Models and Mechanisms for Tangible User Interfaces

Brygg Anders Ullmer

Bachelor of Science, University of Illinois,
Urbana-Champaign, Illinois, January 1995

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Media Arts and Sciences at the
Massachusetts Institute of Technology

June 1997

Author
Brygg Anders Ullmer
Program in Media Arts and Sciences
May 9, 1997

Certified by
Hiroshi Ishii
Associate Professor of Media Arts and Sciences
Thesis Supervisor

Accepted by
Stephen A. Benton
Chair, Departmental Committee for Graduate Students
Program in Media Arts and Sciences

Models and Mechanisms for Tangible User Interfaces

Brygg Anders Ullmer

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning, on May 9, 1997
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Media Arts and Sciences at the
Massachusetts Institute of Technology

Abstract

Current human-computer interface design is dominated by the graphical user interface approach, where users interact with graphical abstractions of virtual interface devices through a few general-purpose input “peripherals.” The thesis develops models and mechanisms for “tangible user interfaces” – user interfaces which use physical objects, instruments, surfaces, and spaces as physical interfaces to digital information. Prototype applications on three platforms – the metaDESK, transBOARD, and ambientROOM – are introduced as examples of this approach. These instances are used to generalize the “GUI widgetry,” “optical,” and “containers and conduits” interface metaphors. The thesis also develops engineering mechanisms called proxy-distributed or “proxdist” computation, which provide a layered approach for integrating physical objects with diverse sensing, display, communication, and computation capabilities into coherent interface implementations. The combined research provides a vehicle for moving beyond the keyboard, monitor, and pointer of current computer interfaces towards use of the physical world itself as a kind of computationally-augmented interface.

Thesis Supervisor:
Hiroshi Ishii
Associate Professor of Media Arts and Sciences

Thesis Committee

Thesis Supervisor
Hiroshi Ishii
Associate Professor of Media Arts and Sciences
Massachusetts Institute of Technology

Reader
Mitchel Resnick
Associate Professor of Media Arts and Sciences
Fukutake Career Development Professor of Research in Education
Massachusetts Institute of Technology

Reader
Terry Winograd
Professor of Computer Science
Stanford University

Acknowledgements

I am grateful for the guidance and support of many people who have made this work possible.

Special thanks to:

Hiroshi Ishii, for his enormous enthusiasm and vision, for innumerable hours of discussion and guidance that gave content and shape to the thesis, and for the opportunity to participate in the creation of the Tangible Media Group.

Thesis readers Mitchel Resnick and Terry Winograd, for their guidance and ideas in helping shape this document.

AT&T and Mitsubishi, my fellowship sponsors during the 1995-96 and 1996-97 academic years, and especially to fellowship mentors Will Hill and Joe Marks.

All of the people at Interval Research who made the summers of 1993-95 a wonderful first "graduate school." Thanks especially to:

Terry Winograd, for providing guidance and inspiration throughout the duration;
David Liddle, for allowing my continuing exploration of proxdist computation ideas;
Debby Hindus, for helping me find my way to and through the Media Lab;
Paul Freiberger, for believing in me and re-opening the doors for second and third rounds; and
Bill Verplank, for his continuing guidance and the idea of the rotation-constraint instrument.

Ron MacNeil, my initial research advisor, for opening the doors to the Media Lab and introducing the concept of "design" into my intellectual vocabulary.

UROPs Chris Fuchs, Chris Lopes, Tom Rikert, Dylan Glas, Minpont Chien, and Craig Wisneski for their work on the transBOARD, metaDESK, and ambientROOM prototypes. Special thanks to ChrisL and Min for their hard work on Tangible Geospace.

Fellow Tangible Media graduate students Matt Gorbet, Scott Brave, and Andrew Dahley, for their camaraderie and collaboration, as well as fellow D+I students, for breathing life and energy into the space and having patience with me during my times of large writings.

Suguru Ishizaki, for much-valued discussions and guidance during my first months at the Lab.

David Small, Lisa Strausfeld, Earl Rennison, and Suguru Ishizaki, for providing a living bridge into the work and passion that was the VLW.

Durrell Bishop, whose work first inspired my interest in physical objects as digital interface.

Thad Starner, for much-valued computer vision assistance and many stimulating discussions along the way.

Jennifer Glos, for her collaboration with POEMs.

David Morgan, for his interest and input along the way.

Betty Lou McClanahan, Linda Peterson, Santina Tonelli, and Laurie Ward, for their guidance and support on innumerable details large and small.

Hannes Vilhjálmsson, for his companionship, support, and long nights of rich discussion.

My family and especially parents, for their unwavering love, support, and guidance in pulling everything into perspective

Contents

Note: minor corrections have been made to this document between May 27, 1997, and its May 9, 1997 submission.

1 Introduction

The graphical user interface (GUI) has proven both a successful and durable model for human-computer interaction which has dominated the last decade of interface design. At the same time, the GUI approach falls short in many respects, particularly in embracing the rich interface modalities between people and the physical environments they inhabit. Systems exploring augmented reality and ubiquitous computing have begun to address this challenge, moving the locus of interface away from the desktop and into the physical environment. However, these efforts have often taken the form of exporting the GUI paradigm to more world-situated devices, falling short of much of the richness of physical-space interaction they seek to augment.

In this thesis, I present research developing the notion of “tangible user interfaces” or TUIs: user interfaces which use physical objects, instruments, surfaces, and spaces as physical interfaces to digital information. TUIs emphasize the direct physical manipulation of physical-world “tangibles,” producing interfaces which capitalize on multiple human senses and kinesthetic engagement of the human body in space. TUIs engage users both with graspable physical objects amplified by graphical, audible, and haptic augmentations, as well as with ambient displays which mediate information display at the periphery of users’ attention.

The thesis research is grounded upon three interface platforms. First, the metaDESK is a graphically intensive horizontal surface driven by interaction with graspable physical objects and instruments. Secondly, the transBOARD is an instrumented physical whiteboard using physical objects as “containers” and “conduits” for whiteboard activity. Lastly, the ambientROOM is an augmented room environment using ambient light, sound, airflow, and water movement as ambient information interfaces.

Drawing from prototype applications on these platforms, the thesis generalizes several interface models by which TUI design may be approached. In particular, the thesis discusses the interface roles played by physical objects, instruments, surfaces, and spaces as TUI “tangibles.” This is used to motivate the “GUI widgetry,” “optical,” and “containers and conduits” metaphors, specific approaches for the use of physical-world tangibles as digital-world interfaces.

In addition, the thesis considers engineering mechanisms by which tangible user interfaces can be designed and built. Principal among these is proxy-distributed or “proxdist” computation, which establishes digital-world proxies operating on behalf of physical-world objects. With proxdist computation, even passive physical objects can be considered to have the capacity for sensing, display, communication, and computation. Proxdist mechanisms address both the process by which complex TUIs can be built, as well as providing conceptual “tools for thought” by changing assumptions about how computation can be manifested in the physical environment.

Finally, it is important to note from the outset that this thesis content is the product of a close collaboration with Professor Hiroshi Ishii as part of the formation of the MIT Media Lab Tangible Media Group. This effort directly gave rise to the platforms, prototypes, and many of the ideas in this thesis, and will be evidenced by frequent use of the pronoun “we” and references to “Ishii and I.”

1.1 Motivation

The art and science of interface design is perhaps one of the oldest human endeavors, dating back to the first emergence of tool use by primitive man. The study of interface design has grown to be an issue of increasing import as people engineer ever more complex devices, environments, and bodies of abstract information, perhaps peaking with the advent of the computer during the last half-century.

The computer, seen as a kind of digital information processor, connects with the challenge of interface design in several major respects. The first of these approaches the challenge of interface design to the computer’s information processing capacities. For the most part, this challenge of digital information interface has been approached with highly visual interfaces serviced by a limited set of general-purpose physical input and output “peripherals” – most commonly, a keyboard, monitor, and pointer hosting a graphical user interface (GUI). This approach has formed the basis for modern human-computer interface (HCI) design, and has become so pervasive that the distinction between the computer as information processor and computer as terminal-based user interface in popular usage has become substantially blurred.

The second interlinkage between computers and interface design relates to the embedding of computers within a broad array of physical-world devices whose primary purpose is not that of abstract information interface. Incorporated within devices ranging from automobiles, water faucets, and telephones, to elevators, lamps, and washing machines, computers here serve a large and increasing number of generally invisible roles. Throughout this process, the physical form and operational metaphor of these devices has remained largely constant, for their functions are rooted in specific physical-world tasks, and their physical interface affordances have evolved for generations to address these concrete demands.

While computers are used in these two capacities in ever-greater numbers, it is interesting to consider whether there is some new territory of computationally-augmented interface design to be found. Instead of the general-purpose GUI terminals of mainstream HCI, it is compelling to look towards the physical legibility and greater sensory engagement of more specialized physical forms which may be applied to the domain of digital information interface. In particular, two new interface domains hold strong promise. First, numerous pre-existing physical objects, instruments, surfaces, and spaces hold the potential to exhibit substantially new functionality in the presence of computational augmentation, while leveraging existing physical affordances and metaphors. Secondly, there remain to be invented an even more diverse range of novel "tangibles" which first become meaningful possibilities in the presence of computational augmentation, but support physical interaction in a fashion distinct from the intangible GUIs of current HCI.

These latter cases form the motivation and target domain for the thesis research. While the GUI terminal theme of modern HCI design is neither backwards nor ill-conceived, it represents only a fragment of a much larger possible design space, the realm of the human body interacting within physical space. The physically-decoupled manipulation of graphical widgetry is successful in escaping certain physical-world limitations and constraints. However, this gain is purchased at the cost of sacrificing many powerful affordances of the physical world which render it ripe for human habitation. In short, rather than structure user interface around the screen, pointer, and keyboard underlying the modern GUI, the thesis research attempts to transform the physical world itself into a form of partially mediated human interface.

1.2 Thesis scope and overview

This thesis attempts to develop the notion of the tangible user interface in two complementary and intimately interdependent respects. First, the thesis develops interface models for TUIs – concrete design solutions by which the physical "tangibles" of TUIs may be used to interact with digital information. Secondly, the thesis develops engineering mechanisms by which TUIs can be functionally built – concrete engineering approaches through which mixed collections of passive and active physical objects may be seamlessly combined into functional user interfaces.

While it is possible to discuss TUI design models without simultaneously considering engineering mechanisms, perhaps the largest contribution aspired to by the thesis is intimately dependent on the support of new engineering approaches. In particular, the very phrase "human-computer interaction," interpreted perhaps hyper-literally, indicates human interaction with a device integrally composed (in the current technological regime) of a silicon-based microprocessor. While a reasonable concept in principle, it is also a limiting one. The microprocessor itself is rarely the user's focus of interest. By postulating the central role of the microprocessor in each would-be user interface, the underlying issues of physical-world user interface are easily obscured.

In contrast, the user interface designs developed in this thesis begin with the hypothesis that *any* physical object, instrument, surface, or even partially ethereal space can be considered to possess diverse capabilities for sensing, display, computation, and communication, provided that certain technological capabilities are present in the environment. This is a fairly ambitious hypothesis that demands substantiation. The engineering mechanisms of this thesis are intended to demonstrate both the engineering feasibility of this hypothesis, as well the conceptual role the hypothesis is intended to serve.

In the following pages, Chapter 2 will introduce major areas and instances of research related to the thesis work. Chapter 3 will introduce two early thesis prototypes, Tangible Infoscapes and POEMs, as well as the primary metaDESK, transBOARD, and ambientROOM platforms on which the thesis is grounded. Chapter 4 will generalize TUI interface models from these prototypes, beginning with consideration of TUI "tangibles" and continuing with introduction of the "GUI widgetry," "optical," and "containers and conduits" interface metaphors. Chapter 5 will introduce the proxy-distributed or "proxdist" computation mechanisms on which the thesis is technically grounded. Finally, Chapter 6 will return to detailed discussion of the Tangible Geospace prototype application of the metaDESK, grounded in the TUI models and mechanisms of earlier chapters.

2 Related Work

The research of this thesis draws from a range of work in user interface design and also underlying computational protocols and architectures. This chapter will briefly survey this territory. First, several broad models for user interface design will be discussed, including augmented reality, ubiquitous computing, and "things that think." Next, a number of particular instances of user interface coupling the physical world with digital information will be introduced. Finally, several threads of research related to computational mechanisms for tangible user interfaces will be surveyed.

2.1 Models

2.1.1 Augmented Reality

Perhaps the broadest area of related work lies in the area of "augmented reality." Augmented reality is broadly concerned with computationally augmenting human interaction with the physical world, and exemplified by works such as KARMA [Feiner 1993] and NaviCam [Noma 1996]. These works also illustrate augmented reality's common theme of visual augmentation of the physical world through head-mounted or hand-held graphical displays. Here, user interface is generally realized as virtual manipulation of graphical abstractions of realspace, rather than physical manipulation of the physical environment itself. Some work such as Wellner's pioneering DigitalDesk [Wellner 1993] has begun to push this boundary, augmenting direct physical manipulation of real-world objects.

2.1.2 Ubiquitous Computing

Another important stream of related research is that of ubiquitous computing. Ubiquitous computing, introduced with the pioneering Xerox PARC research of Weiser et al. [Weiser 1991], explores ways in which many computational devices can be distributed and "transparently" integrated within the physical environment. With an array of devices originally including the Tab, Pad, and Board, the PARC research demonstrated many differently-afforded devices filling different niche interface roles in the physical environment. At the same time, while demonstrating devices of many form-factors and functions, these early ubiquitous computing examples largely adopted the GUI approach of button/stylus interaction with virtual widgetry on a graphical surface.

Later PARC work includes the notion of "calm technology," exemplified by the Live Wire prototype [Weiser 1995], with a goal of user interfaces better integrating into the physical periphery of user activity. Promising related research outside of Xerox PARC includes the Reactive Environments research of Cooperstock [Cooperstock 1995], which develops approaches for seamless computational coordination and augmentation of physical-world user interaction with a range of distributed devices.

2.1.3 TTT

"Things that Think" or "TTT" is a research theme crystallized by consortium research at the MIT Media Lab, within which this thesis research has been conducted. TTT broadly explores the notion of physical things as possessors of computational capabilities and behaviors. One stream of such work (predating the formation of TTT) includes the Behavior Construction Kit, Dr. Legohead, and other research by Resnick, Borovoy, et al. [Resnick 1993, Borovoy 1996], which explores computational capabilities embedded within and attributable to physical objects in an education and learning context. Other related research includes work relating to "information appliances" in the PIA group, work relating to computationally-augmented toys by Glos, Umaschi, and others [Glos 1997, Umaschi 1997], and other threads of tangible user interface research in the Tangible Media Group within which this thesis has been executed.

2.1.4 Affordances, Metaphors, and Legibility

The thesis body makes numerous references to "physical affordances," "interface metaphors," and "legibility of interface." In the context of user interface, the notion of "physical affordances" is often attributed to Norman's text "The Psychology of Everyday Things." Here, Norman defines "affordances" as

> …the perceived and actual properties of the thing, primarily those fundamental properties that determine just how the thing could possibly be used…. Affordances provide strong clues to the operations of things. Plates are for pushing. Knobs are for turning … When affordances are taken advantage of, the user knows what to do just by looking…. [Norman 1988]

In the context of tangible user interfaces, we often refer to physical affordances relating to the physical appearance, mechanical constraints, etc. of various interface tangibles.

The concept of metaphor and its roles in everyday interaction as well as human-computer interface draws from discussion in [Lakoff 1980], [Erickson 1990], among other places. Our usage of the term "legibility" draws from [Lynch 1960], where Lynch defines legibility as "the ease with which [a thing's] parts can be recognized and can be organized into a coherent pattern."

Where Lynch's discussion of legibility is grounded in the visual comprehensibility of architectural space, our discussion focuses on the comprehensibility of physically-instantiated user interface.

2.2 Instances

A number of particular research instances have explored aspects of physical-world interaction as computational interface. In the case of the metaDESK, the Bricks work of Fitzmaurice, Ishii, and Buxton [Fitzmaurice 1995] is most directly inspirational. The Bricks research involves placing one or more bricks – abstract physical blocks tracked with six degrees of freedom – onto some screen-based virtual object, b-spline control point, etc. Bricks can then be used to physically rotate, translate, or (using multiple bricks in combination) scale and deform the "attached" virtual entities by manipulating the proxying brick devices. In more broadly-based doctoral research, Fitzmaurice generalized Bricks and other research instances and studies into a broader framework for "graspable user interface." [Fitzmaurice 1996]

Several other efforts relate to desk-style user interfaces. Wellner's DigitalDesk [Wellner 1993] supports augmented interaction with physical paper documents on a physical desktop, identifying and augmenting these with overhead cameras and projectors. The Responsive Workbench [Krueger 1994] and Immersadesk [Czernuszenko 1996] provide another interesting approach, supporting interaction with a large graphical surface viewable in stereo 3D using LCD shutterglass eyepieces. However, both platforms limit interaction to the virtual graphical space, with no support for physical-world interaction beyond use of a spatial position tracker.

Several research systems make use of devices relating to the metaDESK's lens-style optical devices. The PDDM device [Noma 1996] is an arm-mounted flat-panel display augmented with haptic force-feedback used in the proximity of a wall projection display. The small display of the PDDM is used for "grasping" virtual objects with force feedback in a VR scene. Another system, the NaviCam [Rekimoto 1995], includes a hand-held lens-like device for navigating physical spaces such as a bookshelf. Fitzmaurice's Spatially Aware Palmtop Display [Fitzmaurice 1993] provides an earlier example of a hand-held display used to spatially navigate more abstract information. The Magic Lens and Toolglass research [Bier 1993] provides a compelling example of virtual widgetry which graphically transform windowed regions of a GUI desktop to provide alternative representations of screen-based content.

Several other works focus on manipulation of physical objects as interface. The passive interface props work of Hinckley et al. [Hinckley 1994] uses physical props (e.g. head viewing prop and cutting-plane selection prop) as tools for manipulating 3D models within a GUI surgical interface. Bishop's marble answering machine [Crampton Smith 1995] makes compelling use of marbles as a physical embodiment of voice messages. Finally, Stifelman's paper-based audio notebook [Stifelman 1996] provides an excellent example of employing an augmented physical object (in her case, a paper notebook) in an unusually natural, legible, and useful fashion.

With respect to the transBOARD prototype, numerous systems have explored computationally-augmented physical whiteboards, including [Stafford-Fraser 1996], [Stasior 1993], and others. Most directly inspirational, though, is the ClearBoard work of Ishii and Kobayashi at NTT [Ishii 1994]. More than a model for seamless physical-world collaboration through the vehicle of a shared drawing surface, the ClearBoard developed a vision of interactive surfaces in architectural space that has been directly inspirational to both the transBOARD, metaDESK, ambientROOM, and broader conceptions of tangible user interfaces.

Finally, relating to the notion of ambient media explored in the ambientROOM, two particular works provided direct inspiration to our research. First, the "Fields and Thresholds" work of Raby and Dunne [Dunne 1994] presented the example of two "telematically" coupled steel benches which used the transmission of bench-grounded body warmth as an ambient vehicle for mediating remote audio communications. Secondly, the Live Wire work of Jeremijenko [Weiser 1995] used a twirling length of plastic cord to ambiently communicate the activity of local-area network traffic. These subtle uses of non-graphical media for display at the periphery of human awareness proved powerful inspirations for first ambientROOM iterations.

2.3 Mechanisms

The creation of mechanisms for building tangible user interfaces is supported by numerous engineering and scientific disciplines. At some level, these include areas such as digital signal processing, control systems, sensor fusion, sensor scheduling and management, and robotics, even extending onwards to applied arts like theatre design and Disney-style environmental design. However, the mechanisms focus of the thesis lies at the boundary between distributed systems and user interface design, approaching the task of cleanly integrating the distributed sensor, display, communication, and computation resources composing TUIs.

Here, the repertoires of communications protocols and layered abstractions developed for managing the complexities of computer networking and internetworking hold special relevance. Protocols and mechanisms such as DNS (domain name service), ARP (address resolution protocol), and BOOTP (bootstrap protocol) [Comer 1991] are all relevant to identity establishment and name resolution for the distributed physical objects of TUIs. Similarly, the layered OSI reference model [Tanenbaum 1996] has special relevance to our approach for layered abstraction of complexity in TUIs. These and other examples will be returned to in Chapter 5.

Other domains of relevance include the related area of distributed operating systems, as well as mechanisms for disconnected operation faced by systems supporting mobile communication. At a higher level of abstraction, distributed simulation environments such as SIMNET [Calvin 1993] and SPLINE [Waters 1996] face related issues of maintaining coherent state for distributed processes spanning nodes of varying computational capacity and linked by varying-quality communications paths. Lastly, notions of multi-agent systems, particularly those distributed across environments such as the Internet, develop concepts of distributed agency relevant our notions of distributed proxies operating on behalf of potentially passive physical objects.

3 First Prototypes

The TUI models and mechanisms of the thesis were grounded in the implementation of five tangible user interface prototypes. These include Tangible Infoscapes and Tangible Geospace, both based on the metaDESK; POEMs; the transBOARD; and the ambientROOM. These are briefly presented here to map out the interface design space explored by the thesis. The Tangible Infoscapes and POEMs interfaces were early design exercises, and were supported with only partial implementations. The Tangible Geospace prototype on the metaDESK was most extensively fleshed out, and will be returned to in Chapter 6 for more detailed discussion.

The metaDESK, transBOARD, and ambientROOM platforms were developed in close collaboration with Prof. Hiroshi Ishii as the first research platforms of the Tangible Media Group. These were implemented with valuable support from numerous undergraduate assistants, as well as major contributions to the ambientROOM by other graduate students.

3.1 Tangible Infoscapes and the metaDESK

Tangible Infoscapes was the first prototype application based on the metaDESK platform, and a first attempt at exploring notions of tangible user interfaces. Tangible Infoscapes considered physical media like photographic slides and printed books as physical indices into digital content. These physical objects could be used both as “containers” for associated digital media, as well as handles for manipulating this information.

Tangible Infoscapes was structured around several physical surfaces. The first of these was the metaDESK. The metaDESK is a nearly-horizontal, back-projected graphical surface based on the VisionMaker™ product of Input Technologies. Inspired by the interface vision of the ClearBoard, Ishii and I regarded the metaDESK as a kind of “interactive surface” which might both sense and respond to physical-world stimuli.

The second physical surface of Tangible Infoscapes was the *active lens*. Considered as a physical embodiment of the GUI windows metaphor, the active lens was based on the physical-world arm-mounted jeweler’s magnifying lens, with the optical lens replaced by a graphical “portal” into digital space. Used in combination with the metaDESK, the active lens could also be considered as a type of interactive surface.

3.1.1 Scenario

In our scenario for Tangible Infoscapes, Ishii and I considered two types of digital content: video sequences represented by graphical keyframes, and Japanese "Renga" art, collaborative sequences of "chained" paintings. We represented this digital content with business card sized cardboard objects called "hypercards," each printed with a representative image, textual annotation, and barcode URL/ID for the virtual "contents" of the object. These hypercards act as containers and handles for image sequences – in our case, keyframes from video clips and image sequences from "Renga" artworks. Hypercards were used in place of normal slides and photographs partly because of technical issues involving barcode IDs, and partly due to their simultaneous use with the transBOARD platform.

By placing a hypercard onto the desk and registering it with a barcode wand, a *digital shadow* of the object – a composite of annotating textual and graphical information reflecting the object's identity – was displayed on the desk's surface surrounding the hypercard (Figure 3-2). At the same time, the arm-mounted active lens positioned above the desk displayed a dynamic three-dimensional graphic representing each physical object as a stack of semi-transparent images, layering keyframes of the video-objects and iterations of the Renga-objects (Figure 3-1). These image stacks were thought of as a kind of three-dimensional projective digital shadow of the physical hypercard objects.

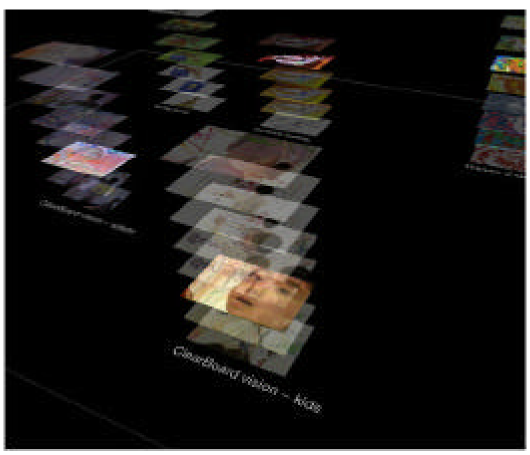

Figure 3-1: Tangible Infoscapes, active lens view

Figure 3-2: Tangible Infoscapes in "use" *(photo by Web Chappell)*

3.1.2 Implementation

The Tangible Infoscapes prototype was built on the VisionMaker™-based metaDESK platform, using a barcode wand for hypercard identification. The active lens was prototyped with an arm-mounted magnifier lens, spatially tracked with an Ascension Flock of Birds six degree-of-freedom sensor. 3wish, an [incr Tcl]-based suite of 3D graphics extensions originally implemented for earlier thesis work in the VLW [Ullmer 1995, 1997b], was used to implement the three-dimensional active lens graphics stacks and the metaDESK's "digital shadows." Computer vision in the visible-light regime was explored for tracking hypercards, but proved problematic. Computer vision for tracking LED-tagged hypercards was implemented, but only for a single unidentified hypercard. In practice, a light pen was used to drag along digital shadows to manually synchronize them with their physical "hosts." Networking support was not yet integrated into 3wish, so the 3D digital shadows displayed by the active lens were decoupled from their 2D desk-surface peers.

3.1.3 Design Lessons

The Tangible Infoscapes prototype was a useful first design iteration on the metaDESK, but also presented many issues requiring further development. We were unhappy with the use of the barcode wand to interact with TUI elements, and also were unable to track the location of the hypercards in this iteration. The choice of application domains – that of 3D visualization and interaction with abstract image collections – was also difficult in that it involved both the open hard challenge of abstract information visualization, in addition to our focus challenge of TUI design. As a result, in selecting a domain for our next metaDESK prototyping effort, we chose to focus on an interaction with geographical space, providing room for rich visual representations of a more well-understood nature while letting us focus on the core TUI concerns.

3.2 POEMs

A second project exploring physical objects as containers of digital information was POEMs. An abbreviation for "physical objects with embedded memories," POEMs was a joint project with graduate student Jennifer Glos exploring physical objects as containers for digital "memories" surrounding their history and use. POEMs was presented as a vision video [Glos 1996] illustrating the problem domain and two prototype interactions.

3.2.1 Motivation and Scenario

POEMs explores the linking of digital "memories" to physical objects. It assumes the existence of digital recording devices (cameras, audio and video recorders, etc.) which tag their output with time, place, and creator, as well as wearable computers which monitor their users' interactions with the physical environment.

Glos and I developed two scenarios: an interaction with a seashell picked up on the beach, and the augmented perusing of a well-thumbed book. In the case of the seashell, we imagine the beachgoing user records a variety of photographs, audio, and video clips, and collects physical souvenirs of her visit – among these, a seashell. We imagine each of these physical interactions as setting off a ripple of digital associations, linking the physical seashell with other digital and physical media within spatial or temporal proximity to the gathering of the shell.

Figure 3-3: POEMs, seashell interaction

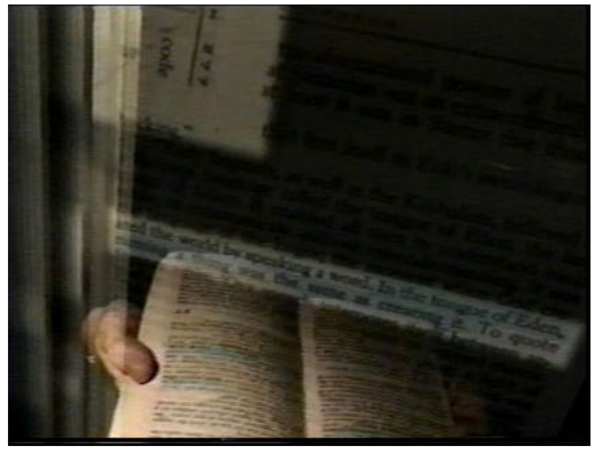

Figure 3-4: POEMs, book interaction

The video illustrates use of the physical seashell object to navigate an audio- and image-collage of the “memories” of the shell, as displayed by a device such as the metaDESK. We use the metaphor of the seashell as a focusing “lens,” causing each image of the collage to fade between transparency and opacity as the seashell is re-oriented over the metaDESK. This interaction was inspired by geologists’ analysis of mineral thin-sections with polarized-light microscopes. Here, reorienting the sample causes different mineral facets to grow bright and dim, an interaction through which the material’s character may be assessed. A still from the video presentation of this interaction is shown in Figure 3-3.

The seashell illustrates a POEM with a history of physical association. The second video scenario, that of augmented interaction with a well-read book, considers POEMs with histories of interaction. In the book example, we digitized each page of a book which the owner had dog-eared, highlighted, or otherwise marked. By viewing the physical book through the active lens device of the metaDESK, we imagined users being able to view with the metaDESK’s active lens the dog-ears and marks of interactions by friends or perhaps famous personalities with other copies of the same book. In this fashion, we complemented the book’s traditional use as a physical annotation carrier with the digitally-augmented role of a collaborative vehicle for sharing this physical annotation with friends, students, or perhaps a more remote subscribing audience.

3.2.2 Implementation

The 3D graphical bindings to the physical seashell and book objects of the video were prototyped with 3wish. The seashell example was driven by a Flock of Birds tracker. Rotation of the tracker caused images from the 3D texture-collage to change their level of transparency, visually foregrounding subsets of the image space corresponding with specific tracker orientations. Translation of the tracker allowed the user to spatially navigate the collage-space. The book example was similarly implemented, such that translation of the tracker caused a transparency-enhanced visualization of the book to shift pages, allowing the user to view virtual augmentations to the book's physical content. An image of this interaction appears in Figure 3-4. The notions of binding digital media to passive, untagged physical objects was supported by the idea of proxdist computing [Ullmer 1996a], discussed at length in Chapter 5.

3.2.3 Design Lessons

Our informal presentations of the POEMs vision video gathered enthusiastic response, and as a vision video, POEMs seemed to successfully illustrate compelling computationally augmented interactions with everyday physical objects. Also, POEMs served as another grounding example for notions of physical objects as containers for digital content.

At the same time, a variety of uncertainties arose from the exact nature of interaction with the individual physical object as POEM, both in the "authoring" of new associations and in the viewing of associated "contents." For instance, the seashell interaction of POEMs illustrated a carefully laid-out spatial collage of images and objects collected from a beach trip. To what extent was the system expected to automatically arrange this content, and to what extent would this in practice reflect explicit authoring by the user?

Moreover, if the seashell indeed "automatically" integrates links to "associated" digital interactions (a potentially plausible, if technically ambitious goal), how is recording-interaction to be distinguished from retrieval-interaction with the object? Are these recording- and retrieval-interactions driven purely by the POEM object and some augmenting interactive surface? If so, what languages of gesture and reference might be required for the POEM, and if not, what manipulating instruments and interactions might be necessary?

In conclusion, the POEMs example seemed an evocative vision for augmented physical interaction with history-enriched objects, but consistent with its vision-video origins, left many difficult unanswered questions about how such interactions might be orchestrated in practice.

3.3 Tangible Geospace and the metaDESK

Tangible Geospace was the first fully implemented application of the metaDESK platform, and the first fully operational implementation of tangible user interfaces in the thesis and the Tangible Media group. As such, it forms the most fully-developed thesis implementation, and will be considered in the remaining document at some length. The operational scenario of Tangible Geospace will be presented below, while a full discussion of its implementation will be presented later in Chapter 6.

Tangible Geospace presents a prototype interaction with geographical space driven by physical objects. The name and content domain was partially inspired by the earlier geographic visualization research titled "GeoSpace" [Lokuge 1995] by Lokuge and Ishizaki in the MIT Media Lab VLW. An overview image of the Tangible Geospace prototype on the metaDESK is shown in Figure 3-5.

Figure 3-5: Tangible Geospace, overview

The base metaDESK platform and active lens devices were introduced earlier with the Tangible Infoscapes prototype. Tangible Geospace introduces a number of additional physical interface devices. First, the earlier notion of physical objects as containers for digital content is given new conceptual form as the "physical icon" or *phicon* – a physical instantiation of the GUI icon metaphor. Phicons act both as containers and handles for physical interaction with digital content.

Another interface element introduced in Tangible Geospace is the *passive lens*. The active lens was introduced earlier with the Tangible Infoscapes prototype. The passive lens is a variant which employs a passive transparent surface for use on the metaDESK's surface. Augmented by the desk's graphical back-projection, the passive lens acts as an independent display device like the active lens, even though it is a passive physical object.

Lastly, Tangible Geospace introduces two new physical instruments: the rotation-constraint instrument and the digital flashlight. The rotation-constraint instrument mechanically constrains physical scaling and rotation interactions on the metaDESK, resolving ambiguities presented by certain later-described phicon interactions. The rotation-constraint instrument was first suggested to Ishii and I by Bill Verplank of Interval Research. The digital flashlight allows a kind of "digital light" to be projected into the Tangible Geospace scene, realized in Tangible Geospace as a translucent graphical overlay layer on the metaDESK's surface. In addition, the "spectrum" (layer of interpreted contents) of this digital light projection can be manipulated through rotation of the digital flashlight's front bezel.

3.3.1 Scenario

Several physical objects and instruments for interacting with geographical space sit in a translucent holding tray on the metaDESK surface. By placing a small physical model (phicon) of MIT's Great Dome onto the desk, a two-dimensional map of MIT appears on the desk beneath the object, bound to the Dome object at its location on the map. (Figure 3-6) The Dome phicon was constructed out of transparent machined acrylic, designed to minimize occlusion of the desk surface, while aesthetically enhancing continuity between the physicality of the phicon and the virtuality of the desk-based map display.

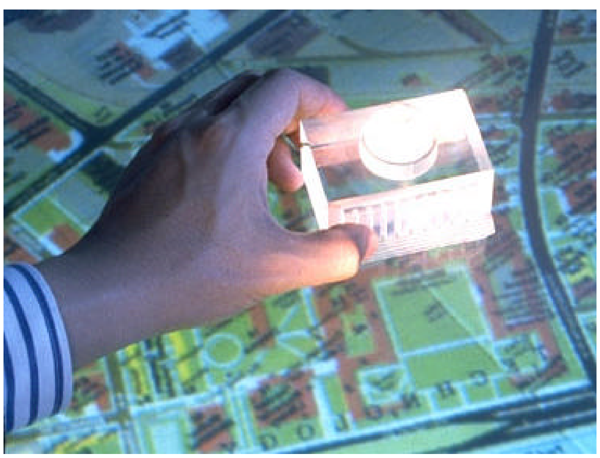

Figure 3-6: Great Dome phicon in Tangible Geospace

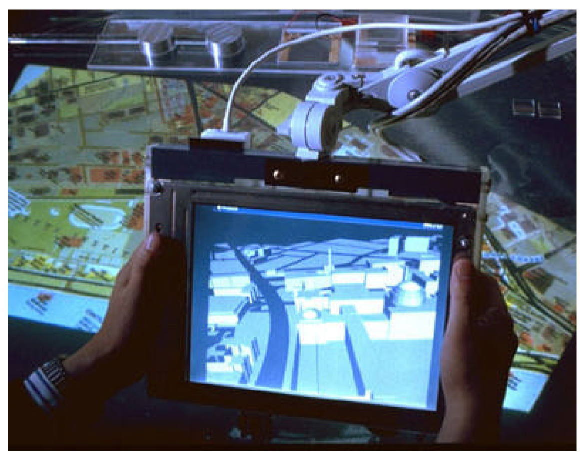

Figure 3-7: Active lens in Tangible Geospace

Simultaneously, the arm-mounted active lens (Figure 3-7) displays a three-dimensional view of MIT with its buildings in perspective, coupled to the Dome model and desk-based 2D map such that movement of the active lens drives a navigation of 3D space consistent with an optical lens metaphor (allowing viewpoint translation, zooming, etc).

The Dome phicon acts both as a container for the digital content of the MIT space, and as a physical handle for manipulating the map. By rotating or translating the Dome object across the desk's surface, both the 2D desk-view and 3D lens-view are correspondingly transformed. The user is thus interacting visually and haptically with three spaces at once – the physical space of the Dome object; the 2D graphical space of the desk's surface; and the 3D graphical space of the active lens.

The user then takes a second physical icon from the holding tray, this time representing the Media Lab, and places it onto the surface of the desk. Now there are two physical constraints and handles on the MIT space, allowing the user to scale or rotate the map by moving one or both objects with respect to each other. The user may grasp and manipulate both phicons simultaneously with two hands. Alternatively, two users may independently grasp one of the building phicons, cooperatively manipulating the geospace.

In this manner, unlike the GUI use of the mouse, Tangible Geospace has no one solitary locus of control. Rather, the phicon interaction is constrained by the physics of the physical environment, supporting multiple pathways of single- and multi-user interaction likely unrealizable with the mouse-based paradigm. The two-phicon interaction is similar to an interaction within the Bricks work [Fitzmaurice 1995] co-authored by Ishii, but differs in the phicon object-semantics of the interaction.

Figure 3-8: Passive lens, with Great Dome phicon

Figure 3-9: Rotation-constraint instrument

With the geospace active on the metaDESK, we have implemented several physical instruments for manipulating and viewing the space. First, we have created an instrument called the "passive lens," (Figure 3-8) a wood-framed transparent surface that functions as an independent display when augmented by the back-projected desk. Since passive lens devices are passive transparent surfaces, many variously afforded lenses might be used simultaneously with no additional active display resources (i.e., additional computer-driven screens).

Using the passive lens, the user may interact with an additional inline 2D overlay view of the MIT campus – potentially a satellite view, schematic of campus infrastructure, or alternate view consistent with the physical instantiation of the Magic Lens interaction metaphor [Bier 1993]. We displayed a processed version of the same 2D campus map, but our implementation supports any properly registered 2D image overlay. We implemented a single passive lens overlay, but our infrastructure supported multiple alternate views.

When manipulating Dome and Media Lab phicons on the desk to scale and rotate the geospace, the simultaneous rotation of both phicons poses an ambiguous interaction. The software resolution of this ambiguity is discussed in section 6.4. As an alternative to the two-phicon scaling/rotation interaction, we implemented a rotation-constraint instrument made of two cylinders mechanically coupled by a sliding bar (Figure 3-9). This instrument allows manipulation of the two control point interaction with intrinsic scaling and rotation constraints.

Finally, we have implemented a physical flashlight instrument which follows an optical metaphor consistent with the active lens and passive lens devices. The flashlight projects onto the desk a translucent image with additional 2D overlay views of MIT campus. By turning the rim of the flashlight, the "frequency" of the light may be varied, actuating fades between multiple overlay layers.

The implementation and design lessons of the Tangible Geospace prototype are discussed in detail within Chapter 6.

3.4 The transBOARD

In addition to the metaDESK-based Tangible Infoscapes, POEMs, and Tangible Geospace prototypes, Ishii and I developed in parallel two other platforms which emphasized different physical-world interface modalities. The first of these is the transBOARD, a digitally-augmented physical whiteboard. The transBOARD is based on a Microfield Graphics Softboard™ product donated by Steelcase Corp. The Softboard product instruments a physical whiteboard surface with scanning infrared lasers that can track the motion of specially-tagged whiteboard markers. As a result, the Softboard supports "natural" pen-based interaction with the whiteboard surface, in addition to digitally transcribing pen strokes in realtime.

Interaction with the transBOARD differs from the metaDESK in several important fashions. First, while the transBOARD possesses a rich and expressive input pathway, it has no intrinsic capacity for computationally-augmented output and feedback save perhaps the coarse beeps and blips of the built-in speaker. This makes for a challenging tangible interface task, for in the absence of additional augmenting displays, interface objects must support sufficient “legibility” of function and state to function independently from computational feedback.

Secondly, the transBOARD is a vertical surface. As a result, physical-world gravity renders object interactions with the transBOARD differently constrained than interactions with the metaDESK’s near-horizontal surface. While gravity makes the metaDESK surface a largely persistent physical space – objects placed on the DESK tend to remain in place until moved again by the user – the vertical transBOARD surface affords these properties to a much narrower repertoire of relatively flat, magnetic objects.

3.4.1 Scenario

In designing the transBOARD, Ishii and I thought of the board as a kind of interactive surface supporting interaction with two kinds of physical objects: whiteboard markers, with which contents of the transBOARD could be authored; and container/conduit hypercard objects, which could be used either to store or transmit whiteboard activity. The transBOARD prototype is illustrated in Figure 3-10.

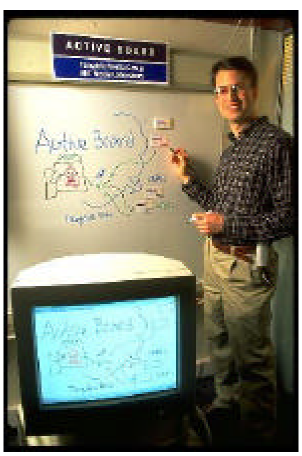

Figure 3-10: transBOARD "in use" *(photo by Web Chappell)*

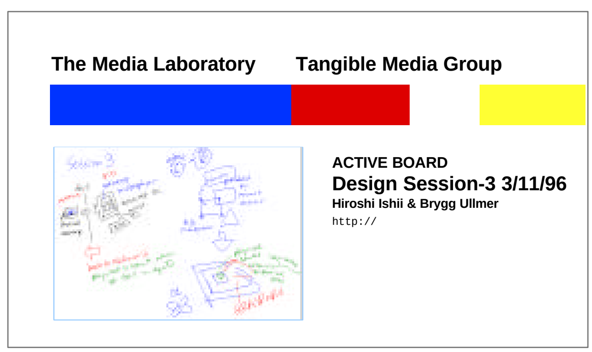

Figure 3-11: "Hypercard" example

For the original transBOARD prototype, we used a magnetically backed version of the hypercard objects introduced by Tangible Infoscapes, pictured in Figure 3-11. With the support of research assistant Chris Fuchs, we prototyped two transBOARD interactions. In the first, the user places a hypercard onto the board's surface, and registers the card's arrival with the swipe of a barcode wand. This action is acknowledged with an audio cue, and initiates recording of stroke activity by the transBOARD system. When the whiteboard session completes, the user wands the hypercard barcode a second time. This is again audibly acknowledged, and causes the transBOARD to store the stroke session on a Internet-linked Web server under the ID of the hypercard barcode. By bringing this hypercard to a remote interactive surface such as the metaDESK, the hypercard can be used to retrieve the "contained" stroke session for playback.

A second prototype used hypercards as a gateway or conduit for the broadcast of live whiteboard activity. After registering a hypercard, the transBOARD begins transmission of live session data onto the Internet. By bringing a copy of the transmitting hypercard to a remote interactive surface, users can remotely access the ongoing whiteboard session.

3.4.2 Implementation

The original transBOARD prototype was implemented by research assistant Chris Fuchs. The "container" hypercard application worked through scripted control of Microfield's Softboard stroke recording software. Activation and storage of stroke recording was triggered by barcode wand events. Resulting stroke sessions were automatically stored on a web server, keyed off the hypercard's barcode ID. Subsequent scanning of the hypercard on a remote computer would invoke the Web-based retrieval and playback of the stroke session, as virtually contained by the physical hypercard. The "conduit" application employed a C++ language "stroke server" which broadcast live transBOARD stroke data to remote network clients. A Java-based stroke client allowed transBOARD sessions to be viewed by remote network machines in a session invocable by the conduit hypercard's barcode ID.

3.4.3 Design Lessons

The transBOARD provided an interesting platform for interaction with a simple augmented surface, and our resulting prototype applications spawned the fruitful "containers and conduits" metaphor discussed later in the Models section. However, many aspects of our first generation hypercard-based transBOARD interface proved unsatisfactory in practice. As with the simultaneously-developed Tangible Infoscapes prototype, the barcode interaction with hypercards proved cumbersome and distracting. Also, the accessing of hypercard contents through traditional PC displays was unappealing as the primary viewing modality of a tangible user interface.

At a deeper level, the metaphorical role of the hypercard objects proved challenging. While the container/conduit metaphor was straightforward in principle, how should this translate in practice to physical manipulation of the hypercard? For instance, what granularity of abstraction is most meaningful for the content-associations of hypercards? Are hypercards most meaningfully mapped to stroke data content on a per-session, per-user, per-project, or some other basis? As abstract physical "containers," hypercards could in principle map to any of these approaches. However, without resolving these distinct roles and the different interface approaches they represent, it is difficult to argue cogently about hypercards and their transBOARD interface role.

Similarly, the command vocabulary of hypercards is another outstanding issue. Is the act of placing the hypercard onto the transBOARD ideally equivalent to activating recording "within" the hypercard, and the removal of a hypercard an embodied "stop recording" request? If so, how does this approach resolve ambiguous cases (e.g., the arrival of a hypercard already containing stored content) or support additional hypercard functionality? If not, what additional TUI tangibles are necessary to successfully interact with the transBOARD?

3.5 The ambientROOM

The last TUI platform explored in the thesis is the ambientROOM, pictured in Figure 3-12. The ambientROOM is based on the Personal Harbor™ product from Steelcase Corp., a 6'x8' enclosed mini-office installation. Ishii and I thought of the Harbor as a strong complement to the repertoire of interface prototypes developed on the metaDESK. With the metaDESK, users maintain a more traditional "external" relationship, with users interacting outside and around a graphically intensive, cognitively foreground surface. The ambientROOM offered the possibility of surrounding the user within an augmented environment – as our Steelcase representative suggested, "putting the user inside the computer" – with the potential for providing more subtle, cognitively background augmentations to activities conducted within the room.

Figure 3-12: ambientROOM, overview

Ishii and I developed the ambientROOM with undergraduate assistant Craig Wisneski as a platform supporting the expression of "ambient media" – ambient light, sound, airflow, water movement, and physical motion used as peripheral displays at the background of user attention. The individual ambientROOM interfaces introduced in the following section were implemented by Wisneski and graduate students Matt Gorbet, Scott Brave, and Andrew Dahley upon their arrival in the group. At the same time, these were developed as a part of Ishii and my original ambientROOM design introduced within [Ishii 1997], along with early input from Gorbet and Maribeth Back. Thus, while grounded upon prototypes individually implemented by other students, the interface instances within the ambientROOM formed part of the broader thesis research, and serve as strong design examples in later discussion of models for tangible user interface.

3.5.1 Scenario

Inspired by the Fields and Thresholds work of Raby and Dunne [Dunne 1994] and the Live Wire of Jeremijenko [Weiser 1995], the first-generation ambientROOM explored a theme of peripheral awareness for external activity, especially of activity attributable to people. The first-generation ambientROOM prototypes broadly fit into the categories of activity displays and activity controls.

Four types of activity displays were explored. The first two were initially coupled to the activity of a resident hamster in our laboratory, for which cage temperature, light level, and wheel motion had already been Web-instrumented by graduate student David Small. This hamster telemetry was chosen as a simple activity source for display within the ambientROOM.

First, hamster wheel motion was coupled to the vibration of an ambientROOM phicon symbolizing the hamster. A fairly foreground display, the hamster phicon could be grasped by the user and used to physically transfer activity to a more ambient display. This second display transformed each quanta of hamster motion into the activation of a plunger in an illuminated water tank. The resulting water ripples reflections onto the ambientROOM's ceiling formed a subtle display of light and shadow experienced at the periphery of the ambientROOM space.

Another ambientROOM display was coupled to the activity of people in an adjoining space of our laboratory. The movement of people within the space was reflected in the changing frequency and pattern of illuminated patches projected onto the ambientROOM's wall. This kind of "active wallpaper" served as another subtle peripheral display. A last ambient display synthesized shifting sounds of birds and rainfall and modulations of ambient room lighting, simulating ambient displays reflecting e-mail arrivals and other computer network activity.

In addition to these ambient activity displays, several activity controls were deployed in the ambientROOM. First, in the face of many possible activity sources the user might wish to occasionally monitor, some vehicle for managing ambient activity sources seemed necessary. Here, a physical bottle was designed as a container for evolving digital content. By uncorking the bottle, contained bits were "let out" into the room, manifesting themselves in the prototype as sounds of vehicle traffic (potentially representing a high-traffic condition on a monitored computer network). Closing the bottle would still these displays, returning the bits back inside its interior.

A second activity control was built around an analog clock. A user recently absent from the room might wish to review activity displays from the past few hours, or skim forward in time to peruse anticipated events. Where uninstrumented clocks serve as physical instruments for passively monitoring time's passing, the hands of the ambientROOM clock could be grasped and turned to actively navigate and manipulate time-based ambient displays. In response to manipulation of the clock's time, the ambientROOM prototype would shift through the ambient sound and lighting displays of past hours, graphically projecting the actual time and other augmenting data onto the physical clock's face. We also imagined force-feedback instrumentation of the clock's hands such that it might haptically display augmenting information during the user's navigation, as well as returning hands to the correct time on completion of a user's interaction.

3.5.2 Implementation

Sensing and display in the ambientROOM were based upon MIDI instrumentations coordinated with OpCode's MAX MIDI control software. MIDI-controlled dimmers adjusted room lighting, MIDI-based rotation and electrical contact sensors monitored manipulation of the clock and bottle, and a MIDI sampler managed sound playback. A MIDI-based electric field sensing "Fish" unit and doppler radar device from the Media Lab's Physics and Media group were used to monitor human movement in space. PIC microcontroller-based “Crickets” from the Media Lab’s Epistemology and Learning group were used to monitor hamster wheel movements, drive hamster phicon vibrations, and actuate the water tank’s solenoid. Intel-based PCs were used to control video projectors augmenting the room’s clock, walls, and desk surface.

3.5.3 Design Lessons

The ambientROOM prototype raises many interface issues. First, how should the many potentially interesting sources of activity and more limited range of ambient displays be associated? Similarly for the case of the bottle, how should specific repertoires of activity be mapped to the abstraction of the empty bottle, allowing it to serve as more than an elegant switch to an arbitrary digital binding?

Another issue involved what style, intensity, and orchestration of ambient displays would be optimal in the course of normal human habitation. Running as a demonstration in our space, the incessant chirping of the ambientROOM’s bird songs grew annoying, and the sounds of highway traffic initially mapped to the bottle could hardly be considered “subtle.” Some observers also noted that modern workspace design has been driven in large part by the effort to shield inhabitants from environmental “ambience.” Thus, while well-received as a concept demonstration, the ambientROOM motivates many challenging design issues awaiting further study.

4 TUI Models

4.1 Introduction

The thesis has introduced the notion of "tangible user interfaces" or "TUIs" as user interfaces which use physical objects, instruments, surfaces, and spaces as physical interfaces to digital information. However, this definition is in need of narrowing. Clearly keyboards, mice, television remote controls, and so forth are all "physical objects," even though these do not meet our conception of tangible user interface.

One way TUIs can be understood is in comparison with graphical user interfaces (GUIs), the current dominant modality of human-computer interaction. Graphical user interfaces are based on graphical representations of virtual user interface elements – disembodied widgetry such as the windows, icons, menus, and other elements pervasive in modern computer interfaces. Physical input peripherals such as the mouse are used to indirectly access and manipulate these virtual interface elements, but it is fundamental to the GUI that graphical widgetry exists independently from the physical world.

In contrast, tangible user interfaces are based on tangible entities which physically embody the digital information and interfaces they represent. For example, in the instance of the physical icon or "phicon" introduced earlier, the phicon physically embodies both the digital content and the means for manipulation of its digital associations – thus seamlessly coupling virtuality and physicality by embodying both at once.

The distinction of TUIs from other models of user interface is also in significant part defined by the interaction metaphors invoked by the interface and perceived by the user. This chapter explores several models which bear out the notion of TUI. First, the chapter reviews the notion of TUI "tangibles" – the physical objects, instruments, surfaces, and spaces which compose tangible user interfaces. Three TUI models motivated by the thesis prototypes are then developed. The first of these explores physical instantiation of GUI widgetry such as windows, icons, and controls. This model was dominant in the design of the Tangible Geospace prototype on the metaDESK. Next, an optical metaphor for tangible user interface is developed. Here, notions of digital light, shadow, and lenses are explored. A last interface metaphor is that of "containers and conduits." In this case, tangibles are used to virtually contain and communicate physically-associated digital information.

It is worth noting that the following models are intended to be illustrative, while remaining grounded within the TUI abstractions explored by the thesis prototypes. The models presented are neither formal, exhaustive, nor prescriptive, all levels of discussion which remain to future work.

4.2 TUI "Tangibles:" Physical objects, instruments, surfaces, and spaces

The thesis has previously introduced the TUI notion of physical objects, instruments, surfaces, and spaces as physical interfaces to digital information. This section reviews and abstracts the kinds of objects, instruments, surfaces, and spaces developed in the thesis, along with the interface roles these serve.

First, though core to the subject of the thesis, the word "tangible" has not yet been defined. Webster defines "tangible" as that which is "capable of being perceived, esp. by the sense of touch." [Webster 1997] As an adjective describing *tangible* user interfaces, our use of "tangible" refers to interfaces leveraging human sensation of and engagement with the physical world. The realm of the tangible includes both haptic, visual, and audible engagement with graspable physical things, as well as ambient senses of light, warmth, vibration, motion, and so forth.

The word "tangible" appears in English not only as adjective, but also as noun. Where I before have verbosely referenced the range of "physical objects, instruments, surfaces, and spaces used as physical interfaces to digital information," I will now more compactly describe these in the aggregate as "tangibles." The distinction of our use from that of common English is of physical substance *employed as user interface to digital information.* It worth noting that a related application of the "tangibles" term has been in earlier use at Interval Research Corporation.

It is also important to note the thesis subdivision of tangibles as objects, instruments, surfaces, and spaces is neither formally rigorous, exhaustively inclusive, nor necessarily disjoint. Both object and instrument tangibles can contain interactive surfaces, as demonstrated by the active and passive lens instruments. Similarly, the distinction between object tangibles and instrument tangibles is sometimes unclear; e.g., an object tangible used as physical constraint may reasonably be considered an instrument. Finally, as will be discussed in the containers and conduits section, the object/instrument/surface/space categories do not clearly provide for the use of liquids or gases as tangibles, even though such applications exist. While a "substance" category could be added, this use has not been developed in the thesis, and so will be left for future discussion. Other similar distinctions and extensions also likely await discovery.

4.2.1 Object Tangibles

The first kind of tangible in the thesis is the physical *object*. Object tangibles can be thought of as discrete physical things which may be either man-made or naturally occurring, and may or may not be instrumented in electronic, mechanical, chemical, or other fashions. Thesis examples of object tangibles include the hypercards of Tangible Infoscapes and the transBOARD; the seashell and book of POEMs; the Media Lab and Great Dome phicons of Tangible Geospace; and the hamster phicon and bottle of the ambientROOM.

Object tangibles can be considered in three respects: their interface role within TUIs; their physical instantiation; and their technologies of augmentation. Object tangibles may be used as physical embodiments of or indices into digital information. Phicons are broadly demonstrative of this capacity. Object tangibles may also be used as physical constraints to computational processes, e.g. the Tangible Geospace two-phicon rotation and scaling interactions. Object tangibles may either embody or be interpreted as commands or instructions, e.g. hypercard placement on the transBOARD activating the storage or transmission of strokes. Object tangibles may also serve as “containers” or “conduits” for digital content, a role demonstrated by transBOARD hypercards and the ambientROOM bottle which will be discussed in section 4.5.

Object tangibles may be embodied in a broad range of physical forms. They may be naturally occurring or man-made, like the seashells and books of POEMs. They may be easily manipulable like the phicons of Tangible Geospace, naturally rooted in the physical environment like a boulder or tree, or mounted as fixtures like a coat rack or kitchen cabinet. Object tangibles may be transient like an ice cube; synthetically derived on demand, as from a 3D printer; or of longstanding human wear and use, like an old wooden chair or table. Lastly, object tangibles may be embodied as the actuality or representation of pre-existing physical entities, like the seashell of POEMs and building-miniatures of Tangible Geospace, or synthesized as new physical forms original to TUIs, like the Triangles of Gorbet and Orth [Gorbet 1997] and multi-location objects of Brave and Dahley [Brave 1997].

Object tangibles may be technologically augmented in many fashions. They may be actively instrumented with electronic sensors, motors, and displays; passively instrumented with optical or RFID tags; or remain completely uninstrumented, with sensing and augmentation by environmental capabilities. In the latter cases, passive objects may be mediated by projecting onto them, like the walls and clock of the ambientROOM; projecting through them, like the passive lens of Tangible Geospace; intermediated with displays like the active lens; or even physically actuated through techniques like the Phantom Chess mechanism [Fitzmaurice 1996]. Objects may be passively sensed with technologies like computer vision, as well as inertially tracked through the “conservation of impetus” (see [Ullmer 1996a]).

4.2.2 Instrument Tangibles

The second kind of tangible in the thesis is the physical *instrument.* The American Heritage dictionary defines "instrument" as an "implement used to *facilitate* work" (italics added) [Heritage 1994]. Where object tangibles often serve to physically *instantiate* digital information, instrument tangibles are used to *operate on* digital information embodied via object, surface, space, and even other instrument tangibles. Thesis examples of instrument tangibles include the active lens, passive lens, constraint instrument, and flashlight of the metaDESK, as well as the clock of the ambientROOM.

Instrument tangibles can be used to provide augmenting views of TUI objects, surfaces, spaces, or other instrument tangibles, as demonstrated by the active lens, passive lens, and digital flashlight thesis examples. These instances also demonstrate instruments for physically navigating digital content, leveraging kinesthetic senses of the human body in space. Instrument tangibles can be used to mechanically constrain physical interactions with digital content, as demonstrated by Tangible Geospace's constraint instrument. Instrument tangibles also afford the physical probing of digital content, discussed as a possibility of constraint instrument servo-mechanization. Finally, instrument tangibles provide possibilities for intimately interlinking input-space with output-space, as demonstrated by the hands of the ambientROOM's clock.

As with object tangibles, instrument tangibles may be realized in many physical forms. Instrument tangibles may be crafted of "inert" physical substances like the constraint instrument of Tangible Geospace, or of more actively-mediating materials such as the metaDESK's active lens. Their structures may be ground-referenced like the active lens, "surface-referenced" like the constraint instrument and passive lens, or a combination of body- and surface-referenced like Tangible Geospace's digital flashlight. Instrument tangibles may also be grounded in the form and function of pre-existing physical devices, like the thesis instrument examples thus-far cited, or may take entirely new TUI-specific forms. The latter is partially exemplified by the inTouch device of Brave and Dahley [Brave 1997] and the ambient fixtures of Dahley. However, these perhaps bear more resemblance to object tangibles, which (at least as interestingly) leaves the case of novelly-conceived instrument tangibles without strong precedent.

Instrument tangibles may be technologically augmented in many fashions. The object tangibles discussion referenced some of the augmentation modes of the active lens, passive lens, and digital flashlight. It is also interesting to note that the passive lens, constraint instrument, and digital flashlight examples all operate in close "collaboration" with the interactive surface of the metaDESK. This example of interdependency between TUI tangibles and environmental capabilities is a recurring theme in the thesis, a theme formalized in some respects within the Mechanisms chapter.

4.2.3 Interactive Surfaces

The notion of surfaces in the context of tangible user interface draws from the ClearBoard discussion of [Ishii 1994]. Here, Ishii described a vision of architectural space where ceilings, walls, desktops, floors, and other *interactive surfaces* could each serve as gateways between physical and digital space. This "surface" concept is highly evocative given its relation to "interface." In defining "interface," the American Heritage dictionary reads "a *surface* forming a common boundary between adjacent regions," (italics added) – in the TUI context, the regions of physical and digital space. Thus, in a sense the interactive surface can be imagined the most direct physical manifestation of physical-digital interface. Thesis examples of interactive surfaces include the metaDESK and transBOARD surfaces; the walls, ceiling, shelves, and desk of the ambientROOM; and the faces of the active lens and passive lens devices.

Interactive surfaces play a major role as a locus and vehicle for bridging digital and physical space. Interactive surfaces may serve as graphical mediators for opaque or transparent objects and instruments, as applied to the passive lens, phicons, and digital flashlight of the metaDESK. They may also serve as locales of object sensing, identification, and tracking, as with the metaDESK, transBOARD, and shelf+desk surfaces of the ambientROOM. Interactive surfaces may serve as augmenting reference- and base-planes for "surface-referenced" instruments, like the metaDESK's passive lens and constraint instruments. They may serve as mobile visual and haptic intermediaries between digital and physical space, as with the metaDESK's active lens and passive lens devices. Finally, interactive surfaces may serve as the loci for ambient mediations of digital light, shadow, and motion, as with the projections on the ambientROOM's ceiling and walls.

Interactive surfaces may be deployed in many forms. They may be realized as architectural surfaces, such as the ceiling, walls, and floor explored with the ambientROOM and transBOARD. They may appear on tables, desks, shelves, and other surfaces of furniture and fixtures, as with the metaDESK and ambientROOM. They may also be integrated as faces on object and instrument tangibles, as with the active and passive lens devices. They be intensively visually mediated, like the metaDESK; ambiently mediated, as with the ambientROOM's ceiling and walls; or without visual mediation, as with the transBOARD. They may also be opaque, as with surfaces of the ambientROOM and transBOARD, or functionally transparent, as with the ClearBoard and the LCD surfaces explored in [Kobayashi 1992], which frost under digital control to bear graphical projection, or shift to transparency as unmediated windows.

Interactive surfaces can be augmented in various fashions. Visual augmentations may be front-projected, as in the ambientROOM; back-projected, as with the metaDESK and passive lens; or endowed with integrated graphic displays, as with the active lens. From a sensing standpoint, interactive surfaces may integrate internal sensors, like antenna coils of the ambientROOM desk and shelf RFID readers; external scanners, like the transBOARD; or internal or external monitoring like the computer vision of the metaDESK.

4.2.4 Interactive Spaces

Interactive spaces are concerned with the spatial relations between multiple surface, object, and instrument tangibles, as well as the spatial expression of ethereal qualities such as air movement, ambient temperature and aroma, and diffuse or spatialized sound. Thesis prototypes of interactive spaces include the operating volume of the metaDESK platform, as well as the spatial enclosure of the ambientROOM.

As illustrated by the metaDESK and ambientROOM spaces, TUI tangibles are not used in isolation, but rather as elements of cohesive spatial frames of reference. The Tangible Geospace example clearly illustrates this with the unified spatial volume of the geospace. The geospace is anchored at the locii of the surface-based phicons, but also extends upwards into the volume probed by the active lens and digital flashlight, and is manifested cross-sectionally by the desk and passive lens surfaces.

Here, the interactive space of the desk's operating volume binds together the operation of each metaDESK tangible. At the same time, the interactive space of the metaDESK is a kind of "thin space." To a blind person, the metaDESK would be perceived no differently when fully active than when devoid of mediation. In contrast, the interactive space of the ambientROOM can be seen as a "thick space," where space itself is rendered a mediated tangible. Air movement; ambient temperature, aroma, and light; and diffuse and spatialized audio are all clearly "tangible phenomena." Moreover, these are often perceived as properties of space itself, and not necessarily localized to any individual object, instrument, or surface tangible. In a sense, then, interactive spaces can be considered as realizing a kind of information-bearing "digital ether" which may be manifested through the induced properties of material tangibles, or through ambient properties of the physical space itself.

It is also worth noting that for "thin spaces," interactive spaces are more an outgrowth of the relations between tangibles than the presence of any specific tangible*.* For instance, the transBOARD prototype as presented in section 3.4 employs strong use of an interactive surface, but makes little invocation of a surrounding interactive space. However, Ishii and I discussed implementing an active lens device for the transBOARD, such that movement of the lens orthogonal to the transBOARD's surface could be graphically viewed as time-movement through the corresponding region of past board state. This addition would have clearly grounded the transBOARD in the notion of interactive space, without involving any change to the transBOARD surface itself.

4.2.5 Discussion

The discussion of TUI "tangibles" serves to tease apart different classes of possible physical-world interfaces. In addition, the tangibles discussion addresses dependencies between different kinds of tangibles, such as the role of interactive surfaces in augmenting object and instrument tangibles, and the role of interactive spaces for knitting together collections of interoperating tangibles. At the same time, the tangibles discussion has not captured all possible distinctions and cases of potential relevance. For instance, even relatively simple physical forms like furniture and fixtures stretch the applicability of these categories. More complex environments like physical vehicles and buildings raise the possibility of additional complicating factors. Nonetheless, the basic object, instrument, surface, and space categories provide a functional first-level grammar for discussing the tangibles of this thesis, as well as reasoning about novel TUI forms.

4.3 GUI widgetry as TUI metaphor

The first TUI model explored in the thesis grew from an interest by Ishii in rendering metaphors from the GUI paradigm into physical form. Simply stated, the notion was to physically instantiate elements of GUI widgetry pervasive in current UI design, especially the devices of icons, windows, and handles. Implicit in this interest was prior success of the Bricks work by Fitzmaurice, Ishii, and Buxton [Fitzmaurice 1995], and the notion that the graspable physical handles provided by Bricks might be profitably extended into a wider range of physically-instantiated widgetry.

4.3.1 The Desktop Metaphor

A second expression of this interest was to take the GUI desktop metaphor itself as a kind of metaphor for physically-instantiated user interface. The desktop metaphor was originally derived from a metaphor of the physical office desktop and associated objects and equipment. Quoting from an article on the Xerox Star, the first system to instantiate the desktop metaphor,

> Every user's initial view of Star is the Desktop, which resembles the top of an office desk, together with surrounding furniture and equipment. It represents a working environment, where current projects and accessible resources reside. On the screen are displayed pictures of familiar office objects, such as documents, folders, file drawers, in-baskets, and out-baskets. These objects are displayed as small pictures, or icons.

> The Desktop is the principle Star technique for realizing the physical office metaphor. The icons on it are visible, concrete embodiments of the corresponding office objects. Star users are encouraged to think of the objects on the Desktop in physical terms. You can move the icons around to arrange your Desktop as you wish. (Messy Desktops are certainly possible, just as in real life.) You can leave documents on your Desktop indefinitely, just as on a real desk, or you can file them away. [Smith 1982]

It is important to be clear that the graphical user interface is not equivalent to the desktop metaphor, and their relation is important in clarifying exactly what was taken as metaphor in the first TUI incarnations. The GUI interface can be seen as having its first roots in Ivan Sutherland's Sketchpad system in 1962 [Sutherland 1962], and its first generalized embodiment in the Xerox PARC Alto system of 1972. The Alto had a mouse, licensed from Douglas Engelbart's SRI NLS project in 1971, and used windows on a bit-mapped graphic display to allow interaction with multiple simultaneous applications. [Johnson 1989]

However, the Alto did not offer or impose one single user interface metaphor. In developing the Star successor to the Alto system, begun in 1975 and offered as product in 1981, the original designers said "… it was a real challenge to bring some order to the different user interfaces on the Alto." [Smith 1982] Additionally, it is interesting to note that the GUI notion of icons was not a pre-existing interface concept common throughout Alto user interfaces. Rather, icons were introduced by the Pygmalion Ph.D. thesis work of David C. Smith in 1975, where Pygmalion was the first large program to be written with Smalltalk on the Alto. [Johnson 1989]

The desktop metaphor of the Star grew out of an effort to realize a unified GUI metaphorical framework for interacting with documents. It is noteworthy to observe this specificity of the Star's targeting towards document manipulation, elaborated on by its designers:

> *The document is the heart of the world and unifies it.* Most personal computers and workstations give no special status to any particular application. Dozens of applications are available, most incompatible with each other in data format as well as user interface.
>
> Star, in contrast, assumes that the primary use of the system is to create and maintain documents. The document editor is thus the primary application. All other applications exist mainly to provide or manipulate information whose ultimate destination is a document. [Johnson 1989]

Thus, while graphical user interfaces per se have no inherent specificity to document manipulation, the desktop metaphor was carefully crafted to map a physical metaphor of document interaction into the virtual domain of the computer.

4.3.2 The GUI widgetry metaphor

Given this background, it is possible to explore the initial mapping of GUI widgetry into the physical domain of the TUI. Figure 4-1 shows a prototypical paralleling of GUI metaphorical objects with TUI devices. The GUI notion of "windows" is mapped to the TUI notion of the *lens*, physically manipulable windows which provide graphical views into physically-anchored digital information. GUI icons are transformed into "physical icons" or *phicons* in the TUI, physical objects which act as both containers and handles for digital content.

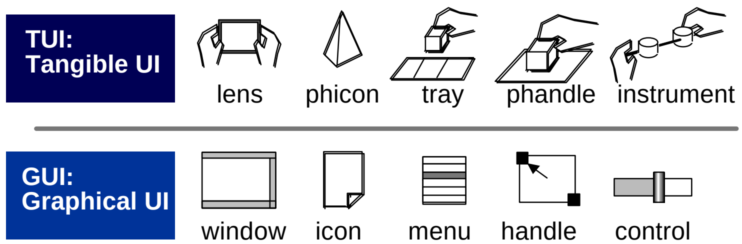

Figure 4-1: TUI instantiations of GUI elements

The menus of GUIs and textual environments are instantiated into TUI *trays*, physical compartments to which digital content and properties can be bound. GUI handles are paralleled by TUI "physical handles" or *phandles*, tangible tools by which digital content may grasped and manipulated. Finally, GUI controls such as scales and scrollbars are physically instantiated as TUI *instruments*, tangible devices by which digital information can be physically manipulated with multiple degrees of freedom and physical constraint.

These tangibles of the GUI widgetry metaphor are of mixed ancestry. The concept of the physical handle comes directly from the Bricks work of Fitzmaurice, Ishii, and Buxton [Fitzmaurice 1995]. The notion of the tray was also introduced as a part of Bricks. Further, the graspable lens relates to Fitzmaurice's Chameleon work [Fitzmaurice 1994], the PDDM, and other graspable display portals introduced in the Chapter 2, though the active and passive lens contents derive more strongly from the Magic Lens work of Xerox PARC [Bier 1993]. Likely the most significant widgetry-metaphor tangible from an interface perspective is the phicon (physical icon). The phicon notion relates most closely to Durrell Bishop's use of physical objects in the Marble Answering Machine and subsequent work prototyping physical instantiations of software applications. [Crampton Smith 1994], [Bishop 1996]

While indebted to this prior work, the GUI widgetry metaphor introduces several threads original to this thesis. First, the GUI widgetry metaphor systematically develops a whole system of user interface metaphors – lenses, phicons, instruments, and so forth – from the disembodied virtuality of the GUI back into a physical-world context. Secondly and more importantly, the TUI tangibles introduced in this discussion as following the GUI widgetry metaphor are not only physical adjuncts to the virtual devices of an underlying graphical user interface. Rather, the tangibles of TUIs *are the interface itself* – interfaces which may include graphical or other mediated contents on lenses and interactive surfaces, but where the interface itself exists in physically-instantiated form.

4.3.3 Phicons and Icons

It is now useful to return to the Desktop metaphor introduced earlier in the chapter. As previously discussed, the GUI with its widgetry is not equivalent to the Desktop metaphor. The Desktop metaphor is a particular (and perhaps the first) systematic application of graphical user interface, but follows a more specific thread of organization focused towards a more specific domain (the manipulation of documents) than the GUI writ large.

As the metaDESK formed our original platform for exploring the GUI widgetry metaphor, analogs between tangible interfaces on the metaDESK's physical desktop and graphical interfaces on the metaphorical GUI Desktop proved stimulating. One particularly interesting instance involves the relation of TUI phicons to GUI icons. As related by the developers of the Desktop metaphor on the Xerox Star,

> As bitmap-, window-, and mouse-based systems have become more common, the use of the term "icon" has widened to refer to any nontextual symbol on the display. In standard English, "icon" is a term for religious statues of pictures believed to contain some of the powers of the deities they represent. It would be more consistent with its normal meaning if "icon" were reserved for objects having behavioral and intrinsic properties. Most graphical symbols and labels on computer screens are therefore not icons. In Star, only representations of files on the Desktop and in folders, mailboxes, and file drawers are called icons. [Johnson 1989]

This clarification of icons as conceived in the Star's desktop metaphor is interesting to apply towards our use of TUI phicons. On the one hand, we have considered TUI phicons to rather broadly serve as physical embodiments of digital content. In some respects, this usage seems related to the popular conception of GUI icons as "any nontextual symbol on the display." This broad conception of phicons could be for better or for worse, remaining simultaneously potent as a thought tool for bridging virtuality and physicality, while perhaps positioned unnecessarily vulnerable to conceptual dilution from the outset.

On the other hand, the physical substance of phicons positions them in an interesting realm. Unlike GUI icons, phicons cannot spontaneously disappear or "dematerialize," cannot instantly change position or instantly morph into different physical forms (where "instantly" leaves certain leeway for advances in animatronics and other clever mechanizations). As such, phicons *do* have certain intrinsic properties resulting from their physicality, and their mediated interpretation by TUI can be seen (at least in the instance of the metaDESK) as endowing phicons with certain behavioral properties.

4.3.4 Discussion

The GUI widgetry metaphor was the first TUI interface metaphor explored in the thesis work. In a sense, it serves as a bridge from the well-explored regime of graphical user interface into the new ground of tangible user interface. The GUI widgetry metaphor leverages off powerful GUI notions like icons and windows, themselves with an original basis in physical forms, and applies them to promising physical constructions like the TUI phicon and lens. Some of these interface approaches have prior precedent – e.g., the field of augmented reality has produced many lens-like devices. Some of the GUI widgetry mappings to TUI are inexact – again with the lens, the frequent contextual independence of individual GUI windows differs in character from the physical affordances of the lens, leaving many lens instantiations with more similarity to the magic lens work of [Biers 1991].

Lastly, the natural persistence of physical-space TUI elements is partially at odds with the virtual devices of GUI, making some powerful GUI constructions intractible (e.g., rooms) and others impractical. At the same time, the GUI widgetry metaphor lends itself particularly well towards other interface regimes, such as those supporting interaction with spatial content bearing multiple dimensions of contextual information. In summary, the GUI widgetry metaphor's mapping of the GUI domain at large to the physical instantiations of TUIs provides a valuable tool for thought in projecting known virtual interface techniques into the physical realm of tangible interface.

4.4 The Optical metaphor

The second model for tangible user interface explored by the thesis is the "optical metaphor." The optical metaphor grew out of prototyping efforts with the metaDESK platform. Given our motivating interest in considering physical objects as digital interface, Ishii and I wished to link the physical manipulation of tangibles to graphical imagery on the metaDESK's interactive surface. Framed in this context, we began to think of graphical displays on the metaDESK as the visible representation of "digital light," and the content of these displays to be the interaction between digital light and information-bearing physical objects.

4.4.1 Digital Shadows

Ishii and I first considered the metaDESK's surface as a locus for transforming digital light into graphical imagery. Here, we thought of the digital information caught up within physical-world tangibles as casting "digital shadows" reflecting their content onto the surface of the metaDESK. In the Tangible Infoscapes prototype, this was realized as textual and graphical augmentations appearing underneath and surrounding the hypercard objects.

In POEMs, Glos and I explored another variant inspired by geologists' analysis of mineral thin-sections with polarized-light microscopes. For geologists, reorienting these mineral samples causes differently-oriented crystalline facets to grow bright and dim, an interaction through which the material's character may be assessed. For POEMs, Glos and I explored a seashell as an analogous optically-active tangible. Here, reorienting the shell above the metaDESK's surface caused each image-facet of the seashell's digital contents to fade between transparency and opacity. A still from POEMs' video portrayal of this interaction appears as Figure 4-2.

Figure 4-2: Digital shadows in POEMs

4.4.2 Digital Lenses

Ishii and I were also interested in the use of "digital lenses" for viewing "optical" information. As one approach, we constructed the metaDESK's "active lens" based upon an arm-mounted optical lens of the jewelers magnifying-glass variety. Our idea was that in replacing the optical lens of the magnifier with a flat-panel computer screen, we could realize a type of "semantic lens" through which the combined virtual and physical contents of the metaDESK might be viewed.

With the Tangible Infoscapes prototype, the active lens was used to display 3D graphical stacks of images associated with hypercard objects. In this case, we thought of the virtual stacks also as a type of digital shadow, but in this case projected into the third dimension through the "optical" transformation of the active lens. A screenshot of such an active lens view is displayed in Figure 4-3, where an image of the active lens in action within Tangible Geospace is shown in Figure 4-4.

Another exploration of the lens notion was the metaDESK's "passive lens." Here, we used a passive piece of transparent optical fiber cluster material for our display, back-illuminating this with the metaDESK surface. The passive lens then serves as a surface-based digital lens realizing a kind of physical embodiment of the "magic lens" interface metaphor. [Bier 1] One such use is pictured in Figure 4-5.

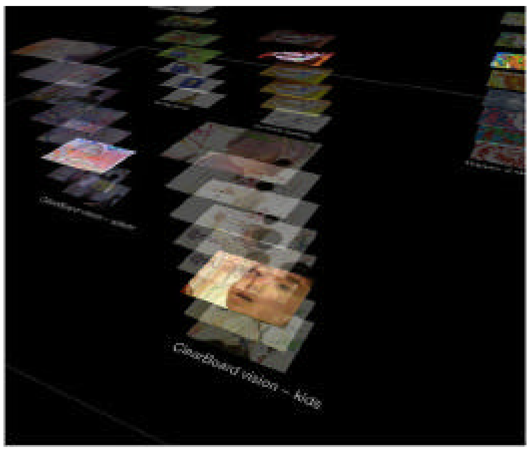

Figure 4-3: Active lens view, Tangible Information Landscapes

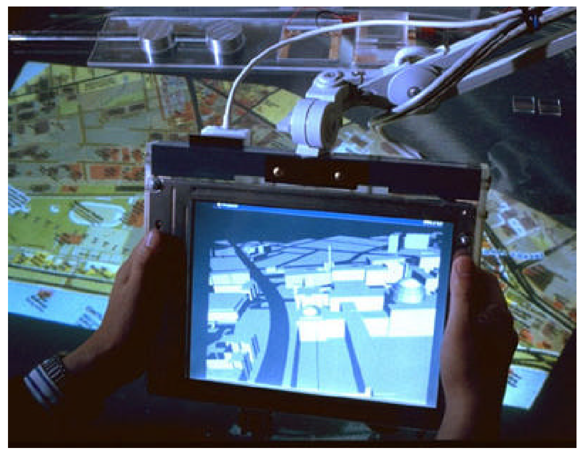

Figure 4-4: Active lens in Tangible Geospace

Figure 4-4: Passive optical prop for active lens

Figure 4-5: Passive lens, with Great Dome phicon

4.4.3 Digital Light

Reflecting on the notion of digital shadows in Tangible Geospace, we were struck by considerations of where sources of semantic, virtual, or “digital light” might originate. So inspired, we explored the use of an instrumented physical flashlight which projects adjustable “wavelengths” of digital light into the Tangible Geospace scene.

In Tangible Geospace, this digital light might cast digital shadows of various physical and virtual objects. For instance, the 3D building geometries of the MIT campus could be imagined not only to render "optically-constrained" shadows of buildings’ physical forms, but also to generate shadows as a function of non-physical parameters of the geographical landscape, say, as a function of the research publications or sponsorship inflows of various buildings. As another example, we can think of projecting other "optical" qualities into the scene such that the rendering style changes from photorealistic to impressionistic.

4.4.4 Projective Lights

A last variant of the optical metaphor involves extending optical notions to the bulk of physical space, painting all architectural surfaces with digital light. A variant on this approach was pioneered in the early 1980’s by media artist Michael Naimark through a series of works called Displacements. In this work, Naimark filmed architectural spaces along with their physical contents and inhabitants using a centrally-positioned rotating camera. Then, he would physically paint the room and objects white, and replace the camera with a rotating projector. Under this “displaced light,” the couch in Naimark’s white room regained its texture, the globe virtually resumed its spinning, and “ghosts” of previous visitors limply slip across physical chairs and walls, blithe to the intervening time and substance.

This illustration of optical projection into and onto physical space, while first demonstrated with analog recording media, has great potential for digitally “reinterpreting” the contents of physical space. Working developments of this approach have been demonstrated by the DigitalDesk of [Wellner 1992] and the Luminous Room of [Underkoffler 1997], and are introduced with respect to TUIs in [Ullmer 1996b].

This approach is attractive given its simultaneous support for digital shadows, passive lens and digital flashlight-style devices, ambient lights, and other digital projections onto realspace objects and surfaces. Issues of occlusion, addressing, and competition with natural light and uninstrumented lighting fixtures remain challenges, and back-projected approaches to devices like the passive lens may always yield more flexibility in the confines of suitably-instrumented surfaces. However, projective lighting of realspace has many yet-unexplored potentials compelling for continuing study and development.

4.4.5 Discussion

The optical metaphor leverages a powerful physical-world metaphor possessing both a kind of natural legibility, as well as a legacy of millennia of instrument and apparatus design which has evolved well-afforded modes and metaphors of physical interaction. Perhaps the most compelling aspect of the optical metaphor is its seamless consistency with the physics of real space. By not only invoking but also obeying the optical constraints metaphorically imposed on our physical interface prototypes, we are able to maximize the *legibility of interface* of our creations. People know what to expect of a flashlight, know what to expect of lenses. By satisfying these expectations, we can realize truly seamless, "invisible" integration of our technologies with the physical environment.

Finally, the optical metaphor is a strong source of inspiration for future work. Ideas about the interaction of digital light with physical space opens many new possibilities for literal and metaphorical uses of mirrors, prisms, virtual and physical transparency and opacity, and light of different spectrums and powers of penetration in the context of both foreground and background information display.

4.5 Containers and Conduits

A last TUI model to be considered is the “containers and conduits” metaphor. In the first uses of hypercards with the transBOARD, Ishii and I used physical hypercards to virtually store the digital contents of a whiteboard session. Additionally, we demonstrated the use of hypercards as physical gateways for the transmission of live whiteboard activity. The prototype of the ambientROOM bottle provided an even more evocative physical instantiation of the containment metaphor. These design instances gave rise to the idea that TUI tangibles might act as physical *containers* for digital information, as well as physical *conduits* for directing the flow of streaming content. This section explores the physical embodiments and interface roles of the “containers and conduits” metaphor as motivated by thesis prototypes.

4.5.1 Containers

The container metaphor considers physical objects as having the capacity to “contain” digital content. While most clearly demonstrated by the transBOARD hypercard and ambientROOM bottle, the container metaphor is also more loosely invoked by the tangibles of POEMs and the phicons of Tangible Geospace. Evocative external instances of the container metaphor include the marble answering machine of Bishop [Crampton Smith 1995], the voiceholder rocks of Laurel, Strickland, et al.'s Placeholder virtually reality piece [Laurel 1994], the voiceboxes of Jeremijenko, and the badges and buckets work by Borovoy, McDonald, et al. [Borovoy 1996]

Realization of the container metaphor has many different facets. First is the issue of how digital content may be stored and retrieved. One approach is by attaching/detaching the container to a content source/sink, with the source/sink embodied either as a surface, object, or instrument tangible. The duration of storage/retrieval activity may be expressed by the duration of contact with the content source/sink. This model is loosely illustrated with the addition and removal of phicons to the metaDESK in Tangible Geospace. Alternatively, the container can be touched to the content source/sink to activate storage/retrieval, then later touched again to toggle deactivation. The barcode-wanded hypercards of the transBOARD loosely approximate this approach. Lastly, some containers like the ambientROOM's bottle can be physically opened and closed, providing a natural metaphor for the container's readiness for storage/retrieval, without intrinsically suggesting what is to be stored or retrieved.

Another container access approach is the insertion or extraction of some physical substance which proxies for digital content. Containable substances include both discrete solids, like the marbles of Bishop's marble answering machine (though the marbles might also themselves be considered as containers), or more continuous liquid, power, or gaseous forms. Containment can also be activated by the triggering of some associated event, as with the container-tangibles of POEMs. Lastly, containment might be activated by the indirect application of a TUI instrument, e.g. aiming a content-bearing flashlight-style instrument onto a container tangible.

Another issue involves how contained content is to be monitored and observed. Contained content may be passively observed through sight and sound – e.g., through viewing the "digital shadow" of a container tangible, or listening for sounds of contained content, perhaps proxied as the sounds of a bubbling fluid. Container contents may also be actively observed through touching, shaking, or application of a mediating instrument. For example, a design mock-up was built by Ross at the RCA, where shaking a physical object (e.g., hollow book) allowed invisible attributes (quantity/volume/etc.) to be experienced as rattling pebbles. [Ross 1995] Ishii and I have explored related interactions for TUI containers, where shaking a container allows "contained" digital information to be haptically and audibly observed.

As another dimension, TUI containers may passively contain digital content, or actively interact with contents and their surrounding environment. In the case of functionally passive containers, the entrance and exit of content is explicitly expressed by the user, and digital contents entering the container are faithfully reproduced upon exit. Functionally active containers may express more interesting roles. For instance, content entrances and exits need not necessarily be explicit; e.g., “leaky” or “hungry” containers may dispense or acquire content spontaneously or in association with some external event. Also, containers may actively transform contents, either applied individually to digital contents (inverting color, etc.), or through aggregate “chemistry” of multiple contained digital substances.

Finally, TUI containers offer other dimensions for exploration. TUI containers may have a unified entrance/exit, as with the ambientROOM’s bottle, or have multiple distant entrances and/or exits (e.g., in fashions perhaps analogous to the many outlets of bagpipes). Containers may have a single physical instance, or may also exist as virtually-coupled remote objects, reminiscent of the multi-location objects of Brave and Dahley [Brave 1997]. Many other design dimensions can also be imagined, indicative of the richness of the physical container metaphor.

4.5.2 Conduits

Conduits act upon flows of streaming digital content. They may be used as portals for streaming transport between remote content sources and sinks, or may be used to express the functional transformation of streaming content. The association of conduit sources and sinks has similar range of dimensions to the container metaphor, with the distinction that conduits simultaneously both absorb and emit content. These associations may again be established with transient or continuous contact with information sources and sinks. They may also be expressed by either discrete or continuous flows of proxying physical substances, as with water through a pipe, or dominoes and marbles through various paths of conveyance. Finally, again as with containers, conduits may passively conduct flows of digital information, or actively transform these flows in a rich variety of dataflow and filter-like fashions.

The development of conduit behaviors was limited in the research contributing to this thesis document, but is the focus of active effort in continuing research.

4.6 Summary

This chapter has presented models and metaphors by which physical-world tangibles may be used as interfaces to digital-world content. The chapter began with discussion of object tangibles, instrument tangibles, interactive surfaces, and interactive spaces, in the process both teasing apart distinguishing characteristics of TUI elements, as well as exploring the relations and interdependencies between these devices.

Three particular metaphors for TUIs were then introduced. First, the “GUI widgetry” metaphor developed a repertoire of metaphorical parallels between the virtual widgetry of GUIs and the physical tangibles of TUIs, thus providing a bridge from the well-explored regime of GUI into the new territory of tangible user interface. Secondly, the "optical" metaphor of digital light, shadow, and "optical" intermediaries taps a powerful physical-world metaphor possessing both a kind of natural legibility, as well as a rich legacy of physical instruments which provide useful interface models for TUI design. Lastly, the "containers and conduits" metaphor provides a promising avenue towards new forms of tangibles which borrow from a familiar metaphor of containment and conduction in the abstract, but also open the door to evocative new physical forms specific to digital content.

5 TUI Mechanisms

The previous chapter discussed a range of models by which tangible user interfaces can be designed. Significantly, the word "computer" seldom arose in this discussion, and in none of the interface instances was it straightforward to point "here is the computer" and "here are the peripherals." In some fashion, the models chapter implicitly considered the capacity for computationally-mediated user interface to be *independent* from the technological instrumentations of individual tangibles. As one articulation of this, the thesis interface models invoke the following assertion:

Hypothesis: *The apparent capabilities of a tangible to sense, maintain, evolve, communicate, and display state can be considered independently from the tangible's intrinsic capabilities, provided the presence of proxying capabilities in the surrounding environment.*

This assertion is the basis of an architecture called proxy-distributed or "proxdist" computation, which will be the primary subject of this chapter. The goal of proxdist computing is to allow tangibles to be addressed in terms of apparent sensing, display, communication, and computation "capabilities," regardless to which of these are intrinsic to the thing in question. This goal is realized by creating digital-world proxies which operate on behalf of each physical-world tangible, hosted across a pool of distributed processing elements present within the surrounding environment. Sensing, display, computation, and communications capabilities are proxied on behalf of tangibles, abstracting away whether individual capabilities are local to the tangible or derived from the capabilities of neighboring tangibles, user-borne technologies, or the surrounding environment.

The mechanisms chapter plays two important roles, addressing issues of both the technical process and underlying concept of tangible user interfaces. First, proxdist computation provides a strong design framework for building complex TUI applications without necessarily considering which underlying devices contribute to a tangible's augmentation, or even whether the individual tangible is highly instrumented or completely passive. Secondly and equally important, proxdist computation provides a powerful "tool for thought" in exploring new tangible user interfaces. By changing the assumptions we can make about the computational and behavioral capacities of physical things, proxdist computing allows us both to reconsider new roles for pre-existing objects and instruments, as well as design entirely new generations of physical interfaces leveraging these capabilities.

The chapter begins with a motivation of proxdist computation drawing from the field of computer networking, continuing with discussion of related computer networking concepts. The basic model of proxdist computation is then presented. This basic model is extended by the introduction of the proxdist flow diagram. The chapter concludes with discussion of the contributions and limits of the proxdist computation approach.

5.1 Motivation

The proxdist computation approach has been strongly motivated by techniques from the field of computer networking. In part, this reflects my earlier work with distributed information interfaces and architectures discussed in [Ullmer 1995], [Ullmer 1994], and [Ullmer 1993], among other places, which developed both technical and metaphorical motivation for the proxdist approach. Additionally, this is in response to technical and metaphorical similarities between the challenges of computer internetworking and tangible user interface.

The modern field of computer networking is marked by the engineering and conceptual ramifications of approaches framed in layered abstractions. By way of example, the Ethernet-based local-area network (LAN) was first developed at Xerox PARC in 1974. While the core technology underlying Ethernet (as one particularly successful networking instance) has survived the intervening quarter-century remarkably intact, a great distance has been spanned in reaching towards the late 1990's Internet and World Wide Web.

First, this distance is marked by an engineering issue of process. The modern Internet forms a vast web of perhaps 50 million computers spanning an enormous range of processor types, local-area network technologies, and long-haul communications infrastructures literally circumscribing the planet. Without the repertoires of layered abstractions for teasing apart this complexity delivered by TCP/IP and accompanying protocol suites, the existence of the Internet as it is known today would not be possible. The gulf of complexity from the simple packet passing of the Ethernet LAN to the convoluted information pathways of the modern Internet is simply too large to be resolved by brute force.

Secondly, the distance from early Ethernet to modern Internet involves a metaphorical issue of concept. While Bush's vision of Memex is now more than half-century old [Bush 1945], the technical genesis of the modern World Wide Web stemmed from the enormous repertoire of layered abstractions embodied by the Internet. The ability to abstractly invoke a connection-based "socket" between any two named devices on the planet, regardless of the wends and whiles of underlying address resolution, packet routing, session management, and physical transport, is an immensely powerful notion. The reality of a system in which it is both conceptually and technically as simple to read a text file or image from an obsolete mainframe on the other side of the planet as from a mail-order PC a few offices away is a conceptual tool of the highest order, directly giving rise to the birth of the modern Web.

Comparable metaphors of process and concept apply directly to tangible user interfaces. From a process standpoint, proxdist computation provides a repertoire of layered abstractions for coming to terms with implementational complexity. Even relatively simple applications like Tangible Geospace provide formidable implementational challenges. Tangible Geospace required the integration of up to a half-dozen computers, each hosting different sensor and display capabilities; each running different operating systems; each resource-bound by different graphics, computation, and network performance constraints; and all to be seamlessly integrated within the interactive space enveloping the metaDESK. The implementational realities of this task left early design efforts drowning in complexity. The computing approaches developed in this chapter have contributed greatly to making subsequent implementational iterations a manageable task, as well as opening the door to the rational engineering of even more interesting and complex interface challenges.

Similarly, much as the mechanisms of the Internet made possible new interface domains such as the World Wide Web, proxdist mechanisms aspire to providing new eyeglasses for viewing the world of possible physical-world interface designs. Given the presence of certain critical-mass concentrations of environmental computation, sensing, and display capabilities – assumptions borne out by the metaDESK and ambientROOM – proxdist computation provides interface designers the tools to work not in terms of which tangibles are outfitted with which processors, sensors, and displays, but rather in terms of what physical artifacts and interaction metaphors are most meaningful disjoint from the mediating limitations of individual tangibles. Thus, the hope for this chapter is to present a computational framework which is not only technically useful, but also conceptually supportive of yet-unimagined models for physical-world user interface.

5.2 Mechanisms from computer networking

The notions of proxdist computing find inspiration in several technologies from the fields of computer networking and internetworking. Among others, these include the OSI reference model for computer networking; the URN (universal resource name) Internet resource addressing proposal; and the BOOTP protocol for diskless workstations.

5.2.1 OSI reference model

The rash of publicity received by the Internet in recent years has made computer networking a concept of household familiarity. The concept of computer networking extends back nearly half a century, addressing the basic need for providing communications capabilities between separate computers. During this interim, a number of technologies for computer networking have come and gone, ranging from early coaxial and twinaxial copper-based technologies, to the more recent deployment of optical fiber and wireless RF approaches.

As underlying networking technologies come and go, it has been important for software protocols to maintain insulation between the low-level physical implementation of data transfer and high-level abstractions of distributed inter-application communication. Perhaps the most well-known model for this successive layering of abstraction is the OSI reference model for computer networking. The OSI model posits a seven-layer stack of abstractions for digital communications, ranging from the lower-level physical data layer to the highest-level application layer, illustrated in Figure 5-1a. TCP/IP, the pervasive protocol underlying the Internet, employs a related layered protocol approach, though the lowest "layer" of its reference model folds together several distinct networking tasks (Figure 5-1b). Tanenbaum also presents a hybrid or rationalized reference model as Figure 5-1c, combining the respective contributions of the OSI and TCP/IP models. [Tanenbaum 1996]

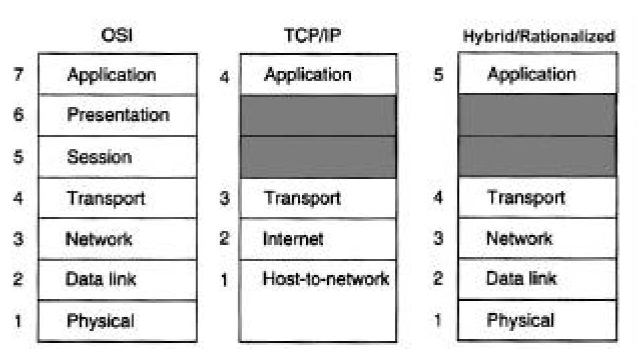

Figure 5-1 a,b,c: OSI, TCP/IP, and hybrid reference model network stacks *(from Tanenbaum 1996, modified)*

While the OSI reference model and its rationalized variants are specific to computer communications, this approach for providing layered abstraction from the details of physical-world communication is directly relevant to TUI mechanisms design. In the TUI case, there is not one but four distinct processes to be abstracted from physical-world implementation: that of sensing, display, computation, and communication. Our approach in managing these TUI processes is inspired by and closely parallels the OSI/rationalized reference model.

5.2.2 URNs

Another core issue in computer networking is that of naming. Clearly, naming spans a far greater set of issues than of computer networking alone, issues explored within [Carroll 1985] among many others. However, names hold special import in computer networking as well as proxdist implementations. In the context of a computer network linked to the Internet, for instance, a name can be thought of as having "unbounded" capacity for storage and computation, limited only by the aggregate distributed resources to which it resolves. For example, the uniform resource locators or "URLs" made ubiquitous by the World Wide Web each may map to a paragraph of text or a terabyte of astronomical data, in the process invoking a computational task ranging from a simple data retrieval to an enormous processing task on the Human Genome Project database.

While URLs have become household identifiers, they are only one instance of a broader concept known as the uniform resource identifier or URI. URLs are fairly brittle constructions, encoding the protocol, computer host name, port, and path of a retrievable information resource. With the constant turnover of Internet hosts and Web site contents, the median lifetime of the current-day URL has been estimated at less than 100 days.

Under the URI scheme, two other addressing and reference schemas have been proposed. One of these is the uniform resource characteristic or URC, a structure used to bundle descriptive meta-information about online resources. The other approach is the uniform resource name or URN. The idea of the URN is to leverage the existence of standardized reference schema such as the ISBN (International Standard Book Number) and UPC (Universal Product Code) codes for use in online resource identifiers which remain independent from protocol, host, port, or path dependencies. While URNs are ultimately resolved into URLs for network retrieval, this resolution can be arbitrated by a trusted name registry service, gaining in the process a measure of abstraction, reliability, and permanence of reference.

This notion of URNs applies directly to proxdist mechanism design. Where Internet URNs can be considered as abstract handles for distributed storage and computation capabilities, URN-style proxdist names are used as abstract handles to distributed sensing, display, communication, and computation "capabilities," limited only by the grounding capabilities of the surrounding physical and network environment. Such names will be returned to later in the chapter and within the discussion of section 6.3.3.

5.2.3 BOOTP

BOOTP addresses the challenge of how a diskless machine may identify itself and retrieve operating instructions for execution. Following the BOOTP protocol, a newly-powered machine broadcasts its hardware (e.g., Ethernet) address to all listeners on the LAN with a message effectively saying "Who am I?" The BOOTP server replies with the diskless workstation's IP address, the BOOTP server's address, and the name of a boot file associated with the diskless client. In the second phase of the BOOTP transaction, the client retrieves and executes the operating system boot image from the BOOTP server, in a sense answering the question "What do I want to do?" [Croft 1985]

The BOOTP approach is interesting to consider in the context of a computationally-augmented physical environment, where tangibles are substituted in place of diskless workstations. Here, the interest is for each physical-world tangible to bind with its digital-world proxy. Some tangibles may possess intrinsic computation and communication capabilities, and thus may be able to initiate and complete a proxy-binding dialog in a manner similar to the diskless workstations of BOOTP. On the other hand, many tangibles may not possess local computation or an operable communications link. Here, upon detection of some physical-world distinction indicating the presence of a new object, we can imagine the augmenting environment both requesting, resolving, and hosting the identity and proxy on behalf of the passive physical object.

This BOOTP analogy for proxdist computation can be considered as something of a technical metaphor, or it can be taken fairly literally. As will be considered later in the discussion, the BOOTP analogy is complicated by the fact that tangibles generally have less easily resolvable identities that the hardcoded 48-bit address of an Ethernet card. This concern is partially addressed by the behavioral approach to proxy-binding introduced in [Ullmer 1996a], but remains for future work to resolve in practice.

5.3 Proxdist computation

As previously introduced, the goal of proxdist computing is to allow tangibles to be addressed in terms of native sensing, display, computation, and communication "capabilities," regardless of which of these are in fact local to the tangible in question. The model is to parallel physical-world tangibles with digital-world proxies whose methods reflect the attributed sensing and display properties. Then, these proxied methods are progressively mapped through sets of layered abstractions to actual sensing and display capabilities distributed throughout the physical environment.

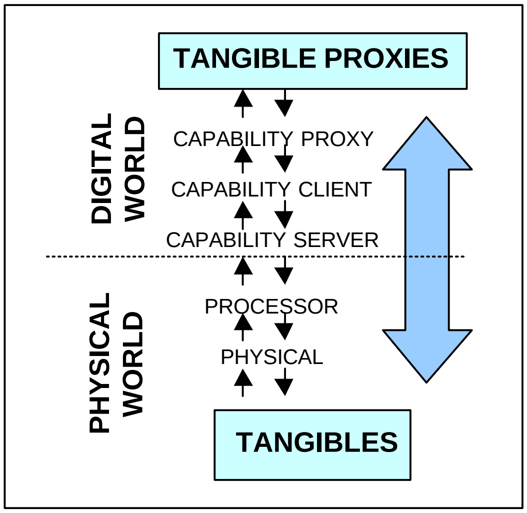

Figure 5-2: Layered Proxdist Computation Approach

A visualization of this approach is presented in Figure 5-2. In the abstract, each physical-world tangible is shadowed by a digital-world tangible proxy. Implementationally, each tangible and its proxy are separated by five layers of abstraction – or more precisely, two such layered stacks, one for input from the physical world, and one for output into the physical-world. These stacks include the physical, processor, capability server, capability client, and capability proxy layers.

The lowest layer of the proxdist stack is the *physical layer*, paralleling the physical layer of the OSI networking reference model. This layer includes the physical sensor and display hardware used to monitor and alter tangibles' physical state. Each sensor and display represents a physical *capability* of the proxdist architecture. These capabilities may map directly to the state of an individual tangible, or they may contribute to the state of a number of different tangibles.

The second layer is the *processor layer*, which includes the computer elements which acquire and express sensor/display state, as well as hosting the computation and communication of upper proxdist stack layers. These computer elements may be embodied as microcontroller elements embedded within tangibles or the surrounding physical environment; as processors worn by human users; or as remote processors linked to sensors and displays in wired or wireless fashions. Also, the computer devices of the processor layer are assumed to have the ability to intercommunicate with peer processors (e.g., form part of a local-area network), even though the bandwidth and latency of these connections may vary dramatically from processor to processor.

The third layer of the stack is the *capability server layer*. The capability servers of this layer serve to acquire and express the state of physical capabilities (i.e., sensors and displays) local to their host computer, and to coordinate these activities with multiple connecting network clients. Some capability servers consume substantial local computation. For instance, the capability servers for thesis applications of computer vision consume on the order of 100 MIPS of local processor computation. Other capability servers act more as caches and network relays for the state of local physical capabilities.

Fourth on the proxdist stack is the *capability client layer*. Capability clients communicate with capability servers to provide networked access to physical capability state. Capability clients locally cache state for immediate access, and in some cases interpolate or (potentially) predictively filter capability state to provide high responsiveness in the face of variable network latency and bandwidth. Some capability clients are engineered to support different communications models for actively accommodating different network loads and throughputs. Most generally, though, capability clients allow the uppermost proxdist stack layers to access capability state independently of the physical capability's locality.

Uppermost on the proxdist stack is the *capability proxy layer*. Capability proxies map the sensor- and display-centric state of physical capabilities into the derived implicit state of individual tangibles. While many tangibles require the combination of multiple sensor and display technologies to derive accurate state (e.g., combined vision-based tracking and LEGO tag-based identification), capability proxies provide each tangible with a clean set of I/O access methods. In this fashion, all the complexities of lower-layer proxdist levels are rendered transparent to the code invoking the access methods of tangible proxies.

A practical example of proxdist computation for the case of the metaDESK's passive lens device is presented in section 6.3.3, providing a concrete illustration of these abstractions.

5.4 Proxdist flow

The last section introduced the notion of the layered architecture of proxdist computation, visually summarized in Figure 5-2. This section introduces a more precise visualization of proxdist architectures called the proxdist flow diagram, and uses this representation to more carefully characterize the operation of existing and speculated tangible user interfaces.

A proxdist flow diagram representing the Tangible Geospace application on the metaDESK system is presented as Figure 5-3. As is apparent, a number of alterations have been made to the earlier illustration of Figure 5-2. First, we have separated the single "capability" stack of Figure 5-2 into independent sensor and display stacks. Secondly, we have "rotated" the resulting cyclical diagram such that the physical layers of the proxdist flow are contained within the stack of the left, while the digital layers are contained in a rightwardly stack.

Thirdly, we have divided each stack into a series of subunits, each with a pair of inlet and outlet nodes. These subunits represent individual tangibles, sensors, clients, proxies, etc. The links between these nodes represent dependencies upon nodes of adjacent proxdist stack layers. An exception is made in Figure 5.3 at the processor layer to better illustrate distributed computation; separate computers here are represented as partitions of the processor layer.

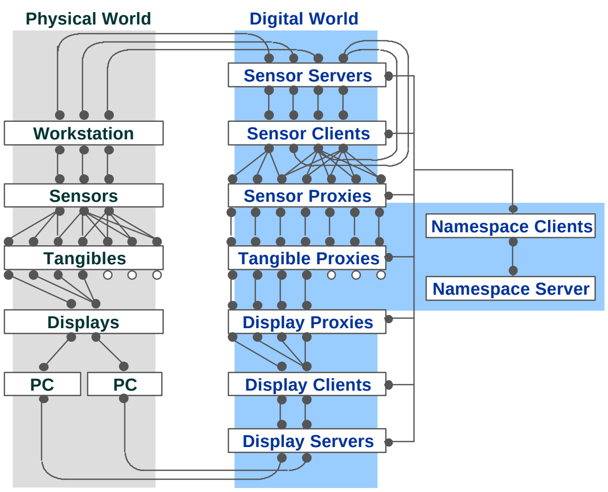

Figure 5-3: Proxdist flow diagram for Tangible Geospace on metaDESK

5.4.1 Proxdist flow and complexity

There are several observations we can make of the proxdist flow diagram. First, the diagram clearly illustrates where complexity is to be found within the proxdist flow, as well as where complexity is alleviated. In the center of Figure 5-3, we find the physical-world tangibles paralleled by digital-world tangible proxies. In the physical world, there is a high degree of complexity surrounding the sensing and augmentation of tangibles. Sensors like the camera input of computer vision are affected by multiple physical targets whose state must be "deconvolved" from a composite signal to be used. Similarly, displays like the metaDESK's interactive surface must synthesize composite output simultaneously mapping to multiple independent physical-world tangibles (phicons, passive lens, etc.). In short, the interface between the unaugmented physical world and augmenting sensor/display capabilities is rife with implementational complexity.

While the interdependencies between physical-world tangibles and sensor/display layers are visibly complex, the comparable dependencies between tangible proxies and the supporting sensor/display proxy layers map very cleanly. Each tangible proxy is individually supported by a corresponding sensor and display proxy. Tangible proxies are insulated from both the physical and network localities of underlying capabilities and computers, as well as from the physical capabilities which contribute to the operation of a given sensor/display proxy. In short, in the midst of a chain of dependencies with substantial local and global complexity, tangible proxies deliver a clean world-model to the interface designer.

5.4.2 Mechanically-derived visualization of system state

It is interesting to note that the proxdist flow diagram is mechanically derivable as a simple visual summarization of underlying software and hardware dependencies. Our proxdist implementations have developing support for live synthesis of proxdist flow diagrams. This will allow live inspection of physical and digital capabilities incorporated within the system; visualization of points of failure and disconnect; and graphical representation of system bandwidth, latency, and other performance metrics. Additionally, this live visualization may provide a useful tool for experimentally manipulating system state, with the assertion of additional proxies for phantom tangibles, manual constraint of communication bandwidth and sensor update rates, and so forth.

5.4.3 Analysis of computation and communication latency and bandwidth

The proxdist flow representation also helps us begin to consider analytic interpretations of the operations represented by individual proxdist layer nodes and interconnects. The proxdist flow diagram can be decomposed into the inlet/outlet pairs and connecting links associated with the nodes of each stack layer (individual sensors, servers, etc.). The inlet/outlet pairs are associated with the internal computation executed by each node, while links represent inter-node communication pathways.

Each of these elements of computation and communication is characterized by certain latencies and bandwidths. The communications latency is straightforwardly manifested as propagation delay in passing data between adjacent layers. Similarly, communication bandwidth is reflected as the time-capacity of a communications channel. With the proxdist stack representing a kind of distributed dataflow system, computational latency can be seen as the propagation delay between which a layer’s inlet stimulus is processed to an outlet response. Finally, computational bandwidth can be considered as the time-capacity of the inlet-to-outlet computational channel; i.e., the volume of inlet data per unit time which can be computational processed.

The analysis of these operational factors is beyond the scope of this thesis. However, their exploration and analysis is a potentially interesting exercise for several reasons. The latency and bandwidth of computation and communication are the factors determining performance, and, ultimately, functional viability of distributed systems. These abilities to assess computational and communications latency and bandwidth support the analysis, optimization, and manipulation of real-world proxdist implementations. Secondly, bandwidth/latency analysis and modeling provides groundwork for analytical reasoning about speculated proxdist systems, even in domains other than the implementation of tangible user interfaces. As one prospective instance, the structuring and coordination of large-scale nanocomputing systems is imagined potentially approachable through variants of these proxdist approaches.

5.5 Discussion

This chapter has introduced the notion of proxy-distributed or “proxdist” computation as a repertoire of engineering mechanisms supporting the construction of tangible user interfaces. The chapter has considered proxdist mechanisms both as issues of process and concept. From a process standpoint, proxdist computation is graphically summarized by Figure 5-2 and Figure 5-3, representing a system of layered abstractions for structuring the implementation of tangible user interfaces. From a conceptual standpoint, proxdist computation is framed in terms of the chapter’s opening hypothesis, asserting the proxy-mediated independence between a tangible’s intrinsic and apparent behavioral capabilities. This hypothesis paves the way for considering even passive physical objects in terms of proxied sensing, display, communication, and computation capabilities.

5.5.1 Proxdist computation and computer networks

In the context of the chapter’s comparative discussion of computer networks and proxdist computation, several observations can be made. Despite the attractive analogies between the layered interfaces of computer networking and proxdist computation, the physical world of TUI is far messier than the digital realm of computer networks. No matter how error-prone the connection or marginal the latency/bandwidth, nodes on (traditional) computer networks are still unambiguously labeled with digital identity; exchange messages of well-defined content and interpretation; and possess mathematically-definable mechanisms for mediating communication and error recovery.

In the case of TUIs, tangibles may not be discernible from their surrounding environment or be resolvable to definite digital identities, especially if tangibles are passive and untagged. The “messages” expressed by human interaction with tangibles also have the potential to be substantially ambiguous, obscured both through the intrinsically noisy signals of physical-world sensing and through human imprecision of effected action. Nonetheless, even if taken as no more than technical metaphor, the abstractions of layered interface models, distributed namespaces, and proxied behaviors delivered by computer networking are productive analogies for proxdist computation.

An additional observation made by Tanenbaum with respect to computer networks may be projected onto our proxdist mechanisms. Speaking of the TCP/IP reference model pictured in Figure 5-1b, Tanenbaum is critical of TCP/IP's confusion of "host-to-network" as a "layer" rather than an "interface," saying "the distinction between an interface and a layer is a crucial one and one should not be sloppy about it." [Tanenbaum 1996] Such criticism may well be applicable to Figure 5-3's construction, as many different kinds of structures – from physical objects to sensors to computers to software abstractions – are joined together. Here, I have tried to consistently position nodes of process, action, and computation within layers, and bridge these layers with paths of communication and interdependency. However, it is likely that some sources of confusion remain unresolved.

What can be said of the proxdist mechanisms in Figure 5-2 and Figure 5-3 is that they demonstrably constitute a workable approach for engineering tangible user interfaces, having been successfully implemented as the underlying infrastructure for the Tangible Geospace prototype (discussed further in section 6.3.3). Still, details of the proxdist computation model may well require modification. For instance, the capability proxy layers may pull together too much functionality within one layer; the capability server and client layers may be better restructured as capability pre-processing, transport, and post-processing layers; and other abstractions may require addition or alteration. Nonetheless, the proxdist mechanisms as presented have been found a valuable approach for the engineering of complex tangible user interfaces.

5.5.2 Proxdist computation and passive tangibles

The layered proxdist model of Figure 5-3 was developed as the technical foundation for the chapter's hypothesis of independence between the apparent and intrinsic behavioral capabilities of tangibles. By pushing implementation-specific computer, sensor/display, and network dependencies beneath the capability proxy layers of the proxdist stack, it is transparent to TUI applications whether tangibles are directly instrumented or proxied by environmental capabilities. Having made these assertions earlier in the chapter, though, it is the role of the discussion to weigh their feasibility.

In support of this assertion is the claim that such independence "works" for the proxdist implementation of Tangible Geospace on the metaDESK. It is "transparent" to Tangible Geospace which tangibles are passively tracked with computer vision, and which actively tracked with Flock of Bird sensors. Similarly, it is (at least potentially) transparent whether Tangible

Geospace's "lens" devices directly integrate active displays, as with the active lens, or are indirectly proxied through environmental capabilities, as with the passive lens.

However, the Tangible Geospace application of the metaDESK is marked by several key characteristics. First, the metaDESK supplies a certain critical mass and density of sensor, display, and computational resources which make possible its operation. Secondly and more importantly, Tangible Geospace is marked by sufficient constraint of the physical world problem-space to render the interface challenge it presents tractable. While the Tangible Geospace/metaDESK combination function reasonably in sensing and augmenting four or five passive objects, these mechanisms would rapidly break down for 10 or 12 passive objects, and would fail completely for 40 or 50 such tangibles. In short, the function of passive objects within Tangible Geospace is framed in a kind of "microworld in which the set of objects, properties, and relations are fixed and limited in an obvious way," [Winograd 1986] with a corresponding retinue of programmer-optimized heuristics to make tasks like computer vision tractable.

As our employment of computer vision within Tangible Geospace might suggest, the strong case of "active behaviors by passive objects" might rapidly grow to resemble the machine understanding problem of unbounded physical-world computer vision, a classic problem of strong artificial intelligence study. To be clear, the task of TUI design (and proxdist computation, as one grounding mechanism) is *not* the construction of machine understanding systems, but rather the construction of physical-world user interfaces. As such, where the developers of the GUI-based Star advised "don't be dogmatic about the Desktop metaphor and direct manipulation," [Johnson 1989] it is impractical and unproductive for us to become dogmatic about the use of passive physical objects within tangible user interfaces.

Rather, by demonstrating that the domain of proxdist TUI support extends even so far as to include passive tangibles, our intention is to illustrate manners by which all kinds of physical-world things can be considered as physical interface to digital information, no matter how diverse their intrinsic capacities for mediation. In point of practice, it remains to the TUI designer to infuse augmenting capabilities within interface tangibles as necessary, and craft microworld environments supporting special interface assumptions where appropriate.

6 Tangible Geospace, revisited

The Tangible Geospace prototype was briefly presented in the prototypes section of the thesis. As the most thoroughly developed implementation of the thesis research, the Tangible Geospace metaDESK prototype will be developed more thoroughly in the context of the TUI models and mechanisms material of previous chapters. In addition, the implementation of Tangible Geospace (and the metaDESK platform upon which it operates) will be discussed in some detail.

6.1 Review of components

As presented earlier, Tangible Geospace is an interface supporting manipulation of graphical views of the MIT campus, driven by the tangible user interface of the metaDESK and supporting tangibles. The earlier scenario for Tangible Geospace involved seven major interface elements:

(a) metaDESK

> The back-projected near-horizontal desk surface of the metaDESK is the central locale of interaction around which other devices of the metaDESK and Tangible Geospace are managed. Internally, it uses back-illuminated infrared computer vision to track passive objects on its surface. Also, it hosts a Flock of Birds transmitter for two six degree of freedom (6DOF) position trackers used on and above the desk's surface.

(b) Phicons

> The Great Dome and Media Lab phicons (physical icons) of Tangible Geospace act both as containers for the MIT campus content, as well as handles and physical constraints for manipulating the campus map. The transparent acrylic phicons are tracked by the desk's infrared vision system.

(c) Tray

> When not in use on the metaDESK's surface, phicons are stored in the clear acrylic tray located at the top of the desk's graphical surface. The tray serves to identify phicons as they enter or depart its compartments (by virtue of attached LEGO Dacta sensors). Tray compartments may also host properties which may be bound to phicons placed within their bounds.

(d) Active lens

The active lens is the arm-mounted flat panel display mounted above the metaDESK. Tracked with a Flock of Birds sensor, it is used to display a three-dimensional view of MIT campus from the virtual viewpoint of the lens' physical position.

(e) Passive lens

The passive lens is the passive fiber-cluster device used to view graphical campus overlays on the surface of the metaDESK. Tracked with a Flock of Birds device, its graphical contents are back-projected through the metaDESK's surface.

(f) Rotation-constraint instrument

Made of two aluminum cylinders mounted on a sliding transparent acrylic axle, the rotation-constraint instrument is used in place of phicons to manipulate the MIT campus geometry. The instrument's intrinsic mechanical constraints prohibit the ambiguous case possible with simultaneous rotations of both phicons. The rotation-constraint instrument is tracked with the desk's infrared vision system.

(g) Digital flashlight

The digital flashlight is used to "project" semi-transparent graphical overlays onto the metaDESK's surface. Tracked with a Flock of Birds sensor, the overlay may be selected from among several layers by turning the flashlight's "spectrum" bezel-dial.

6.2 TUI Models

Chapter 4 introduced four models for structuring tangible user interfaces: the notions of augmented objects, instruments, surfaces, and spaces; the GUI widgetry metaphor; the optical metaphor; and the containers and conduits metaphor. Tangible Geospace makes invocation of each of these models.

First, Tangible Geospace is centered around the interactive surface of the metaDESK which senses the presence and state of graspable tangibles, and graphically augments these tangibles both individually and as an integrated interface (e.g., individually augments the Media Lab and Great Dome phicons, while also augmenting their union as the 2D MIT campus map). The Media Lab and Great Dome phicons serve as object tangibles on the metaDESK, while the active lens, passive lens, rotation-constraint instrument, and digital flashlight all serve as instrument tangibles. In addition, the active and passive lens instruments also integrate interactive surfaces of their own, thus serving a dual capacity. Finally, the three-dimensional volume bounded by the desk surface and flashlight + active lens' operable range forms a kind of interactive space, within which the desk, active lens, and passive lens display surfaces all operate in a spatially-continuous fashion.

Secondly, Tangible Geospace developed as a kind of grounding instance for the GUI widgetry metaphor. The Great Dome and Media Lab objects we think of as physical icons or phicons, the physical instantiation of the GUI icon metaphor. The passive and active lens devices parallel the GUI windows metaphor, while the physical metaDESK surface parallels the metaphorical GUI “desktop” or workspace. Lastly, the desk’s tray begins to serve as a counterpart of the GUI menu device, though this is not actively leveraged in the thesis, while the physical-constraint instrument serves as a first physical instantiation of GUI controls like scales and thumb wheels.

Tangible Geospace’s application of the active lens, passive lens, digital flashlight, and to a lesser extent phicons and desk surface all leverage off of the optical metaphor. The active lens literally was created as the replacement of an arm-mounted magnifier’s optical lens with a graphical augmented surface, and its function within Tangible Geospace was designed to perpetuate this operational metaphor. Similarly, the passive lens and digital flashlight both leverage off the physical form and metaphorical function of their optical magnifying glass and flashlight counterparts. Tangible Geospace’s usage of phicons and the desk surface make weaker invocation of the optical metaphor, but to a degree work to perpetuate the “digital shadow” notion first explored with Tangible Infoscapes.

Lastly, the “tray” device and phicons of Tangible Geospace partially invoke the container notion from the “containers and conduits” metaphor. The metaDESK’s tray draws its origin from the tray of the Bricks work [Fitzmaurice 1995], and carries over the general metaphor that the tray compartments “contain” digital associations to which phicons “dipped” within are bound. In Tangible Geospace, the tray was primarily used as a sensor for identifying the passive Media lab and Dome phicons, but also both metaphorically and literally contributed to the binding of their digital “contents.”

While the Media Lab and Dome phicons were presented as fixed containers for their respective MIT campus bindings, the metaDESK’s interactive surface was unable to distinguish their identity without the tray’s assistance. The phicons’ bindings were actually resolved by sensing their departure from their respective holding compartments on the tray, and heuristically mapping this identity to the next physical “carrier” sensed on the metaDESK’s surface. While in Tangible Geospace this was primarily an implementational necessity, this binding illustrates digital content mapped from a digital “container” to a physical “carrier.”

6.3 Implementation

The metaDESK hardware architecture is illustrated in Figure 6-1. The largest component is the desk itself, a back-projected near-horizontal graphical surface used to display 2D geographical information within the Tangible Geospace prototype. Above the desk, an arm-mounted flat-panel display serves as the "active lens" used to display 3D geographical information. In addition, several passive physical objects are manipulated by the user on the surface of the desk. Sensing in the system is performed by a computer-vision system inside the desk unit, along with magnetic field position sensors and electrical contact sensors. Three networked computers are used to coordinate the system as a whole.

6.3.1 Display Architecture

The desk back-projection unit is based on the VisionMaker™ product from Input Technologies. It uses a three-tube projector to display computer graphics via two internal mirrors onto a plexiglass diffuser surface, which forms the near-horizontal display surface of the desk. The display resolution is 1280x1024 pixels. While a completely horizontal display surface is desirable for supporting physical objects without slippage, the VisionMaker has a minimum display angle of 12 degrees from the horizontal. We cover the display surface with a clear plastic film to minimize object slippage.

Mounted alongside the desk is the active lens, an arm-mounted flat panel display. The active lens is supported with the arm-mount of a jeweler's magnifying lens. The optical lens was removed and replaced with a specially mounted 25-cm 640x480 pixel LCD TFT color flat panel display. The display-mount has three degrees of freedom (DOF), while the arm support itself has another three DOF. An Ascension Flock of Birds™ 6DOF magnetic-field position sensor is attached to the flat-panel display for tracking its spatial position and orientation.

The desk and active lens displays are driven by two Intel Pentium Pro 200MHz-based computers using Intergraph GLZ6 and Intense3D-T graphics systems, respectively. Both platforms provided hardware texture-accelerated graphics running under the Windows NT operating system. The graphics were driven by 3wish [Ullmer 1997b], which among other things provided a wrapper for the TGS Open Inventor 3D graphics toolkit Intel port.

The active lens is used to display a navigable 3D model of MIT campus buildings (~30K polygons), while the desk displays a 2D color map of MIT campus. The 2D map is displayed as a texture-mapped polygon; while only manipulated as a 2D image, at 1140x500 pixels or 570K square pixels, realtime rotation and scaling require hardware texture acceleration. The active lens and desk displays, along with all system sensors, share a common coordinate system.

The 3D active lens view is from a virtual camera positioned orthogonal to the physical active lens surface. Because it is desirable to move the active lens scene-camera both closer to and further from the MIT campus scene than allowed by the arm-assembly's physical constraints, the distance of the virtual camera from the active lens surface follows an exponential curve with empirically-determined coefficients.

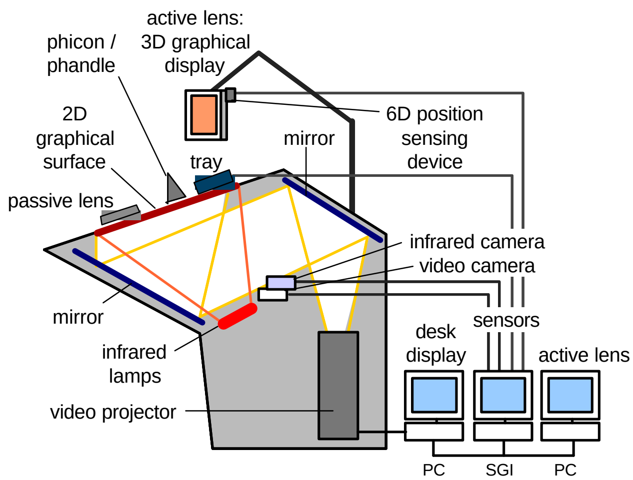

Figure 6-1: metaDESK system hardware architecture

The passive lens is made of a 12cm diameter, 1cm thick circle of fiber-optic cluster material, ringed with a wooden frame and handle. Early iterations of the passive lens were tested both with a plexiglass "lens" and with an empty frame without interior surface. However, the intention with the passive lens was to give the illusion of an independent display surface, i.e. a separate screen. Neither the plexiglass nor empty-frame approaches sustained this illusion.

By raising the image of the back-projected display to the upper surface of the passive lens, the fiber-optic cluster material succeeded in visually simulating an independent display surface. The wooden frame assisted this illusion by masking minor alignment errors between the back-projected passive-lens screen image and the fiber-cluster viewing portal. During movement of the passive lens across the desk, the severity of these errors is a function of the selected tracking technology, network communications, and graphics display frame rate.

Originally computer vision was used to track the passive lens, but later a Flock of Birds sensor was substituted to provide faster, more precise graphics updates. The passive lens display is generated from a second 2D map of MIT campus, texture mapped onto a near-circular polygon on the metaDESK with dynamically-adjusted texture coordinates calculated to fit the position and scale of the phicon-manipulated campus map.

The last hardware "display" of the metaDESK system is the digital flashlight. In our implementation, the flashlight was constructed using a standard battery-driven flashlight modified to include the mounting of a Flock of Birds 6DOF position sensor. Our Flock of Birds system had only two sensors, one of which was attached to the active lens, the other in use by the passive lens. To accommodate the flashlight as a third Flock-tracked device, the passive lens and flashlight devices were altered such that one Flock sensor could be quickly swapped between them. When the flock sensor was registered within a few centimeters of the desk surface, the passive lens was assumed in use, and when at higher altitudes, the flashlight was assumed the active device.

6.3.2 Sensor Architecture

Part of our design aesthetic with the metaDESK was to imagine every object in the physical environment as a potential container or handle for digital information, and thus a potential interface on the metaDESK. Consequently, we were especially interested in using passive minimally tagged physical objects as TUI controls whenever possible. Towards these ends, the Media Lab and Great Dome phicons and the rotation-constraint instruments were designed to be computationally passive (free from active electronics).

Tracking of these passive interface objects is achieved with computer vision. The desk is augmented with two cameras mounted inside its chassis aimed at the back-projected diffuser display surface from underneath. This is illustrated in Figure 6-1. This camera geometry allows objects to be monitored as a largely 2D vision problem free from hand and body object-occlusions. In addition, this approach realizes a modest user-privacy gain in that objects more than ~10cm above the desk surface are invisible to the cameras because of the diffuser coating.

One of the two internal desk cameras is a computer-controlled pan-tilt-zoom Canon VC-C1, used for initial tests at visible-light object tracking and identification. Objects on the surface of the desk were illuminated with pixels from the back-projected desk display, in a sense transforming the display+camera geometry into a flatbed scanner.

This proved unsatisfactory for general object tracking because of interference between graphical output and camera input. To allow physical objects to be tracked, the second camera is used for computer vision in the infrared optical regime. We illuminate objects on the desk's surface with two security-camera infrared LED arrays mounted within the desk, and monitor the resulting scene with a monochrome video camera outfitted with an infrared filter.

This approach cleanly filters the projected computer graphics from camera view because of the uniform infrared component of the projected graphics. In addition, controlled object illumination is generated by the invisible infrared lamps, while external fluorescent room lighting produces a minimum of interfering infrared emissions.

The software of the infrared vision system has two layers: a “tag-track” low-level vision layer, and an upper layer which filters noise artifacts, labels object identities, and tracks objects from frame to frame. The tag-track vision layer was implemented with background-subtraction vision software by Thad Starner of the MIT Media Lab Vision and Modeling group, executed on an SGI Indigo2+Galileo R4400-250MHz at seven frames per second. This software provides for each vision frame a list of unidentified “blobs” extracted from the scene.

Our selection of transparent acrylic Media Lab and Great Dome phicons, while interesting given their visual continuity with the desk display, was challenging because the objects were (not surprisingly) nearly invisible to both infrared and visible-light cameras. We addressed this by backing the phicons with “hot mirrors,” material which is reflective to infrared but transparent to visible-light. This approach allowed objects to be tracked in the infrared while retaining visible transparency.

Objects are alternatively identified with a resistor-tag electrically identifiable by an SGI-based LEGO Dacta Control Lab™ attached to compartments of the tray; or by assigning objects to fixed tray compartments which are monitored for contents with electrical switches or the in-desk vision system. By recording when these (labeled) objects leave the tray and applying this label to the next unlabeled object appearing in the infrared vision scene, we are able to successfully identify and track physical objects on the desk.

6.3.3 Software architecture/mechanisms

The metaDESK demonstrates the use of objects as interface with devices including the active lens, passive lens, and phicons. These TUI elements illustrate diverse realizations of sensing and display capabilities. The active lens “object” contains a real sensor and real display. The passive lens incorporates a real sensor and a “virtual display.” Finally, our phicons in a sense employ “virtual sensors” and virtual displays. This is illustrated in Table 6-1.

As this example implies, the technical realization of TUI elements can become complex. In implementing Tangible Geospace on the metaDESK, we faced the software architecture challenge of cleanly realizing an “objects as interface” implementational metaphor while managing the complexity of a working system.

TUI objects	sensor	display
active lens	real (Flocks of Birds)	real (flat panel LCD)
passive lens	real (Flocks of Birds)	virtual (desk-augmented fiber-cluster surface)
phicons	virtual (computer vision)	virtual (desk-augmented display)

Table 6-1: Real and virtual object sensors/displays

We approached the metaDESK software architecture in two fashions. First, we implemented a platform-independent scripting meta-language, 3wish [Ullmer 1997b], with extended display, sensor, and distributed computing support. Secondly, we implemented a 3wish-based software infrastructure realizing the proxdist computation mechanisms introduced in Chapter 5.

The metaDESK software is implemented with 3wish [Ullmer 1997b], a set of physical sensor and actuator, 3D graphics, and distributed computing libraries built with the “[incr Tcl]” object-oriented extensions to the Tcl scripting language. 3wish currently operates on SGI Irix and Intel-based Windows NT platforms. Cross-platform graphics are implemented using the TGS port of Open Inventor, while Tcl’s shared object loading facilities are used for accessing platform-dependent sensor and computer-vision libraries. 3wish’s Tcl-interpreted graphics implementation was valuable in allowing interface code to be executed transparently across multiple platforms.

Secondly, the proxdist computing approach allows us to provide proxied sensing and display capabilities for Tangible Geospace’s interface objects. This approach is illustrated in Figure 6-2 and Figure 6-3. Tangible Geospace is built around seven interface objects: the active and passive lenses, desk, tray, Dome and Media Lab phicons, and rotation-constraint instrument. While all act as input devices and four serve as displays (the lenses, desk, and tray), these “capabilities” for sensing and display are normally dispersed across the multiple computers driving the metaDESK.

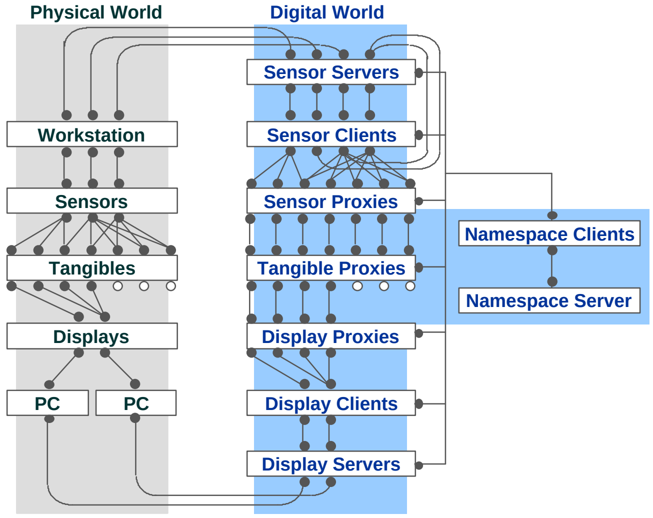

Figure 6-2: Proxdist metaDESK architecture

Figure 6-2 illustrates the proxdist architecture for the metaDESK, as discussed in Chapter 5. To briefly review this diagram, the metaDESK's physical architecture is divided into a series of physical objects, sensors, displays, and computers, each represented as layers on the diagram's left. On the right side of the diagram, the software architecture is shown. The software architecture includes client/server layers for networked communication with each hardware sensor and display technology. In addition, sensor/display proxy layers render transparent to the application interface the physical sensing/display technologies used to realize the virtual sensing/display capabilities of each interface object. Finally, the capability client/server pair is used to manage naming abstraction and to coordinate distributed systems resources.

To provide a concrete example, we will consider the case of the passive lens object on the metaDESK. The passive lens device is tracked with a physical Flock of Birds sensor, and derives its "virtual" display through back-projection by the metaDESK. The physical sensors/displays are addressed by the names

```
tmg:metadesk::sensor:flock
```

and

```
tmg:metadesk::display:desk
```

These names, representing (in the Flock case) zone "Tangible Media Group," encapsulation "metaDESK;" "sensor" capability of type "flock", are associated with [incr Tcl]-based sensor/display API's, and resolved to server host/port TCP/IP addresses through the capability server.

The Tangible Geospace application does not directly reference the above sensor/display clients. Instead, references are made to

```
tmg:metadesk::sensorproxy:plens
```

and

```
tmg:metadesk::displayproxy:plens
```

sensor and display proxies.

The sensor proxy abstracts whether computer vision, Flock of Birds, or some alternate technology is used for passive lens position tracking. In addition, the Flock sensor is physically mounted at a different position than the functional center of the passive lens device. The sensor proxy returns the functional position, again providing insulation from the complexities of physical-world implementation.

Similarly, the display proxy abstracts the display device to which passive lens imagery is rendered. This (ideally) makes it invisible to the application whether the passive lens device is a fully independent display like the active lens, or a device generating its functionality from environmental displays as with the passive lens.

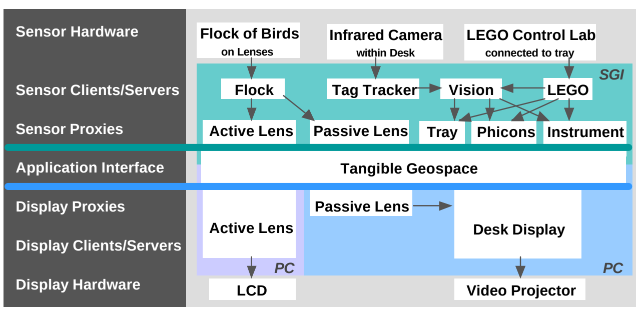

Figure 6-3: Tangible Geospace software architecture

The above-described sensor client/server support has been fully implemented, and simple sensor proxies have been coded for the Tangible Geospace prototype. Sensor clients communicate with servers using a non-blocking “request value/provide value” protocol, with sensors objects caching update values locally.

Sensor clients optionally interpolate server value updates. This is used in the case of computer vision to smooth slow updates, at the cost of slightly increased latency. More sophisticated future clients will hopefully employ dead reckoning or predictive filtering techniques.

The display server/client/proxy layers are more complex than for sensors. Sensors can be thought of as open-loop data sources, streaming sensor readings through servers to sensor clients and proxies. On the other hand, the metaDESK's graphical displays are closed-loop data sinks which input scene geometries and slave updates to the completion of each graphics frame rendering.

Therefore in our implementation, messages to display proxies include 3wish code bound for server-side execution. This is considerably more complex than the sensor proxy case. Instead, the display server/client/proxy hierarchy is currently approximated by executing Tangible Geospace display proxies locally on the desk and active lens display computers. Thus, while the sensor server/client/proxy layers are fully implemented, completion of the display server/client/proxy layers remains to future work.

6.4 Discussion

Tangible Geospace provided the first grounding for the GUI widgetry and optical metaphors, along with the first instance of the physical icon or "phicon" notion. Tangible Geospace produced the passive lens, active lens, and digital flashlight devices and interface concepts. It illustrated issues of interface ambiguity arising within TUIs, along with an example physical constraint device for resolving certain ambiguous interactions. Tangible Geospace also illustrated the use of proxdist mechanisms for coordinating distributed user interfaces, linking distributed sensors and displays, and supporting mixed passive and active devices like the active and passive lens instruments. In short, Tangible Geospace was broadly successful in providing concrete grounding for many ideas underlying tangible user interface.

On the other hand, Tangible Geospace was also marked by several shortcomings. Principal among these, Tangible Geospace was conceived and executed as an exploration of TUI interface concepts, and not in response to any individual real-world problem. As a result, its implementation succeeded in demonstrating numerous concepts and embodiments of physical-world user interface, but fell short of providing a solution to any particular real-world problem.

In part, this mixed success was a consequence of overly ambitious scope. The Tangible Geospace prototype attacked a range of technically demanding tasks, including the construction of many new interface tangibles (active lens, phicons, passive lens, constraint instrument, flashlight, etc.); execution of multiple generations of computer vision software; and construction of new distributed sensor, display, and interface software infrastructures. While each of these efforts was successful, their execution left little time for focusing on evolution of the interface itself. Additionally, the Tangible Geospace prototype was developed while simultaneously attending to many aspects of the transBOARD and ambientROOM efforts, making for a precariously broad distribution of effort.

Beyond these meta-issues, the Tangible Geospace prototype raised numerous specific issues of interest. First, the issue of interface ambiguity, e.g. the simultaneous two-phicon rotation, seems a concern of broad consequence for TUI design. The example of the rotation-constraint instrument illustrates one partial approach, but addresses only part of a much larger challenge. Approaches to TUI ambiguity framed in terms of context, constraint, and closure are introduced in [Ullmer 1997a], but for the most part remain to development by future work.

Another area of concern relates to the choice of relatively literal forms for the Media Lab and Great Dome phicons. This design decision obviously raises issues of scale and clutter. While a GUI desktop can grow cluttered just as readily, the conceptual and literal economics of sweeping into the trash a dozen icons vs. phicons differ substantially. Along related lines, Tangible Geospace's use of phicons for navigating the geospace raises issues of "reference in absence;" how does the user navigate into a space for which no physical embodiment exists? All of these concerns have prompted explorations of more abstract phicons which still retain contextualizing physical affordances. One such area of promising work is the Triangles research of Gorbet and Orth [Gorbet 1997].

Finally, the physical-world persistence of TUI phicons as used in Tangible Geospace raises questions about the utility of platforms like the metaDESK for multiple simultaneous applications. Where complex GUI applications can be saved and restored in mid-stream without loss of "state," TUIs involving more than a few simultaneously-active tangibles face a dilemma when frequent context-shifts are required. While one could imagine decoupling tangibles from their "digital shadows" to alleviate this concern, such approaches might rapidly transform into graspable manipulations of graphical user interfaces, as with the work of Fitzmaurice [Fitzmaurice 1996]. In staying true to the TUI approach, coming to terms with object persistence demands balancing the selection of tangibles between abstract and specific forms, and balancing the expression of application state between physical configurations and digital augmentations.

7 Conclusion

7.1 Summary

In this thesis I have introduced the notion of “tangible user interfaces” or “TUIs” – user interfaces which use physical objects, instruments, surfaces, and spaces as physical interfaces to digital information. I have presented specific physical-world interface models and examples which break free from the image of the computer as a monitor, keyboard, and pointer-endowed terminal. I have also developed a new computational architecture, proxdist computation, supporting the creation of complex physical-world user interfaces.

In particular, I have presented early work on a series of grounding platforms – the metaDESK, transBOARD, and ambientROOM – developed with Ishii and others over the term of thesis research. Special emphasis has been placed on TUI instances, models, and mechanisms on the metaDESK system. Here, with Ishii I have introduced the notion of the physical icon or “phicon,” a physical instantiation of digital information paralleling the GUI icon metaphor. The phicon is further augmented by the idea of the “digital shadow,” the computationally-augmented trace of a physical-world object.

I have presented the active lens, the transformation of a physical-world optical instrument into a computationally-augmented interface, while maintaining the device’s native physical form and metaphorical function. I have developed the passive lens and digital flashlight interfaces, passive physical devices which taken on active behaviors in the presence of an augmenting interactive surface. I have also introduced a physical constraint instrument which illustrates the resolution of certain TUI ambiguities by mechanically constraining the range of possible interactions.

These concepts stand individually, but also combine to form larger models and metaphors for tangible user interface extending beyond the metaDESK. In particular, with Ishii I have introduced the GUI widgetry metaphor, a paralleling of the virtual widgetry of graphical user interfaces with the physical objects and instruments of tangible user interfaces. I have also presented the optical metaphor, a physical metaphor integrating the individual interface concepts of the digital shadow, active and passive lens, and digital flashlight into a broader unifying metaphor, a metaphor which suggests many additional interface possibilities for combining digital light, shadow, and “optical” intermediaries.

At the same time, having co-developed in parallel the transBOARD and ambientROOM platforms, I have discussed additional models for tangible user interface. The transBOARD served as a strong TUI instantiation of Ishii's earlier "interactive surfaces" notion. Also, the application of "hypercards" on the transBOARD provided the first grounding for the "containers and conduits" TUI metaphor. The ambientROOM developed an early TUI instance of augmented spaces, in particular spaces employing ambient environmental displays. The ambientROOM provided grounding for the notion of "ambient media," the use of ambient displays to present information at the periphery of human attention.

Finally, the thesis work has developed and grounded a larger computational framework called "proxy-distributed" or "proxdist" computation, which provides specific mechanisms supporting the thesis TUI platforms and prototypes. Drawing structurally from layered computer networking approaches, proxdist computation provides a strong engineering basis for seamlessly integrating diverse sensor, display, communication, and computation technologies under a unified TUI software framework.

This proxdist computing research serves two significant end effects. First, proxdist computation provides a strong engineering framework for building complex TUI applications without necessarily considering which underlying devices contribute to a tangible's augmentation, or even whether the individual tangible is highly instrumented or completely passive. Secondly and equally important, proxdist computation provides a powerful "tool for thought" in exploring new tangible user interfaces. By changing the assumptions we can make about the computational and behavioral capacities of physical things, proxdist computing allows us both to reconsider new digital interface roles for pre-existing physical objects and instruments, as well as to design entirely new generations of physical interfaces leveraging these augmented capabilities.

7.2 Future Directions

In introducing early explorations of a new research theme in physical-world user interface, the thesis has only begun to touch upon many issues which seem promising for future research. First, the thesis has focused on tangibles which have roots in pre-existing physical forms, exemplified by the active lens, passive lens, and flashlight of Tangible Geospace. While a promising approach, new work with Triangles [Gorbet 1997] and multi-location objects [Brave 1997] demonstrates promising new tangibles without direct physical precedent, physical devices whose existence is framed entirely in terms of computational augmentation. Especially in conjunction with new possibilities raised by the containers and conduits metaphor, I look forward to exploring new generations of tangibles grounded both in novel and pre-existing physical forms.

Another direction for future exploration is the development of tangibles whose function spans multiple augmenting environments. For instance, the augmenting environments demonstrated by the metaDESK and ambientROOM provide quite different mediating capabilities, yet it seems that many tangibles must be able to function meaningfully in many such environments. Proxdist computation provides a technical basis for grounding such efforts. A new research prototype called mediaFlow uses Triangles to channel activity between multiple TUI prototypes using the containers and conduits metaphor, providing a grounding design instance for further exploration.

Finally, while the thesis' simultaneous development of interface models, engineering mechanisms, and multiple design instances has been valuable for broadly surveying new territory in interface design, this breadth makes it difficult to gain full purchase on any one factor of study. In continuing work, I look forward to concentrating efforts either by framing both models and mechanisms in terms of a single design instance, or by developing several design instances illustrating a single interface and engineering approach.

Appendix A: Foreground and background

The thesis introduced three TUI platforms: the metaDESK, transBOARD, and ambientROOM. The metaDESK and ambientROOM, in particular, have a special relationship to each other as originally designed. The metaDESK is a graphically intensive, “hands-on,” and broadly *foreground* interface for physically interacting with digital information. The ambientROOM, in contrast, was originally designed to follow a different interaction model. The ambientROOM’s original conception was an environment where users would concentrate on tasks *other than* direct engagement with the ROOM – say, reading a paper document or conversing on the telephone. The ambientROOM then uses “ambient media” such as ambient light, sound, waterflow, and air movement to make users peripherally aware of changing information and activity. The ROOM thus serves to mediate the *background* display of peripheral information.

The roles presented here for the metaDESK and ambientROOM are clearly different in character. While the metaDESK aspires to strongly engage the user’s senses towards the execution of some foreground task, the ambientROOM serves to complement human senses to provide background awareness of peripheral information.

At the same time, these notions of “foreground” and “background” do not refer to a state of the physical environment, but rather a state of changing human attention. When an eater of breakfast picks a coffee cup from the table, sips reflectively, then replaces the cup to continue with the meal, the cup does not undergo a physical transformation from a “background” to a “foreground” object. Rather, the cup lives in a transitional zone where it seamlessly migrates between these attentional states.

If the user is, say, engrossed in reading the newspaper or devouring an omelet and the coffee is in the anticipated location and state, he may grasp the cup, drink, and replace it without fully consciously engaging in the action. However, if the cup is empty, too hot or sour, drips in the drinking or has been jostled into an unexpected state, it may quickly engage the consumer’s full attention.

Thus, the foreground and background characteristics of the metaDESK and ambientROOM platforms represent not a dichotomy of purpose and function, but points along a continuum between the extremes of total human concentration and obliviousness. Well-affording both foreground and background interface modalities are key tasks of these platforms, but of equal import are provisions for seamlessly transitioning between foreground and background throughout the course of interaction. The design space integrating these interaction modalities is illustrated in Figure 4.

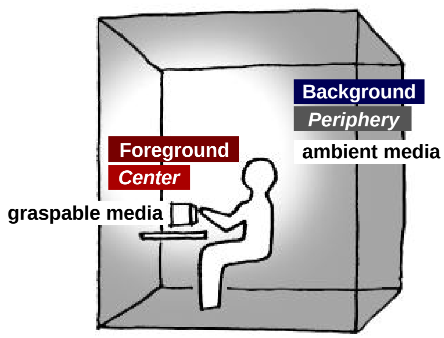

Figure 4: Center and Periphery of User's Attention within Physical Space

Notions of foreground and background are not entirely alien to traditional GUI-based HCI design. However, in each case the GUI represents a somewhat impoverishing influence. The keyboard, pointer, and monitor of the GUI serve functionally as a foreground interface, true; but they miss much exploitation of haptic and tactile interaction and engage vision only in a static center of the visual field. Further, GUIs poorly exploit human spatial senses, not merely in the sense of 2D vs. 3D graphic display, but in leveraging human kinesthesia and knowledge of the surrounding physical environment.

In the background case, GUIs serve more poorly still. Visuals limited to a static small rectangular viewport offer little resource for display in the periphery of the human visual field. GUI's minimal spatial engagement in the physical environment again poorly affords spatial multiplexing of information at the periphery. Finally, peripheral displays with rich roles in the natural environment such as airflow, temperature, vibration, and so forth are lost to the GUI.

Discussion

The foreground/background discussion is not a metaphor for tangible user interface, but rather an identification of dual interface roles that TUIs can serve. TUIs are well positioned to serve both foreground and background interface roles – likely inherently better suited than peer GUI approaches, given the increased range of interaction modalities, media, and configurations spanned by TUIs physical-world environment. Additionally, the potential of TUIs to seamlessly couple both background and foreground interaction modalities provides exciting interface opportunities which have seen limited exploration in the thesis, but serve as promising avenues for future work.

Clearly, the discussion of foreground and background interactions is in need of concrete psychological and human factors grounding to be considered with any level of rigor. Given the breadth of this thesis, such consideration has thus far been outside our scope of study. Nonetheless, the interface notions of foreground and background are important concepts which appear throughout the thesis, and thus merit discussion with at least this limited qualitative treatment.

References

[Bier 1993] Bier, E., Stone, M., Pier, K. et al. Toolglass and Magic Lenses: The See-Through Interface. *Computer Graphics (SIGGRAPH'93 Proceedings),* pp. 73-80.

[Bishop 1996] Bishop, D. Personal communication.

[Borovoy 1996] Borovoy, R. Genuine Object-Oriented Programming. MS Thesis, MIT Media Laboratory.

[Borovoy 1996] Borovoy, R., McDonald, M., Martin, F., and Resnick, M. Things that blink: Computationally augmented name tags. *IBM Systems Journal*, v35 n3-4, 1996. pp. 488-495.

[Brave 1997] Brave, S., and Dahley, A. inTouch: A Medium for Haptic Interpersonal Communication. In *Proc. of CHI'97 Extended Abstracts*, pp. 363-364.

[Bush 1945] Bush, V. As We May Think. In *The Atlantic Monthly,* July 1945.

[Buxton 1995] Buxton, W. Integrating the Periphery and Context: A New Model of Telematics, in *Proceedings of Graphics Interface '95*, 239-246.

[Calvin 1993] Calvin, J., Dickens, A., et al. The SIMNET virtual world architecture. In *Proc. of the IEEE Virtual Reality Annual International Symposium*. Los Alamitos CA: IEEE Computer Society Press. pp. 450-455.

[Carroll 1985] Carroll, J. M. *What's in a Name?* New York: W. H. Freeman & Co., 1985.

[Cooperstock 1995] Cooperstock, J.R., Tanikoshi, K., Beirne, G., Narine, T., Buxton, W. Evolution of a Reactive Environment, *Proceedings of CHI'95*, 170-177.

[Cooperstock 1996] Cooperstock, J.R. Reactive Environments and Augmented Media Spaces. Ph.D. Thesis. University of Toronto, 1996.

[Comer 1991] Comer, D.E. *Internetworking with TCP/IP, v1, 2e.* Englewood Cliffs: Prentice Hall, 1991.

[Croft 1985] Croft, B., and Gilmore, J. *RFC 951: Bootstrap Protocol (BOOTP).* http://www.hin.no/~srk/rfc951.htm

[Crampton Smith 1995] Crampton Smith, G. The Hand That Rocks the Cradle. *I.D.*, May/June 1995, pp. 60-65.

[Czernuszenko 1996] Czernuszenko, Marek. Immersadesk web page. http://www.eecs.uic.edu/-~mczernus/Immersadesk/immersadesk.html

[Dunne 1994] Dunne, A. and Raby, F. Fields and Thresholds. Presentation at the Doors of Perception 2, November 1994, http://www.mediamatic.nl/Doors/Doors2/DunRab/DunRabDoors2-E.html

[Erickson 1990] Erickson, T. Working with Interface Metaphors. In *The Art of Human-Computer Interface Design*, B. Laurel, ed. Reading, MA: Addison-Wesley, 1990.

[Feiner 1993] Feiner, S., MacIntyre, B., and Seligmann, D. Knowledge-based augmented reality. *Comm. of the ACM*, 36(7), July 1993, 52-62.

[Fitzmaurice 1993] Fitzmaurice, G., Situated Information Spaces and Spatially Aware Palmtop Computers. *Comm. of the ACM*, July 1993, Vol. 36, No. 7, 38-49.

[Fitzmaurice 1995] Fitzmaurice, G., Ishii, H., and Buxton, W. Bricks: Laying the Foundations for Graspable User Interfaces. *Proc. of CHI'95,* pp. 442-449.

[Fitzmaurice 1996] Fitzmaurice, G. Graspable User Interfaces. Ph.D. Thesis. University of Toronto, 1996.

[Glos 1996] Glos, J., and Ullmer, B. *POEMs: Physical Objects with Embedded Memories.* Unpublished video.

[Glos 1997] Glos, J., and Cassell, J. Rosebud: Technological Toys for Storytelling. In *Proc. of CHI'97 Extended Abstracts*, pp. 359-360.

[Gorbet 1997] Gorbet, M., and Orth, M. Triangles: Design of A Physical/Digital Construction Kit. To appear in *Proc. of DIS'97.*

[Heritage 1994] *The American Heritage Dictionary, 3e.* New York: Dell Publishing, 1994.

[Heschong 1979] Heschong, L. *Thermal Delight in Architecture.* Cambridge, MA: MIT Press, 1979.

[Hill 1992] Hill, W., Hollan, J., Wroblewski, D., and McCandless, T. Edit Wear and Read Wear. In *Proc. of CHI'92*. pp.3-9.

[Hill 1993] Hill, W., and Hollan, J. History-Enriched Digital Objects. In *Proc. of CFP'93*, http://www.cpsr.org/dox/conferences/cfp93/hill-hollan.html.

[Hinckley 1994] Hinckley, K., Pausch, R., Goble, J., and Kassel, N. Passive Real-World Interface Props for Neurosurgical Visualization. *Proc. of CHI '94*, pp. 452-458.

[Houde 1993] Houde, S., and Salomon, G. Working Towards Rich & Flexible File Representations. *Proc. of INTERCHI'93, Adjunct Proc.*, pp. 9-10.

[Ishii 1994] Ishii, H., Kobayashi, M. and Arita, K. Iterative Design of Seamless Collaboration Media, *Comm. of the ACM*, Vol. 37, No. 8, August 1994, 83-97.

[Ishii 1997] Ishii, H., and Ullmer, B. Tangible Bits: Towards Seamless Interfaces between People, Bits, and Atoms. *Proc. of CHI'97,* pp. 234-241.

[Johnson 1989] Johnson, J., Roberts, T., Verplank, W., et al. The Xerox Star: A Retrospective. *IEEE Computer* 22(9), September 1989, pp. 11-29.

[Kobayashi 1992] Kobayashi, M. and Ishii, H. DispLayers: Multi-Layer Display Technique to Enhance Selective Looking of Overlaid Images. CHI '92 Interactive Poster, ACM, Monterey, May 1992.

[Krueger 1994] Krueger, W., and Frohlich, B. The Responsive Workbench. *IEEE Computer Graphics and Applications*, 14(3), pp. 12-15.

[Lakoff 1980] Lakoff, G., and Johnson, M. *Metaphors We Live By.* Chicago: University of Chicago Press, 1980.

[Laurel 1994] Laurel, B., Strickland, R., and Tow, R. Placeholder: Landscape and Narrative in a Virtual Environment. *ACM Computer Graphics* 28:2, May 1994. pp.118-126.

[Lokuge 1995] Lokuge, I., and Ishizaki, S. GeoSpace: An Interactive Visualization System for Exploring Complex Information Spaces. *Proc. of CHI'95,* pp. 409-414.

[Lynch 1960] Lynch, K. *The Image of the City.* Cambridge, MA: MIT Press, 1960.

[Noma 1996] Noma, H., Miyasato, T., and Kishino, F. A Palmtop Display for Dextrous Manipulation with Haptic Sensation. *Proc of CHI'96*, pp. 126-133.

[Norman 1988] Norman, D. *The Psychology of Everyday Things.* New York: Basic Books.

[Rekimoto 1995] Rekimoto, J., and Nagao, K. The World through the Computer: Computer Augmented Interaction with Real World Environments. *Proc. of UIST'95,* pp. 29-36.

[Resnick 1993] Resnick, M. Behavior Construction Kits. In *Comm. of the ACM.* 36(7), July 1993.

[Ross 1995] Ross, L. Personal communications, 1995.

[Smith 1982] Smith, D., et al. Designing the Star User Interface. *Byte*, April 1982, pp. 242-282.

[Starner 1995] Starner, T., Mann, S., et al. Wearable Computing and Augmented Reality. Vismod Tech Report #355, November 1995.

[Stasior 1993] Stasior, W. Visual Processing for Seamless Interactive Computing. MIT Ph.D. Thesis Proposal, 1993. http://www.tns.lcs.mit.edu/~wstasior/phd-proposal/proposal.html

[Stifelman 1996] Stifelman, L. Augmenting Real-World Objects: A Paper-Based Audio Notebook. *Proc. Of CHI'96 Companion,* pp. 199-200.

[Stafford-Fraser 1996] Stafford-Fraser, , Q., and Robinson, P. BrightBoard: A Video-Augmented Environment. *Proc. of CHI'96,* pp. 134-141.

[Tanenbaum 1996] Tanenbaum, A. *Computer Networks, 3e.* Englewood Cliffs: Prentice Hall, 1996.

[Ullmer 1993] Ullmer, B. End-User Customization Mechanisms for Distributed Information Resources. http://www.media.mit.edu/~ullmer/papers/pgopher/, March 1993.

[Ullmer 1994] Ullmer, B. Multiscale Spatial Architectures for Complex Information Spaces. http://www.media.mit.edu/~ullmer/papers/multiscale/, May 1994.

[Ullmer 1995] Ullmer, B. Towards "Virtual Realty:" Explorations in Spatializing Web Content. http://www.media.mit.edu/~ullmer/papers/urbcyber/, December 1995.

[Ullmer 1996a] Ullmer, B. Behavioral Realizations of Proxy-Distributed Computation. http://www.media.mit.edu/~ullmer/courses/agents/paper1.html, March 1996.

[Ullmer 1996b] Ullmer, B. Physicality, Virtuality, and the Switch that Lights. http://www.media.mit.edu/~ullmer/courses/tat/paper1.html, March 1996.

[Ullmer 1997a] Ullmer, B., and Ishii, H. The metaDESK: Models and Prototypes for Tangible User Interfaces. Submitted to *Proc. of UIST'97.*

[Ullmer 1997b] Ullmer, B. 3wish: Distributed [incr Tcl] Extensions for Physical-World Interfaces. To appear in *Proc. of Tcl/Tk'97.*

[Umaschi 1997] Umaschi, M. Soft Toys with Computer Hearts: Building Personal Storytelling Environments. In *Proc. of CHI'97 Extended Abstracts*, pp. 20-21.

[Underkoffler 1997] Underkoffler, J. "A View from the Luminous Room", *Personal Technologies*. 1(2), June 1997 (to appear).

[Waters 1996] Waters, R., Anderson, D., et al. *Diamond Park and Spline: A Social Virtual Reality System with 3D Animation, Spoken Interaction, and Runtime Modifiability.* MERL TR96-02a. http://www.merl.com/reports/TR96-02a/index.html

[Webster 1997] *Merriam-Webster's Collegiate, Tenth Edition.* Online edition, http://www.eb.com:180/.

[Weiser 1991] Weiser, M. The Computer for the 21st Century. In *Scientific American*, 265(3), pp. 94-104.

[Weiser 1995] Weiser, M. and Brown, J.S., Designing Calm Technology. http://www.ubiq.com/hypertext/weiser/calmtech/calmtech.htm, December 1995.

[Wellner 1993] Wellner, P. Interacting with paper on the Digital Desk. In *Comm. of the ACM.* 36(7), pp. 86-96.

[Winograd 1986] Winograd, T., and Flores, F. *Understanding Computers and Cognition.* Reading, MA: Addison-Wesley, 1986.

[Winograd 1996] Winograd, T. *Bringing Design to Software.* New York: ACM Press, 1996.