THE HUMAN BODY AS AN INTERACTIVE COMPUTING PLATFORM

CHRIS HARRISON
CMU-HCII-13-101
December 2013

Human-Computer Interaction Institute
School of Computer Science
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

chris.harrison@cs.cmu.edu
http://www.chrisharrison.net

COMMITTEE
Scott Hudson (Chair), Carnegie Mellon University
Jodi Forlizzi, Carnegie Mellon University
Anind Dey, Carnegie Mellon University
Desney Tan, Microsoft Research

Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

This work was supported by Disney Research, Intel Research Council, General Motors, Center for the Future of Work (Heinz College, CMU), and the National Science Foundation under Grants IIS-0713509, IIS-0803733, IIS-0840766, and IIS-1217929, as well as by fellowships from Microsoft Research, Google, and Qualcomm.

KEYWORDS

Human-Computer Interaction, User Interfaces, Mobile Devices, Input Techniques, Touchscreens, Gestures, Sensors, Bio-Acoustics, Computer Vision, Ubiquitous Computing Appropriated Surfaces, On-Body Computing.

ABSTRACT

Despite their small size, mobile devices are able to perform tasks of creation, information and communication with unprecedented ease. However, diminutive screens and buttons mar the user experience, and otherwise prevent us from realizing the full potential of computing on the go. In this dissertation, I will first discuss strategies I have pursued to expand and enrich interaction. For example, fingers have many "modes" – they do not just poke, as contemporary touchscreen interaction would suggest, but also scratch, flick, knock, rub, and grasp, to name a few. I will then highlight an emergent shift in computing: from mobile devices we carry to using everyday surfaces for interaction, including tables, walls, furniture and even our skin, bringing computational power ever closer to users. This evolution brings significant new challenges in sensing and interaction design. For example, the human body is not only incredibly irregular and dynamic, but also comes in more than six billion different models. However, along with these challenges also come exciting new opportunities for more powerful, intuitive and intimate computing experiences.

ACKNOWLEDGEMENTS

I owe my academic success to several influential individuals who steered me towards science and engineering at an early age, and later in life, towards Human-Computer Interaction. I am particularly grateful to Gene and Karen Brewer, who put up with my endless questions, fed my curiosity, and instilled many of my better values. As a teenager, Ernest Vesce and John DuBuque gave me the benefit of the doubt and real jobs. The professional skills I picked up under their watch profoundly impacted how I pursue my research. At the same time, Nancy Williams and Wilfredo Chaluisant opened my eyes to research and gave me the confidence to pursue it. During my time at New York University, Dennis Shasha and Brian Amento showed me how to unite my interests in design and computer science, and convinced me to get a Ph.D. in Human-Computer Interaction.

My advisor, Scott E. Hudson, has also been instrumental in my success. I thank him for his six years of support, mentorship and friendship. From day one, he has given me the freedom – both intellectually and financially – to pursue research that I was passionate about. With his guiding hand, I grew from a researcher into a proper scholar, and now, a mentor for my future students. I must also acknowledge two unofficial advisors who have also had a significant impact on my thought process, research trajectory and skill set: Ivan Poupyrev and Desney Tan. Thank you!

Many organizations have helped to fund my research projects and graduate education, for which I am humbled and grateful. These include The National Science Foundation, Microsoft Research, Google, Qualcomm, and Disney Research.

I am also indebted to many co-authors and colleagues, including Tom Bartindale, Olivier Bau, Hrvoje Benko, Anind Dey, Haakon Faste, Jodi Forlizzi, Gary Hsieh, Ali Israr, Brian Lim, Peter Kinnaird, Dan Morris, Stephen Oney, Scott Saponas, Munehiko Sato, Aubrey Shick, Julia Schwarz, Jason Wiese, Karl D.D. Willis, Andy Wilson, Svetlana Yarosh, Zhiquan Yeo and Robert Xiao.

Finally, I am forever grateful to my parents, Terri and Terry, and my brother Ben, who have always rooted for me. Enumerating a list of lessons learned, stories shared and values instilled would easily exceed the length of this dissertation. And then there is Amy, the love of my life, who is my ardent champion and toughest critic. She keeps me in check and also on my toes, where I am most nimble.

TABLE OF CONTENTS

1 INTRODUCTION

1.1 Overview

Computing has evolved repeatedly and dramatically in its short history. In the 1980's, the mainframe era transitioned to a focus on desktop computers. The latter brought computational power much closer to the user, enabling a high level of customization and computational freedom, and sparked the *personal* computer revolution. Today, mobile computers have moved to the forefront, bringing computation ever closer to the user – into our pockets and bags. Despite their diminutive size, they are able to perform tasks of creation, information and communication with unprecedented ease. It is undeniable that they have forever changed the way we work, learn, and play. However, mobile interaction is far from solved. Diminutive screens and buttons mar the user experience, and otherwise prevent us from realizing the full potential of mobile computing.

In this dissertation, I highlight and explore an emergent shift in computing: from mobile devices we carry to using the human body itself as an interactive platform. This brings computational power ever closer to users - out of pockets and onto the skin (Figure 1.1). This evolution brings significant new challenges in sensing and interaction design. Not only is the human form incredibly dynamic and irregular, but also comes in more than six billion different models. Moreover, unlike all other computing platforms, we have no control over the form – we can augment the body in very careful ways, but not modify it. However, along with these challenges also comes exciting new opportunities for more powerful, intuitive and intimate computing experiences.

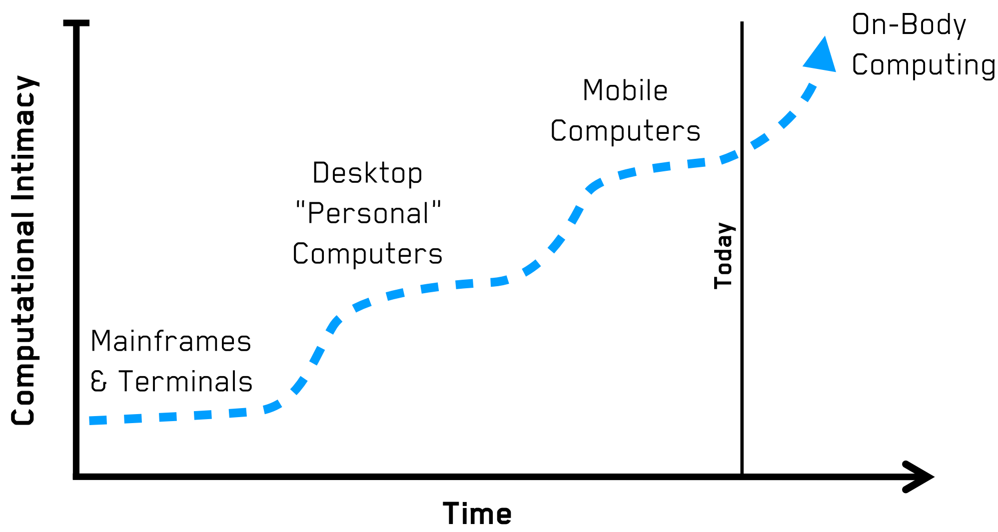


Fig. 1.1 **As computing has evolved, there has been a steady increase in the "computational intimacy" - where computing potential moves closer to the user, and along with it, improved customizability, security and reliability. Drawing this trend out, it is not hard to imagine that soon we will transition beyond today's laptops, tablets and smartphones, and move computation one step closer to the user, and onto the body itself.**

1.2 Organization

For the rest of this Chapter, I discuss where mobile computing has been successful, along with its key limitations. I suggest there are two important avenues for alleviating the mobile input/output bottleneck: increasing interactive surface area and improving input richness. To elucidate my research progression, I provide a brief overview of my research pursing these two approaches. These efforts help to ground my dissertation work, as they provided the necessary perspective to appreciate the human body as a compelling computing platform.

In Chapter 2, I summarize research from disparate, but related domains, including touch input, interaction techniques, biological sensing, brain computer interfaces, hand gestures, body tracking, ad hoc input surfaces, computer vision, wearable computers, and on-body interfaces. In Chapters 3, 4 and 5, I discuss three on-body interactive systems I developed: Skinput, OmniTouch and Touché. Each employs and explores a different sensing approach – acoustic, optical and electrical respectively. In Chapters 6, 7 and 8, I set aside technical issues and describe three additional projects that consider questions relating to on-body design, including interaction design, visual accessibility, and social appropriateness of touch. Finally, in Chapter 9, I conclude with a summary of key contributions, exciting avenues of future work, and some final thoughts.

1.3 On Being Small

The fundamental usability issue with mobile devices is apparent to anyone who has used one: they are small. Achieving mobility through miniaturization has been both their greatest success and most significant shortcoming. Because we have yet to figure out a good way to miniaturize devices without simultaneously shrinking their interactive surface area, mobile computing typically implies diminutive screens, cramped keyboards, tiny jog wheels and similar - all of which diminish usability and prevent us from realizing the full potential of computing on the go. Although input is an outstanding challenge across all forms of computing, the problem is particularly acute for mobile interaction. In particular, being small has two significant implications, both of which have proven difficult to overcome.

Foremost, a small form generally means there is a limited area for graphical output, by far the most predominant and high-bandwidth means of computer-to-human communication. This significantly reduced computer-human bandwidth instantly contracts the possible application space.

Secondly, surface area for direct manipulation user input is equally constrained on small devices. Compounding this problem is that our fingers are "fat" compared to pixels [Siek 2005; Holz 2010], yielding an inescapable lower bound. Buttons simply cannot be made smaller, not because of limitations in sensing or display resolution, but because users would not be able to accurately press them. Moreover, humans are not particularly dexterous at such small scales, and the trend towards touchscreens has removed many of the physical affordances that support fine-grained manipulation tasks, further exacerbating the problem.

For example, it is now standard for touchscreen keyboards to feature real-time spelling correction, simply because it is assumed that users are unable to accurately hit such small targets. Other interactors that cannot benefit from word and language models are inevitably larger. Menus, ribbons, toolboxes, and similar that are commonplace in desktop-class applications must be shrunken down (making them harder to press), tucked away in menu hierarchies (requiring more presses), or simply eliminated (mobile applications often have reduced functionality compared to their desktop counterparts). In general, the mobile input bottleneck dramatically influences how mobile device interfaces are designed and more importantly, what tasks we can perform.

Research in human-computer interaction can often pursue a “time machine” approach, knowing that technology will inevitably advance. What is 1 frame per second today might be 30 frames per second in a year. Unfortunately, a paucity of surface area is a problem that will not solve itself by waiting for technology to advance. While computer processors will get faster, LCD screens thinner, and hard drives larger, added surface area will not come without increased size – it is a physical constraint.

This has trapped users and designers in a device size paradox: We want bigger, more useable devices – but without losing the primary benefit of small size and mobility. Simultaneously, we want smaller, more mobile devices, but without sacrificing usability. In response, device manufactures have walked a fine line for at least a decade, striking a careful balance between usability and size.

This effect is readily apparent to anyone with a laptop. Dedicated number keypads were shed long ago, and layouts featuring squished arrow and function keys are common. Netbooks have gone so far as to shrink every key to accommodate a full layout. Users do not love small keyboards - quite the opposite, but they accept them, mostly because they would not tolerate a larger device. This is true of smartphones as well - if only we could have a full-sized keyboard and a pocket-sized device.

1.4 From Taps to Riches

So far, I have focused on limitations inherent in mobile computers given their diminutive size. However, there is a second, considerably more subtle issue that is equally significant: a paucity of richness. Assuming for a moment that being small is inescapable, the blow could be dampened if interaction was particularly expressive. Unfortunately, far from making maximal use of devices’ limited surface area, contemporary mobile interfaces use the most simplistic user input dimensions.

For many decades, mobile device interaction meant pressing buttons, jogging wheels, and thumbing joysticks, which were binary or coarse. The evolution to “smart” touch-centric mobile interaction exemplifies an increase in richness. Devices gained new capabilities not by growing in size, but by allowing for more powerful interactions in the same physical space.

However, even touchscreen interaction suffers from a paucity of richness. For example, fingers are typically digitized as single X/Y positions on a touchscreen. Even much lauded multitouch gestures are fairly simplistic, the most popular being a two-finger pinch (i.e., two X/Y positions). Compare this to the desktop computer mouse, which provides X/Y translation, up/down scrolling, and two or more buttons – all in a single hand. This provides considerable input bandwidth, and as a result, there are many things we can do on a desktop computer that are cumbersome on a touchscreen device.

Compounding this issue, contemporary touchscreens treat fingers as a single class of input. As a result, there is nothing immediately analogous to a “right click”, an input paradigm that has proven powerful and popular in desktop computing. The problem with treating fingers as a single class is that there is no modality. Modes can either be provided with buttons (but take valuable screen real estate) or fairly unintuitive chording of fingers and tap-and-hold interactions [Li 2005; Lepinski 2010]. Scaling beyond primary and secondary actions gets increasingly unwieldy (e.g., double-tap-and-hold? Index and pinky finger tap?).

This lack of richness would not be so apparent if it were not for the large disparity between touchscreen input and the true capabilities of our fingers. In addition to translating to an X/Y position, our fingers can vary their angle of attack, bend, twist, and apply different pressure and shear forces (at least six additional analog dimensions). Fingers also have many “modes” – they do not just poke, as today’s touchscreen interactions would suggest, but also pinch, scratch, flick, knock, rub, and grasp, to name a few. Combinatorially speaking, our fingers are capable of forming hundreds of poses [Kendon 1988; Mulder 1996]. For reference, American Sign Language (which includes motion) has several thousand signs [Valii 2006] (Figure 1.2). If these additional dimensions of touch could be digitized, there are tremendous opportunities for enriching touchscreen interaction, potentially even alleviating the lack of screen real estate.

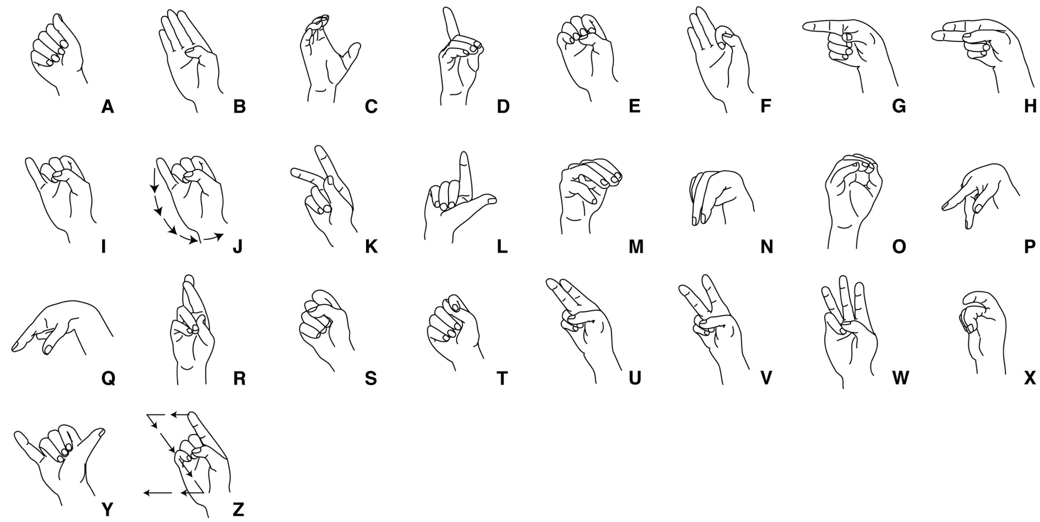


Fig. 1.2 **The American Sign Language alphabet illustrates a small subset of the possible richness of our hands [Valii 2006].**

1.5 Formulating Input Power

As described above, mobile input is constrained by two dimensions. The first is the area available for input – put simply, the number of places we can put our fingers. The more area a device has, the more things our fingers can confortable and accurately target. This dimension can be viewed as the quantity of input space. The second dimension is how rich interactions can be in a given space, for example, the number of different actions one can perform. This dimension can be through of as analogous to the quality of the input space.

The total input power of a device is a combination of quantity and quality. A very tiny, but input rich device might have the same power as a larger, but input poor device. For example, a touchscreen watch and a buttoned cell phone might have roughly equivalent functionality and accessibility despite differences in size. This suggests the following schematic formulation:

*input power = input richness * input area*

Although simple, this formulation has important implications, and elucidates several ways forward. Foremost, it suggests further miniaturization is possible, without losing capability, if we can correspondingly increase richness. Second, increasing input area or richness independently will increase input power, and thus are valuable pursuits individually. And finally, increasing both richness and area will yield multiplicative gains in input power. In other words, a gain in richness yields benefits over all existing surface area, and vice versa.

1.6 Initial Explorations

My dissertation research is the result of an extensive exploration in the area of mobile interaction. These efforts largely fall under the broad categories introduced above: increasing input richness and increasing input area. Although I do not discuss these projects in great detail in this document, they were instrumental in providing the perspective needed to identify the unique benefits on-body interaction affords, which as we will read, is the focus of this dissertation. I now briefly summarize these research efforts (Figure 1.3).

My earliest efforts in improving mobile interaction started with increasing the input richness of devices. Lean and Zoom [Harrison 2009] captured the distance of a user's face from the screen as an additional input dimension. This could be used for example, to adjust the size of content on the screen automatically (in response to human visual acuity constraints). I also developed a novel fiducial marker that allowed optical multitouch surfaces to resolve the order of tangibles in a stack [Bartindale 2009]. Stacking, an action we perform regularly to organize physical objects in the real world, provides an intuitive way to group items and describe ordering, without consuming additional screen real estate. Finally, SurfaceMouse [Bartindale 2011] is a virtual mouse implementation for multitouch screens. In a single hand, users can perform clutched X/Y translation, up/down scrolling, and "click" primary and secondary buttons – a collection of actions that are unwieldy to perform on current touchscreens.

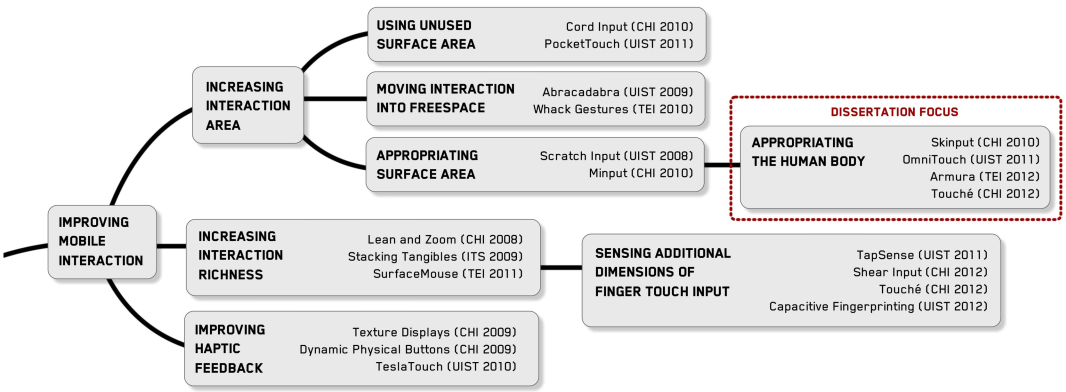


Fig. 1.3 **My explorations in improving mobile interaction.**

I also engaged in a more focused effort to enrich finger-on-touchscreen interactions. As noted earlier, contemporary touchscreens generally simplify finger touches to a 2D coordinate. However, there are many other dimensions of touch input beyond spatial location. For example, through acoustic sensing, TapSense [Harrison 2011] can distinguish among small sets of passive tools as well as discriminate different parts of the finger: pad, tip, knuckle and nail. Shear Input [Harrison 2012] suggests “tangential forces” can operate as a supplemental 2D input channel, enabling e.g., shear gestures and in situ high CD gain manipulations. Lastly, Touché [Sato 2012] uses swept frequency capacitive sensing to capture, for example, how a user is touching or grasping a device, and also enables several new pinching gestures that can enrich interactions.

Concurrent with the previous efforts to enrich interaction, I also pursued research that aimed to increase the input area of mobile devices. I pursued three distinct strategies: 1) utilizing unused surface area, 2) moving interaction into free space, and 3) appropriating surface area.

The first strategy is the most straightforward – take maximal advantage of the surface area a device already has. Typically less than half of a device’s surface area is used for the screen and physical controls. This has led researchers to propose using the reverse side and bezel of devices ([Baudisch 2009] and [Ashbrook 2008] respectively). One largely overlooked opportunity was cords (e.g., headphones, power plug). A cord, although simple in form, has many interesting physical affordances that make it powerful as an input device – which we explored in Cord Input [Schwarz 2010]. Not only can a length of cord be grasped in different locations, but also pulled, twisted and bent - four distinct and expressive dimensions that could potentially operate in parallel. A second project, PocketTouch [Saponas 2011], demonstrated that adaptive capacitive sensing allowed for eyes-free multitouch input while devices resided in pockets (e.g., pants, jacket, shirt) and bags (e.g., backpack, purse). Like Cord Input, PocketTouch enables a rich set of gesture interactions on a surface not previously utilized. Unfortunately, the most obvious surfaces have already been identified, and even 100% utilization is still very little surface area for many classes of device, so this approach seems unlikely to fully alleviate the current woes of mobile input.

If a device is of a particular size, and input has to occur within these dimensions, then logic would suggest that input can only be as large as the device. It is this maxim that illuminates a possible way forward: the only way to have input larger than the device, is to get off of the device. In other words, it is necessary to decouple input from the small physical constraints of the mobile device. One option is to move interaction into the unused “air” space around the device. This volume is many orders of magnitude larger than any mobile device. The core challenge here is technical – unlike a finger touching a touchscreen, monitoring activity in the air requires sensing from afar. Further, mobile devices have to be self-sufficient and self-contained, relying on no external infrastructure.

My initial foray into mobile free space interaction was Abracadabra [Harrison 2009], a magnetically-driven sensing technique that provided wireless finger tracking (and “clicking”) without requiring powered external components (though users must wear a small permanent magnet). Although the mobile device we used was small (1.5” diagonal LCD), users could provide finger input within a roughly 4” radius, providing 50in^2 of input area (fifty times larger than the screen). This provided a high control-device (CD) gain and reduced screen occlusion by moving finger interaction off of the display. I also worked on a technique called Whack Gestures [Hudson 2010], in which users interact by coarsely striking (“whacking”) a mobile device with an open palm or heel of the hand. This enables interaction without getting out, grasping, or even glancing at the device. Although the gesture is performed in free space, the motion (i.e. acceleration) is only captured at the point of contact with the device.

Free space interaction alleviates the immediate problem of limited interaction space. However, we lose many of the physical affordances that make real, hard surfaces great and, for the most part, unbeatable in terms of input precision and speed. Not only do in-air targets provide no tactile feedback, but also generally lack strong reference points. Furthermore, there is something pleasant and intuitive about physical interactions – tangibles we can grasp, motions with friction, buttons we can "click", and so on. Finally, physical surfaces often allow for projection of coordinated graphical feedback (e.g., icons, buttons, and menus), allowing for an interactive area many times the size of the device and an almost endless array of applications.

This realization underscores the key advantage of the third strategy I have pursued – to opportunistically appropriate surface area from other objects and the environment. Most surfaces are orders of magnitude larger than small devices, are often within arms' length (e.g., walls, tables) or otherwise approachable (e.g., doors, floors), and sometimes even ergonomically engineered (e.g., chairs, desks). By temporarily "stealing" surface area from everyday things, we can make small devices big, while retaining the benefits of mobility and interaction on physical surfaces. Similar to free space interaction, this is chiefly a sensing challenge.

My first project to capitalize on the potential of ad hoc appropriated surfaces was Scratch Input [Harrison 2008]. Using a small microphone integrated into the back of a mobile device, scratches (e.g., with a fingernail) on solid surfaces can be detected. This allows whatever surface a device happens to be resting on (e.g., desk) to be used as an ad hoc gestural input canvas. Next was Minput [Harrison 2010], which proposed integrating two optical tracking sensors onto the back of a device. This allows the whole device to be manipulated for input on any convenient surface, such as a door, book, and wall. The use of two tracking elements enables not only conventional X/Y tracking and 2D gestures in general, but also rotation, providing a more expressive design space. With Scratch Input and Minput, the device could be very small - potentially lacking a screen and even buttons - but the interaction space could be very large.

These successful endeavors into increasing input area for mobile devices, and in particular, appropriating surface area, culminated in a specialized thread of research, which is the focus of this dissertation: On-Body Computing. Although there was much power and in appropriating tables, wall, books and similar, the utility hinges on the availability of those surfaces. In Minput, we briefly suggested the palm as a possible fall back option. Although the palm was small, it was at least as large as typical smartphone. It was this realization that ultimately led me to consider appropriating the human body for interactive purposes, as it always travels with us. However, users have little tolerance for being instrumented with electronics, so sensing would have to be minimally invasive. Fortunately, these were the characteristics of the acoustic sensing approach used in Scratch Input. Over the summer of 2009, I developed Skinput (Chapter 3), which brought acoustic sensing to the body and allowed the skin to be appropriated for interactive experiences.

1.7 Skintillating Possibilities

The promise of On-Body Interfaces lies in their unique ability to overcome key limitations inherent in mobile devices, while simultaneously retaining the key benefit of mobility. If we set aside the sensing and interaction complexities and consider the body as a device, we can see it offers several unique qualities.

Foremost, the body provides considerable surface area - one hand's area alone exceeds that of typical smart phone; in total we have roughly 2m^2 of skin. Further, skin provides a natural and immediate surface for dynamic digital projection. Although skin introduces some color and physical distortion, the resolution, frame rate, and overall quality can be high [Mistry 2009; Yamamoto 2007; Harrison 2010, 2011].

Secondly, as the colloquialism "like the back of your hand" suggests, we are intimately familiar with our own bodies. Indeed, our body is the only "device" we receive training with from birth and every waking moment thereafter. Because of this, we are incredibly dexterous; our kinesthetic senses allow us to rapidly and accurately position our body, limbs, and digits - without external tactile feedback and even with our eyes closed [Mine 1997; Gallagher 2005; Wolfe 2006; Shumway-Cook 2011]. We also develop finely tuned muscle memory and hand-eye coordination. This immediately and naturally provides a high level of interactive performance, especially in finger input precision and gesturing – two powerful interaction modalities.

Moreover, the body has dozens of additional degrees of freedom that could be captured for interactive purposes [Laakso 2006; Warren 2003]. On-body interfaces can also unify cognition and bodily action, increasing agency [Coyle 2012] and offering tremendous potential to outperform other interaction modalities, since the interface being touched is the users' own body [Noë 2005; Valera 1991; Wilson 1998]. Finally, the rise of tangible computing has demonstrated that object-specific manipulations such as shaking, squeezing and rotating physical artifacts, align embodiment with physical representation and embeddedness in space [Hornecker 2006].

Finally, our bodies are always with us and often immediately available [Tan 2010; Saponas 2009]. This stands in contrast to conventional mobile devices, which typically reside in pockets or bags, and must be retrieved to access even basic functionality [Ashbrook 2008; Saponas 2011; Hudson 2010]. This generally demands a high level of attention - both cognitively and visually - and is often socially disruptive. Further, physically retrieving a device incurs a non-trivial time cost, and can constitute a significant fraction of a simple operation's total time [Ashbrook 2008].

2 RELATED SYSTEMS

My research on on-body computing draws from a variety of fields, including interaction techniques, touch sensing, bio-sensing, surface computing, free-space gesturing, computer vision, wearables, cybernetics, and ubiquitous computing. Here I focus on related work that was most influential.

2.1 Additional Dimensions of Touch Input

As noted previously, touch is traditionally used for positional input - put simply, we use our fingers as pointing devices. To augment positional input, researchers have developed interaction techniques to aid mode switching and contextual operations, including touch-and-hold [Li 2005] and multi-finger chording actions [Lepinski 2010].

However, even without spatial or temporal overloading, our fingers are capable of providing several additional dimensions of input. For example, touchscreens able to capture finger pressure and shear forces date back at least as far as 1978 [Herot 1978]. Exploration of the interaction space was limited, given that it predated popular use of graphical user interfaces. Recently, Heo and Lee [2011] have taken the idea mobile, augmenting an iPod Touch with pressure and shear sensing. The interactive implications of pressure input have been extensively explored by Ramos et al. [2004, 2005, 2007].

Finger angle of attack and orientation are two additional useful dimensions. Wang and Ren [2009] proposed a detection method for optical surfaces, along with several interaction techniques enabled by the additional information, for example, a pie menu that can be navigated by twisting the finger. Rogers et al. [2011] introduced an approach based on capacitive sensing, and suggest, for example, the additional input dimensions could be used for 3D map navigation. Related is MicroRolls [Roudaut 2009], which looked at in situ rolling of the fingers.

There have also been interaction techniques that use the contact area or shape of the hands and fingers for triggering different interactive modes [Cao 2008]. Paradiso [2000] and Harrison [2011] demonstrated acoustic sensing can allow for different parts of the finger to be recognized in touch applications (e.g., finger pad vs. knuckle). Dietz et al. [2005] showed that touches can be attributed to a particular user (when uniquely grounded) in DiamondTouch.

Finally, because mobile devices are typically handheld, the device itself can be manipulated, providing an input channel that can operate in concert with touch. For instance, device tilt can be used for panning [Hinkley 2011], advancing pages in an ebook [Harrison 1998], or even for text entry [Wigdor 2003]. Squeezing and bending has also been investigated as an input means [Schwesig 2004; Lahey 2011]. Lastly, where the device is being held relative to the body can also have compelling applications [Li 2009].

2.2 Bio-Sensing

Biosignals are a class of signal that can be measured and monitored from biological beings. Biological functions can be both voluntary and involuntary, the latter sometimes operating at subconscious levels. Biosignals are captured by sensors traditionally used in diagnostic medicine and psychology, and have been applied in HCI domains. For example, Mandryk et al. [2006, 2007] used heart rate and skin resistance to assess user experience factors. Moore et al. [2004] suggest skin resistance could be used for simple binary (i.e., yes/no) control. Breathing rate has also been used to enhance entertainment experiences [Marshall 2011].

There has also been much research on brain-sensing technologies, including electroencephalography (EEG), electrocorticography (ECoG) functional near-infrared spectroscopy (fNIR), magnetoencephalography (MEG) and functional magnetic resonance imaging (fMRI). Sensor data has been used to assess cognitive and emotional state [Grimes 2008; Hirshfield 2009, Lee 2006], as well as direct input for use by paralyzed patients [Fabiani 2004; McFarland 2003], including 2D cursor control [Schalk 2008]. However, in general, contemporary brain-computer interfaces (BCIs) lack the bandwidth required for everyday computing tasks, and generally require high levels of training and concentration.

Researchers have used electromyography (EMG), which can detect electrical signals generated by muscle activation (e.g., arm movement), for interactive purposes. For instance, Rosenberg [1998] demonstrated control of a 2D cursor; Benko et al. [2009] built a multitouch table able to sense what finger was being used for input. Interactions built using muscle sensing could operate eyes free, without the use of graphical feedback, as proposed in [Saponas 2008, 2009]. For example, a music player could be controlled with finger-to-finger pinches. Until recently, EMG typically required expensive amplification systems and the application of conductive gel for effective signal acquisition. However, newer armband systems have been developed that are gel-free and also wireless [Saponas 2010].

Bone conduction microphones and headphones – now common consumer technologies – represent an additional bio-sensing technology that is relevant to the present work. These leverage the fact that sound frequencies relevant to human speech propagate well through bone. Bone conduction microphones are typically worn near the ear, where they can sense vibrations propagating from the mouth and larynx during speech. Bone conduction headphones send sound through the bones of the skull and jaw directly to the inner ear, bypassing lossy transmission of sound through the air and outer ear. The mechanically conductive properties of human bones are also employed by [Zhong 2007] for transmitting information through the body, such as from an implanted device to an external receiver.

Finally, bio-acoustics have also been leveraged for computer input. Amento et al. [2002] placed contact microphones on a user's wrist to detect finger movement. The Hambone system [Deyle 2007] employed a similar setup, and through a Hidden Markov Model (HMM), yields classification accuracies around 90% for four gestures (e.g., raise heels, snap fingers). Performance of false positive rejection remains untested in both systems at present. Both techniques require placement of sensors near the area of interaction (e.g., the wrist).

2.3 Hand and Body Sensing

Body-driven and hand-driven input has received attention for decades, and scores of advanced systems are able to detect, track and recognize hands and limbs for a variety of purposes. Capturing all of the degrees of freedom of the hands has proven particularly challenging. On high fidelity approach is to instrument the hands directly with mechanical sensors [Struman 1994], though it is fairly invasive. Alternatively, remote sensing through, e.g., computer vision, avoids having to instrument the user (see [Erol 2007] and [Wachs 2011] for an excellent survey).

The hands can be used in many ways. Foremost is gestural input, as demonstrated in [Starner 1998] with real-time American Sign Language recognition. Cho et al. [2002] introduce the idea of body-inspired metaphors, for example, pinching one's ear to adjust volume or pointing at one's eye to activate a graphical display. Alternatively, the hands can be digitized primarily for positional input, at seen in [Wilson 2006], which uses a single pinching gesture for activation, along with 3D positional tracking. Finally, gestural and positional data can be in concert, for example, controlling a computer mouse in freespace (including left and right "clicks") [Argyros 2006; Yamamoto 2009].

Due to the complexity and invasiveness of instrumenting the whole human body, motion capture typically uses some form of computer vision (often assisted with e.g., infrared or retroreflective markers). Moeslund et al. [2006] provide a comprehensive overview of human motion capture research efforts. Approaches include modeling the optical flow of limbs [Kim 2008], using silhouettes to train support vector machines [Agarwal 2004], and leveraging skin color and body geometry to identify limbs [Siddiqui 2006]. Researchers have also considered interaction design questions, for example, how full-body gestures could be used for interactive purposes [Laakso 2006; Warren 2003]. There are also many perceptual and psychophysical issues in free-space spatial input, including reference points, absolute vs. relative movement, and one- vs. two-handed manipulation [Hinckley 1994].

2.4 Appropriating Ad-Hoc Surfaces

Researchers have extolled the virtues of mobile devices "opportunistically annexing" computational resources sprinkled around the environment [Pierce 2004]. However, given the prodigious advance of electronics, it is my view that mobile devices are computationally capable - the need is not for additional CPU power or memory, but rather area for interaction.

Appropriating surfaces for digital projection is generally classified as augmented reality (see [Zhou 2008] for a review). For better control, many systems semi-permanently or permanently instrument the environment in some manner. The Everywhere Displays project [Pinhanez 2001] uses ceiling mounted projector with an articulating mirror, allowing for dynamic projections nearly anywhere in a room, including walls, tables and even objects. RFIG Lamps [Raskar 2006] uses photo-sensors placed on objects or distributed around the environment; a handheld projector interacts with these sensors through structured light, which in turn provides object identification and geometry. With this information, the projector can then render coordinated graphical feedback onto objects and the environment. Other projects have used fiducial markers [Beardsley 2005] or fixed infrastructure to track the projector in 3D space [Cao 2006, 2007; Blasko 2005]. Other efforts have attempted to be infrastructure-less, and self-contained. Willis et al. [2011a, 2011b] investigated handheld projectors with integrated buttons and accelerators, and later, invisible projected fiducial markers, enabling two or more users to interact.

Appropriating surfaces for input requires different technologies and approaches. Tomasi et al. [2003] described a projected keyboard that can appropriate any flat surface for typing; sensing was achieved with an infrared camera and line laser. SideSight [Butler 2008] was a mobile device ringed by infrared proximity sensors. When laid flat on a surface (e.g., on a table), the area surrounding the device could be used for 2D multi-finger input (within a limited radius). Bonfire [Kane 2009] attached projectors and cameras to the rear side of a laptop screen, enabling interactive projected areas on either side of the computer; finger inputs were digitized through computer vision. Wilson constructed the PlayAnywhere system [2005], which, using IR illumination and shadow shape analysis, could track hands (including "clicks") on everyday surfaces.

Recently, low cost depth cameras have opened new opportunities for ad hoc input, particularly touch sensing on the environment. LightSpace [Wilson 2010] utilized a fixed overhead array of depth cameras and projectors to augment a room with multi-touch capability (e.g., walls, tables). Touches were detected by creating synthetic planar cameras, which could be processed with traditional 2D computer vision techniques. Wilson [2010] demonstrated a single camera could provide conventional touch events by using a per-pixel depth threshold determined from a histogram of the static scene. Both approaches work on a variety of surfaces, but require careful calibration before they can operate.

2.5 On-Body Input

The primary goal of on-body interfaces is to provide an always-available mobile input system – that is, an input system that does not require a user to carry or pick up a device [Tan 2010; Saponas 2009]. To support this class of interaction, a number of approaches have been proposed. The most straightforward is to take conventional physical computing elements and place them on the body. Iconic examples include a one handed keyboard [Lyons 2004] and a wrist-bound touchpad [Thomas 2002]. A similar approach involves input devices built in a form considered to be part of one's clothing [Post 1997; Cho 2002; Mann 1997]. However, taking this approach to always-available input necessitates embedding technology in all clothing, which is currently prohibitively expensive.

It is also possible to instrument the user more directly. For example, glove-based input systems [Sturman 1994] can capture a great deal of expressivity, and allow users to retain most of their natural hand movements. However, such systems are cumbersome, uncomfortable, and sometimes disruptive to tactile sensation. Research has also looked at instrumenting just the tip of the fingers, as in [Mascaro 2004], which could sense finger posture and shear forces. Gemperle et al. [1998] recommend a list of body locations that are suitable for instrumentation, considering attachment means, weight, accessibility, thermal constrains, and several other dimensions.

Researchers have also looked at wearable computer vision systems. Starner et al. [1998] demonstrated real-time American Sign Language recognition from a down-facing camera worn on a hat. Ni and Baudisch [2009] considered micro mobile devices (e.g., the size of a rice grain) and how users might interact with them. The interaction techniques they propose are largely supported by optical sensing and computer vision. Gustafson et al. [2010, 2011] have investigated "imaginary interfaces" – systems with computer-vision-driven finger input, but lacking graphical output. Sensing is achieved by a sensing pendant, similar to [Starner 2000], which can recognize hand gestures. It is even possible to resolve the skeletal movements of the user's body (i.e. motion capture) using an array of body-mounted cameras [Takaaki 2011].

Finally, speech input has been an active research area for decades. Like cameras, microphones are small and can sense from afar, allowing them to be worn, minimally invasive and even hidden [Cohen 1994; Starner 2000, Lakshmipathy 2003; Lyons 2005]. However, speech recognition is limited in its precision, inherently sequential, and unpredictable in noisy environments. Further, it suffers from privacy and scalability issues in shared environments. Starner et al. [2002] suggest it may even interfere with cognitive tasks significantly more than manual interfaces. However, a key benefit of speech is its ability to operate in parallel with physical input [Bolt 1980]. Researchers have also considered using non-speech sounds for input. One example is [Harada 2006], which enabled cursor control with elongated vowel sounds.

2.6 On-Body Projection

Although many projects have employed mobile projectors, few have taken advantage of the body as a projection surface. Unsurprisingly, the art community was among the first to embrace the fusion of the human form with projected digital media. Examples include the opening sequence to Guy Hamilton's “Goldfinger” (1964) and Peter Greenway’s “The Pillow Book” (1996), both of which projected text and images onto actors’ bodies (Figure 2.1). More recently, an interactive installation by Sugrue [2007] allowed people to touch a screen with virtual “bugs”, which could move out onto people’s hand and arms. Barnett [2009] provides a comprehensive summary of many of these artistic efforts.

Fig. 2.1 **Left: Frame from the open credits of Guy Hamilton's "Goldfinger". Right: Promotional poster of Peter Greenway's "The Pillow Book".**

There has also been academic and commercial interest in on-body or worn projection systems. Using overhead projection and hand tracking, TenoriPop [NTT 2010], designed to enhance the retail experience, can render interactive elements onto the palms of users. Interactive Dirt [McFarlane 2009] is a shoulder mounted projector and camera system. Fingers are tracked using infrared retroreflective markers; alternatively, the user can use an infrared laser pointer. The authors propose projecting visual content onto the ground, trees, cars and other surfaces in the environment (Figure 2.2). The projection geometry is fixed, so users must physically orient their bodies or the device.

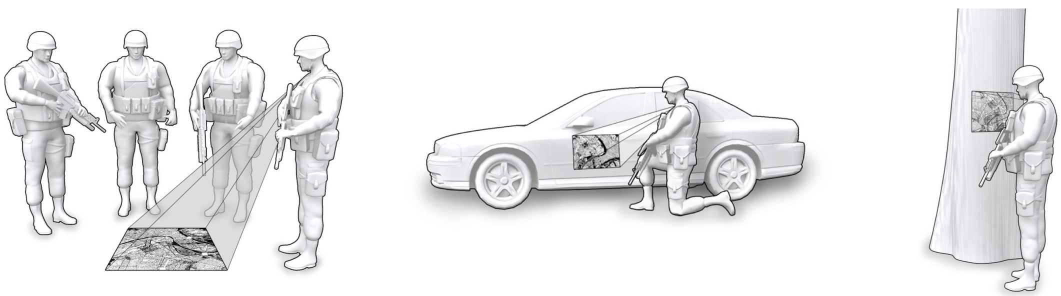


Fig. 2.2 **Proposed example uses for McFarlane and Wilder's [2009] Interactive Dirt system.**

On-body projections have also seen interest in the medical domain. For example, Gavaghan et al. [2011] propose using calibrated 3D projections to assist in complex surgical procedures, for example, displaying the location of tumors and blood vessels. Projection of anatomical data has also been suggested for educational purposes, for example, projecting onto students bodies in order to provide a personalized understanding of the relative location of organs [Patten 2007; Donnelly 2009].

2.7 On-Body Interfaces

Rarest and most recent are systems that attempt both input and graphical output on the body. It is this unique combination that defines *On-Body Interfaces*, and enables a range of sophisticated interactions and applications not possible with input or output alone. I now briefly describe the most notable research efforts.

Karitsuka and Sato [2003] constructed a backpack worn system; an infrared camera and visible light projector operate over-the-shoulder (Figure 2.3 left). Retroreflective markers are used to track objects held in the hands, which allows for graphics to track with objects and also appear correctly rectified. The fingers can be used for input, but must be instrumented with an infrared LED. Although the authors do not describe projecting onto the body, all of the components of an on-body interactive system are present.

Fig. 2.3 **Left: [Karitsuka 2003]. Right: PALMbit [Yamamoto 2007a, 2007b].**

PALMbit [Yamamoto 2007a, 2007b] is a shoulder-worn projector and camera system (Figure 2.3 right). Interfaces are projected onto the palm of the user, which is actively tracked at real-time speeds without the need for markers. Users provide input to the system by pressing (e.g., with their dominant hand index finger) one of the five fingers in their non-dominant (projection) hand. This essentially provides five "buttons" for interaction. The authors demonstrate photo album navigation as an example application. Importantly, the fingers must be spread apart for successful tracking and cannot be occluded by the other hand.

Sakata et al. [2009] describe a "palm top" projection system (Figure 2.4 left). Tracking is achieved using two fiducial markers worn on the wrist – one palm side and one on the back of the hand. This provides two modes and also the 3D posture and position of the user's hand. The authors suggest these projections are ideal for glance-able information; input is not supported. However, the spatial location of the hand could be readily used for positional input and gestures.

SixthSense [Mistry 2009] is a proposed pendent-like device, containing a camera and pico-projector (Figure 2.4 right). Through computer vision, the authors suggest that different objects in the environment could be recognized. Using colored markers, fingers and hands could be tracked for input and gestures. The authors touch on a variety of projected interactions, mostly onto the environment (e.g., walls and objects). Proposed on-arm examples include the dialing of a phone with the fingers and summoning a watch on the wrist with a finger circling motion. Additionally, several static gestures are suggested, including a finger "square" for capturing a photograph.

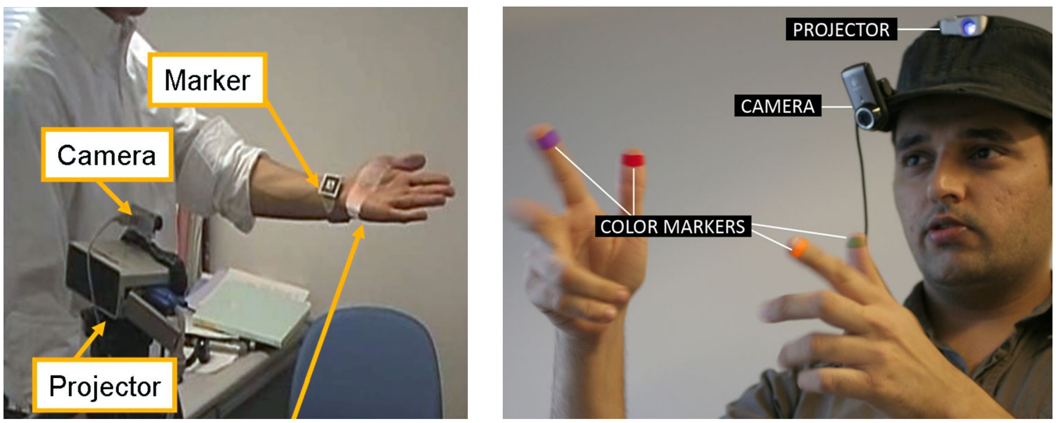


Fig. 2.4 **Left: "Palm top display for glance information" [Sakata 2009].**
Right: SixthSense [Mistry 2009].

There is also great value in fixed infrastructure that can augment spaces with on-body capabilities. This allows any user occupying the room to utilize the features and also eliminates the need for electronics to be worn. LightSpace [Wilson 2010] is one such system, utilizing a fixed overhead array of depth cameras and projectors. Interactions enabled by this system are diverse; relevant to the proposed work is an on-hand "spatial menu". Users can move their hands over a specific "menu" location, and then move their hands in the Z-axis (up/down) to select from various menu items. Selection is achieved by dwelling for a brief period. The authors also experimented with virtual items, which can be "held" in the hands, and then transferred to other projected surfaces in the environment through touch.

3 SKINPUT: INTERACTIVE SKIN USING BIO-ACOUSTICS

My research into on-body interfaces began with Skinput: a non-invasive, bio-acoustic input technique that allowed the skin to be used as an input surface, much like a touchscreen. It works by listening to the sound of finger taps on the skin. The resulting ensemble of vibrations is different for different locations, which can be learned by a classifier. Coupled with a pico-projector, direct manipulation touch interfaces could be rendered on the skin – the first system to achieve this result.

3.1 Bio-Acoustics

When a finger taps the skin, several distinct forms of vibro-acoustic energy are produced. Some energy is radiated into the air as sound waves; this energy is not captured by the Skinput system. Among the vibro-acoustic energy transmitted through the arm, the most readily visible are transverse waves, created by the displacement of the skin from a finger impact (Figure 3.1). When shot with a high-speed camera, these appear as ripples, which propagate outward from the point of contact (like a pebble thrown into a pond). The amplitude of these ripples is correlated to the tapping force and the volume and compliance of soft tissues under the impact area. In general, tapping on soft regions creates higher-amplitude transverse waves than tapping on boney areas (e.g., wrist, palm, fingers), which have negligible compliance.

In addition to the vibro-acoustic energy that propagates on the surface of the skin, some energy is transmitted inward, toward the skeleton (Figure 3.2). These longitudinal (compressive) waves travel through the soft tissues of the arm, exciting the bone, which is much less deformable than the soft tissue, but can respond to mechanical excitation by rotating and translating as a rigid body. This excitation vibrates tissues surrounding the entire length of the bone, resulting in new longitudinal waves that radiate outwards, towards the skin.

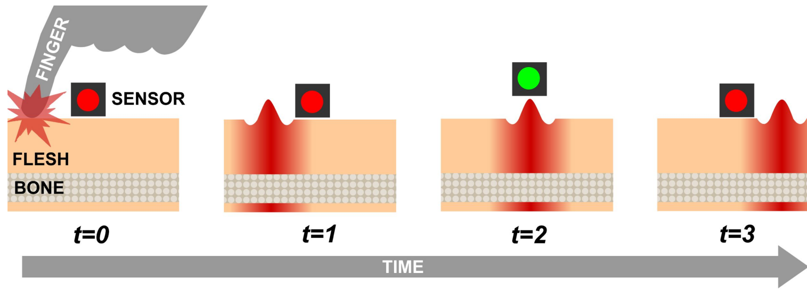


Fig. 3.1 **Transverse wave propagation: Finger impacts displace the skin, creating transverse waves (ripples). The sensor is activated as the wave passes underneath it.**

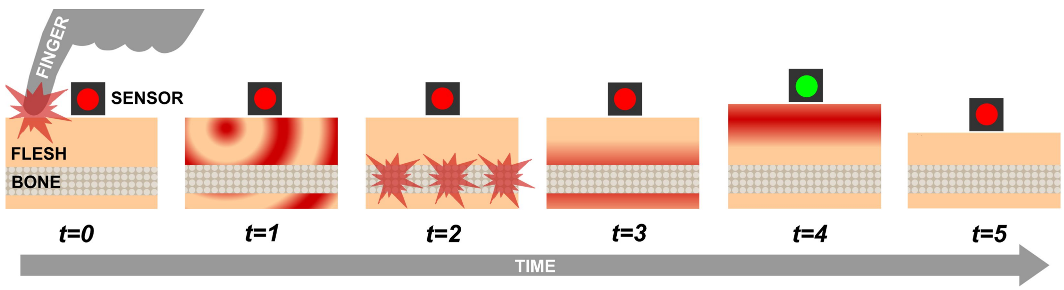


Fig. 3.2 **Longitudinal wave propagation: Finger impacts create longitudinal waves that cause internal skeletal structures to vibrate. This, in turn, creates longitudinal waves that emanate outwards from the bone (along its entire length) toward the skin.**

We highlight these two separate forms of conduction – transverse waves moving directly along the arm surface, and longitudinal waves moving into and out of the bone through soft tissues – because these mechanisms carry energy at different frequencies and over different distances. Roughly speaking, higher frequencies propagate more readily through bone than through soft tissue, and bone conduction carries energy over larger distances than soft tissue conduction. While we do not explicitly model the specific mechanisms, or depend on these mechanisms for our analysis, we do believe the success of our technique depends on the complex vibro-acoustic patterns that result from mixtures of these modalities.

Similarly, we also hypothesize that joints play an important role in making tapped locations acoustically distinct. Bones are held together by ligaments, and joints often include additional biological structures such as fluid-filled cavities. This makes joints behave as acoustic filters. In some cases, these may simply dampen acoustics. In other cases, these will selectively attenuate specific frequencies, creating location-specific acoustic signatures. Finally, muscle contraction may also affect the vibro-acoustic patterns recorded by our sensors [Matheson 1997], including both contraction related to posture maintenance and reflexive muscle movements in response to input taps.

3.2 Sensing

To capture the rich variety of vibro-acoustic information described in the previous section, we evaluated several sensing technologies, including bone conduction microphones, conventional microphones coupled with stethoscopes [Harrison 2008], piezo contact microphones [Amento 2002], and accelerometers. However, these transducers were engineered for very different applications than measuring vibro-acoustics transmitted through the human body. As such, we found them to be lacking in several significant ways. Foremost, most mechanical sensors are engineered to provide relatively flat response curves over the range of frequencies that is relevant to our signal. This is a desirable property for most applications, where a faithful representation of an input signal – uncolored by the properties of the transducer – is desired. However, because only a specific set of frequencies is conducted through the arm in response to finger tap, a flat response curve leads to the capture of irrelevant frequencies and thus to a high signal-to-noise ratio.

While bone conduction microphones might seem a suitable choice for Skinput, these devices are typically engineered for capturing human voice, and filter out energy below the range of human speech (whose lowest frequency is around 85Hz). Thus most sensors in this category were not especially sensitive to lower-frequency signals (e.g., 25Hz), which we found in our empirical pilot studies to be vital in characterizing finger taps.

To overcome these challenges, we moved away from a single sensing element with a flat response curve, to an array of highly tuned vibration sensors. Specifically, we employ small, cantilevered piezo films (MiniSense100 [Measurement Specialties]). By adding small weights to the end of the cantilever, we were able to alter the resonant frequency, allowing the sensor to be responsive to a unique, narrow, low-frequency band of the vibro-acoustic spectrum. Adding more mass lowers the range of excitation to which a sensor responds. We weighted each sensor such that it aligned with particular frequencies that pilot studies showed to be useful in characterizing bio-acoustic input.

Figure 3.3 shows the response curve for one of our sensors, tuned to a resonant frequency of 78Hz. The curve shows a ~14dB drop-off ±20Hz away from the resonant frequency.

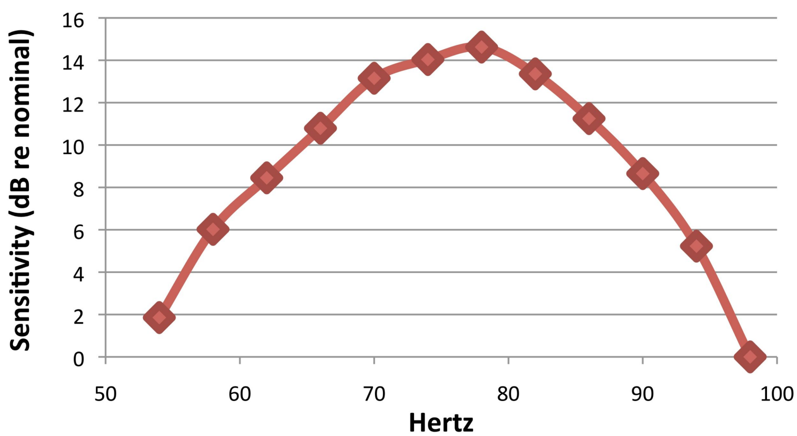


Fig. 3.3 **Response curve (relative sensitivty) of the sensing element that resonates at 78 Hz.**

Additionally, the cantilevered sensors were naturally insensitive to forces parallel to the skin (e.g., shearing motions caused by stretching). Thus, stretching of the skin induced by many routine movements (e.g., reaching for a doorknob) tends to be attenuated. However, the sensors are highly responsive to motion perpendicular to the skin plane – ideally suited for capturing transverse surface waves (Figure 3.1) and longitudinal waves emanating from interior structures (Figure 3.2).

Finally, our sensor design is relatively inexpensive and can be manufactured in a very small form factor (e.g., a micro electro-mechanical device), rendering it suitable for inclusion in future mobile devices (e.g., an arm-mounted audio player).

3.3 Prototype Hardware

Our final prototype, shown in Figures 3.4 and 3.5, features two arrays of five sensing elements, incorporated into an armband form factor. The decision to have two sensor packages was motivated by our focus on the arm for input. This is an attractive area to appropriate as it provides considerable surface area for interaction, including a contiguous and flat area for projection. Furthermore, the forearm and hands contain a complex assemblage of bones that increases vibro-acoustic distinctiveness of different locations.

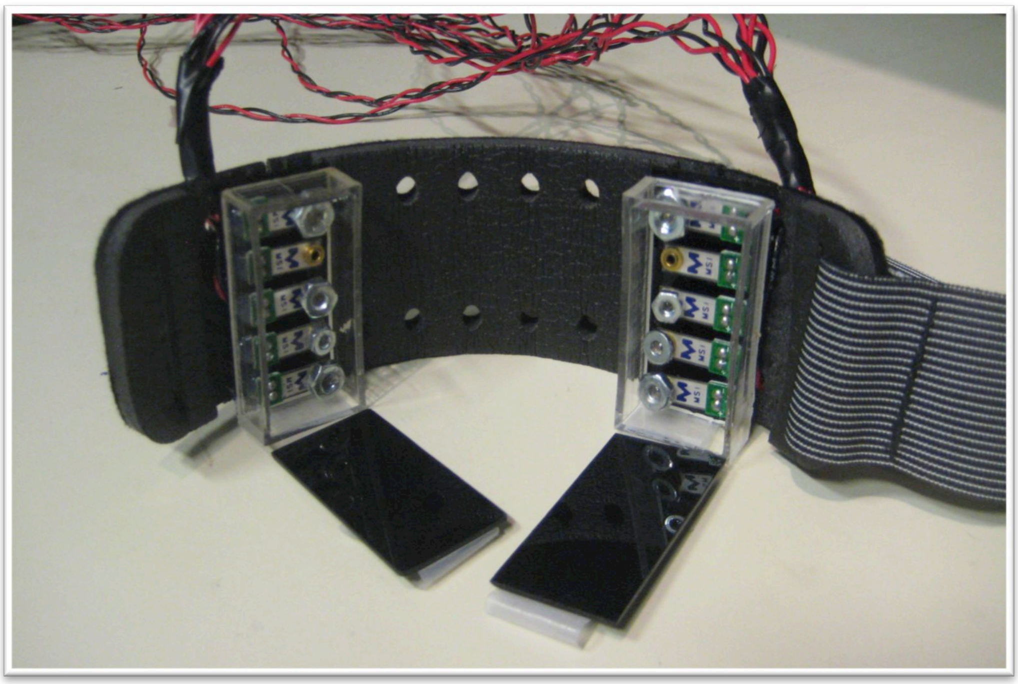


Fig. 3.4 **Our wearable, bio-acoustic sensing array built into an armband. The two sensor packages shown above each contain five, specially weighted, cantilevered piezo films, responsive to a particular frequency range.**

In particular, when placed on the upper arm (above the elbow), we hoped to collect acoustic information from the fleshy bicep area in addition to the firmer, underside of the arm, with better acoustic coupling to the *Humerus*, the main bone that runs from shoulder to elbow. When the sensor was placed below the elbow, on the forearm, one package was located near the *Radius*, the bone that runs from the lateral side of the elbow to the thumb side of the wrist, and the other near the *Ulna*, which runs parallel to this on the medial side of the arm closest to the body. Each sensor package thus provided slightly different acoustic coverage and information, helpful in disambiguating input location.

Based on pilot data collection, we selected a different set of resonant frequencies for each sensor package (Table 3.1). We tuned the upper sensor package to be more sensitive to lower frequency signals, as these were more prevalent in fleshier areas. Conversely, we tuned the lower sensor array to be sensitive to higher frequencies, in order to better capture signals transmitted though (denser) bones.

UPPER ARRAY	25 Hz	27 Hz	30 Hz	38 Hz	78 Hz
LOWER ARRAY	25 Hz	27 Hz	40 Hz	44 Hz	64 Hz

Table 3.1 Resonant frequencies of individual elements in the two sensor packages.

Although our bio-acoustic input approach is not strictly tethered to a particular output modality, coordinated graphical output is invaluable for complex applications. In response, we outfitted our armband with a [MicroVision] PicoP laser pico-projector (Figure 3.5). Skinput does not perform any real-time tracking (computer vision or otherwise) of the arm's location. However, there are two nice properties of wearing a projection-capable device on the arm that permitted us to largely sidestep calibration issues. Foremost, the arm is a relatively rigid structure, with one degree of freedom in the elbow. The projector, when attached appropriately, will generally track with the arm naturally. Second, since we have fine-grained control of the arm, making minute adjustments to align the projected image with the arm is trivial (e.g., projected horizontal stripes for alignment with the wrist and elbow).

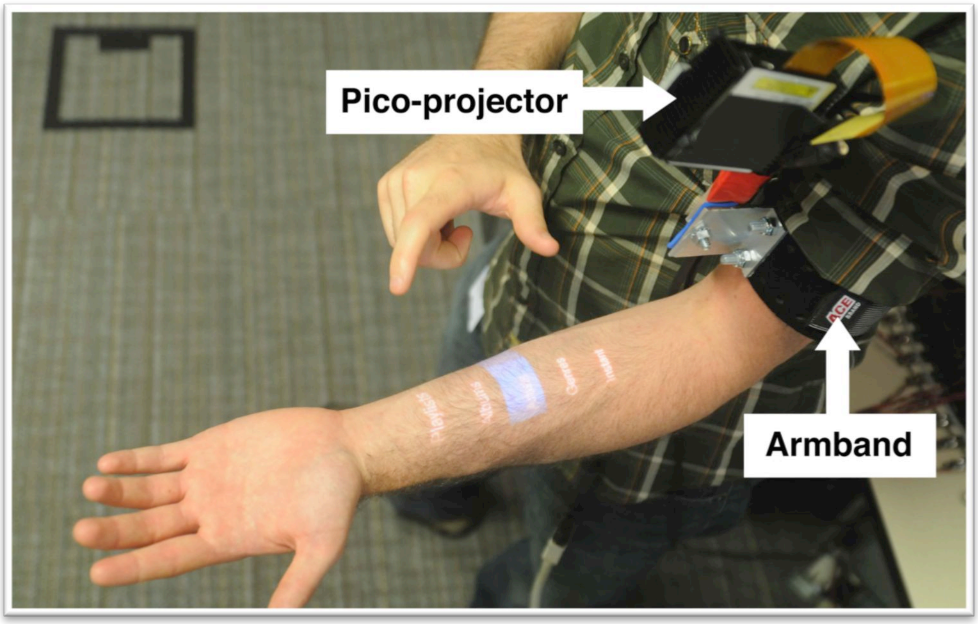


Fig. 3.5 Our Skinput armband outfitted with a pico projector. Here a scrollable list interface is projected onto the arm.

3.4 Processing

In our prototype system, we employ a Mackie Onyx 1200F audio interface to digitally capture data from the ten sensors (http://mackie.com). This was connected via Firewire to a conventional desktop computer, where a thin client written in C interfaced with the device using the Audio Stream Input/Output (ASIO) protocol. Each channel was sampled at 5.5kHz, a sampling rate that would be considered too low for speech or environmental audio, but was able to represent the relevant spectrum of frequencies transmitted through the arm. This reduced sample rate (and consequently low processing bandwidth) makes our technique readily portable to embedded processors. For example, the ATmega168 processor employed by the Arduino platform can sample analog readings at 77kHz with no loss of precision, and could therefore provide the full sampling power required for Skinput (55kHz total).

Data was then sent from our thin client over a local socket to our primary application, written in Java. This program performed three key functions. First, it provided a live visualization of the data from our ten sensors, which was useful in identifying acoustic features (Figure 3.6). Second, it segmented inputs from the data stream into independent instances (taps). Third, it classified these input instances.

The audio stream was segmented into individual taps using an absolute exponential average of all ten channels (Figure 3.6, red waveform). When an intensity threshold was exceeded (Figure 3.6, upper blue line), the program recorded the timestamp as a potential start of a finger tap. If the intensity did not fall below a second, independent "closing" threshold (Figure 3.6, lower purple line) between 100ms and 700ms after the onset crossing (a duration we found to be the common for finger impacts), the event was discarded. If start and end crossings were detected that satisfied these criteria, the vibro-acoustic data in that period (plus a 60ms buffer on either end) was considered an input event (Figure 3.6, vertical green regions). Although simple, this heuristic proved to be robust, mainly due to the extreme noise suppression provided by our sensing approach.

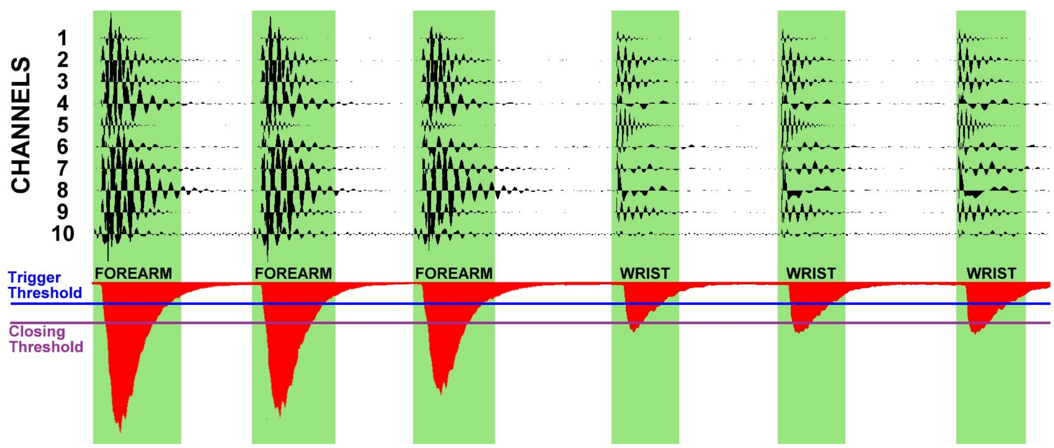


Fig. 3.6 **Ten channels of acoustic data generated by three finger taps on the forearm, followed by three taps on the wrist. The exponential average of the channels is shown in red. Segmented input windows are highlighted in green. Note how different sensing elements are actuated by the two locations.**

After an input has been segmented, the waveforms are analyzed. The highly discrete nature of taps (i.e. point impacts) meant acoustic signals were not particularly expressive over time (unlike gestures, e.g., clenching of the hand). Signals simply diminished in intensity overtime. Thus, features are computed over the entire input window and do not capture any temporal dynamics.

We employ machine learning for classification, computing 186 features in total, many of which are derived combinatorially. For gross information, we include the average amplitude, standard deviation and total (absolute) energy of the waveforms in each channel (30 features). From these, we calculate average amplitude ratios between pairs of channels (45 features). We also include an average of these ratios (1 feature). We calculate a 256-point FFT for all ten channels, although only the lower ten values are used (representing the acoustic power from 0Hz to 193Hz), yielding 100 features. These are normalized by the highest-amplitude FFT value found on any channel. We also include the center of mass (centroid) of the power spectrum within the same 0Hz to 193Hz range for each channel, which provides an approximation of the fundamental frequency of the signal displacing each sensor (10 features). Subsequent feature selection established the all-pairs amplitude ratios and certain bands of the FFT to be the most predictive features.

These 186 features are passed to a Support Vector Machine (SVM) classifier. Our software uses the SMO implementation provided in the Weka machine learning toolkit [Witten 2005]. Before the SVM can classify input instances, it must first be trained to the user and the sensor position. This stage requires the collection of several examples for each input location of interest. When using Skinput to recognize live input, the same 186 acoustic features are computed on-the-fly for each segmented input. These are fed into the trained SVM for classification. We use an event model in our software – once an input is classified, an event associated with that location is created. Any interactive features bound to that event are fired.

3.5 Evaluation

3.5.1 Participants

To evaluate the performance of our system, we recruited 13 participants (7 female) from the Seattle metropolitan area. These participants represented a diverse cross-section of potential ages and body types. Ages ranged from 20 to 56 (mean 38.3), and body mass indexes (BMIs) ranged from 20.5 (normal) to 31.9 (obese).

3.5.2 Experimental Conditions

We selected three input groupings from the multitude of possible location combinations to test. We believe that these groupings, illustrated in Figure 3.7, are of particular interest with respect to interface design, and at the same time, push the limits of Skinput's sensing capability. From these three groupings, we derived five different experimental conditions described below.

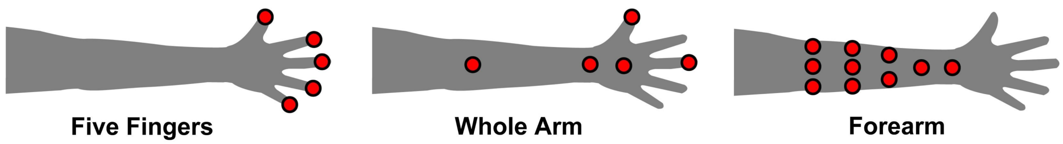


Fig. 3.7 The three input location sets evaluated in the study.

3.5.2.1 Fingers (Five Locations)

One set of gestures we tested had participants tapping on the tips of each of their five fingers (Figure 3.7, "Five Fingers"). The fingers offer interesting affordances that make them compelling to appropriate for input. Foremost, they provide clear, discrete interaction points, which are well-named (e.g., "ring finger"). In addition to five finger tips, there are 14 knuckles (five major, nine minor), which, taken together, could offer 19 readily identifiable input locations on the hands alone. Second, we have exceptional finger-to-finger dexterity, as demonstrated when we count by tapping on our fingers. Finally, the fingers are linearly ordered, which is potentially useful for interfaces like number entry, magnitude control (e.g., volume), and menu selection.

At the same time, fingers are among the most uniform appendages on the body, with all but the thumb sharing a similar musculoskeletal structure. This drastically reduces vibro-acoustic variation and makes differentiating among them difficult. Additionally, acoustic information must cross as many as five (finger and wrist) joints to reach the forearm, which further dampens the signal. Despite these difficulties, pilot experiments showed measureable vibro-acoustic differences among fingers, which we theorize is primarily related to finger length and thickness, interactions with the complex structure of the wrist bones, and variations in the acoustic transmission properties of the muscles extending from the fingers to the forearm. For this experimental condition, we decided to place the sensor arrays on the forearm, just below the elbow.

3.5.2.2 Whole Arm (Five Locations)

Another task investigated the use of five input locations on the forearm and hand: arm, wrist, palm, thumb and middle finger (Figure 3.7, “Whole Arm”). We selected these locations for two important reasons. First, they are distinct and named parts of the body (e.g., “wrist”). This allowed participants to accurately tap these locations without training or markings. Additionally, these locations proved to be vibro-acoustically distinct during piloting, with the large spatial spread of input points providing further variation.

We used this location set in three different conditions. One condition placed the sensor above the elbow, while another placed it below. This was incorporated into the study to measure the accuracy loss across this significant articulation point (the elbow). Additionally, participants repeated the lower placement condition in an eyes-free context: participants were told to close their eyes and face forward, both for training and testing. This condition was included to gauge how well users could target on-body input locations in an eyes-free context (e.g., driving).

3.5.2.3 Forearm (Ten Locations)

In an effort to assess the upper bound of our approach’s sensing resolution, our fifth and final experimental condition used ten locations on just the forearm (Figure 3.7, “Forearm”). Not only was this a very high density of input locations (unlike the whole-arm condition), but it also relied on an input surface (the forearm) with a high degree of physical uniformity (unlike, e.g., the hand). We expected that these factors would make acoustic sensing difficult. Moreover, this location was compelling due to its large and flat surface area, as well as its immediate accessibility, both visually and for finger input. Simultaneously, this makes for an ideal projection surface for dynamic interfaces.

To maximize the surface area for input, we placed the sensor above the elbow, leaving the entire forearm free. Rather than naming the input locations, as was done in the previously described conditions, we employed small, colored stickers to mark input targets. This was both to reduce confusion (since locations on the forearm do not have common names) and to increase input consistency. As mentioned previously, we believe the forearm is ideal for projected interface elements; the stickers served as low-tech placeholders for projected buttons.

3.5.3 Design and Setup

We employed a within-subjects design, with each participant performing tasks in each of the five conditions in randomized order: five fingers with sensors below elbow; five points on the whole arm with the sensors above the elbow; the same points with sensors below the elbow, both sighted and blind; and ten marked points on the forearm with the sensors above the elbow.

Participants were seated in a conventional office chair, in front of a desktop computer that presented stimuli. For conditions with sensors below the elbow, we placed the armband roughly 3cm away from the elbow. For conditions with the sensors above the elbow, we placed the armband roughly 7cm above the elbow, such that one sensor package rested on the biceps. Right-handed participants had the armband placed on the left arm, which allowed them to use their dominant hand for finger input. For the one left-handed participant, we flipped the setup, which had no apparent effect on the operation of the system. Tightness of the armband was adjusted to be firm but comfortable. While performing tasks, participants could place their elbow on the desk, tucked against their body, or on the chair's adjustable armrest; most chose the latter.

3.5.4 Procedure

For each condition, the experimenter walked through the input locations to be tested and demonstrated finger taps on each. Participants practiced duplicating these motions for approximately one minute with each location set. This allowed participants to familiarize themselves with our naming conventions (e.g. "pinky", "wrist"), and to practice tapping their arm and hands with a finger on the opposite hand. It also allowed us to convey the appropriate tap force to participants, who often initially tapped unnecessarily hard.

To train the system, participants were instructed to comfortably tap each location ten times, with a finger of their choosing. This constituted one training round. In total, three rounds of training data were collected per input location set (30 examples per location, 150 data points total). An exception to this procedure was in the case of the ten forearm locations, where only two rounds were collected to save time (20 examples per location, 200 data points total). Total training time for each experimental condition was approximately three minutes.

We used the training data to build an SVM classifier. During the subsequent testing phase, we presented participants with simple text stimuli (e.g. "tap your wrist"), which instructed them where to tap. The order of stimuli was randomized, with each location appearing ten times in total. The system performed real-time segmentation and classification, and provided immediate feedback to the participant (e.g. "you tapped your wrist"). We provided feedback so that participants could see where the system was making errors, as they would if using a real world application. If an input was not segmented (i.e. the tap was too quiet), participants could see this and would simply tap again. Overall, segmentation error rates were negligible in all conditions, and not included in further analysis.

3.6 Results

In this section, we report on the classification accuracies for the test phases in the five different conditions. Classification had an across-conditions average accuracy of 87.6%.

3.6.1 Five Fingers

Despite multiple joint crossings and roughly 40cm of separation between the input locations and sensor armband, classification accuracy for the five-finger condition averaged 87.7% (SD=10.0%) across participants. Segmentation, as in other conditions, was essentially perfect.

3.6.2 Whole Arm

Participants performed three conditions with the whole-arm location configuration. The below-elbow placement performed the best, posting a 95.5% (SD=5.1%) average accuracy. This is not surprising, as this condition placed the armband closer to the input locations than any other condition. Moving the sensor above the elbow reduced accuracy to 88.3% (SD=7.8%), a drop of 7.2%. This is almost certainly due to the vibro-acoustic attenuation and obfuscation at the elbow joint, as well as the additional 10cm between the armband and input location.

The eyes-free input condition yielded lower accuracies than other conditions, averaging 85.0% (SD=9.4%). This represents a 10.5% drop from its vision-assisted (but otherwise identical) counterpart condition. It was apparent from watching participants complete this condition that targeting precision was reduced. In sighted conditions, participants appeared to be able to tap locations with a ~2cm radius of error. Although not formally captured, this margin of error appeared to double or triple when the eyes were closed. We believe that additional training data, which better captures the increased input variability, would remove much of this accuracy deficit. However, we also caution designers developing eyes-free, on-body interfaces to carefully consider the locations participants can tap accurately.

3.6.3 Forearm

Classification accuracy for the ten-location forearm condition was 81.5% (SD=10.5%), a surprisingly strong result for an input set we purposely devised to tax our system's sensing accuracy.

Using our experimental data, we considered different ways to improve accuracy by post-hoc collapsing the ten locations into input groupings. The goal of this exercise was to explore the tradeoff between classification accuracy and number of input locations on the forearm, which represents a particularly valuable input surface for application designers. We grouped targets into sets based on what we believed to be logical spatial groupings (Figure 3.8, A-E and G). In addition to exploring classification accuracies for layouts that we considered to be intuitive, we also performed an exhaustive search (programmatically) over all possible groupings. For most location counts, this search confirmed that our intuitive groupings were optimal. However, this search also revealed one plausible (although irregular) layout with high accuracy at six input locations (Figure 3.8, F).

Unlike in the five-fingers condition, there appeared to be shared vibro-acoustic traits that led to a higher likelihood of confusion with adjacent locations. This effect was more prominent laterally than longitudinally. Figure 3.8 illustrates this with lateral groupings consistently out-performing similarly arranged, longitudinal groupings (B and C vs. D and E). This is unsurprising given the morphology of the arm, with a high degree of bilateral symmetry along the long axis.

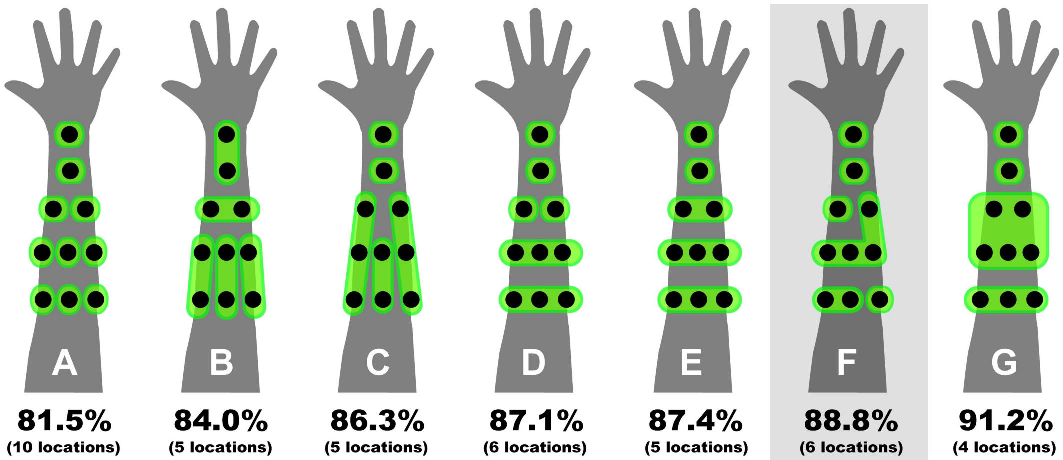


Fig. 3.8 Higher accuracies can be achieved by collapsing the ten input locations into groups. A-E and G were designed to be spatially intuitive. F was created following an analysis of per-location accuracy data.

3.7 Supplemental Experiments

We conducted a series of smaller, targeted experiments to explore the feasibility of our approach for other applications. In the first additional experiment, which tested performance of Skinput while users walked and jogged. We recruited one male (age 23) and one female (age 26) for the latter experiment. For the rest of the experiments, we recruited seven new participants (3 female, mean age 26.9) from within our institution. In all cases, the sensor armband was placed just below the elbow. Similar to the previous experimental procedures, each additional experiment consisted of a training phase, where participants provided between 10 and 20 examples for each input type, and a testing phase, in which participants were prompted to provide a particular input (ten times per input type). As before, input order was randomized; segmentation and classification were performed in real-time.

3.7.1 Walking and Jogging

With sensors coupled to the body, noise created during other motions is particularly troublesome, and walking and jogging represent perhaps the most common types of whole-body motion. To better understand Skinput in these conditions, two participants trained and tested the system while walking and jogging on a treadmill. Three input locations were used to evaluate accuracy: arm, wrist, and palm. Additionally, the rate of false positives (i.e., the system believed there was input when in fact there was not) and true positives (i.e., the system was able to correctly segment an intended input) was captured. The testing phase took roughly three minutes to complete (four trials total: two participants, two conditions). The male walked at 2.3 mph and jogged at 4.3 mph; the female at 1.9 and 3.1 mph, respectively.

In both walking trials, the system never produced a false-positive input. Meanwhile, true positive accuracy was 100%. Classification accuracy for the inputs (e.g., a wrist tap was recognized as a wrist tap) was 100% for the male and 86.7% for the female. In the jogging trials, the system had four false-positive input events (two per participant) over six minutes of continuous jogging. True-positive accuracy, as with walking, was 100%. Considering that jogging is perhaps the hardest input filtering and segmentation test, we view this result as extremely positive. Classification accuracy, however, decreased to 83.3% and 60.0% for the male and female participants respectively.

Although noise generated from the jogging almost certainly degraded the signal (and in turn, lowered classification accuracy), we believe the chief cause for this decrease may be the quality of the training data. Participants only provided ten examples for each of three tested input locations. Furthermore, the training examples were collected while participants were jogging. Thus, the resulting training data was not only highly variable, but also sparse – neither of which is conducive to accurate machine learning classification. We believe that more rigorous collection of training data could yield stronger results.

3.7.2 Single-Handed Gestures

In the experiments discussed thus far, we considered only bimanual gestures, where the sensor-free arm, and in particular the fingers, are used to provide input. However, there are a range of gestures that can be performed with just the fingers of one hand. This was the focus of [Amento 2002], although this work did not evaluate classification accuracy.

We conducted three independent tests to explore one-handed gestures. The first had participants tap their index, middle, ring and pinky fingers against their thumb (akin to a pinching gesture) ten times each. Our system was able to identify the four input types with an overall accuracy of 89.6% (SD=5.1%). We ran an identical experiment using flicks instead of taps (i.e., using the thumb as a catch, then rapidly flicking the fingers forward). This yielded an impressive 96.8% (SD=3.1%) accuracy in the testing phase.

This motivated us to run a third and independent experiment that combined taps and flicks into a single gesture set. Participants re-trained the system, and completed an independent testing round. Even with eight input classes in very close spatial proximity, the system was able to achieve 87.3% (SD=4.8%) accuracy. This result is comparable to the aforementioned ten-location forearm experiment (which achieved 81.5% accuracy), lending credence to the possibility of having ten or more functions on the hand alone. Furthermore, proprioception of our fingers on a single hand is very accurate, suggesting a mechanism for high-accuracy, eyes-free input.

3.7.3 Segmenting Finger Input

A pragmatic concern regarding the appropriation of fingertips for input was that other routine tasks would generate false positives. For example, typing on a keyboard strikes the finger tips in a very similar manner to the finger-tip-input we proposed previously. Thus, we set out to explore whether finger-to-finger input sounded sufficiently distinct such that other actions could be disregarded.

As an initial assessment, we asked participants to tap their index finger 20 times with a finger on their other hand, and 20 times on the surface of a table in front of them. This data was used to train our classifier. This training phase was followed by a testing phase, which yielded a participant-wide average accuracy of 94.3% (SD=4.5%, chance=50%).

3.8 Example Applications and Interactions

With input capability somewhat analogous to a touchscreen, the application space of Skinput could encompass many of the interactions and applications seen in today's touchscreen devices. To demonstrate this ability, we built several prototype interfaces that demonstrated Skinput's ability to appropriate the human body, in this case the arm, and use it as an interactive surface. In the first interface, we projected a series of buttons onto the forearm, on which a user can tap to navigate a hierarchical menu (Figure 3.9). In the second interface, we project a scrolling menu (Figure 3.5), which a user can navigate by tapping at the top or bottom to scroll up and down one item. Tapping on the selected item activates it. In a third interface, we project a numeric keypad on a user's palm and allow them to e.g., dial a phone number (Figure 3.10). Finally, as a true test of real-time control, we ported Tetris and Frogger to the hand, with controls bound to different fingertips (Figure 3.11).

Fig. 3.9 **Button based hierarchical menu.**

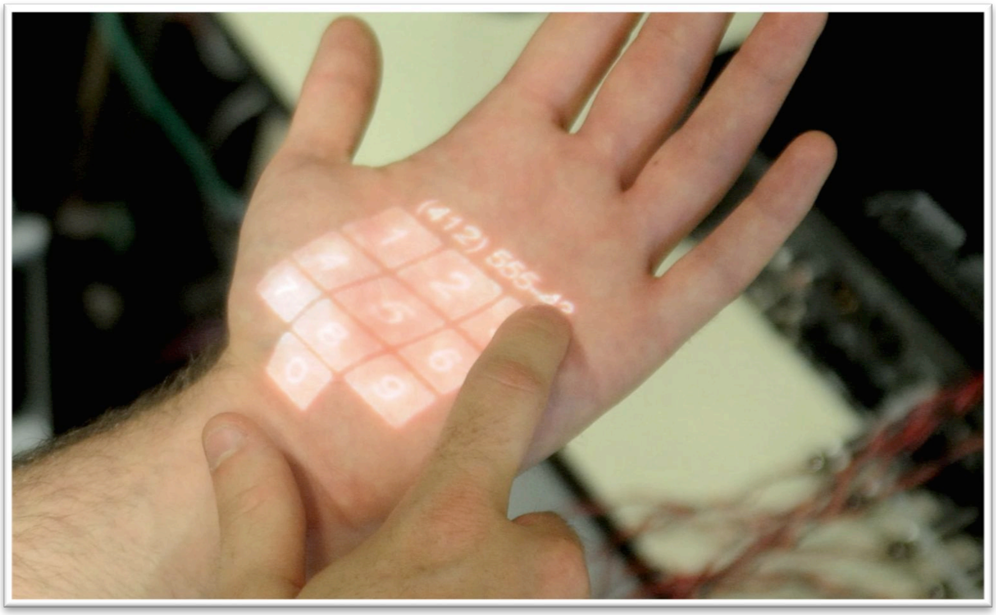


Fig. 3.10 **A numeric keypad for entering a phone number.**

Fig. 3.11 **Skinput is able to support real-time games, including Tetris (left) and Frogger (right).**

3.9 Conclusion

In this chapter, I presented a novel approach to appropriating the human body as an input surface. I described a wearable bio-acoustic sensing array that I built in the form of an armband. This setup could detect and localize finger taps on the forearm and hand in real time. Results from our experiments demonstrate that Skinput can operate even when the body is in motion. However, accuracy is insufficient for practical use at present. I concluded with brief descriptions of several prototype applications that demonstrate the rich design space that Skinput enables.

4 OMNITOUCH: MULTITOUCH INTERACTION EVERYWHERE

A central component of my dissertation work was to explore sensing methods beyond the bio-acoustic approach used in Skinput, which while successful, had significant limitations with respect to input accuracy and capability. My next project – OmniTouch – sought to expand the scope and capability of on-body interfaces, as well as improve their robustness. Whereas Skinput used acoustic sensing, OmniTouch is instead driven by computer vision, taking advantage of a special short-range depth camera and pico-projector worn on the upper body. A key contribution is a novel, depth-driven, elastic template matching and clustering approach to multitouch finger tracking. This enables on-the-go interactive capabilities, with no calibration, training or instrumentation of the environment or the user, creating an always-available interface.

4.1 Prototype Hardware

OmniTouch is a wearable system that enables graphical, interactive, multitouch input on arbitrary, everyday surfaces. Our shoulder-worn implementation allows users to manipulate interfaces projected onto the environment (e.g., walls, tables), held objects (e.g., notepads, books), and their own bodies (e.g., hands, lap).

Our proof-of-concept implementation (Figures 4.1 and 4.2) consists of three principal components. First is a custom, short-range depth camera, which provides a 320x240 depth map at 30 FPS [PrimeSense]. Objects as close as 20cm can be imaged by this sensor, with error in the Z axis (depth) of approximately 5mm. Depth accuracy decreases and noise increases at larger distances. However, for our application, which chiefly considers interaction within a 1m "bubble" in front of the user, noise and accuracy loss was minimal.

The second key component is a Microvision ShowWX+ scanned-laser pico-projector [Microvision]. This projector has the important property of wide angle, focus-free projection of graphical elements regardless of depth (i.e., distance from projector). Finally, the depth camera and projector are tethered to a conventional computer for prototyping purposes.

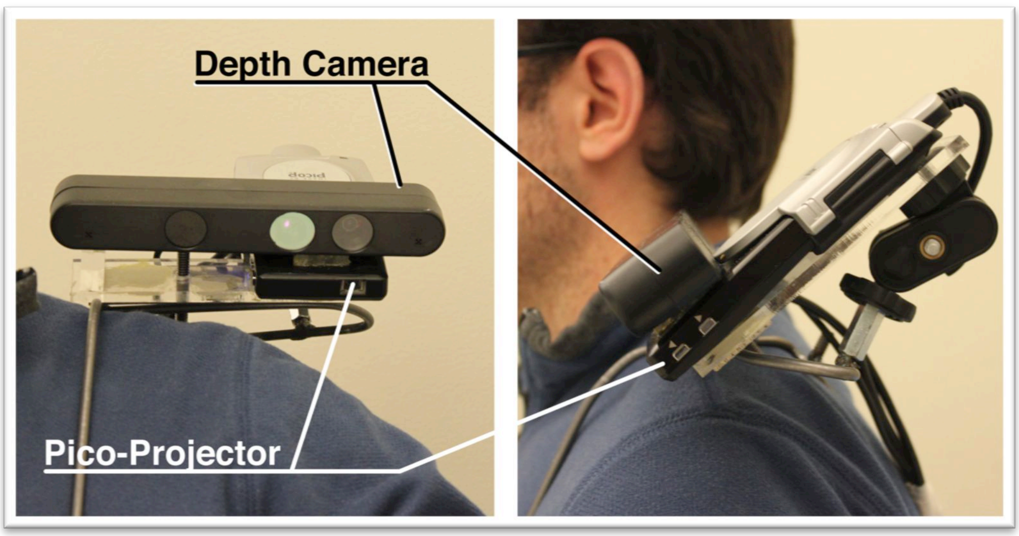


Fig. 4.1 **Our prototype shoulder-worn OmniTouch System. This setup was used for the evaluation.**

Initially, the depth camera and projector were rigidly mounted to a form-fitting metal frame, which was worn on the shoulders, and secured with a chest strap (Figure 4.1). Later, we constructed an updated, smaller version that attached to a shoulder strap of a bag (Figure 4.2). We chose the shoulder as it provides a good vantage point (both for sensing and projection) of the arms and held objects, as well as proximate fixed surfaces, such as walls and tables. However, our approach is amenable to other locations, including the upper arm [Harrison 2010], chest [Starner 2000], and wrist [Ni 2009].

Additionally, the shoulders tend to be very stable, allowing for projected interfaces with minimal sway and jitter.

The first person body-stabilized perspective is desirable for sensing and processing, as many simplifying assumptions can be made about the location and orientation of fingers and hands. For example, it is physically impossible for the user's arms to enter the image from the top. Additionally, the system's field of view naturally translates with the wearer. Moreover, camera and projection occlusion issues are minimized, as their fields of view roughly coincide with the wearer's line of sight.

Fig. 4.2 **A later OmniTouch prototype that attached to the strap of a bag.**

4.2 Multitouch Finger Tracking

We present a unique approach to ad hoc finger tracking that enables multitouch input on arbitrary surfaces, both flat and irregular, with no calibration or training. We can resolve the 3D position of fingers, and whether they are touching or hovering over a surface. Thus, OmniTouch produces input events similar to that of mice or touchscreens, enabling a wide variety of applications.

4.2.1 Finger Segmentation

Identifying finger input is a multistep process. First, we take a depth map of a scene (Figure 4.3 A) and compute the depth derivative in the X- and Y-axes using the average depth of a sliding 5x5 pixel window (Figure 4.3 B; X and Y derivative visualized using blue and red channels respectively). We then iterate over this derivative image, looking for vertical slices of cylinder-like objects. This is similar to template matching, but with some dynamic parameters. Put simply, for a slice of pixels to be a candidate, it must show a steep positive derivate, followed by a region of relative smoothness, and finally closed by a steep negative derivative (Figure 4.4). This ordering is critical, otherwise, concave features (e.g., gaps between fingers) would also be recognized. Also significant is that the depth camera we use represents sensing errors, out-of-range surfaces, and occlusion boundaries as (infinite) holes in the depth image. As such, they appear as concavities in the derivative, which our process ignores.

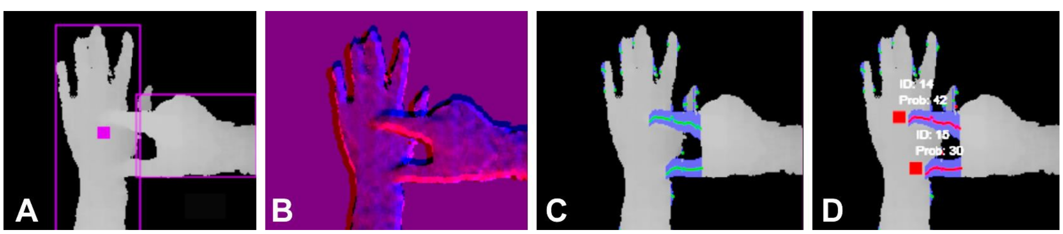


Fig. 4.3 **Left to right: depth map, derivative of depth map, finger slices overlaid in blue, path finding and tip estimation.**

To primarily isolate fingers, candidate slices must be between 5 and 25mm long – a range we found to cover typical finger diameters, including the critical fingertip. Pixel distances can be converted into real world distances (mm) because the depth value is also known. The result of this finger-slice identification process is shown in Figure 4.3 C.

Using the derivative of the depth map has several benefits that make it a key component of our sensing approach. Foremost, this approach suppresses absolute depth information, allowing the scene to be treated as a conventional 2D image, which is easier to process with standard computer vision techniques. Additionally, regardless of the surface the finger is operating on, the derivative profile is mostly invariant, greatly simplifying recognition.

Once all candidate finger slices are identified, we then greedily group proximate slices into contiguous paths. Paths that are shorter or longer than probable fingers are discarded. Even in noisy scenes, this process yields few false positives. The output, seen in Figure 4.3 D, resembles a skeletal model of the fingers. Like other computer vision techniques, fingers that are occluded are not detected. Additionally, and usefully, fingers that are "tucked in" are not tracked. However, our technique is sensitive to approach angle (can neither be too steep nor too shallow) and generally requires fingers be outstretched for reliable recognition.

Many approaches are possible for disambiguating which end of the path is the fingertip. In our proof-of-concept system, we assume a right-handed user, and thus, in almost all cases, the leftmost point in a path is the fingertip. This worked well in practice for our left-shoulder mounted configuration. To eliminate sensing noise and pixel-boundary flicker, fingertip positions are smoothed by a Kalman filter.

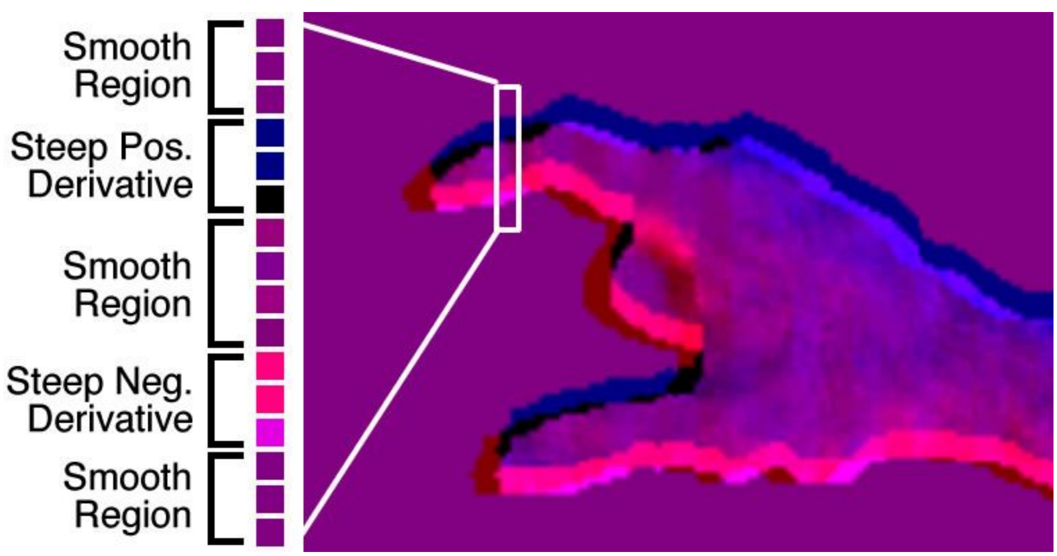


Fig. 4.4 **Close up example of a candidate finger slice.**

4.2.2 Finger Click Detection

The finger segmentation process, described above, yields the spatial location (X, Y and Z) of fingers. A secondary process is used to determine whether these fingers - specifically the tips - are in contact with a surface (i.e., a "click").

We start by computing the midpoint of the finger path, which roughly equates to the location of the minor knuckle. From this point, we flood fill towards the fingertip (i.e., all directions but rightward). This operation is performed on the depth map using a tolerance of 13mm in depth to determine if neighboring pixels can be filled. When the finger is hovering above a surface or in free space, the flood fill expands to encompass the entire finger (Figure 4.5, left). However, when the finger contacts a surface, the fill operation floods out into the connecting object (Figure 4.5, right). If a pixel count threshold is passed (e.g., 2000 pixels), the flood fill discontinues and the finger is determined to be clicked. Note that if the surface is very small or lies outside the camera's view, the threshold may not be passed, and the click missed.

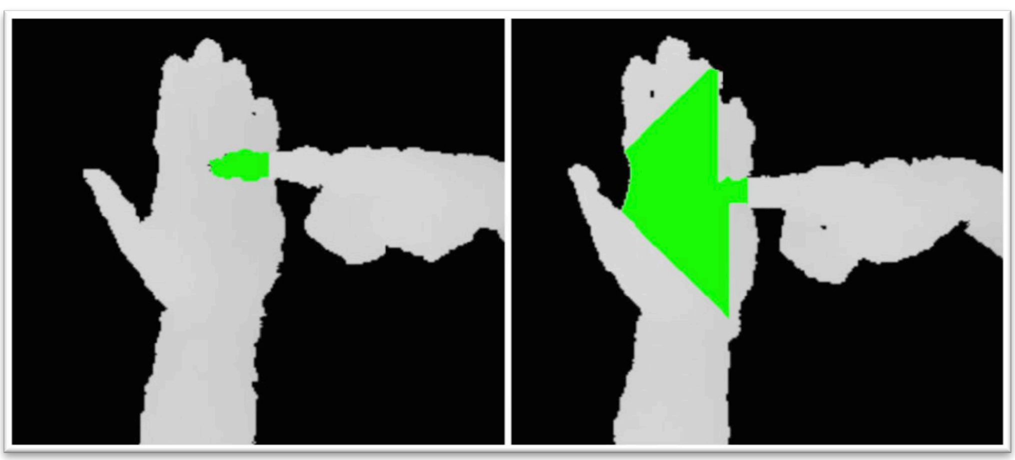


Fig. 4.5 Flood filling result when finger is hovering (left) and "clicked" (right).

This process detects finger clicks robustly, and also maintains a clicked state when dragging a finger across a surface, including irregular ones. In practice, a finger will be seen as "clicked" when its hover distance drops to 1cm or less above a surface; above 2cm is reliably seen as hovering. Hover distances between 1 and 2cm are ambiguous, and largely depend on local noise; we apply hysteresis to reduce flickering between click states. Anecdotally, users did not notice the ambiguity and generally "clicked through" this region on the way to their desired target.

4.3 On-Demand Projected Interfaces

With finger tracking alone, it is possible to support interfaces lacking graphical feedback, or "invisible interfaces" [Gustafson 2010]. For example, it would be possible to sketch simple figures or perform graffiti-like text entry on one's palm.

Infusing interactive graphical feedback expands the application space considerably. However, the inherent dynamic nature of the human body and objects in the real world makes this complex. Not only must interfaces track with surfaces they are rendered on, but they must also be projected in such a way as to account for their host surface's position and orientation in 3D space (Figure 4.6). Without these considerations, interfaces would be rendered with inappropriate position, orientation and size, and be subject to perspective visual distortions.

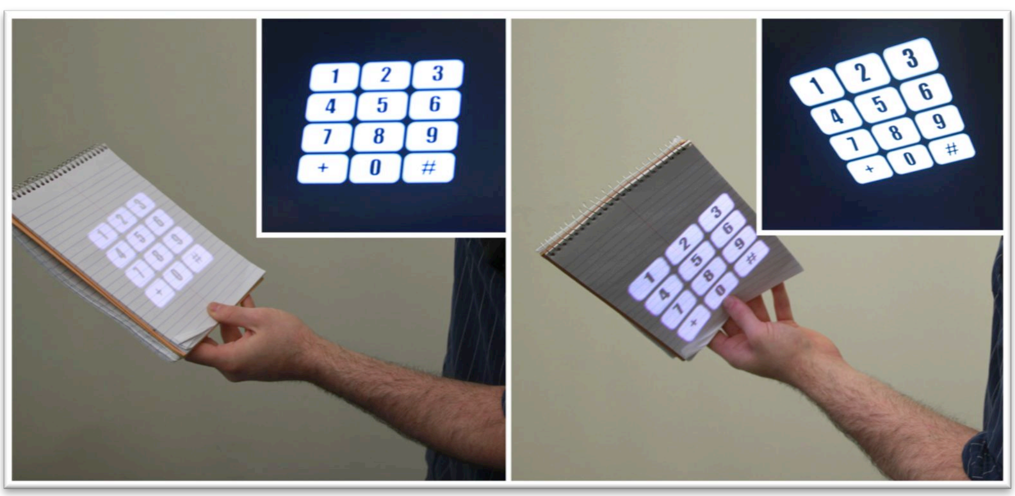


Fig. 4.6 In order for interfaces to appear visually aligned and correct when projected onto moving surfaces, the projected image must be dynamically pre-distorted (see inset images).

4.3.1 Surface Segmentation and Tracking

In addition to finger tracking, the depth video stream is also used to track surfaces suitable for projection in front of the user. First, distinct surfaces are segmented by performing a 3D connected components operation on the depth map (Figure 4.7, right). Surfaces smaller than hand size are discarded.

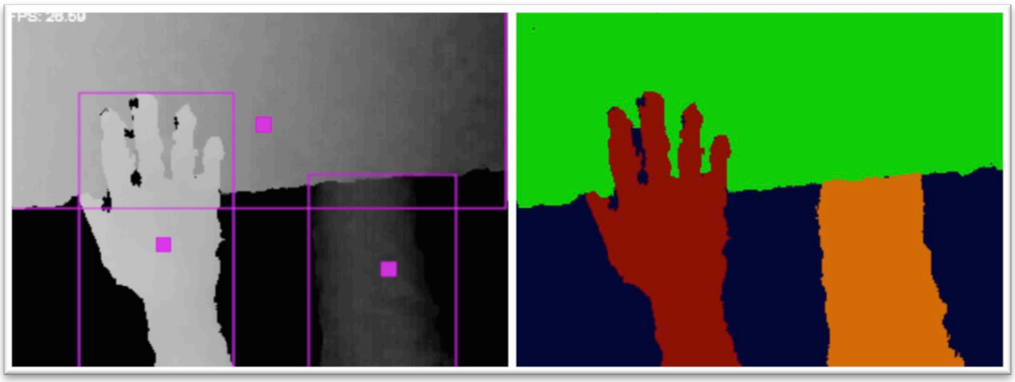


Fig. 4.7 3D connected components and their lock points.

For each surface, we compute the orientation about the Z-axis (orthogonal to the camera) by taking the covariance of the component's pixels in space, and computing the eigenvectors. Orientation about the X- and Y-axes is estimated using the distribution of surface normals, which typically peak over the primary orientations.

We also generate a central X/Y/Z "lock point", to which an interface can be attached (Figure 4.7, left, purple dots). This point must be stable regardless of translation and rotation in 3D space. One approach is to take the centroid of an object's pixels. However, because part of the surface may be occluded when the user is interacting with their fingers, this is not reliable. Instead, we move inwards 10cm along the surface's major axis from its upper extent, centered on the midpoint of the minor axis (Figure 4.8, red). Although more sophisticated techniques are possible, this solution worked well. Finally, a Kalman filter is used to smooth all six degrees of freedom.

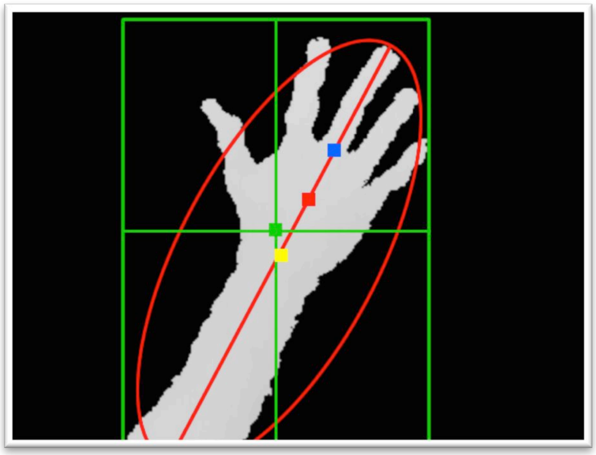


Fig. 4.8 **Possible lock points on the hand. Green: absolute center of surface bounds. Yellow: centroid of surface's pixels. Red: 10cm offset along major axis from upper extent. Blue: midpoint between wrist and middle finger tip.**

4.3.2 Projector/Camera Calibration

To enable authoring and interaction with projected interfaces, it is necessary to calibrate the projector and camera in a unified 3D space. Since our depth camera reports real-world depth values (mm), we chose that as our target coordinate system and calibrate the projector using camera values.

The process requires the intrinsic parameters of the projector, such as the field of view and the center of projection. To find the extrinsic projector parameters we require four non-coplanar calibration points. These four points must be identified by the depth camera and located in the projector image. Once the correspondence of the 2D points in the projected image and their actual 3D location in space (depth camera value) is established, we use the POSIT algorithm [DeMenthon 1995; Wilson 2010] to find the position and orientation of the projector. Note that this calibration only needs to be performed once, since the spatial relationship between the projector and the camera is fixed (i.e., both are mounted to a rigid frame).

4.3.3 Summoning and Defining Interactive Areas

Determining where to place an interface and how large it should be is non-trivial. For example, consider the hand: Do we center the interface in the middle of the palm, or the centroid of the surface? Or the midpoint between the wrist and finger tips? Or the absolute center of the bounds of the hand? Figure 4.8 illustrates these four (of many possible) options. Sizing interfaces has similar challenges: do we fit an interface to just the palm (which is attractive due to its relative flatness), the hand minus the thumb, or the full extent of the hand?

Previous on-body projected interfaces [Mistry 2009; McFarlane 2009], including Skinput, used a fixed-sized interface at a fixed image location. In order to use such an interface, a user must raise a physical object into this region at a specified distance, or walk up to a wall. This places the interface localization burden entirely on the user and is ill suited for many on-the-go mobile scenarios. In contrast, OmniTouch implements three distinct approaches to define, present, and track interactive areas:

4.3.3.1 One Size Fits All

OmniTouch can use a surface's lock point and orientation to provide an interface that tracks with a surface. However, because the bounds of the object in 3D space are unknown, the interface can only be as big as the smallest likely surface (generally the hand). Thus, even when a projecting on a large table, the interface will still be hand-sized. Additionally, every surface must use a generic lock point, which can lead to sub-optimal centering on asymmetric and organic surfaces, such as the hands. These drawbacks motivated us to explore more sophisticated options.

4.3.3.2 Classification-Driven Placement

Classification-driven placement consists of two stages. First, the system differentiates between a small set of surfaces by performing surface classification. Second, the system automatically sizes, positions and tracks an interface given the available projection area and heuristics describing the appropriate location for that surface.

We perform surface classification among a set of five common surfaces (hand, arm, pad, wall, and table) by considering a variety of features derived from each surface's depth image. For example, to distinguish between planar and organic surfaces, we calculate the standard deviation of the surface normals. Planar objects inherently have a majority of their normals pointing in a common direction, yielding a low standard deviation. On the other hand, organic surfaces tend to be more "rounded" (often symmetrically so), leading to diverse distributions and higher standard deviations. Size is very also descriptive; depth data allows for reasonable approximation of real world size - a notepad is easily distinguished from a table. Additionally, aggregate surface orientation immediately disambiguates tables from walls. These simple features worked well in our prototype implementation given the small set of surfaces to distinguish, but a more general solution would require more sophisticated features (e.g., see [Lai 2011] for depth-driven object recognition).

Each class of surface defines a unique interface placement heuristic (an offset vector from the surface's lock point) and default size. For example, a hand has a hand-sized interface, while a wall has a wall-sized interface. Lastly, once the surface is identified and the interface is placed, we track the surface change frame-to-frame and accordingly adjust the interface to reflect this change. This mimics the expectation of the user: once an interface is established, it should remain "glued" to the surface it is projected on. Optionally, given real-world depth data, the interface can be further refined and fitted to the available area on the surface, by performing depth-constrained flood filling from the interface's placement point.

Unfortunately, this classification-driven approach suffers from scalability issues, since it is simply not possible to build a classifier for every conceivable surface. However, for common surfaces that have unique placement considerations, this approach is attractive and viable.

4.3.3.3 User-Specified Placement

An entirely different approach is to let the user define the interactive area. This sidesteps much of the complexity described above, as users have a good innate sense of where interfaces should be centered and how big they should be. This exposes a high level of customization to users. However, this flexibility comes at the expense of requiring additional user interaction before an interface can be utilized.

In our prototype system, we provide two mechanisms for user specified placement, although many options are possible. The simplest is for a user to "click" on a surface, causing a generically sized interface to be centered at that location. Alternatively, a user can click and drag to position and size in one continuous action (Figure 4.9). As with the classification-driven approach, once the interface is established, we update its location and orientation frame-by-frame.

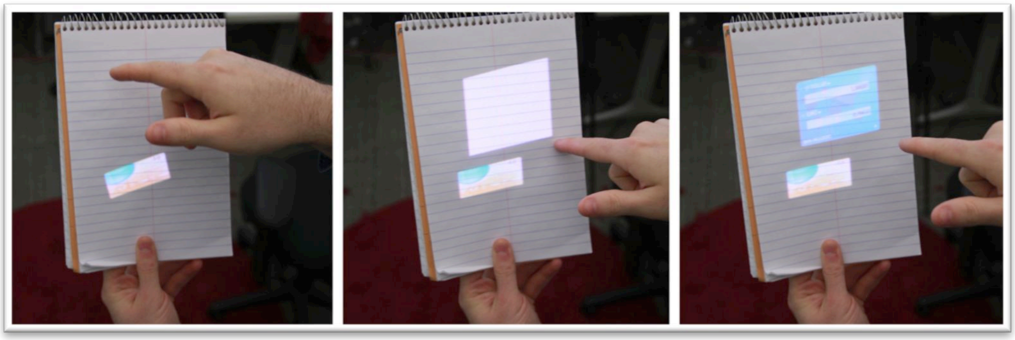


Fig. 4.9 To sidestep complexities in automatically positioning and sizing interfaces, users can simply "click-and-drag" interfaces wherever desired.

4.3.4 Compositing Interfaces in 3D Space

In our proof-of-concept implementation, we model interfaces as planar 2D surfaces, which are positioned and oriented in 3D space. Their 3D placement is computed in relation to the aforementioned lock points and surface orientations, so that they are correctly updated as surfaces move. Displaying such interfaces on top of any available surface is straightforward since our projector is precisely calibrated to the depth camera coordinate system. We simply create a 3D scene containing all active surfaces and then render this scene from the perspective of the projector using the projector/camera calibration discussed earlier (Figure 4.6). Although we currently render only planar interfaces, our approach easily lends itself to experimenting with 3D interfaces that take into account the true geometry of the projected surface.

By defining our interfaces in the 3D world space (i.e., using millimeters), they are projected with correct scale and distortion regardless of where the surface is with respect to the camera (as long as it is visible). Our aim is that the interactive surfaces appear to the user as “glued” to the physical surface. 3D rendering also automatically takes into account the Z-ordering of our interfaces.

Simultaneously, we use the 3D scene to ray cast fingertip positions onto our planar interfaces. Finger inputs are reported as X/Y coordinates in their local 2D space, which simplifies interface development and enables detection and tracking of finger hover. Although it is possible to use Z distance for click detection, we found our flood-fill heuristic approach to be most accurate.

4.4 Evaluation

To evaluate and demonstrate the feasibility of our approach, we conducted a user study that sought to quantify the key performance characteristics of OmniTouch. At a high level, can the system correctly register touch events and how accurately can they be localized? At a meta-level, how large would interface elements have to be to enable reliable operation of an ad hoc interface rendered on the hand? To place our system's performance in context, we compare our method to the gold standard - capacitive touch screens - drawing performance results from the literature [Holz 2010, Lewis 1993, Sears 1991].

4.4.1 Participants

We recruited 12 participants from our local metropolitan area (6 female), ranging in age from 23 to 49, with a mean of 34. All participants were right handed and were required to have some experience with touchscreen devices. The study took approximately one hour and included a gratuity.

4.4.2 Test Surfaces

Our goal with OmniTouch was to support interaction on three classes of surface: 1) on-body, 2) objects held in the hands, and 3) fixed surfaces in the environment. For our user study, we included one example from each class: the hand, a note pad held in the hand, and a wall. Additionally, we included the forearm (arm), as on-body interaction was a particular focus of the work and also challenging from a sensing perspective. Moreover, the arm served as a nice contrast to the hand, which, although highly irregular, is still fairly planar. Finally, these four surfaces, seen in Figure 4.10, represent ad hoc surfaces our system would likely use.

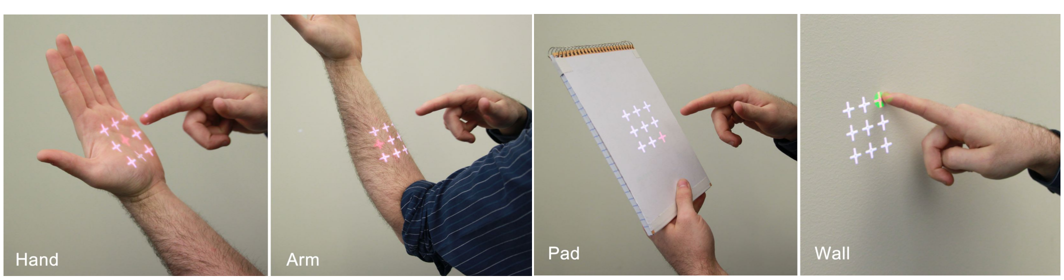


Fig. 4.10 The four surfaces we tested and user click distributions.

4.4.3 Procedure

We first fit participants with our shoulder-mounted system (Figure 4.1). Once the frame was secured and comfortable, participants were allowed to play with a simple, phone keypad example application (Figure 4.15). This let them find comfortable positions to hold their arms, both for being projected on and for pointing, and also to practice using the system. During this period, the experimenters provided feedback to help participants become more accurate. This training period lasted a maximum of 10 minutes, though most participants felt confident using the system after just a few minutes of use.

Our primary user study interface consisted of nine crosshair targets, laid out in a 3x3 pattern (Figure 4.10). Columns and rows were spaced 3cm apart; the crosshairs were 2x2cm in size. In each trial, a single crosshair was rendered in red. Users "clicked" this crosshair as accurately as they could with a finger. If a click was detected, the system would beep and a green circle was placed around the target crosshair (see Figure 4.10, wall). The experimenter advanced the interface to the next trial after each click attempt, regardless of whether or not it was detected. Each of the nine crosshair locations was repeated four times, for a total of 36 click trials; presentation order was randomized.

This interface and procedure was used for each of the four test surfaces: hand, arm, pad and wall. Before each surface, users were allowed to briefly practice before data collection began. For the wall condition, participants were asked to stand approximately 30cm from the wall. For the other three surfaces, users found a comfortable position. The ordering of the surfaces was randomized to compensate for any order effects. We ran two rounds of data collection to investigate if there were any effects from learning, fatigue, or slight variations in posture. This produced 288 trials (2 rounds x 4 surfaces x 36 click trials) per participant.

To quantify how our system performed at different distances, we included two additional rounds of data collection. Participants were asked to hold their hands at arm length (far), at an "average and comfortable" distance (average), and as close to the system as possible, while still being able to click with their other hand (close). We also tested these three distances with the pad surface; the ordering of the pad and hand distance trials was alternated between participants. This procedure produced 216 trials (2 surfaces x 3 distances x 36 click trials) per participant.

The tests described above were primarily designed to isolate click segmentation and spatial accuracy. A key feature of OmniTouch is its ability to track fingers while dragging. To better understand the spatial performance of finger drags, we created a drawing experiment interface (Figure 4.14 I). In this application, users were presented one of six possible shapes: up line, down line, left line, right line, clockwise circle, counterclockwise circle. Each shape was repeated four times, for a total of 24 drawing trials per participant.

Direction of the stroke was indicated using a green arrow and a red "stop" mark. Participants were asked to draw as closely to the white path as possible, balancing speed and accuracy. Unlike in the crosshair experiments, users received graphical feedback in the form of a red path illustrating their stroke. This allowed participants to compensate for any inaccuracies in their movement and the system's fingertip estimation. We chose to conduct this experiment on the pad, as a flat surface minimized external confounds (e.g., user inaccuracy caused by the irregular surface of the hands).

4.5 Results

Our 12 participants produced 3456 click trials on our four surfaces, a further 2592 in our distance experiment, and 288 drawn shapes. No effect was found between the two rounds of crosshair trials (e.g., from fatigue or learning). Additionally, there were no significant performance differences between participants within any surface. Thus, participant was removed as a factor in our analyses. Data from the distance and the dragging trials was kept separate for independent analysis. Ultimately, these results should be considered a performance baseline, as significant improvements in depth camera resolution and sensitivity are forthcoming.

4.5.1 Finger Click Detection

We combined data from all of our crosshair-clicking experiments (two rounds of four surfaces and two rounds of three distances) – a total of 6048 click trials. Of these, 96.5% correctly received exactly one finger click event. Regarding errors, 50 trials (0.8%) had no click event (i.e., the system missed the participant's finger click), 154 trials (2.5%) had two click events (i.e., the system incorrectly thought the user clicked twice, or believed a secondary finger to have clicked), and 8 trials (0.1%) had three click events.

For the user study, we configured OmniTouch to record all input events, without any high-level mechanism for click rejection, as typically found in interactive systems. Of the 162 trials receiving double and triple clicks, 94.8% percent occurred within 500ms of the first click event. Thus, with a simple timeout, single finger click segmentation accuracy would be 98.9%.

Of the 50 trials (0.8%) with missed clicks, 33 were contributed by the three left-most crosshairs in the arm condition (see Figure 4.10, arm). As noted in [Roudaut 2011], participants tend to hook their fingers when targeting items on reverse slopes, which is the case for the right hand targeting the left most side of the left forearm. One possible explanation for this increased error is that hooking occludes the contact point and also shortens the finger's profile from the camera's perspective, which can cause tracking loss. Otherwise, the distribution of missed-click and multi-click errors was evenly spread over all crosshair positions and surface conditions.

Finally, we compared click segmentation performance at the three distances tested in the user study (hand/pad surfaces at close/average/far distances). However, no significant effects were found.

4.5.2 Finger Click Spatial Accuracy

Importantly, our results represent the cumulative error of the system and the user. There are three primary sources of error: 1) misalignment and non-linearities in the projector/camera calibration (e.g., a button is projected somewhere slightly different from where the camera believes it to be), 2) inaccuracy in the fingertip estimation, especially when the tip fuses with the surface during clicks, and 3) user inaccuracy when clicking targets, (e.g., due to "fat fingers" and varying perception of one's finger input point [Holz 2010]). Although some of these factors are outside of our control, they model the real-world performance of our system.

There are two important and independent measures for analyzing targeting performance: offset and spread [Chapanis 1951; Holz 2010; Sears 1991].

4.5.2.1 Finger Click Spatial Offset

Analysis revealed there was a small systematic offset between where OmniTouch believed the user clicked and where the user believed they clicked. Specifically, we found an average offset of 11.7mm to the left of targets across all conditions and participants, in agreement with previous findings in the touchscreen literature [Lewis 1993; Sears 1991]. Y-offset for the hand, arm and pad surfaces was similarly an average of 1.1mm above the true target. Finger touches on the wall, however, were offset downwards an average 10.0mm, possibly due to its extreme angle (at roughly chest height, and oriented vertically). Finally, distance appears to have no significant effect on offset.

Because offsets are systematic across-users and across-surfaces, we simply apply a single post-hoc X/Y offset to our subsequent data analysis. These offset values could be trivially added to our system's real-time finger point estimations. The only case we handle specially is the wall, which is recognized by OmniTouch using size and orientation information (see section 4.3.3.2). With the wall, points are shifted upward 10.0mm. For maximum generality, we did not compute or apply any per-user offset, though this has been shown to significantly increase accuracy [Holz 2010; Wang 2009].

4.5.2.2 Finger Click Spatial Precision

The spatial precision of OmniTouch is visualized in Figure 4.11, which depicts 95% confidence ellipses for the nine crosshair targets on our four test surfaces. For analysis, we removed 93 outliers (1.5% of our click trials) in order to plot our results side-by-side with those in [Holz 2010] (Figure 4.12). Outliers were defined as points lying greater than three standard deviations away from the mean difference between points and the intended target. Similar to [Holz 2010], outliers were a mix of user error, user inaccuracy, and tracking errors.

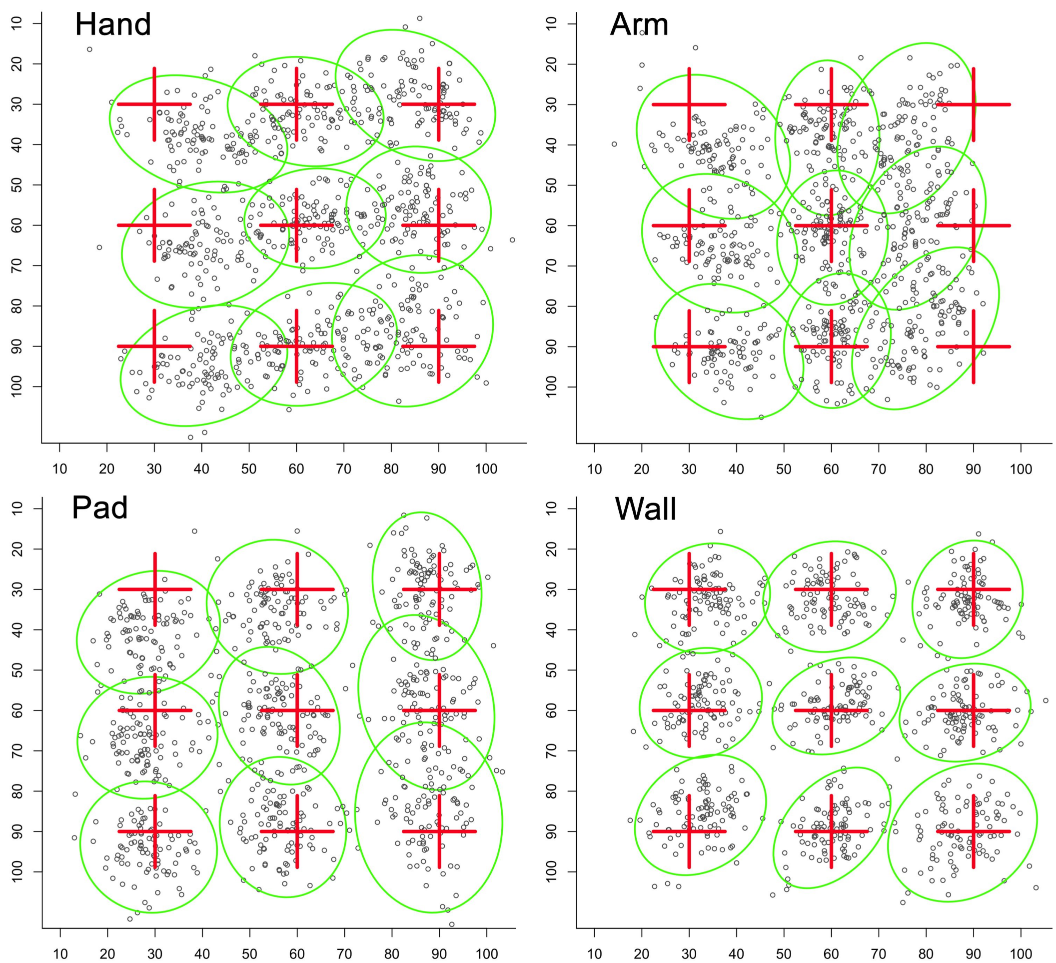


Fig. 4.11 Distribution of clicks from all users. Crosshairs, true to size and location, are shown in red. 95% confidence ellipses are shown in green. Axes units in mm.

Figure 4.12 displays the minimum button diameter necessary to correctly capture 95% of touches for each surface. We also include two points of comparison from [Holz 2010] - an estimation of conventional touch input (derived from a capacitive touchpad) and results from crosshair trials using a high-resolution optical fingerprint scanner. Exceeding our expectations, OmniTouch on a wall appears to be nearly as accurate as conventional touchscreens (16.2mm vs. 15.0mm). The hand requires buttons to be 22.5mm in diameter; the pad performs similarly to the hand.

The arm is our least accurate surface, requiring targets be approximately 70% larger than a conventional touchscreen to achieve the same 95% touch accuracy (25.7mm vs. 15.0mm). This degradation in error comes chiefly from buttons located on the sides of the arms, where curvature is high (see Figure 4.10 and 4.11). Given that the arm is well suited to narrow, tall interfaces (Figure 4.14 C), we also computed the accuracy using only the center column of crosshair targets, which proved to be quite accurate (20.5mm, SD=5.1mm).

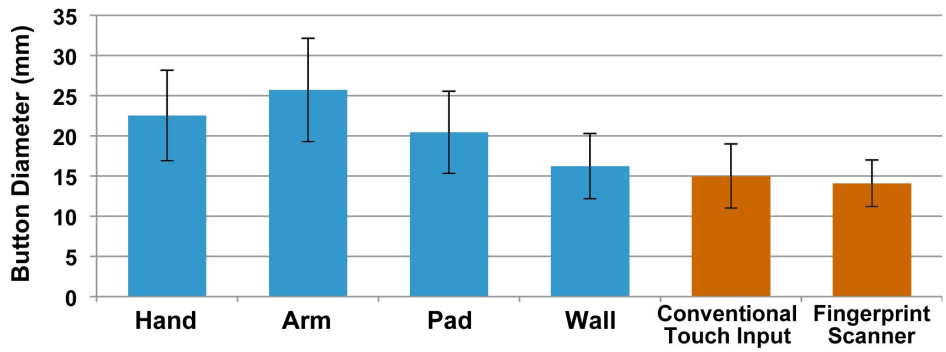


Fig. 4.12 **Button diameter necessary to encompass 95% of touches. Error bars denote standard deviation across all trials. Results in orange from [Holz 2010].**

4.5.2.3 Effects of Distance on Spatial Precision

We previously reported that distance had no significant effect on click segmentation accuracy or on spatial offset. However, there does appear to be a significant loss of precision when interacting at far distances. Using a Bonferroni-corrected all-pairs t-test, we found no significant difference in performance between the hand and pad at our three test distances. We then combined hand and pad distance trials into aggregated far, average and close data sets. Overall, the far condition is significantly worse performing than both average and close distances (both $p<.001$) – requiring buttons to be roughly 60% larger to operate at around arm's length. There is no significant difference between average and close distances. Figure 4.13 illustrates these results as button diameters necessary to encompass 95% of touches.

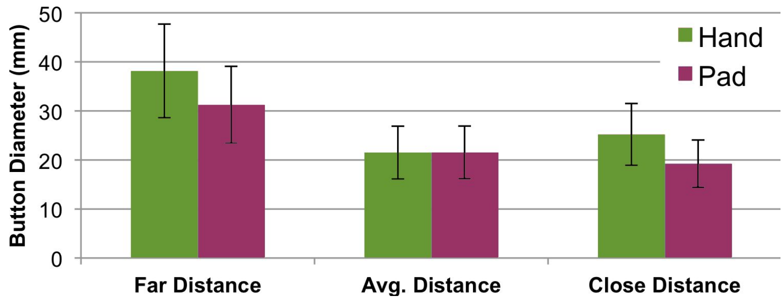


Fig. 4.13 **Button diameter needed to encompass 95% of touches. Error bars show standard deviation across trials.**

4.5.3 Finger Drag Spatial Accuracy

In our dragging (drawing) experiment (Figure 4.14 I), we included vertical and horizontal lines, as well as circles, in order to assess X, Y and X+Y dragging performance. For both lines and circles, we use the absolute Euclidian distance from the closest point on the desired path as our error distance metric. We did not apply our global X/Y offset to stroke points as live graphical feedback was provided in situ. Such feedback allowed participants to compensate for any system inaccuracies as they performed the task.

On average, participants deviated from the desired path by just 6.3mm (mean SD=3.9mm). There is no significant performance difference between shapes, 1D or 2D trials, or in X- or Y-axes.

4.6 Example Applications and Interactions

With OmniTouch providing capabilities similar to that of mice and touchscreens, the application space is expansive. Over the course of development, we created many small, interactive applications that ran on top of our OmniTouch engine (Figure 4.14). These served both as proof of concept and also as a gauge of input accuracy and real world applicability. We briefly describe a few exemplary applications.

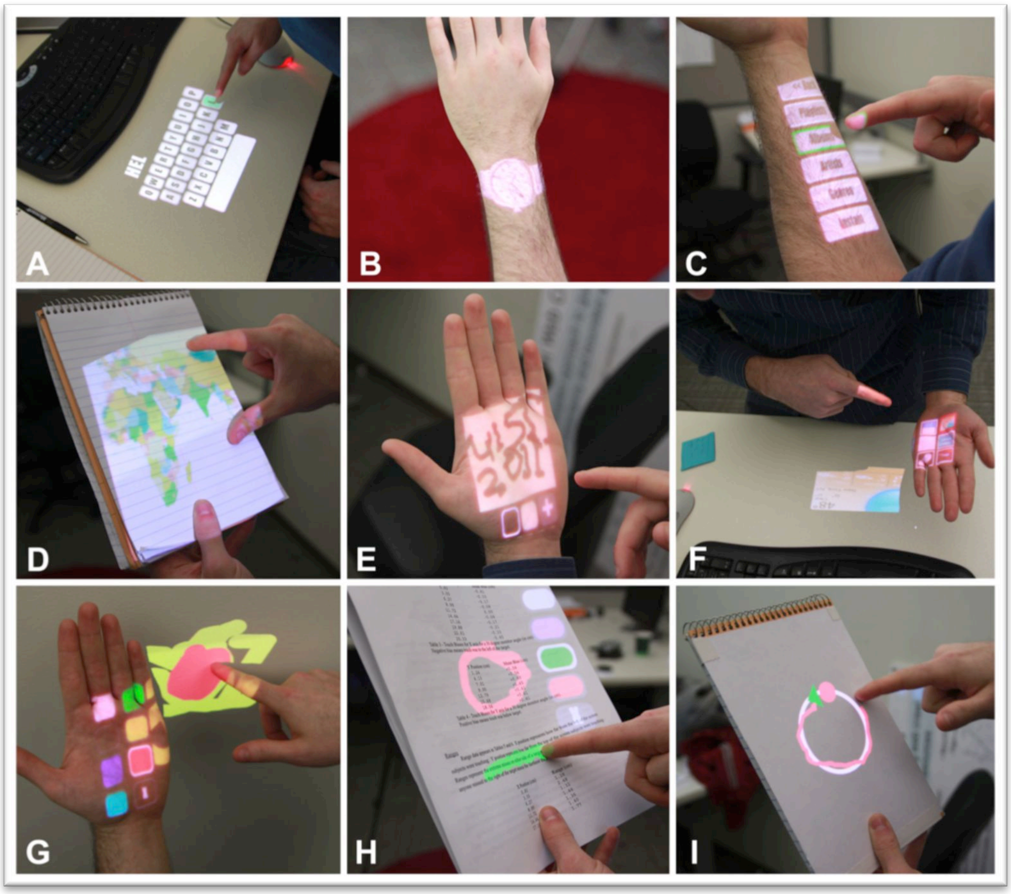


Fig. 4.14 **Some of the applications we developed that run on top of OmniTouch. See text for descriptions.**

4.6.1 Conventional Interaction

With the ability to click, buttoned interfaces are immediately possible; dragging enables interfaces to be scrolled. As a simple demonstration, we built a phone keypad application (Figure 4.15). To prevent accidental dialing, a “slide to unlock” feature is included. We also experimented with hierarchical menu navigation, visualized as a scrollable list with clickable items (Figure 4.14 C). This class of interface is prevalent in contemporary mobile devices. Additionally, we built a full keyboard, potentially allowing for text entry on the go (Figure 4.14 A). Lastly, a “post-it” application allows users to write quick notes on their palm (Figure 4.14 E).

To showcase our system’s multitouch capabilities, we developed a simple annotation application, where each finger generates a stroke; a small palette of highlight colors is provided (Figure 14.4 H). We also implemented the ubiquitous map panning and zooming demo, which is controlled by finger drags and pinching respectively (Figure 4.14 D).

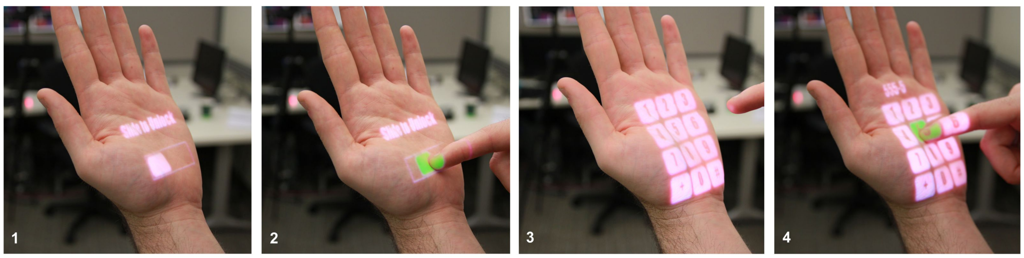


Fig. 4.15 **We created a simple phone keypad application. 1) User raises their hand and is presented a "slide to unlock" interface. 2) User drags the unlock widget. 3) A number keypad is presented. 4) User can dial a number.**

4.6.2 Multi Surface Interaction

As OmniTouch can track multiple objects within its field of view, it is possible to support interaction on multiple surfaces and levels. As a proof of concept, we created a painting application for walls. We use the left hand as the color pallet, which can be raised and lowered as needed (Figure 4.14 G). Similarly, when working at a table, the hand may serve as an application switcher (Figure 4.14 F).

4.7 Conclusion

In this chapter, I described and evaluated our proof-of-concept implementation of OmniTouch, a wearable depth-sensing and projection system that enables interactive multitouch applications on everyday surfaces. Beyond the shoulder-worn system, there is no instrumentation of the user or environment. The system allows the wearer to use their hands, arms and legs as graphical, interactive surfaces. Users can also transiently appropriate surfaces from the environment to expand the interactive area (e.g., books, walls, tables). On such surfaces - without any calibration - OmniTouch provides capabilities similar to that of a mouse or touchscreen: X and Y location in 2D interfaces and whether fingers are “clicked” or hovering, enabling a wide variety of interactions. Reliable operation on the hands, for example, requires buttons to be 2.3cm in diameter. With OmniTouch, it is conceivable that anything one can do on today's mobile devices, could now be done in the palm of one's hand.

5 TOUCHE: CAPACITIVE TOUCH SENSING ON THE BODY

So far, we have discussed Skinput and OmniTouch, which employed acoustic- and vision-driven approaches to appropriating the body for interactive purposes. We now round out our technical explorations with a third method that is electrical in nature.

Modern capacitive touchscreens, like those found in smartphones, rely on a (transparent) conductive layer covered by a thin insulator. Coincidently, humans are composed similarly – we have a conductive interior (liquids and soft tissues) surrounded by an insulating layer (skin). As such, we can use the body itself as a crude capacitive sensor to enable a variety of on-body gestures, which could be used to control, for example, a worn music player. To achieve this, we developed a novel and specialized form of capacitive sensing called Swept Frequency Capacitive Sensing. Experimental results show that gesture classification accuracies of around 95% are achievable. Although we do not couple Touché with projected graphical output, it has many of the properties necessary to support full-featured on-body interfaces.

5.1 Conventional Capacitive Sensing

The importance of touch and gestures has been long appreciated in the research and practice of human-computer interaction. There is a tremendous body of previous work related to touch, including the development of touch sensors and tactile displays, hand gesture tracking and recognition, designing interaction techniques and applications for touch, and building multitouch, tangible and flexible devices (see e.g., [Bau 2010; Dietz 2001; Lee 1985;

Paradiso 1999; Philipp 1999; Poupyrev 2003; Wimmer 2011] for a subset of previous work on touch).

The foundation for all touch interaction is touch sensing, i.e., technologies that capture human touch and gestures. This includes sensing touch using cameras or arrays of optical elements [Matsushita 1997], laser rangefinders [Cassinelli 2005], resistance and pressure sensors [Rosenberg 2009] and acoustics [Paradiso 2000], to name a few. The most relevant technology to Touché is capacitive touch sensing, a family of sensing techniques based on the same physical phenomenon: capacitive coupling.

The basic principles of operation in most common capacitive sensing techniques are quite similar: A periodic electrical signal is injected into an electrode forming an oscillating electrical field. As the user's hand approaches the electrode, a weak capacitive link is formed between the electrode and conductive physiological fluids inside the human hand, altering the signal supplied by the electrode. This happens because the user body introduces an additional path for the flow of charge, acting as a charge "sink" [Zimmerman 1995]. By measuring the degree of the signal change, touch events can be detected.

There is a wide variety of capacitive touch sensing techniques. One important design variable is the choice of signal property that is used to detect touch events. For example, changes in signal phase [Hinkley 1999] or signal amplitude [Barrett 2010; Philipp 1999; Rekimoto 2002] can be used for touch detection. The signal excitation technique is another important design variable. For example, the earliest capacitive proximity sensors in the 1970s oscillated at a resonant frequency and measured signal dumping; additional capacitance would affect the resonant frequency of the sensing circuit [Skulpone 1973], which could be measured. Variations in the topology of electrode layouts, the materials used for electrodes and substrates, and the specifics of signal measurement resulted in a multitude of capacitive techniques, including charge transfer, surface and projective capacitive approaches, among others [Barrett 2010; Philipp 1999].

Capacitive sensing is a malleable and inexpensive technology. Consequently, we find capacitive touch in millions of consumer device controls and touchscreens. It has, however, a number of limitations. One important limitation is that capacitive sensing is not particularly expressive. In general, it can only detect when a finger is touching the device and sometimes infer finger proximity. To increase the expressiveness, matrices of electrodes are scanned to create a 2D capacitive image [Dietz 2001; Lee 1985; Rekimoto 2002; Song 2011]. Such space multiplexing allows the device to capture spatial gestures, hand profiles [Rekimoto 2002] or even rough 3D shapes [Smith 1996]. However, this precludes ad hoc instrumentation of everyday objects and environments, as well as the human body.

Touché advocates a different approach to enhancing the expressivity of capacitive sensing – by using frequency multiplexing. Instead of using a single, pre-determined frequency, we sense touch by sweeping through a range of frequencies. We call this technique *Swept Frequency Capacitive Sensing* (SFCS). We refer to the resulting set of capacitive values as a *capacitive profile* and demonstrate their ability to support small gesture vocabularies using single electrodes. Further, the technique is sufficiently robust, that it can use the human body as an electrode, providing capacitive touch sensing on the skin.

5.2 Swept Frequency Capacitive Sensing

The human body is conductive, e.g., the average internal resistance of a human trunk is ~100 Ω [Webster 2010]. Skin, on the other hand, is highly resistive, ~1M Ω for dry undamaged skin [Webster 2010]. This would block any weak constant electrical (DC) signal applied to the body. Alternating current (AC) signal, however, passes through the skin, which forms a capacitive interface between the electrode and ionic physiological fluids inside the body [Foster 1996]. The body forms a charge "sink" with the signal flowing though tissues and bones to ground, which is also connected to the body through a capacitive link [Philipp 1999; Zimmerman 1995].

The resistive and capacitive properties of the human body oppose the applied AC signal. This opposition, or electrical impedance, changes the phase and amplitude of the AC signal. Thus, by measuring changes in the applied AC signal we can 1) detect the presence of a human body and also 2) learn about the internal composition of the body itself. This phenomenon, in its many variations, has been used since the 1960s in medical practice to measure the fluid composition of the human body [Foster 1996], in electro-impedance tomography imaging [Cheney 1999] and even to detect the ripeness of nectarine fruits [Harker 1994].

The amount of signal change depends on a variety of factors. Foremost, it is affected by how a person touches the electrode, e.g., the surface area of skin touching the electrode. It is also affected by the body's connection to ground, e.g., wearing or not wearing shoes or having one or both feet on the ground. Finally, it strongly depends on signal frequency. This is because at different frequencies, the AC signal will flow through different paths inside of the body [Foster 1996]. Indeed, just as DC signal flows through the path of least resistance, the AC signal will always flow through the path of least impedance.

The human body is anatomically complex and different tissues, e.g., muscle, fat and bones, have different resistive and capacitive properties. As the frequency of the AC signal changes, some of the tissues become more opposed to the flow of charges, while others less, thus changing the path of the signal flow (see [Foster 1996] for an overview of the bioelectrical aspects of human body impedance). Therefore, by sweeping through a range of frequencies, we obtain a wealth of information about 1) how the user is touching the electrode, 2) how the user is connected to ground and 3) the current configuration of the human body and individual body properties.

One of the reasons why SFCS techniques have not been investigated before could be due to computational expense: instead of sampling a single data point at a single frequency, SFCS requires a frequency sweep and analysis of hundreds of data points. Only recently, with the advance of fast and inexpensive microprocessors, has it become feasible to use SFCS in touch interfaces. Another challenge in using SFCS is that it requires high-frequency signals, e.g., ~3 Mhz. Designing conditioning circuitry for high-frequency signals is a complex problem. We will discuss these challenges and solutions in detail in the next section of this paper.

5.3 Implementation

The overall architecture of Touché is presented in Figure 5.1. The user interacts with an object that is attached to a Touché sensor board via a single wire. If the object itself is conductive, the wire can be attached directly to it. Otherwise, a single electrode has to be embedded into the object and the wire attached to this electrode.

Touché implements SFCS on a compact custom-built board powered by an ARM Cortex-M3 microprocessor (Figure 5.2). The on-board signal generator excites an electrode with sinusoid sweeps and measures returned signal at each frequency. The resulting sampled signal is a capacitive profile of the touch interaction. We stress that in the current version of Touché we do not measure phase changes of the signal in response to user interaction. We leave this for future work.

Finally, the capacitive profile is sent to a conventional computer over Bluetooth for classification. Recognized gestures can then be used to trigger different interactive functions. While it is possible to implement classification directly on the sensor board, a conventional computer provided more flexibility in fine-tuning and allowed for rapid prototyping.

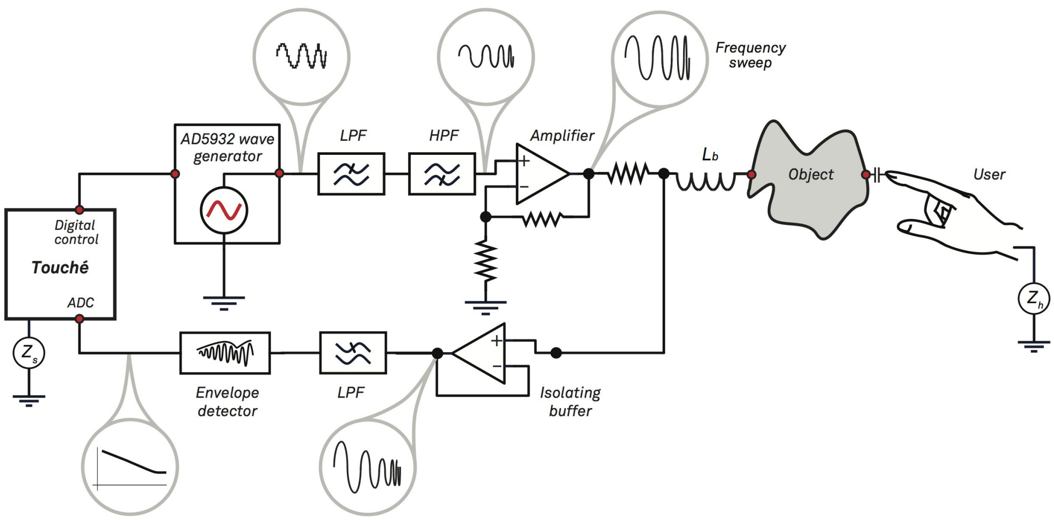


Fig. 5.1 **Swept Frequency Capacitive Sensing with Touché**

5.3.1 Sensor Board Design

An ARM microprocessor, NXP LPC1759 running at 120 MHz, controls an AD5932 programmable wave generator (Figure 5.2) to synthesize variable frequency sinusoidal signal sweeps from 1 KHz to 3.5 MHz in 17.5 KHz steps (i.e., 200 steps in each sweep, see Figure 5.1). The signal is filtered to remove environmental noise and undesirable high frequency components and is also amplified to 6.6 Vpp (Figure 5.3, A), which is then used to excite the attached conductive object. In the current design we tune Touché to sense very small variations of capacitance at lower excitation frequencies by adding a large bias inductor Lb (~100 mH), a technique used in impedance measurement. By replacing it with a bias capacitor, we can make Touché sensitive to very small inductive variations, e.g., copper coil stretching.

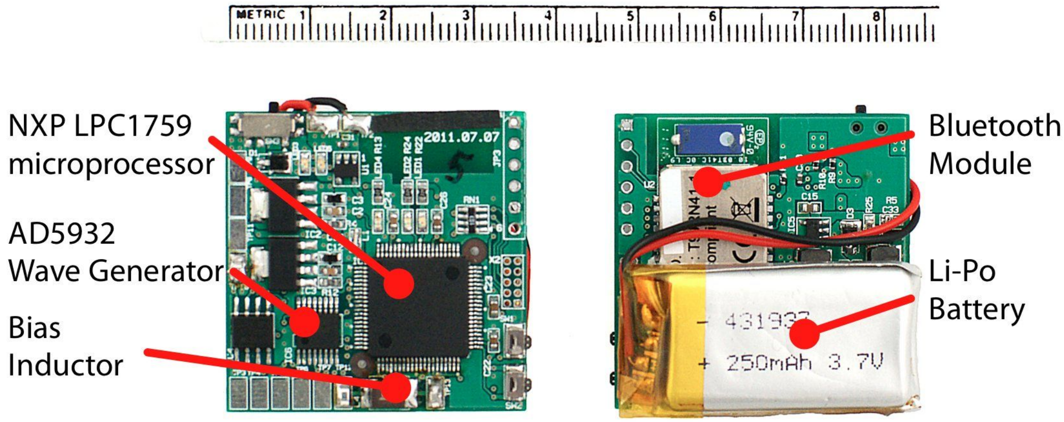


Fig. 5.2 Touché sensing board created by Munehiko Sato and Ivan Poupyrev, measuring 36x36x5.5 mm and weighing 13.8 grams.

The return signal from the object is measured by adding a small sensing resistor, which converts alternating current into an alternating voltage signal (Figure 5.3, B). This signal is then fed into a buffer to isolate sensing and excitation sections; an envelope detector then converts the AC signal into a time-varying DC signal (Figure 5.3, C). The microprocessor samples the signal at a maximum of 200kHz using a 12-bit analog-digital converter (ADC). A single sweep takes ~33ms, translating to a 33Hz update rate.

Currently, the sampling rate of ADC is a main limiting factor for speed: a dedicated ADC with a higher sampling rate would significantly increase the speed of Touché. Sampling is much slower at low frequencies, as it takes longer for the analogue circuitry to respond to a slowly varying signal. In applications where an object does not respond to low frequencies, we swept only in the high frequency range, tripling the sensor update rate to ~100Hz.

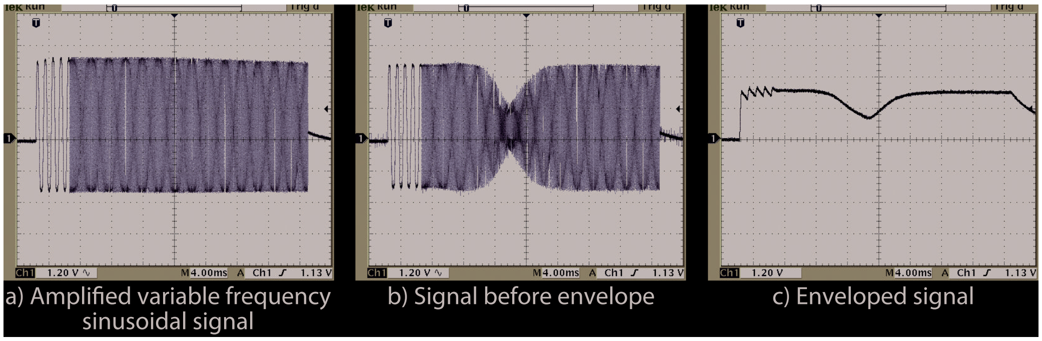


Fig. 5.3 Variable frequencies sweep and return signal.

5.3.2 Communication and Recognition

For classification, we use a Support Vector Machine (SVM) implementation provided by the Weka Toolkit [Hall 2009] (SMO, C=2.0, polynomial kernel, e=1.0) that runs on the aforementioned conventional computer. Each transmission from the sensor contains a 200-point capacitive profile, from which we extract a series of features for classification.

The raw impedance values from the frequency sweep have a natural high-order quality. As can be seen in Figures 5.5 and 5.6, the impendence profiles are highly continuous, distinctive and temporally stable. Therefore, we use all 200 values as features without any additional processing. Additionally, we compute the derivative of the impedance profile at three different levels of aliasing, down-sampling capacitive profiles into arrays of 10, 20, and 40, yielding another 70 features. This helps to capture shape features of the profile, independent of amplitude. For example, it is easy to see the peaks and minima in Figures 5.5 and 5.6 – more difficult to see is the visually subtle, but highly discriminative slope of the peaks. Moreover, using the derivative increases robustness to global variations in impendence, e.g., an offset of signal amplitude across all frequencies due to temperature variations. As a final feature, we include the capacitive profile minima, which was found to be highly characteristic in pilot studies (see Figures 5.5 and 5.6). Once the SVM has been trained, classification can proceed in a real-time fashion.

5.4 Example Applications and Interactions

Touché has many applications, both on and off the body. Relevant to this dissertation are two body-centric gesture sets that we created to explore interactions with Touché. We describe an evaluation of these two sets in Sections 5.5 and 5.6.

5.4.1 Touché Sensing Configurations

Touché has two basic sensor configurations. First is when the user is simply touching an object or electrode (Figure 5.4, A). This is the classic capacitive sensor configuration that assumes that both the sensor and the user are sharing common ground, even through different impedances. For example, if the sensor were powered from an electrical outlet, it would be connected to the ground line of a building. The user would be naturally coupled to the same ground via a capacitive link to the floor or building structure. Although this link may be weak, it is sufficient for Touché.

In the second case, the sensor is touching two different locations of the user body with its ground and signal electrodes (Figure 5.4, B). In this configuration, Touché measures the impedance between two body locations [Foster 1996].

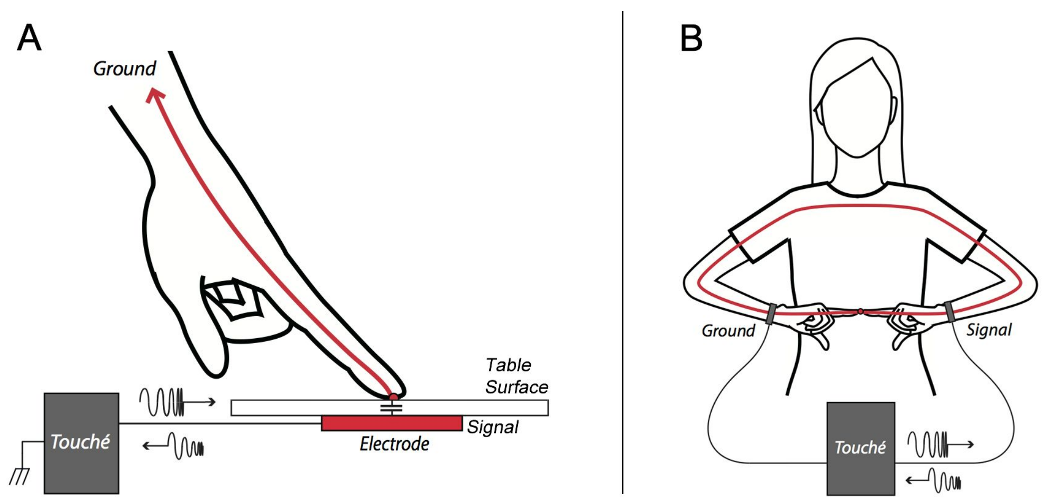


Fig. 5.4 Body-centric configurations of Touché.

5.4.2 Body Configuration Sensing

Touché can be used to sense the configuration of the human body. For example, a door could sense if a person is simply standing next to it, if they have raised their arm to knock on it, are pushing the door, or are leaning against it. Alternatively, a chair or a table could sense the posture of a seated person – reclined or leaning forward, whether the arms are on the armrests, if one or two arms operating on the surface, as well as their configuration (Figure 5.5). Importantly, sensing can occur without instrumenting the user.

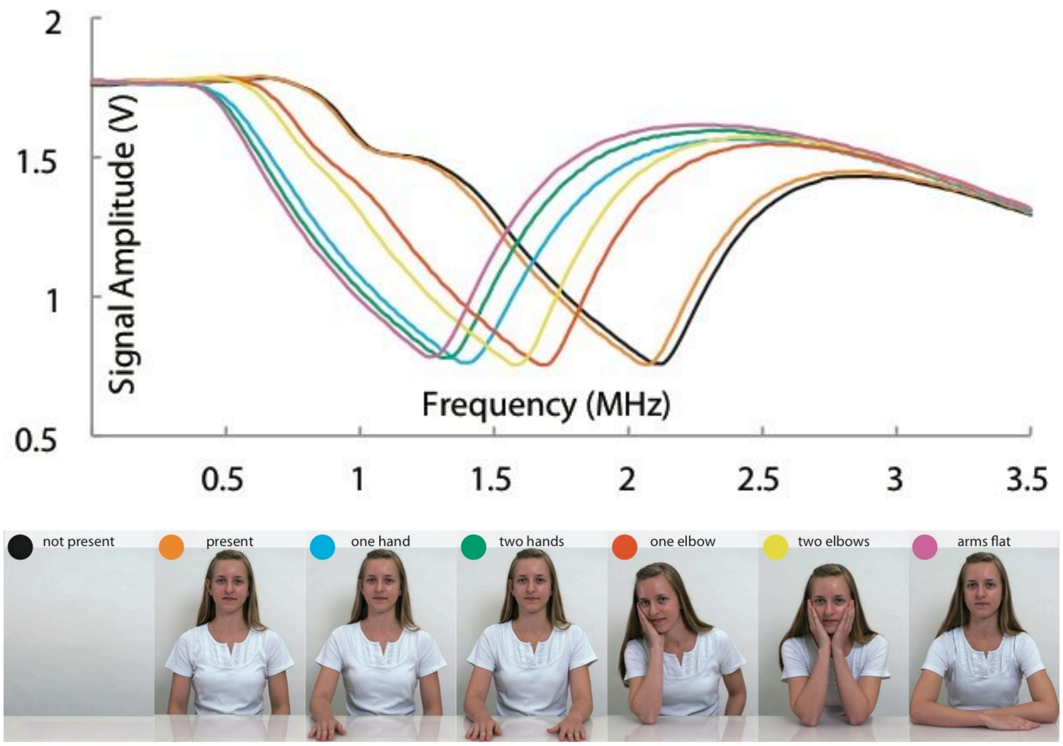


Fig. 5.5 **Capacitive profiles for sensing body postures on a Touché-augmented table.**

Sensing the pose of the human body without instrumenting the user has numerous compelling applications. Posture-sensing technologies are an active area of research, with applications in gaming, adaptive environments, smart offices, in-vehicle interaction, rehabilitation and many others [Forlizzi 2005]. We view Touché as one such enabling technology, with many exciting applications.

To evaluate the performance of Touché as a body sensor embedded in the environment, we constructed a sensing table, which could discern body posture (Figure 5.5). This consisted of a conventional table with a thin copper plate on top of it, covered with a 1.6 mm glass fiber and resin composite work board. A single wire connected the copper plate to the

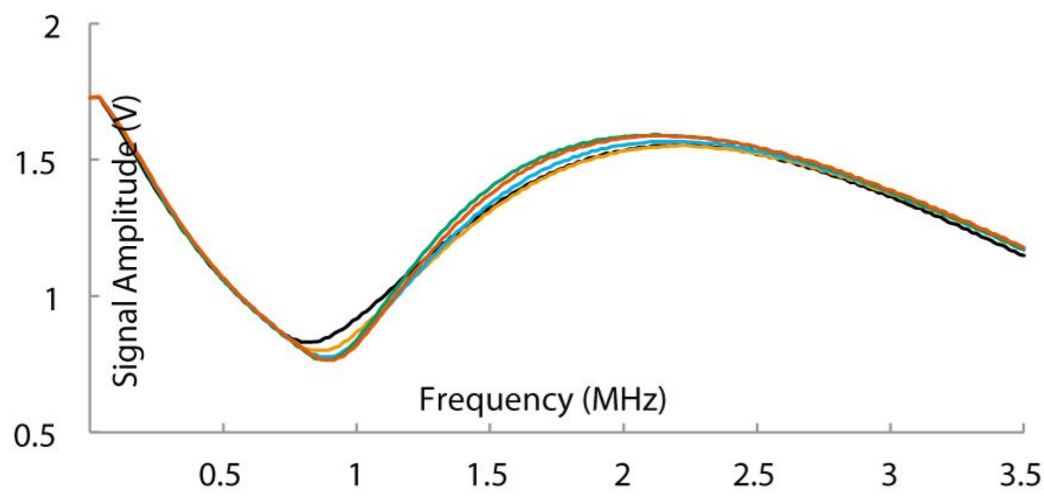
Touché sensor board. The static nature of a table meant that we could ground the sensor to the environment in this configuration (Figure 5.4 A).

5.4.3 On-Body Gesture Sensing

Touché is able to take advantage of the conductive properties of the human body and appropriate the skin as a touch sensitive surface. Because humans are inherently mobile, it is necessary to have the signal source and charge sink located on the body. This configuration is illustrated in Figure 5.4, B.

As our hands serve as our primarily means of manipulating the world, they are a logical location to augment with Touché. In this case, the source or sink is placed near the hands, for example, worn like a wristwatch. The other electrode can be placed in many possible locations, including the opposite wrist, the waist, collar area, or lower back [Gemperle 1998]. As a user touches different parts of their body, the impedance between the electrodes varies as the signal flows through slightly different paths on and in the user's body. The resulting capacitive profile is different for each gesture, which allows Touché to recognize a variety of hand-to-hand or hand-to-body gestures. For our evaluation, described in the next section, we chose to place electrodes on each wrist (Figure 5.6).

A wide array of applications could be built on top of the body. One example is controlling a mobile phone using a set of hand-to-body gestures. For example, making a "*shh*" gesture with the index finger touching the lips, could put the phone into silent mode. Putting the hands together, forming a book-like gesture, could replay voicemails.



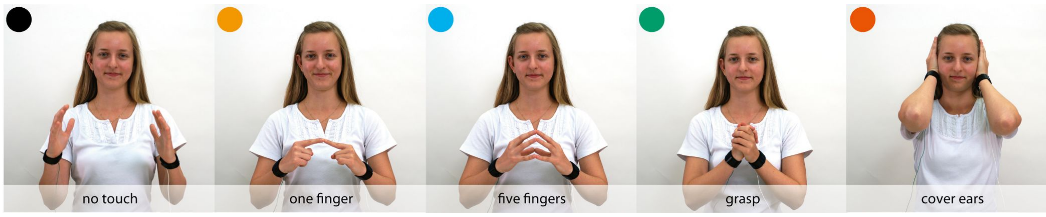


Fig. 5.6 Capacitive profiles for on-body sensing with wrists-mounted Touché sensors.

5.5 Evaluation

We evaluated the two example gesture sets described in the previous section. This served several purposes: 1) to demonstrate different sensor configurations and interactions enabled by Touché, 2) to underscore the immediate feasibility of our approach, 3) to explore the potential richness of gesture vocabularies our system could support, and 4) to establish the baseline performance of the classification engine.

5.5.1 Participants

Twelve participants were recruited (3 female) with a mean age of 27.6. Each participant and condition was run independently, allowing us to distribute data collection over approximately a seven-day period. This permitted us to capture natural, real-world variations in e.g., humidity, temperature, user hydration and varying skin resistance. Although we do not specifically control for these factors, we show that our system is robust despite their potential presence. In fact, our "walk-up" general classifiers were specifically designed to model these temporal and inter-participant variances.

5.5.2 Procedure

Both studies followed the same structure described below. Each study was run independently; the experiment took approximately 25 minutes to complete.

5.5.2.1 Training

Participants were shown pictographically a small set of gestures and asked to perform each sequentially. While performing gestures, the participants were told to adjust their gestures slightly, e.g., tighten their grip. This helped to capture additional variety that would be acquired naturally with extended use, but impossible in a short experiment.

While the participants performed each gesture, the experimenter recorded 10 gesture instances by hitting the spacebar. The experimenter then advanced the participant to the next gesture until all gestures were performed and captured. This procedure was repeated three times providing 30 instances per gesture per participant. In addition to providing three periods of training data useful in post-hoc analysis, this procedure allowed us to capture variability in participant gesture performance, obtaining more gesture variety and improving classification.

5.5.2.2 Testing

Following the training phase, collected data were used to initialize the system for a real-time classification evaluation. Participants were requested to perform one of the gestures from the training set, which was randomly selected and presented on a computer monitor. The system – invisible to both the experimenter and participants – made a classification when participants performed each gesture. A true positive result was obtained when the requested gesture matched the classifier's guess. The experimenter used the spacebar to advance to the next trial, with five trials for each gesture.

5.5.3 Accuracy Measures

Our procedure follows a per-user classifier paradigm where each participant had a custom classifier trained using only his or her training data. This produces robust classification since it captures the peculiarities of the user. Per-user classifiers are often ideal for personal objects used by a single user, as would be the case with, e.g., a mobile phone, desktop computer, or car steering wheel.

To assess performance dimensions that were not available in a real-time accuracy assessment, we ran two additional experiments post-hoc. Our first post-hoc analysis simulated the live classification experiment with one fewer gestures per set. The removed gesture was the one found to have the lowest accuracy in the full gesture set. Accuracy typically improves as the gesture set contracts. In general, we sought to identify gesture sets that exceeded the 95% accuracy threshold.

Our second post-hoc analysis estimated performance with “walk up” users – that is, classification without any training data from that user - a general classifier. To assess this, we trained our classifier using data from eleven participants, and tested using data from a twelfth participant (all combinations, i.e., 12-fold cross validation). This was the most challenging evaluation because of the natural variability of how people perform gestures, anatomical differences, as well as variability in clothes and shoes worn by the participants. However, this accuracy measure provides the best insight into potential real-world performance when per-user training is not feasible, e.g., a museum exhibit or theme park attraction. Moreover, it serves as an ideal contrast to our per-user classifier experimental results.

5.6 Results

Figures 5.5 and 5.6 illustrate the physical setup and accompanying touch gesture sets for the two gesture sets we tested. “Walk -up” accuracies with different-sized gesture sets are shown in Figure 5.7, left. Real-time accuracy results for all five studies are summarized in Figure 5.7, right.

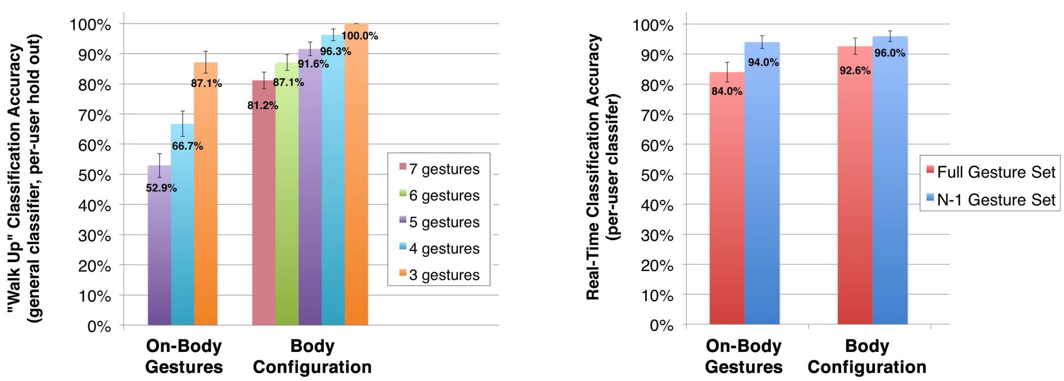


Fig. 5.7 **Left: "Walk Up" classification accuracy for five example applications. Right: Real-time, per-user classification accuracy for five example applications.**

5.6.1 Body Configuration Sensing

A set of seven gestures was evaluated: *not present*, *present*, *one hand*, *two hands*, *one elbow*, *two elbows*, *arms* (Figure 5.5). Average real-time classification performance with seven gestures was 92.6% (SD=9.4%). Eliminating the *two elbows* gesture boosted accuracy to 96.0% (SD=6.1%). Walk-up accuracy at seven gestures stands at 81.2%. As seen in Figure 5.7, accuracy surpasses

90% with five gestures (*not present*, *present*, *one hand*, *two hands*, *two elbow;* 91.6%, SD=7.8%). With only three gestures (*presence*, *two hands*, *two elbow*), accuracy is 100% for every participant.

5.6.2 On-Body Gesture Sensing

Our on-body gesture set consisted of five gestures: *no touch*, *one finger*, *five fingers*, *grasp*, and *cover ears* (Figure 5.6). Real-time, per-user classification accuracy was 84.0% (SD = 11.4%) with five gestures. Removing a single gesture – *one finger* – improved accuracy to a more useable 94.0% (SD=7.4%). In contrast, walk-up accuracy with a general classifier does significantly worse, with all five gestures yielding 52.9% accuracy (SD=13.8%). Reducing the gesture set to three (*no touch*, *five fingers*, *grasp*) only draws accuracy up to 87.1% (SD=12.5%) – stronger, but still too low for robust use.

This divergence in accuracy performance between per-user and general classifiers is important. The results suggest that for on-body gestures where the user is the "device", per-user training is most appropriate. This result is not particularly surprising – unlike man made objects, such as tables, the individual differences between participants are very significant, not only in gesture performance, but also in their bodies' composition. A per-user classifier captures and accounts for these per-user differences, making it more robust.

5.6.3 Anatomical Factors

Touché is sensitive to variations in users' anatomy. To test if anatomical variations have a systematic effect on classification accuracy, we ran several post hoc tests. We found no correlation between accuracy and height (1.6 ~ 1.9m), weight (52 ~ 111kg), BMI (19.6 ~ 32.3), or gender. This suggests the sensing is robust across a range of users.

5.7 Conclusion

Touché investigated an entirely different sensing mechanism to enable on-body interfaces: electrical capacitance. This approach can not only detect whether a user is touching or not touching a surface, but also recognize complex configurations of the human hands and body. Further, its flexible nature brings capacitive sensing to new contexts and materials, including the human body. In this chapter, we demonstrated two body-centric gesture sets implemented using Touché, which could be used for interactive control.

6 ARMURA: MOVING BEYOND FINGERS CLICKING BUTTONS

The outcomes of Skinput, OmniTouch and Touché were primarily technical. Although considerable development work remains, on-body interfaces are no longer artifacts of science fiction – prototypes have been successfully built, evaluated, and demonstrated publically. However, these systems largely borrowed existing, successful interaction paradigms – especially from touchscreen devices – and brought them to the body. Fortunately, on-body interfaces allow us to take advantage of unique and extra dimensions of input our bodies naturally afford us.

In this chapter, we consider how the arms and hands can be used to enhance on-body interactions, which has traditionally been finger input centric (i.e., "fingers clicking buttons"). To explore this opportunity, we developed Armura, a novel interactive on-body system, supporting both input and graphical output. Using this platform as a vehicle for exploration, dozens of applications and interactions were prototyped. This helped to confirm chief use modalities, identify fruitful interaction approaches, and in general, better understand how interfaces might operate on the body. Here, we highlight the most compelling techniques we uncovered. Our explorations also revealed that several "desktop class" interactions are portable to and relevant in the on-body domain. Further, several new interaction techniques were discovered, many unique to on-body computing.

6.1 The Arms as an Input/Output Platform

The sum expressive capabilities of the arms are enormous. The shoulder and elbow allow the hands to be translated and rotated in 3D space (six degrees of freedom). The wrist provides a separate mechanism for rotating the hand on two axes. Each finger has one major and two minor knuckles, which allow them to bend and have limited rotation on two axes. Our thumbs, in contrast, have a much greater rotational range and can also bend. In total, there are more than twenty independent degrees of freedom. Combinatorial speaking, our fingers are capable of forming hundreds of poses [Kendon 1988; Mulder 1996].

We choose to focus on the arms over e.g., the legs or torso, because they serve as our chief appendages for manipulating the physical world. The arms are also an ideal ad hoc display platform given their ready availability and proximity to the head. The skin provides a natural and immediate surface for dynamic digital projection. Line of sight is required, but in practice, is easily accommodated when the arms are active and in front of the user. Furthermore, laser projection (which is focus-free) could one day allow for highly oblique projection, e.g., from a bracelet. Although we focus on graphical output in this work, as that is the predominant output means for computing, on-body interfaces could also readily incorporate auditory or haptic feedback.

6.2 Implementation

To explore how the arms and hands can be used to supplement on-body interactions, we built Armura. Our system is a combination of hardware and software that enables real-time, on-body, graphical interaction. In many ways, it is a realization and extension of the interactions described in previous systems, including Skinput and OmniTouch. As we will see, Armura also advances on-body interaction, offering capabilities not seen in previous systems. As such, it provides a unique vehicle for exploration, allowing us to consider and develop several novel arm-driven interfaces.

The chief feature of our system is the ability to track the location of the arms and hands, as well as recognize their gestural state, in real-time. Additionally, we can simultaneously project coordinated graphical feedback onto the body. We primarily consider a motion- and gesture- driven approach supported by computer vision. Our prototype hardware consists of a ceiling-mounted DLP projector. A mirror is used to project downwards and also expand the interactive area. An infrared camera featuring a wide-angle lens is affixed to this setup, also facing downwards. It is capable of sampling a 640x480 image at 60 FPS. Three infrared illuminators, set apart in a triangle configuration, help to increase contrast (skin is reflective to infrared).

On top of this hardware, we developed a custom Java-based application for detecting, tracking and recognizing arms, hands, and gestures. The heart of the software is a hybrid feature- and template-based gesture recognizer driven by a Support Vector Machine classifier provided by the Weka Machine Learning Toolkit [Hall 2009]. Our system can identify when a user enters the field of view of the camera, the location of the arms if raised (X and Y; Z can be estimated using hand size, but requires calibration – a moot point with depth cameras), and the state (i.e., gesture) of the arms and hands. The later is supported by two independent classifiers, one for arm configurations and another for hand gestures. It is assumed the arms and hands are always holding some gesture, even if just flat or loose (e.g., arms by the side, open palm). Rectification between camera and projector space is achieved with a four-point projective transformation.

The screenshot seen in Figure 6.1 illustrates the result of our user segmentation process. Also pictured are the seven hand gestures our proof-of-concept system supports. The red dot is the estimated top of the hand, used for X/Y position. Although not a comprehensive gesture set, designing around a small range of motions and gestures can also have human advantages in terms of memory and attention load [Hudson 2010], and overall is a good place to begin explorations.

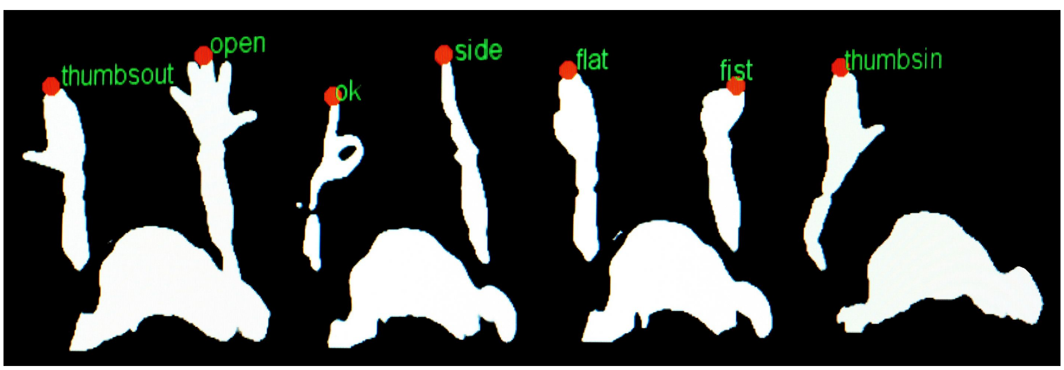


Fig. 6.1 **Four top-down views of a user following the segmentation process. Each hand demonstrates one of seven gestures our exploratory system supports.**

Armura also tracks and classifies 10 synergistic "arm-level" gestures, seen in Figure 6.2. Importantly, hand gestures can only be inferred when the hands are visible and distinct (i.e., not touching other elements). X/Y location is produced using a bounding box of the user (red outline in Figure 6.2).

In practice, our system performs well, tracking at interactive speeds and achieving acceptable classification accuracies (during piloting, ten-fold cross-validation accuracies were in excess of 96% with our seven hand gestures). This proved more than adequate for our exploration of the design space. Most importantly, it served as a good "in-lab" analog (that could be built today) for more complex and/or future technologies that can operate on the go (e.g., with depth cameras and pico-projectors).

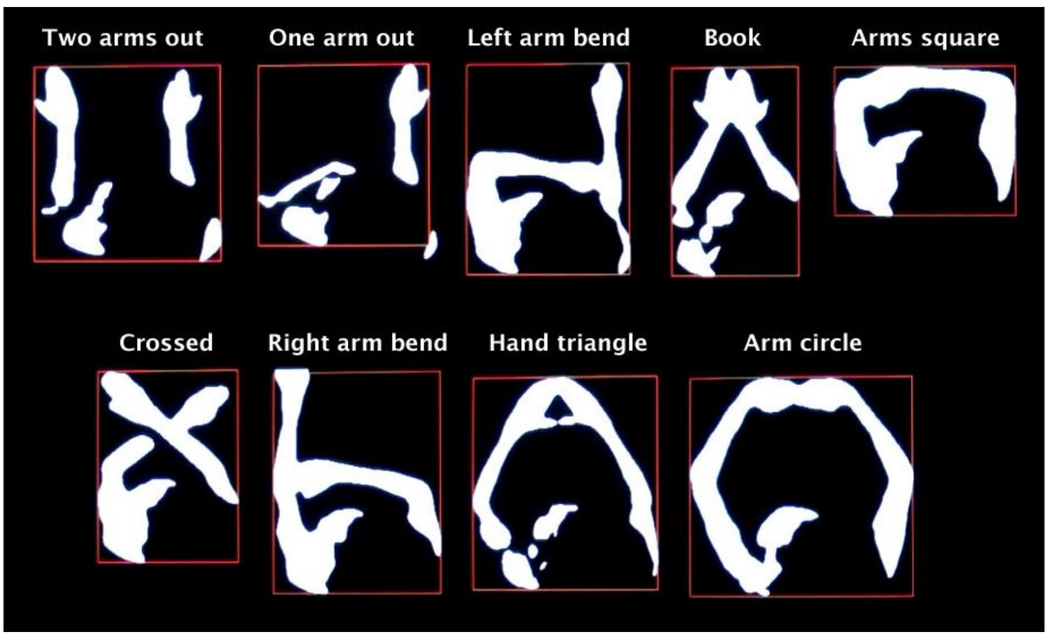


Fig. 6.2 Our exploratory system can also recognize a series of synergistic arm gestures.

6.3 Chief Use Modalities

Over the course of Armura's development, as well as iterative development on a suite of demonstration applications, four chief use modalities became apparent.

6.3.1 Single Arm

The simplest use modality is a single arm (and hand). Although input and output is heavily constrained, it has the important property of leaving the other arm entirely free. This allows the user to engage in another independent task, such as writing or taking a phone call. It could prove especially useful for accessing information related to the other task, such as looking at ones calendar, getting directions, or checking the time. Anecdotally, the elbows tucked in, hands held in front, and palms up appeared most intuitive to users, and also comfortable (i.e., reduced “gorilla arm”).

6.3.2 Single Arm + Other Arm for Display

In this modality, one hand is used for input, while the other hand is used for graphics. The input capability of this modality is identical to that of *Single Arm*, since still only one arm is providing input. The graphical output is simply projected on the “display” arm. This simple change, however, has several beneficial implications.

Foremost, the visual content can now be displayed on a less-dynamic hand, which is easier to view. Not only can this hand be positioned to provide a superior view of the content, but also located in a more comfortable and sustainable position, in case of extended interaction. Equally important, the user no longer has to look at their hand driving the input. Thus, it is free to operate more quickly, at greater distances from the user, and even out of the user’s visual field. In our experiences, this modality is the most comfortable and intuitive to use. In many ways, the input hand operates “off to the side” much like a mouse, requiring no visual attention to control. The eyes can remain fixed on the interactive graphical output, which feels very much like a handheld screen.

6.3.3 Two Single Arms

The previous two modalities rely on the expressive power of a single arm for input. Using two arms for input doubles the expressive power (e.g., 2 x 3 DOF arms = 6 DOF). Note that we do not call this modality “two arms”. Instead we prefer the term “two single arms”, which stresses their individuality. In particular, they simply operate in parallel, gaining no new joint input dimensions. In the next section, we will consider how two arms can be used in conjunction, yielding more synergistic uses.

Bimanual manipulation has been shown to be extremely powerful and intuitive [Buxton 1986; Guiard 1987; Hinckley 1994; Kabbash 1994; Kabbash 1993]. Having extra degrees of freedom available also enables more sophisticated interactions and also provides additional graphical space if needed. The classic 4 DOF action is simultaneously scaling and rotating an image – typically demonstrated on multi-touch surfaces with a two finger "pinch" gesture. This action could be achieved on-body using two hands and a "grasping" gesture. A more practical example might be two-dimensional panning of a map with one hand, while the second hand controls the zoom level and other modal functions, such as a traffic overlay.

6.3.4 Synergistic Use of Both Arms

The final modality is the use of both arms in a synergistic fashion. They no longer operate as two single arms, but rather as a new, combined entity. Uniquely, synergistic configurations, which intersect the arms or hands at fixed points, have the potential to reduce the degrees of freedom. For example, in our *arms square* gesture, the arms become rigidly fixed to one another, naturally limiting the degrees of freedom. This reduction in expressiveness of the input space makes synergistic gestures exceptionally well suited for high-level mode switching. Anecdotally, this appears to match peoples' intuition - initiating an interaction with a coarse arm-level gesture, and then switching to the more dexterous abilities of the hands (and fingers) for manipulation within that state.

Motion constrains in synergistic configurations have some interesting design implications. For example, the *crossed* gesture (Figure 6.2), defined by the intersection of the two limbs at their midpoints, allows for 1 DOF scissor-like movement. The *arm circle* gesture has the interesting property of expansion in one dimension causing the contraction of the other dimension. However, the *book* gesture, which joins the two hands, retains all three axes of motion, as well a unique "page turning" action (Figure 6.3). Scores of other synergistic examples exist, many with interesting and potentially powerful innate motion constraints.

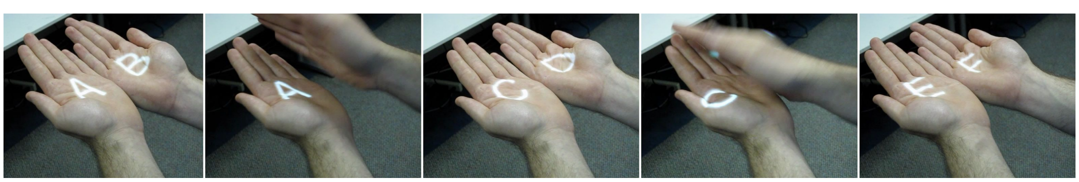


Fig. 6.3 Two hands can create a book interaction metaphor. "Pages" can be turned by flipping the appropriate hand, which transitions the interface. Time advances from left to right.

6.4 Example Applications and Interactions

A central objective of this work was to explore the design space of arm- and hand-driven interaction. Put plainly: if you have two arms, what on-body interactions are possible?

To explore this area, we drew inspiration from three distinct sources. First and most straightforward, was to look at existing on-body systems (see section 2.7) and how the arms and hands were utilized in the proposed interactions. Secondly, we identified promising interactions and metaphors found in contemporary computing environments. Finally, the iterative rapid development of many small applications that run on top of Armura also proved invaluable - we discuss the most interesting of these.

As noted previously, Armura is position- and gesture-centric. It does not capture all possible dimensions of on-body input (e.g., bend angle of the arms) – this full level of richness will slowly become available with future advances in processing and sensors. However, importantly, our system does afford us the ability to explore continuous and discrete input approaches, as well as interactions that take advantage of both simultaneously. Finally, the interactions we describe are not intended to preclude or replace finger input. Indeed, they can be used in concert with fingers to enhance the overall on-body experience.

6.4.1 Presentation Only

The very simplest interaction is to use the arms as a projection surface with no input capability. That is, the location of the arms adds no expressive power to the interaction - they are simply tracked so as to provide a projection surface. This could be used, to e.g., project a map of a user's surroundings, which could prove useful to visitors of museums, hospitals and similar. The system could also display context-sensitive information, such as the time, upcoming meetings, and office directories – all of which could be valuable and shown without the need for input.

6.4.2 Modal Gestures

With even the addition of a single gesture (beyond the required presentation gesture for activation), the interaction space expands significantly. In general, this enables a user to fire an event. If the user does not wish to activate this option, the arms can simply be lowered (i.e., acting as a cancel or escape). For example, a *fist* hand gesture could cycle between different modes - each opening of the palm would reveal a different application, e.g., clock, calendar, stocks. Or if running late for a meeting, an *OK* hand gesture could trigger an apologetic email.

With multiple gestures, the design space is very large and the uses diverse. Different gestures could trigger different modes. For example, a bent "watch-on-wrist" arm gesture could render the time on the wrist. Forming a *hand triangle* gesture and moving it up or down could control the dimming of lights in a room. Gestures are also well suited to instantiation of interaction; for example, using a *two arms out* gesture to "summon" interactive capability. To end an interaction, or navigate backwards through a menu, an *arm cross* gesture could be used.

6.4.3 Modal Layout and Control

A special use of gestures (either hands, arms, or synergistic) is to select a preferred layout and control mechanism for an interactive experience. As discussed in Synergistic Uses of Both Arms (section 6.3.4), some configurations have innate motion constrains, providing natural dimensions for input.

Consider the simple example of an office directory. Presenting one arm might render a compact list of names and offices on the hand, which can be scrolled up and down by translating the arm. When the user presents two hands, the secondary hand might display a headshot and bio of the currently selected individual in the list (rendered on the primarily hand). As an alternative to list navigation, a *book* metaphor could be used, allowing the user to turn "page by page" (Figure 6.3) through the directory, or simply to provide more contiguous surface area for projection.

6.4.4 Position Considerations

In addition to gestures, which are discrete, Armura can track the arms in free space – giving us three dimensions of continuous input. There are four important ways this positional information can be calculated for on-body systems.

First, the body can serve as an anchor point, making positions relative with respect to e.g., the user's head [Hinkley 1994; Mine 1997]. This has the potential to leverage muscle memory, if for example, we know our calendar is accessible to the left of us [Li 2009]. Second, one arm can provide a reference point, from which the location of the second arm can be derived (see e.g., [Gustafson 2010]); Third, movement can be tracked relative to the arm or hand's entry position or state change (e.g. switching from a *flat* to *side* hand gesture). Lastly, as shown in LightSpace [Wilson 2010], the interface can be made relative to the environment – a hand over a desk might have a different function than one over the floor.

6.4.5 Modal Positions

If desired, position can be used for modal input. This was the basis for [Li 2009], which laid out mobile device functions in space around the user. Looking again at the very simplest of interactions, we could imagine a navigation application that displays what lies ahead based on how the hand is positioned relative to the user or environment. For example, directions to different buildings could be shown when the hand is to the left, right and in front of the user (Figure 6.4). This requires no further interaction; the user simply drops their hand when finished.

Fig. 6.4 **Moving the hand to different positions relative to the body reveals what buildings lay in that direction.**

6.4.6 Menuing

Navigation of "menus" is core to the modern computing experience, and allows for graphical browsing and selection of desired functionality. Through a use of gestures, positional tracking, and a combination of the two, on-body interfaces can readily support intuitive menu navigation, and thus enable a wide variety of interactive applications.

6.4.6.1 Gestures

Simple menuing can be supported though a small vocabulary of gestures. For example, a yes/no dialog could be answered by flipping the hand palm-side-up or down (Figure 6.5). With as little as three gestures, navigation of hierarchical menus becomes possible (although not practical for long lists). Indeed, a simple vocabulary of arm and hand gestures working in concert could provide sufficient richness for a simple, but powerful on-body personal organizer.

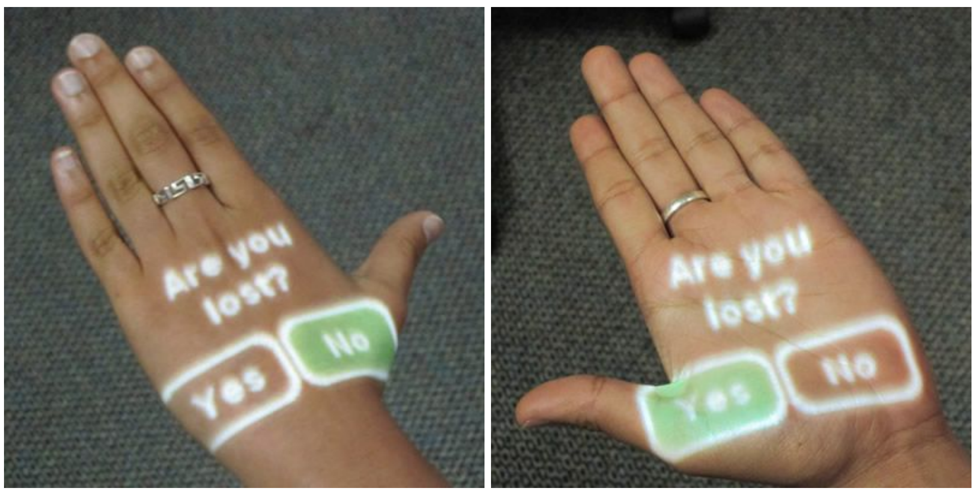


Fig. 6.5 A user can toggle between yes and no by flipping their hand (thumb in and thumb out gestures respectively), pointing their thumb at the desired target. To select, the thumb is tucked in (flat gesture).

6.4.6.2 Position Tracking

A single axis of movement could be used for continuous manipulation – to, for instance, scroll an arm-projected office directory. As in previous examples, when the interaction is complete, the user can simply drop their arms to a resting position.

Two or more axes of positional data enable unique interactions. For instance, we can use one dimension for modal control and another for manipulation. As an example, we built a music player interface with five modes: seek, volume, next song, previous song, pause/play. The modes were traversed by moving the hand from left to right. The song position and volume level were manipulated by the relative forwards/backwards motion of the hand (Figure

6.6). The three binary actions (next song, previous song, pause/play) could be activated by briefly rocking the hand forwards.

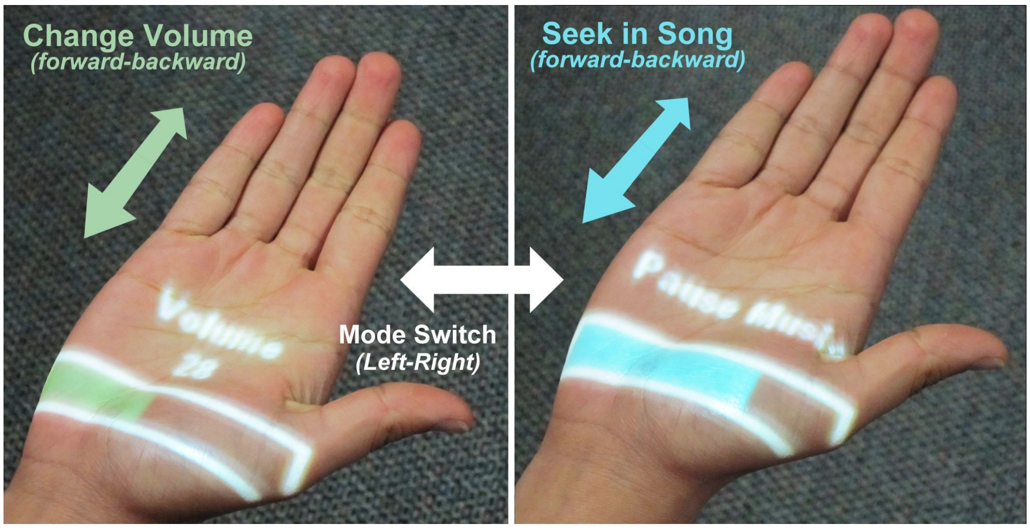


Fig. 6.6 **Moving the hand left-to-right toggles between different audio modes (X-axis). The level can be manipulated by rocking the hand forwards and backwards (Y-axis).**

6.4.6.3 Gestures and Positional Tracking

Unsurprisingly, combining gestures and position enables the richest and most intuitive set of interactions. Even a single degree of positional freedom in concert with one "action" gesture allows users to navigate hierarchical list and menus in a practical manner (users can traverse up the hierarchy by using "back" menu items, see Figure 6.7 left). This immediately enables rich applications, such as an address book, calendar, or music player. With one additional gesture, the "back" menu items become unnecessary; a dedicated gesture can be used instead.

Fig. 6.7 **Left: Hierarchical menu navigation on the hand. Right: demonstration tattoo painting application.**

6.4.7 Crossing Gestures

Skinput and Omnitouch relied on direct finger-to-arm contact for driving their button-centric interfaces. However, it is also possible to use a specific gesture to serve as an analog to a mouse-down action, with pinching appearing most popular [Gustafson 2010; Wilson 2006]. However, it is possible to support on-body selection without "clicks" or discrete gestures though the use of crossing gestures [Accot 2002; Apitz 2004]. As seen in Figure 6.8, a list of actions is rendered on the arm; selection occurs by swiping the free arm through a desired target.

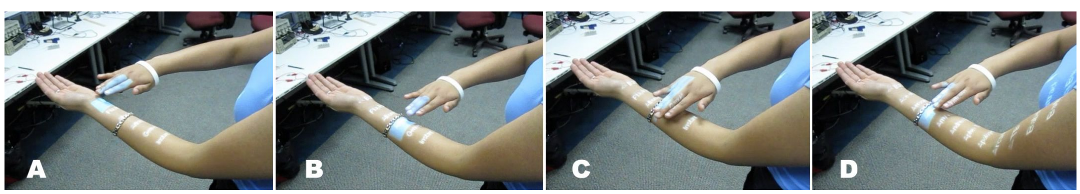


Fig. 6.8 **A linear menu rendered on the left arm. Sliding the right arm forwards and backwards along this axis changes the selection (A & B, blue highlight). Selection is performed with a crossing gesture (C), which advances the interface (D).**

6.4.8 Defining Axes of Input

The arms and hands can also be synergistically used as reference points, bounding features, or to delineate an area for interaction. Such a method is described in [Gustafson 2010], which uses a ‘L’ gesture to provide a two-dimensional plane for drawing in free space. The interaction illustrated in Figure 6.8 offers another example: an extended arm is used to provide an intuitive linear axis of input movement, which also serves as a large graphical canvas for navigating a list. The other hand can moves along this axis for selection.

6.4.9 Cursor Control

As demonstrated by the venerable mouse, X/Y movement and a few buttons, combined with appropriate graphical feedback, enables endless interaction possibilities. The hands can be readily digitized to provide such functionality, allowing for control of on-body, point-and-click interfaces.

To demonstrate cursor control, but in a more fitting domain, we built a digital tattoo painting application (Figure 6.7, right). The left arm serves as the canvas (no input, tracking only). The right hand controls three modes. The first is the nominal mode, where movement of the right hand correspondingly moves a brush cursor on the opposing arm (using a control-device gain of roughly 1:5 for precision painting). If the user forms a “brush holding” gesture (Figure 6.1, *OK*), the cursor paints as it moves. Finally, closing the thumb (Figure 6.1, *flat*) on the right hand switches to brush selection mode. The user can select a paint color by moving their hands forward and backwards, while brush thickness is controlled by left-and-right motion. Visual feedback is provided with an enlarged “brush preview” rendered on the right hand.

6.4.10 Peephole / Spatially-Aware Displays

The arms, and especially the hands, can be used like a peephole or lens [Fitzmaurice 1993; Hang 2008; Yee 2003]. This has been shown to be a powerful metaphor for small displays - for all practical purposes, we can treat the hand as a small display. Jumping back to an earlier example, peephole displays could allow users to not only see a map of their current surroundings, but also move their hands to view other areas (Figure 6.9).

As a spatially-aware display, the selection of an appropriate anchor point is important (see Position Considerations, section 6.4.4). Using, for example, the body for relative positioning [Mine 1997], could allow the hands to show an “X-ray view” inside our bodies. Or, if anchored to the environment or an object [Wilson 2010], such as a map, the hand could act as an additional layer, providing supplemental geo-spatial information.

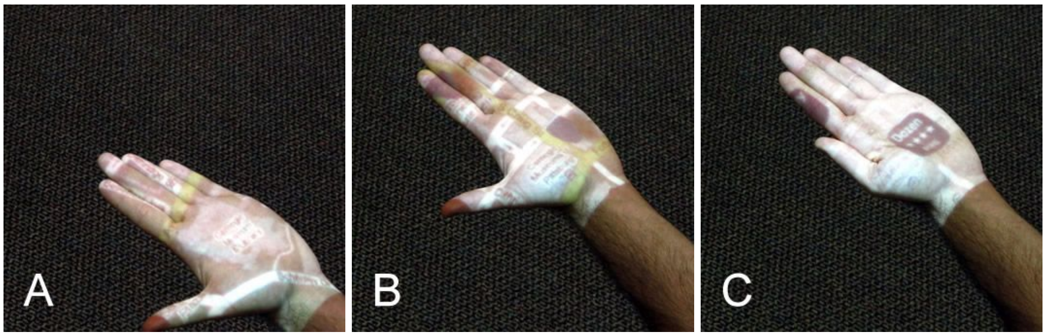


Fig. 6.9 **The hand can be used like a peephole display. Translating the hand in two-dimensions also translates the map view (A to B). "Clicking" the thumb (C) switches to an alternate layer; in this example, restaurant reviews.**

6.4.11 Control-Device Gain and Clutching

Unlike, e.g., a scroll wheel on a mouse, the arms have a maximum pan-able distance: one arm's length. To overcome reach limitations that occur with cursor control and peephole displays, we experimented with two ways to expand the interactive space.

First we tried manipulating the control-device (CD) gain (with the device being one's own hand). This appeared to work conceptually for an "on-body cursor" like we used in our tattoo application. However, with peephole displays, the effect was somewhat disorienting. For example, if seeing an object on the periphery of their hand, there was a tendency to move the hand to that spot so as to center that item – as they would do when manipulating the real world (which is inherently 1:1). However, a higher CD gain broke this useful physical metaphor, and anecdotally caused overshooting. Overall, we found a CD gain of 1:1 to be most natural and intuitive for peephole displays. In other words, when the hand is moved 5cm, the graphics should also track 5cm. Our second approach is clutching via hand gestures. This can greatly expand the interactive space while also preserving a 1:1 CD gain if desired.

We built a simple map application to help users locate near-by restaurants, navigated by moving a hand in the X/Y plane (1:1 CD gain). For clutching, we used a *fist* gesture, a natural choice for a “grab” action. The result feels very much like how one would manipulate a large map on a table. When a restaurant of interest is located, the thumb can be tucked in to see metadata (e.g., star rating; Figure 6.9).

6.5 Conclusion

In this chapter, we have provided an overview of our explorations with the Armura system. We found that the interaction design space for on-body interfaces is large and underexplored. Indeed, not only are conventional desktop interactions such as cursor control and crossing gestures relevant, but also new avenues, such as innate motion constraints and dynamic modal layout. We endeavored to describe and taxonomize these different design and interface control aspects. We hope this can both guide future research efforts as well provide a vocabulary to help describe such systems.

7 IMPLICATIONS OF BODY LOCATION FOR ON-BODY VISUAL ALERTS

On-body interfaces will need to notify users about, for example, new emails, upcoming meetings, changes in weather, excessive caloric intake, and other aspects of our lives mobile computers will be able to monitor directly or have access to. A critical dimension of design to consider is body location. For successful development and user acceptance, a body location needs to be visually and physically accessible, socially acceptable, sufficiently large and stable for interaction, and also effective at conveying information. In this chapter, we describe a focused experiment that sought to understand the reaction time performance of visual alerts distributed across the body. In the future, systems like OmniTouch and Armura will be able digitally project such alerts.

Our experimental results can be used to inform the design of future on-body interfaces (and wearable displays in general). For example, an incoming phone call must be answered within perhaps fifteen seconds. Locating the corresponding alert at a place on the body that is rarely viewed will result in many missed calls. On the other hand, consider the example of a notification indicating that six months has elapsed since one's last dentist appointment. This does not require a response with fifteen seconds, or potentially even days. Locating this alert in a highly visually salient location is clearly inappropriate. This would unnecessarily interrupt the user from their present task or social engagement for an item of little immediate consequence. Thus, this research allows researchers and practitioners to best align their application with areas of the body that have the necessary attention demand and reaction time characteristics. Applications that apply this information may be less disruptive and reduce overall information burden.

7.1 Alert Methods

One way to draw a wearer's attention is through vibrotactile stimulation, commonly employed in mobile devices to alert users (e.g., to an incoming call). Cholewiak et al. [2000] investigated the sensory performance characteristics of vibrotactile patterns at different body locations (device access time is studied in [Ashbrook 2008]). Also popular, although not tied to any particular body location, are audio alerts.

Visual alerts offer an alternative notification method. These could be digitally projected (as in Skinput, OmniTouch and Armura) or physical (e.g., small worn displays [Williams 2006; Lim 2011]). Unfortunately, there has been little research into optimal body placement, despite being an unobtrusive, lightweight, and low-powered information delivery mechanism. Furthermore, visual stimuli have the added benefit of being able to work alone or in concert with conventional methods, including auditory and vibrotactile alerts (see, e.g., [Brewster 2000] and [Brown 2006] respectively).

7.2 Study Hardware

To explore aspects of optimal body placement, we developed a distributed array of small, self-contained sensor-displays (Figure 7.1). These devices were attached to different parts of a participant's body, where they would initiate visual stimuli and capture participant reaction times. The final design, employing a PIC microprocessor, measures 2.3x3cm and weighs 11g, including battery. Being both small and wireless meant that participants' mobility was not restricted. An integrated safety pin allowed the devices to be easily affixed to a participant's clothing. A velcro strap was used for the wrist if the participant did not wear long sleeves.

A red LED flashing at approximately 10Hz acted as the visual stimulus. The LED was frosted to provide better, omni-directional light dispersion. Users interacted with the device using a single, large, surface mounted button. The device's logic is as follows: (1) Suspend for a random period between 2 and 16 minutes. (2) Wake up and begin flashing the LED. Begin tracking the elapsed time using an internal counter. Continue flashing the LED until the user responds by pressing the button. (3) When this occurs, save the elapsed time (i.e., reaction time) into non-volatile memory. (4) Turn off the LED and go to step 1.

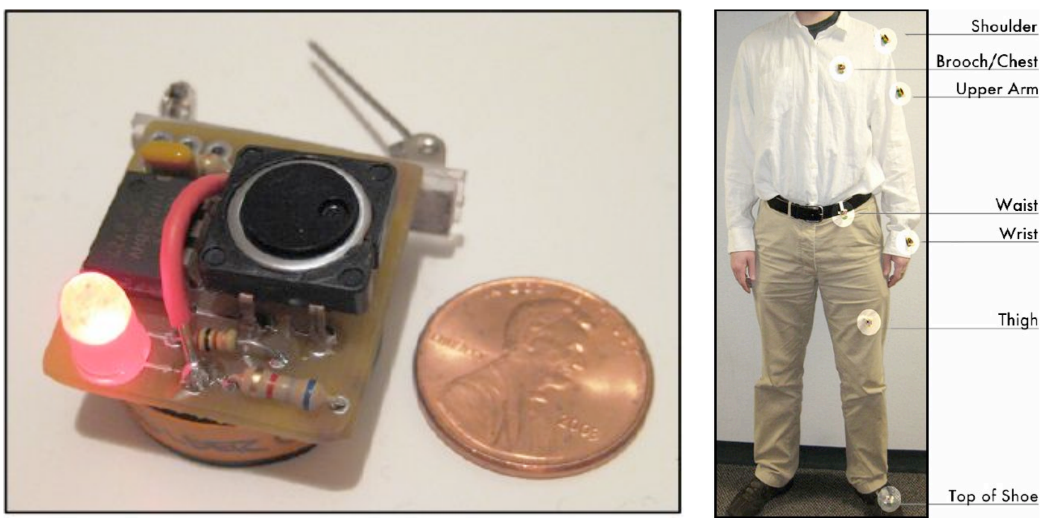


Fig. 7.1 **Above: one of seven sensors worn by participants to determine reaction time to visual stimuli. Right: The seven body locations selected for evaluation.**

7.3 Evaluation

To prevent overwhelming the user with an incessant need to react, we had to strike an important balance between the number of devices and the frequency of their activation.

We ultimately selected seven locations per participant based on suggestions by Gemperle et al [1998] (Figure 7.1). Their investigations looked at issues such as accessibility for interaction and aesthetics – two vital factors to consider for both worn and projected interfaces. Additionally, these locations encompass the most common placements for interactive devices today - upper arm, wrist, and waist – as well as all of the locations used in [Ashbrook 2008]. To be comprehensive, we included four additional locations that represented the most significant, although unconventional, remaining areas – feet, legs, and torso. Although other locations exist (e.g. fingers), we believe the distribution and density we employ sufficiently covers and captures data for the most useful and likely areas of the human body.

Twenty-five participants (12 female) with a mean age of 23.3 (min=18, max=50) were recruited using a recruiting website and posted flyers. Each received $15 for their involvement. To keep device (and interruption) count down, we deployed sensors to only one side of the participant's body. We compensated for experimental effects associated with differences in laterality (i.e., dominant side) [Porac 1981; Holder 1997] by balancing the side of the body we deployed our sensors to within right- and left-handedness groups (n=22 and 3 respectively).

Participants were given a brief explanation of how the devices work, and that they were to press the device button as quickly as possible once they noticed the light blinking. Then, with the help of the participant, the devices were affixed in the specified locations. The devices were worn for 2-3 hours, allowing for about ten data points to be collected per location. Participants completed an exit survey at the end of the study period.

During the study period, participants were told to go about their normal routine. We purposely avoided engaging participants in an artificial task, as this would have created a fictitious attentional scenario, both cognitively and visually (and thus rendered our sensitive reaction time results equally fictitious). We simultaneously hoped this flexibility would capture a variety of use contexts. However, although all participants left our lab, they spent most of the study period (91.5%) seated and working (e.g., using a laptop, doing homework, reading). All but three participants were observed from afar during the study period. This provided a considerably more contextualized and intimate source of information than the raw reaction time data, especially about how different locations perform in various postures and settings (e.g., working with or without a desk). Finally, although this seated and working context is narrow, it is perhaps the most prevalent backdrop for device interaction and information exchange in the digital age, and is thus of significant interest and importance.

7.4 Results

We found that the reaction times in our study roughly conform to an exponential distribution. To derive our statistical measures, we took the log of reaction times, which transformed the data into a more normal distribution, allowing us to use paired, two-sided t-tests. A closer inspection of the data revealed a slight bimodal distribution. This appears to be caused not by two types of participants (e.g., fast and slow), but rather an effect within participants. We suggest that this is a result of two distinct ways users react to visual alerts. In particular, we believe the first peak (reaction times under four seconds) is caused by people noticing the device switch state (i.e., begin flashing), prompting them to react immediately. However, if the wearer does not catch this initial change, their reaction time is roughly modeled by a log-normal distribution, with means between 32 and 128 seconds, depending on the location.

There was no significant difference (α=.05) between left- and right-handed groups; we combine the results for brevity. There was also no significant difference (Bonferroni-corrected two-sided t-test) in dominant vs. non-dominant body side placement, except in the waist location ($p<.05$). However, the data suggests dominant side placement outperforms non-dominant side placement by almost 40% on average. Figure 7.2 displays the average reaction time for each body location across participants. Table 7.1 contains p-values from a Bonferroni-corrected, all-pairs, two-sided t-test. Figure 7.3 illustrates how these reaction times are distributed in time. Cumulative reaction time likelihood (Figure 7.4) can assist in selecting body locations - one can look up, for example, that 90% of reaction times for the arm location occur within 64 seconds.

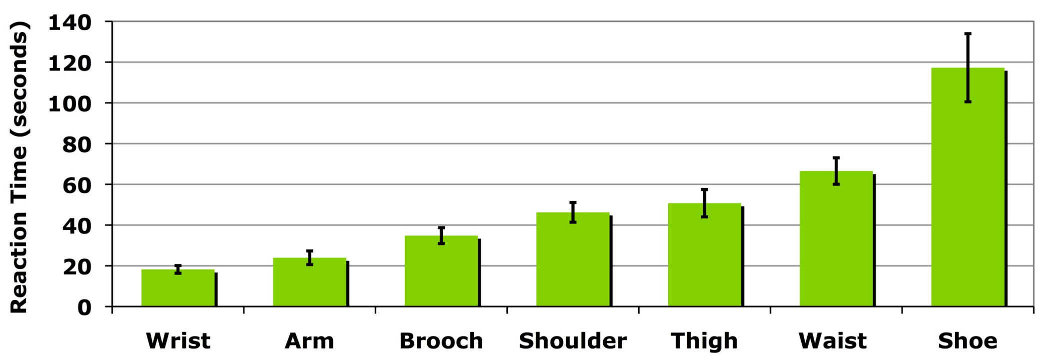


Fig. 7.2 **Average reaction time performance for the seven tested body locations.**

	Wrist	Arm	Brooch	Shoulder	Thigh	Waist
Arm	n.s.					
Brooch	< .05	n.s.				
Shoulder	< .001	< .001	n.s.			
Thigh	< .001	< .001	n.s.	n.s.		
Waist	< .001	< .001	< .05	n.s.	n.s.	
Shoe	< .001	< .001	< .001	< .001	< .001	< .001

Table 7.1 **Statistical significance of reaction time performance differences (all location combinations).**

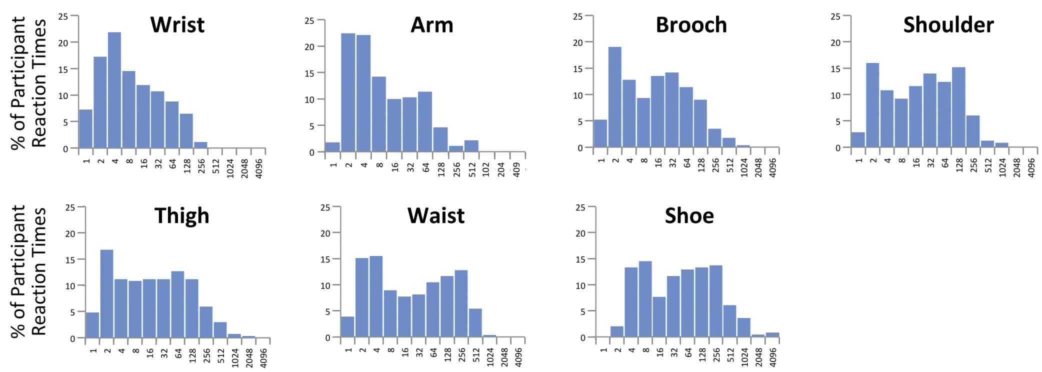


Fig. 7.3 **Distribution of reaction times for each of the seven tested body locations. The X-axis is reaction time in seconds (log scale). The Y-axis is percentage of reaction times that occurred in that period.**

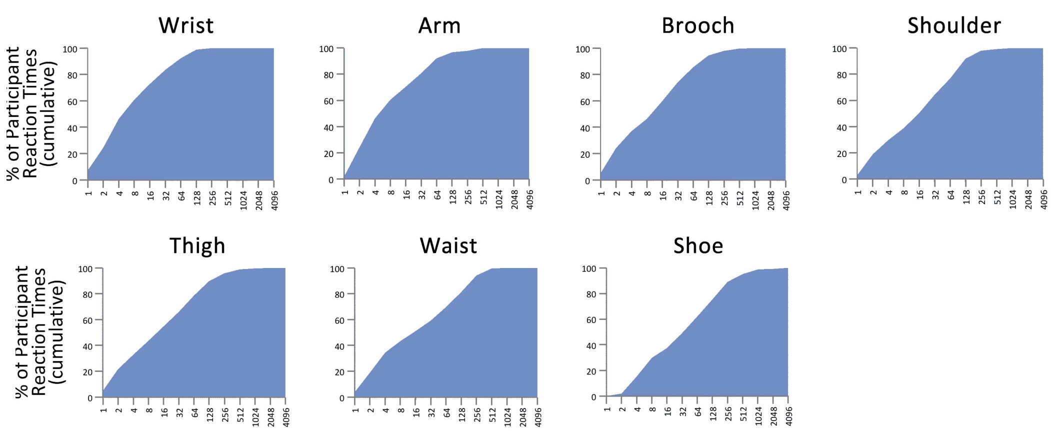


Fig. 7.4 **The cumulative reaction time likelihood for each of the seven tested body locations. The X-axis is reaction time in seconds (log scale). The Y-axis is cumulative likelihood of the user reacting within in elapsed period.**

7.5 Discussion

There were distinct differences in how participants reacted to visual stimuli in each of the seven locations. In addition to strong variations in reaction time captured by our distributed sensor system (Figure 7.2), there were many interesting observations made regarding when and how different locations were particularly effective and ineffective.

7.5.1 Wrist

Participants reacted not only fastest to the visual stimuli on the wrist, but also more consistently (smallest standard error among all locations). It outperformed all other locations ($p<.05$), except the arm. This result aligns well with field observations. The wrist appeared to be most visually accessible during hand motor activities, such as writing, reading, typing, and conversational gesturing. Because most of our subjects were engaged in one of the latter activities, wrists performed exceptionally well. However, this performance might drop significantly in other contexts such as walking. Nonetheless, wrists offer the unique opportunity to deliver alerts and information most saliently during activities when we are using our hands, which could be leveraged to great effect in activities like typing (e.g., at work), building (e.g., at construction site), and cooking.

7.5.2 Arm

With performance almost matching that of the wrist, it is curious why so few devices take advantage of this salient area for visual alerting. It significantly outperformed the thigh, shoulder, waist and shoe locations ($p<.001$). Located on the exterior of the upper arm (tricep), our devices appeared to be just within most participants' peripheral vision, allowing for almost immediate reaction when the visual stimulus began. Participants were 41% faster at reacting to stimuli when the device was placed on their non-dominant side (although, as noted previously, this result was not statistically significant). Use of their dominant (more agile) hand to press the button would explain some of this performance gain. However it seems unlikely to account for the full, 12-second difference.

7.5.3 Brooch

Unlike the wrist and arm, the brooch location seems to be located outside the wearer's peripheral vision. From field observations, we learned that when participants were engaged in activities in which their head was level with the horizon (e.g., walking, socializing, talking on a cell phone), brooch-located visual stimuli often went unnoticed. It was only when users tilted their heads down did the location seem to enter the visual field (users would suddenly notice the stimulus and press the button). The brooch area, like the arm, benefited from being placed on the non-dominant body side, a performance gain of 36%.

The respectable performance of this location (third best) is likely due to the high percentage of participants working on the laptops or reading books, both of which tend to orient the head downwards. This characteristic could be useful, however, by allowing information to be pushed to the user when engaged in "heads down" activities, such as eating, reading, and drawing. Meanwhile, "heads up" activities, such as socializing, walking, and driving, could benefit from reduced interruption.

7.5.4 Shoulder

Participants repeatedly commented that they believed the shoulder was the most noticeable location, in contrast to the collected data. It is possible that this location's proximity to the eyes increases the saliency of the stimuli when it is noticed. Interestingly, this proximity did not guarantee the stimulus would fall within the wearer's field of view. Observations suggest that natural panning or tilting of the head caused the location to periodically drift into the peripheral vision, allowing the device to catch the wearer's attention. Changes in posture, which shift how clothes sit on the body, also contributed to this effect. Lastly, some participants noted that the stimulus was occasionally visible as a reflection in their glasses.

Data collected from the reaction time sensors show the shoulder only significantly outperformed a single location, the shoe ($p<.001$). Although not statistically significant, it fares slightly better than thigh and the waist. This is interesting as the waist is a popular location to situate mobile electronics, where they can be attached to the belt. However backpack straps (and similar) could serve a similar function and offer improved reaction time.

7.5.5 Thigh

The performance of this location suffered heavily due to occlusion by tables - it was only readily noticed when participants were leaning back. Additionally, similar to the brooch position, the thigh is typically not visible when standing - the wearer must look down to see it. However, it is possible that the thigh could be especially useful in contexts like TV viewing or seated socialization. It is in these settings that users tend to be especially reclined, providing line of sight to the thigh and easy access to it. The thigh, like other locations, benefited from being placed on the non-dominant side of the body, with average reaction times 45% faster - a difference of more than 28 seconds.

7.5.6 Waist

This location was particularly sensitive to how participants were seated. When leaning back, the waist was usually visually accessible. However, participants leaning forward or sitting up against a work surface tended not to notice alerts. In some cases, the area was completely occluded by the body (e.g., elbows on knees, arms in lap, large body mass) - one participant forgot about the position entirely. Additionally, this position does not fall within the wearer's field of view when facing forward. This was the only location that had a significant difference between dominant and non-dominant body side placement, ($p<.05$) with reaction times of 83.8 and 48.1 seconds respectively.

The poor performance of this location, second only to the shoe, stands in contrast to its popularity for device placement ([Ashbrook 2008] indicates the area is not particularly physically accessible either). Although the waist (belt) is convenient for attaching mobile devices today, future smaller devices are likely to render this location less useful, especially considering its visual reaction time performance deficit. Furthermore, this result clearly demonstrates why we are so reliant on auditory and vibrotactile alerts for devices worn in this area.

7.5.7 Top of Shoe

Beyond the obvious physical obstacles involved in interacting with a shoe-bound interface, the location also offers the worst reaction time performance of any body location we tested. From our observations, this seemed entirely due to the fact that feet are either tucked under tables, hidden by knees when seated, or obscured by books or laptops (and many other shoe-obscuring contexts are possible). Interestingly, the shoe also demonstrated the largest dominant vs. non-dominant placement effects, with means of 157 and 78 seconds respectively.

However, the poor visual accessibility of the shoe might also be considered its strongest attribute. Feet tend to be hidden when people are working (e.g., using a laptop, sitting at a desk), a context where interruption may be particularly expensive. Information presented via the shoe is unlikely to catch the wearer's attention, reducing disruption. It seems the shoe becomes most accessible when in an upright location, and in particular, when walking or running, when the feet tend to kick out in front, and may enter the peripheral vision. Thus, shoes might be the ideal platform from which to deliver information when users are mobile between tasks (e.g., walking to a meeting, to the bathroom, or to get lunch). Because of the significant variation in reaction time, information would likely be limited to types that are time-insensitive (e.g., pick up milk on your way home, schedule a dentist appointment).

7.6 Conclusion

In this chapter, we have shown that there are significant differences in how visual alerts distributed on the human body capture our attention. This reaction time performance is not only influenced by the innate properties of each location, such as physical distance or visual accessibility, but also outside factors, such as occlusion by furniture. We believe these data and observations can help inform the design of future on-body systems.

8 IMPLICATIONS OF BODY LOCATION FOR ON-BODY INTERFACES

Skinput, OmniTouch, Touché, Armura and other on-body systems primarily focused on bringing interactive capabilities to the arms and hands. As described in Section 7.1, arms have many benefits that make them excellent candidates for on-body interactions. However, the body has other surfaces that are capable of supporting interactive experiences. Thus, an invaluable question to explore is where else on the body is applicable, and what are the associated benefits and drawbacks with such body locations.

To understand this expansive space, we employed a two-part, mixed-methods exploratory process. Participants from a variety of backgrounds were asked to reflect on their openness to interacting with touch interfaces projected onto various parts of their body. Our investigations started with high spatial resolution, but low detail crowdsourced data. We then complimented this model with low spatial resolution, but high-detail qualitative feedback from a diverse set of experts, including a tattoo artist, massage therapist, and dance instructor. The results of this structured exploration reveal both limitations and opportunities, which point the way towards for more comfortable, efficacious, and enjoyable on-body user experiences.

8.1 Social Dimensions of On-Body Touch

Because on-body computing systems require direct interaction with the body, their design necessitates an examination of the social acceptability of touch. The importance of touch is widely acknowledged in childhood development (honing essential motor skills, spatial and emotional perception, and psychological and physical health [Field 2001; Hertenstein 2002; Montagu 1986]) and also adulthood, given its importance in interpersonal communication. Work has also examined the distinct emotional signals that are conveyed through touch and how touch enhances the multimodal perception of emotion [Thayer 1982; Knapp 2006]. Studies of identification of emotion from being touched by an unseen stranger on the arm, or when watching someone else being touched on the arm, demonstrate that a tactile modality can accurately signal at least six emotions to a degree comparable to facial and vocal communication [Hertenstein 2006].

One of the most comprehensive studies on the meanings people assign to touch was a study by Jones and Yarbrough [1985]. Data from 37 participant observers trained to directly notate touch events in which they were involved demonstrated the variety of precise cultural meanings communicated by touch. For every social touch event they experienced over three days, participants immediately recorded: who initiated it; the number and location of body parts that came into contact; whose space the encounter occurred in; who said anything and the timing of the verbalization in relation to the touch; the participant's interpretation of the meaning of the touch; and other relevant data.

Nearly 1,500 touches were collected and analyzed, revealing twelve distinct categories of meaning for touch: support, appreciation, inclusion, sexual interest or intent, affection, playful affection, playful aggression, compliance, attention-getting, announcing a response, greetings, and departure. They also identified two general regions of the body: *non-vulnerable body parts*, including hands, arms, shoulders, and upper back, and *vulnerable body parts*, comprising everywhere else—the head, neck, torso, lower back, buttocks, legs, and feet. They observed that vulnerable parts were touched in *close relationships*, while non-vulnerable parts were socially available for *others to touch*.

Perhaps the most significant aspect of this study is the finding that, not only does interpersonal touch have specific and widely ranging symbolic meanings in adult communication, contextual factors (such as verbal and nonverbal behaviors and relational and situational factors) are critical to the meanings of touch events. For on-body interfaces to seamlessly integrate into real-world use they will need to understand and leverage these highly variable social contexts in adaptive ways.

Recent work examining some of the contextual nuances of interpersonal touch from disciplines such as cognitive and social psychology, neuroscience, and cultural anthropology, demonstrates that it appears capable of powerfully affecting people's behavior. Gallace and Spence [2010] review work on how even casual interpersonal touch can powerfully affect people's attitudes towards a particular service, their compliance with requests, and their degree of bonding with others.

However, in many social contexts, interpersonal touch is actively discouraged for cultural and legal reasons, norms that will need to be carefully considered by on-body interface designers. Some guidance may be found in a survey of nearly 500 college students [Tomita 2008] of beliefs about where it is acceptable to touch and be touched by other students in casual social interactions with the intent of informing college sexual harassment and health policies. A hierarchical cluster analysis was used to form *touch zones* (Public, Discretionary, and Private) by gender showing the ratio of acceptability of touching a zone in each condition. Yet while this approach classifies where students believe it is acceptable to touch and be touched in casual social interactions, such studies do not address the real-time question of why such touching may or may not be contextually appropriate.

One possible means of guiding the acceptance of interpersonal touch is through the design of embodied interfaces that facilitate social interaction. In a project titled Mediated Body, Hobye and Löwgren [2011] present an experimental interactive performance that uses a body suit and audio feedback to justify bare-skin touch between strangers. This paradigm is a striking example of how mediated environments can provide open-ended prompts for social interaction that can extend across bodies. Such mixed-reality interfaces heighten the potential for interpersonal communication, and indicate a largely untapped area of research.

In general, on-body interactions provide the potential for heightened sensory perception, and consequently, associated improvements in emotional awareness and human-centered creativity [Faste 1995; Löwgren 2009; Milgram 1994]. Although research into the social implications of on-body interfaces is clearly in an early stage, our aim with this project is to provide guidance on how such development might best proceed.

8.2 Crowdsourcing a Baseline Model

Before we could begin asking questions about *why* locations were good or bad for on-body interfaces, we first had to know *where.* Thus, our first step was to build a baseline model for the appropriateness of locations across the human body. Of course, there are a multitude of biological, personal and cultural factors that influence this model, including body image issues, greeting customs, and musculoskeletal constraints. To build a model that illuminated generalizable recommendations, it was important to solicit input from a wide spectrum of people. To achieve this, we designed a task for Amazon's Mechanical Turk (http://mturk.com). The population of participants (workers) is a reasonable proxy for the population at large (see [Ross 2010] for more discussion).

8.2.1 Developing a Set of Poses

The human body can form an almost limitless set of possible poses. In these different poses, the appropriateness of on-body interfaces change, due to factors such as physical and visual accessibility (as we touched on in Chapter 7). In response, we felt it was important to investigate a set of poses, such that these differences – potentially opportunities – could be explored. Through informal observations at a cafeteria, coffee shop and library, a set of seven commonplace poses was decided upon. In addition to standing, there were six varieties of sitting: legs together, reclined and legs together, legs apart, legs crossed at knee, legs crossed ankle-to-knee, and crossed legged on floor (Figure 8.3). A vast majority of typical peoples' days are spent in these various poses.

8.2.2 Procedure

After filling out basic demographics information, participants viewed a thirty second, silent video montage of on-body interfaces from Skinput, OmniTouch and Armura. A numerical code appeared briefly in the middle of the video as a reliability check; participants had to enter this code later in order to continue the study. We first we considered using storyboards and mocked-up photos. However, piloting proved very useful in determining that the inclusion of a video was absolutely necessary. On-body interfaces are something most people have never encountered. A video quickly and accurately conveyed the main idea in a way that text and even images could not.

Participants were then presented with a randomly selected black silhouette, depicting one of the seven poses described in the previous section. A single small green dot was superimposed onto a random body location (Figure 8.1 left). If visualized all at once, these dots would form a regular grid covering the silhouette, spaced apart by ten pixels both horizontally and vertically (Figure 8.1 right). As an additional postural cue, a small inset image depicted a stick figure in side profile in the same pose.

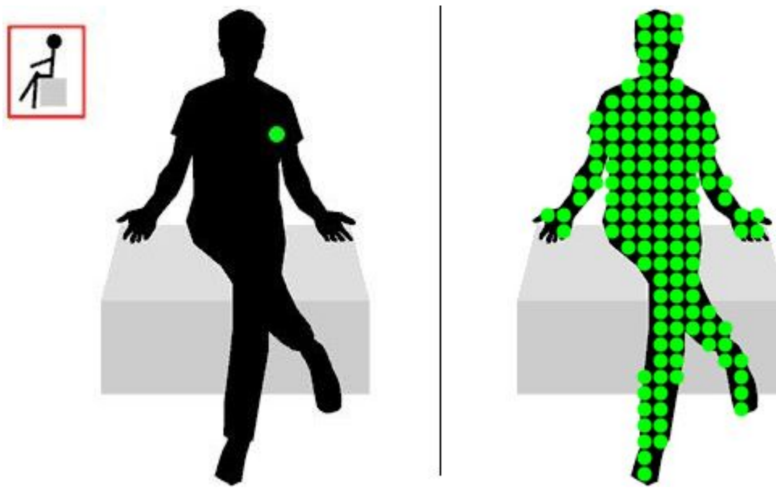


Fig. 8.1 **Left: an example silhouette and dot location. Right: Visualization of all dot locations if rendered altogether (136 total for this pose). Across all seven poses, seen in Figure 8.3, there were 952 dot locations.**

Participants were told to "imagine this silhouette is your body" and also to "imagine you are in a public place and there is an interface projected onto the location of the green dot." Below the silhouette, a single question was asked. Participants were randomly assigned to one of four conditions upon starting the experiment, each featuring a different question, found in Table 1. Participants answered this question using a five point Likert scale: 1) very uncomfortable, 2) uncomfortable, 3) neither uncomfortable nor comfortable, 4) comfortable, and 5) very comfortable. In total, participants entered responses for 20 silhouettes, each with a random pose and dot location.

We chose to exclude the back of the body. Not only would this have doubled the number of points we needed to test, but more importantly, we found the back confused people in piloting, as it is not a surface that can be practically used (e.g., hard to reach, hard to see). Additionally, the decision to use the somewhat ambiguous scale of "comfort" was difficult, but deliberate. Of the many adjectives we considered, comfort best captured the multi-dimensional and highly personal nature of touch [Hertenstein 2006; Manning 2006; Wilson 1998]. In piloting, users interpreted "comfort" in both the physical sense and also with respect to social ease – paramount aspects of our exploration. Likewise, we also elected to use the somewhat indistinct phrase of *someone you are close with* (borrowed from [Jones 1985]), as there are many different types of relationships that do not necessarily follow a linear progression of intimacy. In all three cases, these were calculated experimental compromises aimed to control combinatory explosion (e.g., separating friends, family, co-workers, strangers, etc.).

Finally, we instituted two reliability checks to ensure a high level of answer integrity. First, questions 1 and 10 were repeated in the set of 20. A participant's data was dropped entirely if answer pairs differed by more than one Likert point. Further, participants with low answer variance were also dropped (e.g., all 3's).

	Self	Others
Touching	*How would you feel about **touching** an interface projected on this location with your fingers?*	*How would you feel about someone you are close with **touching** an interface projected on this location with his or her fingers?*
Looking	*How would you feel about **looking** at an interface projected on this location?*	*How would you feel about someone you are close with **looking** at an interface projected on this location?*

Table 8.1 Participants were assigned to one of four question conditions, exploring self vs. others and touching vs. looking. Bold and underline used to only here to highlight differences.

8.2.3 Participants

2,496 Mechanical Turk workers completed the study, who were paid $0.25 for their involvement. Participation was limited to those residing in the United States. Data from 484 participants (19.4% of the data) was dropped as per the quality check procedure noted in the previous section (a proportion inline with other mturk-based studies). Subsequent discussion is based on data from the remaining 2,012 participants (63% male, mean age 27.0, SD=9.1). 84.8% reported being right-handed, 12.8% left-handed, and 2.4% ambidextrous, matching the general population [Porac 1981; Holder 1997].

8.2.4 Generating Heatmaps

Our participants produced 36,216 responses to the 952 silhouetted interface location questions. Within each question condition (Table 8.1), we ensured that each body location had no fewer than five responses – this was important to reduce noise in our models (the average location had 13.2 responses). Combining all four question conditions, each location had no fewer than 23 responses (mean 52.7). As noted before, locations formed a grid with ten-pixel spacing (Figure 8.1). To produce filled heatmaps, we used bilinear interpolation (see Figure 8.2 for color scale). Figure 8.3 displays general models (i.e., combing condition, gender, etc.) for the seven poses.

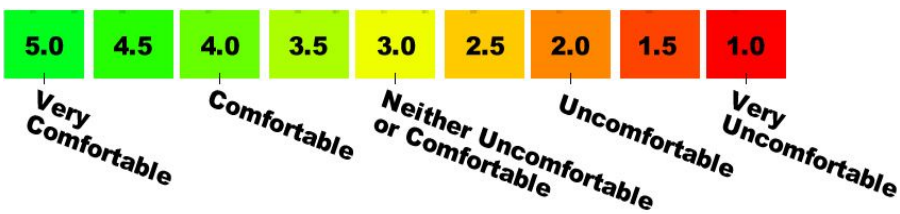


Fig. 8.2 Color scale used for the crowdsourced heatmaps.

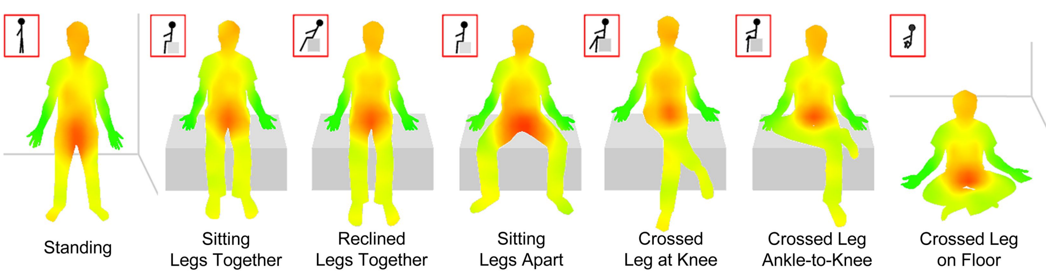


Fig. 8.3 The seven poses included in our exploration. Heatmaps superimposed here are combined models, including all study factors.

8.3 Expert Interviews

The crowdsourced heatmaps provided a high resolution, tip-to-toe visualization of appropriateness. However, standing alone, they were not particularly descriptive, offering only a thin shell of information. It was easy to identify *where*, but not *why* a body location was well suited to on-body interfaces. Also missing were special considerations, for example, transient factors that might enable a body surface for a passing moment. To fill out our *where* model with an equally rich understanding of *why* prompted us to initiate in-depth interviewers with experts.

8.3.1 Expert Participants

Unfortunately, due to the nascence of the field, there are no on-body interface experts; few people have had hands-on experiences, even inside of research institutions. Thus, we solicited expert participants from related backgrounds, which broadly covered three main areas of interest: 1) kinesiology and position, 2) ergonomics and physical comfort, and 3) aesthetics and social aspects. We found that most of the experts we interviewed touched other people as part of their work. This allowed our experts to not only comment professionally and personally on aspects of touch, but also provide considerable insight into how others feel and react to being touched on different parts of their bodies. Brief descriptions of our ten experts follow, including abbreviations used subsequently:

Boutique Manager (BM) - Female, 24. BA in Fashion Design and Merchandising. Three years experience in corporate retail, and three years experience managing a boutique. Helps to select clothing lineup each season and directly assists customers.

Dance Instructor (DI) - Female, 28. Has been dancing since age 13. Four years experience as instructor at a dance studio. Specializes in west coast swing, a social dance.

Accessory Designer (AD) - Female, 49. Started work in bridal shop making accessories. Now has almost 20 years experience as accessory designer (e.g., hats, handbags, scarves). Costume design and seamstressing on the side.

Chiropractor (CP) - Male, 72. "Touch is what I do." Has a deep understanding of where and why parts of the body hurt. 46 years experience; owns practice.

Jewelry and Metalsmith (JM) - Female, 35. Completed a BFA and MFA. In addition to producing jewelry, has been teaching jewelry and metalsmithing classes for 12 years.

Massage Therapist (MT) - Female, 26. After completing university, enrolled in 700+ hour massage therapy course. Licensed and practicing for almost three years.

Yoga Instructor (YI) - Female, 29. Three years experience as an instructor. Deep awareness of body and positional issues.

Tattoo Parlor Owner and Body Modification Enthusiast (ME) - Female, 33. Wanted to be a "tattoo lady" from early age; now owns and manages a parlor. Has tattoos over much of her body, including face, as well as stretched ears, lip piercings, and microdermal implants.

Tattoo Artist (TA) - Male, 41. Fifteen years professional experience; has worked on thousands of tattoos.

Ergonomics Compliance Manager (EC) - Male, 70. Certified ergonomic compliance manager, certified safety manager, and an OSHA-authorized trainer. 18 years experience. Extensive hands-on knowledge of injuries relating to technology use.

8.3.2 Procedure

Interviews lasted approximately one hour; expert participants were paid $40 for their time. The interviews started by discussing what their jobs involved, training or experience required, interesting professional experiences, and similar background matter. Following this general discussion, participants were shown the same 30-second video montage of on-body interfaces used in the crowdsourced study. Following this, they were allowed to ask questions and comment on what they had seen. The interviewer elicited additional discussion by asking participants if they would use such a technology and also to imagine possible benefits and drawbacks, especially in relation to current mobile devices.

Next, participants were presented with an outlined silhouette of a person standing. They were told to imagine this was themselves, and that on-body interfaces like they had seen in the video could appear anywhere on their body. Three marker pens were made available: green for locations that were most appropriate, red for poor locations, and blue for locations that were somewhere in between, or neutral. Participants were told to color the silhouette and think aloud during the process, commenting both personally and from their professional experiences. This exercise (see Figure 8.4 for examples) proved highly effective in getting participants to systemically consider all parts of the body, and vocalize the multitude of reasons why different areas would be good and/or bad. Frequently, participants touched and reconfigured their own bodies to help with their assessment.

After completing the standing silhouette, participants duplicated this exercise for the remaining six poses. It became clear after the first two interviews that appropriateness for much of the body remained unchanged across poses, the torso in particular (reflected in Figure 8.3). Thus, participants were told to focus primarily on differences from the standing pose, which was left on the table for comparison.

After completing all seven poses, participants were once again presented with an unmarked, fresh silhouette of a person standing. Participants were told to imagine again this was their body, however this time, instead of touching themselves, it was someone close to them touching. When appropriate, this was interleaved with discussion about how they would feel if a perfect stranger asked to use their body for an interaction (e.g., to phone a taxi). After the interviews were completed, they were transcribed, coded and synthesized using affinity diagrams [Beyer 1999] to identify common themes.

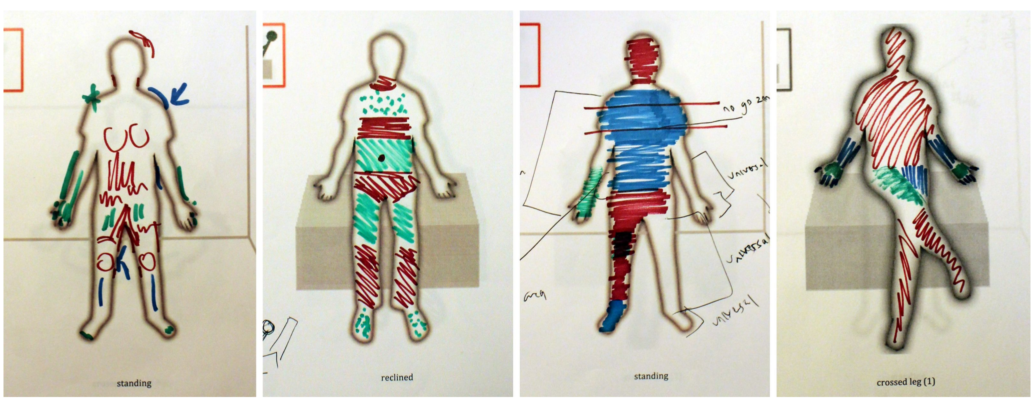


Fig. 8.4 **Four example silhouettes "filled in" by experts during the think aloud exercise. Three pens were provided: green for 'good' places, red for 'bad', and blue for neutral.**

8.4 Results and Discussion

Our crowdsourced heatmaps provided a high spatial resolution view of appropriateness. However, the explanatory and descriptive power of any one location was low given the underlying Likert-style data. Conversely, our expert interviews had low spatial resolution - perhaps only ten broad areas of the body were discussed vs. 150 points. However, for these areas, we gathered rich and detailed notes. In this section, we combine these complimentary results, highlighting core issues regarding appropriateness across the human form. In particular, this enables us to discuss the *where* and *why* in tandem for the first time.

We include several heatmap models to help guide and ground our discussion. Breaking these out by pose, gender, age, touching vs. looking, and self vs. others yields hundreds of different heatmaps, which is overwhelming to process and discuss. Thus, we have elected to combine results where appropriate in order to highlight key findings, along with a digest of quotes from our expert interviews.

8.4.1 Expert Reactions

Most of the experts interviewed found the technology vision appealing, and many said they would use it today if available. Others were less enthralled by the thought of having touch interfaces projected onto their bodies, but nonetheless articulated many potential benefits. Universally noted was the immediate accessibility: “Access would be easier” (BM); “So easy, it's always right there!” (AD); “Hands free, less to carry, its always with you” (EC). “If you need something, it's right there, you don't have to remember it.” (ME). “It's just convenient, you're not carrying a lot of stuff, it's always there” (TA).

Less apparent was that the “display takes up no space” (DI), “it's not something you have to stuff into your pocket and lug around with you” (TA) and potentially could offer even “bigger space” (CP). The potential for eyes free operation was also remarked on: “Skin is really comfortable. [Interviewer: ‘as comfortable as a touch screen?’] Maybe more. Your hands are super sensitive - you can even do it without looking as much” (JM). Finally, EC noted, “I don't see [such interfaces] being a big ergonomic issue.”

8.4.2 Body Overview

Experts were fairly decisive and universal in their preferences for locations across the body. With very few exceptions, experts strongly aligned with the heatmap data. Most locations were quickly brushed off as being inappropriate, in favor of the arms and legs (discussed in depth subsequently). The biggest reason to disqualify a location was visual accessibility.

“When standing, there’s not enough places on the body you can see really well” (JM). “It'd be weird to have interfaces on your legs, or body. How are you going to see it?” (AD) The head draw the obvious criticism: “[you] can't see your own head” (JM), “you'd have to look at a mirror” (TA). Similarly, the back is “inconvenient” (YI).

The front of the body is also problematic: “Chest is lot of surface area, but looking down at it, you’re not going to see it all” (BM). “How can you see it? [Even reclined] you would have to crane your neck” (CP). With specific respect to the upper chest: “I can't read here.” (DI). The stomach area: “wouldn't find that area that useful, and it's kinda weird” (JM); “Weird to project on the stomach” (DI).

Figure 8.5 shows heatmaps in a sitting pose separated by touching and looking at interfaces at one’s own body. It is readily apparent the trunk of the body is considered visually inaccessible or awkward (right), with mean comfort ranges between 1.52 and 2.72 (mean 2.3). More areas, however, are physically accessible (left) by reaching with the hands.

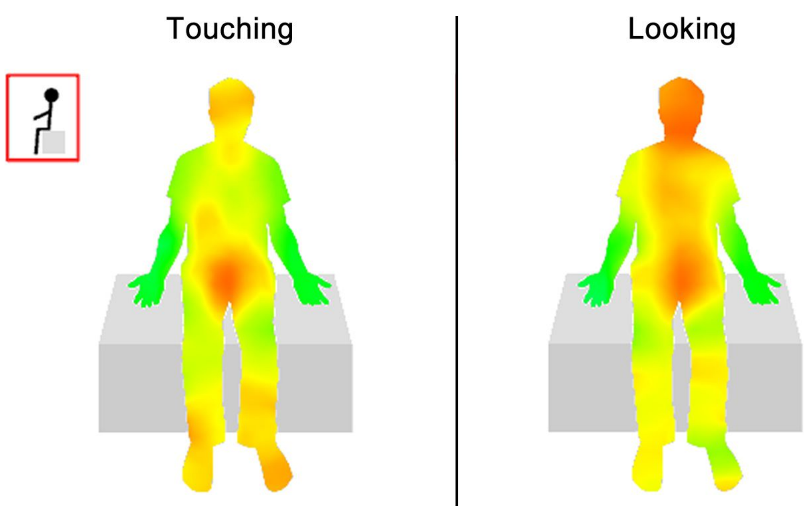


Fig. 8.5 Heatmap models in a seated pose separated by touching and looking at interfaces on oneself.

8.4.3 Hands and Arms

By far and away, the most favorable and universally accepted area of the body for projected touch interfaces was the arms (across all poses, mean 4.3, see Figure 8.3). Not surprisingly, this is where past on-body systems have focused their efforts and previous research has confirmed as the least venerable [Jones 1985]. Among a multitude of reasons, visual and physical accessibility were the most commonly cited, matching results in Section 7. "You can flip it into whatever angle you need" (BM); "Hands are most visible" (JM); "The arm - its right there!" (AD). "[Arms] are always there in front of you, you can bring them up" with visibility "similar to a book or newspaper" (EC).

Also commanding much discussion were aspects of control. "I feel like hands would be pretty intuitive… two [hands] working together" (TA). "You have a lot of control on your hand over other places" (JM). "You are used to touching things with your hands. That's how you touch [the world]" (CP). "You can feel much more small gradations in movements [on your palm]" (TA).

Pose comfort did not appear to be a significant issue. "You can tuck [the arm] in close and not support it" (TA). "Most natural is straight out like this [*palm up, forearm supinated*]… Keep it in a plane, elbow at 90 degrees." (EC). On sustaining the latter arm pose, CP noted, "I can't imagine that would be a problem for many people," even older users. Additionally, two experts commented that the hands could be supported by "resting arm on the thigh" (EC), or perhaps a table or armrest. "If you have a support surface, you can be kinder to your body" (DI).

These benefits did not extend above the elbow; the upper arm was generally disliked from a visibility standpoint (see Figure 8.5). "It's closer to focus, and has a weirder angle" (MT); "Above the elbow, how are you going to look at it?" (CP); "Angle is kind of bad" (BM); "Forearms up to elbow. Not above, awkward posture … very close to your eyes. [Lower arms are] better from a comfort and ergonomics stand point" (EC). DI noted, however, that the upper arms could serve as a "peripheral" display.

The second most common arm posture demonstrated was similar to how one would read a wristwatch (arm held up, elbow flexed, forearm pronated). While convenient, and exposing the desirable back of forearm (discussed next), it is "not something people can sustain" (EC) for extended periods. However, repeated use, analogous to training, might mitigate this issue: "There are styles of belly dance where your arms don't go below shoulder level for half an hour. You can train to do that" (DI).

The irregularity of the hand was identified as a potential usability issue. “[Hand is better for] applications that aren't so visual… like an instagram on your hand probably wouldn't be that great,” but for typing an email “it'd be fine” (TA). The “inside [of the arm] is best, no hair on it, smoother, clearer, flat. Next best is the palm, then back of arm, then back of hand” (AD). Although experts were generally positive about the inner arm, it was tempered with sensitivity issues. For example, the inner arm “skin is more sensitive to touch” (YI) as “veins and tendons are closer to the surface” (MT). Similarly, “the back of the hand… skin is thinner, has more nerve endings than forearm” (MT). The outside of the forearm drew less concern “because the skin is thicker” (YI) and less sensitive. “Better off on the outside. Poke in here [inside of arm] you could hurt someone” (CP). Overall, however, “people can usually tolerate pretty firm pressure [touches] in the arm” (MT).

Fingers were notably absent from commentary during the think aloud exercise. When prompted: “I don't really see a huge advantage from being on the fingers” (TA). Although providing a contiguous extension from the palm when together, it is not a neutral pose: “I can hold them together, but it takes some force” (EC).

8.4.4 Thighs and Legs

Though the response was less enthusiastic, the thighs were also frequently mentioned. They were universally appreciated as affording more room for interaction, assuming a seated pose. The “tops of your thighs ... have a lot of room” (TA), “would have more surface area, if you are reading something with longer text” (YI), and “are relatively flat” (CP) for better projection.

Using the legs as the interactive surface frees the hands, enabling bi-manual interactions. Of particular interest was the ankle-to-knee crossed-legged pose, which opens a large, contiguous and smooth surface for interaction on the calf, which is oriented perpendicular to the body. “The inside of your calf… this is funny because when you tattoo yourself … that’s where you do it, because you need both hands ... and I can see that working here… hmmm, you could type [with both hands] on your leg” (TA), “like a keyboard” (EC), “the mid section of your lower leg crossed over, like a screen” (YI). This opportunity was also reflected in the heatmaps (area mean of 3.02), the highest rating of any leg region except when cross-legged on the floor.

Two of the seated poses we explored had legs that were crossed. Several experts raised ergonomic issues: “Not healthy to cross your legs… I don't like the crossed [one] where your legs are really tight, because that is bad against blood flow" (YI). EC noted that "anytime you have contact, bone to bone, but more importantly, compressed nerves, compressed blood flow" there is going to be a fatigue and comfort issue. Experts noted that people "tend to switch legs constantly" (DI) in this pose. This isn't necessarily detrimental, but does impede prolonged use. In regards to looking at such interfaces on the legs, "you'd really have to hold your neck down … after a few minutes I'd have to [*stretches neck*]… the neck and upper back will start to feel it" (EC). This was much less of an issue with the arms.

Likewise, there were points raised regarding touch sensitivity. "Almost everybody is more sensitive in the legs" (MT). In particular, "knees are very sensitive" and "shins can be tender" (CP). Experts tended to highlight the ideal location for interaction starting at the mid thigh and ending just before the knee: "[with] the length of the arm, your arm naturally goes to your knees" (MT). "Maybe halfway down the thighs, beyond that you're going to be pulling your arms back" (EC), forming an acute elbow angle and putting stress on the shoulders and upper arm. Simultaneously, this arm pose also keeps the hands away from the groin (discussed next). This preference towards the lower thigh is confirmed in the heatmaps (Figure 8.3).

Finally, feet (and lower legs) were quickly dismissed as practical locations for on-body interfaces. When discussed, they were always portrayed as being input poor. "You can use the feet for any sort of display that didn't require something else to poke at it … it's not going to be as detail oriented as you can be with your hands" (DI). "It'd be kind of far away to do anything" (TA) and "text is going to have to be larger [to see]" (DI).

8.4.5 Groin Considerations

Unsurprisingly, interaction on the "crotch would not be the most appropriate. In public, it is always a bad area" (BM). "You don't want to have to poke yourself in the crotch to hit OK" (DI). "I don't know if you'd want to be manipulating anything on your crotch, because it just looks bad. But as far as, like from a technical stand point, it's not, I don't think… if you are me, you can't even see your crotch if you are sitting down, but if you don't have a pop belly.. then.. I think that would be more of a social thing" (TA).

DI explained: "There two things that affect the crotch areas availability and sensitivity. The thing that affects availability the most is the angle between the legs and torso. If you are standing, there is more area here, and if you sit, there is less area that is accessible. The thing that affects the sensitivity is the angle between the two legs - the further your legs are spread, the more dicey it feels to be touching there" (DI). When the legs are crossed, the groin "doesn't exist as much, tucked underneath you" (BM). This effect is easily seen in the heatmaps (Figure 8.3): e.g., legs apart (min 1.29, mean 1.72) vs. crossed legs at knee (min 1.83, mean 2.10).

8.4.6 Gender and Body Image

There was a great deal of commentary on gender differences and the "politics of touch" (MT). Beyond the obviously taboo groin area, neither male nor female experts described any front-facing, reachable areas that would be totally off limits for men. On the other hand, women produced a litany of locations that would be off-limits on the female body. "Arms, whole arms, are more universally [acceptable] for men, no matter what shape your arm is in. Women think their arms can be too skinny or too fat - men don't seem to think about it, it's just there" (BM). "Women are sensitive about their legs. Some people are weird about feet. There is a whole degree of women out there that won't wear a short sleeve shirt" (AD). "[Chest] area is off limits" (DI). "I don't know many women that really like to look at their thighs and study them, because they are so critical already, or belly or breasts, like all these hot spots" (YI).

Concern stemmed primarily from drawing attention to sensitive areas. "It all really depends on what relationship you have with your body" (YI). "People who are self conscience or cautious about those part of their body wouldn't want to be touching them, wouldn't want screens on them." (BM) "For a woman in particular, I wouldn't want something glowing on my chest, drawing attention, in the same way I wouldn't want something super tight" (BM). These "hot spots" are not universal, and are influenced by factors such as body type: "For example, say if you were an apple shaped person, larger on the top. I wouldn't want anything clinging to me, touching me, lighting up here – stomach, torso, chest, and shoulder. If I were really pear, like big around the hips, I wouldn't want anything lighting up around there" (BM).

Looking at gender-split heatmaps (Figure 8.6), this conservatism is reflected strongly, particularly with the upper torso and shoulder areas. These perceived gender complexities are nicely illustrated by a male/female pair of filled silhouettes (Figure 8.6, right) completed by BM, who has to accommodate body issues when assisting customers at her boutique. Unlike woman, “the chest region on a guy, and shoulder area, would be ok,” though the “hips, love handley area, is more sensitive “ (BM).

Of particular note, the hands and lower arms appear to sidestep these body image complexities, with both men and women. “Hands are universal - elbow down. Above elbow is difficult, more sensitive” (BM). “The [lower arm] is totally safe. Nobody ever says: ‘oh my, my forearms are so fat!’” (AD). This again is strongly reflected in the heatmaps.

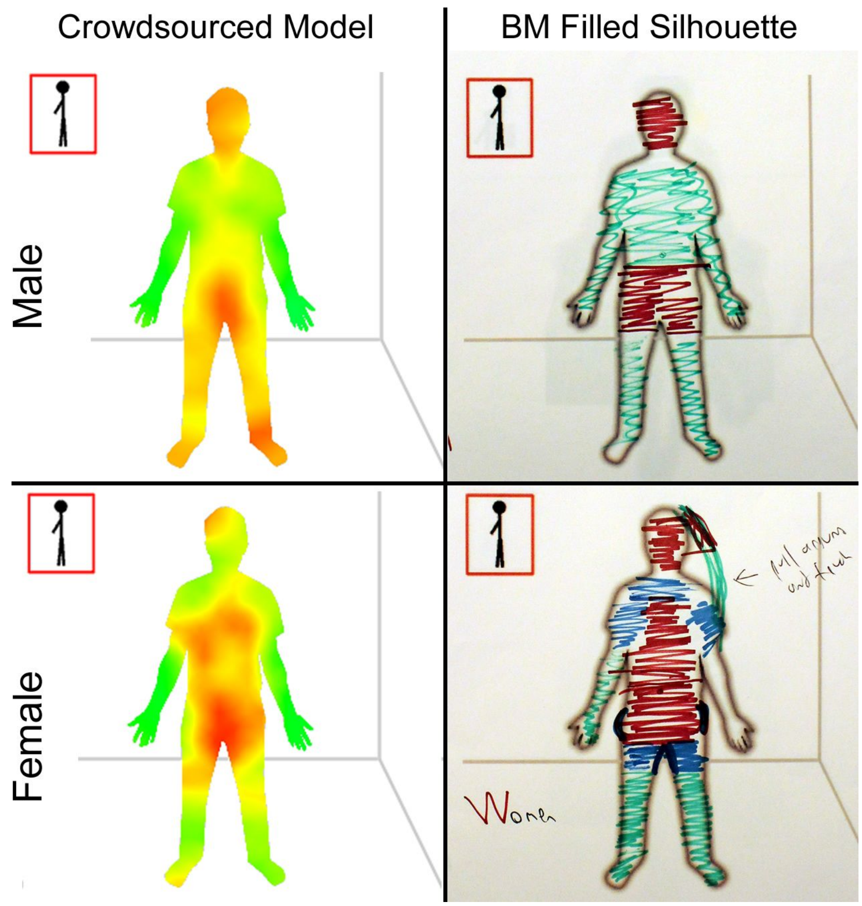


Fig. 8.6 Comfort of touching interfaces appearing on ones' own body, broken down by gender.

8.4.7 Others Touching

The second half of the think-aloud exercise was concerned with how people would feel if another person asked to use an interface on their body. Questions surrounded what part of themselves would they "lend" first, and last, and why? This turned out to be an incredibly complex subject, fueled by opinions on e.g., personal boundaries, gender power imbalances, and physical [Hobye 2011; Jones 1985; Tomita 2008].

Foremost, it is important to note that resistance to lending something to a stranger is not unique to the body or on-body interfaces. "I am so surprised by how people do not share cell phones. It is rude to ask someone at the airport, can I please use your cell phone? [A phone is] very personal" (YI). Giving someone a communications device is act of trust: "You can give them a pen, and you don't necessarily need it back. A phone, you definitely need back. Also [with a pen] they don't have access to your personal information, your contact information" (MT).

Overall, people seemed comfortable with close friends and family using an interface on their body. "If it were my friend, my hand would be ok, my arm would be ok" (CP) – matching how people would use their own bodies. However, reactions to unfamiliar people touching were much stronger. "Feels like loosing control.. vulnerable" (YI); "I would really hesitate" (JM); "It's an invasion into that persons space" (CP); "I would feel that my personal space would be invaded" (MT).

When dealing with strangers (e.g., someone at a bus stop), two body locations were repeatedly suggested by our group of experts. Foremost was the arm, which is also popular for oneself. "The outside of the arm [is an] impersonal area" (JM); "It's not unusual for people to touch your arms" (TA); "Back of the forearms: combination of most practical to use and lease invasive… It's not up here, close to where the breasts would be for a woman. Its not the hand, which that can feel really intimate - holding hands is a particular type of touch." (MT). Other experts were more favorable to the hands: "Just my hand. It seems less personal. And I touch people everywhere. Back of hand is preferable, it seems less germy" (AD); "Only hands. That's what you'll give a stranger, you'll shake their hand." (TA); "Hands - it's where you are least vulnerable" (EC).

The other body area frequently suggested by our group of experts, which did not elicit much commentary prior, was the back and shoulder blade area. "I wouldn't mind lending my shoulder" (BM), "[it] feels like it's really impersonal" (JM). "If I were going to tap that guy on the shoulder, that's where I feel like I can touch that person there without making them mad" (JM). "It's almost like strangers can have here, it's like showing someone the cold shoulder. It's good on the bus. It's protecting me more." (YI). "Just like you turn your back and someone with a paper writes on you. You don't do that on your front, or head, or somewhere else" (DI). Further, from a touch sensitivity standpoint: "that's pretty dead back there relatively speaking" (CP).

Regardless of preference for hand or back (which matches [Jones 1985; Tomita 2008]), the other commonality was a desire not to have to face the person. "Back of shoulder, because they are not interacting with me, they are interacting with my surface. Having someone facing me and interacting with something on me is too social" (DI). "For a stranger: shoulder first… don't want them facing me" (YI). "[It's like being] face to face in an elevator, it's uncomfortable" (EC). "Front is a little more personal. Back is less personal" (CP). "[back] might be better than the forearm for strangers because you don't have to look at them" (MT). Similar comments were echoed about being constrained. "I'm happy to give my body as a display for communication, but I don't want to be constrained in my movement" (YI). "A strange man grabbing your arm or wrist is a very threading gesture. Someone touching your back over your clothes ... doesn't have the cultural association of power and threatening" (MT).

Others touching the legs for interactive purposes was generally viewed as impractical. When standing, "everywhere on the legs other than the upper thighs are awkward [for others to reach]" (MT). However, people were not keen on lending their thighs: "Wouldn't let them feel any higher than my knee" (BM); "The thighs are an awkward place to touch someone" (MT). Both YI and DI mentioned rules of thumb in their respective fields, as they often have to touch students as part of their instruction. These reinforce the primacy of the hand of shoulder as acceptable areas of touch between less familiar people. "Where is it boney, it's ok. But, where it is soft, that wouldn't be really good [for touching]" (YI). "This is a well established concept in social dance. There are five connection points that are acceptable: Shoulder, elbow, hands, hips and knees" (DI).

8.4.8 Pose Switching

In a few instances, experts noted they might switch poses to accommodate an interaction task. For example a person might sit "if [his/her] hands were carrying something" (EC). The reclined position in particular drew comments: "[I] would prefer to sit up and use my legs" (JM) "Just get up. If we've gotten to a point in society where we can't fucking sit up for five seconds, and we need another area [for input], just kill me" (ME).

8.4.9 Symmetry

Two experts noted the importance of "symmetrical activities" (CP). "You wouldn't want to make any tool that makes the body sided" (YI). "Asymmetrical sends me business, either long time or repeated" (CP). For this reason, CP suggested the hands raised in front of you is "a natural place … you don't have to turn your neck," whereas turning the head towards either side is "uncomfortable for neck" (YI).

8.4.10 Final Thoughts on the Acceptability of On-Body Interfaces

Several experts mused on the acceptability of such on-body interfaces if there were to be introduced. It was generally accepted the barriers surrounding touching oneself would begin to loosen if such technologies went mainstream. "I guest it depends on how commonplace this stuff becomes. If everyone is doing it then you aren't going to feel odd. If you're the first guy, you're probably not going to want to start with the inner thighs" (TA). "If this becomes significant for writing emails, touching even in the inside [of the groin] would become fine" (YI). Experts also made allusions to the apprehension surrounding the transition to mobile phones. "It's like when you got rid of your land line, and everyone just went all cellphone, there was a little while there me and other people my age where like, what if something goes wrong, you should have this backup landline. Which, of course, it didn't matter it turns out" (TA).

With respect to learning how to use an on-body system for the first time, DI put it eloquently: "When you start dancing with a partner you've never danced with before, you don't bust out the crazy moves right away, you start with the simple stuff and see how they react. You sort of build up a familiarity of the technical capabilities of your partner to find a good match. I feel that with a projection system like that, there is definitely going to be the first song, the first pancake - it's going to be kind of awkward."

8.5 Conclusion

In the previous two chapters, my aim has been to shift the question away from *how* to make on-body interfaces (Chapters 3 through 6), and towards *where* they should manifest on the body. These are the class of questions that inform the design of future systems and validate some design decisions used in prior systems. In this chapter specifically, we have shown there are a variety of complex considerations – ranging from biomechanical to social – that affect where interfaces might manifest on the body.

9 CONCLUSION AND FUTURE WORK

The previous eight chapters have formed an extended arc. We started by exploring fundamental issues in mobile computing, which illuminated the skin as an interesting and underexplored interactive surface. We then discussed a series of systems, each pursuing a radically different sensing approach for appropriating the skin for interactive purposes. We then shifted our focus to questions about how on-body interfaces should manifest on the body, irrespective of the technology used. Although each chapter has offered insights that stand alone, I believe the concepts, design recommendations, study results, and technical advances are synergistic, and help lay a solid foundation for future efforts in the budding domain of on-body computing. To conclude, I now summarize key contributions of this dissertation and also where I believe important areas of future work remain.

9.1 Summary of Contributions

Above all, this dissertation contributed three novel sensing approaches to appropriating the human body as an interactive computing platform: bio-acoustic, vision-driven, and capacitive. These efforts significantly advanced the level of interactive sophistication over previous systems, bringing many proposed interactions to reality. Moreover, these systems achieved a new level of robustness, demonstrating the feasibility of on-body interfaces. Although more technical work is required, a path towards popular adoption is now illuminated. Technologists will be able to build on top of the systems described here, interaction designers will be able to incorporate and study them, and one day, end users will be able to enjoy them.

Equally significant is the contribution to interaction design. This work endeavored to establish a series of elemental principles and example interactions to ground and guide the design of future on-body systems. This served as an important first step, as the interaction design community typically does not have access to early prototypes, such as those described in this dissertation. Further, the work provides a springboard for interaction designers to take on-body interfaces to the next level. Finally, technically-oriented researchers can apply the knowledge described in this document to on-body systems they are developing, pushing the field forward into new areas.

9.2 Future Work

Since work in the area of on-body interfaces has only recently begun, there are many areas for future work. These include exploration of the properties of new sensors and overall sensing modalities. Further, it is not known how much of the full dexterity of arms, hands and fingers can be reliably recognized and how much of it can be usefully brought into an interactive system. There are also ergonomic issues, such as flexibility and fatigue. An immediate step is to delve deeper into the many issues and avenues uncovered in this dissertation work, such as on-body clutching, supported by rigorous user studies. Topics range from the commonplace (how does proprioception affect Fitts' law performance?) to the amusing (how do people feel about getting their buttons pushed?). We describe two main thrusts of future work below.

9.2.1 On-Body Technologies

Skinput, OmniTouch, Touché and Armura served as proof of concept that the human body could serve as an interactive platform. It answered the fundamental question of whether or not it was possible from a technical standpoint. The challenge now is how to move forward – making on-body technologies and interactions more mobile, more robust, and able to support richer applications. This "polish" is necessary if these interfaces are to move out of the lab and into the real world.

One important area of future work is combining sensing methods – a *sensor fusion* approach. For example, the techniques employed in Skinput, OmniTouch and Touché are not mutually exclusive, and could be combined. The resulting ensemble would be superior to any one technique, yielding a more capable and robust experience. For example, Skinput was exceptionally robust at segmenting input, but had low spatial resolution (e.g., five predefined locations). On the other hand, OmniTouch had excellent spatial resolution and supported multitouch; however, it's optical approach suffered from false positive touches if the fingers hovered too close to the skin. A system combining these methods would inherit the strengths, but not the weaknesses. Thus, it is important to look for ways to combine sensing technologies as well innovate new ones. Although Skinput, OmniTouch and Touché explored a significant swath of the sensing landscape, more sensing approaches remain to be implemented and tested.

One underexplored sensing approach is passive time difference of arrival (TDoA) [Ishii 1999, Paradiso 2000]. Like Skinput, this approach listens for mechanical vibrations created by finger touches. Instead of using acoustic fingerprinting (and machine learning), a distributed array of vibro-acoustic sensors attempts to localize the source of the impact based on timing (similar to GPS triangulation). This approach has different pros and cons, which makes it a useful compliment to the sensing approaches already explored. Foremost, it could allow for X/Y localization of touches without training (though finger drags would be challenging, if not impossible). Secondly, the technique does not require cameras, which removes occlusion and line of sight issues. However, sensors would have to be distributed on the arms or hands to capture the acoustic information, which is more invasive. It remains to be seen if the technique could be integrated into a small form factor such a wristwatch.

Another area ripe for future exploration is improving on-body *output*, particularly projected graphics. Skinput, OmniTouch and other related work [Mistry 2009; McFarlane 2009, Yamamoto 2007] placed projectors relatively orthogonal to possible projection surfaces. For example, Skinput used a cantilevered pico-projector strapped to the upper arm, while OmniTouch was shoulder-mounted. Although yielding high quality graphical output with minimal distortion, it also makes the hardware fairly cumbersome and socially awkward. Further, because the projector is decoupled from the projection surface, active tracking is required to keep graphics locked to the body, which is computationally expensive and potentially error prone.

An ideal approach would use projectors operating near to the skin, operating at oblique angles. For example, a wristwatch with integrated projector might one day allow for the whole arm to be illuminated with graphics. This is a much more palatable form factor for users, though it comes with significant technical questions that will require investigations in optics, sensing, electronics, as well as human factors.

9.2.2 On-Body Interaction Design

The outcomes of Skinput, OmniTouch, and Touché were primarily technical, grazing only the surface of interactions enabled by bringing touch interaction to the body. Armura, and the studies described in chapters 7 and 8, began to investigate basic aspects of interaction design, such as visual accessibility. However, there is a standing opportunity to explore the rich interaction space enabled by on-body interfaces.

A key component in the success of mobile computing has been reduction of interaction barriers. Instead of being tethered to a computer, the user is free to move about. Combined with mobile connectivity, one is able to rapidly take out their smartphone, look up something on Wikipedia, and settle a dinner debate with incredible ease and fluidity. The larger vision of on-body computing is to reduce this interactive viscosity again, by literarily having computing power in the "palm of your hand" and always available. This, fundamentally, is an interaction design challenge, supported by technical advances.

The human body is interesting to consider from a design standpoint. Unlike most computing devices, we cannot control the form; we can augment it, but not change it. However, this also means its capabilities are fixed – humans (in general) are a known quantity. For example, there are only so many ways we can move our limbs, defined by set of biomechanical constraints. Output is limited by our senses (barring dramatic BCI advances). Since the I/O capabilities are predefined for us, the question then becomes, what can we do with the human body? This is somewhat analogous to being handed a ball and told to invent a new game. The design possibilities are constrained in some dimensions, and limitless it other ways. There is no one answer to this complex question; millennia of dance tradition shows us innovation continues to be possible in something as fundamental as human motion.

For example, it will be interesting to explore if and how the unique contours and affordances of the body can be leveraged to make interfaces more natural and powerful. Chapter 6 (Armura) demonstrated that the arms and hands can be used in creative ways; Chapters 7 and 8 suggest other areas of the body area applicable for interactive experience. However, we do not yet know if the unique motor and musculoskeletal dimensions of these locations cam be utilized for interactive purposes. Further, it is likely that different parts of the body carry different connotations with respect to interactive functionality. For example, consider a music player application rendered on the arm. It seems logical (though this would have to be tested) that core functionality would reside near the hand, if not directly in the palm or on the fingers, and that secondary presentation of information (e.g., songs) might be rendered in a list running down the arm.

Thus, it is an open question about how to layout and structure interactive functionality on the body. There are well-established principles for GUI design in desktop and mobile applications (as well as a huge body of work on automatic layout and interface reflowing). Do these principles and findings hold true in the unique context of the body? Is the standard desktop widget set applicable on the body? Skinput and OmniTouch followed the "fingers poking buttons" interaction seen in conventional touchscreens. Given the unique expressivity of our bodies, do we need buttons? Can more actions be gestural, as Touche and Armura suggest? Moreover, Skinput, OmniTouch and Armura focused on graphical interfaces that were 2D and primarily rectilinear – our bodies, however, are neither. If we are to have interfaces on organic surfaces [Holman 2008], it will be valuable to rethink classic interface paradigms, and consider how our unique form can contribute to the computing experience.

9.3 Final Remarks

For some, on-body interfaces may seem like an uncomfortable direction to take computing. However, I believe on-body interfaces can feel natural and intuitive if the design is informed. Much of the comfort will lie in implementation specifics. For example, there is nothing particularly natural about grasping a small rectangular device in one's hand and poking fingers at it. Yet, good design has made interacting with mobile devices second nature.

Moving interfaces onto the human body is a dramatic a jump – perhaps similar in magnitude as the transition from desktop to handheld computing.

Given the enormous volume of work dedicated to mobile device interaction, it is not hard to imagine the human form will spark even more research. Interfaces on the body are constrained in unique dimensions and unbounded in other ways. Although this expansive design landscape is intimidating, it also signals the tremendous interactive possibilities that await future on-body systems. I hope this dissertation can serve as an early step in that direction.

10 BIBLIOGRAPHY

1. Accot, J. and Zhai, S. More than dotting the i's -- foundations for crossing-based interfaces. In Proc. CHI '02. 73-80.
2. Agarwal, A. and Triggs, B. Learning to track 3D human motion from silhouettes. In Proc. ICML '04. 2-9.
3. Amento, B., Hill, W. and Terveen, L. The Sound of One Hand: A Wrist-mounted Bio-acoustic Fingertip Gesture Interface. CHI '02 Ext. Abstracts. 724-725.
4. Apitz, G. and Guimbretière, F. CrossY: a crossing-based drawing application. In Proc. UIST '04. 3-12.
5. Argyros, A. A. and Lourakis, M. I. A. Vision-based Interpretation of Hand Gestures for Remote Control of a Computer Mouse. In Proc. ECCV '06 Workshop on Computer Vision in HCI. LNCS 3979. 40-51.
6. Ashbrook, D. L., Clawson, J. R., Lyons, K., Starner, T. E. and Patel, N. 2008. Quickdraw: the impact of mobility and on-body placement on device access time. In Proc. CHI '08. 219-222.
7. Ashbrook, D., Lyons, K. and Starner, T. An investigation into round touchscreen wristwatch interaction. In Proc. MobileHCI '08. 311-314.
8. Barnett, A. The dancing body as a screen: Synchronizing projected motion graphics onto the human form in contemporary dance. Comput. Entert. 7, 1 (2009). 1-32.
9. Barrett, G. and Omote, R. Projected-Capacitive Touch Technology. Information Display, (26) 3, 2010. 16-21.
10. Bartindale, T. and Harrison, C. Stacks on the Surface: Resolving Physical Order with Masked Fiducial Markers. In Proc. ITS/Tabletop '09. 57-60.
11. Bartindale, T. Harrison, C., Olivier, P. L. and Hudson, S. E. SurfaceMouse: Supplementing Multi-Touch Interaction with a Virtual Mouse. In Proc. TEI '11. 293-296.
12. Bau, O., Poupyrev, I., Israr, A. and Harrison, C. Teslatouch: electrovibration for touch surfaces. In Proc. UIST'10. 283-292.
13. Baudisch, P. and Chu, G. Back-of-device interaction allows creating very small touch devices. In Proc. CHI '09. 1923-1932.
14. Beardsley, P., Baar, J. V., Raskar, R. and Forlines, C. Interaction using a handheld projector. IEEE Computer Graphics and Applications, 25, 1 (2005), 39-43.
15. Benko, H. Saponas, T. S., Morris, D., and Tan, D. Enhancing input on and above the interactive surface with muscle sensing. In Proc. ITS '09. 93–100.
16. Beyer, H. and Holtzblatt, K. Contextual Design. Interactions, 6, 1 (January 1999). 32-42.

17. Blasko, G., Feiner, S. and Coriand, F. Exploring interaction with a simulated wrist-worn projection display. In Proc. ISWC '09. 2-9.

18. Bolt, R. A. “Put-that-there”: Voice and gesture at the graphics interface. SIGGRAPH Comput. Graph., 14, 3 (July 1980). 262-270.

19. Brewster, S. and Murray, R. Presenting Dynamic Information on Mobile Computers. Pers. Ubi. Computing, 4, 4 (Jan. 2000). 209-212.

20. Brown, L.M., Brewster, S.A. and Purchase, H.C. Multidimensional tactons for non-visual information presentation in mobile devices. In Proc. MobileHCI '06. 231-238.

21. Burges, C.J. A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery, 2, 2 (June 1998). 121-167.

22. Butler, A., Izadi, S. and Hodges, S. SideSight: multi-"touch" interaction around small devices. In Proc. UIST '08. 201-204.

23. Buxton, W. and Myers, B. A study in two-handed input. In Proc. CHI ’86. 321-326.

24. Cao, X. and Balakrishnan, R. Interacting with dynamically defined information spaces using a handheld projector and a pen. In Proc. UIST '06. 225-234.

25. Cao, X., Forlines, C. and Balakrishnan, R. Multi-user interaction using handheld projectors. In Proc. UIST '07. 43-52.

26. Cao, X., Wilson, A., Balakrishnan, R., Hinckley, K., and Hudson, S. E. ShapeTouch: Leveraging contact shape on interactive surfaces. In Proc. TABLETOP ‘08. 129-136.

27. Card, S. K., English, W. K. and Burr, B. J. Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT. Ergonomics, 21 (1978). 601-613.

28. Cassinelli, Á., Perrin, S. and Ishikawa, M. Smart laser-scanner for 3D human-machine interface. In CHI EA'05. 1138-1139.

29. Chapanis, A. Theory and methods for analyzing errors in man-machine systems. Annals of the New York Academy of Science 51, Human Engineering (1951). 1179-1203.

30. Cheney, M., Isaacson, D. and Newell, J. C. Electrical impedance tomography. SIAM Review, 41, 1 (1999). 85-101.

31. Cho, C. and Yang, H. Body-Based Interfaces. In Proc. ICMI '02. 466-472.

32. Cholewiak, R. W. and Collins, A. A., (2000). The generation of vibrotactile patterns on a linear array: Influences of body site, time, and presentation mode. Perception & Psychophysics, vol. 62. 1220–1235.

33. Cohen, P. R. The role of natural language in a multimodal interface. In Proc. UIST '94. 143-149.

34. Coyle, D., Moore, J., Kristensson, P., Fletcher, P. and Blackwell, A. I did that! Measuring users' experience of agency in their own actions. In Proc. CHI '12. 2025-2034.

35. DeMenthon D. and Davis, L. S. Model-based object pose in 25 lines of code. International Journal of Computer Vision, 15 (1995). 123–141.

36. Deyle, T., Palinko, S., Poole, E. S. and Starner, T. Hambone: A Bio-Acoustic Gesture Interface. In Proc. ISWC '07. 1-8.

37. Dietz, P. and Leigh, D. DiamondTouch: a multi-user touch technology. In Proc. UIST '01. 219-226.

38. Donnelly, L., Patten, D., White, P. and Finn, G. Virtual human dissector as a learning tool for studying cross-sectional anatomy. Med Teach., 31, 6 (2009). 553-555.

39. Erol, A., Bebis, G., Nicolescu, M., Boyle, R. D. and Twombly, X. Vision-based hand pose estimation: A review. Computer Vision and Image Understanding, 108 (2007). 52-73.

40. Faste, R. The Role of Aesthetics in Engineering. Japan Society of Mechanical Engineers (JSME) Journal, 28 (1995). 385.

41. Field, T. (2001) Touch. MIT Press, Cambridge, MA.

42. Forlizzi, J., Disalvo, C., Zimmerman, J. and Hurst, A. The SenseChair: The lounge chair as an intelligent assistive device for elders. In Proc. DUX '05. 1-13.

43. Foster, K. R. and Lukaski, H. C. Whole-body impedance - what does it measure? The American journal of clinical nutrition, 64 (3). 1996. 388S-396S.

44. Gallace, A. and Spence, C. (2010). The science of interpersonal touch: An overview, Neuroscience and Biobehavioral Reviews, 34. 246-259

45. Gallagher, S. (2005). How the body shapes the mind. Clarendon Press, Oxford.

46. Gavaghan, K. A., Peterhans, M., Oliveira-Santos, T. and Weber, S. A Portable Image Overlay Projection Device for Computer-Aided Open Liver Surgery. IEEE Transactions on Biomedical Engineering, 58, 6 (June 2011). 1855-1864.

47. Gemperle, F., Kasabach, C., Stivoric, J., Bauer, M. and Martin, R. Design for wearability. In Proc. ISWC '98. 116-122.

48. Grimes, D., Tan, D., Hudson, S.E., Shenoy, P. and Rao, R. Feasibility and pragmatics of classifying working memory load with an electroencephalograph. In Proc. CHI '08. 835-844.

49. Guiard, Y. Asymmetric Division of Labor in Human Skilled Bimanual Action: The Kinematic Chain as a Model. Jour. of Motor Behavior, 19, 4 (1987). 486-517.

50. Gustafson, S., Bierwirth, D. and Baudisch, P. Imaginary interfaces: spatial interaction with empty hands and without visual feedback. In Proc. UIST '10. 3–12.

51. Gustafson, S., Holz, C. and Baudisch, P. Imaginary phone: learning imaginary interfaces by transferring spatial memory from a familiar device. In Proc. UIST '11. 283-292.

52. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. and Witten, I. H. The WEKA Data Mining Software: An Update. SIGKDD Explorations, 11, 1 (2009). 10-18.

53. Harada, S., Landay, J. A., Malkin, J., Li, X. and Bilmes, J. A. The vocal joystick: evaluation of voice-based cursor control techniques. In Proc. Assets '06. 197-204.

54. Harker, F. R. and Maindonald, J. H. Ripening of Nectarine Fruit. Plant physiology, 106 (1994). 165-171.

55. Harrison, B., Fishkin, K., Gujar, A., Mochon, C. and Want, R. Squeeze Me, Hold Me, Tilt Me! An Exploration of Manipulative User Interfaces. In Proc. CHI ’98. 17-14.

56. Harrison, C. and Dey, A. K. Lean and zoom: proximity-aware user interface and content magnification. In Proc. CHI '08. 507-510.

57. Harrison, C. and Hudson, S. E. Abracadabra: Wireless, High-Precision, and Unpowered Finger Input for Very Small Mobile Devices. In Proc. UIST '09. 121-124.

58. Harrison, C. and Hudson, S. E. Minput: Enabling Interaction on Small Mobile Devices with High-Precision, Low-Cost, Multipoint Optical Tracking. In Proc. CHI '10. 1661-1664.

59. Harrison, C. and Hudson, S. E. Scratch Input: Creating Large, Inexpensive, Unpowered and Mobile finger Input Surfaces. In Proc. UIST '08. 205-208.

60. Harrison, C. and Hudson, S. E. Utilizing Shear as a Supplemental Two-Dimensional Input Channel for Rich Touchscreen Interaction. In Proc. CHI '12. 3149-3152.

61. Harrison, C., Benko, H. and Wilson, A.D. OmniTouch: wearable multitouch interaction everywhere. In Proc. UIST '11. 441-450.

62. Harrison, C., Lim, B. Y., Shick, A. and Hudson, S. E. Where to Locate Wearable Displays? Reaction Time Performance of Visual Alerts from Tip to Toe. In Proc. CHI '09. 941-944.

63. Harrison, C., Schwarz, J. and Hudson S. E. TapSense: Enhancing Finger Interaction on Touch Surfaces. In Proc. UIST '11. 627-636.

64. Harrison, C., Tan, D. and Morris, D. Skinput: appropriating the body as an input surface. In Proc. CHI '10. 453-462.

65. Heo, S. and Lee, G. Force gestures: augmented touch screen gestures using normal and tangential force. In Proc. UIST’11. 621-626.

66. Herot, C. and Weinzapfel, G. One-Point Touch Input of Vector Information from Computer Displays. In Proc. SIGGRAPH '78. 210-216.

67. Hertenstein, M. J. (2002). Touch: Its communicative functions in infancy. Human Development, 45. 70-94.

68. Hertenstein, M. J. and Keltner, D. (2006). Touch Communicates Distinct Emotions, Emotion, 6 (3). 528-533.

69. Hinckley, K. and Song, H. Sensor synaesthesia: touch in motion, and motion in touch. In Proc. CHI '11. 801-810.

70. Hinckley, K., Pausch, R., Goble, J. C. and Kassell, N. F. A survey of design issues in spatial input. In Proc. UIST '94. 213-222.

71. Hinkley, K. and Sinclair, M., Touch-sensing input devices. In Proc. CHI '99. 223-230.

72. Hirshfield, L. M., Solovey, E. T., Girouard, A., Kebinger, J., Jacob, R. J., Sassaroli, A. and Fantini, S. Brain measurement for usability testing and adaptive interfaces: an example of uncovering syntactic workload with functional near infrared spectroscopy. In Proc. CHI ’09. 2185-2194.

73. Hobye, M. and Löwgren, J. Touching a Stranger: Designing for Engaging Experience in Embodied Interaction. International Journal of Design, 5, 3 (2011). 31-48.

74. Holder, M. K. (1997). “Why are more people right-handed?” Sciam.com. Scientific American. Retrieved 2012-9-14.

75. Holman, D. and Vertegaal, R. Organic user interfaces: designing computers in any way, shape, or form. Comm. of the ACM, 51, 6 (2008). 48–55.

76. Holz, C. and Baudisch, P. The generalized perceived input point model and how to double touch accuracy by extracting fingerprints. In Proc. CHI '10. 581-590.

77. Hornecker, E. and Burr, J. Getting a Grip on Tangible Interaction: A Framework on Physical Space and Social Interaction, In Proc. CHI ‘06. 437-446.

78. Hudson, S. E., Harrison, C., Harrison, B. L. and LaMarca, A. Whack gestures: inexact and inattentive interaction with mobile devices. In Proc. TEI '10. 109-112.

79. Ishii, H., Wisneski, C., Orbanes, J., Chun, B. and Paradiso, J. PingPongPlus: design of an athletic-tangible interface for computer-supported cooperative play. In Proc. CHI ’99. 394-401.

80. Jones, S. and Yarbrough, A. E. (1985). A Naturalistic Study of the Meanings of Touch. Comm. Monographs, 52 (1). 19-56.

81. Kabbash, P., Buxton, W. and Sellen, A. Two-Handed Input in a Compound Task. In Proc. CHI ’94. 417-423.

82. Kabbash, P., MacKenzie, I. S. and Buxton, W. Human performance using computer input devices in the preferred and non-preferred hands. In Proc. CHI '93. 474-481.

83. Kane, S., Avrahami, D., Wobbrock, J., Harrison, B., Rea, A., Philipose, M. and LaMarca, A. Bonfire: A nomadic system for hybrid laptop-tabletop interaction. In Proc. UIST '09. 129-138.

84. Karitsuka, T. and Sato, K. A Wearable Mixed Reality with an On-Board Projector. In Proc. ISMAR '03. 321-322.

85. Kendon, A. (1988). How Gestures Can Become Like Words. In Crosscultural Perspectives in Nonverbal Communication. Toronto, C.J. Hogrefe. 131-141.

86. Knapp, M. L. and Hall, J. A. (2006). Nonverbal Communication in Human Interaction (7th ed.), Wadsworth Publishing.

87. Laakso, S. and Laakso, M. Design of a body-driven multiplayer game system. Comput. Entertain., 4, 4 (2006). 1544-3574.

88. Lahey, B., Girouard, A., Burleson, W. and Vertegaal, R. PaperPhone: understanding the use of bend gestures in mobile devices with flexible electronic paper displays. In Proc. CHI '11. 1303-1312.

89. Lai, K., Bo, L., Ren, X., and Fox, D. Sparse Distance Learning for Object Recognition Combining RGB and Depth Information. In Proc. ICRA '11.

90. Lakshmipathy, V., Schmandt, C. and Marmasse, N. TalkBack: a conversational answering machine. In Proc. UIST ’03. 41-50.

91. Lee, J. C. and Tan, D. S. Using a low-cost electroencephalograph for task classification in HCI research. In Proc. CHI ’06. 81-90.

92. Lee, S. K., Buxton, W. and Smith, K. C. A multi-touch three dimensional touch-sensitive tablet. In Proc. CHI '85. 21-25.

93. Lepinski, J. G., Grossman, T. and Fitzmaurice, G. The design and evaluation of multitouch marking menus. In Proc. CHI '10. 2233-2242.

94. Lewis, R. J. Literature review of touch-screen research from 1980 to 1992. IBM Technical Report, 54.694. Aug 20, 1993.

95. Li, F. C. Y., Dearman, D. and Truong, K. N. Leveraging proprioception to make mobile phones more accessible to users with visual impairments. In Proc. ASSETS '10. 187-194.

96. Li, F. C., Dearman, D. and Truong, K. N. Virtual shelves: interactions with orientation aware devices. In Proc. UIST '09. 125-128.

97. Li, Y., Hinckley, K., Guan, Z. and Landay, J. A. Experimental analysis of mode switching techniques in pen-based user interfaces. In Proc. CHI '05. 461-470.

98. Lim, B. Y., Shick, A., Harrison, C. and Hudson, S. E. Pediluma: Motivating Physical Activity Through Contextual Information and Social Influence. In Proc. TEI '11. 123-180.

99. Löwgren J. Towards an Articulation of Interaction Aesthetics. The New Rev. of Hypermedia and Multi. 15, 2 (2009). 129-146.

100. Lyons, K., Skeels, C., Starner, T., Snoeck, C. M., Wong, B. A. and Ashbrook, D. Augmenting conversations using dual-purpose speech. In Proc. UIST '04. 237-246.

101. Lyons, K., Starner, T., Plaisted, D., Fusia, J., Lyons, A., Drew, A. and Looney, E. W. Twiddler typing: one-handed chording text entry for mobile phones. In Proc. CHI '04. 671-678.

102. Mandryk, R. L. and Atkins, M. S. A Fuzzy Physiological Approach for Continuously Modeling Emotion During Interaction with Play Environments. Intl Journal of Human-Computer Studies, 6, 4 (2007). 329-347.

103. Mandryk, R. L., Inkpen, K. M. and Calvert, T. W. Using Psychophysiological Techniques to Measure User Experience with Entertainment Technologies. Behaviour and Information Technology, 25, 2 (March 2006). 141-58.

104. Mann, S. Smart Clothing: The Wearable Computer and WearCam. *Personal Technologies*, 1, 1 (1997). 21-27.

105. Manning, E. (2006). Politics of Touch: Sense, Movement, Sovereignty. Univ. Of Minnesota Press.

106. Marshall, J., Rowland, D., Egglestone, S. R., Benford, S., Walker, B. and McAuley, D. Breath control of amusement rides. In Proc. CHI '11. 73-82.

107. Mascaro, S. A. and Asada, H. H. Measurement of finger posture and three-axis fingertip touch force using fingernail sensors. In IEEE Trans. on Robotics and Automation, 2004.

108. Matheson, G. O., Maffey-Ward, L., Mooney, M., Ladly, K., Fung, T. and Zhang, Y. T. Vibromyography as a quantitative measure of muscle force production. *Scand J Rehabil Med*. 29, 1 (Mar. 1997). 29-35.

109. Matsushita, N. and Rekimoto, J., HoloWall: designing a finger, hand, body and object sensitive wall. In Proc. UIST'97. 209-210.

110. McFarland, D. J., Sarnacki, W. A. and Wolpaw, J. R. Brain–computer interface (BCI) operation: optimizing information transfer rates. Biological Psychology, 63, 3 (Jul 2003). 237-51.

111. McFarlane, D. and Wilder, S. Interactive dirt: Increasing mobile work performance with a wearable projector-camera system. In Proc. UbiComp ’09. 205-214.

112. Measurement Specialties, Inc. MiniSense100. Retrieved December 20, 2011. http://meas-spec.com/product/t_product.aspx?id=2474

113. MicroVision, Inc. http://www.microvision.com

114. Milgram, P. and Kishino, F. A Taxonomy of Mixed Reality Visual Displays. IEICE Transactions on Information Systems, vol. E77-D, no. 12 (1994).

115. Mine, M. R., Brooks, F. P. and Sequin, C. H. Moving objects in space: exploiting proprioception in virtual-environment interaction. In Proc. SIGGRAPH '97. 19-26.

116. Mistry, P., Maes, P. and Chang, L. WUW - wear Ur world: a wearable gestural interface. In CHI ‘09 EA. 4111-4116.

117. Moeslund, T. B., Hilton, A. and Krüger, V. A survey of advances in vision-based human motion capture and analysis. Comput. Vis. Image Underst. 104, 2 (November 2006).

118. Montagu, A. (1986). Touching: the human significance of the skin (3rd ed.). Harper and Row.

119. Moore, M. and Dua, U. A galvanic skin response interface for people with severe motor disabilities. In Proc. ACM SIGACCESS Accessibility and Comp. ’04. 48-54.

120. Mulder, A. (1996). Hand Gestures for HCI, School of Kinesiology, Simon Fraser University, Technical Report 96-1.

121. Ni, T. and Baudisch, P. Disappearing mobile devices. In Proc. UIST 2009. 101-110.

122. Noë, A. (2005). Action in Perception. MIT Press.

123. NTT IT Corp. 2010. TenoriPop: http://tenoripop.com

124. Oulasvirta, A., Tamminen, S., Roto, V. and Kuorelahti, J. Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile HCI. In Proc. CHI ’05. 919-928.

125. Paradiso, J. A., Hsiao, K. and Benbasat, A. Tangible music inter-faces using passive magnetic tags. In Proc. NIME '01. 1-4.

126. Paradiso, J. and Hsiao, K., Swept-frequency, magnetically-coupled resonant tags for realtime, continuous, multiparameter control. In CHI EA '99. 212-213.

127. Paradiso, J., Hsiao, K., Strickon, J., Lifton, J. and Adler, A. Sensor Systems for Interactive Surfaces. IBM Sys. J., 39(3/4), Oct ‘00, 892-914.

128. Paradiso, J.A., Leo, C. K., Checka, N. and Hsiao, K. Passive acoustic knock tracking for interactive windows. In CHI ‘02 EA. 732-733.

129. Patten, D. What lies beneath: the use of three-dimensional projection in living anatomy teaching. The Clinical Teacher, 4 (2007). 10-14.

130. Philipp, H. Charge transfer sensing. Sens. Review, 19, 1999. 96-105.

131. Pierce, J. S. and Mahaney, H. E. (2004). Opportunistic Annexing for Handheld Devices: Opportunities and Challenges. Human-Computer Interface Consortium 2004.

132. Pinhanez, C.S. The Everywhere Displays projector: A device to create ubiquitous graphical interfaces. In Proc. UBICOMP ‘01. 315–331.

133. Porac, C. and Coren, S. Lateral Preferences and Human Behavior. New York: Springer-Verlag, 1981.

134. Post, E. R. and Orth, M. Smart Fabric, or “Wearable Clothing.” In Proc. ISWC ’97. 167-168.

135. Poupyrev, I. and Maruyama, S. Tactile interfaces for small touch screens. In Proc. UIST '03. 217-220.

136. PrimeSense Ltd. http://www.primesense.com

137. Ramos, G. and Balakrishnan, R. Pressure marks. In Proc. CHI '07. 1375-1384.

138. Ramos, G. and Balakrishnan, R. Zliding: fluid zooming and sliding for high precision parameter manipulation. In Proc. UIST '05. 143-152.

139. Ramos, G., Boulos, M. and Balakrishnan, R. Pressure widgets. In Proc. CHI '04. 487-494.

140. Raskar, R., Beardsley, P., van Baar, J., Wang, Y., Dietz, P., Lee, J., Leigh, D. and Willwacher, T. RFIG lamps: interacting with a self-describing world via photosensing wireless tags and projectors. In Proc. SIGGRAPH ’04. 406–415.

141. Rekimoto, J. SmartSkin: An Infrastructure for Freehand Manipulation on Interactive Surfaces. In Proc. CHI '02. 113-120.

142. Rogers, S., Williamson, J., Stewart, C. and Murray-Smith, R. AnglePose: robust, precise capacitive touch tracking via 3d orientation estimation. In Proc. CHI '11. 2575-2584.

143. Rosenberg, I. and Perlin, K. The UnMousePad: an interpolating multi-touch force-sensing input pad. In Proc. SIGGRAPH'09, Article 65. 65:1-65:9.

144. Rosenberg, R. The biofeedback Pointer: EMG Control of a Two Dimensional Pointer. In Proc. ISWC ’98. 4-7.

145. Ross, J., Irani, L., Silberman, M., Zaldivar, A. and Tomlinson, B. Who are the crowdworkers?: Shifting demographics in mechanical turk. In CHI EA '10. 2863-2872.

146. Roudaut, A, Pohl, H. and Baudisch, P. Touch input on curved surfaces. In Proc. CHI ’11. 1011-1020.

147. Roudaut, A., Lecolinet, E. and Guiard, Y. MicroRolls: expanding touch-screen input vocabulary by distinguishing rolls vs. slides of the thumb. In Proc. CHI '09. 927-936.

148. Sakata, N., Konishi, T. and Nishida, S. Mobile Interfaces Using Body Worn Projector and Camera. In Proc. VMR '09. 106-113.

149. Saponas, T. S., Harrison, C. and Benko, H. PocketTouch: Through-Pocket Capacitive Touch Input. In Proc. UIST '11. 303-308.

150. Saponas, T. S., Tan, D. S., Morris, D. and Balakrishnan, R. Demonstrating the feasibility of using forearm electromyography for muscle-computer interfaces. In Proc. CHI ’08. 515-524.

151. Saponas, T. S., Tan, D. S., Morris, D., Balakrishnan, R., Turner, J. and Landay, J. A. Enabling always-available input with muscle-computer interfaces. In Proc. UIST '09. 167-176.

152. Saponas, T. S., Tan, D. S., Morris, D., Balakrishnan, R., Turner, J., and Landay, J. A. Making muscle-computer interfaces more practical. In Proc. CHI 2010. 851-854.

153. Sato, M. Poupyrev, I, and Harrison, C. Touché: Enhancing Touch Interaction on Humans, Screens, Liquids, and Everyday Objects. In Proc. CHI ’12. 483-492.

154. Schalk, G., Miller, K. J., Anderson, N. R., Wilson, J. A., Smyth, M. D., Ojemann, J. G., Moran, D. W., Wolpaw, J. R. and Leuthardt, E. C. Two-dimensional movement control using electrocorticographic signals in humans. Journal of Neural Engineering, 5, 1 (2008). 75-84.

155. Schwarz, J., Harrison, C., Mankoff, J. and Hudson, S. E. Cord Input: An Intuitive, High-Accuracy, Multi-Degree-of-Freedom Input Method for Mobile Devices. In Proc. CHI '10. 1657-1660.

156. Schwesig, C., Poupyrev, I. and Mori, E. Gummi: a bendable computer. In Proc. CHI '04. 263-270.

157. Sears, A. Improving touchscreen keyboards: Design issues and a comparison with other devices. IEEE Computer, 3 (1991). 253–269.

158. Shumway-Cook A. and Woollacott M. (2011) Motor Control: Motor Control: Translating Research into Clinical Practice. Lippincott Williams and Wilkins, Baltimore, MD.

159. Siddiqui, M. and Medioni, G. Robust real-time upper body limb detection and tracking. In Proc. VSSN '06. 53-60.

160. Siek, K. A., Rogers, Y. and Connelly, K. H. Fat Finger Worries: How Older and Younger Users Physically Interact with PDAs. In Proc. INTERACT ’05. 267–280.

161. Skulpone, S., Dittman, K. Adjustable proximity sensor. US Patent 3,743,853, 1973.

162. Smith, J., White, T., Dodge, C., Paradiso, J., Gershenfeld, N. and Allport, D. Electric Field Sensing For Graphical Interfaces. IEEE Comput. Graph. Appl. 18, 3 (1998). 54-60.

163. Smith, J.R. Field mice: Extracting hand geometry from electric field measurements. IBM Systems Journal, 35 (1996). 587-608.

164. Starner T., Pentland A. and Weaver, J. Real-Time American Sign Language Recognition Using Desk and Wearable Computer Based Video, IEEE Trans. on Pattern Analysis and Machine Intelligence, 20, 12 (1998). 1371-1375.

165. Starner, T. The Role of Speech Input in Wearable Computing. IEEE Pervasive Computing, 1, 3 (July 2002). 89-93.

166. Starner, T., Auxier, J., Ashbrook, D. and Gandy, M. The Gesture Pendant: A self-illuminating, wearable, infrared computer vision system for home automation control and medical monitoring. In Proc. ISWC ’00. 87–94.

167. Sturman, D.J. and Zeltzer, D. A Survey of Glove-based Input. IEEE Comp. Graph. and Appl., 14, 1 (1994). 30-39

168. Sugrue, C. 2007. "Delicate Boundaries" (art installation).

169. Shiratori, T., Park, H. S., Sigal, L., Sheikh, Y. and Hodgins, J. K. Motion capture from body-mounted cameras. In Proc. SIGGRAPH 2011, Article 31, 10 pages.

170. Tan, D., Morris, D. and Saponas, T. S. Interfaces on the go. ACM XRDS, 16, 4 (2010). 30-34.

171. Thayer, S. (1982). Social Touching. In Tactile perception: a sourcebook. Cambridge University Press.

172. Thomas, B., Grimmer, K., Zucco, J. and Milanese, S. Where Does the Mouse Go? An Investigation into the Placement of a Body-Attached TouchPad Mouse for Wearable Computers. Personal Ubiquitous Computing, 6, 2 (January 2002). 97-112.

173. Tomasi, C., Rafii, A. and Torunoglu, I. Full-size projection keyboard for handheld devices. Comm. of the ACM, 46, 7 (2003). 70–75.

174. Tomita, M. An Exploratory Study of Touch Zones in College Students on Two Campuses, Californian Journal of Health Promotion, 6 (1), 2008. 1-22.

175. Valera, F., Thompson, E. and Rosch, E. (1991). The Embodied Mind: Congitive Science and Human Experience. MIT Press.

176. Valii, C. (ed). (2006). The Gallaudet Dictionary of American Sign Language; 1st edition. Gallaudet University Press.

177. Wachs, J. P. Kölsch, M., Stern, H. and Edan, Y. Vision-based hand-gesture applications. Commun. ACM, 54, 2 (February 2011). 60-71.

178. Wang, F. and Ren, X. Empirical evaluation for finger input properties in multi-touch interaction. In Proc. CHI '09. 1063–1072.

179. Wang, J., Zhai, S. and Canny, J. Camera phone based motion sensing: interaction techniques, applications and performance study. In Proc. UIST '06. 101-110.

180. Warren, J. (2003). Unencumbered Full Body Interaction in Video Games. MFA Design and Technology Thesis, Parsons School of Design.

181. Webster, J. G. (ed). (2010). Medical instrumentation: application and design. Wiley.

182. Wigdor, D. and Balakrishnan, R. TiltText: using tilt for text input to mobile phones. In Proc. UIST '03. 81-90.

183. Williams, A., Farnham, S., and Counts, S. Exploring wearable ambient displays for social awareness. In CHI '06 Extended Abstracts. 1529-1534.

184. Willis, K. D. D., Poupyrev, I. and Shiratori, T. Motionbeam: a metaphor for character interaction with handheld projectors. In Proc. CHI '11. 1031-1040.

185. Willis, K. D. D., Poupyrev, I., Hudson, S. E. and Mahler, M. SideBySide: ad-hoc multi-user interaction with handheld projectors. In Proc. UIST '11. 431-440.

186. Wilson, A. and Benko, H. Combining multiple depth cameras and projectors for interactions on, above and between surfaces. In Proc. UIST '10. 273–282.

187. Wilson, A. Robust computer vision-based detection of pinching for one and two-handed gesture input. In Proc. UIST ’06. 255-258.

188. Wilson, A. PlayAnywhere: a compact interactive tabletop projection-vision system. In Proc. UIST '05. 83–92.

189. Wilson, A. Using a depth camera as a touch sensor. In Proc. ITS '10. 69–72.

190. Wilson, F. (1998). The Hand: How its use shapes the brain, language, and human culture. Pantheon.

191. Wimmer, R. and Baudisch, P., Modular and Deformable Touch-Sensitive Surfaces Based on Time Domain Reflectometry. In Proc. UIST '11, 517-526.

192. Witten, I. H. and Frank, E. (2005). Data Mining: Practical machine learning tools and techniques, 2nd Edition, Morgan Kaufmann, San Francisco.

193. Wolfe, J. M., Kluender, K. R., Levi, D. M., Bartoshuk, L. M, Herz, R. S., Klatzky, R. L. and Lederman, S. J. (2006). Sensation and Perception. Sinauer Associates, Sunderland, MA.

194. Yamamoto, G. and Sato, K. PALMbit: A PALM Interface with Projector-Camera System. In Adj. Proc. UbiComp ’07. 276-279.

195. Yamamoto, G. and Sato, K. PALMbit:A Body Interface utilizing Light Projection onto Palms. The Journal of The Institute of Image Information and Television Engineers, 61, 6 (2007). 797-804.

196. Yamamoto, G., Xu, H., Ikeda, K. and Sato, K. PALMbit-Silhouette: A User Interface by Superimposing Palm-Silhouette to Access Wall Displays. In Proc. HCI International '09. LNCS 5611, 2009. 281-290.

197. Zhong, L., El-Daye, D., Kaufman, B., Tobaoda, N., Mohamed, T. and Liebschner, M. OsteoConduct: wireless body-area communication based on bone conduction. In Proc. ICST ’07. 1-8.

198. Zhou, F., Duh, H. B. and Billinghurst, M. Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR. In Proc. ISMAR '08. 193-202.

199. Zimmerman, T. G. Personal area networks: near-field intrabody communication, IBM Systems Journal, 35, 3-4 (1996). 609-617.

200. Zimmerman, T.G., Smith, J.R., Paradiso, J.A., Allport, D. and Gershenfeld, N., Applying electric field sensing to human-computer interfaces. In Proc. CHI '95, 280-287.