large displays. In Proceedings of Graphics Interface 2010, 269-276, Canadian Information Processing Society. Retrieved from http://dl.acm.org/citation.cfm?id=1839261 on May 6, 2016.

Jude, A., Poor, G. M., & Guinness, D. (2014). An evaluation of touchless hand gestural interaction for pointing tasks with preferred and non-preferred hands. In Proceedings of the 8th Nordic Conference on Human-Computer Interaction, 668–676, ACM. doi:10.1145/2639189.2641207

Kabbash, P., MacKenzie, I. S., & Buxton, W. (1993). Human performance using computer input devices in the preferred and non-preferred hands. In Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems, 474–481, ACM. doi:10.1145/169059.169414

Kabbash, P., Buxton, W., & Sellen, A. (1994). Two-handed input in a compound task. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 417–423, ACM. doi:10.1145/191666.191808

Kajastila, R., & Lokki, T. (2013). Eyes-free interaction with free-hand gestures and auditory menus. *International Journal of Human-Computer Studies*, *71*(5), 627–640. doi:10.1016/j.ijhcs.2012.11.003

Kamuro, S., Minamizawa, K., Kawakami, N., & Tachi, S. (2009). Ungrounded kinesthetic pen for haptic interaction with virtual environments. In Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication, 436–441, IEEE. doi:10.1109/ROMAN.2009.5326217

Karam, M. & Schraefel, M. C. (2005). A taxonomy of gesture in human computer interactions. Technical Report, ECSTR-IAM05-009, Electronics and Computer Science, University of Southampton. Retrieved from http://eprints.soton.ac.uk/261149/ on May 6, 2016.

Kimura, D., & Vanderwolf, C. H. (1970). The relation between hand preference and the performance of individual finger movements by left and right hands. *Brain: A Journal of Neurology*, *93*(4), 769–774. doi:10.1093/brain/93.4.769

Kirsh, D. (2013). Embodied cognition and the magical future of interaction design. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *20*(1), 3. doi:10.1145/2442106.2442109

Kister, U., Reipschläger, P., Matulic, F., & Dachselt, R. (2015). BodyLenses: Embodied Magic Lenses and Personal Territories for Wall Displays. In Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces,117–126. ACM. doi:10.1145/2817721.2817726

Klapp, S. T. (1975). Feedback versus motor programming in the control of aimed movements. *Journal of Experimental Psychology: Human Perception and Performance*, *1*(2), 147. doi:10.1037/0096-1523.1.2.147

Klapp, S. T., & Jagacinski, R. J. (2011). Gestalt principles in the control of motor action. *Psychological Bulletin*, *137*(3), 443–462. doi:10.1037/a0022361

Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In C. Freksa & C. Habel (Eds.), Spatial cognition. An interdisciplinary approach to representing and processing spatial knowledge, 1–17, Heidelberg: Springer-Verlag. doi:10.1007/3-540-69342-4_1

Koffka, K. (1922). Perception: An introduction to the “Gestalt-theorie”. *Psychological Bulletin*, *19*(10), 531–585. doi:10.1037/h0072422

Kulshreshth, A., & LaViola Jr, J. J. (2014). Exploring the usefulness of finger-based 3D gesture menu selection. In Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems, 1093–1102, ACM. doi:10.1145/2556288.2557122

Kurtenbach, G., & Buxton, W. (1994). User learning and performance with marking menus. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 258–264, ACM. doi:10.1145/191666.191759

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Lécuyer, A., Burkhardt, J. M., & Etienne, L. (2004). Feeling bumps and holes without a haptic interface: The perception of pseudo-haptic textures. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 239–246, ACM. doi:10.1145/985692.985723

Lécuyer, A. (2009). Simulating haptic feedback using vision: A survey of research and applications of pseudo-haptic feedback. *Presence: Teleoperators and Virtual Environments*, *18*(1), 39–53. doi:10.1162/pres.18.1.39

Lee, J. C. (2010). In search of a natural gesture. *ACM Crossroads*, *16*(4), 9–12. doi:10.1145/1764848.1764853

Lehtinen, V., Oulasvirta, A., Salovaara, A., & Nurmi, P. (2012). Dynamic tactile guidance for visual search tasks. In Proceedings of the ACM Annual Symposium on User Interface Software and Technology, 445–452, ACM. doi:10.1145/2380116.2380173

Leigh, J., Johnson, A., Renambot, L., Peterka, T., Jeong, B., Sandin, D. J., ... & Sun, Y. (2013). Scalable resolution display walls. *Proceedings of the IEEE*, *101*(1), 115–129. doi:10.1109/JPROC.2012.2191609

Lenman, S., Bretzner, L., & Thuresson, B. (2002). Using marking menus to develop command sets for computer vision based hand gesture interfaces. In Proceedings of the 8th Nordic Conference on Human-Computer Interaction, 239–242, ACM. doi:10.1145/572020.572055

Lepinski, G. J., Grossman, T., & Fitzmaurice, G. (2010). The design and evaluation of multitouch marking menus. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2233–2242, ACM. doi:10.1145/1753326.1753663

Liu, C., Chapuis, O., Beaudouin-Lafon, M., Lecolinet, E., & Mackay, W. E. (2014). Effects of display size and navigation type on a classification task. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 4147–4156, ACM. doi:10.1145/2556288.2557020

Macaranas, A., Antle, A. N., & Riecke, B. E. (2015). What is intuitive interaction? balancing users' performance and satisfaction with natural user interfaces. *Interacting with Computers*, *27*(3), 357–370. doi:10.1093/iwc/iwv003

MacKenzie, I. S., & Buxton, W. (1992). Extending Fitts' law to two-dimensional tasks. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 219–226, ACM. doi:10.1145/142750.142794

Magill, R. A., & Anderson, D. (2007). *Motor learning and control: Concepts and applications, 11*. New York: McGraw-Hill.

Malfait, N., & Ostry, D. J. (2004). Is interlimb transfer of force-field adaptation a cognitive response to the sudden introduction of load? *The Journal of Neuroscience*, *24*(37), 8084–8089. doi:10.1523/JNEUROSCI.1742-04.2004

Malik, S., Ranjan, A., & Balakrishnan, R. (2005). Interacting with large displays from a distance with vision-tracked multi-finger gestural input. In Proceedings of the ACM Annual Symposium on User Interface Software and Technology, 43–52, ACM. doi:10.1145/1095034.1095042

Malizia, A., & Bellucci, A. (2012). The artificiality of natural user interfaces. *Communications of the ACM*, *55*(3), 36–38. doi:10.1145/2093548.2093563

Marr, D. (1982). *Vision: A computational investigation*. New York, NY: Freeman.

Markussen, A., Jakobsen, M. R., & Hornbæk, K. (2014). Vulture: a mid-air word-gesture keyboard. In Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems, 1073–1082. ACM. doi:10.1145/2556288.2556964

Markussen, A., Jakobsen, M. R., & Hornbæk, K. (2013). Selection-based mid-air text entry on large displays. In Human-Computer Interaction–INTERACT 2013, pp. 401–418, Springer Berlin Heidelberg. doi:10.1007/978-3-642-40483-2_28

Mentis, H. M., O'Hara, K., Sellen, A., & Trivedi, R. (2012). Interaction proxemics and image use in neurosurgery. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 927–936, ACM. doi:10.1145/2207676.2208536

Merleau-Ponty, M. (1996). *Phenomenology of perception*. Motilal Banarsidass Publishe.

Merleau-Ponty, M., & Baldwin, T. (2004). *The world of perception*. London: Routledge.

Microsoft Corporation. (2013). Human Interface Guidelines (HIG) version 1.8. Retrieved from http://go.microsoft.com/fwlink/?LinkID=247735 on May 6, 2016.

Microsoft. (2014). Retrieved from https://support.xbox.com/en-US/xbox-360/kinect/kinect-sensor-setup on May 6, 2016.

Microsoft, Kinect for Windows Human Interface Guidelines v1.8.0. 2016. Retrieved from http://go.microsoft.com/fwlink/?LinkID=247735 on April 28, 2016

Miller, V. (2011). *Understanding digital culture*. Sage Publications.

Miller, T., & Zeleznik, R. (1998,). An insidious haptic invasion: adding force feedback to the X desktop. In Proceedings of the ACM Annual Symposium on User Interface Software and Technology, 59–64, ACM. doi:10.1145/288392.288573

Mine, M. R., Brooks Jr, F. P., & Sequin, C. H. (1997). Moving objects in space: exploiting proprioception in virtual-environment interaction. In Proceedings of the Annual Conference on Computer Graphics and Interactive Techniques, 19–26, ACM. doi:10.1145/258734.258747

Monnai, Y., Hasegawa, K., Fujiwara, M, et al. (2014). HaptoMime: mid-air haptic interaction with a floating virtual screen. In Proceedings of the ACM Annual Symposium on User Interface Software and Technology, 663–667, ACM. doi:10.1145/2642918.2647407

Morris, M. R. (2012). Web on the wall: insights from a multimodal interaction elicitation study. In Proceedings of the International Conference on Interactive Tabletops and Surfaces, 95–104, ACM. doi:10.1145/2396636.2396651

Morris, M. R., Danielescu, A., Drucker, S., Fisher, D., Lee, B., Schraefel, M. C, & Wobbrock, J. O. (2014). Reducing legacy bias in gesture elicitation studies. *Interactions*, *21*(3), 40–45. doi:10.1145/2591689

Morrison, C., Huckvale, K., Corish, B., Dorn, J., Kontschieder, P., O’Hara, K., ... & Sellen, A. (2016). Assessing multiple sclerosis with Kinect: designing computer vision systems for real-world use. *Human–Computer Interaction*, 1–36. doi:10.1080/07370024.2015.1093421

Mullaney, T., Yttergren, B., & Stolterman, E. (2014). Positional acts: using a Kinect™ sensor to reconfigure patient roles within radiotherapy treatment. In Proceedings of the International Conference on Tangible, Embedded and Embodied Interaction, 93–96, ACM. doi:10.1145/2540930.2540943

Nancel, M., Wagner, J., Pietriga, E., Chapuis, O., & Mackay, W. (2011). Mid-air pan-and-zoom on wall-sized displays. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 177–186, ACM. doi:10.1145/1978942.1978969

Nancel, M., Chapuis, O., Pietriga, E., Yang, X. D., Irani, P. P., & Beaudouin-Lafon, M. (2013). High-precision pointing on large wall displays using small handheld devices. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 831–840, ACM. doi:10.1145/2470654.2470773

Nansen, B., Vetere, F., Robertson, T., Downs, J., Brereton, M., & Durick, J. (2014). Reciprocal habituation: a study of older people and the Kinect. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *21*(3), 18. doi:10.1145/2617573

Nebeling, M., Huber, A., Ott, D., & Norrie, M. C. (2014). Web on the Wall Reloaded: Implementation, Replication and Refinement of User-Defined Interaction Sets. In Proceedings of the International Conference on Interactive Tabletops and Surfaces, 15–24, ACM. doi:10.1145/2669485.2669497

Newell, A., & Card, S. K. (1985). The prospects for psychological science in human-computer interaction. *Human-Computer Interaction*, *1*(3), 209–242. doi:10.1207/s15327051hci0103_1

Ni, T., Schmidt, G. S., Staadt, O. G., Livingston, M. A., Ball, R., & May, R. (2006). A survey of large high-resolution display technologies, techniques, and applications. In Virtual Reality Conference, 223–236, IEEE. doi:10.1109/VR.2006.20

Ni, T., McMahan, R. P., & Bowman, D. A. (2008). Tech-note: rapMenu: Remote menu selection using freehand gestural input. In IEEE Symposium on 3D User Interfaces, 55–58, IEEE. doi:10.1109/3DUI.2008.4476592

Ni, T., Bowman, D. A., North, C., & McMahan, R. P. (2011). Design and evaluation of freehand menu selection interfaces using tilt and pinch gestures. *International Journal of Human-Computer Studies*, *69*(9), 551–562. doi:10.1016/j.ijhcs.2011.05.001

Norman, D. A. (1986). *Cognitive engineering. User centered system design: New perspectives on human-computer interaction*, 3161.

Norman, D. A. (1988). *The psychology of everyday things*. Basic books.

Norman, D. A. (2010). Natural user interfaces are not natural. *Interactions*, *17*(3), 6–10. doi:10.1145/1744161.1744163

Notebaert, W., Houtman, F., Opstal, F. V., Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: an orienting account. *Cognition*, *111*(2), 275–279.

Oakley, I., McGee, M. R., et al. (2000). Putting the feel in ‘look and feel’. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 415–422, ACM. doi:10.1145/332040.332467

Oehlberg, L., Simm, K., Jones, J., Agogino, A., & Hartmann, B. (2012). Showing is sharing: building shared understanding in human-centered design teams with Dazzle. In Proceedings of the Designing Interactive Systems Conference, 669–678, ACM. doi:10.1145/2317956.2318057

O'hara, K., Kjeldskov, J., & Paay, J. (2011). Blended interaction spaces for distributed team collaboration. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *18*(1), 3. doi:10.1145/1959022.1959025

O’Hara, K., Harper, R., Mentis, H., Sellen, A., & Taylor, A. (2013). On the naturalness of touchless: Putting the “interaction” back into NUI. ACM Transactions on Computer Human Interaction (TOCHI), 20(1), 5:1–5:25. doi:10.1145/2442106.2442111

O'Hara, K., Gonzalez, G., Sellen, A., Penney, G., Varnavas, A., Mentis, H., ... & Carrell, T. (2014). Touchless interaction in surgery. *Communications of the ACM*, *57*(1), 70–77. doi:10.1145/2541883.2541899

O’Hara, K., Gonzalez, G., Penney, G., Sellen, A., Corish, R., Mentis, H., ... & Carrell, T. (2014). Interactional order and constructed ways of seeing with touchless imaging systems in surgery. *Computer Supported Cooperative Work (CSCW)*, 23(3), 299-337. doi:10.1007/s10606-014-9203-4

Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive psychology*, *9*(4), 441–474. doi:10.1016/0010-0285(77)90016-0

Pasquero, J., Stobbe, S. J., and Stonehouse, N. A haptic wristwatch for eyes-free interactions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 3257–3266, ACM. doi:10.1145/1978942.1979425

Pedersen, E. R., McCall, K., Moran, T. P., & Halasz, F. G. (1993). Tivoli: An electronic whiteboard for informal workgroup meetings. In Proceedings of the INTERACT and CHI conference on Human factors in Computing Systems, 391–398, ACM. doi:10.1145/169059.169309

Peirce, C. S. (1931-58): *Collected Writings* (8 Vols.). Charles Hartshorne, Paul Weiss & Arthur W Burks (Ed.). Cambridge, MA: Harvard University Press.

Pook, S., Lecolinet, E., Vaysseix, G., et al. (2000). Control menus: execution and control in a single interactor. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems EA, 263–264, ACM. doi:10.1145/633292.633446

Potter, L. E., Araullo, J., & Carter, L. (2013). The leap motion controller: a view on sign language. In Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration, 175–178. ACM. doi:10.1145/2541016.2541072

PrimeSense Labs. (2013). http://openni.org/files/grab-detector/.

Pyryeskin, D., Hancock, M., & Hoey, J. (2012). Comparing elicited gestures to designer-created gestures for selection above a multitouch surface. In Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces,1–10, ACM. doi:10.1145/2396636.2396638

Rateau, H., Grisoni, L., & De Araujo, B. (2014). Mimetic interaction spaces: Controlling distant displays in pervasive environments. In Proceedings of the 19th International Conference on Intelligent User Interfaces, 89–94, ACM. doi:10.1145/2557500.2557545

Ren, G., & O'Neill, E. (2012). 3D marking menu selection with freehand gestures. In IEEE Symposium on 3D User Interfaces (3DUI), 61–68, IEEE. doi:10.1109/3DUI.2012.6184185

Richardson, I. (2007). Pocket technospaces: the bodily incorporation of mobile media. *Continuum: Journal of Media & Cultural Studies*, *21*(2), 205–215. doi:10.1080/10304310701269057

Richter, H., Loehmann, S., Weinhart, F., & Butz, A. (2012). Comparing direct and remote tactile feedback on interactive surfaces. In Haptics: Perception, Devices, Mobility, and Communication, Springer Berlin Heidelberg, 301–313. doi:10.1007/978-3-642-31401-8_28

Roberts, J. C., Ritsos, P. D., Badam, S. K., Brodbeck, D., Kennedy, J., & Elmqvist, N. (2014). Visualization beyond the desktop--the next big thing. *Computer Graphics and Applications*, IEEE, *34*(6), 26–34. doi:10.1109/MCG.2014.82

Robertson, G., Czerwinski, M., Baudisch, P., Meyers, B., Robbins, D., Smith, G., & Tan, D. (2005). The large-display user experience. *IEEE Computer Graphics and Applications*, *25*(4), 44–51. doi:10.1109/MCG.2005.88

Robles-De-La-Torre, G., & Hayward, V. (2001). Force can overcome object geometry in the perception of shape through active touch. *Nature*, *412*(6845), 445–448. doi:10.1038/35086588

Rosa, G. M., & Elizondo, M. L. (2014). Use of a gesture user interface as a touchless image navigation system in dental surgery: Case series report. *Imaging science in dentistry, 44*(2), 155–160. doi:10.5624/isd.2014.44.2.155

Rovelo Ruiz, G. A., Vanacken, D., Luyten, K., Abad, F., & Camahort, E. (2014). Multi-viewer gesture-based interaction for omni-directional video. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 4077–4086, ACM. doi:10.1145/2556288.2557113

Ruppert, G. C. S., Reis, L. O., Amorim, P. H. J., de Moraes, T. F., & da Silva, J. V. L. (2012). Touchless gesture user interface for interactive image visualization in urological surgery. *World journal of urology*, *30*(5), 687–691. doi:10.1007/s00345-012-0879-0

Saunders, J. A., & Knill, D. C. (2004). Visual feedback control of hand movements. *The Journal of neuroscience*, *24*(13), 3223–3234. doi:10.1523/JNEUROSCI.4319-03.2004

Scheidt, R. A., Conditt, M. A., Secco, E. L., & Mussa-Ivaldi, F. A. (2005). Interaction of visual and proprioceptive feedback during adaptation of human reaching movements. *Journal of Neurophysiology*, *93*(6), 3200–3213. doi:10.1152/jn.00947.2004

Schofield, W. N. (1976). Hand movements which cross the body midline: Findings relating age differences to handedness. *Perceptual and Motor Skills*, *42*(2), 643–646. doi:10.2466/pms.1976.42.2.643

Schwarz, L. A., Bigdelou, A., & Navab, N. (2011). Learning gestures for customizable human-computer interaction in the operating room. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011, 129–136, Springer Berlin Heidelberg. doi:10.1007/978-3-642-23623-5_17

Schwaller, M., Brunner, S., & Lalanne, D. (2013). Two handed mid-air gestural HCI: Point + command. In Human-Computer Interaction. Interaction Modalities and Techniques, 388–397, Springer Berlin Heidelberg. doi:10.1007/978-3-642-39330-3_41

Seixas, M., Cardoso, J., & Dias, M. T. G. (2015). One Hand or Two Hands? 2D Selection Tasks with the Leap Motion Device. Retrieved from http://hdl.handle.net/10400.14/17159 on May 6, 2016.

Shadmehr, R., Smith, M. A., & Krakauer, J. W. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annual review of neuroscience*, *33*, 89–108. doi:10.1146/annurev-neuro-060909-153135

Shoemaker, G., Tang, A., & Booth, K. S. (2007). Shadow reaching: a new perspective on interaction for large displays. In Proceedings of the Annual ACM Symposium on User Interface Software and Technology, 53–56, ACM. doi:10.1145/1294211.1294221

Shoemaker, G., Tsukitani, T., Kitamura, Y., & Booth, K. S. (2012). Two-Part Models Capture the Impact of Gain on Pointing Performance. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *19*(4), 28. doi:10.1145/2395131.2395135

Shotton, J., Sharp, T., Kipman, A., Fitzgibbon, A., Finocchio, M., Blake, A., ... & Moore, R. (2013). Real-time human pose recognition in parts from single depth images. *Communications of the ACM*, *56*(1), 116–124. doi:10.1145/2398356.2398381

Sigrist, R., Rauter, G., Riener, R., & Wolf, P. (2013). Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. *Psychonomic bulletin & review*, *20*(1), 21–53. doi:10.3758/s13423-012-0333-8

Smart TV: Smart Interaction. (2013). http:// reviews.cnet.com/8301-33199_7-57411497-221/samsung-smart-interaction-hands-on-with-voice-and-gesture-control/

Smith, S. L., & Thomas, D. W. (1964). Color versus shape coding in information displays. *Journal of Applied Psychology*, *48*(3), 137. doi:10.1037/h0045919

Sodhi, R., Poupyrev, I., Glisson, M., & Israr, A. (2013). AIREAL: Interactive tactile experiences in free air. *ACM Transactions on Graphics (TOG)*, *32*(4), 134. doi:10.1145/2461912.2462007

Song, P., Goh, W. B., Hutama, W., Fu, C. W., & Liu, X. (2012). A handle bar metaphor for virtual object manipulation with mid-air interaction. In Proceedings of the

SIGCHI Conference on Human Factors in Computing Systems, 1297–1306, ACM. doi:10.1145/2207676.2208585

Sridhar, S., Feit, A. M., Theobalt, C., & Oulasvirta, A. (2015). Investigating the dexterity of multi-finger input for mid-air text entry. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 3643–3652, ACM. doi:10.1145/2702123.2702136

Stokes, D. E. (1997). *Pasteur's quadrant: Basic science and technological innovation*. Brookings Institution Press.

Suchman, L. A. (1987). *Plans and situated actions: the problem of human-machine communication*. Cambridge university press.

Sülzenbrück, S. (2012). The impact of visual feedback type on the mastery of visuo-motor transformations. *Zeitschrift für Psychologie*, *220*(1), 3. doi:10.1027/2151-2604/a000084

Sung, J., Ponce, C., Selman, B., & Saxena, A. (2012). Unstructured human activity detection from rgbd images. In 2012 IEEE International Conference on Robotics and Automation (ICRA), 842–849, IEEE. doi:10.1109/ICRA.2012.6224591

Swaminathan, K., & Sato, S. (1997). Interaction design for large displays. *Interactions*, *4*(1), 15-24. doi:10.1145/242388.242395

Taele, P., & Hammond, T. (2014). Developing sketch recognition and interaction techniques for intelligent surfaceless sketching user interfaces. In Proceedings of the companion publication of Intelligent User Interfaces Conference, 53–56, ACM. doi:10.1145/2559184.2559185

Tan, D. S., Gergle, D., Scupelli, P., & Pausch, R. (2003). With similar visual angles, larger displays improve spatial performance. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 217–224, ACM. doi:10.1145/642611.642650

Tan, D. S., Gergle, D., Scupelli, P., & Pausch, R. (2006). Physically large displays improve performance on spatial tasks. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *13*(1), 71–99. doi:10.1145/1143518.1143521

Tan, J. H., Chao, C., Zawaideh, M., Roberts, A. C., & Kinney, T. B. (2013). Informatics in radiology: Developing a touchless user interface for intraoperative image control during interventional radiology procedures. *Radiographics*, *33*(2), E61–E70. doi:10.1148/rg.332125101

Teixeira, L. A. (2000). Timing and force components in bilateral transfer of learning. *Brain and Cognition*, *44*(3), 455–469. doi:10.1006/brcg.1999.1205

Todor, J. I., & Doane, T. (1978). Handedness and hemispheric asymmetry in the control of movements. *Journal of Motor Behavior*, *10*(4), 295–300. doi:10.1080/00222895.1978.10735163

Valkanova, N., Walter, R., Vande Moere, A., & Müller, J. (2014). Myposition: Sparking civic discourse by a public interactive poll visualization. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1323–1332, ACM. doi:10.1145/2531602.2531639

Van Mensvoort, I. (2002). What you see is what you feel: Exploiting the dominance of the visual over the haptic domain to simulate force-feedback with cursor displacements. In Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques, 345–348, ACM. doi:10.1145/778712.778761

Van Mensvoort, I. (2008). PowerCursor. http://www.powercursor.com

Vatavu, R. D., & Pentiuc, S. G. (2008). Multi-level representation of gesture as command for human computer interaction. *Computing and Informatics*, *27*(6), 837–851. Retrieved from http://www.cai.sk/ojs/index.php/cai/article/view/16/3 on May 6, 2016.

Vatavu, R. D., & Zaiti, I. A. (2014). Leap gestures for TV: insights from an elicitation study. In Proceedings of the 2014 ACM International Conference on Interactive Experiences for TV and Online Video, 131–138, ACM. doi:10.1145/2602299.2602316

Vatavu, R. D., & Wobbrock, J. O. (2015). Formalizing agreement analysis for elicitation studies: new measures, significance test, and toolkit. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 1325–1334, ACM. doi:10.1145/2702123.2702223

Verwey, W. B., & Heuer, H. (2007). Nonlinear visuomotor transformations: locus and modularity. *The Quarterly Journal of Experimental Psychology*, *60*(12), 1629–1659. doi:10.1080/17470210601100472

Vogel, D., & Balakrishnan, R. (2005). Distant freehand pointing and clicking on very large, high resolution displays. In Proceedings of the ACM Annual Symposium on User Interface Software and Technology, 33–42, ACM. doi:10.1145/1095034.1095041

Vogt, K., Bradel, L., Andrews, C., North, C., Endert, A., & Hutchings, D. (2011). Co-located collaborative sensemaking on a large high-resolution display with multiple input devices. In Proceedings of the INTERACT Conference, 589–604, Springer. doi:10.1007/978-3-642-23771-3_44

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological Bulletin*, *138*(6), 1172. doi:10.1037/a0029333

Wagemans, J., Feldman, J., Gepshtein, S., Kimchi, R., Pomerantz, J. R., van der Helm, P. A., & van Leeuwen, C. (2012). A century of Gestalt psychology in visual perception: II. Conceptual and theoretical foundations. *Psychological Bulletin*, *138*(6), 1218. doi:10.1037/a0029334

Walter, R., Bailly, G., & Müller, J. (2013). Strikeapose: revealing mid-air gestures on public displays. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 841–850, ACM. doi:10.1145/2470654.2470774

Walter, R., Bailly, G., Valkanova, N., & Müller, J. (2014). Cuenesics: using mid-air gestures to select items on interactive public displays. In Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices & Services, 299–308, ACM. doi:10.1145/2628363.2628368

Weiser, M. (1993). Some computer science issues in ubiquitous computing. *Communications of the ACM*, *36*(7), 75–84. doi:10.1145/159544.159617

Werner, R., Armstrong, T. J., Bir, C., & Aylard, M. K. (1997). Intracarpal canal pressures: the role of finger, hand, wrist and forearm position. Clinical Biomechanics, 12(1), 44–51. doi:10.1016/S0268-0033(96)00044-7

Wertheimer, M. (1912). *Experimentelle studien über das sehen von bewegung*. JA Barth.

Wertheimer, M., & Riezler, K. (1944). Gestalt theory. *Social Research*, *11*, 78–99.

Wigdor, D., Williams, S., Cronin, M., Levy, R., White, K., Mazeev, M., & Benko, H. (2009). Ripples: utilizing per-contact visualizations to improve user interaction with touch displays. In Proceedings of the ACM Annual Symposium on User Interface Software and Technology, 3–12, ACM. doi:10.1145/1622176.1622180

Wigdor, D. (2010). Architecting next-generation user interfaces. In Proceedings of the International Working Conference on Advanced Visual Interfaces, 16–22, ACM. doi:10.1145/1842993.1842997

Wigdor, D., & Wixon, D. (2011). *Brave NUI world: Designing natural user interfaces for touch and gesture*. Burlington, MA: Morgan Kauffman.

Window shopping, virtually. 2011. Retrieved from http://phys.org/news/2011-01-window-virtually.html on May 6, 2016

Winograd, T., & Flores, F. (1986). *Understanding computers and cognition: A new foundation for design*. Intellect Books.

Wright, C. E., & Lee, F. (2013). Issues related to HCI application of Fitts's law. *Human–Computer Interaction*, *28*(6), 548–578. doi:10.1080/07370024.2013.803873

Wobbrock, J. O., Morris, M. R., & Wilson, A. D. (2009). User-defined gestures for surface computing. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1083–1092, ACM. doi:10.1145/1518701.1518866

Wolfe, J. M. (1998). Visual search. *The Handbook of Attention, 27.*

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, *5*(6), 495–501. doi:10.1038/nrn1411

Wolfe, J. M., Birnkrant, R. S., Kunar, M. A., & Horowitz, T. S. (2005). Visual search for transparency and opacity: attentional guidance by cue combination? *Journal of Vision*, *5*(3), 9. doi:10.1167/5.3.9

Wu, M., Shen, C., Ryall, K., Forlines, C., & Balakrishnan, R. (2006). Gesture registration, relaxation, and reuse for multi-point direct-touch surfaces. In IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TableTop), 183–190, IEEE. doi:10.1109/TABLETOP.2006.19

Yi, J. S., Ah Kang, Y., Stasko, J. T., & Jacko, J. A. (2007). Toward a deeper understanding of the role of interaction in information visualization. *IEEE TVCG*, *13*(6), 1224–1231. doi:10.1109/TVCG.2007.70515

Yun, K., Honorio, J., Chattopadhyay, D., Berg, T. L., & Samaras, D. (2012). Two-person interaction detection using body-pose features and multiple instance learning. In 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 28–35, IEEE. doi:10.1109/CVPRW.2012.6239234

Zelaznik, H. N., Hawkins, B., & Kisselburgh, L. (1983). Rapid visual feedback processing in single-aiming movements. *Journal of Motor Behavior*, *15*(3), 217–236. doi:10.1080/00222895.1983.10735298

Zhao, S., & Balakrishnan, R. (2004,). Simple vs. compound mark hierarchical marking menus. In Proceedings of the 17th annual ACM symposium on User interface software and technology, 33–42, ACM. doi:10.1145/1029632.1029639

Zhou, H., & Hu, H. (2008). Human motion tracking for rehabilitation—A survey. *Biomedical Signal Processing and Control*, *3*(1), 1–18. doi:10.1016/j.bspc.2007.09.001

CURRICULUM VITAE

Debaleena Chattopadhyay

Education

Ph.D. in Informatics (Human-Computer Interaction track)	2016
Human-Centered Computing Department, School of Informatics and Computing	
Indiana University, Indianapolis, Indiana, USA	
M.S. in Computer Science	2011
Computer Science Department	
Stony Brook University, Stony Brook, New York, USA	
B.Tech. in Computer Science & Engineering	2009
Department of Computer Science and Engineering	
West Bengal University of Technology, Kolkata, West Bengal, India	

Honors, Awards, and Fellowships

Indiana University Graduate School IUPUI Chancellor's Scholar	2016
CHI 2016 Late-Breaking Work (LBW) Best Paper Honorable Mention	2016
Best graduate student in the IU School of Informatics and Computing	2015
Premiere 10 Award, IUPUI	2015
Elite 50 Award, IUPUI	2015
IUPUI Graduate Office Travel Fellowship Award x 2	2015
NSF Travel Grant, TEI 2015 Doctoral Consortium	2015
Microsoft Travel Award, ACM SRC, GHC	2013
Xerox-Foundation Scholarship, GHC	2013
Research Support Funds Grant, IUPUI	2013
IUPUI Fellowship	2012
IUPUI SoIC Travel Award, ACM TAPIA Conference	2012
Carnegie Mellon University Honorarium, Art && Code Conference	2011
ACM scholarship, CRA-W Graduate Cohort Workshop	2011
Computer Science Chair Fellowship, Stony Brook University	2009

Publications

[1] Chattopadhyay, D. & MacDorman, F., K. Familiar Faces Rendered Strange: Why Inconsistent Realism Drives Characters into the Uncanny Valley. *Journal of Vision*. Forthcoming.

[2] Chattopadhyay, D., O'Hara, K., Rintel, S., & Rädle, R. (2016). Office Social: Presentation Interactivity for Nearby Devices. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI, 2487–2491, ACM.

[3] Chattopadhyay, D., Duke, J., D., & Bolchini, D. (2016). Endorsement, Prior Action, and Language: Modeling Trusted Advice in Computerized Clinical Alerts. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI, 2027–2033, ACM.

[4] MacDorman, K., F., & Chattopadhyay, D. (2016). Inconsistency in Human Realism, and not Category Uncertainty, elicits the Uncanny Valley Effect. *Cognition*, *146*, 190–205.

[5] Chattopadhyay, D., Rohani Ghahari, R., Duke, J., D., & Bolchini, D. (2015). Understanding Advice Sharing among Physicians: Towards Trust-Based Clinical Alerts. *Interacting with Computers*. Forthcoming.

[6] Chattopadhyay, D., & Bolchini, D. (2015). Motor-Intuitive Interactions Based on Image Schemas: Aligning Touchless Interaction Primitives with Human Sensorimotor Abilities. Special Issue on Intuitive Interactions, *Interacting with Computers*, *27*(3), 327–343.

[7] Chattopadhyay, D. (2015). Toward Motor-Intuitive Interaction Primitives for Touchless Interfaces. In Proceedings of the Tenth International Conference on Interactive Tabletops and Surfaces, ITS, 445–450, ACM.

[8] Chattopadhyay, D. (2015). Exploring Perceptual and Motor Gestalt in Touchless Interactions with Distant Displays. In Proceedings of the Ninth International Conference on Tangible, Embedded and Embodied Interaction, TEI, 433–436, ACM.

[9] Chattopadhyay, D., & Bolchini, D. (2014). Touchless Circular Menus: Toward an Intuitive UI for Touchless Interactions with Large Displays. In Proceedings of the International Working Conference on Advanced Visual Interfaces, AVI, 33–40, ACM.

[10] Chattopadhyay, D., Achmiz, S., Saxena, S., Bansal, M., Bolchini, D., & Voida, S. (2014). Holes, Pits, and Valleys: Guiding Large-Display Touchless Interactions with Data-Morphed Topographies. Ext. Abstracts, UbiComp, 19–22, ACM.

[11] Chattopadhyay, D., Pan, W., & Bolchini, D. (2013). A `Stopper' Metaphor for Persistent Visual Feedback in Touchless Interactions with Wall-Sized Displays. International Symposium on Pervasive Displays, PerDis, Mountain View, California, USA.

[12] Chattopadhyay, D., & Bolchini, D. (2013). Laid-Back, Touchless Collaboration around Wall-size Displays: Visual Feedback and Affordances. Position paper at the International Workshop on Interactive, Ultra-High-Resolution Displays (POWERWALL), CHI, Paris, France.

[13] Yun, K., Carrillo, J., H., Chattopadhyay, D., Berg, T., L., & Samaras, D. (2012). Two-person Interaction Detection using Body-Pose Features and Multiple Instance Learning. In Proceedings of Computer Vision and Pattern Recognition Workshops, CVPR, 28–35, IEEE.

[14] Berg, T., L., Chattopadhyay, D., Schedel, M., & Vallier, T. (2012). Interactive Music: Human Motion Initiated Music Generation using Skeletal Tracking by Kinect. In Proceedings of Society for Electro-Acoustic Music in the United States, SEAMUS, Wisconsin, USA.

[15] Bhowmick, B., & Chattopadhyay, D. (2009). Shot Boundary Detection Using Texture Feature Based on Co-occurrence Matrices. In Proceedings of International Conference on Multimedia, Signal Processing and Communication Technologies, IMPACT, 165–168, IEEE.

Patents

Content Navigation Control 2016

US Provisional Patent, (date filed January 2016), Co-inventors: Kenton O'Hara, Gavin Smyth, Sean Rintel, and Debaleena Chattopadhyay.

Shot Boundary Detection Based on Co-occurrence Matrices 2009

Government of India Provisional Patent Application No. 2124/MUM/2008 A, Date filed: 03/10/2008, Date published: 30/07/2010. Co-inventors: Debaleena Chattopadhyay and Brojeshwar Bhowmick.

Professional Experience

Research Intern (Supervisor: Dr. Kenton O'Hara) — Summer, 2015
Microsoft Research, Cambridge, UK

Instructor — Spring, 2015
Department of Human-Centered Computing
Indiana University School of Informatics and Computing, IUPUI

Co-Instructor — Spring, 2014
Department of Human-Centered Computing
Indiana University School of Informatics and Computing, IUPUI

Research Assistant (Supervisor: Dr. Davide Bolchini) — 2012–2016
Department of Human-Centered Computing
Indiana University School of Informatics and Computing, IUPUI

Teaching Assistant — 2013–2014
Indiana University School of Informatics, IUPUI

Research Assistant (Supervisor: Dr. Amy S. Lu) — 2011–2012
Indiana University School of Informatics, IUPUI

Research Assistant (Supervisor: Dr. Tamara L. Berg) — 2010
Computer Science Department, Stony Brook University, New York

Teaching Assistant — 2009–2010
Computer Science Department, Stony Brook University, New York

Research Intern (Supervisor: Aniruddha Sinha) — Summer, 2008
Innovation Lab, Tata Consultancy Service Ltd., Kolkata, India

Teaching

Indiana University School of Informatics and Computing, IUPUI

- Instructor, Introduction to Informatics — Spring, 2015
- Co-Instructor, Introduction to Informatics — Spring, 2014
- Teaching Assistant, User Experience Architectures — Summer, 2014, 2015
- Teaching Assistant, Psychology of HCI — Fall, 2013
- Teaching Assistant, Introduction to Research in Informatics — Spring, 2012
- Teaching Assistant, Serious Games — Fall, 2011
- Teaching Assistant, Psychology of Media — Fall, 2011

Computer Science Department, Stony Brook University, New York

- Teaching Assistant, Introduction to Programming — Spring, 2010

Teaching Assistant, Computer Science I — Spring, 2010
Teaching Assistant, Introduction to Computer Science — Fall, 2009

Guest Lectures

Informatics Research Design, Empirical Research in HCI — November, 2015
Seminar in Health Informatics-I — July, 2013
Serious Games, Introduction to Behavioral Theories — October, 2011
Psychology of Media, Introduction to Persuasion Theories — October, 2011

Service

Peer review

ASSETS 2016: ACM SIGACCESS Conference on Computers and Accessibility
AVI 2016: International Working Conference on Advanced Visual Interfaces
CHI 2016 Late Breaking Work: ACM Conference on Human Factors in Computing Systems
INTERACT 2015: 15th IFIP TC.13 International Conference on Human-Computer Interaction
IUI 2015: ACM International Conference on Intelligent User Interfaces
CHI 2015: ACM Conference on Human Factors in Computing Systems
UbiComp 2014: ACM International Joint Conference on Pervasive and Ubiquitous Computing
NordiCHI 2014: Nordic Conference on Human-Computer Interaction
ENTER 2013: eTourism Conference
ACM Multimedia, 2010
The Visual Computer
Interacting with Computers
Cognition
Frontiers in Psychology
International Journal of Social Robotics
PLoS ONE

Administrative experience

Indiana University-Purdue University, Indianapolis, USA

Chair, ACM-W Chapter — 2013–2015
Graduate Vice-president — 2012–2014

Women in Technology (WiT) student organization

School of Informatics and Computing, Indiana University, Indianapolis, USA

Human-Centered Computing Tenure Track Search Committee 2014–2015

Informatics Student Government (ISG) 2012–2013

Stony Brook University, Stony Brook, New York, USA

Secretary, ACM-W Chapter, Women in Computer Science (WiCS) 2009–2011

Juror, Hearing Board of the Academic Judiciary, CS Department 2010–2011