Rapid Development, Real-World Deployment, and Evaluation of Projected Augmented Reality Applications

Natan Linder
B.A, Computer Science, The Interdisciplinary Center Herzliya Israel (2003)
S.M, Media Arts and Sciences, Massachusetts Institute of Technology (2011)
Fluid Interfaces Group, MIT Media Lab
linder@media.mit.edu

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
in partial fulfillment of the requirements for the degree
of Doctor of Philosophy at the Massachusetts Institute of Technology

August 2017 [September 2017]

Signature redacted

Author

Natan Linder
Program in Media Arts and Sciences
August __, 2017

Signature redacted

Certified by

Pattie Maes
Professor of Media Arts and Sciences
August __, 2017

Signature redacted

Accepted by

Pattie Maes
Academic Head
Program in Media Arts and Sciences
August __, 2017

Rapid Development, Real-World Deployment, and Evaluation of Projected Augmented Reality Applications

Natan Linder

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning on August __, 2017
in partial fulfillment of the requirements for the degree
of Doctor of Philosophy at the Massachusetts Institute of Technology

August 2017

Abstract

Current interactive projected augmented reality systems are not designed to support rapid development and deployment of applications beyond the confines of research labs. I developed a series of self-contained interactive projector–sensor systems (collectively LuminAR devices) and a web-based software development framework. The design goal of this research work was to advance the state of the art of projected AR interfaces and to explore how they can manifest in day-to-day objects. This novel, tightly integrated approach allows developers who are not versed in computer graphics, vision algorithms, and augmented reality techniques to implement projected AR applications rapidly. In this work, I review several real-world uses of the system for retail presentation, desktop interaction and collaboration applications, manufacturing, and education. The work is evaluated through extensive use of the hardware and software by developers as well as two user studies that specifically explored applications for manufacturing and education. The evaluation methodology focused both on basic interaction and system usability as well as the implications of using augmented interfaces in the specific application domains of education and manufacturing. I also discuss the results of the first large-scale user studies of projected augmented reality rapid application development. Finally, I provide a set of design principles for projected augmented reality applications, and recommendations concerning how to deploy such applications in the real world.

This dissertation work was partially supported by research grants from Intel, Steelcase and Pearson.

Dissertation Supervisor
Pattie Maes
Professor of Media Arts and Sciences
Program in Media Arts and Sciences

Rapid Development, Real-World Deployment, and Evaluation of Projected Augmented Reality Applications

Natan Linder

Dissertation Committee:

Signature redacted

Pattie Maes

Dissertation Supervisor
Professor of Media Technology
Director, Fluid Interfaces Group, MIT Media Lab

Signature redacted

Joseph A. Paradiso

Dissertation Reader
Alexander W. Dreyfoos (1954) Professor in Media Arts and Sciences
Director, Responsive Environments Group, MIT Media Lab

Signature redacted

Steven K. Feiner

Dissertation Reader
Professor of Computer Science
Director, Computer Graphics and User Interfaces Lab, Dept. of Computer Science, Columbia University

Totum factum.

Table of Contents

Acknowledgments

This dissertation is a result of intensive collaboration with numerous colleagues, UROPs, faculty, staff, Media Lab members, and external collaborators. First and foremost, I would like to thank my advisor, Pattie Maes, for her guidance and support throughout the research process. Pattie has let me roam free and break some rules while keeping me focused on doing good work. Her guidance, inspiration, support and patience is the reason this work was started and the reason I was able to complete it. I would also like to thank my dissertation readers Joe Paradiso and Steve Feiner for their critique and thoughtful comments. Joe is the ultimate hacker and Steve is one of the godfathers of applicable augmented reality, I could not imagine a better team to help me put this work together.

I started this work back in 2009 when I joined the Media Lab, but it would have never come so far without my research partner and good friend Rony Kubat. Rony is the best collaborator one can have. I am beyond lucky and extremely grateful to have had his support. Next, special thanks go to the LuminAR team over the years, but especially my close collaborators: Ben Weissmann, Yihui Saw, Yoav Reches, Matt Berlin, Jesse Gray, Craig Broady, Rebecca Kleinberger, Michael Poon, Marcus Boorstin and Mason Glidden.

Special thanks for the members of the Fluid Interfaces group for their continued critique, advice and friendship, In particular I would like to thank Roy Shilkrot, Sajid Sadi, Marcelo Coelho, Niaja Farve and Tal Achituv.

A special thanks to the great support I have gotten from the Media Lab staff and faculty, In particular: Andy Lippman, Hiroshi Ishii, Paula Aguilera, Jonathan Williams, Felice Gardner, Bob Swartz, Steve Otis, and Alexandra Kahn. This is also the point to mention that without the coaching of Linda Peterson and the editing work of Stacie Slotnick this dissertation would not exist. Thank you!

I'd also like to thank the generous support I got from my Media Lab members companies: Steelcase, Intel, Best Buy, Panasonic, Pearson and Qualcomm.

Finally, I would like to thank Emmy. There are not many words that can describe the support I got at home. Emmy did all that while we actually became a family. Romy, Yotam, and Dannie were born while I was working on this research and inspired it heavily. My heart goes to them, and I hope this leaves them with some inspiration how humanizing technology may become.

This work is dedicated to my grandparents: Miriam and Dr. Alexander Guterman, both teachers who taught me how to learn and also how to teach. My very much alive and kicking grandfather Meir Linder who never finished high school, but is one of the best engineers I ever worked with. He taught me endless curiosity and how to love life.

1 Introduction

Augmented reality (AR) interfaces have captured the imagination of generations of human – computer interaction researchers and designers. The principles of physical-world augmentation were laid out by science fiction writers in the early 20th century. It is likely that L. Frank Baum unknowingly catalyzed the ideas in his 1901 novel *The Master Key* **[1]**. In the book, he describes the 'Character Marker' – an electronic spectacles device that allows to the wearer to view personality traits of people as marks upon their foreheads: "while you wear them everyone you meet will be marked upon the forehead with a letter indicating his or her character. The good will bear the letter 'G,' the evil the letter 'E.' The wise will be marked with a 'W' and the foolish with an 'F.' The kind will show a 'K' upon their foreheads and the cruel a letter 'C.'" The protagonist describes his use of devices such as the Character Marker as "being a century ahead of the times." The idea of enhancing the real world with information was set loose. As those ideas proliferated in science fiction literature, a body of groundbreaking research emerged, yielding impressive advances in display, sensing, computer vision, and interaction techniques.
With time, the ideas manifested in the science fiction writing of Baum and others, trickled down to our everyday life as designers and consumers of digital technology and media. In this era, the digital world has become an integral aspect of everything we do, whether work or play. We take for granted the ability to connect to information anywhere and anytime, and have become almost completely technology-dependent. Battery-powered electronic devices with digital interfaces now connect us to vast networks of information. In under 30 years, we have miniaturized our personal computers to a handheld size. We can package enough computation in a matchbox to carry out massive processing that once required a room full of computers. We expect wireless connectivity as we expect electricity.

User-interface (UI) design is central to this perception, creation, and adoption of technology. In the past decade, web-based applications transformed into a cloud of services. At the same time, embedded computing, wireless communication, and novel User Interface technologies matured quickly to a point where we *think* we have the Internet at the tip of our fingers at all times—but do we really? Most of our interactions with the wealth of information and functionality are dominated by screens that constrain the user experience. This fact underlines one of the key themes I am exploring in this thesis—the Digital–Physical Interface Problem.

1.1 The Digital-Physical Interface Problem

This dissertation raises this question and asserts that there is, in fact, a discontinuity between our experiences in the physical world and our interactions with the digital world. Pioneers of ubiquitous computing and augmented reality have described a future where computing and interface components are integrated into the physical environment, creating a seamless experience using all the human senses [2].

In our human-tangible reality, our world interactions are predominantly physical. Our current work environments–lab benches, drafting tables, retail counters–are examples of how we really work: multiple devices, notes, paper, various tools and objects. The practical reality is that the digital world and the physical world are not at all disconnected. A note we write finds its way to an email message, and a measurement we take will end up in a CAD model, often through a laborious, often manual process.

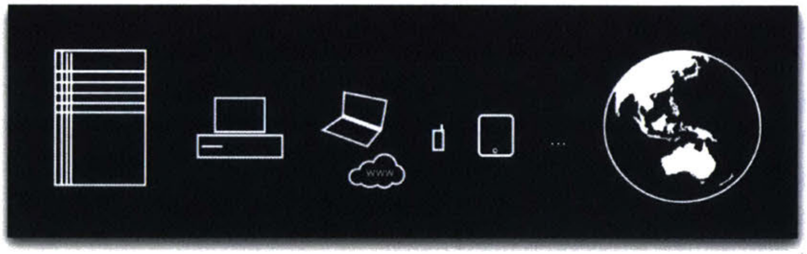


Figure 1-1 The evolution of computer interfaces. From the mainframe to the physical world.

More than 40 years ago computers introduced a display centric user interface paradigm. The user interfaces for the early computers were extremely physical. They included elements such as dials, cranks, switches and patch cords. Along the displays several lights were used to indicate status to the user.

With time, displays took over. The display centric paradigm prevailed as interfaces miniaturized over the years. In the next era of user interfaces, our physical world will become our interface, where objects, surfaces, spaces become interactive, offering relevant information, based on context and interests of user.

As consumers and producers of information using computer interfaces, we as users are responsible for connecting the digital and the physical worlds. We humans did not evolve as quickly as our technology, and fundamentally, we have a very limited cognitive ability to process the vast amounts of information we can now reach. Our brains and sensory systems define a biological I/O bottleneck that we work hard to overcome, mainly because as users we have adapted ourselves to interfaces and not vice versa.

At the time of writing this dissertation work, it is apparent that in the near future, augmented reality (AR) interfaces will become commonplace. We describe the advancement in the field in the chapters that follow. It suffices to say that such interfaces hold the promise to truly amplify our natural physical environment with digital information. AR technologies have been in the center of the human–computer interaction research and as such hold the promise to radically change for the better how we integrate information interfaces into our daily lives. With that, we can focus the discussion on the current state of augmented reality interfaces.

1.2 Augmented Reality Interfaces

Augmented reality (AR) is a relatively young field, though its origins can be mapped to the early works of Sutherland in the 1960s **[3]**, **[4]**, and Myron Krueger's artificial reality and the VIDEOPLACE project of the 1990s **[5]**. Augmented reality, or AR, refers to a set of technologies that allow computer-generated virtual media to overlay physical objects in real time. AR has been the goal of numerous research projects that have tried to create environments to seamlessly blend digital information with the real world. The domain of augmented reality interface research relies heavily on techniques for visual scene reconstruction and understanding, display technologies, and user interaction methods **[6]**.

The origins of Augmented Reality (AR) lie in manufacturing. The actual term was coined in 1992 by Caudell and Mizell **[7]**. Their research was motivated by trying to improve complex aircraft manufacturing processes using Head Mounted Displays. Milgram positioned and classified AR as a subset within the reality-virtuality (RV) continuum **[8]**. He defined mixed reality as a taxonomy encompassing both augmented reality (AR) and virtual reality (VR), juxtaposing fully virtual and fully real environments. Azuma provided a more formal definition of AR as a technology that (1) combines the real and the virtual, (2) is interactive in real time, and (3) is registered in 3D **[9]**. Bimber and Raskar provided an excellent classification **[10]** of the types of AR interfaces, dividing them into (1) head-attached (2) hand-held and (3) spatial.

1.2.1 Hand Held

Currently there are many commercially available see-through AR systems that use a mobile device (e.g., smartphone, tablet) screen-bounded interaction; to interact with an AR system, users typically must actively use a display as a mediator of the actual experience.

Hand held AR interfaces have a very obvious limiting factor: the user's hands are not free, with at least one hand allocated to hold the device or the augmentation object. In many cases the user holds the AR interface in one hand, using the other hand to interact. Clearly, it is desirable in the case of situated AR to be able to use both hands. To some degree, this drawback could be addressed by systems that use sensing technologies embedded in the environment (e.g., Sensetable [11]), but such systems are more complicated to deploy as they naturally require hardware installation in the physical space.

1.2.2 Head or Body Attached

Head-worn projective displays is another well researched area. While such an AR interface can provide a rich interactive AR experience, those typically have a 'personal' user experience which fundamentally provides an individual experience, over a 'spatial' or space-related experience that can be experienced by multiple users at the same time. Cakmakci et al. provided a comprehensive review of such systems [12]. Most of the head-worn displays use retroreflective material to increase brightness and reduce crosstalk between eyes and users.

Only recently have consumer-grade form factors emerged for wearable AR systems; a notable example is Google Glass, though the AR experience it currently provides is very limited and a source of much debate in the research community [13]. In the case of Google Glass, specifically utilizes a small monoscopic field of view that makes geometrically registered augmentations impractical in many situations.

Recently additional AR eyewear such as Microsoft's HoloLens [14] and the META 2 headsets [15] have demonstrated impressive new technical capabilities in sensing and augmentation, setting a new bar for AR experiences. We discuss further in Chapter Two.

This phenomenon is just one indicator that the process of adopting augmentation technology has already begun. I review some of the latest advances further later in this chapter.

1.2.3 Spatial

Originally proposed by Raskar, spatial augmented reality (SAR) describes AR systems that make use of non-traditional display technologies such as digital projectors, holographic displays, and transparent screens [10], [16]. SAR systems have a clear advantage, as they enable an immersive blending of digital media and physical space. This approach also decouples the display from the user, as it does not require the user to wear a head-mounted display or use a hand-held device. However, most SAR systems to date lack portability and are hard to relocate due to a fixed projector–camera setup. They can also suffer from problems such as self-occlusion and ambient lighting conditions

1.2.4 Discussion

The discussion on the advantages and drawbacks of the various types of AR systems is still very much in debate within the research community. Kruijff et al. provide an excellent classification of perceptual issues in augmented reality with respect to head-worn displays, handheld mobile devices, and projector-camera systems [17].

To date, mobile based AR is common-place. The techniques to use see-through AR using a mobile device display and camera are fairly known and well commercialized. Key examples are Apple's ARKit [18] and PTC's Vuforia [19].

In recent years wearable AR interfaces (HMD or HUD displays) have leaped forward. Examples of such AR interfaces are already appearing at mass consumer form factors. The current day interfaces are mainly wearable devices such as Google Glass [20], HoloLens [14] and the Meta Glasses [15] . We can also find AR in more complex products. One dominant example of that are Heads Up displays in vehicles. That said, such wearable AR interfaces require the user to adapt by wearing additional computing components that facilitate the interaction.

Finally, mobile-wearable systems such as the SixthSense [21] device developed by the MIT Media Lab Fluid Interfaces Group address this problem using a novel wearable form factor and a gestural interface that enables immersive mobile AR.

In my research, I've been drawn to the domain of projected augmented reality. The promise of using light as the key medium for interaction was compelling and exciting. I was intrigued by the fact that while the research domain is rich and deep, and the technological building blocks are generally available – there are very few examples of real world product deployment. Inspired mostly by Wellner's seminal work, the DigitalDesk **[22]**, **[23]** I embarked on my own path to transform light to a viable medium for interaction.

1.3 Light based Interfaces

The use of interactive light as a form of user interaction is nothing short of magic. As humans, with a peripheral vision system, we respond very well to light, shadows and occlusions. This fact provided plenty of motivation for conducting research in projected augmented reality, or 'light based interfaces'.

Early projector-camera system research dates to the 1980s. IBM published patents **[24]** that used shadow parallax phenomena to create an interactive display using projectors (see Figure 1-2).

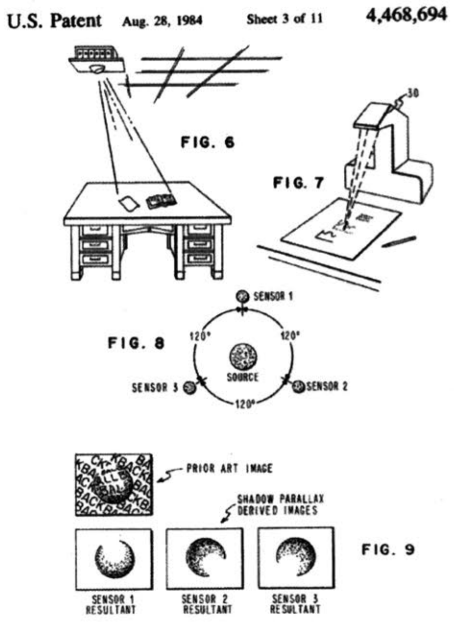


Figure 1-2 IBM Edgar's Patent – one of the earliest fully realized projector based interface with a full feedback loop that provides interaction.

In the early 1990s, Wellner's seminal work the Digital Desk **[22]**, **[23]**, **[25]** at XEROX Parc introduced for the first time the use of a projector camera in an office environment, for general information processing applications. The research community introduced numerous advances, including correction for complex geometry of surfaces and optics, support for multiple overlapping projectors, and tracking and projection onto objects **[26]**. Bimber and Raskar captured the evolution of the field in their book *Spatial Augmented Reality* **[10]**. Despite advances in the field, projected augmented reality interfaces, or "light based interfaces," are not widely adopted and rarely exist outside of research labs or experiments. There are few examples of commercial deployments of interactive projector-camera systems, none of which have reached the scale of mass-market consumer electronics.

Projected AR systems project digital information and interfaces onto the physical world and are typically implemented using interactive projector-camera systems. Such systems have the potential to truly fulfill the promise of ubiquitous computing, providing a transparent and embedded substrate of digital interfaces, enabling seamless experiences across devices and physical environments. Projected AR interfaces generally fit under the definition of "spatial augmented reality" (SAR), described above.

Figure 1-3 The original Apple Macintosh ad 1984 (left); The time of the Ubiquitous Light-based Interface has yet to come, light based interfaces are not yet commonplace (right). (Picture Source: Apple, Luxo).

Projected AR Interfaces are far from perfect. Like any interaction modality they have clear pros and cons. The use of projection-based spatial displays in interactive systems also introduces several new problems **[10]**, I will summarize the key advantages and challenges in the table below:

Table 1-1 Advantages and Constraints of Projected AR Interfaces

Advantage	Why?	Challenge	Why?
Plug & Play	Projected AR Interfaces can be embedded in the architectural environment, much like other light fixtures	*Shadows*	Interface may suffer from shadows cast by physical objects in the environment;
Low footprint	Light fixtures already exist in our spaces or can be wall or ceiling attached Light based interface itself can take up little space and adapt to physical environment	*Occlusions*	Interface prone to self-occlusions by the user or objects, who may obstruct the projected image and complicate interaction;
Works in messy environments	Light can get wet, dirty with no issue	*Display fidelity*	Projector displays core technology will always follow behind in terms of brightness and resolutions compared to other flat panel display technologies
Blending digital and physical	Pixels can be projected on physical objects	*Ambient light*	Interface may be sensitive to ambient light conditions in the target environment. Additional sources of interference may come from global illumination, including reflection and refraction from other parts of the environment. In general, projected interfaces are not useable in outdoor light conditions.
Context-aware	Projected AR interfaces are effectively computers that can sense a state of the interaction	*Sensing/Tracking robustness*	Natural interaction requires tight control loop that is hard to make robust and real time

	environment		
Flexible display size	Projection display technology makes it easy to dynamically (digitally) or physically change the display size	*Requires calibration*	Interface requires adaptive and real-time geometric projection model
Augment anything	Interfaces can appear on surfaces, objects or people	*Physical world dependency*	Requires augmentation target. Projected AR does not work in free space. There are also known constraints from size, shape, projector resolution, and projection surface material properties and color.
Natural Interaction	Use hand gestures, detect common physical action (e.g. object pick up) to create a true natural interface	*Limited development tools*	Very limited application development framework option compared to classic mobile, web environments

1.3.1 Inhibiting Engineering and Usability Factors

Several factors currently inhibit further development and mass adoption of projected AR interfaces. The three factors I believe are the most crucial are:

1. Slow introduction and integration complexity of new required hardware components. Only recently has laser and LED based pico-projection technology became widely available, but both are still very much behind in terms of performance, resolution, and intensity when compared to active displays. In addition, while the computer vision research community and industry rapidly evolved the field of depth-sensing hardware and algorithms, available sensors remain limited in resolution, operation ranges, and computational requirements, and remain relatively expensive.
2. Lack of modern application development frameworks that abstract the complexity of augmentation algorithms away from the developer.

3. Very few of real-world deployment and evaluation. This factor speaks to the lack of a concrete use-case or “killer app” that makes projected AR a viable must-have type of interface. Some examples of real-world products are the AccuVein [29] and VeinViewer [30] systems, which images veins using infrared light and projects the image on a patient’s skin, allowing a nurse to accurately detect the veins.

Due to these constraints, most interactive AR experiences adapted for everyday use typically involve simple single-channel I/O via fiducial marker scanning or image recognition. Immersive AR experiences typically involve the custom setup of projector-camera systems that require location-specific setup and calibration. Such systems normally use dedicated software to support the interaction. The community at Projection Mapping Central [29] is an excellent resource for readers who wish to explore the domain of immersive AR installations.

1.4 Approach

My research on interactive, projected AR systems is located at the intersection of the key problems I describe above and attempts to address them. I believe that the combination of these factors, and the absence of a true projector-camera integrated hardware form factor, coupled with simple and modern software frameworks, limits advances in the field.

It is very likely that the roots of this multi-faceted problem lie in the evolution of the form factors of computing devices. Essentially, we continue to use display-centric devices, and as such they define the key modality of interaction, which hasn’t changed substantially since the PC. I believe that humans are spatial creatures that were not ‘designed’ to spend most of their time in front of computers. And yet it seems that more and more of our work and play hours are spent interacting with a screen. This defines an ‘interface paradox’ from a cognitive, ergonomics and design point of view—we are designing interfaces that change our essence. The direct impact of this phenomenon is at the core of thought and discourse of modern sociologists. Vast research is being carried out to define and measure how human perception, thinking and behavior is altered due to the massive proliferation of technology and interfaces in our lives [30], [31].

Another aspect of this meta-problem is the fact that we are constantly curating several consumer electronic devices that connect us to the digital world: laptops, smart-phones, e-book readers, digital cameras.

We are responsible for updating software, running cables, charging, and syncing. Unfortunately, Weiser's Ubiquitous Computing [32] vision still seems far away, and poses a challenge to Human Computer Interaction research.

Information Technology has yet to become transparent and embedded in our environment. Augmented Reality interfaces—and specifically light-based interfaces—can help make technology embedded and transparent, bridging the digital and physical worlds. Projected, light based interfaces may hold the key to address the interface problem that is core to this dissertation work.

Thus, the research work proposed here is centered on the idea that a viable and widely adopted light-based interface design should address all elements of the system, considering hardware architectures, physical form factor, software framework, and application and interface aspects.

1.4.1 Rapid Development, Real World Deployment

The goal of this research work is to advance the practical use of projected AR interface in the real world. To accomplish that I have done extensive testing and deployment of our systems in real-world environments.

The research approach I have taken over the past seven years has depended heavily on the large body of work on HCI theory. Specifically, my work has relied on the practice of user-centered design (UCD). The term 'user centered design' stems from the work of Donald Norman at the University of California, and is considered a pillar of Human-Computer Interaction research and practice methodologies [33], [34]. This approach encourages the use of short design iterations that begin with deep exploration of a problem domain, and continued effort to generalize as a system or an experience is designed and implemented. To accomplish this, I have conducted numerous field studies and interviews with potential users of the systems described in this work. This fieldwork gave the work a unique starting point, strongly grounding it in the real needs of different sets of target users. These users were participants in the full lifecycle of the projects and contributed immensely across iterations. It is my belief that this approach leads to designs that are more meaningful, but, above all, very different when compared to pure lab-based prototypes that define abstract virtual personas that represent an imaginary user.

1.4.2 In-the-Wild Studies

To complement the approach described above, my work also draws on an emerging school of thought within the HCI community that practices in-the-wild studies. This methodology essentially promotes the design and prototyping of technology in situ [35]-[37]. One of the core ideas stems from the seminal work by Hutchins [38] that articulated the advantages of studying a social phenomenon within its context over isolated and artificial lab conditions. Effectively, this approach promotes interventions in real-world environments early and often in the design process.


(a) (b) (c) (d)

Figure 1-4 (a) Field studies and prototypes at Best Buy retail stores; (b) MARS live production trial at Steelcase's Kentwood West Factory; (c) Deployment of the white:scape prototype at Turnstone's workspace; (d) The Enlight user study deployment at the MIT Media Lab

This dissertation covers several iterations of hardware and software engineering efforts. The combination of both was adapted to use cases in the domains of desktop interactions, kinetic interfaces, retail, collaboration, manufacturing and education.

The reason that in-the-wild studies are particularly applicable to my work (see Figure 1-4) stems from my early goal to design a device that is self-contained and easily deployed. As such, it holds much greater potential to enter and enhance users' everyday lives than a lab-based prototype that exists only in lab conditions. In fact, we can think about this approach as "disruptive," promoting radical behavior change that could be sustained over time. This is different compared to the temporal nature of lab-based experiments.

The literature suggests that in-the-wild studies have demonstrated different outcomes than classic HCI studies [39]. Specifically, such studies have shown how users learn to appropriate novel experiences on their own terms for self-defined purposes. As such, they provide a much-needed critique of classic HCI research practices as well as an alternative to lab-based experiments.

1.5 Design Goals

This dissertation work presents overarching design goals that address the evolution of a novel augmented reality interface, its potential usability, and the technology's applicability to real-world deployment scenarios. The work is focused on two themes:

1. How can we create a self-contained, projected, augmented reality interface that allows for rapid application development and deployment? How can we embed projected AR interfaces into common day-to-day object?
2. What utility do the prototype systems and software frameworks we developed afford to the speed and ease of creating projected augmented reality applications?

To complement the overarching theme of this research–Rapid Development, Real World Deployment, and Evaluation of Projected Augmented Reality Applications I focus on two application domains: Augmented Manufacturing and Augmented Education. I will address the domain-specific research questions considering the high-level design goals presented above, and will also include a brief review of the respective related work for each of the proposed work areas.

1.5.1 Form Factors

I am passionate about product design and the evolution of form factors. Since the very early stages of the work, one of my personal goals was to advance the state of the art of projected AR interfaces and to explore how they can manifest in day-to-day objects.

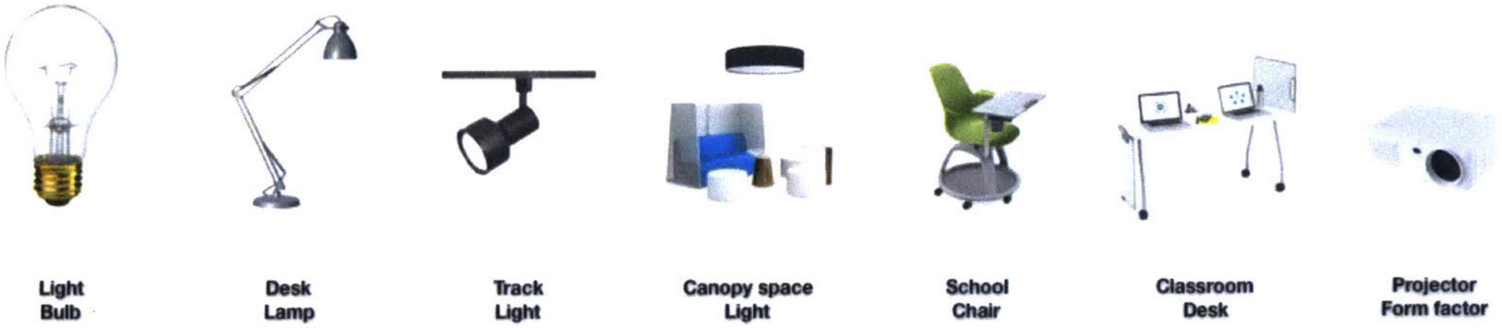


Figure 1-5 Form factors of light based interfaces created during the work on this dissertation.

I have included below in Figure 1-5 the various objects in which we have integrated LuminAR hardware and which were transformed to light-based interfaces. I intentionally leave the mapping of these objects to reveal themselves in the chapters that follow that describe the project work.

1.6 Key Projects Map

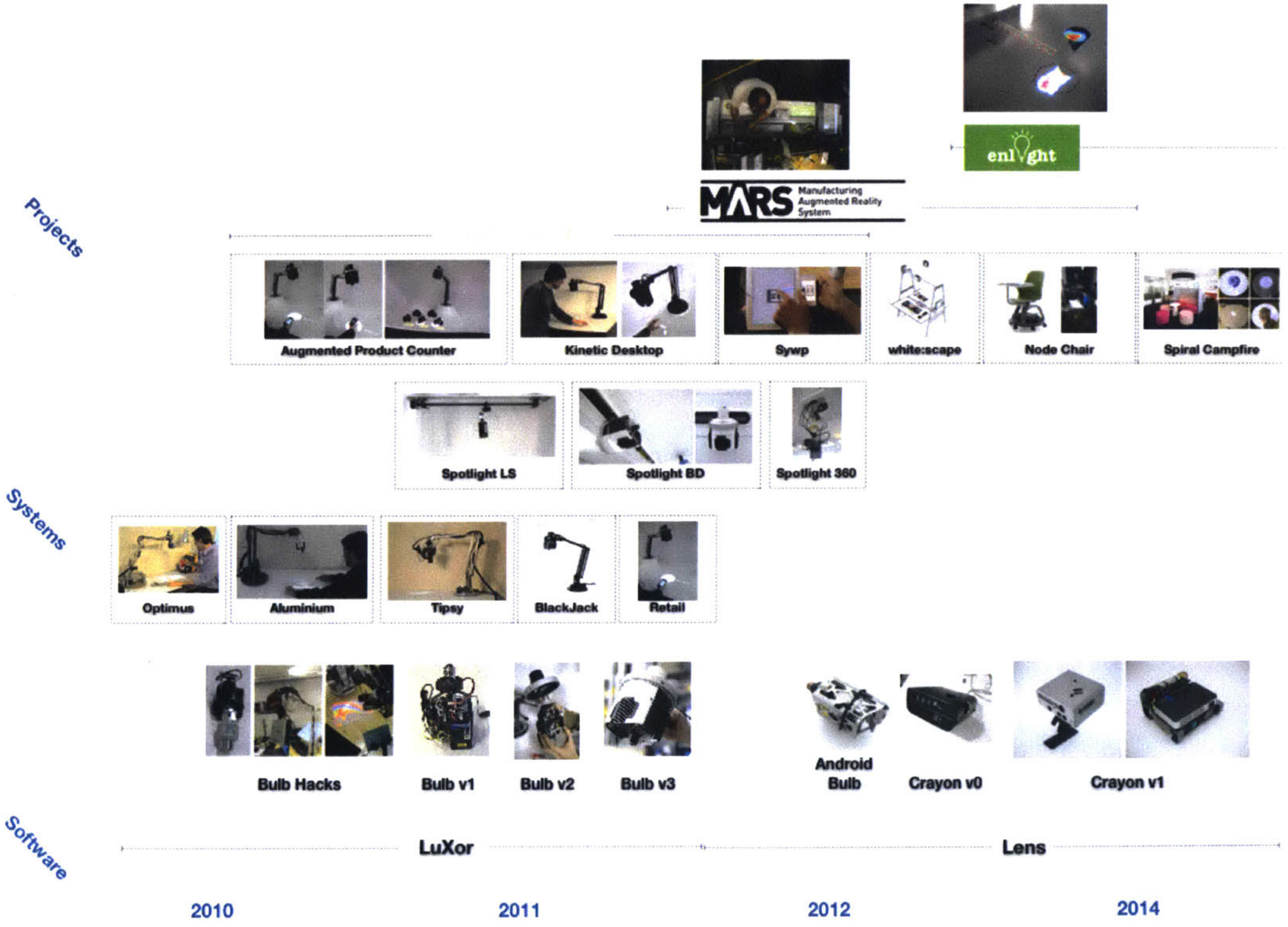


Figure 1-6 Key Project Map

Figure 1-6 provides a visual map for this dissertation work. It represents the various hardware iterations, software frameworks, interaction techniques and the engulfing projects and application domains my work spans. Later in this section, I briefly describe the key projects.

1.6.1 LuminAR

In his Ph.D thesis John Underkoffler introduced a Luxo-format for the I/O bulb, a conceptual evolution of the ordinary lightbulb: one which not only projects high resolution information but also simultaneously collects live video of the region it's projecting onto **[40]**. Implemented in 1999, the system employed joint sensor based forward kinematic calculations that provided the 3D position of the projector-camera pair. The 3D location was used to apply distortion corrections to the projected image. The LuminAR Lamp expands on this work.

LuminAR **[41]**, **[42]** is a compact projected augmented reality interface embodied in familiar everyday objects, namely a light bulb and a task light. It allows users to augment physical surfaces and objects dynamically with superimposed digital information using gestural and multi-touch interfaces. I presented several form factors for this new type of light-based computer that can bridge the gap between the digital realm and our physical world, in an attempt to push the vision of light-based interaction closer to reality. LuminAR was used to prototype several desktop and collaboration interaction techniques.

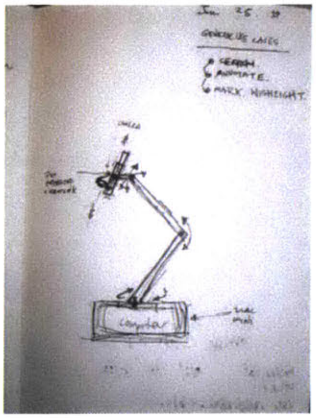


Figure 1-7 LuminAR early concept sketch, January 2010.

1.6.2 LuXor and Lens

In addition, this document describes LuXor and Lens, the software framework and application development tools we developed. LuXor was designed to support the basic projector-sensor functionalities as well as the kinetic aspects of the LuminAR system.

Lens is the evolution of LuXor to a web-based toolkit that allows developers to use standard web-development tools to build projected AR applications. This approach eases the development process and opens up light-based interfaces to all levels of developers.

Using the LuminAR, LuXor and Lens building blocks, this dissertation covers a wide variety of application scenarios implemented using our system, including a retail product counter, a kinetic interactive tabletop, collaboration methods, and manufacturing and education applications.

1.6.3 LuminAR Lamp supporting Kinetic I/O

We have also used the LuminAR hardware to create a kinetic version of the LuminAR interface which I described in my master's thesis **[41]**, **[43]**.

In the course of this research I also conducted several in-the-wild activities; a few examples are captured in *LuminAR Arm and Lamp Supporting Kinetic User Interface and Dynamic Multi-Touch* **[44]**.

Figure 1-8 (right) vintage Luxo Lamp; (left) LuminAR Lamp Blackjack model (source: Luxo, Doron Gild)

The LuminAR Lamp is composed of a custom robotic arm (i.e., LuminAR Arm), which is designed to interface with the LuminAR Bulb. The two together are combined to create a motorized version of the classic Anglepoise task light. We have developed six robotic arm platforms that were used to evolve the LuminAR Lamp concept. We were able to quickly use our insights across multiple iterations. However, our real goal in the design of these objects was to make the technology, specifically the robotics, disappear altogether. A fully integrated LuminAR lamp compared to a Luxo Lamp is shown in Figure 1-8.

Dynamic Multi-touch

Dynamic Multi-Touch **[42]** is an extension for the classic gesture vocabulary of multi-touch that takes advantage of a projected kinetic I/O system. Dynamic Multi-Touch systems utilize actuated DOFs of the projected touchscreen display to support real-time relocating, reorienting, and resizing of the projected display. Dynamic Multi-Touch attempts to extend the spatial limits of the classic screen bounded interface by allowing it to move. It also addresses the interaction space above the display. I have prototyped several gestures that explore the Dynamic Multi-Touch concept. The core principles of Dynamic Multi-Touch used in the LuminAR system were described in the recently granted US patent Kinetic input/output **[43]**.

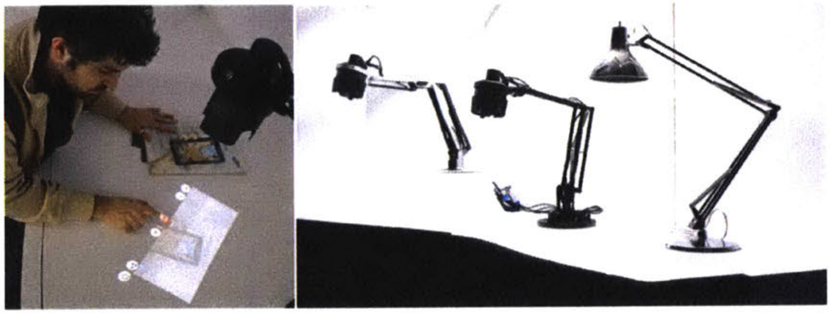


Figure 1-9 LuminAR Lamps and Projected Augmented Reality Interface

1.6.4 MARS

In one extensive application development effort, I present LuminAR in a form factor called the Manufacturing Augmented Reality System (MARS). MARS was designed to promote the adoption of lean manufacturing principles within a production operation. The general hypothesis in this area of the dissertation is that the use of projected AR interfaces in assembly tasks can increase worker productivity, reduce errors, and increase worker satisfaction. In addition, the system can also be used to collect detailed metrics critical to facilitating the continual process improvement that lean manufacturing seeks.

1.6.5 Enlight

In the educational context, I introduce LuminAR in another form factor: Enlight. Enlight is a system for deploying projected AR in education. The system complements existing approaches in science instruction, allowing instructors to incorporate AR applications and simulations with existing physical learning materials.

1.7 Evaluation

LuminAR represents a fully functional, compact, standalone, and fully deployable projected AR interface. It allows developers to easily create and deploy projected AR applications without deep knowledge of the underlying complex computer vision and AR algorithms that make it possible.

The system was evaluated with a series of quantitative and qualitative user studies and workshops to demonstrate the usability and advantages of the system design in terms of quickly developing and deploying real-world projected augmented reality applications. The motivation and purpose behind this evaluation approach is to demonstrate the generality of the system and its applicability to a wide variety of use cases. This dissertation work is based on our experience developing augmented reality applications and doing large-scale studies on the rapid development of projected augmented reality applications using the LuminAR hardware form factors and the LuXor and Lens software framework.

One of the main contributions of the resulting dissertation is demonstrating that light based interfaces are a near reality. My goal was to provide the community with real world examples that pave a path for the rapid development and deployment of projected augmented applications. This dissertation contributes the descriptions and evaluation protocols for such deployments in the domains of manufacturing and education, which served to assess the generality of the system as a solution for rapid development and deployment of projected augmented reality applications.

1.7.1 Methodology

The work completed during this dissertation was evaluated with a series of quantitative and qualitative user studies, as well as workshops that demonstrated the usability and advantages of our system's design in terms of quickly developing and deploying real-world projected augmented reality applications. The motivation and purpose behind this evaluation approach is to demonstrate the generality of the system and its applicability to a wide variety of use cases. I focused on two specific application domains of manufacturing and education. For these domains, it was required to perform deeper studies to expose the actual benefits of the system in a real-world deployment scenario.

The proposed user studies and experiments follow the acceptable practices in the HCI community **[36]**. I have received approval for two human studies plans with MIT COUHES:

- Effects of Augmented Reality Interface on Performance of Specified Tasks (MIT COUHES #1302005532)

- Effects of Augmented Reality in the Learning of Concepts in Physics (MIT COUHES #1405006364)

1.8 Contributions

This work contributes to the evolution of augmented reality interfaces in general and specifically advances the state of the art of projected augmented reality interfaces. This dissertation takes a holistic approach, providing design principles and technical architecture descriptions that could apply to a new class of interactive information devices. During this research, several detailed case studies of specific applications demonstrated the breadth of the technology and its potential. We collect those and distill them into guidelines that could usher in widespread use of such interfaces.

Specifically, the results of this dissertation work include:

- Novel, compact, and self-contained projector-sensor system hardware. Integration of the such system in common-use objects.
- Design and implementation of LuXor and Lens software frameworks for rapid development and deployment of augmented applications.
- Design principles and interaction techniques developed specifically for use with projected interfaces.
- Usability studies of the system focused in the domain of Augmented Manufacturing and Augmented Education.
- Descriptions and evaluation protocols for deployments of our system in manufacturing and education case studies.

1.9 Dissertation Outline

This dissertation is organized into nine chapters:

Chapter 1	**Introduction:** provides a high-level review of the current interface challenges addressed in this dissertation, and a brief overview of the research work.
Chapter 2	**Related Work:** provides motivation for this dissertation work. It covers the relevant human computer interaction theory and the previous work in augmented interactive spaces, multi-touch, and gestural interfaces as well as personal robotics. It also positions the work within the context of projected augmented reality interfaces.

Chapter 3	**Hardware:** provides the implementation details of the various projected AR systems prototypes. This chapter provides rationale for the various form factors I developed, and details the hardware implementation process. Readers who are less interested in hardware development, should feel free to skip this chapter.
Chapter 4	**Software:** describes the architecture and implementation details of the LuXor and Lens augmented reality software stack.
Chapter 5	**Interactions:** provides a review of the proposed interaction techniques we implemented using the various iteration of the LuminAR hardware
Chapter 6	**Applications:** provides a comprehensive review of the suite of projected AR applications created during this research. It mainly covers the early work in the domains of Work and Productivity and the Augmented Product Counter project.
Chapter 7	**Augmented Manufacturing Case Study:** describes the MARS project, an implementation of projected AR in a manufacturing environment.
Chapter 8	**Augmented Education Case Study:** Provides the details of the Enlight project. Enlight was our deep exploration to the use of projected AR in the domain of education.
Chapter 9	**Conclusion and Future Work:** summarizes the lessons learned and provides an outlook for proposed future work and a long-term vision.

2 Related Work

This research is motivated by the desire to evolve new form factors for computers that explore further the as-yet-unfulfilled promise of ubiquitous computing **[2]**. Computers have transformed human society. They carry the power of information, communication and creation to almost every aspect of our lives. However, it seems that we have adapted ourselves to technology and not vice versa. Our interaction with digital information is dictated predominantly by screen-based form factors and various input devices.

Interaction with digital information devices poses a challenge to us humans, as effectively we have become our very own human-computer I/O bottleneck. As we incorporate more devices into our lives, our curation process becomes more complex. The goal of this work is to design new form factors that address this problem space directly from two key directions: (1) create fluid interfaces that seamlessly blend digital media with the real world, and (2) design form factors that embed computation in everyday objects.

Both approaches are still very much considered to be open challenges, though recent years have shown the potential and feasibility of combining the two. To stir the discussion of the related work, this section provides a quick overview of the key streams of HCI research that serve as foundations for the work presented here.

By the early 1980s the personal computing revolution was well underway. PCs were powered by affordable microprocessors and had interfaces that, at the time, seemed reasonable for normal people to use: the keyboard, screen, and mouse. At the same time, another revolution was brewing, and a decade later Tim Berners-Lee completed his work on the World Wide Web **[45]** at CERN to address the needs of large, dispersed scientific collaborators to share diverse data. The Internet became a practical means to communicate and share digital media and information.

Both revolutions are in debt to pioneers like Vannevar Bush and Douglas Engelbart. Engelbart, as early as 1968, demonstrated in the Mother of all Demos **[46]**, what Bush had envisioned in his famous article in *The Atlantic Monthly* technology section: "As We May Think." **[47]** Their work had possibly the most profound impact on the interfaces we actually use to interact with computers. Soon the adaptation of computer-based applications into daily life became a reality: spreadsheets, video games, and word-processing are just a few early examples. Now, it seems hard to find a niche uninvaded by computers and networks. It was not long before HCI researchers and product designers realized that the benefits of computing, networking, and digital media also posed a great interface challenge. Researchers Mark Wieser and Hiroshi Ishii coined the terms Ubiquitous Computing and Tangible User Interfaces respectively **[2]**, **[48]**. Both envisioned a future where computers become invisible, and interfaces to digital information become embedded in everyday objects.

The notion that the physical world is still relevant to how we interact with digital media spawned a mass of research. Several projects pursued post-desktop, post-WIMP (window, icon, menu, pointing device) interactions. Examples of such research domains include: virtual reality (VR), augmented reality (AR), pervasive computing, context-aware computing, multi-touch, and gestural interfaces.

HCI research has been strongly influenced by these ideas of "ubiquitous computing," and vast efforts to implement some of these ideas have emerged over the years. AR is a good example of one such effort. There are several approaches in AR, some using goggles or glasses, others using handheld displays that blend digital information into the visual scene. In this dissertation, I am particularly interested in Projected Augmented Reality systems (PAR) or AR interfaces that are implemented using projector-sensor systems.

Another aspect of the interface challenge was well captured in the famous debate between Ben Shneiderman and Pattie Maes that juxtaposed Direct Manipulation vs. Interface Agents **[49]**. Maes suggested that computer interfaces are no longer self-contained and that computer screens have become windows to vast networks of information, foreseeing the Cloud Computing era. She also discussed how computer interfaces were designed with a 20-year-old set of assumptions that are now out of date. Today, even though much progress has been made, many of our interfaces are still designed with the same set of direct manipulation assumptions, which manifest in interface modalities of display centric devices (e.g., a tablet, a smartphone, or a laptop computer).

This dissertation work was inspired and influenced by several excellent research projects that explored interactive systems that blend the physical world with the digital. We review below some of the most relevant prior work, highlighting similarities and differences.

In the sections that follow, we will focus the related work review in the key domains that provide technical foundations as well as interaction inspiration for this research work. We will focus the review on the topics of augmented reality, interactive spaces, specifically with respect to projected AR techniques, tabletop computing, multi-touch and gestural interfaces.

2.1 Augmented Reality

As discussed above in Chapter one's "Augmented Reality Interfaces" section, while AR is a relatively young field, it does have roots as far back as the 1960s [4]. AR has been the goal of myriad research working to allow computer-generated virtual media to overlay physical objects in real time [6]—all part of the larger goal of creating environments that seamlessly blend digital information with the real world. My research was inspired and influenced by much of this work, including Milgram [8] (see Figure 2-1), Azuma [50], and Bimber and Raskar. Cakmakci et al. provided a comprehensive review of such systems [12].

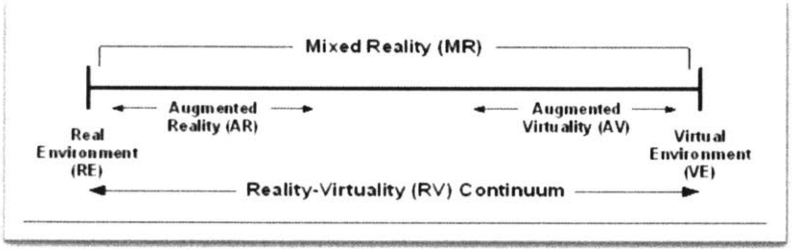


Figure 2-1 Reality-Virtuality Continuum (source: Milgram et al 1994; 1999)

Augmented reality interfaces first evolved via computer and camera based interfaces in the early years to the more generally available forms of mobile see through AR [51]-[53].

Of course, from the dawn of AR research, significant focus was given to AR interfaces in the forms of Head Mounted Display (HMD) or Heads Up Display (HUD) based systems [54]-[56].

Finally, the main domain of interest for the purpose of this work, projected augmented reality is still quite far from wide adoption in real-world commercial scenarios. Although advances in computation power and display systems have made great leaps only in the past few years (see Figure 2-2), current commercial experiences are very limited. Advanced projected AR systems are primarily found in academic and industrial research labs.

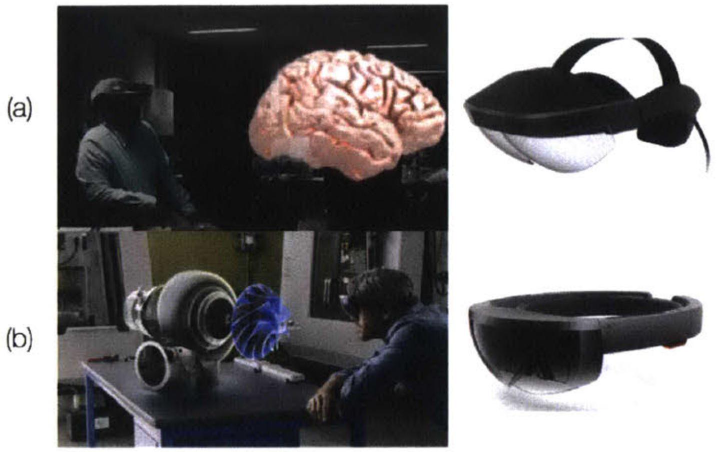


Figure 2-2 Examples modern wearable AR interfaces: (a) Meta 2 Headset (b) Microsoft HoloLens

Generally, projected AR interfaces introduce specific interaction challenges. Bimber et al. provided a good summary [10], and I include the key factors below:

- constraints from size, shape, projector resolution, and projection surface material properties and color
- shadows cast by physical objects in the environment
- sensitivity to ambient light conditions
- self-occlusions by the user, who may obstruct the projected image and complicate interaction
- the need for an adaptive and real-time geometric projection model, and
- unlike optical and video see-through systems, projection-based systems require a surface on which to project, so augmentation must be co-located with the target physical object.

These constraints have led most interactive AR experiences that are adapted for everyday use to typically employ simple single-channel I/O via fiducial marker scanning or image recognition. This is because immersive AR experiences involve custom setup of projector-camera systems requiring location-specific setup and calibration, and normally require dedicated software.

However, in the domain of mobile, see-through AR there are plenty of out-of-the box solutions such as Apple's ARKit [18] and others. I will discuss this theme furthermore later in chapter four that covers the "Lens" software framework.

(a) (b) (c)

Figure 2-3 Examples of classic mobile AR: (a) Daqri app (b) Pokemon GO mixed reality mobile game (c) Apple's AR Kit (source: Daqri, Nintendo, Apple)

There are three major interaction-limiting factors of existing mobile see-through AR systems. To some extent, wearable AR interfaces (HMD or HUD displays) address these problems directly, but they require the user to adapt by wearing additional computing components that facilitate the interaction. The three interaction-limiting factors are:

1. **AR is screen bounded.** Current commercially available see-through AR systems predominantly use a mobile device (e.g., smartphone, tablet, see Figure 2-3) and as such, by definition provide screen-bounded interaction. To interact with the AR interface, users typically must actively use a display as a mediator of the actual experience. As mentioned above, HUD and HMD deal with this limitation, as the AR view port is aligned to the user perspective. However, it requires the user to wear and interact with a head-worn piece of technology.

2. **Context switch is required to complete the experience.** Experiencing AR via a mobile display view-port requires users to context-switch back and forth between the physical space and the screen used to display the augmented scene. This introduces interaction discontinuity and poses challenges to creating natural flow in AR applications. Again, this limitation is less of an issue with wearable devices that do a good job providing a true mixed reality experience, For example, when using Google Glass [20], the user typically need to shift her view to the top side of the viewport to see the actual interface, due to very limited Field-of-View (FOV). Other HMDs such as Meta[15] and HoloLens[14] provide true holographic experience, eliminating the need for a context switch.

3. **The user's hands are not free to interact.** Mobile AR systems address some of the issues described above by allowing the user to carry the AR system in his hand; however, they are still suffering from similar interaction limitations. Users normally have to hold onto or wear (e.g.,

> interaction gloves) marker objects in order to interact with the AR interface. Even more critical is the fact that, in many cases, the user holds the AR interface in one hand using the other hand to interact. Clearly, it is desirable in the case of situated AR to be able to use both hands. This is of course, not a limitation of HMDs or HUDs.

This work is focused on solving these problems in the indoor domain. I am further assuming that the user experience will benefit if we do not impose additional hardware.

2.2 Projected Augmented Reality

For the purpose of this dissertation, and in light of the abundant work in the domain of augmented reality, it is important to position the various LuminAR iterations, MARS and Enlight, within the context of the previous work.

Figure 2-4 DigitalDesk by Pierre Wellner

Several surveys classify and analyze different technical aspects of the actual display and sensing sub-systems, and we refer the readers to the latest survey by D.W.F. van Krevelen and R. Poelman **[56]** for further reading. Bimber and Raskar [10] classified the approaches to augmented reality display technologies into head-attached, hand-held, and spatial displays.

Positioning wearable projectors and steerable/kinetic projector systems in the context of Spatial Augmented Reality, I have updated the original diagram by Bimber and Raskar to include the cases of wearable projector and steerable/spatial projector in blue. They also proposed the framework depicted in Figure 2-4 to summarize this categorization. We have added to their classification two additional cases that complement the definition of spatial augmented reality displays with new approaches that were reported since their work was published: (1) wearable projector, and (2) steerable/kinetic projector. MARS and Enlight clearly fit in the categorization under (3) spatial optical see-through display using projectors. This category is marked in blue in Figure 2-4.

2.3 Projected Interactive Spaces

In recent years, enabling technologies for the design and implementation of augmented interactive spaces have been extensively researched. Some of the past work includes various computer vision based methods for interaction and augmented reality systems [6], [10].

Augmented Interactive Spaces typically use a projector-camera system along with computer vision techniques to augment a physical space. Wellner's DigitalDesk (see Figure 2-5) system is considered one of the first pieces of work in this domain [22], [23], [25]. Rekimoto and Saitoh expanded on this approach by introducing wall, table, and laptop cross-integration, allowing information to propagate between modalities and devices [59]. The EnhancedDesk [60] introduced real-time finger tracking using IR light, enabling the integration of paper and digital information.

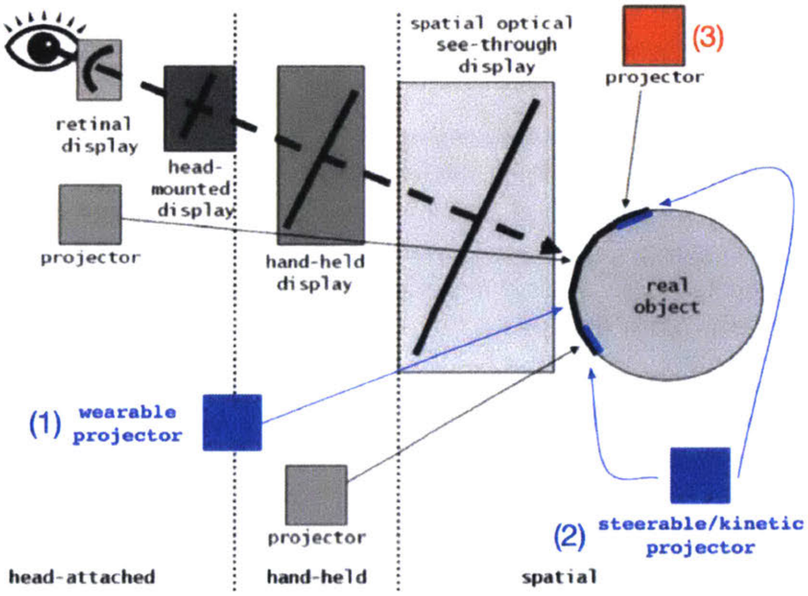
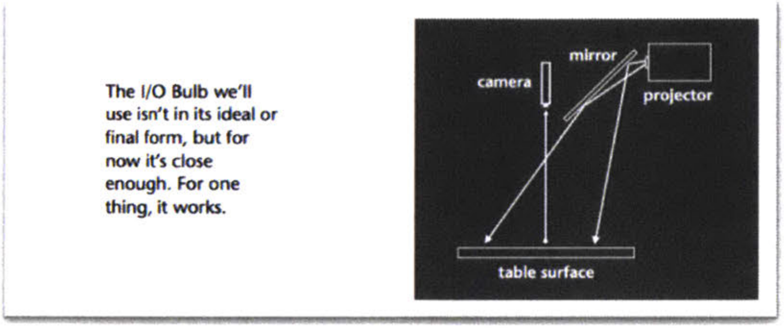


Figure 2-5 Positioning wearable projectors and steerable/kinetic projector system in the context of Spatial Augmented Reality, we have updated the original diagram by Bimber and Raskar to include the cases of wearable projector and steerable/kinetic spatial projector in blue. (source: Oliver Bimber and Ramesh Raskar)

The LuminAR system enables similar flat surface, desk-, and table-based augmentation and interaction, yet it provides the advantage of portability and compact size, as the entire system is self-contained in both the case of the LuminAR Bulb and the LuminAR Lamp form-factors.

In his seminal work, I/O Bulb and Luminous room [61], [62] (see Figures 2-6 and 2-7), and Urp [63], Underkoffler has described a basic two-way optical information device and how it will be used to transform a room into an architectural information space. At the time, technology and cost issues prevented the full realization of this vision. The actual system implemented included spatially separated cameras and projectors. Although a "real" I/O Bulb was described, it was unavailable in research labs or as a commercial product prior to the work of this thesis.



The I/O Bulb we'll use isn't in its ideal or final form, but for now it's close enough. For one thing, it works.

Figure 2-6 I/O Bulb and the Luminous room by John Underkoffler (source: **[40]**)

Pinnhanez created the Everywhere Displays **[64]** that used a stationary projector with a rotating mirror to expand projection to multiple surfaces (see Figure 2-8). He also proposed a portable solution, Everywhere Display Lite **[65]**. This work introduced the concept of steerable interfaces that can appear on objects and surfaces anywhere in space, and can also dynamically adapt to form, function, and location of the interface as well as user context.

However, even though the mirror element presented in Pinhanez's work offers more flexibility, it allows very limited animatronics compared to the articulated degrees of freedom of the LuminAR Lamp, and it also requires a complex setup.

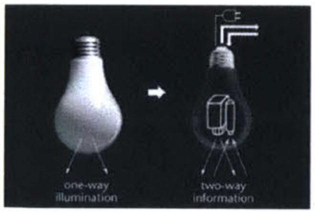


Figure 2-7 I/O Bulb concept by John Underkoffler et al, 1999

One of the challenges of implementing a steerable and interactive projector-camera system lies in the calibration of a generalized optical and mechanical system. Ashdown and Sato presented a method and algorithm for calibrating a steerable projector system, focusing on pan-tilt mirror setups **[66]**. Their work represents an important step, though it is still limited as it is based on the projection of patterns.

Steerable projector systems that use a pan-tilt mirror setup indeed offer more flexibility, but

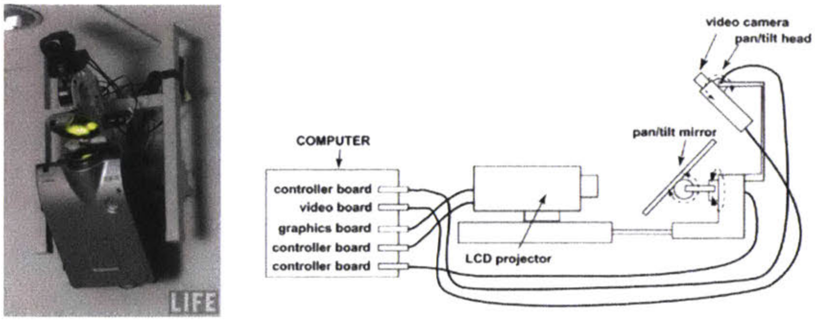


Figure 2-9 EveryWhere Displays by Claudio Pinhanez et al, work done in IBM Research (source: IBM Research)

are constrained in terms of calibration and ease of setup. LuminAR address these drawbacks by using a moving projector approach combined with simple one-time offline calibration.

Figure 2-8 DockLamp by Frederic Kaplan et al

In recent years, important advances in portable and small-scale projector camera systems have been realized. iLamps by Raskar et al. [26] presented an adaptive projection on non-planar surfaces enabling object augmentation with a hand-held projector, including interaction techniques. Wilson outlined a more advanced variation on portable projection systems in his project PlayAnywhere [67]. This project explores paper, hand, and object recognition using a projection system. The system supports various wall and tabletop use cases. More recent work has combined pico-projectors and cameras in a lamp form factor. A prominent example is the DockLamp (see Figure 2-9) by Kaplan, et al. [68]. In addition, Microsoft research has created the Mobile Surface [69] as part of their effort to push forward the domain of Surface Computing. Wilson and Benko also presented very relevant work on interaction above surfaces [70].

Even though these examples demonstrate substantial, yet incremental improvements in portable AR interfaces, unfortunately all the systems described above are completely static once they are installed in a space. This fact limits the scope of interaction and augmentation experience these systems can provide.

Projects like DeskJockey [71] and Bonfire [72] have introduced PC-centric workspace extension using projection (see Figure 2-10). DeskJockey used a fixed projector to superimpose widgets in the space around a PC. Interactions were possible via proxy window on the PC desktop only. The Bonfire project suggested another approach, presenting a nomadic projection system attached to a laptop using two pico-projectors.

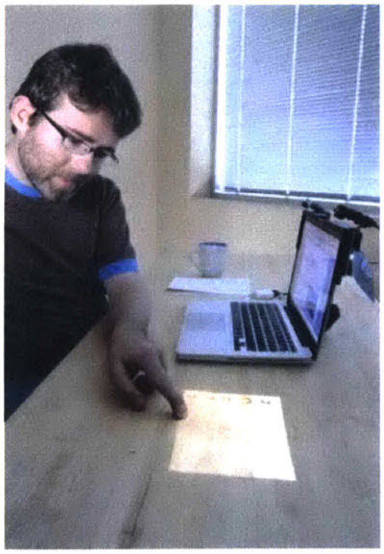


Figure 2-10 Bonfire by Shaun Kane et al developed in collaboration with Intel Lab (source: Intel Labs)

Prior research in augmented interactive surfaces has extended surface-based interactions to more elaborate object-based physical/digital interactions. Systems like metaDesk [73], WeSpace [74], UbiTable [75], and MemTable [76] introduced object-based interactions to retrieve documents or interact with the projection surface. These systems typically use fiducial markers, RFID, or other wireless mediums (e.g., Bluetooth, Zigbee etc.) to register objects in the system. This approach yields an enhanced user experience, as shown in the OASIS project from Intel Research Lab [77]. OASIS introduces a framework for general-purpose object recognition for interactive systems. The framework is integrated with a projected display to create interactive "islands" in everyday environments.

The LuminAR Bulb introduced two major advances over this prior work. First, it realized the I/O Bulb's vision, and efficiently integrated all the required components for the bulb (camera, computer, projector, sensors) into a single system. It is thus a truly portable and scalable solution to implement customized augmented interaction spaces easily. Second, it added actuated degrees of freedom, and robotic control elements that enabled the LuminAR Bulb to dynamically change projection parameters.

The LuminAR approach does not require or assume a PC to augment a space and it is not dedicated to PC-based interaction. The LuminAR system has a dedicated computer that allows it to remain portable and to easily interface to different displays in its vicinity.

2.4 Gestural Interaction

Multi-touch and gestural interfaces provide the foundations of interaction techniques presented by this work. One of the earliest visions of gesture-based interfaces was introduced in the StarFire video prototype [78]. Early gestural-interface systems employed simple computer-vision based freehand-gesture recognition techniques [79]-[81].

In the past decade, the domain of gestural and multi-touch interfaces has been extensively researched. The key element of such systems is the implementation of robust touch-detection mechanisms that allow direct manipulation of data using natural gestures [82]-[84]. Buxton [85] provided an excellent summary of such systems . Multi-touch systems use a large variety of sensing technologies; the most common include: (1) periphery, front-, or back-mounted cameras using custom surface as a display and interaction surface [86]-[88]; (2) IR cameras [79], [89] and multiple (stereo-vision) cameras [80], [90];

(3) embedded capacitive or resistive sensors **[82]**, **[84]**; and (4) acoustic sensors **[91]**.

The fundamental drawback of such multi-touch systems is that they are only able to detect physical touch events, and therefore incapable of supporting touch-independent freehand gestures or to detect arbitrary objects. They also involve a non-trivial setup of several independent components (i.e., computers, projectors, mirrors) that complicates deployment, and in most cases, they require users to perform manual calibration steps. Current gesture-recognition systems discussed below, unfortunately still suffer from similar limitations. Our approach uses computer vision techniques to create a system that supports both touch interactions as well as gesture recognition in a portable form factor, addressing many of the drawbacks described above.

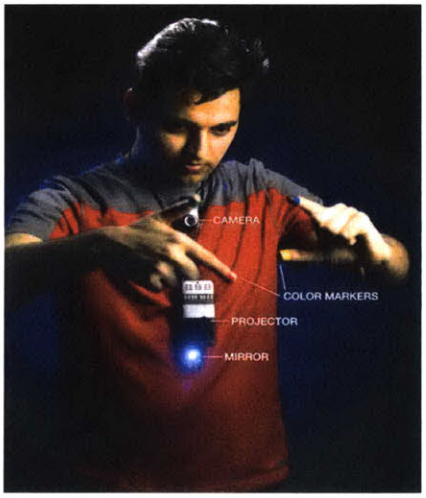


Figure 2-11 SixthSense device by Pranav Mistry and Pattie Maes (source: MIT Media Lab)

In recent years, various advanced gesture interaction paradigms have been explored including hand gestures **[92]**-**[94]**. Additional examples include the HoloWall **[86]** project and Oblong's g-speak system **[95]**, see Figure 2-12. Continuing advances in miniaturization of projection technology, embedded computing, and computer vision algorithms have enabled several portable, mobile, and wearable gestural interface systems in the past few years. Examples of such systems include the SixthSense device (see Figure 2-11) **[21]**, **[96]** and the iLamps project **[26]**; for additional examples see **[97]**-**[100]**.

Figure 2-12 Oblong G-Speak System (source: Oblong Industries)

2.5 Form Factors

The wearable systems discussed above, generally involve a user who either wears or holds a projector, using it to augment a surface or object. Our work is distinct from handheld gestural interfaces as it leaves the user's hands free to interact.

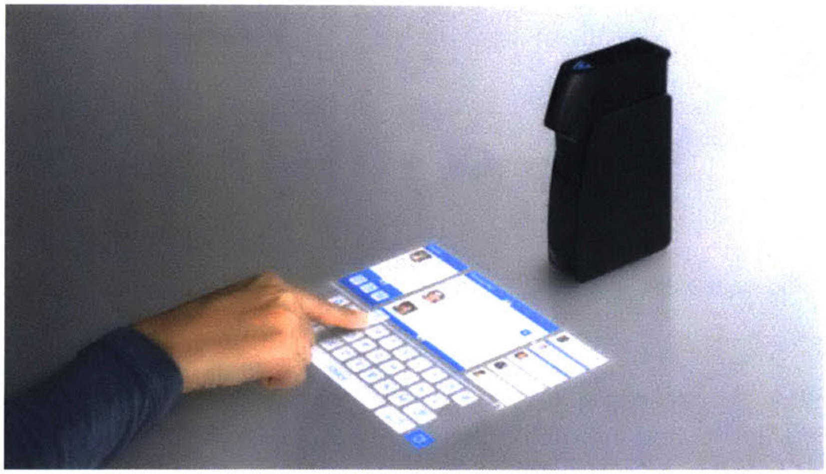


Figure 2-13 Light Touch by Light Blue Optics

It is also distinct from wearable systems, as it does not require the user to wear anything in order to interact with the system. Since LuminAR is a situated AR system, our work emphasizes portability over mobility. By "portability," we refer to the ability to simply unplug the system and move it to a different location, without the need for a complex setup or further calibration. In 2010, Light Blue Optics, a UK-based projection systems start up, introduced Light Touch **[101]** a portable, interactive laser pico-projector platform capable of projecting a virtual touchscreen (see Figure 2-13). Another classic example for projected keyboard was presented in 2002 by Canesta **[102]**, see Figure 2-15.

The work presented in this dissertation was completed in 2014. In the time between the conclusion of the work and writing this thesis, the space of projected interface has evolved, with quite a few new interfaces being introduced commercially. I summarize a few of the notable examples below:

Figure 2-14 HP Sprout (source: HP)

Sprout by HP is a personal computer introduced in 2014 by HP (see Figure 2-14). The sprout has dual interactive touch screen. The vertical screen is a standard touch screen, while the horizontal touch screen is a projected screen. The sprout also includes a camera and a depth sensor.

Bob,
The meeting has been moved to Thursday at 2pm

Figure 2-15 Canesta projected keyboard

Sony has released in 2017 the Xeperia Touch **[103]**, a compact interactive short throw projector (see Figure 2-16). The product is designed mostly for consumer experiences and based on the Android operating system In addition, two notable examples from the Future Interfaces Group at CMU followed LuminAR and provide a great extension to our early work. They introduced World Kit and Desktopgraphy (see Figure 2-17). WorldKit **[104]** is a software development framework that allows a developer to quickly develop projected widget. We cover this work in more detail in chapter four. Like LuminAR Desktopgraphy **[105]** hardware is using a standalone light bulb interface form factor that integrates a projector and a depth sensor. Desktopgraphy defines a set of gestures such as zoom, pinch and attach to object that specifically apply to projected AR applications. We have described similar interactions in our work on Dynamic Multi-touch (see Chapter Five).

.

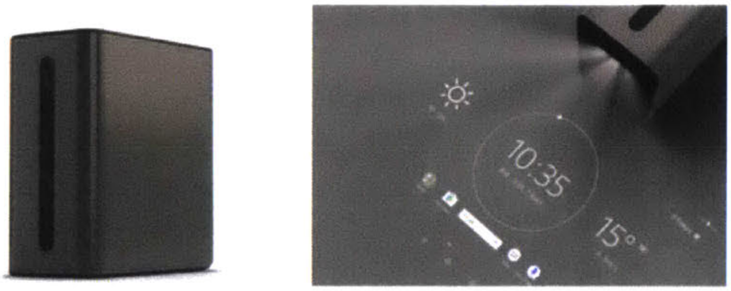


Figure 2-16 Sony Xperia Touch Projector (source: Sony)

In addition, two notable examples from the Future Interfaces Group at CMU followed LuminAR and provide a great extension to our early work.

Figure 2-17 CMU Desktopgraphy Project bulb based projected AR interface (on left); Widget based interface (right) (source: CMU Future Interfaces Group)

3 Hardware

The path that led to this dissertation work included several iterations of projected augmented reality software and hardware prototypes. To complement the theoretical background and the research approach this dissertation presents, I will briefly present my previous work as background to this dissertation work.

The LuminAR prototypes were developed through several design and implementation iterations, aiming to push the technical boundaries of existing augmented reality interfaces, but also to break new ground in terms of the industrial design of new form factors.

In this chapter, we describe the technical implementation details for the family of LuminAR devices we have designed and developed in the course of our research. We begin with discussing the design principles LuminAR devices share, and continue to describe each of the prototype systems we have built.

We begin with the heart of the system - the LuminAR Bulb. We then discuss the prototypes of the LuminAR Lamps, which are comprised of specifically designed robotic arms that carry the LuminAR Bulb. Next, we discuss the LuminAR Spotlight device prototypes. Finally, we describe the LuminAR Projector form factor.

For each of the different systems we developed, I present the design process and goals that guided the work. We also provide a deep hardware development documentation for each of the from factors. The hardware implementation section provides the interested readers detailed information about the hardware components and architecture, as well as various other technical considerations. Readers who are not interested in such level of detail may prefer to skip those sections. Finally, we also refer the readers to the technical appendix of my Master's thesis [41], where we include a complete mechanical specification for the various LuminAR devices described in this chapter.

3.1 Form Factor Design Guidelines

In Chapter Two, we presented and discussed the key properties of projected augmented reality interfaces. This background research was the driving force behind the design of the various AR interface hardware I implemented as part of this research work. My focus as a product designer was to achieve the highest level of integration possible between the various hardware components. In this section, I will describe the design guidelines that inspired and informed the hardware design of LuminAR, MARS and Enlight. It is my belief that these guidelines are necessary to evolve projected augmented reality to a ubiquitous user interface technology that is integrated to real products.

Simple everyday objects that hide technology

Hide computation and interfaces in familiar devices, which have common, everyday form factors. The forms LuminAR devices take on today are those of a light bulb, an Anglepoise task light, and a track light fixture. From a user's perspective, these objects should obfuscate the underlying technology they use. At the same time, when used, they provide known affordances and almost natural, contextual interaction techniques, using hand gestures, multi-touch, and kinesthetic motion (we delve into the details of LuminAR's interaction design principles later in this chapter).

Portable, adaptable compact, and self-calibrating

Design compact and portable devices capable of dynamically forming an interaction space. A non-technical user can easily relocate a LuminAR device from one place to another. The system should function without the need for any prior knowledge of the space it is deployed in. This dictates the use of simple power and data interfaces, but also the need to support automatic calibration of the AR subsystem.

Embedded, connected and modular

Directly derived from the need for portability is the need for the LuminAR devices to be self-contained computing systems. This dictates the design of the LuminAR bulb: the heart of each LuminAR device is a wireless embedded computer with all required peripheral interfaces for display, actuation, and sensing.

The LuminAR bulb can connect to a network connection. The network can provide content, interfaces, and facilitate data exchange with other LuminAR devices. It also provides LuminAR devices with modularity, in two key aspects: (1) off-loading computation to remote devices or other network services, and (2) supporting interaction scenarios where two or more LuminAR devices are used in parallel.

Design for product

Product design balances the need for functionality and aesthetics. One of the goals of the hardware design in this dissertation work was to push the boundaries of current computer form factors, product design methodologies played a key role. We considered issues that go beyond a pure interaction research agenda, and considered core product design issues such as ergonomics, safety, standards, heat dissipation, and work envelopes. Product design also assumes that prototypes mature and down the road are manufactured. This is the reason we tried, as much as possible, to use existing, off-the-shelf available technologies. We selected components that can be mass-produced, making LuminAR devices low cost and commercially viable.

Kinetic I/O

The ability to move makes interfaces dynamic (i.e., not statically situated in space). This principle is juxtaposed with existing display-centric devices that by definition are static once positioned within a workspace. LuminAR devices can dynamically reposition a projected interactive display. They can also use motion to enhance user interaction.

Kinetic I/O can be achieved mechanically, like in the case of the LuminAR lamp, in this case the entire projection area can be shifted and relocated per the reach envelope of the mechanical system.

Kinetic I/O can also be achieved virtually, by having the UI engine render UI in various locations in the physical space. This is bound by the projection real-estate.

3.2 LuminAR: Projected Augmented Reality Interface

LuminAR [42], [44] is a compact projected augmented reality interface embodied in familiar everyday objects, namely a light bulb and a task light. It allows users to dynamically augment physical surfaces and objects with superimposed digital information using gestural and multi-touch interfaces. We also created kinetic versions of the LuminAR interface. We describe the various version in the sections that follow.

The LuminAR prototypes were developed through several design and implementation iterations to push the technical boundaries of existing augmented reality interfaces and to break new ground in terms of the industrial design and new form factors. LuminAR attempts to advance the state of the art of projected augmented reality interfaces. Our prototypes provided designers and engineers with a complete design of a compact and kinetic system that is feasible to build with current technology.

Another goal of this work was to provide a glimpse into future potential applications of this technology in relevant real-world scenarios. The work on these applications enabled us to distill a set of new interaction methods. These new techniques make use of the unique form factor components of the LuminAR system and their respective kinetic capabilities.

Finally, after several iterations of prototyping, we were able to refine our design and reach another goal of this work–embedding interface technology in everyday objects. In this part of the work we emphasize the industrial and product design aspects.

3.2.1 LuminAR Prototypes

We developed four different systems for exploring the space of kinetic projected augmented reality interfaces: LuminAR Bulb, LuminAR Lamp, LuminAR Spotlight and the LuminAR projector. The insights we derived from these prototypes can hopefully stir the discussion on a set of design principles for self-contained, compact kinetic interfaces and associated interaction techniques. We provide a brief description of the prototypes later in this chapter.

3.2.2 Related Work: Kinetic and Robotic Interfaces

The field of personal robotics has advanced at a fast pace since the introduction of microcontrollers. Domestic robots have become a commercial reality in areas such as companion toy robots (Sony's AIBO and Innovo Labs' Pleo) **[106]**, **[107]**, service robots (Roomba and Scooba from iRobot) **[108]**, and personal surveillance and security

[109]. Social robots that serve as robotic coaches have also recently emerged from research labs [110].


(a) (b) (c) (d) (e)

Figure 3-1 Personal Robotics: (a) Innovo Labs Pleo (b) Sony Aibo (c) WooWee Rovio (d) Intuitive Automata Autom (e) iRobot Roomba (image sources from respective companies)

However, personal robots designed as 'information interactive robots' are still in a very embryonic stage (see Figure 3-1). It is clear that the future will make more use of AR techniques to facilitate a more expressive interface for robots themselves. Green et al. detailed the case for AR-enabled robots in a comprehensive review [111], [112]. We are interested in using robotic capabilities to enhance interactive information driven application. This work explores the design of such robots and their use in real-world scenarios.

In 2008, Hoffman created AUR, a robotic desk lamp, which acts as a collaborative lighting assistant [113]. AUR was used to explore the concept of fluency in human-robot collaboration. Finally, Stubbe and Lerner created Outerspace Robot [114], a reactive robotic creature designed to explore the surrounding space and interact with people through touch and behaviors (see Figure 3-2).

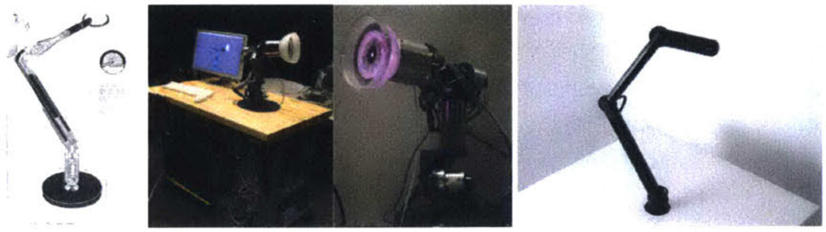

(a) (b) (c)

Figure 3-2 (a) Besselink et. al. Robotic Lamp (b) Hoffman's AUR (c) Stubbe and Lerner Outerspace Robot (source: J. Phys. IV France, MIT Media Lab, A. Stubbe)

These projects are good examples of robotic arms specifically designed for human interaction, but which are very limited in digital media interaction capabilities, yet they have no display and very little spatial sensing.

3.2.3 Characteristics of Steerable and Kinetic Projected Augmented Reality Interfaces

Pingali et al.'s groundbreaking work on steerable interfaces [115] presented an alternate approach for pervasive computing interfaces. In their vision, interfaces to computing should appear where and when the user needs them. They outlined the salient characteristics of such steerable interfaces:

- Support moveable input and output: interactive interface that can move around the physical environment and appear on ordinary objects or surfaces.
- Adapt to user context: interface should respond to some sensing of user needs based on location, physical function, or other data. Freeing the user from a single point of access via a computer with a traditional mouse and keyboard interaction.
- Adapt to environmental context: interfaces that adapt to the characteristics of the environment and change accordingly in terms, of location, size, shape, color, etc. This normally means that the system should be aware of the geometric and ambient parameters of the physical space.
- Device-free and natural interaction: a steerable interface is able to sense and support forms of interaction such as speech, hand gestures, motion of body, and touch, which are based on the human body and do not require special devices.

The LuminAR work proposes new characteristics to complement the set above and possibly drive forward the adaptation and investigation of steerable and kinetic projected augmented reality interfaces:

- System portability: users should be able to reconfigure and change their environments. Projected augmented reality systems should be designed to be compact and simple, such that a user may easily relocate the system in a physical space. To achieve such portability, it is important to design compact systems that support automatic calibration of the projected display. It is also required that such systems are able to sense and map the environment to support interaction.
- Embedding computation in everyday objects: closely related to the need for high portability is the need to design pervasive

interfaces that disappear into the physical environment. To achieve this, it is required to design computation functionality into already-existing everyday objects such as desks, lamps, and light bulbs.

- Integrating user interface modalities: projected kinematic interfaces should be designed to support multiple user interface modalities that utilize traditional multi-touch combined with hand gestures and object manipulation. Users should be able to seamlessly and naturally mix all modalities.
- Using physical motion: physical motion can greatly enhance interfaces, as humans are responsive to kinesthetic sensations. It is possible to create a richer user interface that uses motion cues as part of the interface feedback loop. Furthermore, actuation and animatronics sequences can be used to direct a user or even embody information.

3.3 LuminAR Bulb

The LuminAR Bulb is a compact, kinetic, and interactive projector-sensor system. The system was designed to follow the physical metaphor and design language of a classic light bulb, evolving it into an interactive computing device.

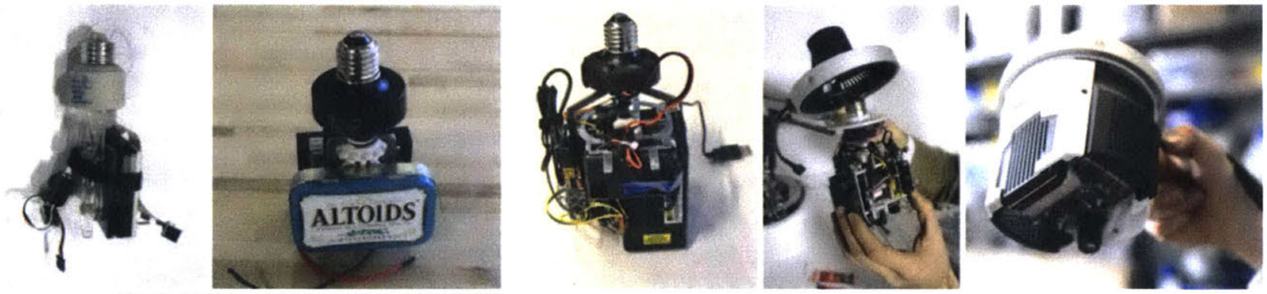


Figure 3-3 The evolution of LuminAR Bulbs

While trying to adhere to the dogma of bulbs that fit into a standard power distribution grid, I was inspired by the conceptual underpinning that Underkoffler provided in his work on I/O Bulb [40].

3.3.1 Design Goals and Process

With this general goal in mind, we also had concrete design goals that guided the design work of the LuminAR Bulb. Figure 3-3 shows the hardware evolution of the LuminAR bulb.

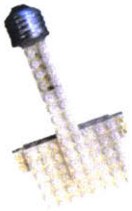


Figure 3-4 The bulb's LEGO bricks mockup model

- Create a fully realized two-way information device requiring only AC power and wireless communication to operate.
- Design the bulb to fit in standard Edison bulb sockets.
- The bulb should be compact and a completely self-contained projector-sensor system. It should integrate a laser pico-projector, cameras, sensors, control electronics, power supply, and an embedded wireless computer.
- The bulb should have its own DOFs for pivot rotation and camera tilt.

Design Process

Design and development involved several iterations. Initial explorations of form and fit to light fixtures were constructed from LEGO bricks, but we quickly moved on to advanced materials in the prototype that followed. The shape and weight of the bulb were critical factors to observe, as we tried to integrate together several components. Fitting the bulb in standard fixtures and lampshades dictated our general circumference and weight constraints. To explore physical constraints, we created a set of conceptual prototypes.

Figure 3-5 Testing of early version of the bulb in a Luxo Lamp

As our work progressed, we were able to refine the design and consider different packaging strategies. We discuss these details in the sections that follow. We focus our discussion on the latest and most relevant iterations of the LuminAR Bulb.

Early Versions

Our first version of the LuminAR Bulb was very much a learning prototype and had limited functionality. Based on our initial mock-ups, we designed a frame to carry a pico-projector and a webcam. The frame was also fitted with a standard male Edison screw using a hacked bulb screw adapter. At the base of the frame we constructed a stationary double disc part. The stationary part was glued to the screw cavity, providing support and pivot-axis rotation. The rotation axis was not actuated, but we could manually set the radial position of the bulb.

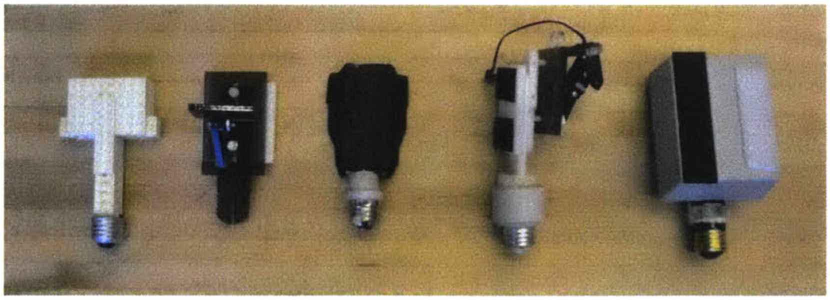


Figure 3-6 Early mock up models and conceptual prototypes of the LuminAR Bulb

This bulb's frame was constructed from laser-cut acrylic parts that were press fitted and glued. The frame was designed to support the adjustment of the projector plane and camera plane individually. This was done to allow us to experiment with different hardware calibration parameters for the computer-vision software. However, this version of the bulb did not include any power circuitry, and we used external power cables to supply the bulb.

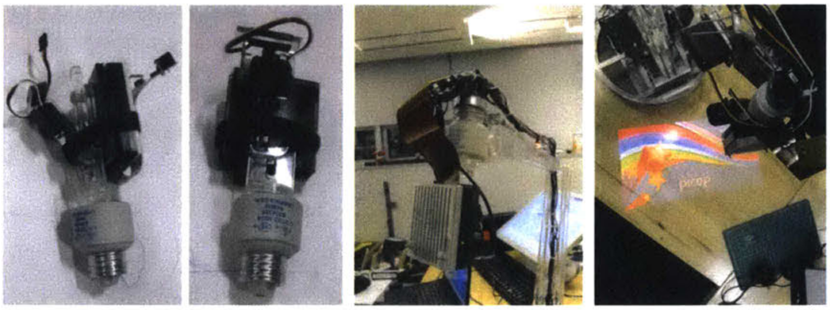


Figure 3-7 LuminAR Bulb early versions. Bulbs are fitted with Microvision PEK-1 laser pico-projector with a custom heat sink

We selected the Microvision PicoP™ Evaluation Kit v1 (PEK-1) as our projection engine–it has a standard VGA display interface, thin profile, and small mechanical dimensions (H 10mm x W 61 mm x L 68mm). It also uses a 5V power supply. Its main advantage is that it is a full-color laser projection device with a resolution of 848x480 (WVGA) with 16bit color depth and a 16:9 aspect ratio. It has a focal range between 200mm-2000mm; in this range the image is always in focus, ideal for a dynamic kinetic platform like LuminAR. In addition, our version of the PEK-1 was upgraded to a Class 3R Laser device, outputting 15 lumens, which was bright enough for many of the applications even in bright daylight conditions

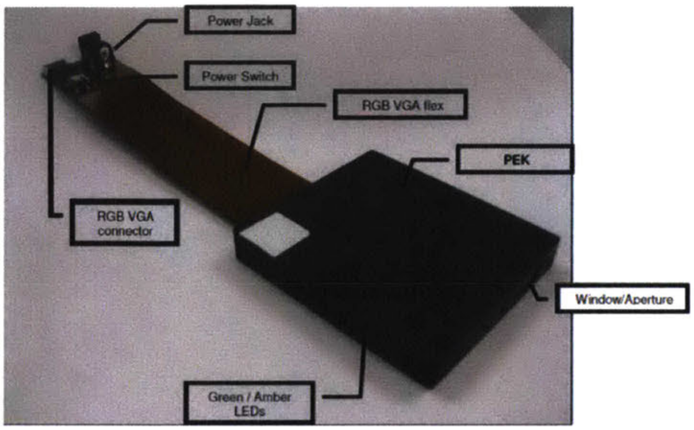


Figure 3-8 Microvision's PicoP™ Evaluation Kit v1 (PEK-1) (source: Microvision)

The PEK-1 has to be fitted with a heat sink element for ongoing operation. To conserve the bulb's small footprint, we replaced the PEK-1 original heat sink that extended the PEK's mechanical dimensions with our own compact design. Our heat sink was designed to fit the PEK-1 enclosure. It was fabricated from three 0.125" aluminum sheets with 11 ribs of approximately 0.1" inch wide.

Since we only had one PEK-1 unit, in some iterations, and for prototyping purposes, we also used Microvision's SHOW-WX Laser Pico Projector, which has similar specifications, except for a low-power laser–Class 2, which makes it a bit less bright.

We used a micro R/C servo fitted with a custom-made horn to provide support for the webcam. We selected the Microsoft LifeCam NX 2000 webcam, mainly as it has a very small footprint for both the optics and the electronics, but also as it was relatively inexpensive considering the quality of the images it provided (two megapixel).

LuminAR Bulb version 1 was mainly used in fixed light fixtures to facilitate software development, but was also the main bulb we used for the LuminAR Lamps Optimus and Aluminum models we describe later in this chapter.

3.3.2 Hardware Implementation

System Integration

The following two LuminAR Bulb iterations focused on subsystem integration; with a goal of developing a fully functional prototype, design details were intentionally left for later revisions. See details in Figure 3-9.

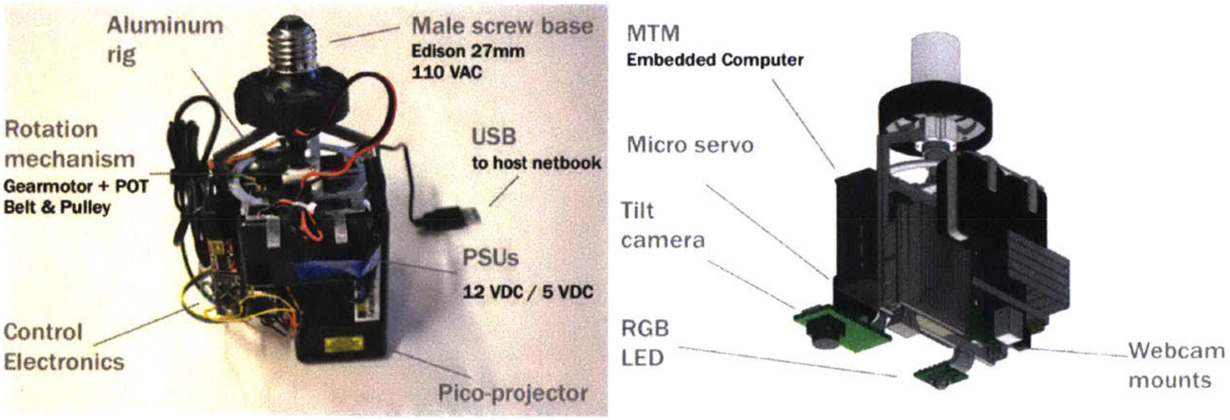


Figure 3-9 LuminAR Bulb integration version: on the right a view of the CAD design and on the left the physical prototype.

Power Circuitry

The bulb was designed for use with standard Edison sockets. To accomplish this, we had to incorporate AC/DC switching power supply units internally. The pico-projector and control electronics were supplied using a 5V/2A line. To support the embedded computer, we needed a 12V/1A line. USB peripherals were fed 5V directly from the USB lines.

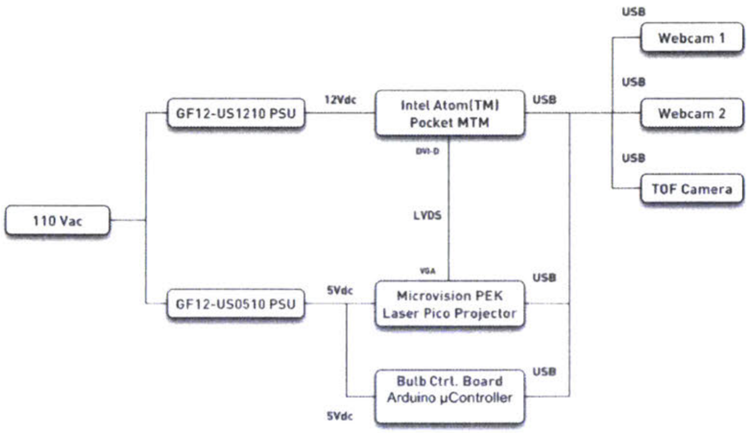


Figure 3-10 LuminAR Bulb power and connections schematics, The block diagram below shows a general overview of the bulb's components and their respective connections

To maintain the bulb's limited space constraints, we selected two small 5V and 12V PSU packages from the Go Forward GF 12 series (GF12-US0510 and GF12-US1210 respectively). The GF12 series were ideal, with mechanical dimensions of LxWxH:45.4x33.3x24.7 (mm).

Integration into the bulb included hacking the units to eliminate the original wall plug and jack connectors, leaving only the plastic case and power and ground wires. Keeping the PSU's plastic case also helped deal with the heat dissipation.

Lightweight Aluminum Rig

First, to support a complex integration of electronics, motors, cables, and components (i.e., webcam, embedded computer, etc.), we had to design a light yet strong rig that would be able to mount all the components in place, and provide us with a platform to test different configurations (see Figure 3-11). The rig was constructed from 0.125 aluminum sheets. The use of aluminum also allowed the frame to serve as an additional heat sink. On one side of the rig we attached the PEK-1, and the other side was reserved for an embedded computer.

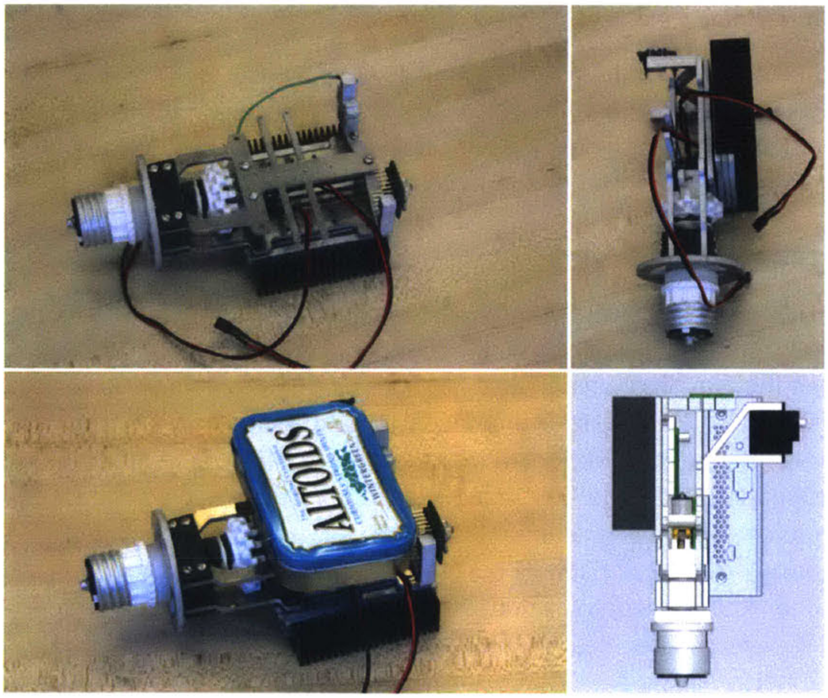


Figure 3-11 LuminAR Bulb integration version lightweight aluminum rig. This sequence shows on of the earlier designs, the Altoids box serves as a placeholder for the MTM Atom embedded computer. This rotation mechanism in this version used a central shaft solution.

Rotation Mechanism

The top section of the rig was used to house a rotation mechanism, which was coupled to the male Edison #27 bulb screw interface. The shaft is rigidly fixed to the screw through a flange. Managing the bulb mass and cabling made supporting pivot-axis rotation a difficult task. We tested two alternatives for the rotation mechanism.

Our initial approach used a central shaft coupled to a micro gearmotor mounted on the rig's top section. We used a Pololu Micro Metal gearmotor, which is extremely small and efficient and thus ideal for use in embedded systems. The motor ran at 5V with a gear ratio of 250:1, capable of an approximate torque of 4kg-cm. The motor was also coupled with a wheel optical encoder. The plastic ribbed wheel was designed to generate readings for the phototransistors on the encoder board. The encoder board was attached to the motor using a custom plastic bracket. The encoder was running at 5V providing 48 counts per revolution, a linear resolution of approximately 3mm.

This initial approach did not prove mechanically robust enough over continuous runs. Moreover, sensing the bulb's position proved challenging using the wheel encoder. However, the biggest hurdle was the size of the wheel encoder. This led us to replace both the rotation and the sensing mechanisms.

In our second approach, the rotational motion was generated by the same gearmotor and transmitted through a belt drive. The bulb's structure rotates below the belt around the central shaft. We selected a shaftless potentiometer mounted directly on the main shaft.

Control, Electronics, and Firmware

The bulb's electrical systems were designed to digitally control the rotation mechanism, sensing, LEDs, and additional servomotors. To do this, we programmed a small microcontroller system (Arduino Pro Mini 328 - 5V/16MHz) to interface with a small and commercially available motor driver board (Pololu Qik 2s9v1 Dual Serial Motor Controller). The board provided us with the means to control the motor's speed and direction. The rotation motion was monitored using a closed loop, continuously reading the potentiometer position. The feedback loop provided relatively accurate positioning in a range close to 350 degrees (taking into account the potentiometer dead zone).

Micro servomotors (Power HD Sub-Micro Digital Servo DS65HB) used to carry the additional webcam were interfaced directly to the microcontroller and controlled using standard PWM signals.

The bulb's firmware provided simple interfaces to control the bulb orientation and query its encoder position. This interface was used by the LuminAR application to define user interface projection orientation.

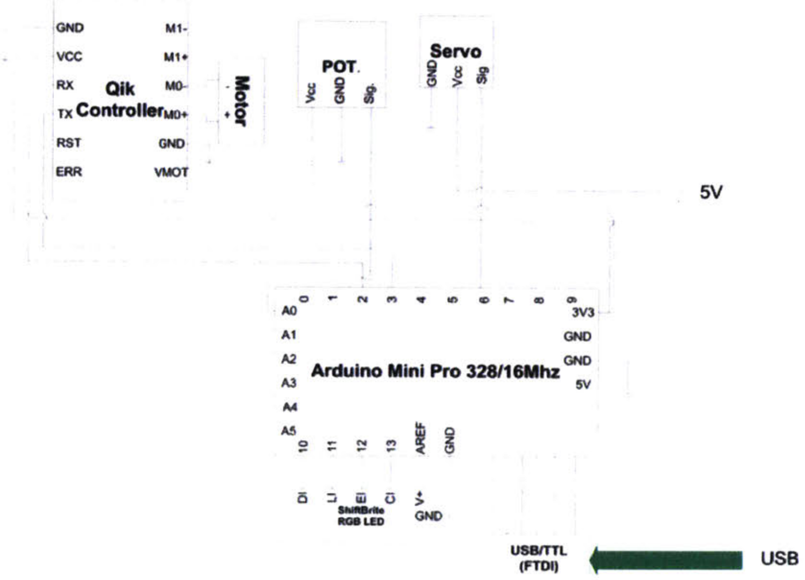


Figure 3-12 LuminAR Bulb control electronics board schematics

The bulb's host computer controlled the bulb's board using RS-232 serial communication over an FTDI USB connection. The diagram below shows the bulb's electrical schematics. This design was also used for future versions but its actual layout was changed to fit the physical dimensions of the new design.

Embedded Computer

The bulb's embedded computer was a pre-release version of Intel's Machine-to-Machine (M2M) reference platform **[116]** based on the Atom™ Processor (see Figure 3-13). The M2M comprises a carrier board and a Kontron COM Express compatible nanoETXexpress-TT Computer-On-Module. The system is packaged in an enclosure that measures 100mm x 67mm, slightly larger than an Altoids box, making it ideal for space-constrained hardware like the LuminAR Bulb.

The M2M supports wireless connectivity, including WiFi, and optionally additional radios (e.g., Bluetooth, ZigBee, or 3G/4G). It also has an accelerometer, dual HDMI display ports, HD audio, GB Ethernet, and

3x USB ports. The nanoETXexpress can carry fanless dual-core Atom Processors that run at 1.6Mhz. The system has one GB of memory and an external MicroSD card that can support upto 32 GB of storage. These specs enabled us to run the LuminAR software stack and perform computationally intensive tasks such as real-time control and computer vision, as well as communication and graphics rendering.

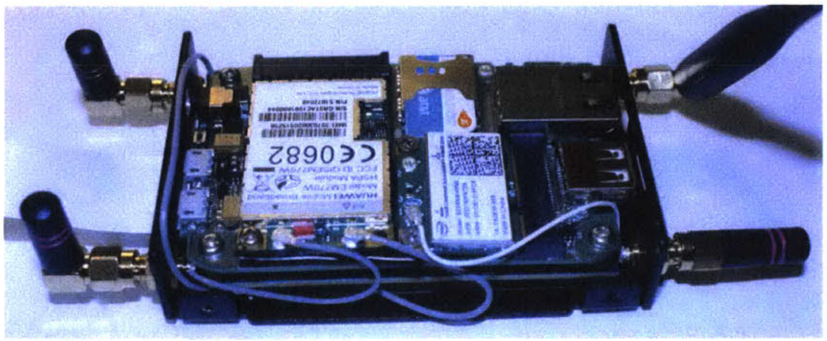


Figure 3-13 Intel Atom MTM platform

Unfortunately, even though our design took into account key embedded-system engineering issues such as heat dissipation, connectors, and mechanical mounting, we stumbled on a major roadblock in the final integration step. The MTM display adapter and the Microvision PEK display input were incompatible. The PEK expected an analog VGA input, while the MTM output was digital signal DVI-S. Given the time constraints of this phase of the research, we decided not to invest in developing our own convertor, and off-the-shelf convertors were simply too big for our design. However, we were aware that future display adapter formats (e.g., DisplayPort) could solve this issue in the future.

Our intermediate solution for development and demo purposes was to use Atom-based netbooks. We selected Asus Eee PC 1015PN and Eee PC 1215N, both Atom-based and with very similar specifications to the MTM. The constraint of having the actual computer external to the bulb mandated additional complexity when managing VGA and USB cables, which was particularly challenging in our final revisions.

Webcams

The bottom part of the rig was designed to carry a stationary webcam, and tilt webcam. The stationary webcam was aligned to the pico-projector aperture. Since the positions of the projector and the webcam were statically fixed with respect to each other, we could support easy semi-automatic calibration when installing the bulb in different locations.

The tilt camera was mounted on a small micro-servo and was positioned to serve both for interaction purposes, but also to capture the user's face or objects in the workspace. With a range of motion of approximately 90 degrees, the camera could be positioned down vertically to view the tabletop, or horizontally to view the space in front of the bulb.

Physical Configurations

We explored several different options for the bulb's component integration configurations. We tried to come up with different designs that would make the bulb feasible to use in different light fixtures. The domain of light bulb design is highly standardized, and we refer the reader to **[117]** for more information. We include some of our design configuration evolution in the Figure 3-14.

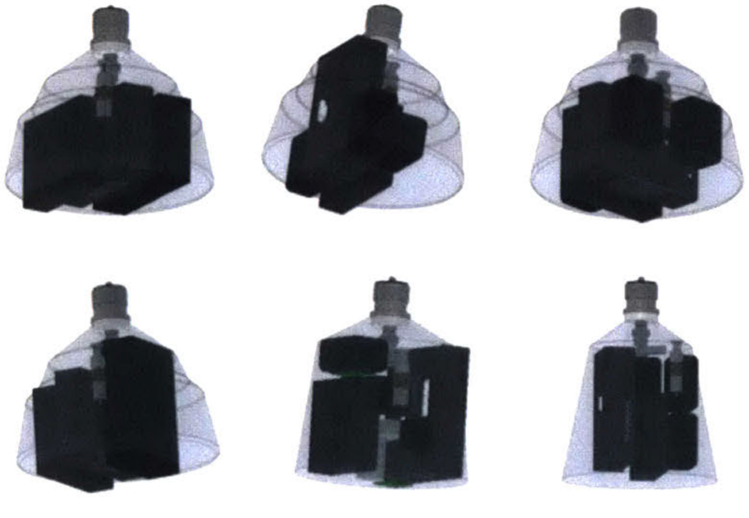


Figure 3-14 Exploring physical configurations for the bulb

3.3.3 Industrial Design Process

The final bulb iteration focused on compacting and streamlining the design and packaging the bulb in product-like covers. Based on our previous integration iteration, we were able to reduce the bulb size by approximately 30% without its covers.

Improved Rig, Electronics, and Cable Management

The rig of the final version was made of 1/16" aluminum sheets. The design follows a layered approach: spacers separate each layer of the frame. The electronics board design was improved to fit internally in one of the bulb's frame layers.

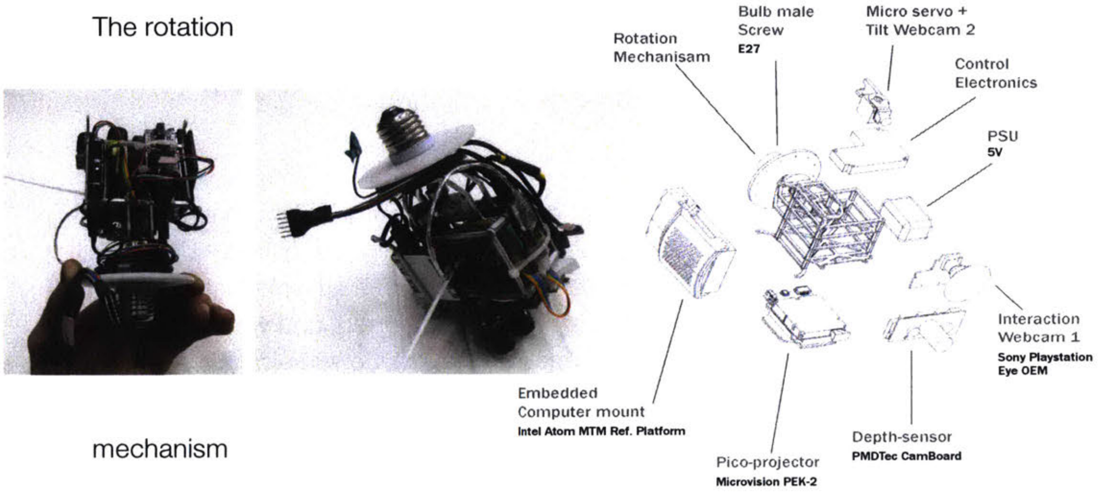


Figure 3-15 LuminAR Bulb – final version

The rotation mechanism was also redesigned for additional robustness. The shaft was reinforced with a collar. The pulley was also reinforced with additional thrust bearing to handle radial load.

Cable management posed an additional challenge, as we aimed to support close to 360 degrees of rotation of the bulb. To accomplish this, we designed a cavity in the top section of the bulb, under the E27 screw, that allowed the cables to rotate freely around bulb's main shaft.

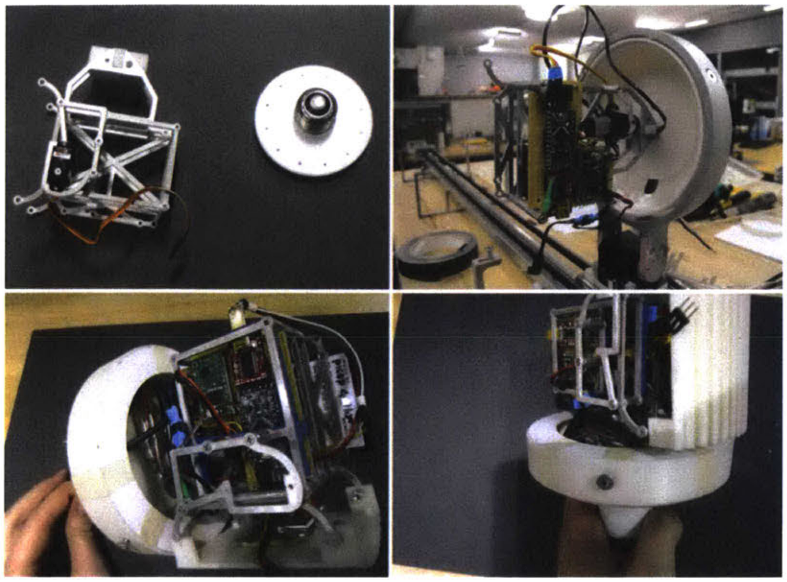


Figure 3-16 The LuminAR Bulb final lightweight aluminum rig design. The rig allowed efficient component mounting. It also took into account the need for efficient cable management. Cables were coiled in a special cavity created above the rotation mechanism, directly below the Edison bulb socket interface.

Even Smaller Projector

The final version of the bulb was designed to use Microvision's PicoP™ Evaluation Kit v2 (PEK-2) Laser pico-projector. The PEK-2 replaces the PEK-1 kit we used in the earlier versions. Its main advantage is its extremely small mechanical dimensions, of W x L x H 55mm x 68 mm x 21 mm. It also has a simpler display adapter interface with no additional Flex cable.

Adding Webcams and Sensors and Putting It All Together

On the bottom section of the rig we mounted webcams and sensors.

Sensor	Manufacturer/Model	Description/Purpose
Interaction Webcam 1	Sony Playstation Eye (PS3) OEM	A stationary webcam, aligned to the pico-projector aperture. Used to capture interaction gestures, capture imagery and detect fiducials.
Tilt Webcam 2	Microsoft LifeCam NX-6000	A rotating webcam, mounted on a servo. Designed to swivel 90 degrees

		between a position looking down, and a front facing position. This camera can be used for teleconferencing or track a user's face.
Depth sensor	PMDTec CamBoard	A USB based TOF sensor. Producing a depth map of 200x200 pixels. Capable of effective detection in a short range of 40cm. Used for hand gesture and multi-touch detection.

LuminAR Bulb Webcams and Sensors

The final design iteration of the bulb was specifically developed to match the LuminAR BlackJack and LuminAR Retail Arms (we describe these systems later in this chapter).

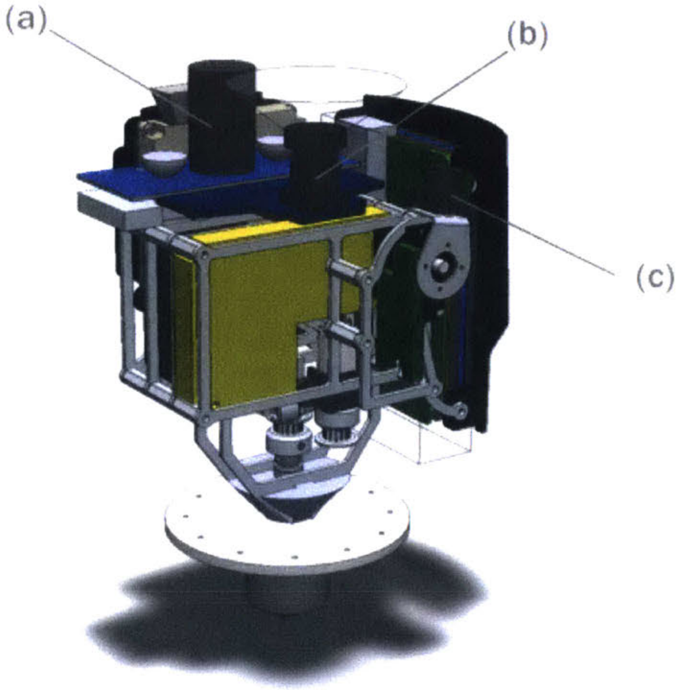


Figure 3-17 LuminAR Bulb sensors: (a) PMDTec CamBoard (b) Sony PS3 Eye Webcam (c) Microsoft Lifecam NX-6000

While dealing with demo deadlines and hardware integration pains, our efforts to fully integrate the PMDTec sensor with the bulb failed, most likely due to electrical ground noise in our design that damaged the PMDTec sensors. To work around this issue and meet our demo deadlines we created another version of the bulb using an additional Sony PS3 webcam instead of the CamBoard.

Designing the Shells

Finally, the bulb was enclosed in four shells: one fixed, attached to the main shaft flange, and the rest connected to the aluminum frame. The pico-projector and the onboard computer both have custom heat sinks used as part of the bulb shell.

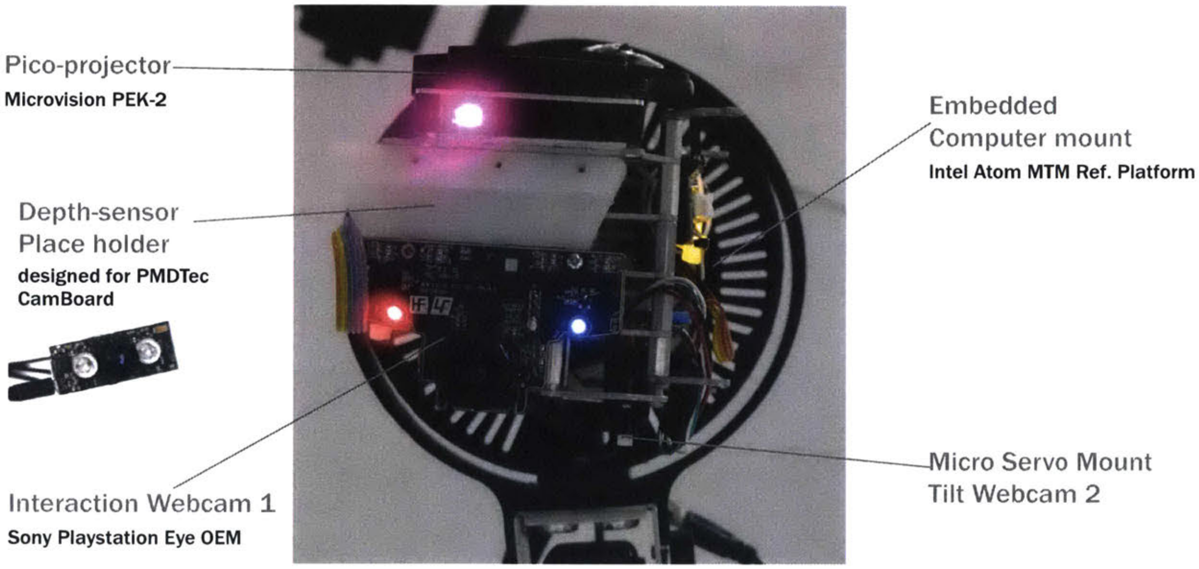


Figure 3-18 LuminAR Bulb webcams and sensors

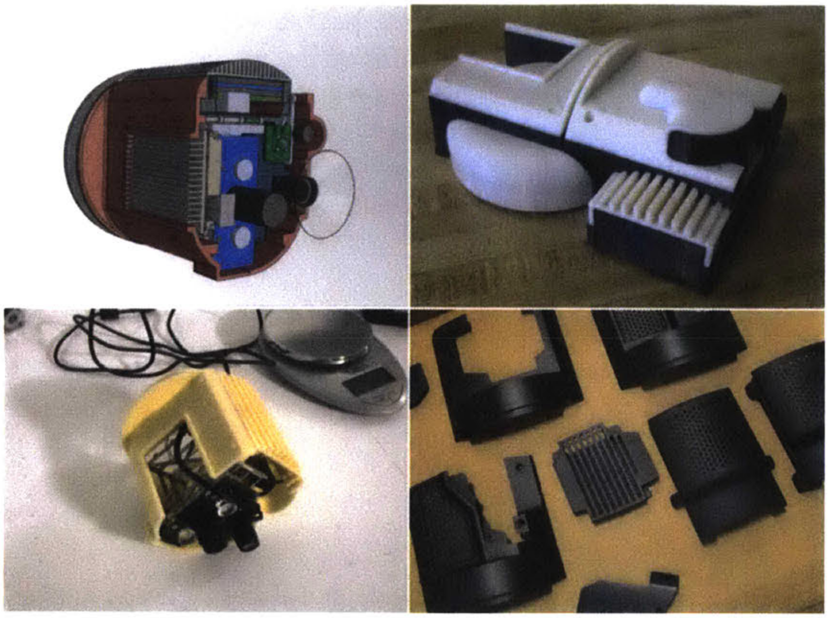


Figure 3-19 Steps in the design of the bulb covers included iterations of rapid prototyping using 3D printed covers. Final covers were fabricated from ABS plastic.

The shells were designed in two iterations; we 3D printed the initial design and completed basic fit testing to the frame. Later on, the design was refined and was fabricated from ABS plastic.

Figure 3-20 Steps in constructing the final prototype of the LuminAR Bulb

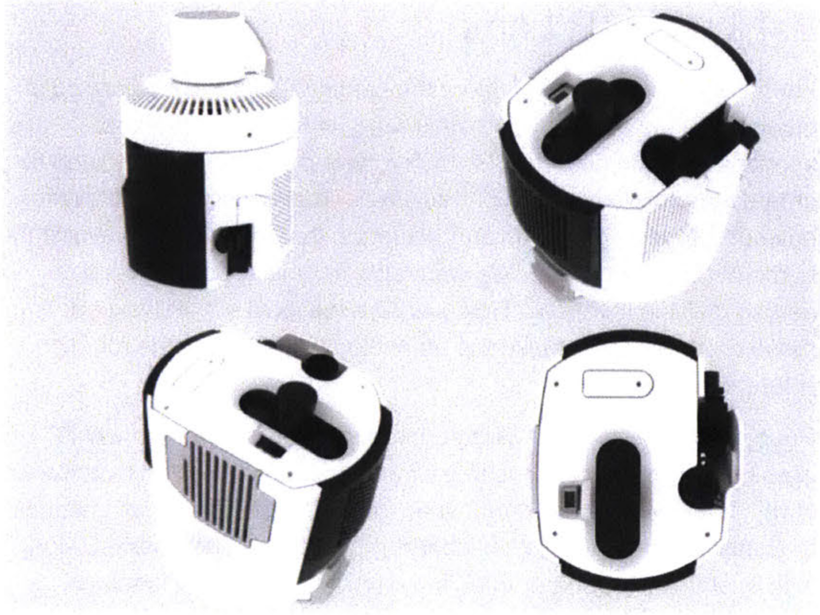


Figure 3-21 Renders of the final design for the LuminAR Bulb (rendering: Jason Robinson)

Steps in the design of the bulb covers included iterations of rapid prototyping using 3D printed covers. Final covers were fabricated from ABS plastic.

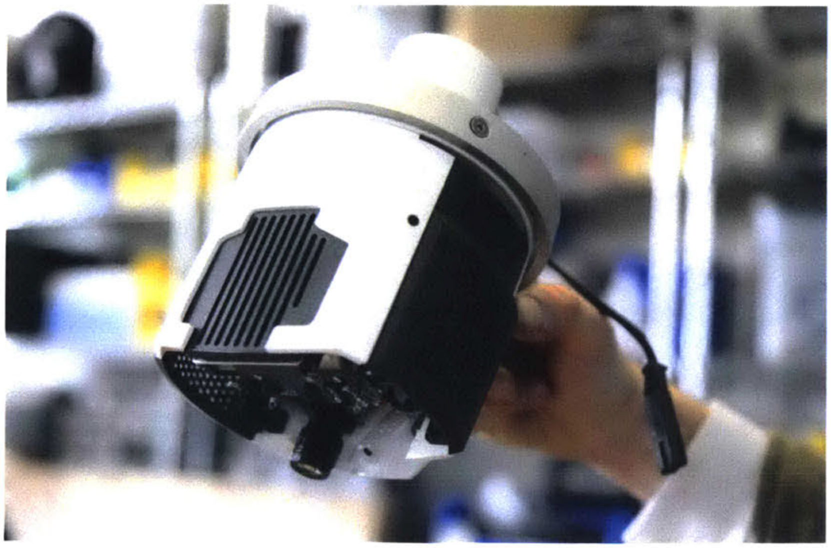


Figure 3-22 The final LuminAR Bulb

3.4 LuminAR Lamp

For the sake of formal definition, the LuminAR Lamp is composed of custom robotic arm (i.e., LuminAR Arm), which is designed to interface with the LuminAR Bulb. The two together, are combined to create a motorized version of the classic Anglepoise task light. We have developed six robotic arm platforms that were used to evolve the LuminAR Lamp concept. We were able to quickly use our insights across multiple iterations. However, our real goal in the design of these objects was to make the technology, specifically the robotics, disappear altogether.

Poupyrev et al. defined Actuated Interfaces as interfaces in which physical components move in a way that can be detected by the user **[118]**. They also specified that such interfaces could employ changes in spatial position of objects, including orientation and position, along with changes in speed or direction of motion. The early LuminAR prototypes introduced approaches to such actuated interfaces. These explorations helped to refine the concept of Dynamic Multi-Touch. Dynamic Multi-Touch is an extension for the classic gesture vocabulary of multi-touch that takes advantage of a projected kinetic I/O system. Dynamic Multi-Touch systems utilize actuated DOFs of the projected touchscreen display to support real-time relocating, reorienting, and resizing of the projected display. Dynamic Multi-Touch attempts to extend the spatial limits of the classic screen-bounded interface by allowing it to move. It also addresses the interaction space above the display.

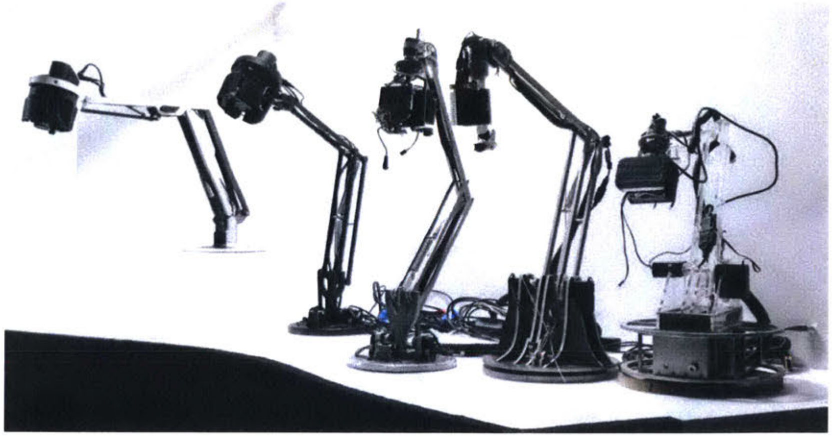


Figure 3-23 The evolution of LuminAR Lamps. Five generation of hardware starting with the earliest iteration on the right and the last version on the left. (photo credit: Doron Gild)

Poupyrev et al. defined Actuated Interfaces as interfaces in which physical components move in a way that can be detected by the user

[118]. They also specified that such interfaces could employ changes in spatial position of objects, including orientation and position, along with changes in speed or direction of motion. The early LuminAR prototypes introduced approaches to such actuated interfaces. These explorations helped to refine the concept of Dynamic Multi-Touch. Dynamic Multi-Touch is an extension for the classic gesture vocabulary of multi-touch that takes advantage of a projected kinetic I/O system. Dynamic Multi-Touch systems utilize actuated DOFs of the projected touchscreen display to support real-time relocating, reorienting, and resizing of the projected display. Dynamic Multi-Touch attempts to extend the spatial limits of the classic screen-bounded interface by allowing it to move. It also addresses the interaction space above the display. In our work, we have prototyped several gestures that explore the Dynamic Multi-Touch concept. We describe the Dynamic Multi-Touch interactions in more detail in Chapter Five. The core principles of Dynamic Multi-Touch used in the LuminAR system were described in the recently granted US patent Kinetic input/output [43].

3.4.1 Design Goals and Process

The design of the LuminAR Lamp is an attempt to embed computation and interfaces in everyday objects. The challenge was in hiding the technology and making it invisible to the user. To do so we had to reinterpret the classic Anglepoise design, evolving it into new yet familiar interactive object. In fact, we can think of the lamp as a new form factor for a computer that uses projected augmented reality and natural interaction techniques as its main UI modality.

Much like the LuminAR Bulb design, the LuminAR arm development process was iterative. In the early stages of the project we employed an agile approach, pushing for very short prototyping iterations. The lessons from one prototype generation fed the next. Overall, we developed and tested six different arm prototypes.

The Anglepoise Type 75 Reference

Figure 3-24 Anglepoise Type 75 (source: Anglepoise)

In order to translate our design into an actual mechanically operational system we decided to use the Anglepoise Type 75 (see Figure 3-24 and 3-25) as our main study model. Early on, the LuminAR arms follow its proportions and to some extent some of its aesthetics.

Figure 3-25 Testing of the first LuminAR arm proportions compared to an Anglepoise Type 75 Lamp

Mechanical System

Before we dive into the description of the various LuminAR Lamp prototypes, it is important to review the basic mechanical operation principles they all share.

LuminAR follows the classic Anglepoise counterbalanced arm design **[119]**, where the force of gravity of the bulb and lampshade is counteracted by extension springs. This allows static positioning of the arm within its effective reach envelope. The arm uses a four-bar linkage system–a parallelogram–to translate the movement of the upper arm of the system to a small linkage closer to the base. This enables the counterbalancing spring and any actuator for the upper arm to apply actuation on the small bar and, as a result, be positioned lower, closer to the base.

For a complete review of the physical, mathematical model and mechanical working principles that govern counterbalance systems we refer the reader to the excellent review by French et al. **[120]**.

Topology

The arm itself has four rotational degrees-of-freedom (DOFs). The first DOF is located in the base of the lamp, allowing the Base-DOF to rotate in the horizontal plane, in a range slightly less than a full 360 degrees. Next, the DOFs are named accordingly with respect to human arm metaphor: Shoulder-DOF, Elbow-DOF, and Wrist-DOF are located on a single plane, orthogonal to the Base-DOF (see Figure 3-26).

The Shoulder-DOF and Elbow-DOF are positioned directly above the base on a mechanical fork structure. The fork is a support structure for the servo motors and shafts that actuate the elbow and shoulder links of the arm. Finally, the Wrist-DOF is located at the end of the Elbow-link, though the actuator can be mounted elsewhere in the structure of the lamp.

The Shoulder, Elbow, and Wrist DOFs are spring loaded, each with its own separate extension coil spring. The springs counter the mass of the links the arm's payload, and the LuminAR bulb. Actuating the balanced arm required the output of enough torque to overcome the friction load. The Wrist-link has an interface for the Edison 27mm female bulb socket. Overall, the lamp has five rotational DOFs, including the additional rotation inside the Bulb.

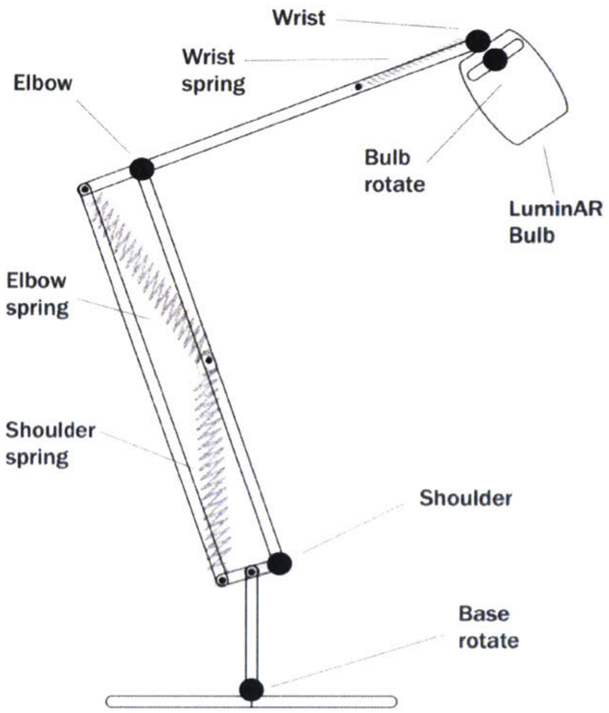


Figure 3-26 LuminAR Lamp mechanical system topology, the diagram shows the lamp's DOFs and extension springs.

Safety and Proactive Behavior

We designed the LuminAR Lamp counterbalanced arm for desktop environment. As a kinetic interface, with motion paths that are generated by relatively powerful servo-motors, it is essential to ensure users safety when interacting with the system. To achieve the required safety baseline, we designed the LuminAR Lamp mechatronic to operate at human-equivalent speeds, which are inherently limited and hence safe to operate alongside humans.

The requirements for robust safety mechanisms becomes crucial if the LuminAR Lamp is designed to support proactive behavior. Currently, the system only supports user programmed motion routines. As a result, the kinematic behavior and motion path are relatively predictable. Proactive behavior is naturally less predictable.

However, currently, our system is an experimental lab prototype. As such, it lacks real product safety requirements. Specifically, any kinetic interface would require mechanisms for contact detection to avoid sudden impact with humans or objects. Typically, such safety mechanism can be implemented with a combination of sensing hardware and robotic behavior software that are able to act on the sensed state, control and stop the robotic system if needed, as well generate cues to warn the user ahead of time.

3.4.2 Hardware Implementation

In following section, we describe the different LuminAR lamps developed over the course of this dissertation work.

LuminAR Optimus

Optimus was our first fully functional prototype (see Figure 3-27). Its main purpose was to serve as a learning platform for the mechanical subsystems and begin exploring concepts of kinetic user interfaces.

Optimus had a heavy round base that was constructed from a heavy 0.5" MDO board, fitted with an aluminum construction (0.25" aluminum sheets). The base rotation mechanism used a central stationary shaft fitted with a gear. A servomotor with a 1:2 gear ratio was engaged to the central shaft gear, giving it 180 degrees of rotation. On top of the base structure we designed an acrylic structure that supported the elbow and shoulder servomotors. The elbow and shoulder springs extended from the four-bar arm linkage to a screw-bolt holder in the bottom of the base. The arm was also constructed from acrylic laser-cut parts that were press-fitted and glued.

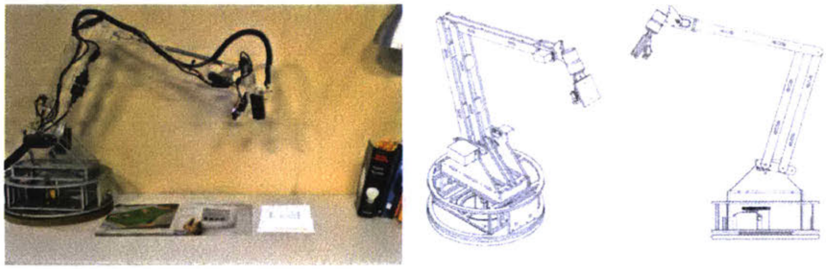


Figure 3-27 LuminAR Optimus

Optimus had a bulb socket construction, but it did not provide power. One of the early LuminAR Bulb rough prototypes was installed on Optimus. Wiring was done through the gap between the arm bars, making sure the cable had enough bend profile. This problem persisted and always required attention in all of the prototypes we created.

Optimus was used extensively in our concept mock-up scenarios, and some of the demos we created are available as a part of the MIT Media Lab LabCast video series [121].

LuminAR Aluminum

Our second prototype was named Aluminum (see Figure 3-28), as it was the main material used for this prototype. The transition to metal was trivial, as wood and acrylic did not provide the robustness and accuracy desired for the kinetic interaction we envisioned. Aluminum also had a refined low-profile base.

Figure 3-28 LuminAR Aluminum

The base was implemented using a simple turntable and a central gear. The servomotor mounting structure and the actual arms were also streamlined, pushing towards a design that resembles a real Anglepoise task light.

Actuation, Control, and Early Software Versions

Both Aluminum and Optimus were actuated using simple R/C servomotors. To control them, we used the commercially available servo control board Maestro Micro from Polulu. Major advantages of the Maestro Micro are its tiny size and simple USB-based programming interface. Aluminum and Optimus were both connected to a netbook running Linux (Ubuntu 10.4). Control was done using RS-232 over USB. Projection was supported directly using the netbook's VGA port. To begin the LuminAR software stack design, we started with quick prototypes using a Python-based control server, while our frontend UI was based on the QT4 GUI toolkit.

LuminAR Tipsy

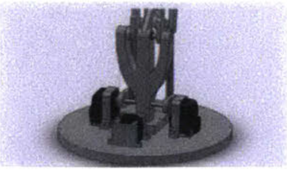


Figure 3-29 LuminAR Tipsy has a new central for element, which enabled us to position the motors on the base plate

Tipsy represents a major revision in many aspects of the LuminAR Lamp evolution (see Figure 3-29). Anecdotally, Tipsy got its name due to a miscalculation of the lamp's base weight, which caused it to tip and lose balance. Even though far from perfect, Tipsy represents our attempt to reach the proportions of a real task light. It also represents a system integration milestone, as it was tested with an almost fully functional LuminAR bulb. Based on our experience with LuminAR Aluminum, we were able to introduce important improvements to the mechanical and product design.

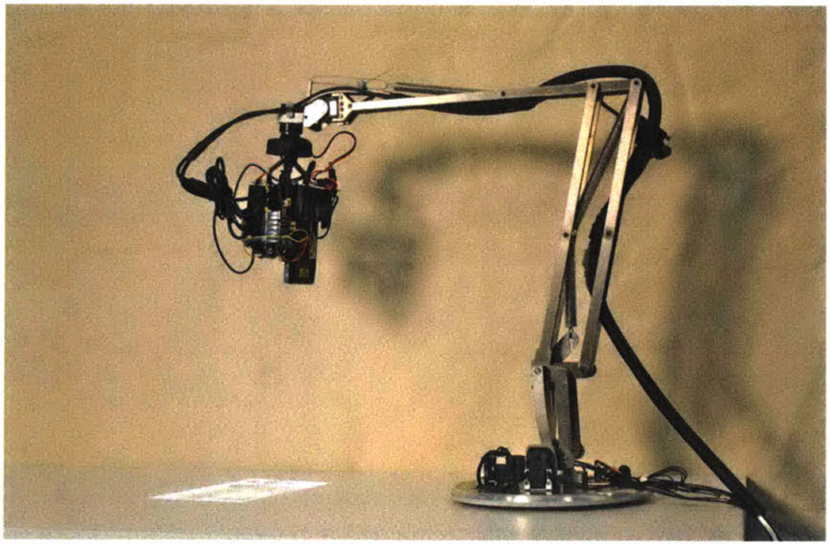


Figure 3-31 LuminAR Tipsy

First, we improved on the base rotation mechanism. Moving away from the use of turntables, we designed our own ball-bearing system; doing so allowed us to reduce the base diameter further.

Next, we designed a new central shaft and a new fork element. The fork allowed us to align and mount the elbow and shoulder servomotors directly on the top of the base plate. This approach saved redundant use of aluminum wings above the servos. To implement this, we added custom servo horns and linkages that provided the actual hinge to the elbow and the shoulder DOFs respectively. It also allowed using a single central shaft in the fork structure.

Figure 3-30 Configurations we created based on Tipsy's design

Finally, we were able to fit the springs for the elbow and shoulder DOFs elegantly in the gap between the lower arm linkages (see Figure 3-30). Mounting the springs there contributed to the overall range of motion as well as to the overall form of the lamp. As the design of Tipsy was clean and simple we could easily explore different configurations using the same basic platform.

Actuation, Electronics, and Control

Tipsy was our first model to incorporate a power supply. 110VAC was supplied to the bulb socket attached to the Wrist, while 15VDC was supplied from a PSU mounted on the base. It was used to feed the arms' servomotors. A small USB-Dynamixel board mounted on the power supply supported the communication to the bulb encoder. In our original design, we also included external position sensing for each DOF. This was to be accomplished using potentiometer that were mounted on the DOFs shaft. All encoders were then connected to a small microcontroller at the base of the lamp.

Tipsy was actuated by four Robotis Dynamixel servos (DX-117). We decided to upgrade the arm's actuators after we damaged or broke several giant scale R/C servos (e.g., Hobbico CS-80). Although clearly more expensive, the Dynamixel servo system [122] has clear advantages in sensing, networking and torques compared to standard R/C servos.

Tipsy was the first arm to be fully controlled by our very own integrated LuminAR bulb version. It was also the first platform on which we developed and tested the LuXor software stack that we describe later in this chapter.

LuminAR SilverJack and BlackJack

One of the main goals of this dissertation work was to embed computation in everyday objects. The main innovation of the SilverJack and BlackJack revision (see Figure 3-32), together with the new LuminAR bulb, is the creation of a new actuated interactive robotic task light. In this iteration, we focused on the industrial and product design details that pushed our prototype further, so it is a 'designed object' and one that could be perceived as such by users.

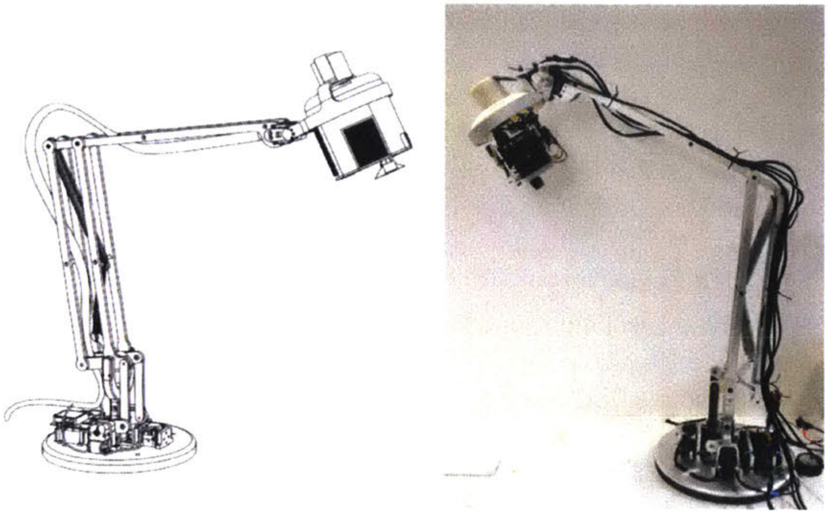


Figure 3-32 LuminAR SilverJack and Blackjack. (photo credit: Doron Gild)

SilverJack and BlackJack were heavily based on the Tipsy design. They were designed for short-run professional manufacturing. The use of professional manufacturing techniques clearly upgrades the final result in terms of product design but also impacting accuracies and tolerances, which has a direct effect on the overall mechanical performance of the system.

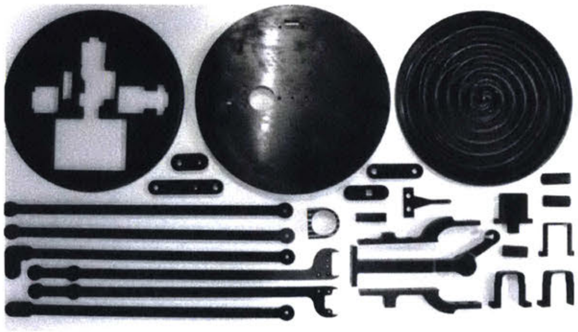


Figure 3-33 LuminAR Blackjack fabricated parts before assembly

The SilverJack and BlackJack revisions also included mechanical improvements. The main fork was revised to provide better reinforcement for the servomotors. The base rotation mechanism was also improved to use anti-backlash gears and a central shoulder screw.

Figure 3-34 LuminAR Blackjack. (photo credit: Doron Gild)

LuminAR Retail

The LuminAR Retail arm design was specifically developed to integrate into the Augmented Product Counter (APC – see: Chapter Six: Applications). This revision explored how our design could be manifested in a future retail environment.

Conceptually, LuminAR Retail is an adaptation of the LuminAR BlackJack tabletop arm design to a countertop lamp for use in public spaces. In addition, the arm's design had to take into account the physical dimensions of the installation space, and to account for the APC projection real-estate requirements.

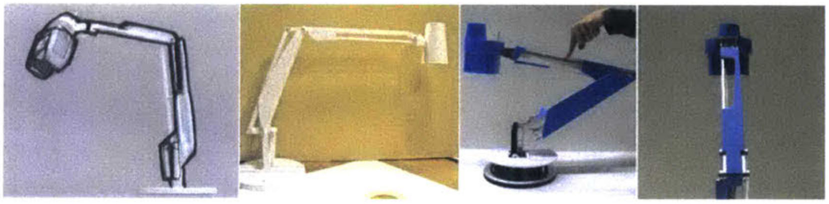


Figure 3-35 LuminAR Retail design process

LuminAR Retail design process used existing lamps as platform for physical modeling and form exploration. To accomplish that, our design process used the existing SilverJack model as a physical prototyping platform. Using this approach, we were able to reuse most of the base components, while developing the new proportions and working envelopes that were extended to cover larger workspace.

In order to streamline the final form of the retail arm we repositioned the wrist servo, locating it in the back section upper arm linkage. The result was a clean and simplified connection point between the bulb interface and the upper arm wrist joint. The motion was transmitted to the wrist DOF using a custom-made horn and linkage that was mounted under the upper arm. In addition, the retail arm makes extensive use of casing and covers. We used black ABS plastic custom-made covers to obfuscate technical components as much as possible. In the final version, we included covers for the arm linkages (top and front), bulb interface and base covers.

Due to its larger size and mass, to maintain an effective work range, we had to upgrade the servomotors for the Shoulder-DOF and Elbow-DOF from Dynamixel DX-117 to Dynamixel EX-106+, which provided the required torques to handle LuminAR Retail's increased dimensions and weight.

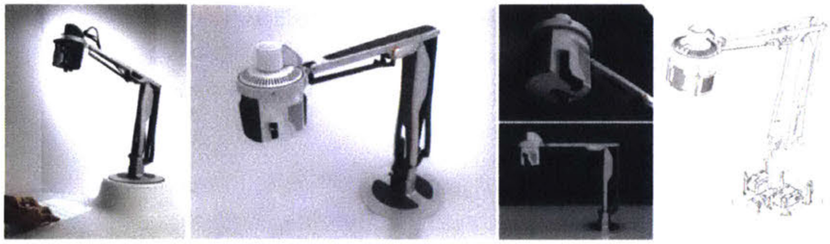


Figure 3-36 LuminAR Retail (photo credit: Doron Gild, rendering: Jason Robinson)

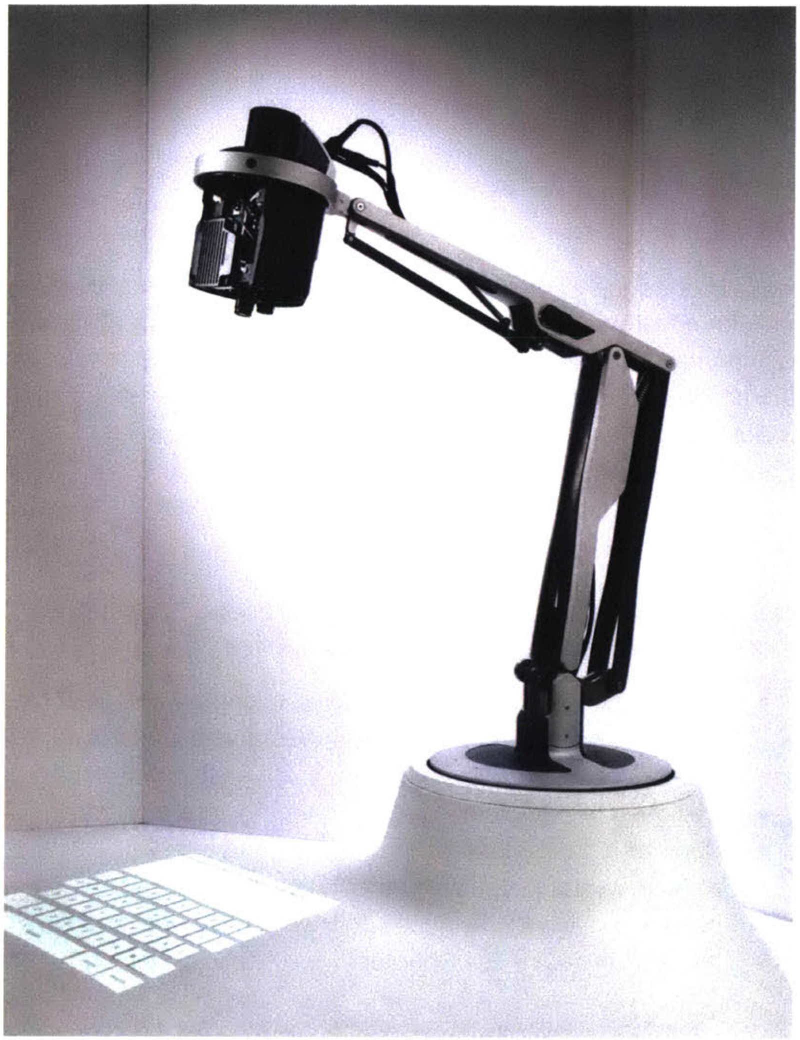


Figure 3-37 LuminAR Retail projecting a keypad (photo credit: Doron Gild)

3.5 LuminAR Spotlight

LuminAR Spotlight is a ceiling- or wall-mounted, actuated carrier fixture for the LuminAR Bulb (see Figure 3-38). It was designed to add articulation capabilities to static track lights and light fixtures. Combined with the projected augmented reality interface that the LuminAR bulb provides, LuminAR Spotlight creates dynamic interactive physical spaces.

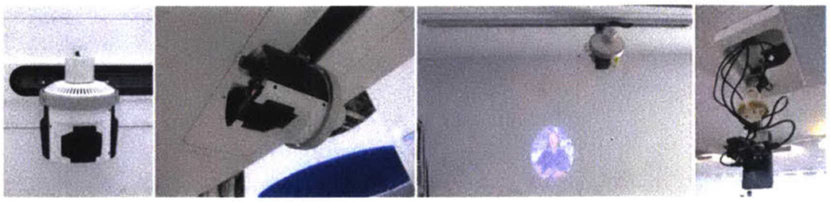


Figure 3-38 LuminAR Spotlight and Spotlight360

Our goal in designing LuminAR Spotlight was to develop additional kinetic form factors, that can be embedded seamlessly in a physical space and explore the interaction techniques it affords. In this section, we describe the Spotlight prototypes and discuss two variations of the Spotlight concept: the first uses an articulated linear track, and the second variation uses a pan/tilt mechanism.

3.5.1 Design Goals and Process

To complement the general design concept of spotlight, we defined the following design goals for the Spotlight system:

- Can be mounted on a ceiling or wall, taking advantage of the additional height and projection angles above and around a workspace;
- Designed to completely blend into the physical environment;
- Support natural interaction modalities; and
- Able to communicate with other LuminAR devices, complementing the use case (this was actually achieved in the APC project – see Chapter Five – Applications).

3.5.2 Hardware Implementation

LuminAR Spotlight LS

Our first approach explored the concept of actuated track lights. We focused on building a quick prototype to be able to tackle challenges like actuation and retractable cable management. The Spotlight LS (see Figure 3-39 and 3-40) system consists of a linear, shaft-based track and a carriage with an R/C servo motor tilting head, giving LuminAR Spotlight two DOFs. The carriage head interconnects with the LuminAR bulb through a standard Edison bulb screw socket.

In this version, motion was transmitted from a stepper motor through a fast lead screw that carried the bulb carriage. It had no position sensing except far and near switches at the end of shaft tracks. We used a small microcontroller and motor driver board to open-loop control the position of the carriage. The microcontroller was then hooked to the LuminAR Bulb netbook computer, giving the bulb control of its location along the track.

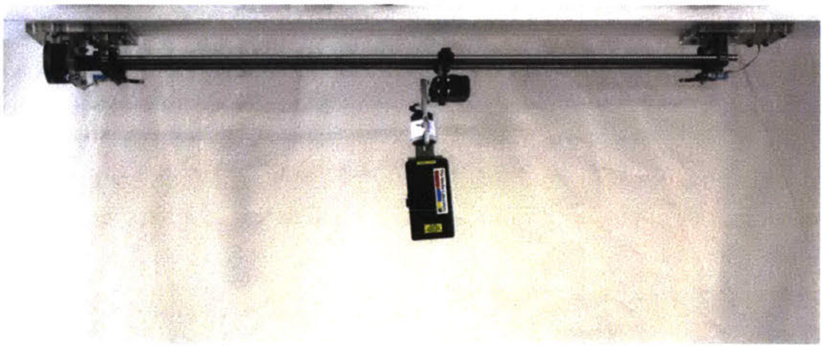


Figure 3-39 LuminAR Spotlight LS

The carriage position could be estimated using a time-based calculation. The actual location of the carriage was calculated using rotation revolution count. Integrating the lead screw pitch with the number of steps the stepper motor performed in a given time slot provided the estimated distance the carriage travelled.

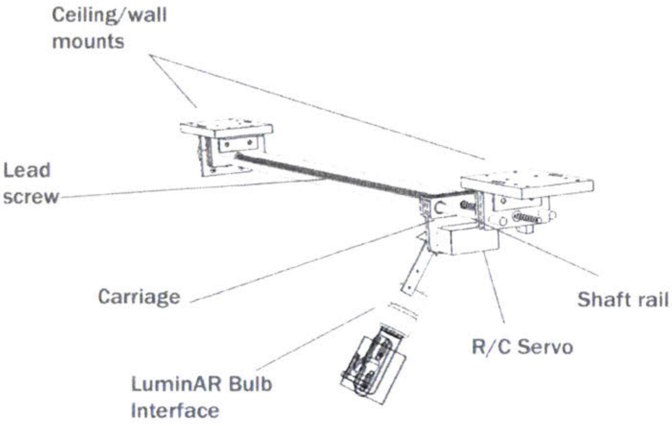


Figure 3-40 LuminAR Spotlight LS - mechanical design

However, this approach gave average results and was one of the reasons the use of a lead screw was abandoned in the next version, the Spotlight BD (belt drive). The other reason we opted to replace the lead screw actuation method were the slow speeds it provided and the high noise levels.

Interaction with Spotlight LS

Spotlight LS also served as a platform to explore different interaction techniques that we already developed for the LuminAR Bulb, and specifically new methods that are unique to Spotlight.

First, we were able to dynamically control the location and orientation of a touch-enabled projected interface. This concept could be used to dynamically choose the projection location. For example, we could decide on a horizontal surface (e.g., a tabletop), and then using a gesture, direct the Spotlight to project on a vertical surface (e.g., a wall).


(a) (b) (c)

Figure 3-41 LuminAR Spotlight interaction techniques include: (a) projected mutlitouch, (b) hand tracking and, (c) tracking fiducial markers

Next, we could have the Spotlight track a fiducial marker. This is used to track an object under the spotlight. We were also able to use similar techniques to track a user's hand.

LuminAR Spotlight BD

Following our early general explorations with Spotlight, we started working on a specific version of the Spotlight for the Augmented Product Counter project. From an interaction point of view, this version was designed to provide an 'Expert Wall' feature; we describe this scenario in Chapter Six – Applications.

The Spotlight BD prototype (see Figure 3-42) represents a major upgrade of the mechanical and electrical systems. No less important, we were able to meet one of our design goals and develop a solution that integrates into a ceiling cavity, hiding all the technology under the hood.

Mechanical Design

Similar to the Spotlight LS, the system consists of a linear track and a carriage with a tilting head, giving LuminAR Spotlight BD the same two degrees of freedom.

The linear motion is generated by a stepper motor, positioned on one end of the rail and transmitted through a belt and pulley drive mechanism. The idler pulley shaft is equipped with a rotational encoder to track the carriage position. The belt and pulley proved a better solution, in terms of ambient noise, and position sensing control accuracy.

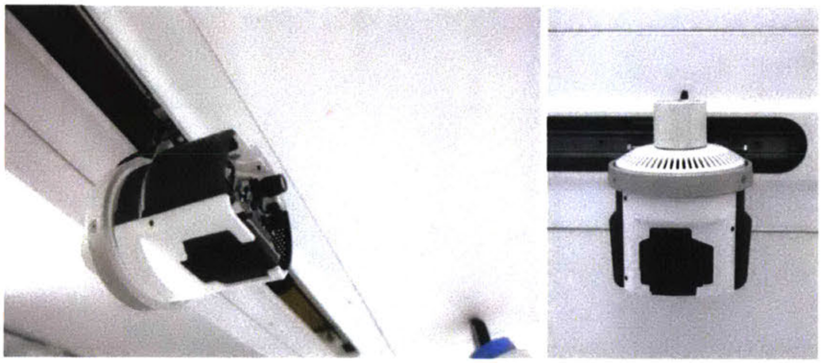


Figure 3-42 LuminAR Spotlight BD

The carriage head tilt is accomplished by a servo mounted on the carriage. The servomotor was upgraded to use a stronger Robotis Dynamixel AX-10 servomotor. The servomotor was fitted with a standard bulb interface.

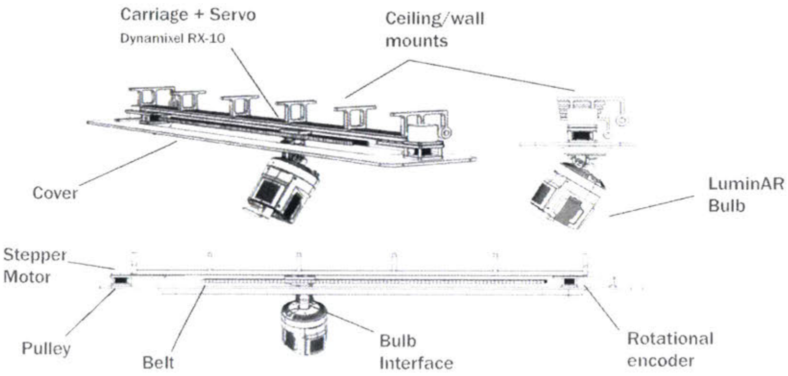


Figure 3-43 LuminAR Spotlight BD – mechanical design

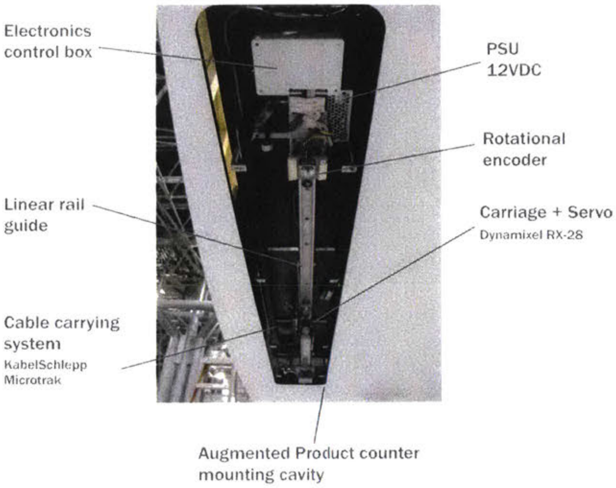


Figure 3-44 LuminAR Spotlight BD – integrtion

In this version, we had to provide a better solution for continuous retractable cable motion. We integrated a KabelSchlepp Microtrak cable carrying system. The Microtrak was an ideal choice as it is lightweight and non-conductive, yet very durable and easy to integrate.

Electrical Engineering

We created a custom control electronic board for the Spotlight BD. It was important to create the board so the entire system could be used stand alone, requiring only power connection and the LuminAR bulb to operate. The entire system was fed by a power supply that provided motors with a 12VDC line, and a 5VDC line for the control electronics.

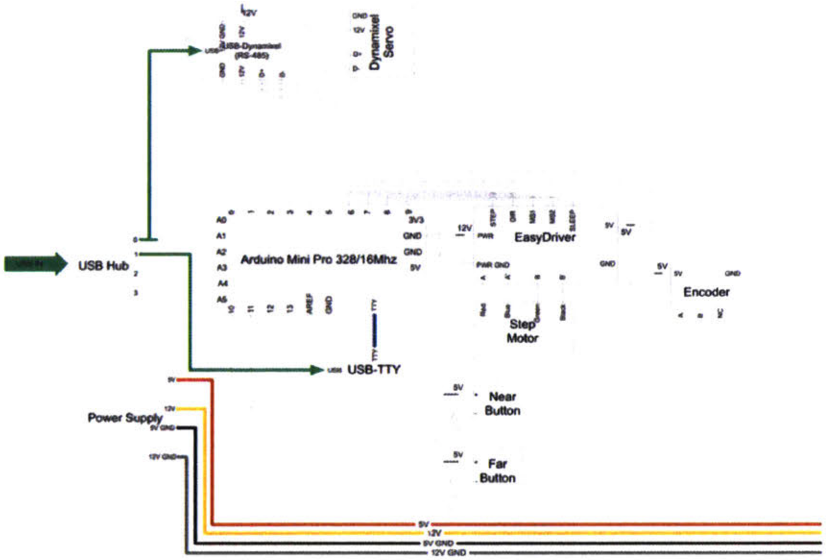


Figure 3-45 LuminAR Spotlight BD control electronics board schematics

The board controlled the stepper motor using a small microcontroller (Arduino Mini Pro 238/16Mhz), connected to a commercially available stepper motor control board from Sparkfun called EasyDriver. The EasyDriver board provided a simple control interface to the motor's step and direction. The near and far switch sensors were connected directly to the microcontroller, providing a stop-switch mechanism when the carriage reaches the end of the linear track. The rotational encoder is also connected directly to the microcontroller, providing a closed-loop control.

The board included a USB hub that provided interfaces for the FTDI RS-232 / USB interface boards, used by the microcontroller and the Dynamixel servo communication interface. The board was placed in a plastic box, mounted on the linear track rig, and fitted with ON/OFF LEDs for indication purposes. The microcontroller on the board ran firmware that provided simple interfaces to initialize the system control and query the location of the carriage on the track.

LuminAR Spotlight360

The final Spotlight exploration we created is called Spotlight360 (see Figure 3-45 and 3-46). It is a simple fixture, built from two Dynamixel DX-117 servos, the first serving as a rotation axis (pan) and the second serving as a tilt servo. The second servo interconnects with the LuminAR bulb through a standard Edison bulb screw socket.

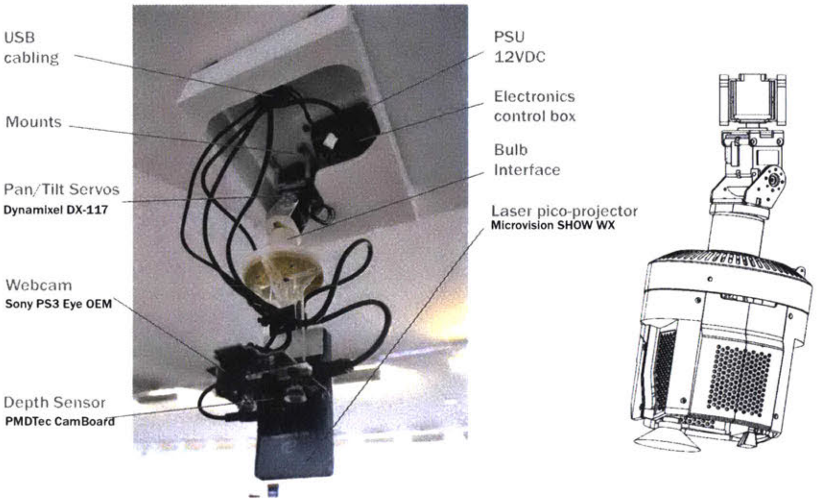


Figure 3-46 LuminAR Spotlight360 – mechanical design

The goal of this version was to provide a simple form factor capable of revolving the LuminAR bulb 360 degrees with respect to its mounting point. In this specific iteration, we used a stripped version of the LuminAR bulb that used a depth sensor. This enabled us to use free hand gestures (e.g., open hand, fist, pointing-finger) in the bulb's field of view to control the bulb projection location and orientation.

3.6 LuminAR Projector

The LuminAR Projector (see Figure 3-47) hardware was designed using off-the-shelf components, allowing for a scalable and inexpensive system. The system integrates computation, sensing, and projection into a single unit. The main processor is the P910 from VIA, an x86-compatible, four cores, 1GHz Pico-ITX form-factor motherboard. The embedded system handles both computer vision and user interface display. A commercial 500 lumen projector was modified to fit the unit's custom housing. The projector is capable of 720p HD video display at a 16:9 aspect ratio.

The overall unit is compact, measuring 7.5" W / 3.5" H / 6.5" D. The projector was coupled with a PrimeSense Carmine 1.09 sensor, providing the system with a 640 x 480 (VGA) depth map at a spatial (X,Y) resolution of 3.4mm-0.9mm and depth resolution (Z) of 1.2cm 0.1cm at 60 fps. The device uses a standard VESA 75mm mounting pattern on the bottom, allowing the unit to be easily installed on various industrial mounting systems.

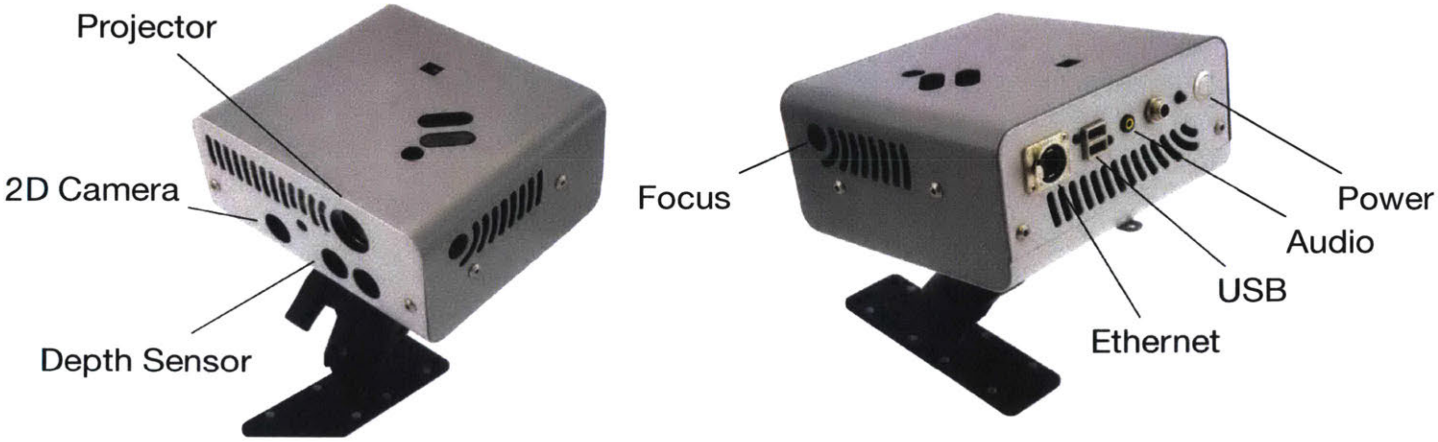


Figure 3-47 The LuminAR Projector component. Front view (left); Back view (right)

3.7 Conclusion

In this chapter, I presented the family of LuminAR devices we designed and developed. It is my hope that our experience and documentation of the hardware development process will inspire the next generation of light-based interface designers. I exemplified through the various prototypes the validity of the approach to embed projected AR interfaces in common-place objects.

First, we developed a functional bulb prototype and were able to test it on both LuminAR Arms and as standalone installations in standard sockets. The bulbs serve as a natural carrier for the projector, and our design process allowed us to derive interesting techniques to deal with real-world engineering factors such as weight and thermal management.

The final design of the LuminAR bulb completed a series of explorations that advances the design and engineering of compact projector-camera systems. It is also to the best knowledge of the author, the first example of a kinetic and computer embedded system.

The evolution of the LuminAR Lamp concept generated six kinetic projected AR interfaces variations. Through the several iterations, we embedded the robotics elements of our design, creating an actuated interface embedded in a common day-to-day object.

We provided a valid first step for the design and implementation of such future actuated task lights. But, we are also aware that our hardware is limited, many improvements can and should be applied mainly to the actuation and sensing subsystems. Additionally, the overall cost of the arm should be reduced, so it becomes commercially viable.

We also embedded the technology in a ceiling mounted Spotlight and a media projector form factor. This allowed us to easily integrate projected interfaces into objects such as school chairs and various tables. The hardware we created is compact and portable. Our designs reduced the inherent complexity of the traditional projector camera system installation, yielding a simple system that is easy to integrate. Such form factors can now be reused across many applications and installation scenarios.

The case for Multiple LuminAR Devices

Finally, in our work, we focused on a single LuminAR scenario, and on interacting between co-located and remote LuminAR devices. We describe such scenarios in Chapter Six, where we discuss the LuminAR collaborative applications white:scape and Swÿp. One interesting area for future research is the case where multiple LuminAR devices are interacting in the same environment. Such scenarios would require the system to support projector display overlap support, kinematic motion synchronization and interface techniques that allow content and GUI elements to span multiple LuminAR devices. While we imagined such scenarios, and to a degree were inspired by Raskar's vision for the office of the future [123], we did not implement multiple device support, and we leave that for future work that would follow.

4 Software

In the related work section, we reviewed several of the AR systems that use projection to augment spaces. While these systems implement the low-level capabilities to allow developers to build applications without having to re-implement basic AR algorithms, they lack simple software interfaces that enable rapid application development and deployment that a typical web-based software frameworks provides. Additionally, they do not provide the level of abstraction required to shield developers from changes in the back-end implementation and to create truly developer-oriented APIs. Finally, tightly coupling developer APIs with the augmented reality implementation reduces the applications' portability and forces them to be tied to a specific platform.

LuXor and Lens are the components of the Projected Augmented Reality Software Framework motivated by the existing systems' drawbacks. My goal was to create a software framework that can support complex projected AR scenarios, while hiding the complexity of the underlying computer vision and augmented reality algorithms from the application developer. This approach allowed developers with basic web development knowledge to create augmented reality application with impressive speed and functionality.

In this section, we provide a high-level review of the software requirements and design considerations that motivated LuXor and Lens. We begin with software architecture overview, and continue to describe the core services and the application framework that enables developers to create projected AR applications.

4.1 High-Level Software Requirements

The software requirements for LuXor were also informed by the design principles presented in the beginning of this chapter. We can extend some of the principles to the software domain, supporting our goal to create an interactive, kinetic projected augmented reality user experience. Finally, to complement the discussion of the high-level requirements presented in this section, the reader should also refer to the Interaction Design Principles presented in Chapter Four – Interaction techniques.

The table below summarizes a set of high-level software requirements that guided LuXor's and Lens software architecture design:

Table 4-1 Luxor and Lens High Level Software Requirements

Embedded	• The entire software stack should be able to run on embedded boards that have a single processor (e.g., Intel Atom MTM or ARM based systems) • The entire code base should have a small footprint so it fits in embedded storage devices • Support close to real-time constraints for computer vision, motor control and graphics • Low runtime memory requirements
Layered and Modular	• The frameworks should provide good abstractions and relevant layers to separate low level drivers, system core services and application interfaces • The system should be developed in a modular fashion, allowing (1) different configuration options and (2) distributing computation to either additional module instances or remote modules
Extensible / Portable	• Provide interfaces to extend the system in its respective layers • Support portability of the codebase across different LuminAR devices, abstracting hardware dependent code
Connected	• The system should enable LuminAR devices to work in a network • Support device addressing and data exchange protocols
Application Model	• Provide a standard web-based application development framework
Standard / Open	• Use open standards, allowing for easy 3rd software integration
Natural Interaction / AR Interfaces	• Provide UI modality fusion. Support standard GUI approaches (e.g., multi touch) with gesture detection and AR interaction techniques
Limited Installation & Calibration	• The system should require minimal configuration and/or hardware-software calibration to run

Interactive Kinetics Control	• Abstract robotics control aspects, providing mechanisms to use kinetics for application, interactive purposes

4.2 LuXor: Kinetic Projected AR Software Framework

LuXor is the LuminAR software framework and execution environment. The LuXor software stack is designed to run on the embedded computer of the LuminAR Bulb. LuXor has specific LuminAR systems drivers that communicate with the firmware of the various LuminAR devices. LuXor also provides the interface for the bulb to communicate with other LuminAR hardware devices. It was developed for the later iterations of the LuminAR Bulb and specifically tested with the various versions of the arms along with the LuminAR Spotlight [79].

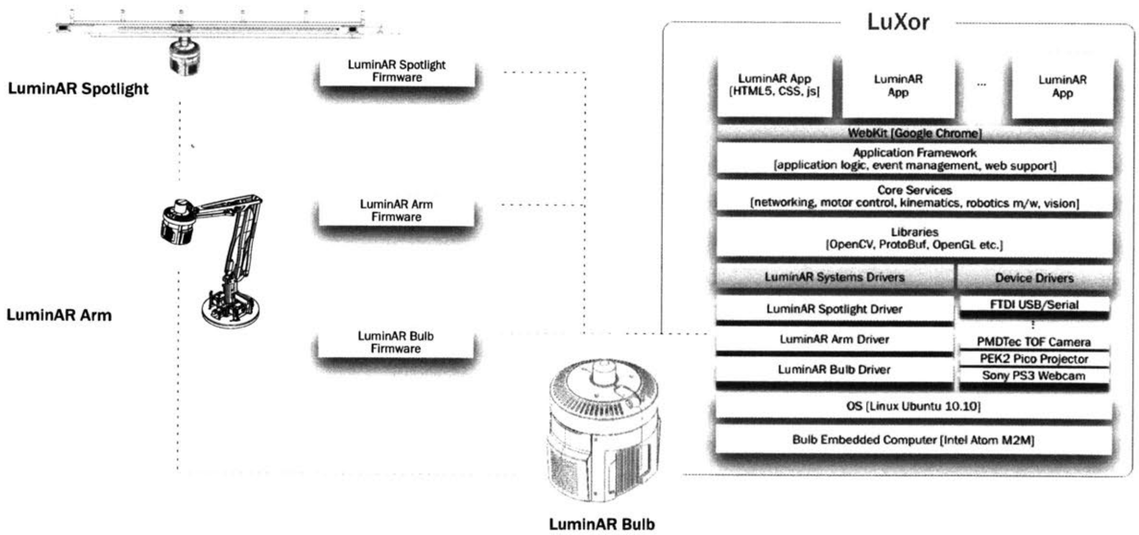


Figure 4-1 The LuXor software stack is designed to run on the embedded computer of the LuminAR Bulb. LuXor has specific LuminAR systems drivers that communicate with the firmware of the LuminAR Arm and LuminAR Spotlight

LuXor combines different software disciplines to enable LuminAR devices to become interactive, kinetic, and connected projected augmented reality interfaces. In broad strokes, LuXor comprises a web-based application framework (see Figure 4-2 for software architecture overview diagram) and a set of core services that provide LuminAR devices with a hardware drivers, robotic middleware, application runtime logic, event management, computer vision services, and projected GUI utilities.

4.2.1 Software Architecture

Two main areas of software design inspired the LuXor software architecture design: mobile platform architectures (e.g., Google Android or Apple iOS **[124]**, **[125]**), and robotic middleware solutions (e.g., ROS from Willow Garage **[126]**). The LuXor architecture is a minimalistic hybrid between these two very different paradigms.

The diagram below provides an elaborated view of the LuXor software stack. In the following sections, we briefly discuss LuXor's major components.

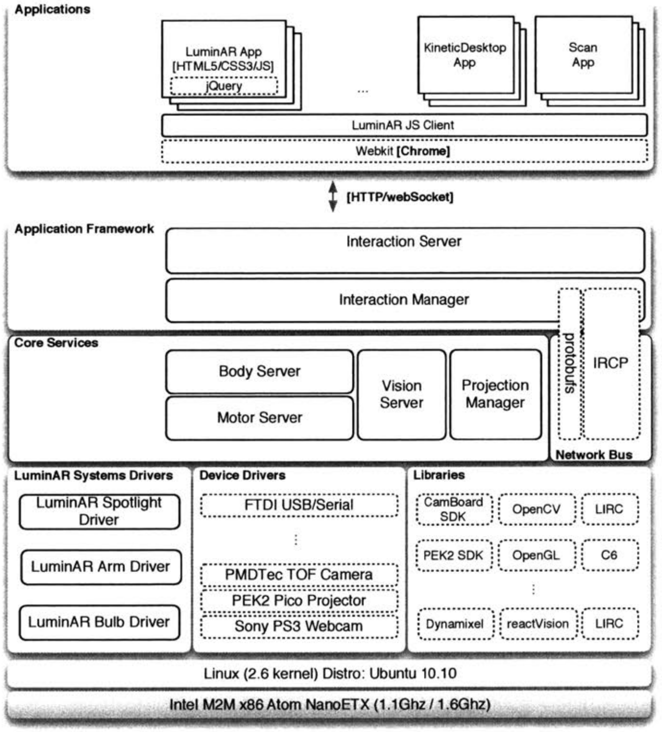


Figure 4-2 LuXor software architecture

4.2.2 Hardware and Operating System

LuXor is designed to run on Intel's M2M Atom hardware. The M2M is an x86 platform and can easily boot Linux. LuXor relies on the Linux kernel for core OS functions and basic services such as network connectivity and windows management. We also rely on Linux driver model to support all the peripherals attached directly to the bulb's computer.

4.2.3 Components Overview

The LuXor infrastructure is designed as a layered architecture, defining the following layers and components:

- **Applications** – front-end, web-based LuminAR applications
- **Application Framework** – runtime execution, interaction and event management
- **Core Services** – a set of infrastructure components, interfacing with low level hardware (motors, sensors), exposing the high level interface for the application framework
- **Network Bus** – IPC communication pipeline and data bus
- **Libraries** – includes all 3rd party dependencies
- **LuminAR Systems Drivers** – drivers to LuminAR systems
- **Device Drivers** – drivers to peripheral components of the LuminAR Bulb

In the following sections, we provide more details on the specific functionality of the various LuXor components.

Core Services

LuXor's core services layer includes a set of independent servers, each dedicated to specific services required by the Interaction Manager. The services interface with low-level hardware drivers and expose simple interfaces to the application framework.

Networking

The LuXor services and application framework uses a shared network bus to communicate. Inter-application communication is critical when implementing a modular, distributed system. The network bus is used to integrate independent software modules, providing simple means to send messages and data. This enables LuXor to become an event driven system throughout its different layers.

Our message description implementation was based on Google's Protobufs open source project [127]. Protobufs stands for Protocol Buffers. It compiles human-readable structured message descriptors to binary protocol buffers. Protobuf also provides interfaces in various platforms, which was important when integrating a low-level C++ server with a high level Python interface.

LuXor uses the Intra-Robot-Communication-Protocol (IRCP) as the transport layer for the network bus. IRCP is a thin layer of networking logic, implemented on top of UDP transport protocol. It provides Python, C++ and Java interfaces. IRCP handles the basic network creation, address resolution and basic error checking. It provides simple interface to send typed data packets or send streams of data. IRCP was developed at the MIT Media Lab Personal Robots Group [128].

Vision Server

The vision server's role in the LuXor system is to manage input from the various cameras in the LuminAR bulb and produce interaction events, such as the pressing of a button, detecting a hand gesture or the presence of an object. Events are propagated to the Interaction Manager, where they can trigger application-level state changes.

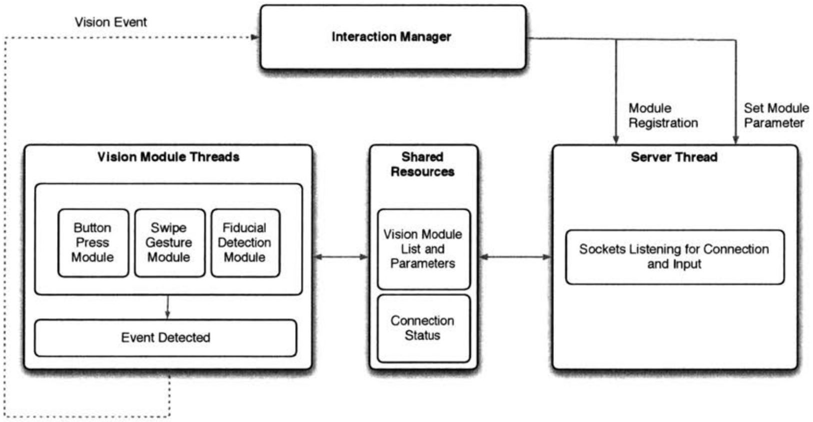


Figure 4-3 LuXor's Vision Server vision events flow

The server was developed in C++ using OpenCV [129] and reactiVision [130], and was optimized for embedded performance. It also manages the access to shared camera resources, providing thread safety. The server modules could also be deployed on multiple machines, thanks to LuXor's network bus. This was useful when the vision tasks were pushing the limits of the Atom based platforms we were using.

The server generalizes and abstracts the interface to cameras and sensors, exposing a functional interface using vision modules. The vision server owns a dynamic array of vision modules, each of which represents an independent thread processing video and producing events, which can be sent asynchronously to the interaction manager.

Vision modules need not only the ability to send events, but also the ability to receive information from the Interaction Manager about the application state, such as the location of buttons on a page. We implemented modules to detect button press, hand swipe gestures and a fiducial marker tracker module. We also experimented with a depth-based touch detection module using PMDTec CamBoard USB Time-of-Flight(TOF) sensor **[131]**.

The Vision Server modules assume homography between the projected display and the camera viewpoint. The homography is accomplished using a one-time simple calibration process, in which camera pixels are correlated to projector pixels. The calibration data is then saved to a configuration file.

Button Press Module

To allow LuminAR to determine when a button in the projected user interface has been pressed, we initially turned to using a time-of-flight depth-sensing camera from PMDTec. The sensor measures the distance of objects by sending out pulses of light and then measuring the amount of time it takes for the light to return to the camera, with a resolution of 200x200 pixels and measurements in meters. After calibrating the sensor, we found the depth value at each button by sampling the depth image and then forming a virtual hemisphere above each button. This technique was also applied recently by Wilson using a Kinect sensor **[90]**. Then, by the number of depth pixels in that 3D volume, we determined whether or not a finger was present on the button. We found, however, that the thickness of the average finger was not significantly larger than the noise level in the depth frames, and so we had a large number of false positives with low values of the threshold and a large number of false negatives for high values of the threshold. That, combined with technical issues regarding the depth sensing hardware, forced us to use simple webcams as an alternative.

As a replacement, we decided to use a webcam (Sony PS3 Eye OEM) with a far simpler scheme. We found that the saturation values for both the white table and the projection on it were very low, while the saturation values for hands over the table and the projection was quite high, so thresholding the values of saturation within the region where the button was located allowed us to reliably determine when a button was present in that area, signifying a click.

Clearly, this approach was aimed at giving us minimal-vision based interaction capabilities, and we can use it to explore many applications that require simple touch capabilities. Using the Vision Server's architecture, it is easy to add additional modules that will improve the detection capabilities with finger tracking and hand gesture detections mechanisms.

Swipe Gesture Module

The Swipe Module allows the user to perform a swipe gesture beneath the LuminAR Bulb's viewport, and determines the direction the swipe was made, either right to left, left to right, top to bottom, or bottom to top. To determine the presence of a swipe and its direction we initially attempted methods involving optical flow, but found that they were not lightweight enough for the Atom hardware considering the vast amounts of additional processing that was necessary. Instead, we chose a far simpler method that used the same hand extraction by saturation thresholding as the button press module.

Fiducial Detection Module

To detect the presence and location of objects in the LuminAR view, we used fiducial markers. reacTIVision [130] is an open source implementation of fiducial marker detection that we opted to use. It sends fiducial information over a socket, so in the Vision Server we created a module that listens on the reacTIVision port and forwards fiducial events up to the Interaction Manager, indicating the presence and location of objects with fiducial tags.

Body Server

The LuXor Body Server is a small-scale robotic middleware layer. It is in charge of controlling the LuminAR device's pose. It also abstracts this function from the application, providing an interface to manage poses. The Body Server also handles the entire kinematics calculations specific to the device (i.e. The LuminAR Arms and LuminAR Spotlight have very different kinematic models). Using the kinematic model, the Body Server is also responsible to load and maintain the Vision Server calibration data. The diagram below shows the main flow of the Body Server.

The Body Server continuously communicates with the Interaction Manager to monitor and control the pose of the LuminAR device.

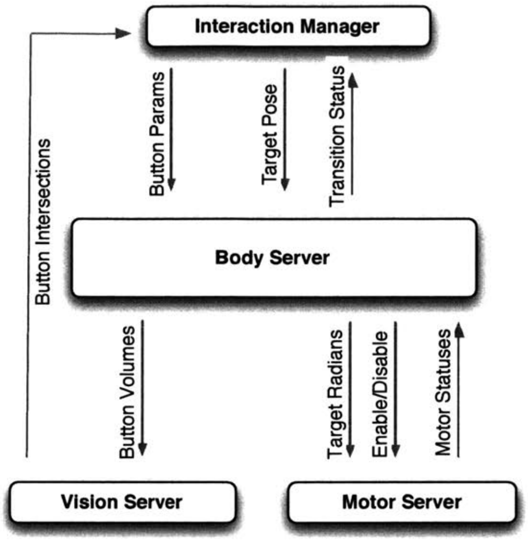


Figure 4-4 LuXor's Body Server event flow and interfaces

Once an application has reached a state in which the application is required to move, the Interaction server would request a logically defined 'Target Pose" from the Body server. The Body server in turn translates the logical target name to actual motor coordinates in radians and forwards them to the Motor Server. Once a motion sequence is completed, the Body Serve returns the interaction control to the Interaction Manager.

Motor Server

The LuXor motor server provides an abstraction for low-level motor control required of LuminAR devices. It provides the means to configure multiple motor configurations per device. It also allows for multiple device configurations. This feature was extremely useful when defining complex devices with separate motor configurations like the LuminAR Spotlight or the LuminAR Lamp.

Projection Manager

The Projection Manager was designed to interface with the application framework to provide display manipulation routines. The PEK2 SDK provides hardware-based keystone and projection angle manipulations that could help deal with geometrically correcting a projected image. Unfortunately, we did not have sufficient time to fully complete and test this manager.

Application Framework

LuXor's application framework provides runtime execution environment, state management and event distribution mechanisms. It is designed to support a simple, yet powerful, web-based development environment.

Interaction Manager

The heart of the Application framework is the Interaction Manager. It is responsible for managing the entire application lifecycle. The manager has direct interfaces to the Body Server and the Vision Server.

Two key mechanisms govern the server's main loop. First, the *Transition Executor* is responsible for maintaining the current logical state of an application. Each state is defined by a set of transitions that corresponded respectively to the logic of the application pages. We discuss the general structure of LuminAR applications in the next section of this chapter.

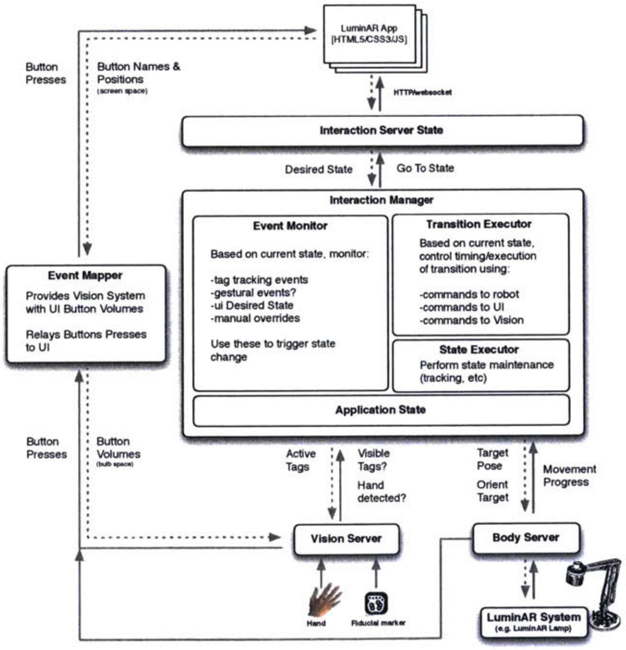


Figure 4-5 LuXor's Interaction Manager logic diagram, event management components and interaction events flow

The second mechanism is the *Event Monitor and Mapper;* it generates triggers based on incoming events in the system. Events can be either I/O driven from the application level or internal system events. Triggers invoke the State Executor, completing the event flow loop. Figure 4-5 summarizes the interaction manager event loop.

Interaction Server

The interaction server is responsible for relaying events, and state changes to and from LuminAR Applications. The server has a direct interface to the Interaction server, and a socket interface for LuminAR application to bind to. The server also has a registry for LuminAR events.

LuXor Applications

LuminAR applications are, in fact, web apps. Modern web browsers like Webkit, Chrome, and others support the powerful new HTML5 and CSS3 specifications, and Javascript provides a platform-independent development environment. These new browser capabilities combined with visual Javascript toolkits such as jQuery, provides an excellent front end GUI development platform.

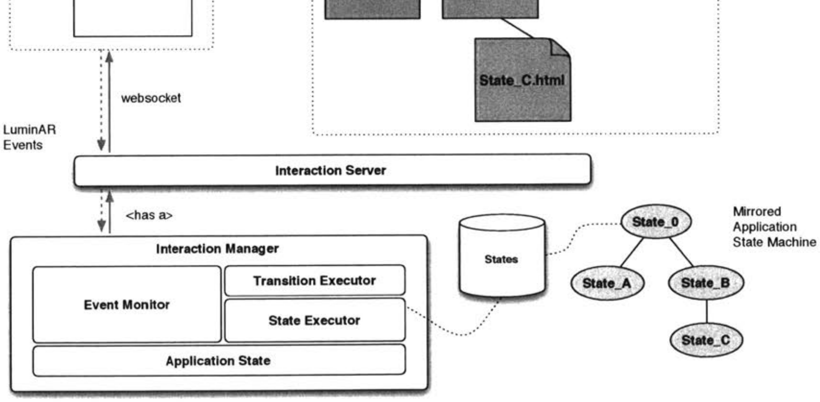


Figure 4-6 The LuXor application runtime model

Developing application code in a document-driven, scripted environment has numerous advantages to the traditional programming alternative. Two advantages are important to note: (1) in a web application, all the resources are packaged with the application, and the browser has a rich set of capabilities to render audio, video and various graphic formats. (2) Web based applications are easy to integrate with practically any web service or API that exists online.

The LuXor Application Runtime Model

The challenge we had to tackle is how to robustly connect the browser front end to the LuXor backend. This was accomplished using a custom JavaScript client we developed. Once the client is imported to the application's main page, it opens a websocket connection to the interaction server and loads the definitions of the supported LuminAR events. With the communication pipeline established, application developers can create interactive web applications that use vision-based interaction and kinematics.

What we did not have sufficient time to implement is the code-generation step that automatically creates LuminAR application state definitions for the front end. In the course of this dissertation work, we hand-coded the application states using simple Python configuration scripts, but there is no reason they cannot be auto-generated.

4.3 Lens: Web-Based AR Software Framework

To support multiple augmented reality systems, we created Lens: A Javascript SDK for Building Web-Based Projected Augmented Reality Applications.

Lens has evolved from the experience we gained creating LuXor. At its core Lens is an independent software library that assumes an underlying computer vision and AR subsystem. LuXor serve that sub-system for Lens. Earlier in this chapter we described LuXor's Vision Server, that effectively provides Lens with various capabilities to support gesture detection, fiducial detection as well as other basic computer vision services. In addition, LuXor provides an API to integrate additional network services as well as data protocol APIs.

Lens served as the key framework we used to develop the application that drive MARS (See Chapter Seven) and Enlight (See Chapter Eight).

There has been substantial research into Augmented Reality (AR) interfaces, but it is still difficult to build applications that leverage these technologies due to the lack of development tools. Lens is a Javascript API that allows developers to use standard web development tools to build projected AR applications. This eases the development process and opens up this interaction modality to all levels of developers.

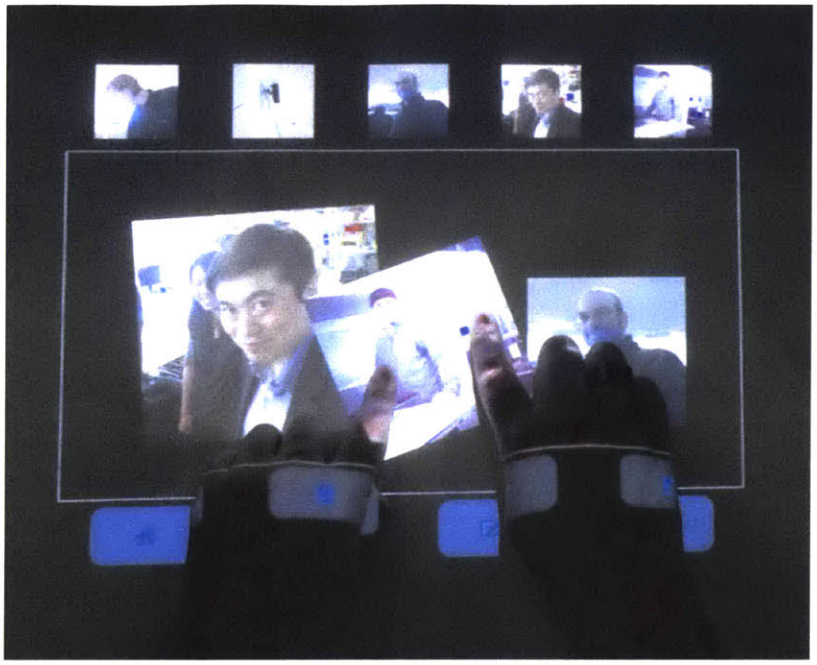


Figure 4-7 A user manipulating photos using multi-touch gestures in the Lens augmented scrapbooking application

At present, Lens supports multi-touch interaction, object identification and tracking, and capturing snapshots from a camera and cropping them based on the position of the web page's context HTML elements. Lens also provides applications with a number of pre-built interface widgets and components specifically designed around the constraints of projected AR interfaces. We have deployed Lens on top of LuminAR to develop a myriad of applications, including augmented scrapbooking, a Pong video game clone, and an augmented manufacturing application deployed and evaluated at an operating assembly line.

Lens provides a high-level abstraction decoupled from the underlying AR platform. It uses a web-based programming environment that allows rapid development of AR applications. Developers can use Lens to create applications with full multi-touch capabilities using a dozen lines of code. More sophisticated applications can leverage the entire ecosystem of web-development tools. This approach makes projected AR interfaces accessible to a wide audience of developers who only need to have basic web-development skills.

4.3.1 Related Work

There is growing interest in projected AR. A substantial number of research and industry projects leverage projection as a means to augment spaces and create new kinds of user interfaces.

Developer-oriented toolkits do exist for other forms of augmented reality. There are a number of commercial offerings targeted primarily at mobile phones, such as the metaio SDK [132], which provides developers with the ability to augment images captured by the mobile device's camera. In addition, the metaio Web SDK is a web-based framework for AR, but is similarly limited to augmenting images captured from a webcam, and focuses on object identification, rather than creating novel interfaces on hardware designed for augmented experiences. Metaio was acquired by Apple and recently have shipped an evolution of the original metaio SDK called ARKit [18] that extend AR development functionalities to iOS devices. Another great example of a contemporary AR development environments is PTC's Vofuria [19].

While these systems implement the low-level capabilities to allow developers to build applications without having to implement basic AR algorithms, it doesn't help them create experiences that integrate well with the rest of the system. The systems also don't provide the level of abstraction needed to shield developers from changes in the back-end implementation in order to create truly developer-oriented APIs. Finally, tightly coupling developer APIs with the augmented reality implementation reduces the portability of applications and forces them to be tied to a specific platform.

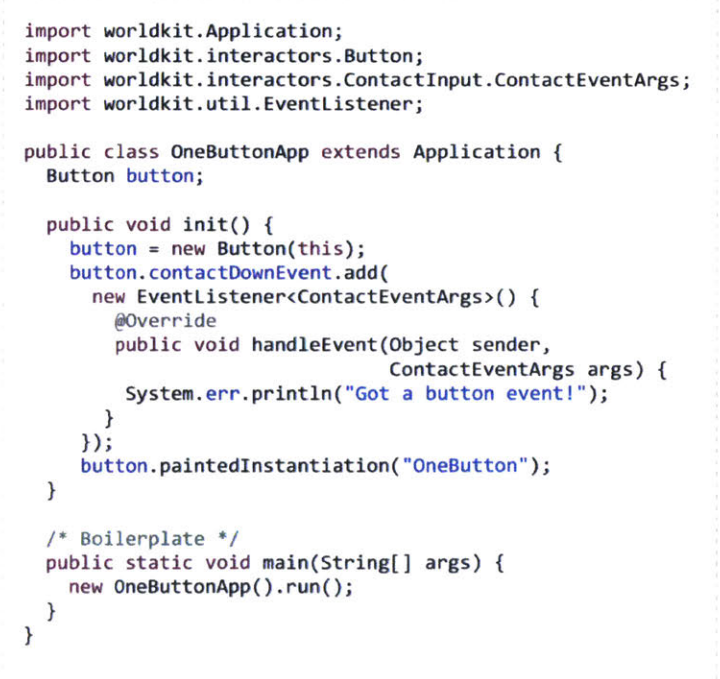
In the field of wearable AR technology based on head-mounted displays, a number of toolkits have been developed, such as the framework of Behzadan et al. [133] and the DWARF framework of Bauer et al. [134]. DWARF, in particular, uses a component-based design much like Lens and draws a clear distinction between the presentation layer and the image processing and recognition layer. DWARF, however, predates modern depth cameras, so it doesn't support many of the interactions enabled by depth sensors. Ubiquitous computing also has a number of development frameworks, including iStuff [135] which focuses on interaction with other wireless devices in the augmented space, rather than gestural interfaces and interaction with arbitrary objects.

The past few years have seen a tremendous burst of new wearable AR hardware and the respective software development kits. Several companies innovated quickly and introduced several form factors for wearable AR headsets: Daqri **[136]**, **[137]**, META **[15]**, Microsoft's HoloLens **[14]** and Google's Glass **[20]**.

There has been a small amount of research into using the web as a platform for developing projected AR interfaces. Hardy and Alexander presented a framework **[138]** that runs applications inside a browser that supports HTML, CSS, and Javascript. It mainly provides multi-touch interaction, and, like Lens, it does so by injecting standard touch events into the page. However, this framework has fewer AR primitives than Lens: it doesn't support object recognition or tracking, and doesn't provide a set of user interface components and styles to create a cohesive visual style across all applications.

Ahn et. al. **[139]** used web-based techniques to separate and structure mobile AR content from the mobile application logic. This approach promotes the adoption of AR as the web has such a clear significance as the key infrastructure for our consumer and business applications.

Some new work has followed LuminAR, LuXor and Lens path and has shown real promise and progress in the interaction and software domain. Robert Xiao has presented WorldKit **[140]**. WorldKit introduced easy to use Java-based interfaces for projected AR functionality.



```
import worldkit.Application;
import worldkit.interactors.Button;
import worldkit.interactors.ContactInput.ContactEventArgs;
import worldkit.util.EventListener;

public class OneButtonApp extends Application {
  Button button;

  public void init() {
    button = new Button(this);
    button.contactDownEvent.add(
      new EventListener<ContactEventArgs>() {
        @Override
        public void handleEvent(Object sender,
                                ContactEventArgs args) {
         System.err.println("Got a button event!");
        }
     });
     button.paintedInstantiation("OneButton");
  }

  /* Boilerplate */
  public static void main(String[] args) {
    new OneButtonApp().run();
  }
}
```

Figure 4-8 WorldKit Example code for a single button application. (Source: **[140]**)

More recently, a few projects have explored creating toolkits along the lines of Lens to ease the development of projected augmented reality applications. ProjectorKit [141] is a closer relative of Lens, providing an open-source developer-oriented toolkit for building interactive mobile projected applications by providing high-level primitives. However, the authors acknowledge that their project depends on high-end, expensive equipment for determining position, rather than using cheap, ubiquitous technologies such as depth cameras. Furthermore, ProjectorKit is implemented in C#, which makes it highly complicated to port to non-Microsoft Windows based platform. In another relevant recent work, Xiao presented Desktography [105] that extended WorldKit to include various advanced gestures.

4.3.2 Motivation

A major barrier to the adoption of augmented reality technology is the difficulty of creating applications for such platforms. Until developers can easily and quickly create compelling applications, augmented reality cannot cross into the mainstream. Fundamentally, the applications built on an AR platform are much more important than the platform itself. In analogy to the web: do users care about which web browser they use? In most cases, no. Some browsers may be faster, more attractive, or provide more features, but the browser pales in comparison to the web page itself. The success of the web is due, primarily, to the ease of developing web applications. HTML, CSS, and Javascript are simple enough to teach to introductory-level students—almost all colleges and many high schools have introductory classes in web development. The simplicity of developing web applications—no compiler, linkers, libraries, or dependencies to install—and the ease of publishing a website to the world make it an ideal platform for software development.

Most developers have a solid grasp of web technologies, and Lens allows them to leverage these existing skills to develop augmented reality applications. In addition, by enforcing a clear separation between application code and the underlying augmented reality implementation, Lens could be ported to other form factors and platforms to allow the same applications to run on hardware better adapted to particular environments. For example, Lens could be adapted for Bonfire or DeskJockey to run its applications in a PC-centric workspace extension, or for DWARF to run the same applications in a wearable form-factor with a head-mounted display.

4.3.3 Software Architecture

Lens is a Javascript library that seamlessly integrates augmented reality capabilities into web development. At the heart of Lens is the Dispatcher, which handles communication over the WebSocket. Other parts of Lens send objects to the Dispatcher, which converts them to JSON and sends them over the WebSocket to the Bridge. In addition, other parts of Lens can subscribe to messages of a particular type. The Dispatcher listens to the WebSocket and then relays messages to subscribers.

The organization of Lens is shown in Figure 4-9. The application-facing part of Lens consists of a number of modules that expose high-level APIs. It then translates these API calls into the low-level events sent to the Dispatcher for transmission to the Bridge. Lens's modules are grouped into three sections: ARCore, Components, and Look-and-Feel toolkit (LAF).

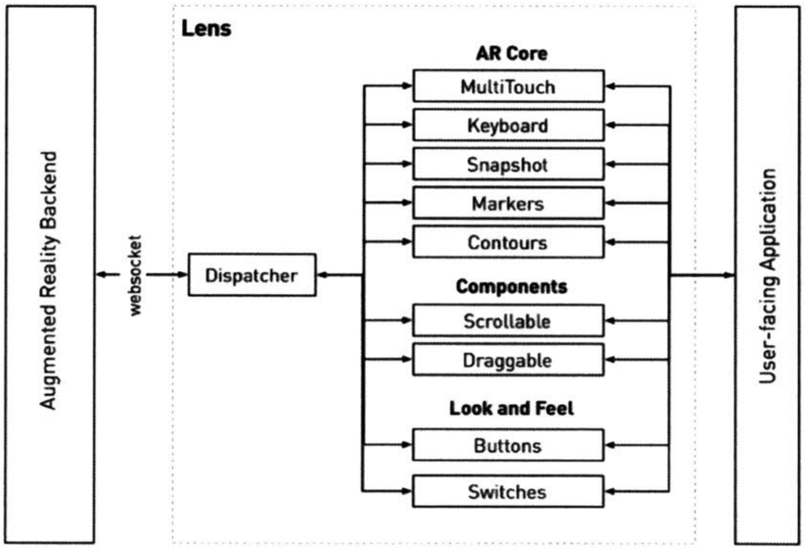


Figure 4-9 The organization of Lens

The lowest-level section, ARCore, directly exposes the AR functionality of the backend software to applications. Modules in ARCore provide these tools in a way that meshes well with the web environment but also provides relatively low-level access to AR capabilities. A higher-level section is Components. Modules in Components implement common interaction paradigms such as multi-finger scrolling and pinch-to-zoom. Components build upon ARCore, but are still opt-in—developers have to explicitly use them;

they don't actively change the page. Finally, the highest-level section is the Look-and-Feel toolkit, which actively modifies the page to make buttons, switches, and input fields fit visually with the rest of the system, and behave appropriately in a projected environment.

Lens applications run in a stock Chrome or Chromium browser. Each application has its own directory in a central applications folder. We use Chrome's `--host-rules` flag to map all requests for domains within the nonexistent `.lens` top-level domain to our own webserver running on the local machine, which uses the domain name to determine which application to serve. For example, a request to `http://myapp.lens/foo.js` would be routed to our webserver, which would serve `/apps/myapp/foo.js`. Lens itself is served from `http://lens.lens/`, so applications can include the Lens library by inserting a script tag to load `http://Lens.Lens/src/lens.js`. This model of loading applications allows us to run untrusted Lens applications, because the applications are run entirely in the browser sandbox: Chrome is executing the application, and treats it as an untrusted webpage. In addition, applications are served from different domains, so they are protected from each other by the same-origin policy.

ARCore

The ARCore component of Lens allows applications to access basic augmented reality functions. In addition, ARCore exposes this functionality in a way that fits well with common paradigms of web development, so it integrates smoothly with existing Javascript libraries and tools. ARCore leverages the Document Object Model (DOM) for event dispatch and subscriptions, allowing Lens to work seamlessly with the popular jQuery library.

ARCore is comprised of five main modules. The Touches module uses the backend system's finger tracking to implement multi-touch events based on the World Wide Web Consortium's draft specification for touch events [142]. Because Lens uses standard touch events, developers can use existing libraries to work with these events.

Lens also extends the specification to allow those developing specifically for Lens to use a complementary object-oriented API in addition to the event-based API: applications can subscribe to `touchstart`, `touchmove`, and `touchend` events, and the handlers for these events receive touch objects that allow for easy subscription to further events. A snippet from a multi-touch drawing application like the one in Figure 4-10 is shown in Listing 4-1. Lens also calculates the rotation, scaling, and motion of multi-touch gestures and extends the touch objects given to event handlers with this information to allow application to easily implement multi-touch gestures such as multi-finger scrolling and pinch-to-zoom.

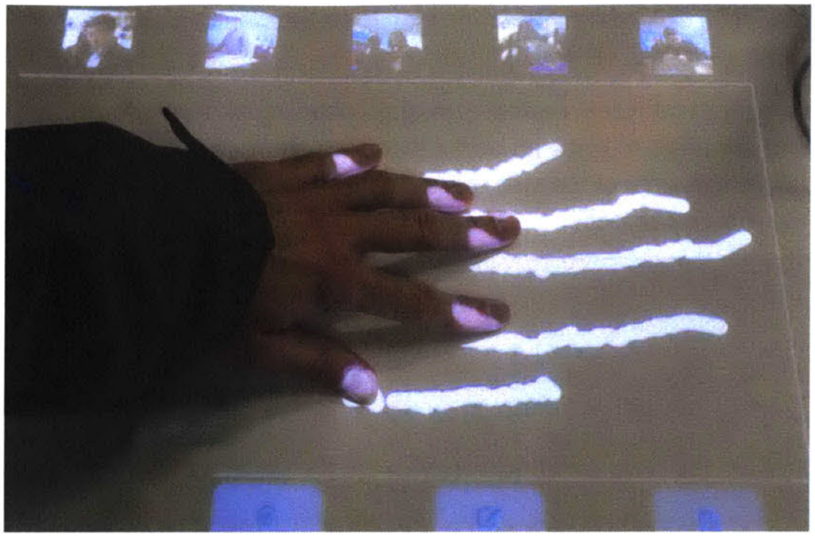


Figure 4-10 A multi-touch drawing & scrapbooking application, like the one from Listing 4-1

```
// lens:touchmove is fired whenever a finger moves
// on the canvas
canvas.addEventListener("lens:touchstart", function(evt){
var prevX = evt.touch.pageX; var prevY = evt.touch.pageY;
evt.touch.move(function() {
        drawLine(prevX, prevY, touch.pageX, touch.pageY);
        prevX = touch.pageX;
        prevY = touch.pageY;
});
});
```

Listing 4-1 Basic multi-touch drawing

ARCore's Contour module works similarly to Touches, but tracks the outline of arbitrary objects, not just fingers. Rather than firing events on DOM elements, developers create a `ContourSearch` which specifies a region of interest, then use an API similar to the object-oriented touch API to receive notifications when objects are detected in the region. An example usage of this API is shown in Listing 4-2.

ARCore also has a Marker module, which tracks texture-based markers, rather than all objects. Given an image of a target texture, the Marker module will recognize and track the target as it moves. In addition, the Marker module can "pin" a DOM element to the tracked object, using 3D CSS transformations to fit the element onto the object, as shown in Listing 4-3.

ARCore also supports taking snapshots from camera. These can either be raw photos directly from the camera, or a transformed and cropped portion of the image to match the location of a particular DOM element (Listing 4-4).

As a more complete example, an application could have a "scan" area and ask users to place an object there; the application would then take a snapshot of the area, pass it to the Markers module, and track the scanned object.

```
// watches for objects in the center of the app, // slightly
above the surface. var search = Lens.Contours.search({

    x: [200, 800],

    y: [200, 800],

    z: [100, 300]

});

search.appear(function(contour) { // the object has appeared
contour.move(function() {

        // the object has moved

});

});
```

Listing 4-2 Watching for object in a region of interest

```
// pin an HTML element to a physical object

// test.png is a picture of the object

var marker = Lens.Marker.fromImageURL("test.png");
marker.pin("#someElement")
```

Listing 4-3 : Pinning a DOM element to a tracked object

```
Lens.Snapshot.capture("#someElement", function(snapshot){ //
do something with the snapshot

});
```

Listing 4-4 : Taking a snapshot of a DOM element

Finally, ARCore has a Keyboard module that allows applications to receive keyboard input. The Keyboard module makes use of the Nearby Server to detect nearby phones. If a phone is found, the Keyboard module uses the phone's keyboard. If no phone is found, the Keyboard module displays an onscreen keyboard (Figure 4-11). In either case, events from the keyboard are fired on the focused page element using the standard W3C keyboard events specification. In addition, if an input element is selected, typed characters appear in this element. ARCore provides access to augmented reality primitives while balancing between low-level control and ease of development.

In addition, by leveraging standard web technologies such as the DOM and conforming to existing event specifications, Lens ensures compatibility with current Javascript libraries, giving developers a rich ecosystem in which to build their applications.

Components

Lens also provides a number of components that use ARCore to enable common interaction paradigms. Lens provides the Draggable component, which uses ARCore's Touches module to allow DOM elements to be dragged with a single finger, as well as rotated and scaled with multiple fingers (Listing 4-5).

In addition, Lens provides a Scrollable component that allows users to scroll a page element using a configurable number of fingers, as well as a TextEntry component that uses the Keyboards module to display a keyboard when a user taps an input element. These components are higher-level than ARCore, ensuring ease of use and consistency across applications. Like ARCore, they are opt-in, and do not actively change the page without a specific call from the developer.

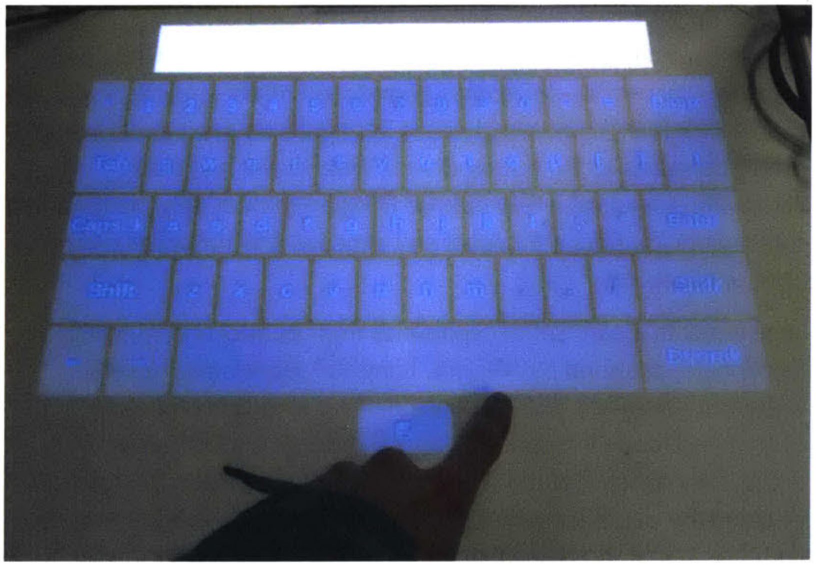


Figure 4-11 The Lens onscreen keyboard

```
// allow a user to use multi-touch gestures to // move, scale,
and rotate a page element
Lens.Components.Draggables.enable("#someElement");
```

Listing 4-5 : Using the Draggable component

User Interface

The highest-level functionality provided by Lens is the User Interface Look-and-Feel (LAF) toolkit, much of which is based on Designmodo's FlatUI toolkit [143]. LAF is a set of CSS styles and a small amount of Javascript code that actively transform page elements to fit with the needs of a projected interface. It applies a common color scheme, typography, and iconography, and adds the necessary functionality to make HTML pages usable as Lens applications. Developers can include either the entire LAF toolkit or specific parts. In addition to applying colors and typography, the LAF toolkit styles buttons, check boxes, radio buttons, toggle switches, text boxes, and scrollbars. It also displays indicators when the user touches the interface, applies TextEntry components to all input elements, and applies the Scrollable component to the body element. LAF is highly configurable, with pluggable color schemes, icons, and fonts using CSS variables. A demonstration of an application that uses LAF is shown in Figure 4-12.

4.3.4 Platform

Lens is currently deployed on LuminAR, a complete hardware and software platform for developing augmented reality applications (Figure 4-9) [144]. However, Lens is not tightly coupled to LuminAR: it simply connects to the backend via a WebSocket and then communicates using a well-defined Bridge Protocol. Any platform that implements the Bridge Protocol and exposes a WebSocket can run Lens applications.

Lens expects a core set of capabilities from its backend to enable the functionality provided by ARCore. LuminAR implements several important algorithms and functions needed to build compelling augmented reality interfaces. Most importantly, it identifies arbitrary object boundaries (called "Contours") within LuminAR's field of view by using the depth camera to locate the dominant background plane and then identifying blobs that intersect this plane. It additionally identifies blobs that are likely to be fingertips to enable multi-touch interaction. The vision server also supports texture-based marker tracking: given a texture image, it can recognize the texture and track its motion. Finally, the vision server can take snapshot images and transform them based on the homography between the camera and projector, allowing applications to take pictures of particular regions of the projection area.

LuminAR also has a Nearby Server, which uses zero configuration networking (zeroconf) **[145]** to automatically discover nearby LuminARs and other devices such as tablets and phones. These other devices can then broadcast their capabilities (such as access to a camera or keyboard) so nearby LuminARs can take advantage of them.

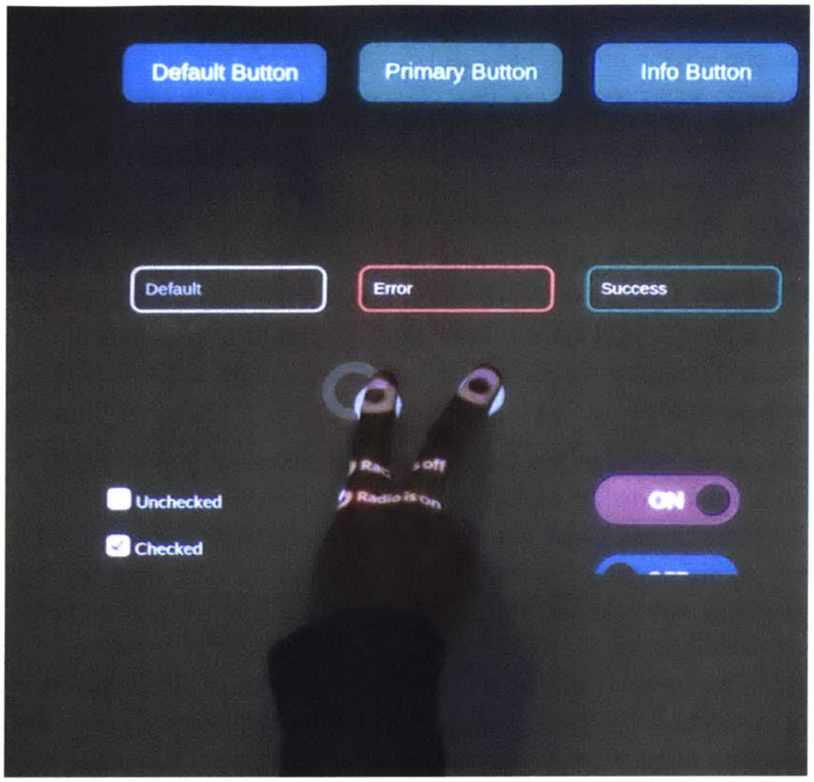


Figure 4-12 The Lens LAF toolkit.

While Lens has been designed to take advantage of the particular capabilities of LuminAR, it does not necessarily need to run on a platform that supports all of these features. ARCore is separated into discrete modules, so a platform could choose to implement only a subset of the API and simply disable those modules in Lens—the rest of the library would continue to work as expected. Moreover, the modular design would allow other platforms to implement ARCore modules for their own unique features. We therefore believe that Lens is suitable for developing applications on a wide range of platforms, and for developing applications that work well on multiple platforms.

4.3.5 Evaluation

Lens has successfully allowed the LuminAR team to develop more than 30 applications, ranging from photo browsing, to augmented scrapbooking (shown in Figure 4-10), to augmented manufacturing (also see Chapter Seven) for an assembly line (shown in Figure 4-13). We've also hosted workshops where developers outside our team have built applications using Lens. Based on feedback from team members and workshop participants, the Lens API has gone through two significant versions. The first version of Lens was a full framework enforcing a model-view-controller organization for applications. The second version of Lens has been designed to operate as a library, allowing developers to use their choice of tools and structures.

In addition, the portions of Lens that actively change a page (the LAF toolkit) have been separated from the core modules of Lens and are used only upon request. We have found that this new design greatly accelerates the speed of development: there is substantially less to learn, and developers can use tools and libraries they already know, rather than picking up new ones. Basic applications such as the drawing application shown in Listing 4-1 are simple, concise, and easy to create.

Overall, we've found that integrating augmented reality functionality consumes a small fraction of a developer's time when building applications with Lens. The web-based development paradigm has substantial advantages: first, most development can be done in a standard web browser without specialized hardware, allowing many developers to share a single device. Second, advances in web technologies and development tools speed up the creation of Lens applications immediately. Finally, because the portion of the code that interacts with Lens is small, developers who are already familiar with building Javascript applications have to learn very little to begin building Lens applications.

Fundamentally, Lens means that anyone who can build web applications can build AR applications. Rather than forcing developers to learn an entirely new toolchain and have in-depth knowledge of computer vision, camera calibration, and projected interface design, Lens provides a complete set of tools that abstract away the most difficult parts of AR application development, allowing developers to work on novel uses for AR.

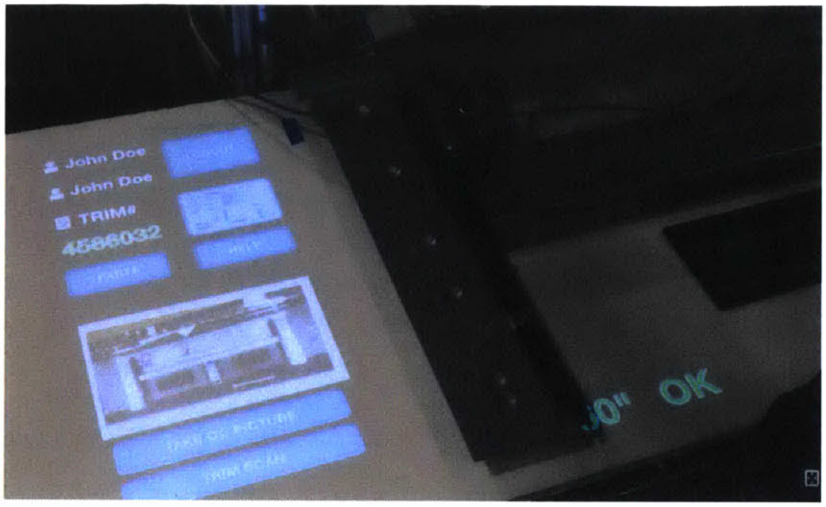


Figure 4-13 The Lens augmented manufacturing application, currently deployed in a operating assembly line.

4.4 Conclusion

We have found the web to be a great platform for developing AR applications. LuXor and Lens provided an experimental playground to assess the potential of web-based AR interfaces. The web provides a wealth of tools and libraries, combined with almost universal familiarity among developers. This approach to software development allowed us to develop sophisticated, attractive AR interfaces with very little effort and almost no learning curve. In addition, we've been able to leverage modern web-design techniques and toolkits to create a cohesive visual consistency between applications.

We continue to actively develop Lens to add support for additional AR primitives, and to bring tighter integration with popular Javascript libraries. We hope to involve more developers and give more people the opportunity to work with the latest AR technology. By providing a web interface to projected AR, Lens accelerates the pace of AR application development and research and provides access to a new and exciting interaction modality.

5 Interactions

LuminAR was designed to deliver a rich and interactive projected augmented reality user experience. The work in this dissertation was carried out in a period of time in which user interface technology emphasis has shifted towards the conceptual realm of Natural Interaction **[146]**.

Projected AR interfaces pose a host of interaction challenges. The known limitations include lack of feedback from a projected surface, display fidelity in different lighting conditions, occlusions and shadows caused by hands or objects, and the need for flat white surfaces. Such limitations many times can be mitigated by interaction design that takes them into account.

While we implemented a simple hand-tracking and basic gesture-detection mechanism, the next step would be to expand and support full multi-finger, multi-hands detection including postures. The technical details of our initial implementation are included in the description of the LuXor software framework Vision Server (see details in Chapter Four), There are many variants in the literature **[79]**, **[147]** that provide solution to these problems, and they could be adapted to the LuminAR software stack as new modules for our computer vision server.

5.1 Natural Interaction

Natural User Interface (NUI) is an umbrella term that encompasses various known techniques for multi-touch user interfaces, gestural interfaces, and other sensing-based interaction. Such Natural Interfaces support direct interaction, where the hands or the body serve as the input device, rendering the need for an intermediary device such as a keyboard or a mouse obsolete. The common denominator for these approaches is the creation of direct, intuitive interfaces that make our interaction with computers seamless and unobtrusive. Natural Interfaces also refer to the interfaces that are able to blend the digital and physical world while responding to context.

The computer mouse, GUI, and the WIMP concepts contributed immensely to the mass adaptation of personal computers. They also contributed to the adaptation of other display-centric computing devices (e.g., smart phones and tablet computers) that essentially used the same interaction paradigm. In the broad sense, Natural

Interfaces are well on their way to becoming a key interaction modality in the years to come, and may very well contribute in a similar fashion to the adoption of new form factors for computing that would use augmented reality as the key interaction modality. Early evidence of this trend can be seen in the emergence of new standards and major open source projects, such as the Microsoft's Kinect SDK [148], OpenNI initiative [149], and work carried out by the Natural User Interface Group [150]. This trend continues with the strong emergence of the wearable headset. Companies like META are redefining the desktop interaction space using a powerful set of mixed reality metaphors [15]. HoloLens is another example where new forms of point-and-click interactions are redefined to control elements of a holographic, virtual user interface [14].

In Chapter Two we outlined some of the challenges from which current augmented reality experiences suffer as they try to provide a natural user experience. In this chapter, we will provide the interaction design principles that guided our work on the projected AR interfaces. We also provide an overview of the various interaction techniques we developed to address some of the current drawbacks. Finally, we propose a set of gestures that take advantage of Kinetic I/O.

5.2 Interaction Design

Before diving into the discussion of the actual LuminAR interaction techniques, it is important to review the underlying guidelines they share.

We summarize our interaction design guidelines below. The sections that follow describe these principles in more detail and provide further insights into the actual interaction techniques we developed.

Natural Interaction	Support natural and direct interaction and heterogeneous input modalities such as: multi-touch, natural hands gestures above surface, object-based gestures, standard input devices, voice inputs etc.
Both Hands are Free	Support immersive spatial augmented reality, namely users should be able to interact with both hands without an intermediate device

No Context Switch	Users do not need to perceive the augmented reality experience through a mediating display. Digital content is directly superimposed on the physical environment
Kinetic / Dynamic	Enable steerable, kinetic interfaces that extend the reach of user interfaces. Support relocating, reorienting and resizing of the projected display
Object Augmentation	Enable the detection, tracking and augmentation of objects using top-projection
Just-in-Time-and-Place (JITAP)	Enable Just-in-Time-and-Place interactions based on application, object, and user context

5.3 Kinetic Input and Output

5.3.1 Kinetic Interfaces

Kinetic interfaces are defined as actuated interfaces, in which physical components move in a way that can be detected by the user and form a meaningful interaction [118]. Such interfaces typically employ changes in spatial position of objects, including orientation and position, as well as changes in speed or direction of motion. The discussion in this section deals with Kinetic I/O interfaces, a subset of the general definition of Kinetic Interfaces.

In the domain of projected AR, kinetic interfaces fall into the category of spatial-steerable AR displays. The seminal work in this domain is the Everywhere display [64] project, which we discussed in more detail in Chapter Two.

5.3.2 Kinetic I/O

Kinetic interfaces are currently in a very embryonic stage. Although various researchers laid important groundwork for kinetic-interactive systems [151], it is still hard to outline clear interaction design guidelines for kinetic interfaces, specifically when kinetics meet augmented reality.

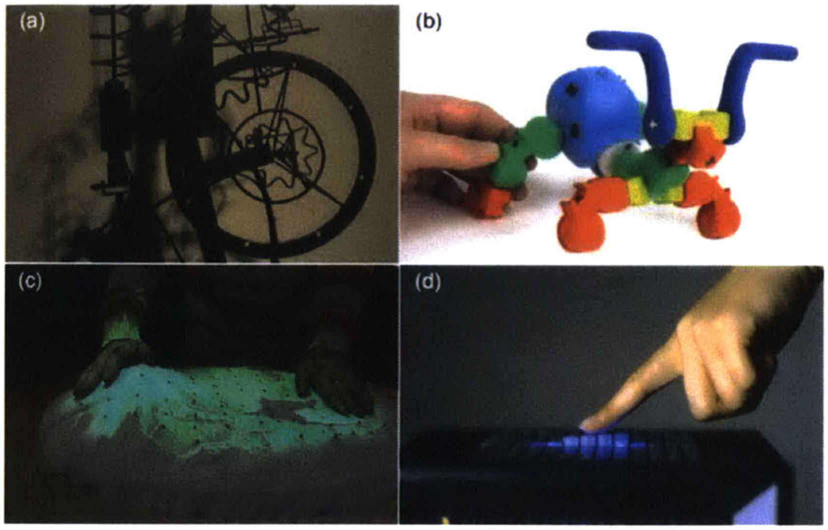


Figure 5-1 Examples of kinetic interfaces: (a) The Dream, kinetic sculpture by Arthur Ganson (b) Topobo by Hayes Raffle and Amanda Parkes **[152]** (c) Relief by Daniel Leithinger **[153]** (d) Lumen by Ivan Poupyrev **[154]**, photograps by Makoto Fujii, courtesy of AXIS Magazine.

However, there are examples for Kinetic I/O systems (see Figure 5-1) that do not involve projection or displays directly. One such example of a haptic kinetic I/O is Salisbury, and Srinivasan's work on Phantom-Based Haptic Interaction with Virtual Objects [155]. The first property of kinetic I/O is simply that interfaces can move, normally in multiple degrees of freedom.

As we described in Chapter Three, the LuminAR Arms have four degrees of freedom that allow them to dynamically manipulate the vertical and horizontal position and orientation of the LuminAR bulb (see figure 5-2). This allows the system to dynamically relocate interfaces easily within the workspace, and increases its interaction vocabulary.

Figure 5-2 Exploring concepts of kinetic I/O: LuminAR Aluminum is used to alternate the projected display between the wall and the tabletop.

From the user's perspective, Kinetic I/O allows natural interaction using hand gestures to position and manipulate the projected display properties and, most importantly, the actual content. Based on that general property we propose in the next section an extension to classic multi-touch interfaces; we call this approach Dynamic Multi-Touch.

Kinetic I/O holds great potential for projected augmented reality interfaces. I have explored this space deeply during my master's thesis work and I refer interested readers there for more information [41].

5.4 Dynamic Multi-Touch

Multi-touch refers to the interaction techniques that implement the detection of hands and simultaneous finger inputs, that allow the control of computer applications. Multi-touch-enabled devices include computer displays, known as "touchscreens," tablet computers, and also projected touch displays. Touch-screen technology is now also in extensive use in mobile devices.

Early examples of multi-touch devices date to the 1980s, when pioneers of the field like Bill Buxton at the University of Toronto developed a multi-touch tablet capable of sensing multiple points of contact [156]. Buxton also provides a good review of the history of multi-touch systems [85]. As multi-touch systems evolved, different hardware solutions were developed, as well as algorithms for finger tracking and gesture recognition. Westerman provides an excellent review in his Ph.D. dissertation [157]. Works like the BiDi Screen [158] and SixthSense [21], represent recent research trends that combine 3D gestures with multi-touch. Work

One of the contributions of this work in terms of interaction techniques lies in the concept of Dynamic Multi-Touch. Dynamic Multi-Touch, an extension of the classic gestures vocabulary of multi-touch, takes advantage of a projected kinetic I/O system. Dynamic Multi-Touch systems utilize actuated DOFs of the projected touchscreen display to support real-time relocating, reorienting, and resizing of the projected display. Dynamic multi-touch has broader application than just pure mechanical driven kinetic UI. It can work also with static (i.e. non-kinetic) projected AR.

Dynamic Multi-Touch tries to extend the spatial limits of the classic screen-bounded interface by allowing it to move; it also addresses the interaction space above the display. In our work, we have prototyped several gestures that explore the Dynamic Multi-Touch concept. We provide details in the sections that follow.

5.4.1 Dynamic Multi-Touch Gestures

The Dynamic Multi-Touch gestures described in this section were initially explored and prototyped using the basic capabilities of the LuXor Vision server modules (see Chapter Three – LuminAR).

It is important to note that while we developed our set of Dynamic multi-touch gestures using the LuminAR Lamp, a mechanically actuated projected AR interface, it is very applicable to non-mechanically actuated configurations. We can loosely define such setups as 'virtually kinetic' systems. However, for such systems to implement the interaction techniques we describe below, the use of a sensors that provide depth information is required. Depth information can provide the required data that describe the 3D volume between the projector and the augmentation surface or object. This information can be used to create virtual kinetic interactions.

Position-Swipe

One of the basic advantages of Dynamic Multi-Touch is the ability to position the display. We have implemented a position-swipe gesture that allows a user to position the projected display using a directional long swipe motion from a source projection area to a destination projection area (see Figure 5-3).

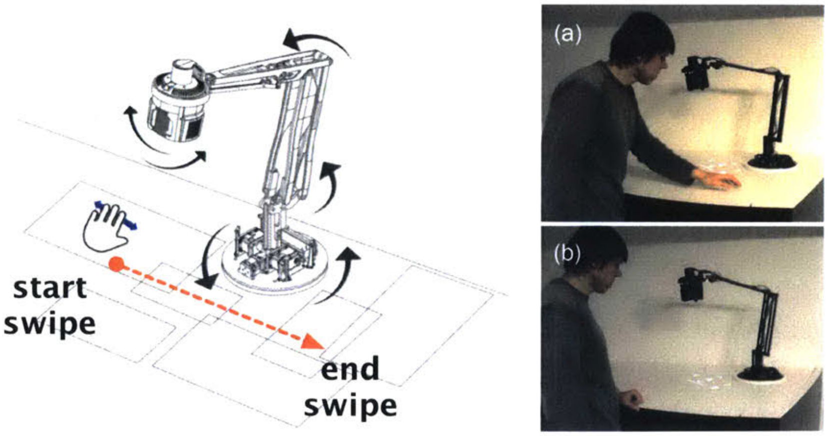


Figure 5-3 Dynamic Multi-touch: Position-Swipe gesture. (a) A user starts the swipe gesture and preforms a long directional swipe (b) after the user removes his hand, the arm follows his swipe trajectory.

To perform this gesture, the user simply selects a start swipe position and begins to move his hand in a steady direction: either right to left, left to right, top to bottom, or bottom to top. The system detects the trajectory of the swipe. The swipe is complete if the user moves his hand away from the viewpoint of LuminAR or if he holds his hand steady under the lamp. The swipe can take advantage of all of the available DOFs of LuminAR. We have used the position-swipe gesture to implement features of the Augmented Desktop applications that we describe in the next chapter.

Swipe-Unlock

Modern operating systems normally implement a login screen that serves as an entry point to the desktop metaphor. This holds for PCs as well as mobile information devices (e.g., smartphones and tablets).

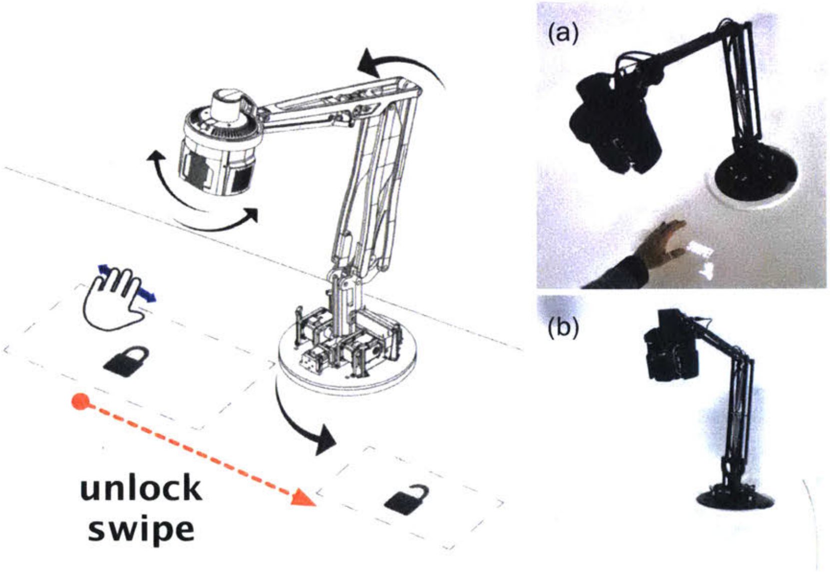


Figure 5-4 Dynamic Multi-touch: Swipe-Unlock gesture. (a) A user starts the unlock swipe gesture and preforms a long directional swipe (b) if the unlock swipe is correctly registered an animatronic sequence is activated and the systems' main menu is displayed.

The swipe-unlock gesture was designed as an interaction entry point gesture. It builds directly on the Swipe-Position gesture we presented above (see Figure 5-4). The gesture makes use of the animatronic capabilities of the LuminAR system; once the unlock event is registered, the LuminAR's can be programmed to perform a motion sequence to alert the user of success or failure of the unlock swipe. When the user swipes his hand under the system, the system responds with a subtle motion combined with unlocking of the projected screen saver.

Touch-Hover

In LuminAR's case, classic multi-touch gestures such as button press and swipe are supported using vision-based techniques that rely on the homography between the projected display and the camera viewpoint.

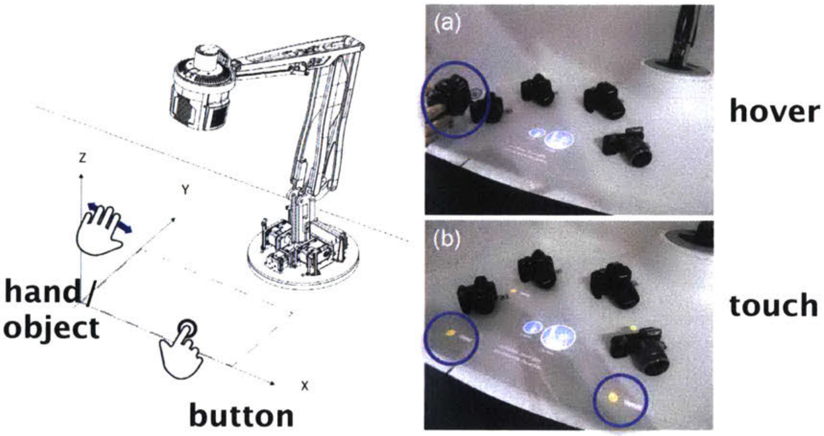


Figure 5-5 Dynamic Multi-touch: Touch-Hover gesture. (a) A user picks up an object, in this case a digital camera. The interaction with the object above the surface triggers contextual JITAP interface (b) that includes projected touchable buttons.

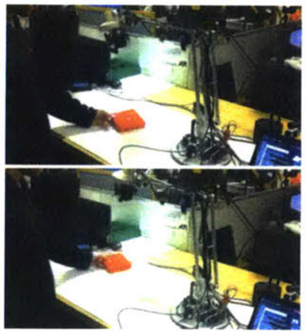


Figure 5-6 LuminAR can also track objects; in this example it is augmenting a red box that serves as a tangible email inbox.

The touch-hover interaction technique (see Figure 5-5) uses the same mechanism. It combines object detection or a gesture that is performed above the surface with standard touch interaction that is carried out on the surface. We have used this technique to implement a specific scenario in the Augmented Product Counter retail application (see Chapter Six). In this use-case, the user holds physical object (e.g., digital camera) above the projection surface, and once the object omission from the tabletop is detected, the system provides contextual JITAP user interface. This interface also includes touchable buttons. Conceptually we are combining interaction with a physical object with digital interfaces. This general concept can be further expanded to facilitate more complex combinations that can make use of the additional Z-axis.

Track Drag & Drop

When we use traditional GUI desktop systems, we take it for granted that we can drag and drop content. We have developed a kinetic-physical drag and drop gesture. The user first selects a virtual object window using a button press or a select gesture; both can be implemented using the LuXor Vision Server modules. Once a selection is made, LuminAR tracks the user's hand.

We have explored different gestures and found that a fist gesture is somewhat natural to use as an indicator for the system to track the user's hand. When the desired location is reached, the user can extend the fist to an open hand gesture to drop the content in the new desired position (see Figure 5-6).

We have used this technique to implement the Kinetic Desktop application we describe in the next chapter. The same technique can also be expanded to a kinetic copy and paste gesture.

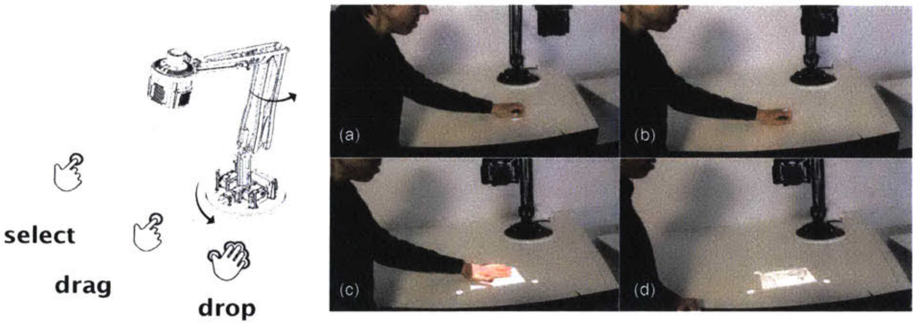


Figure 5-6 Dynamic Multi-touch: Track Drag & Drop gesture (a) A user selects a virtual object by pressing a button, the virtual object is highlighted with a simple border to simplify detection and tracking (b) the user starts a drag motion using a closed fist, the LuminAR arm tracks his hand (c) when the users decided on a desired drop location he opens his hand and hold its position for about a second (d) the system detects the drop gesture and records the new position for the virtual object.

Dynamic Resize/Zoom

Multi-touch systems typically enable multiple DOFs of interaction based on the number of detected fingers. But detecting multiple DOFs is not enough; touch display systems also need to support meaningful gestures. Doing so involves sensing a range of touch beyond simple touch-points. It also involves direction, angle-of approach and vector information [85].

The dynamic resize-zoom gesture (see Figure 5-7) adds additional DOFs to the arsenal of DOFs multi-touch systems already have; not in terms of input (i.e., not additional sensing) but in terms of display configuration output. The gestures below utilize the DOF of the LuminAR Arm or Spotlight to dynamically, and in close to real time, change the geometry, orientation, size, and position of the projected display. It is therefore one of the basic interaction techniques of kinetic I/O as we define it in this dissertation.

The dynamic resize operation works as follows: a user is touching two corners of the projected content, generating a motion vector between the two touch points. The vector 2D orientation and scale determine the physical position and size of the projected display. This technique can be used to resize and reposition the entire display, or to zoom on a specific section of it.

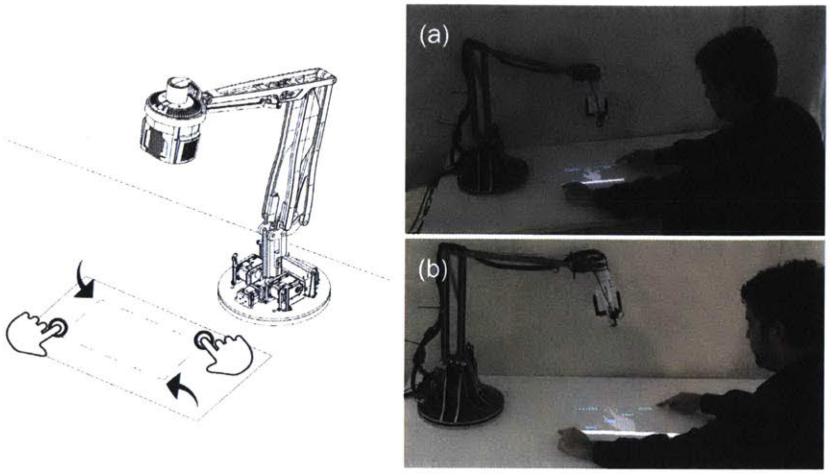


Figure 5-7 Dynamic Multi-touch: Dynamic Resize/Zoom gesture (a) a press down on two corners of the projected display, triggering a resize gesture (b) as he swipes away from corners, the arm moves up in proportion to the distance his fingers moved from the original location, causing the projected display to grow in size.

Dynamic Rotate

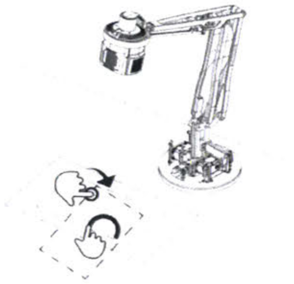


Figure 5-8 Dynamic Multi-touch: Dynamic rotate gesture design

Building on the principle of the dynamic resize-zoom gesture, we can design an additional gesture that would support a common orientation change operation (see Figure 5-8) that is very common and useful. Unfortunately, we did not have time to fully implement this gesture, but we include its design here as one of the Dynamic Multi-Touch possible future gestures.

This gesture starts when a user is pressing down on a pivot point with one hand, while using his other hand to perform an arc-like gesture on the surface. The direction of the arc determines the rotation direction desired. The notion of pivot-finger based gestures should be explored further and extended with gestures that support content scaling, navigation, and positional directions.

Tap-Focus

The tap-focus is another LuminAR position-setting gesture, designed to provide the user with a method to position LuminAR (see Figure 5-9). It is a simple gesture, requiring only a tap or double-tap (if needed) to cause a selection of an area-of-interest that is then magnified.

Double taps can be distinguished from simple button presses using application logic. The LuXor application framework can define software timers and events that can be used to develop application that respond to the different tap events (i.e. double-tap, tap, and long-tap).

Optionally, this gesture could be chained with additional taps that would allow a user to easily choose between two zoom levels with just a single tap. This gesture also provides clear and direct kinesthetic feedback, as the LuminAR arm would move according to the user's desired zoom level up or down.

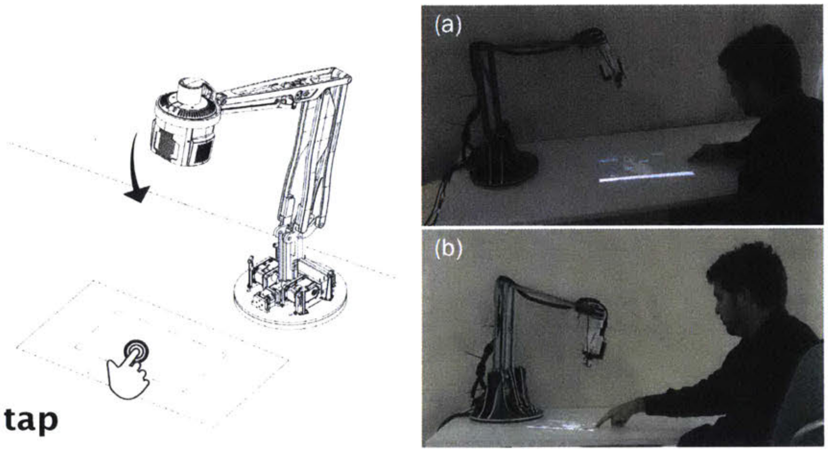


Figure 5-9 Dynamic Multi-touch: Tap-Focus gesture (a) double taps a point of interest on the projected surface (b) the system responds with focusing the projected display around the location of the tap.

5.5 Interaction Explorations

Earlier in this chapter, I presented the interaction design guidelines I developed during this thesis work. I also described Kinetic I/O and Dynamic Multi-touch, two large-scale interface design efforts. As we hardware matured, and with our LuXoR and Lens software frameworks reaching a viable feature set, we were able to explore several interaction techniques that I will describe in the sections that follow.

5.5.1 Just-in-Time-and-Place Interactions

Context-aware interfaces hold great promise for transforming the utility of computing. They encompass the capabilities to detect and react to changes in state of the environment. Without them, computer systems are static and require human users to initiate and manage all interactions. Understanding context allows a system to respond to a specific user or environmental state accordingly.

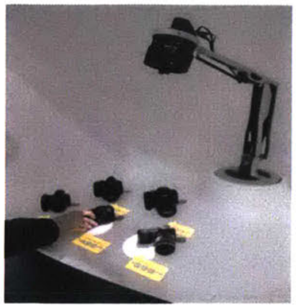


Figure 5-10 (2) Cameras are augmented with price information

Just-in-Time (JIT) *interactions* are a specific branch of context-aware computing **[159]** that attempt to enable computers with contextual knowledge to offer relevant information when and where we need it. However, 'where' in the definition of JIT interaction refers to how we use computers today, namely via display-centric devices and interactions. This means that 'where' is actually on a computer screen window or via a mobile phone's push notification.

LuminAR enables an extension to the classic Just-in-Time interface definition. We are proposing *Just-in-Time-and-Place (JITAP)* interfaces. The notion of 'Place' in our definition refers to a physical location. It is possible for the system to define and recall several locations that embed specific projected information as defined by the user. For example, LuminAR can be programmed to save and retrieve physical locations for different projected applications. Applications can then be invoked in the specific location when they become relevant; for example, an email application can appear in a fixed location whenever a new message is received.

Moreover, the system can utilize its animatronics capacity to alert the user. Naturally, traditional UI modalities such as sound and graphics can also be combined to fully complete the experience. The result is a unique actuated ambient interaction.

In Chapter Six, we describe how this interaction method was put to work in the case of the Augmented Product Counter and the Augmented desktop applications, for example.

5.5.2 Device Integration

The LuminAR bulb is a wireless computer. It is capable of communicating with other devices in its vicinity, allowing for interaction to extend across device modalities. We have explored cross-interaction scenarios between multiple LuminAR bulbs, mobile devices, and laptop computers.

5.5.3 Digital Glue Device

An easy metaphor to consider in relation to cross-device integration is glue: think of LuminAR as a "glue device." LuminAR does not attempt to render laptops or mobile phones obsolete, but rather amplifies and complements their use. According to context, the system can suggest and facilitate data transfer across devices. Rekimoto's Pick-and-Drop **[160]**, Augmented Surfaces **[59]**, and Butz et. al **[161]** EMMIE (Environment Management for Multiuser Information Environments) system, all provide inspiration to this approach and inform the key interaction. Our work contributes the specific usability aspects of a projected dynamic multi-touch display.

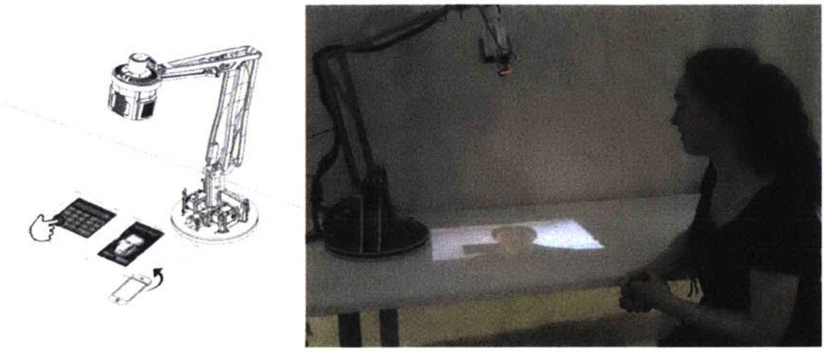


Figure 5-11 LuminAR as a "glue device"

We include a few examples below:

- An application where the user is able to physically gesture to the lamp to transfer a document that is currently open on his laptop to the surface next to his screen, yet flat on the desk. This can serve as a virtual stack of documents that is pending, waiting for user inputs and manipulation.
- When LuminAR detects a smartphone in the near vicinity, the system can then initiate a wireless data exchange session. The mobile device could then stream an application to the LuminAR device.
- A calculator application projected on the tabletop (described in more detail below).

If a call comes in in this mode, the incoming call event can also be projected next to the application while the calculator application is still running. Clearly in some cases, it would be desirable to have mobile device content displayed on a large screen.

5.5.4 Data Sharing

The basic notion we have explored is sharing data and context between LuminAR bulbs–in this case we can program the system to form a network between LuminAR bulbs. Applications can then share state data and respond with relevant application content and actuation. Good examples of this behavior is described in the next chapter when we discuss Spotlight, white:scape and Swyp.

Next, we can clearly identify merit in sharing data captured by the LuminAR bulb between a mobile phone and a laptop, and even between bulbs. Since the bulb software stack uses web-based folders, it is easy to accomplish simply by sharing URLs.

This opens the door for integration with any web application that is relevant for the specific data captured. For example, we can integrate a publishing feature to the Scan Application (see Chapter Five – Applications), scanned images can be automatically uploaded to an online document service like Evernote.com [162].

5.5.5 User-Interface Leeching

Not all interfaces are equal, and not all application scenarios require the same input modalities. This was the guiding principle of a technique we call 'User-Interface Leeching.' Mobile devices and tablets have great input capabilities, and wireless keyboards are great for typing. Since LuminAR is simply a computer it is possible to leech on such input devices and have them function within the context of a LuminAR application.

Figure 5-12 User Interface Leeching

We have tested an example of such an interaction using a touch-enabled smart phone. We used the phone's user interface as a handheld controller for LuminAR. Flicking pages back and forth on the mobile phone touchscreen caused content to change on LuminAR's projected display accordingly.

5.5.6 Object Augmentation

In 1999, Jun Rekimoto et al, presented their work on Augmented Surfaces [59]. They contributed several new interaction techniques, among them the concept of object auras. They focused on augmenting physical everyday objects with digital information. Inspired by this work and other work that followed, the LuminAR user interface's vision and tracking abilities enable it to generally augment objects that are in view.

For example, a printed magazine advertisement could be detected and linked to a special online-based video widget that contains further information about the product and an interface to interact with the content, i.e. save as favorite, send to a friend, or order the product.

Figure 5-13 Early object augmentation explorations using LuminAR Optimus. The lamp detects objects like a magazine or a can of soda using simple fiducial markers, once detected the objects are augmented with relevant web content.

Interactions based on physical object augmentation also open the door for many new physical object-based applications. Detecting and tracking physical objects can help produce valuable metrics, providing insights on usage patterns and user behaviors. This would be desirable for many real-world applications. During this work, we have explored augmentation of books, toys, hand tools, books, and workbenches. In the chapter eight we go deeply to the use of this idea in the context of an educational use case.

It is possible to outline how traditional user interfaces map to augmented interfaces using physical objects. In the list below, we describe some examples of possible future augmented interactions:

- Search, bookmarking, physical copy and paste, and annotations of printed material (note that this could also work for digital inputs).
- Physical hyperlinking: objects and gestures can invoke web access or email composition. For example, if a business card is placed under the LuminAR Lamp, the card can activate the address book application automatically.
- Integrating with passive I/O devices: for example, if a simple pen is registered by the LuminAR Lamp as the invoking object for a note-taking application.

5.6 Projected AR Design Case Study

Before the large user studies are presented, I introduce in this section a projected AR application case study. Developing projected AR applications is not a simple task. In practice, and given the current state of technology, available tools and techniques – application development that spans the digital and the physical worlds requires several iterations. The design considerations are highly dependent on the application context at hand. In sections that follow, I will illustrate some of the design thinking that has to happen when building a real-world application of a projected interface by going through a concrete example of the retail application we built. The full application description is included in Chapter Six. We hope this discussion of the steps and iterations needed can be insightful to others. The documentation and conclusions we present below, can potentially inform similar design efforts for projected augmented reality interfaces in domains other than retail.

The Augmented Product Counter (APC) was developed in collaboration with Intel and Best Buy. Intel provided initial background information for the project with their concept work "The Responsive Store" [163].

5.6.1 Understanding the Interaction Space

Early in the design process it was important to understand the physical constraints we had to consider while designing the projected GUI for APC. To accomplish this, we used several mock ups of test projections and made measurements that informed our design process (see Figure 5-14).

Figure 5-14 Designing the APC interaction space: (a) measuring users' reach (b) measuring product distribution and projection real-estate (c) the APC foam-core mock up

5.6.2 Work envelope sketches

One of the key metrics that was crucial to define for the UI design to complete was the operational envelope of the LuminAR Retail arm in relation to the user. We accounted for parameters like user reach, distance between products, projection angles etc. To fully test our design and integration of hardware, software and interface, a foam-core APC model was built and used extensively.

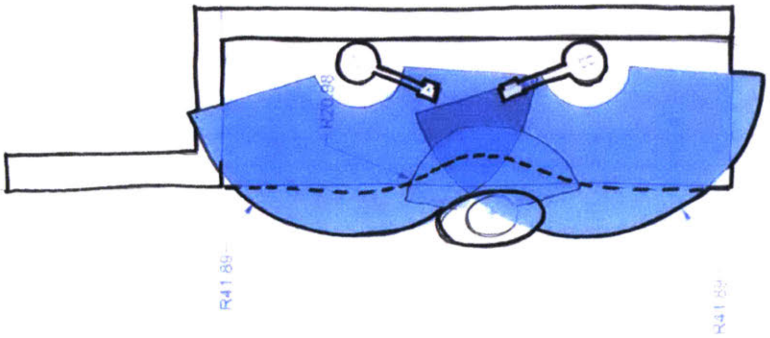


Figure 5-15 Work envelope sketches

5.6.3 Projected GUI Guidelines

Creating effective projected user interfaces is not an easy task. The designer must compensate for many elements, such as legibility, color-clarity, shadows, occlusions, lighting conditions and so on.

We used several different projection test patterns to understand what we can expect from different projection parameters. We also explored non-rectilinear graphics projection, to ensure that the view orientation is always correct regardless of the projection angle.

This is a design approach for solving the skew and keystone problems to which projected user interfaces are prone. In our design process, we defined a set of principles or guidelines that can assist in the design of projected user interfaces, through several testing sessions:

- **Avoid white, pad with black:** projecting white demands the most from the projection hardware; it also takes away luminosity from the rest of the scene, so use it scarcely. However, the color black is your friend. Generally, using black to pad and outline your GUI elements will result in a clearer and brighter projection.
- **Design for dynamic scales:** projected UIs need to be able to scale dynamically with respect to the projected screen size, specifically when the projection setup changes. This is very common when the system is kinetic (e.g., in LuminAR's case).
- **Use non-rectilinear graphics:** this is a design approach to avoid handling complex image skew and keystone correction.

- **Minimize shadows and occlusions:** place UI elements as much as possible in locations that minimize the chance that users would reach out and occlude the interface, typically the edges of the projected screen.
- **Use big and legible fonts:** always test fonts as a function of the projected display. We found that for a 20" display at WVGA resolution, fonts less than 16p would render poorly.

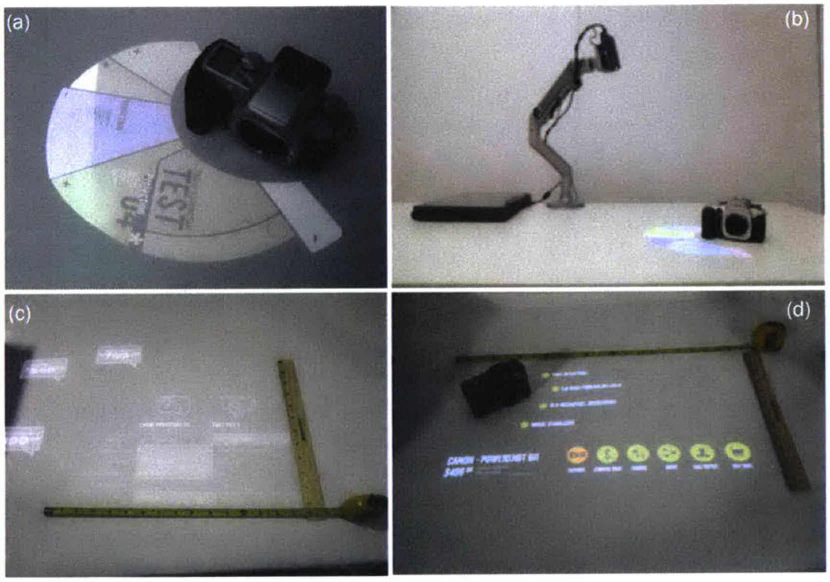


Figure 5-16 Deriving projected GUI guidelines: (a) and (b) show non-rectilinear projected GUI (c) shows the legibility challenges with certain light colors (d) shows a completed 22" projection using big round buttons, white fonts and bright colors over a black background.

5.6.4 Projection Surface Materials

We explored different materials to test their application as a valid projection surface for APC. We were specifically interested in various properties of glass, acrylic and laminates.

Clearly, material parameters such as thickness finish and color impact the quality of the projection. In our experiments, we discovered that matte laminates provided the sharpest images (see Figure 5-17). Reflective and back treated materials had a nice result as well, but with a reduced image quality that blurred with increased thickness. Projection reflective paints are also a good option, but require application and are naturally more expensive.

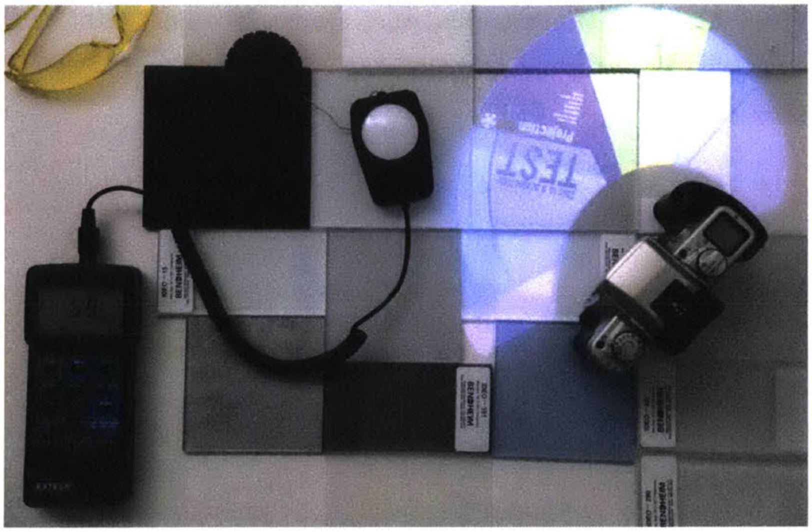


Figure 5-17 Testing projection materials for the APC counter-top

5.7 Conclusion

In this chapter, we introduced the concepts of Kinetic I/O and Dynamic Multi-Touch. We have also included details of the initial explorations designing novel gestures that use projected AR interfaces.

In our informal user testing, a general pattern of behaviors emerged. Initially, users expected the system to behave like a common touch screen. We can attribute this to the strong presence in 2017 of mass-market consumer electronics that use touch screens. The findings in our augmented education case study (See Chapter Eight) seem to support this observation. However, and sometimes to our great surprise, users were very comfortable continuing to interact with LuminAR after they discovered the touch screen can actually move around the workspace, and that gestures can affect the digital information projected on physical objects.

Many of our early users easily picked up the interaction techniques we described in this chapter. However, as with any new interface, we had challenges. For example, some of users did not respond as well to the fact that a moving robotic arm might potentially hit them during the interaction process. Naturally, this is a topic in need of further research.

In addition to the interaction techniques we explored, and based on the feedback we received from our test users, it is clear that additional gestures should be designed to complete the Dynamic Multi-Touch grammar. The list below contains some ideas for such future gestures:

- High-fidelity hover gestures that accurately return X, Y, Z coordinates of the arm, hand, and finger positions.
- Dynamic hand gestures based on training and detection (e.g., open-hand, thumb-up, thumb-down).
- Hover-pinching gestures. Such gestures use a pinch gesture in arbitrary Z height above the projection surface. These types of gestures can facilitate object interaction.
- Finger tracking in all Z-levels of the interaction space.

The work in this area is still preliminary, however, our initial results show promising potential that could be used in the near future, when depth-sensing hardware becomes ubiquitous and computer vision becomes hardware accelerated.

6 Applications

To complement our design and engineering of the LuminAR system, we were constantly and in parallel engaged in creating applications that test the system, its technical function, and overall usability.

Projected AR interfaces can support a wide range of application domains. As AR interfaces emerged from research labs, in many cases the mass consumer applications focused on annotating physical objects with media content, mostly via mobile, hand held devices. The specific applications were in the education, advertisement and gaming domains.

In recent years, as head mounted AR displays became a reality, new and exciting applications of AR emerged. Such applications include immersive experience in the domains of physical simulation, collaboration and education. There are several examples where AR is in use in the general field of human guidance. For example, some modern cars already include Heads-up AR displays that provide the driver with information such as speed and driving directions directly on the wind shield.

At the core of the process driving this thesis was field observation of concrete AR use cases in situ. I was intrigued to discover how projected AR interfaces will be received and utilized in environments where such interfaces never existed. This approach led me to explore several application domains.

In this chapter, I describe in detail the application domains I focused on in the early stage of this thesis work. I first describe our exploration into the domain of Work and Productivity. In this area, we developed new augmented desktop scenarios, collaboration applications, data exchange methods and kinetic interaction techniques. The domain I focused on, prior to the execution of the two large scale studies that conclude this dissertation, was the Augmented Product Counter (APC) project. This project explores the domain of augmented retail. I conclude this chapter with a brief discussion of the user feedback we received when testing out these applications with real users.

Finally, this early work informed the large-scale application development projects and experiments that followed in the domains of augmented manufacturing and augmented education. I discuss both case studies at length in subsequent chapters.

6.1 Work and Productivity

6.1.1 Augmented Desktop

Reading, writing, and interacting with standard computers are the most commonplace tasks for desktops activity, and given the abundance of previous work (which we detailed in Chapter Two), it was natural to explore this domain with our system.

Part of the motivation for this choice also lies in the fact that LuminAR is unique in form factor, as it is embedded in a desk lamp. It is therefore a form of a digital computer that on the one hand is embedded in your space and on the other hand does not take up "desktop real-estate" like laptop computers do, for example.

We first focused on familiar interactions with digital media and information, developing projected touch-enabled widgets. However, immediately after exploring standard use cases, we shifted our focus to scenarios that blend modalities, taking advantage of kinetics, top-projection and object augmentations as well as LuminAR's networking capabilities. In this section, we provide details of the features we developed.

Projected Widgets

LuminAR's first use case was to provide an interactive augmented space. To demonstrate this functionality, we focused initially on developing a set of general purpose projected widgets capable of performing everyday tasks (see Figure 6-1). We created a picture browsing widget (using a similar technique to Apple's OS X Cover Flow feature), a media player and a scrolling text widget. These widgets also serve as the building blocks for the rest of the LuminAR applications.

As with all other LuminAR applications, these widgets are web-based, and designed to incorporate text, video and images. The interaction with the widgets was primarily based on touch events, but also includes hand gestures. For example, a user can scroll a large amount of text by swiping his hand under the lamp.

We envision such projected widgets used in standard desktop environments, working in conjunction with a PC. In such cases, the widgets can be used to enhance an existing software interface or serve as an additional contextual display for information. But not less interesting is the case when LuminAR is installed as a standalone object without a computer. We can image having a LuminAR bulb installed in a kitchen and used to augment the countertop.

Following this logic, in one of our early experiments, we explored a concept of LuminAR as a reading lamp. It was used to augment a magazine, enabling digital-physical cross-media experiences. In our example, a physical printed advertisement invokes a website that allows a user to watch videos of the product.

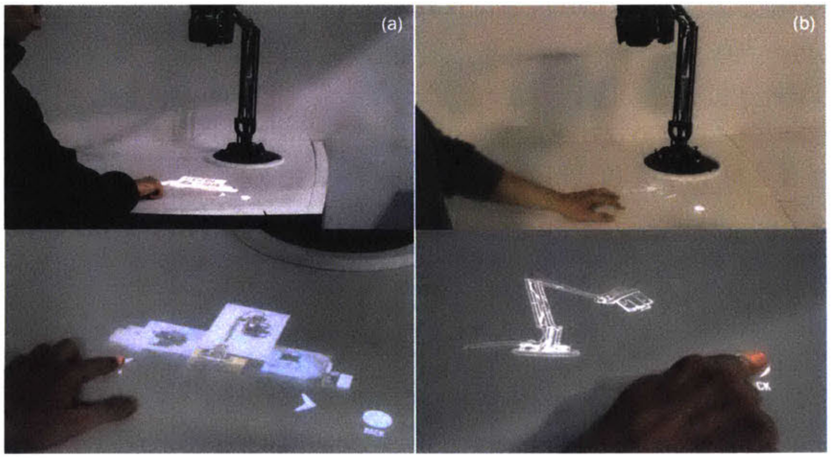


Figure 6-1 Augmented Desktop Projected Widgets (a) Cover flow widget (b) Video player widget

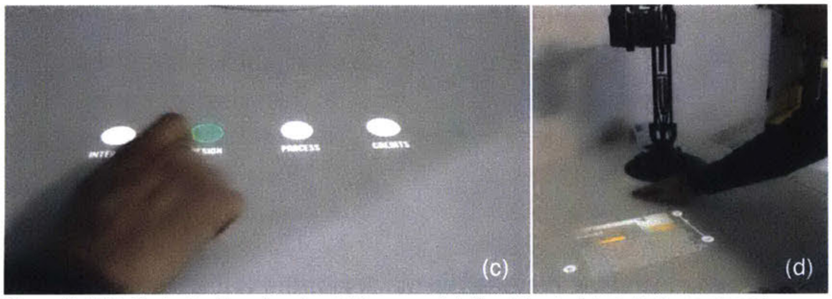


Figure 6-2 (c) Projected touch relocatable menu (d) Scroll text widget that uses hand gestures to scroll large amount of text.

Scan Application

The LuminAR Scan application allows users to capture images of objects and projected images on the tabletop. It is designed to provide an instantaneous scan function that does not involve a dedicated scanner or a relatively complex sync with a digital camera.

Concept Design for the Scan Application

Using a simple interface, a user can place the object under the system and execute a scan (see Figure 6-3). The interface also supports zoom functionality. Once an image is captured, the user can resize the result, and project it on the desktop.

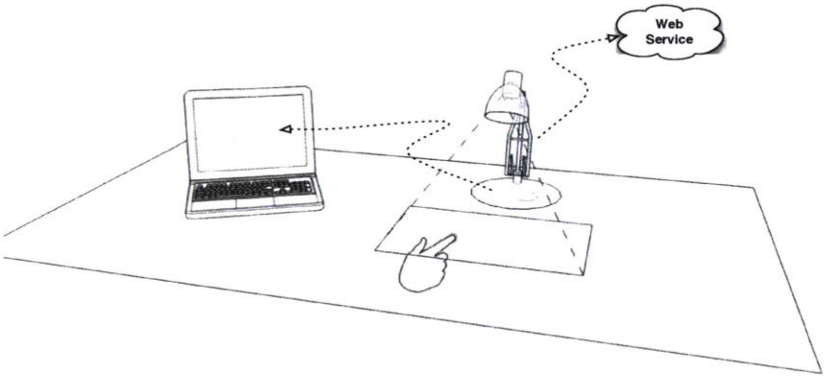


Figure 6-3 Concept design for the Scan Application

This enables the user to place a different object, and scan again. This interaction design allows the composition of complex images based on multiple scans. We call this technique *Scan-Project-Rescan* (See Figure 6-4) .

Finally, the scan application allows for two different LuminAR systems to share the scanned content. We implemented this feature using a simple web-folder as a destination for posting scanned image results.

When a share request is executed, we simply point the destination LuminAR to the desired image URL. The Scan App is very simple example of how LuminAR can be used to support remote collaboration.

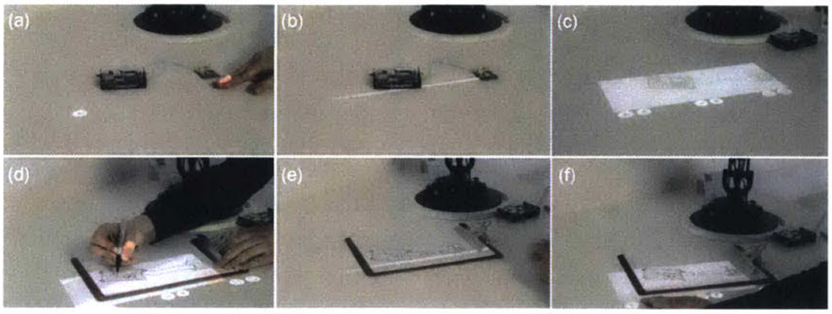


Figure 6-4 Augmented desktop scan application: (a) an object is placed under the LuminAR Lamp, when a user hits the 'scan' button (b) the object is scanned, (c) the user can adjust the result using zoom controls, (d) the image can be projected to a paper and (e) rescanned into the system. (f) Finally the user can press the 'share' button to send the image to a web service or another LuminAR device.

The scan application we developed is a good example for the future of augmented collaborative interfaces. We can imagine how such features create new possibilities for remote communication by virtually sharing a physical desktop and objects while augmenting them with digital information in real-time.

6.1.2 Kinetic-Spatial Augmented Desktop

To fully explore the concept of Kinetic I/O we describe in Chapter Four, we designed a novel type of an augmented desktop system that takes advantage of the LuminAR interaction techniques. We call this application the Kinetic-Spatial Augmented Desktop (see Figure 6-5).

Figure 6-5 Location aware menu system; top-level menu is projected in the 'home' position. Subsequent sub-menus have their own location relative to the home position.

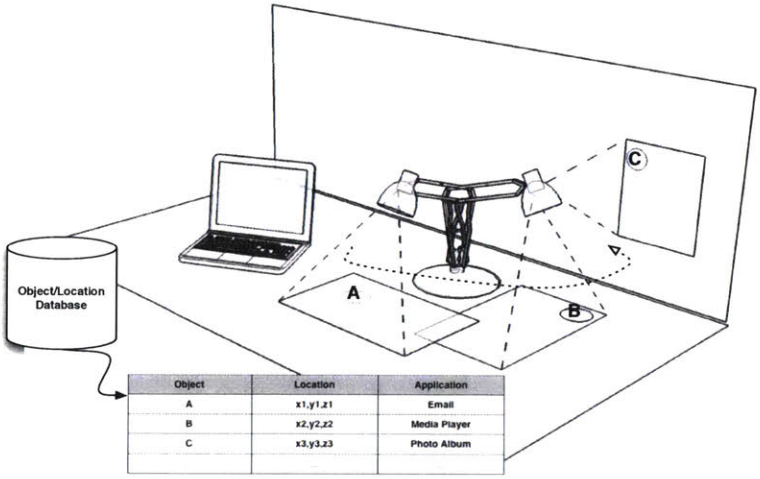
The application attempts to reclaim the desktop metaphor that was claimed for the digital desktop to a physical desktop. In other words, in our system digital content can be spatially located on an actual desktop, thereby blending the physical and digital worlds more closely.

To explore this concept, we implemented a location aware menu system. The interaction begins with a main menu entity that appears in an initial entry point location. When subsequent sub-menus are invoked, the projected display will move accordingly to location relevant to the specific submenu functionality. We can program the system to locate a menu according to context, the position of an object, or the user's hands.

We used this feature to implement a conceptual demo application that showcases the different LuminAR projected widgets and kinetic demo applications, using our own system for this purpose.

We also designed an interface that allows users to relocate specific widgets in the workspace. Our demo setup included different live web apps that streamed content from the web. We developed a weather widget, a YouTube widget and a daily Twitter feed widget. Each widget has a different location assigned in the physical space.

The applications were arranged in a dock element, similar in principle to existing application dock bars that are very common in WIMP-based operating systems (e.g. Mac OS X or Windows 7). The dock enabled the user to invoke widget and load them to the projected desktop. Once a widget is selected, its location information is retrieved, directing the LuminAR arm to the desired rendering location.



Object	Location	Application
A	x1,y1,z1	Email
B	x2,y2,z2	Media Player
C	x3,y3,z3	Photo Album

Figure 6-6 Concept design for the kinetic desktop object/location pairing

Once a widget is located in its new position, the user can then select to activate a different widget. In this case the user selected a YouTube video widget. LuminAR servos to the stored location for the requested widget and displays it (see Figure 6-6).

The user can then choose to relocate the application by pressing the move button. To implement the relocation function, we developed a drag and drop gesture. The user can drag the application to a desired new location using a drag gesture that is tracked by the system. The actual dragging involves the LuminAR arm actually performing visual-servoing while tracking. Once a desired location is reached, the user simply needs to hold this position for about a second to conceptually drop the application in the new location. Finally, we implemented a swipe-gesture that allows the user to use simple swipes to position the projected display.

This exploration provides a glimpse of the potential for actuated interfaces. It builds on the notion that human cognition is spatial, and therefore kinetic enabled interfaces can possibly assist with tasks that involve information retrieval and recall. For example, future applications can support virtual piles of documents on a physical space stored in different locations.

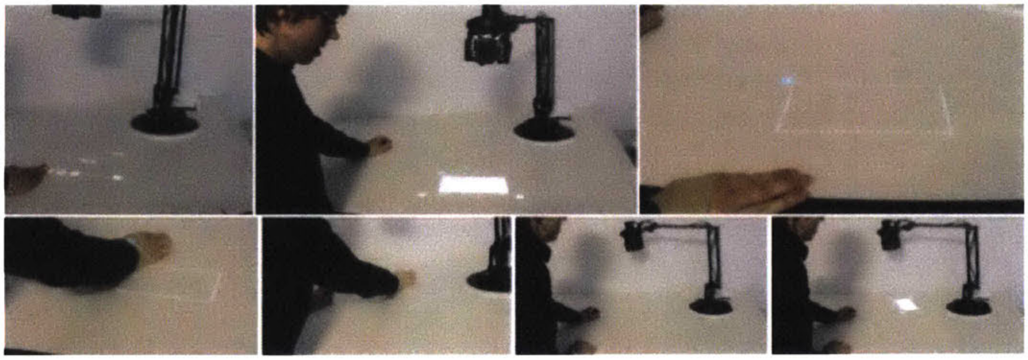


Figure 6-7 Kinetic desktop prototype; the user can select to move a widget (in this example we show live twitter feed widget) by pressing the 'move' button. Once the widget is selected, the user can simply drag the widget frame to a new location in the workspace. The 'drop' can be detected by either removing the hand or detecting a change in the hand posture, for example, a 'fist' could be used to perform the 'drag' function, while an 'open hand' gesture can signal a 'drop' function.

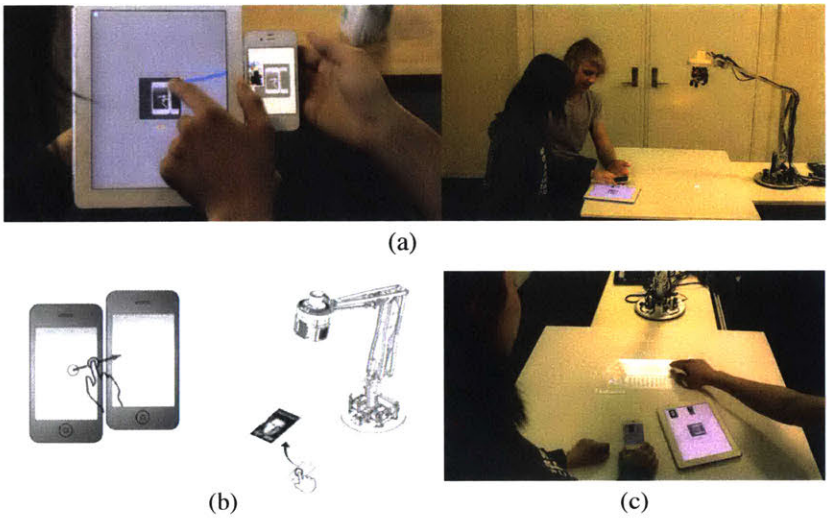
Figure 6-8 Once a widget is located in its new position, the user can then select to activate a different widget. In this case the user selected a YouTube video widget. LuminAR servos to the stored location for the requested widget and displays it.

6.1.3 Swÿp

Dating back to Rekimoto's Pick-and-Drop and Augmented Surfaces projects **[59]**, **[160]**, much work has been done towards pseudo-physical data-exchange between applications and devices. More recently, gestural-kinetic implementations have strived to make users' interaction with content more physical through implementations like LuminAR. Many applications supporting data-transfer and collaboration suffer in either pairing devices, or lack an obvious physical paradigm to easily enable users to transfer data. The transition to collaboration has moreover been recognized as critical for remaining seamless to the user **[164]**. Finally, novel interaction techniques such as SPARSH by Mistry et al. **[165]**, leveraged the information cloud and the user's body as a medium for fun and intuitive data exchange.

Swÿp is a framework facilitating cross-app, cross-device data exchange using physical "swipe" gestures. Our framework allows any number of touch-sensing and co-located devices to establish file exchange and communications with no previous pairing other than a physical gesture. Users can physically pull content across devices: with this inherently physical paradigm, users immediately grasp the concepts behind device-to-device communications. Relevant gestures for natural interaction vary in context and from device to device, but Swÿp supports each as an open framework. Swÿp interaction examples are shown in Figure 6-9 below.

Swÿp can be integrated within any device supporting a network connection and operates by matching network operations with the dissected segments of the swiping gesture between a source and a destination device. Devices can serve as either a source or a destination for swipe events. An out-swiping device registers as a data provider on available networks and global Swÿp servers, and a swipe-in destination device connects to the data channel of the relevant available services.


(a)

(b)

(c)

Figure 6-9 (a) LuminAR + Swÿp interaction scenario; (b) shows the key physical pairing interaction supported by the Swÿp framework across two mobile touch enabled devices; (c) shows a scenario of swiping content to the LuminAR augmented surface interface, users are collaborating to make digital postcards using LuminAR with Swÿp.

The notion of cross-app communication also makes Swÿp compelling. By supporting different content types, all Swÿp-enabled apps can automatically interoperate with no action required on behalf of the user or application developer. To provide this capability, Swÿp simply mandates that all digital content support representation as one common image format. Swÿp was also shared with the community as an open source project **[166]**.

6.1.4 white:scape

white:scape is an exploration of augmenting the workplace. We have built a collaboration tool designed to explore the use of multiple LuminAR projector units. Our goal was to create a workspace that allows users to immerse themselves in a collaborative environment, first with large projected interfaces that are shared compared with devices such as laptops that are inherently personal.

The system was installed on furniture set up from the Steelcase Turnstone line that we have adapted to host three LuminAR devices, the setup is Figure 6-10.

The application we built to test this new setup was designed to facilitate Design Thinking processes [167]. Specifically, we focused on the needs of co-creation of digital content.

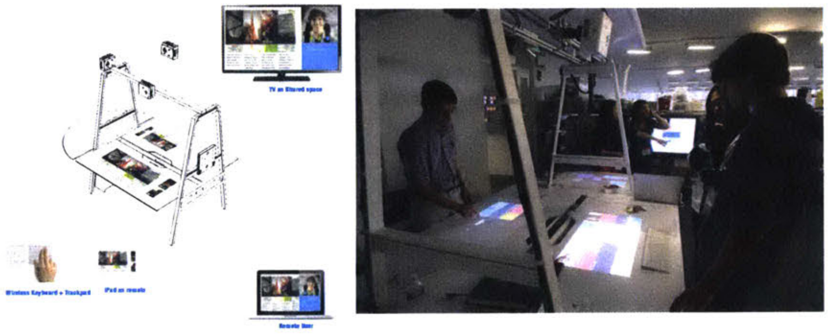


Figure 6-10 The white:scape system design and implementation

The interaction technique we designed uses a key matrix metaphor. The matrix serves as a work space for adding and organizing shared information and among several users. Our design approach was influenced by the popular iPad application Grid [168], and our work attempts to bring the interaction principles to the projected AR domain. In our system, a content matrix can span different types of devices such as the LuminAR devices, tablets, and mobile phones as well as laptop computers running standard browsers. Users would first select an area on the matrix and then add text and images. It was also possible to add snapshots of physical items to the matrix space. Because the data are in a matrix form, we could incorporate concepts from the spreadsheet by adding tags and categorizing the pieces of information you have in the space. We could then analyze the data quantitatively or qualitatively through various design methods.


(a) (b)

Figure 6-11 (a) white:scape installed Bivi table from Steelcase with a shared screes (b) the cross-browser white:scpae app running on LuminAR and COTS tablet, showing design thinking functionality using templates.

The system also allowed the creation and use of templates. For example, a user could select templates for a specific design method, such as a User Persona template, that would be useful for representing different user types when designing a new product offering. After creating the content, users could share it on the common display for everyone in the same work space. It was also possible to incorporate browser-based teleconferencing to the application to further support remote collaboration. We have deployed a full white:scape prototype installation in our research collaborators' office at Turnstone. Two sets of the system were deployed at the headquarters of Steelcase in Grand Rapids, Michigan and were used as demo systems.

6.1.5 Spiral Campfire

LuminAR integrated into the Campfire Big Lamp **[169]** from Turnstone, a furniture brand from Steelcase, to create an interactive and collaborative work environment. While white:scape was designed to mostly explore remote collaboration scenarios, Spiral Campfire was designed as a collocated experience, allowing users to experience a collaborative content creation tool for enhanced projected AR.

Campfire is a unique furniture offering, described by lead designer Kurt Martin as such: "We designed Campfire to create a space within a space . . . to create a familiar place where people can gather and collaborate informally with tools that are not high tech. In fact, their low tech—simple."

Campfire allowed users to gather as if in a circle around a bonfire and to create and explore digital content together. We hacked into Campfire Big Lamp by putting our projector and PrimeSense sensor (for sensing and camera) in the center of the lamp. We also routed the power and cables along the lamp arm and stored the computer in the lamp base. RGB LEDs were also integrated into the lamp (future work will involve synchronizing lamp color with the Spiral Campfire app).

Our exploration was mainly targeted at deploying a LuminAR system on a larger-scale light object, and using our software framework to quickly deploy a circular projected AR interface that matched the physical properties of the host furniture. The work is motivated by the extended body of work on visualizing time-oriented data such as Continuum [170] and TimeZoom [171]. Specifically, Shen's DiamondSpin Toolkit [172] , upon which our work relies for some guiding principles with respect to multi-user around-the-table interaction metaphor implementation. Another very relevant work is PhotoHelix [173], PhotoHelix demonstrated similar spiral-based photo organization interface on using a multi-touch table top. It provides an excellent basis and inspiration to our work.

The Spiral interface provided contextual menus that enable users to view the interface regardless of their position relative to the table. Spiral's main theme was the spiral display of a timeline. Users could zoom in and out of the spiral to see fewer and more events on the timeline. The timeline itself is color coded to indicate night vs. day vs. sunrise/sunset. Colored blocks in the timeline indicate when an active session occurred at the table; orange stripes indicate when a photo was taken of the space. In "canvas" mode, users can lay content out on the table (e.g., by drawing directly on the table, placing physical objects on the table, etc.). Users could "take a picture" of whatever was created in "canvas" mode. Users can view images previously captured using "slideshow" mode; in the future, users should be able to bring up images directly by interacting with the colored blocks in the timeline. Figure 6-12 below shows the system setup and examples of the Spiral Campfire user interface.


(a) (b) (c)

Figure 6-12 (a) Steelcase Campfire Big Lamp; (b) users playing with the Spiral collaboration application; (c) elements of the Spiral user interface

Spiral Campfire enabled users to document and capture any content that they wanted saved while allowing people to still create content in the tangible ways they are used to. Similarly, the system expands the capabilities on content creation, because now users can add to previously recorded content but not necessarily permanently alter it, because each step of the content creation can be saved separately. The timeline puts people's work, ideas, and innovations in the context of time. This is interesting because being able to observe one's use of time can influence how one plans the future use of time.

6.2 Augmented Product Counter

The Augmented Product Counter (APC) was developed in collaboration with Intel and Best Buy. Intel provided initial background information for the project with their concept work "The Responsive Store" **[163]**.

6.2.1 The Need for New Retail Interfaces

In the past ten years, first the web, and soon afterward mobile devices, generated two tidal waves that disrupted the traditional shopping experience. Shopping transformed into a social process of discovery, powered by tools for comparison, sharing and purchasing. Retailers adopted technologies to remain relevant, using the new online channels. However, retailers need to find means to engage customers in physical brick and mortar stores. This new concept was also previously explored by Sukaviriya et al **[163]**.

Figure 6-13 Intel's Responsive Store concept – augmented product counter (source: Intel)

The challenge that the APC project addresses is how could brick and mortar stores evolve to use interface technologies that allow them to remain relevant in an age of online shopping.

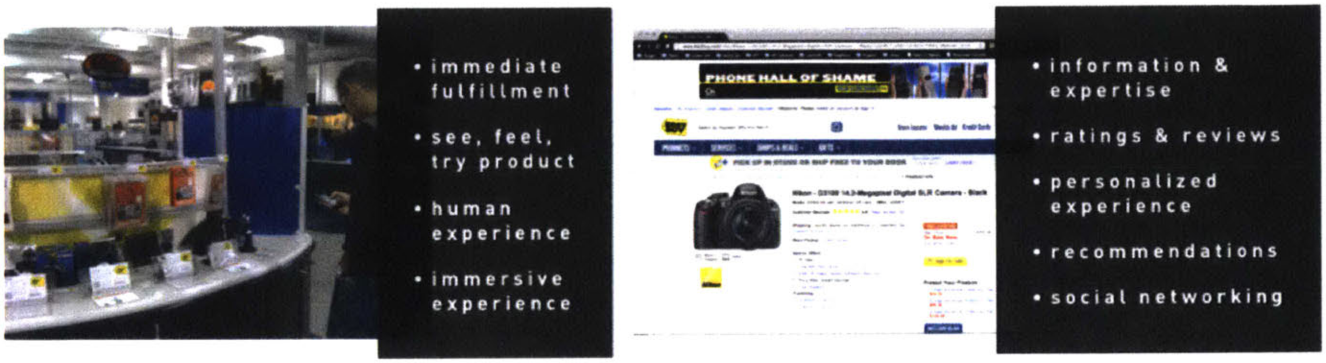


Figure 6-14 Brick and mortar retail experience vs. online shopping/e-commerce experience

During the design process for APC (see Figure 6-13 and 6-14), the author participated in two MIT Media Lab member company workshops led by Andrew Lippman; the workshops revolved around the topic "The Future of Retail" and addressed several aspects of possible solutions to the problem domain described above including:

- **Making technology easier for the retailer** – Introduction of new inexpensive in-store interfaces that can easily integrate to the backend store management systems.
- **Helping customer make their decisions** through connectivity (be it their home or in the store) – Support the customer in the decision-making process during research at home and while browsing at the store.
- **Connecting with others** – Making a physical shopping experience social, fun and connected.
- **Engaging the customer** in a different way – Steering away from standard, static counter displays to dynamic digital solutions.
- **Customer telling** (product and customer insights for sales associates) – Facilitating two-way communication between a customer and sales associate; allowing customers to reach the most relevant sales person at any given time.

The workshops, and the ongoing work with Intel and Best Buy retail experts, helped us gain insights into the key challenges facing retailers today. In general, we found that retailers continue to struggle to stay relevant, and must continue and adopt new technologies to remain competitive.

(a) (b) (c)

Figure 6-15 Examples of new technologies for retail: (a) In-store applications from target, together with shopkick.com who provide a personal deal stream to the user's smart phone. (b) Twelpforce by Best Buy, an online technical help question & answer service based on Twitter. (c) In-store media kiosk for Olay at Wal-Mart.

We summarized their insights in the list below, which served as inputs to our design process:

- **Empower consumers:** The need to design an in-store customer shopping experience that is intuitive and self-driven.
- **Promote trust:** address the phenomenon of "marketization of information." Customers should be able to trust the retailer with information, so that they do not need to search online or verify the quality of the deal offered. One of the solutions is to fully democratize customer access to information on products in the store.
- **The mobile problem:** smartphones provide customers with the means to independently search, compare, and even buy while browsing products in a store. The result? Retailers lose business.
- **Costs of labor:** sales associates are the key differentiator for retailers, but they also represent a huge cost item. Customer experience, is in fact, a highly varied experience based on the quality of the sales associate. In addition, in stores with thousands of products, associates are not real experts. Thus, there is a need to empower associates and to streamline their work process.
- **Solutions vs. products:** In the stiff competition retailers face with online shopping, selling single products is not enough.

> For retailers to become profitable they need to sell multi-product packages. These are called "solutions." For example —a camera with accessories such as a memory card, a digital photo-frame *and* a subscription to an online service is considered a solution.

6.2.2 Field Study

In addition to the workshops, we conducted a field observation study of Best Buy locations. We scouted three locations: Downtown Boston, Cambridge, and Minneapolis. In our study, we tried to better understand the current physical setup of product display counters, specifically how products and product labeling is carried out. We also documented and observed user interaction with actual sales associates and interviewed them to better understand their work environments and how they engage customers (see Figure 6-16).

The information we collected informed our design process directly, and eventually led us to decide to focus the Augmented Product Counter around digital cameras.

Augmented Shopping Experience Concept

To address the needs discovered, we designed a LuminAR based "Augmented Product Counter." Conceptually, any standard product counter can be transformed into an interactive surface, enabling shoppers to get detailed information and conduct research while they play with the real products. Users can also access the web to read unbiased reviews, compare pricing, learn about product features and talk to an expert who may be located remotely to get additional advice.

experience that combines live product interactions of physical environments and vast amount of information available on web in an engaging and interactive manner.

By engaging shoppers in an intuitive, fun, and efficient shopping experience and helping them make informed purchasing decision, this solution can potentially enable retailers to differentiate themselves, resulting in repeat shopper visits and improved profitability.

In the sections that follow we delve into the design process, use case scenarios and offer implementation details of the Augmented Product Counter.

Figure 6-16 Field study of consumer electronics retail spaces was conducted in three different Best Buy locations, and included observation of consumer behavior and interviewing store sales associates.

The Augmented Product Counter delivers a novel in-store shopping

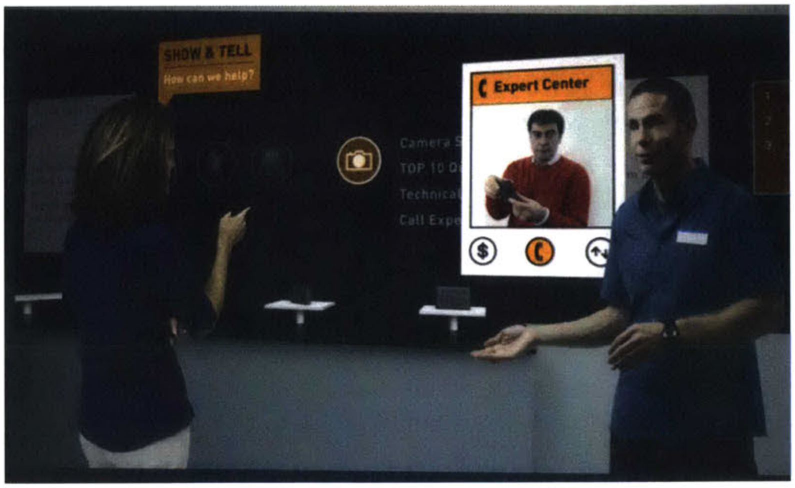


Figure 6-17 Intel's Responsive Store concept – virtual expert (source: Intel)

Design

The Augmented Product Counter was designed originally as part of the Intel Connected Store booth at the National Retail Federation Conference 2011 in NYC. We provide more details of this live demonstration in Chapter Six.

Our design was inspired by the analogy of two intersecting elements; one virtual and the other physical. The result provided the design language for the entire counter.

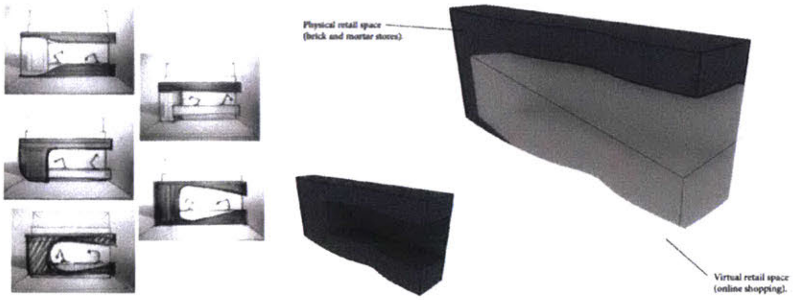


Figure 6-18 Concept design for the Augmented Product Counter (sketches by Jason Robinson and Yoav Reches)

The design concept (see Figure 6-18) had strong horizontal proportions, maximizing horizontal space for projection surface purposes. To come up with the required setup parameters such as dimensions, location of the LuminAR Spotlight, ambient light management, we created a mockup space and conducted several projection mock-up tests.

The countertop was divided into three zones:

- Product display: where all physical products are displayed;
- Interaction: where the user can interact with a specific product; and
- Spotlight: a vertical projection space used for the virtual expert session.

Use cases for Augmented Retail

In this section, we describe the actual use cases supported by the APC, their design, and implementation. The general design requirement called for users to be able to interact with the products on display and receive just-in-time information based on their pre-existing profile, market research or store pushed data. Users are able to use their mobile phone or simply login to the system. We employed gestural and multi-touch interfaces, as well as augmented reality techniques to support various use cases. In the sections that follow, we discuss the entire APC user interaction flow.

Interaction Entry Point

The APC interaction entry point involves being able to identify a customer and access his personal shopping preferences information. Such information may include standard personal information but also product wish lists. To facilitate a "login" to the system a user could simply use his mobile device. In our implementation, we used a Samsung Galaxy I Android handset. We developed a mobile shopping application that provides the user with a unique fiducial marker identifier on their phone. To start interacting with the APC, the user simply scans this marker under the system; the system then connects to the shopping profile and retrieves relevant wish list information that is displayed directly on the products.

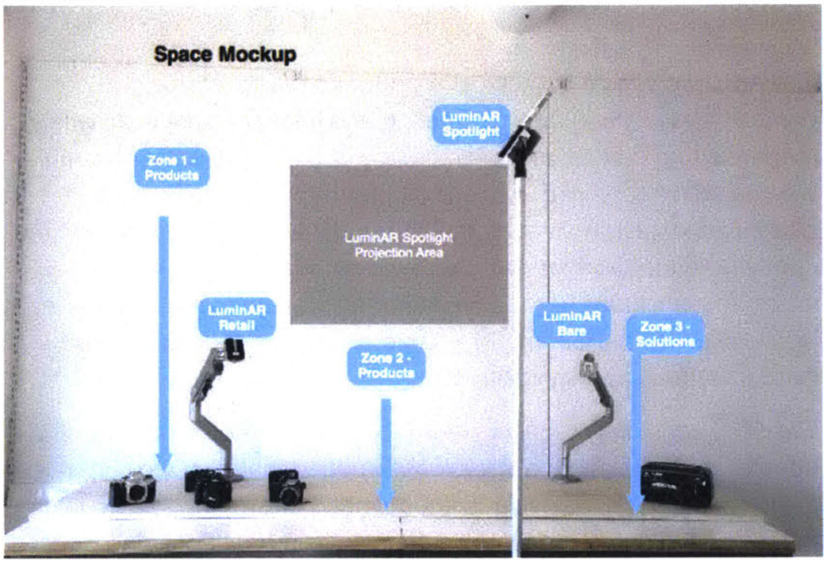


Figure 6-19 Augmented Product Counter space mockup design with different interaction zones.

Figure 6-20 Final APC GUI look and feel

Contextual Product Browsing

When consumers are searching for a new product, they naturally flag products they are interested in, following suggestions from retailers or other customers. They also consult with their direct social network for family and friend advice. Before a purchasing decision is made, consumers spend time reviewing product specifications, comparing it to competing product alternatives, and reviewing other factors like warranty and shipping costs.

(a) (c) (b)

Figure 6-21 Contextual product browsing with Just-in-Time-and-Place information: (a) Design from Intel's responsive store concept. (b) When a user picks a product from the APC, (c) contextual information is displayed directly underneath the product.

The APC provides a contextual product-browsing interface. By contextual, we mean that information and interfaces are displayed in response to a certain user interaction with the product. For example, when a customer picks up a product from the counter's product-display zone, just-in-time information is displayed directly "under" the spot where the product was displayed on the counter. Such information includes key feature information. For example, when the user picked up the Canon G11 camera, the face recognition and optical features are highlighted.

The APC constantly monitors which products are currently not on the counter, thus making the implicit assumption that they are currently in the customer's hand. Using a timeout mechanism, the system infers that since the user is spending a substantial amount of time inspecting product, he would be interested to learn more about this specific product. In such cases, the LuminAR Retail arm servos from the product display zone to the product interaction zone. This is an example of using the kinetic behavior to create a more engaging experience and encourage the user to interact with the product.

Product Interaction Zone

The product interaction zone provides a simple cue for the user to place a product of his choice in a projected target zone.

Once a product is placed in the target zone, the interaction-zone becomes alive with information. The interface updates with real-time information such as: price comparison, product rankings, special sales and video manuals.

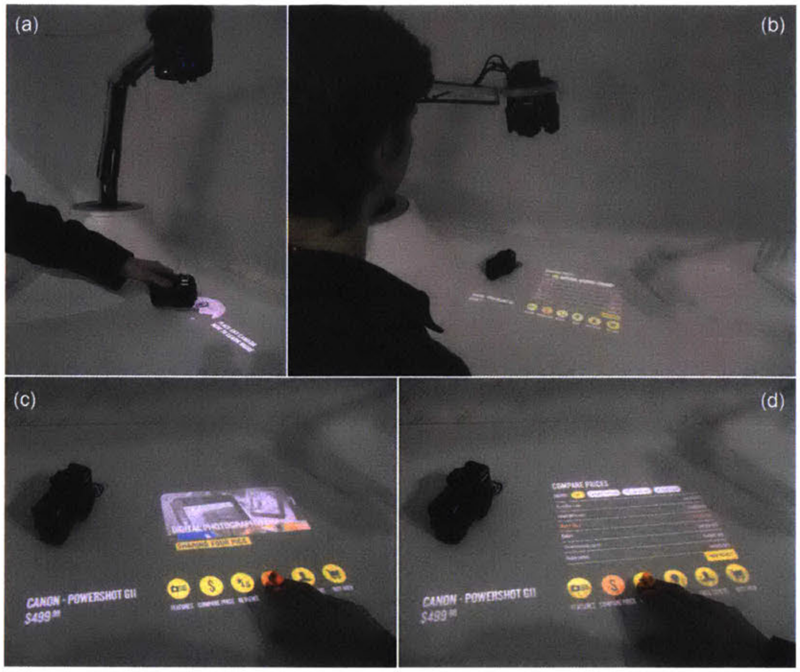
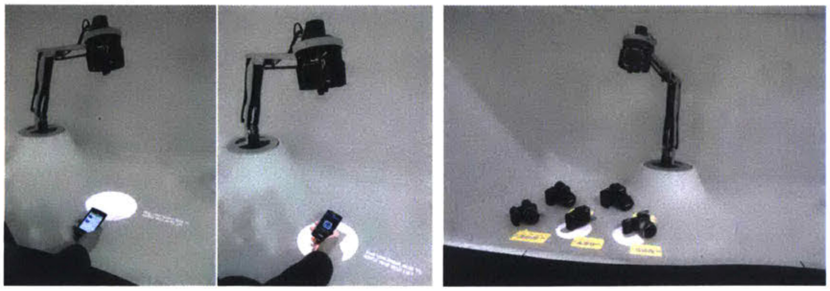


Figure 6-22 APC Product Interaction Zone: (a) The user places the camera in the product interaction zone (b) An online shopping experience is projected around the camera (c) control product features information video (d) In-store price comparison.

Remote Expert

Earlier in this section, we mentioned the labor problem retailers face. The APC tries to address this problem by introducing a remote expert function. The basic idea was to support both the customers and the staff of the retailers, by giving them access to an associate that can be remotely located.



(a) (b) (c)

Figure 6-23 (a) When a user approaches the APC, he can use a mobile application to browse his profile and wish list (b) the user can 'login in' to the APC by scanning a fiducial marker from his phone, that serves as a personal identifier (c) once the login is complete the APC reflects the users wish list by augmenting the products on the counter.

This functionality is used the LuminAR Spotlight form factor to enable a two-way video teleconference between a customer and a virtual associate.

The remote expert can provide product usage examples and manuals, augmenting the interaction-zone surface visible to the customer and also answers any questions that come up. The expert could also potentially see the customer using the tilt webcam in the bulb, although in the current APC implementation, we did not fully implement this feature.

In the example scenario, we created, we focused on customer side interfaces, developing capabilities to push product manual pages as well as additional information about complementary products for the digital camera, such as printers and digital photo frames. This remote augmented reality scenario, can potentially help retailers create actionable sales, but also effectively optimize labor resources.

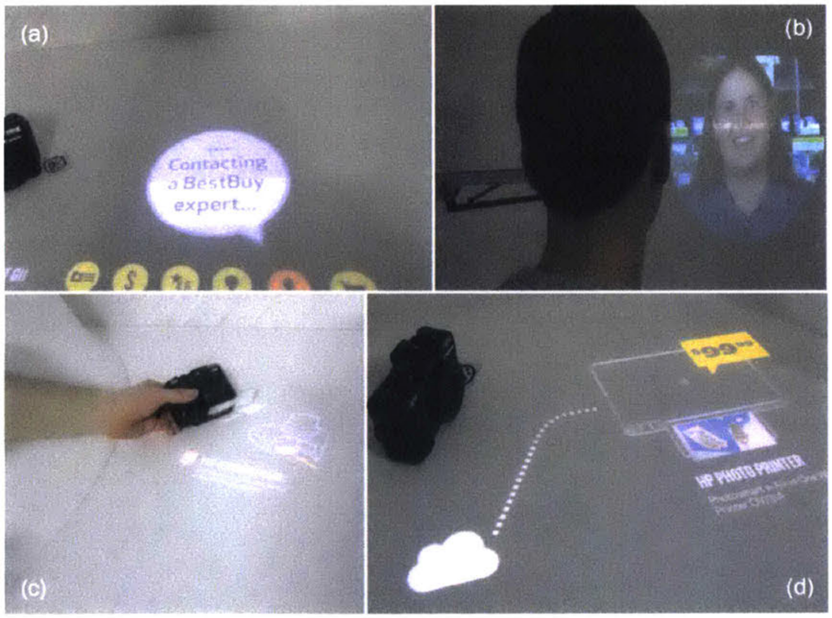


Figure 6-24 (a) A user can start a teleconference with a remote virtual expert by pressing a button (b) the expert can answer product questions but can also recommend additional products (c) the expert can re-augment the counter with information relevant to answer the customer's question, in this example the expert is displaying a manual page (d) The expert can also push additional information about complementing products to the camera the customer is enquiring about.

6.2.3 Conclusion

In this section, we presented the Augmented Product Counter concept. We explored how LuminAR can be utilized to revitalize the traditional retail experience, providing an enhanced shopping experience for consumers in physical stores.

The demo scenarios we explored show how interactive and persuasive interfaces can be used to engage customers. The APC also proposed a design for fluid transitions between online shopping to actual physical product browsing in a store. For retailers, the APC provides insight for the application of augmented reality as a valid in-store technology. The APC builds on today's familiar web and mobile-enabled consumer behavior. Such intuitive and connected interfaces could contribute to increased engagement of customers in the store. Allowing users to freely search and explore products will potentially increase the trust between a customer and a retailer, potentially directly contributing to more actionable sales, cross-sales, and opportunities to up-sell products. It can also help support the retailers' workforce, enabling staff to answer questions and sell like experts while providing real-time metrics.

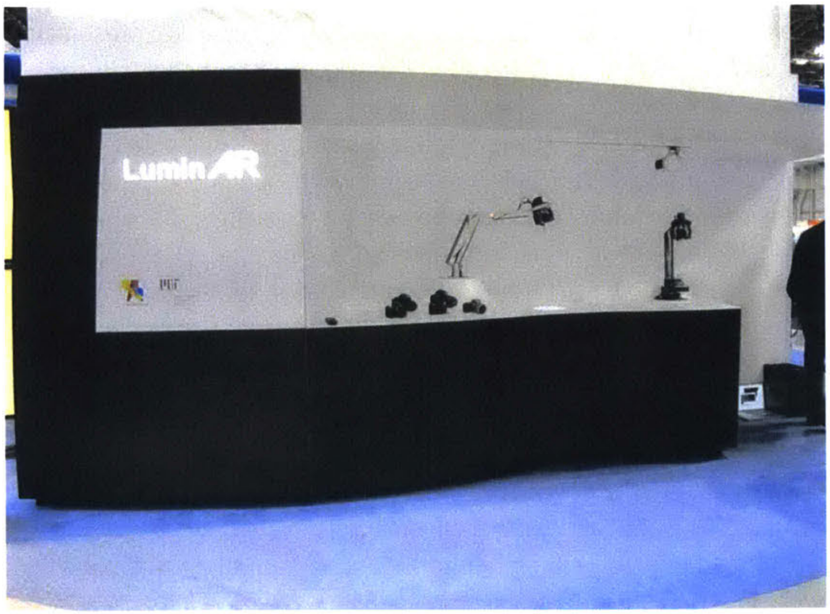


Figure 6-25 Augmented Product Counter at NRF 2011, NYC

6.3 User-Experience Feedback and Qualitative Findings

This section summarizes some of the findings that resulted from testing out the various LuminAR systems in described in this chapter. Our finding informed the user studies design in the domain of manufacturing and education that we describe in chapters seven and eight respectively.

6.3.1 UIST Conference 2010

A demonstration **[44]** of LuminAR was given to approximately 200 participants of the UIST 2010 conference (October 3-6, 2010). About 50 of them briefly had a chance to experience the interface, and it is safe to say that most of the participants were HCI practitioners.

We presented two LuminAR systems; the first demo was of LuminAR Aluminum model running a retail scenario experience prototype. The scenario included downloading product information from Best Buy APIs and providing a projected online store experience. Interaction was supported using fiducial markers as a quick means to prototype interaction. The second demo was a hardware demo that presented our latest design of the LuminAR Tipsy model. This model had proportions and design aesthetics closer to an Anglepoise lamp. It also had the first fully implemented bulb prototype that could rotate. The Tipsy model was mainly used to show the kinematic capabilities of the new hardware, including the new motor system that was implemented as well. This demo also served as a good test case for relocating the LuminAR setup from one location to another. The lamps were easily put in a box and set up at the venue. This was also the first time we tested our auto-calibration mechanism, allowing us to quickly set up the system in a new location, even though the ambient conditions were very different.

Reactions to the concept and implementation of the system were very positive. The idea of a small-scale steerable system seemed compelling to most of the people who interacted. Many mentioned the value in a compact projector-camera setup. We received critical feedback on our interaction techniques, as our demo included basic fiducial marker-based interaction. In later versions, we upgraded our vision server and implemented a hand detection mechanism based on the feedback we got at this conference.

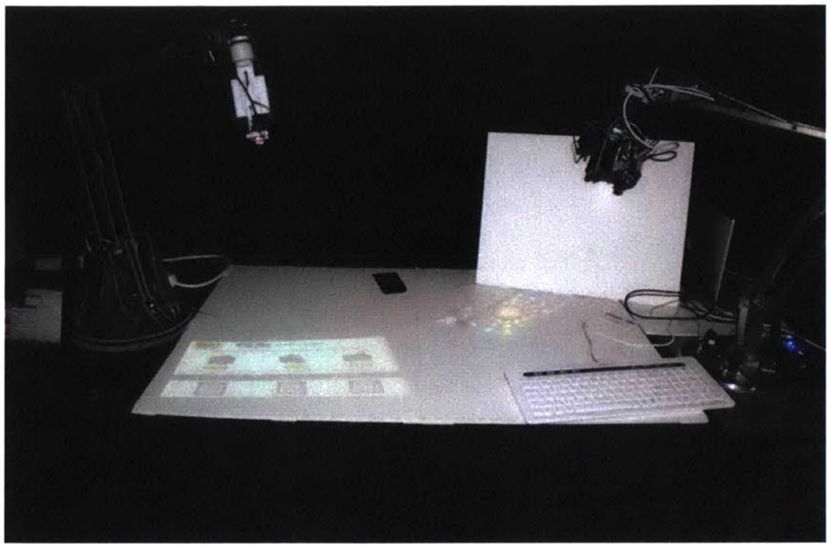


Figure 6-26 LuminAR Aluminum and LuminAR Tipsy prototypes at UIST 2010, NYC

6.3.2 Light Expo at the MIT Museum 2010

LuminAR was also featured in the opening celebration of the MIT Museum's Light Expo and the Luminous Window 2011 exhibition (December 10, 2010). The crowd included about 50 children in various age groups who were very interested in playing around with the new interface. For this demo, an application that responds to physical objects was created. Its main feature was to browse the web, based on association with physical objects. For example, a Coke can would bring up the Coca-Cola website. A mobile phone was integrated as a browsing interface, where flicking through images of fiducial markers would change the website displayed.

Figure 6-27 LuminAR @ the MIT Museum Light Expo 2011 (photo credit: Monica Brandt, Mark Ostow Photography, source: MIT Museum)

The children who experimented with the system immediately understood the basic touch-based interaction and how they can use it to change web pages using the mobile phone. They all asked the same question: "Do you have games we can play?" Gaming indeed seems like a good potential domain for future LuminAR applications.

6.3.3 NRF Conference 2011

As a result of a research collaboration with MIT Media Lab sponsors Intel and Best Buy, LuminAR was showcased prominently in Intel's booth **[174]** at the National Retail Foundation (NRF) 100th Annual Convention and EXPO in New York City (January 10-11, 2011). The conference attracted over 22,000 retail professionals, and The EXPO Hall was an enormous 150,000 sq. ft. with more than 500 vendors.

The LuminAR Retail demonstration included a rich augmented retail scenario, including physical product augmentation, integration of e-commerce features and the LuminAR Spotlight remote expert feature.

The response from the professional retail crowd was very positive. It seems that the LuminAR experience may answer real concerns and challenges the brick and mortar retail experience suffers from, namely the loss of customer base to online shopping. By bridging the physical and digital experiences in a retail environment, retailers can have the best of both worlds.

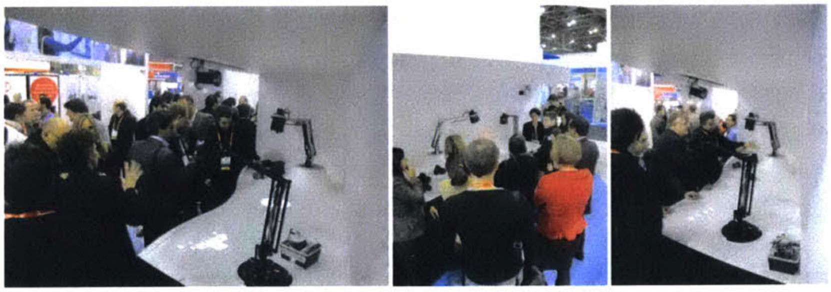


Figure 6-28 Several LuminAR Retail and LuminAR Blackjack demos took place at NRF 2011. LuminAR was featured in Intel's booth

The retail professionals also voiced concerns they had adopting such technologies. Some were skeptical regarding the need for actuation in a retail setting, arguing that much of the interaction we created could be done without kinematics and that a moving system may confuse the customer in the store. Other issues included the ease of integration and fidelity and resolution of the display with respect to the actual projected display size. All of these issues are indeed important and should be considered if and when LuminAR technologies migrate from the lab to the real world.

We collected the feedback from individuals who had a chance to interact with the system, and tried to solve immediate issues. Many of the suggestions and feedback we received in the various demonstrations were implemented in subsequent project iterations.

For example, we include below anonymous responses from our informal user study participants:

> ***"I think it fills an interesting need, and when integrated with physical displays can help create a more immersive shopping experience..."***
>
> ***—SVP & General Manager New Business Customer Solution Group, August 9, 2010.***

> ***"It is very good, interesting, adaptable to the environment. I can imagine it placing in store. One of the thoughts I have is for operations, with doctors, health related function such as perspiration medication distribution. There is no need for tablet (that we use now), we can just project on a wall or other natural places. It should have a quicker response time and possibly a higher-resolution, though for some case this is just good enough."***
>
> ***—Marketing Director, February 2, 2011.***

> ***"The augmented shopping scenario shows a very creative interface applied to shopping. Coordinating the projector motion, camera recognition, and content selection was a big job, I'm sure, but it came together really nicely to show the power of the system. I can see this LuminAR approach also having important application in education."***
>
> ***—Director, Advanced Technology, March 2, 2011.***

> ***"For me LuminAR was really fun and engaging. It was slightly confusing sometimes, mostly due to some of the inconsistencies/bugs in the tech... like when things didn't read location of taps properly, and not at all a problem with the design. Even with the few examples we saw (for example taking picture of the desktop) there was this quite *magical* feeling that the entire desktop was your***

digital oyster (that any flat surface underneath the cone of the projector could become anything else) and it was wonderfully tactile in that way. I wonder what would happen if we got to incorporate tangible tokens into the interaction? It was interesting that the actual projection moved around due to the lamp arm. At times this was confusing ("ahh! where is it going???!") and other times it is ridiculously engaging, since the interface is essentially a spotlight that hypnotizes you with moving information... so you can't help but follow it with your eyes. The idea of directing attention/focus through LuminAR sound particularly interesting..."

—Graduate Research Assistant, MIT, August 8, 2011.

"I've used touch-based interfaces, and several different forms of augmented reality interfaces, but the experiences I had with LuminAR were singularly different. Gesture became bidirectional and surprisingly informative—anthropomorphic gesturing robots feel a poor imitation of the human social fabric, but here is a new kind of partner, for play or work, freed from the demands imposed by our genetic heritage."

—Graduate Research Assistant, MIT, August 8, 2011.

"At first looks, LuminAR made me immediately think about places I personally have my own LuminAR installed and physical interactions that are now becoming possible. (For example I thought about my bedside reading lamp that can turn a blank page of paper to my e-book without having the risk of breaking it when I was falling asleep, same as for the kitchen space and my work environment). What I liked about LuminAR is that it is generic enough to support many communication needs and present information in a very non-generic and non-conventional way tailored to different use cases in the real tangible space. When experiencing the retail space demo with a virtual sales representative I started to think that LuminAR can enrich our current internet based communication with a parallel (optional) layer of communication that the website adds related to me personally and the context of the page. I was a bit frustrated with the fact that the tactile feedback is still missing from this concept

but it is understandable as these limitations may be inherent to many augmented reality concepts."

—CTO, Advanced Technology Lab Director, August 3, 2011.

"LuminAR is at core a smartbulb with built-in display plus camera which seamlessly projects a fluid applications experience onto ordinary surfaces—including tabletops, walls, floors and others—thus turning these otherwise static spaces into as rich an application ecology as you find on your smartphone, and perhaps more so. Since LuminAR has integrated camera, display, connectivity, and computation all bundled together into the form-factor of an ordinary Edison-screw-socket lightbulb, the users experience immersive interactivity in the lit-up area. Multiple LuminAR's overlapping the same surface area can auto-stitch themselves together to create supersized displays. And with the addition of robot-actuation in the lamp fixtures, an entire range of auto-orientation is an API away for the authors of smartbulb apps."

—Lecturer, MIT, August 2, 2011.

"What struck me about LuminAR was that even without the robotics, it enabled computing to be installed in places not usually suitable—up out of the way in dusty, or hygienic environments where a laptop or touchscreen unit would be inappropriate."

—Lecturer, Monash University, August 2, 2011.

"The device you created would have significant impact to the domain of augmented reality by providing a device that is easily installed and could be easily used in the home market. The LuminAR could be deployed with the movable base or installed in a standard six-inch ceiling recessed lighting can above a desk, kitchen work area, bedroom, or anywhere providing augmented support for the user. The kitchen is a natural for news and recipes, and in the living room, office bedroom, televised information, videos, anything the user may want to view. The office would be another rich area for providing secondary display information like family pictures, news information, email prompts, scheduling, wayfinding etc. You have a natural broad use product, and as the

> ***smaller laser projectors improve with time, the product will both evolve and improve."***
>
> ***—Research Manager, August 2, 2011.***

> ***"I love the perspective-taking angle. LuminAR gives me the impression that the computer is changing perspective on how it "sees" me and responds to what I do. I think that gives me the impression that the machine is an empathic being. Maybe it is because I think of it as the little lamp that Pixar uses in their opening titles....Maybe it's the motion. Kinetic things remind us of animals. Awesome industrial design work!! I find laser projectors a bit hard on the eyes. I would love a LuminAR assistant on my desk!"***
>
> ***—Researcher, MIT, August 2, 2011.***

6.3.4 Hardware/Mechanical Improvements

Based on the aggregate feedback received, as well as our own experience with the system, we have collected and outlined the key improvements. We detail such improvements below, categorizing the suggestions into the areas of software, hardware/mechanical and interaction techniques.

- Utilize better joint position sensing for LuminAR Arm. Current sensing relies on potentiometers that suffer from known accuracy, dead-zone and mechanical coupling issues.
- Investigate and implement a low-cost servo motor solution. Current off-the-shelf servos are excellent as a prototyping tool, but are too expensive as a real-world solution.
- Utilize projection hardware dynamic view angle and auto-calibration capabilities. This should reduce computational load from the software process responsible for geometric scene registration running on the main LuminAR processor.
- Add sensing for ambient light conditions, as this will allow LuminAR to automatically compensate for changes in the ambient lighting conditions, contributing to the ongoing real-time calibration of the vision system, and resulting in improved interaction.

- Add capacitive sensing to the arm and bulb's, this will enable natural interaction by detecting when a user is actually touching a LuminAR device. This could also serve as an important safety mechanism, and enable hand positioning of the LuminAR bulb.

6.3.5 Software Improvements

- Improve computer vision algorithms; implementing a more robust detection scheme for hands and fingertip detection.
- Implement full support for two-hand multi-touch.
- Develop high-level Javascript APIs for kinematics control; for example, the API should support commands to define LuminAR device position for a specific user input or an application context.
- Develop a high level Javascript API for abstracting the computer vision coding requirements for interaction. Eliminate the need for an application developer to handle vision-related coding in application code level.

6.3.6 Interaction Techniques Improvements

- Improve overall interaction pipeline response time. This improvement may include processor hardware upgrades, but will very likely involve refactoring the user interface events system.
- Add gestures that seamlessly position the projected display on a desired surface; Support smooth transitions from a tabletop to a wall or a ceiling projection.
- Add user interface feedback mechanisms for poor input conditions. This will improve the overall interaction when the system becomes less responsive.
- Improve integration with other devices such as laptops, tablets and smart phones.
- Implement animatronic feedback for user purposes. Create gestures and postures that can be integrated to application flow and provide cues to the user.

6.4 Conclusion

In this chapter, I have presented the early applications we developed for the LuminAR system. In summary, we were able to develop a wide range of applications using the LuXor and Lens software framework, and deploy them on LuminAR hardware to various real-world environments.

We did so relying on developers that were able to design and implement complex projected AR user experiences, with practically no prior knowledge of AR or computer vision techniques. Many of the developers simply relied on mobile or web development experience.

Our focus was to test the system in situ, and deploy out-of-lab installations. This in-the-wild approach provided us with abundant feedback I summarized earlier in this chapter.

Our projects in the Work and Productivity domain, provided great insights into the requirements of an AR developer-facing software framework. Our work in the Retail domain gave us a blueprint for future real-world deployments, that could be generalized as a methodology to other domains as well. We describe this projected AR design case study at the end of Chapter Five.

Finally, the iterative nature of the work helped to improve our software and hardware integration, and prepare for the final large-scale user studies. In the following chapters, I will introduce the case studies in the manufacturing and education domains.

7 Augmented Manufacturing Case Study

This chapter describes the design, deployment, and evaluation of the Manufacturing Augmented Reality System (MARS). MARS used the new LuminAR hardware design we created to support a higher resolution and brightness version that can fit in industrial environments. It maintained the same approach, a compact and self-contained projected augmented reality interface coupled with a database backend server.

MARS was designed specifically to support Lean Manufacturing Standard Work enforcement requirements and enable data collection and analysis. Previous research widely explored the effectiveness of Spatial Augmented Reality (SAR) interfaces (we introduced the concept of SAR in the Chapter Two) in procedural assembly tasks and to provide Computer Aided Instruction (CAI) to assembly workers. However, the domain of projected SAR interfaces that support lean manufacturing applications are relatively unexplored. There are few empirical results that stem from actual deployments of SAR interfaces in real-world, live production-line environments.

Our prototype augments workpieces and provides task instructions. It provides a direct user interface to the assembly worker while logging all operations. Work order information is pulled from a database server that is integrated into the factory manufacturing planning system. The results of a production trial show that users exhibited a short learning curve, adjusting well to the AR interface without significant alteration to their typical operation or impacting their work performance. It also proved an efficient data gathering tool, collecting worker performance metrics unattainable by current factory methods. Qualitative feedback indicates that assembly workers and other factory stakeholders overwhelmingly acknowledged the usability benefits of the system as a highly reliable means to enforce standard work procedures and measure in high granularity actual work performance. We also found that the system can serve as an effective on-the-job training tool.

7.1 Introduction

The origins of the term Augmented Reality (AR) lie in manufacturing. The actual term was coined in 1992 by Caudell and Mizell **[175]**. Their research was motivated by trying to improve complex aircraft manufacturing processes using Head Mounted Displays.

Our work centered on the design, deployment, and evaluation of MARS, an application of the LuminAR system designed specifically to support assembly tasks in a lean manufacturing environment.

7.2 Lean Manufacturing

Assembly is a standard manufacturing process, ubiquitous across many industries. It requires the manipulation and joining of parts to form a whole product [176]. Research on assembly work in the 1960s as reviewed by Seymour [177], clearly showed learning periods are essential for any in efficient assembly work.

The term lean manufacturing (LM) defines a production practice that puts value creation for the customer at the center and scrutinizes waste and expenditure. The LM philosophy proliferated from the Toyota Production System and was studied extensively by Womack et al. [178], [179] . In a recent article, Capozzi and Sacco describe the immense potential of AR to have a meaningful impact on lean manufacturing. They also propose a framework for an Augmented Factory [180].

LM principles require continuous manufacturing process improvement. However, measuring and analyzing the manufacturing operations is complex, manual and expensive. The MARS system is specifically designed to alleviate such challenges by allowing ongoing data collection and analysis. With regard to Capozzi and Sacco's vision, it is possible that our MARS system is an initial essential link in the implementation of what they describe as an 'Augmented Factory'. At the time of writing this dissertation, the MARS system is one of the few standalone AR interfaces that actually works out of the box. Additionally, it is the only system allowing a developer to use simple web-based tools to create complete AR experiences that can support lean manufacturing environments.

7.3 Augmenting a Real-World Production Line

We specifically focused on the c:scape office furniture production line [181]. The IMO researchers had identified LM challenges that could be well served by employing an AR interface in a production line environment. We held a day-long design workshop at the MIT Media

Lab, and shortly after, MARS developers and researchers made two site visits to the actual manufacturing facilities. We performed several observations that informed the research questions we will present in the next section.

The MARS system was designed to promote the adoption of lean manufacturing principles within a production operation. The general hypothesis of this dissertation is that the use of projected AR interfaces in assembly tasks can increase worker productivity, reduce errors and increase worker satisfaction. In addition, our system can also be used to collect detailed metrics that are critical to facilitate the continuous process improvement LM seeks.

We devised a study to evaluate this general hypothesis and pilot the system in a real-world production environment. From the data gathered in this study, we measured the following quantitative metrics: (1) reduction of manufacturing errors; (2) increased operator productivity (in terms of build time and/or completed units per rotation); (3) increased operator knowledge and proficiency in the standard work protocols; and (4) increased build operation consistency and adherence to standard work. Along with the quantitative analysis we performed, we carried out extensive interviews and feedback sessions with actual production line operators, as well as with other related stakeholders within the KWW manufacturing organization.

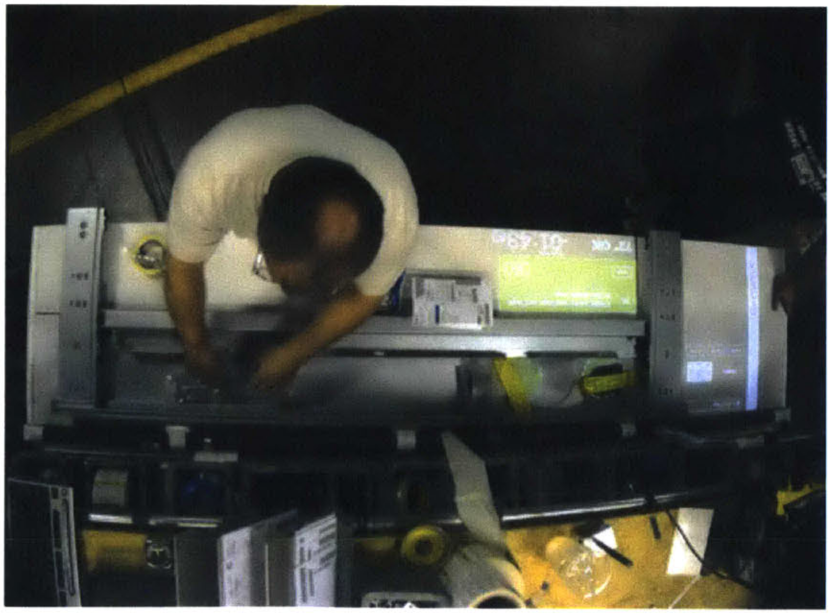


Figure 7-1 MARS enabled AR assembly process in Steelcase Kentwood West Factory

The Augmented Manufacturing work presented in this dissertation is centered on the design, deployment, and evaluation of the Manufacturing Augmented Reality System (MARS). MARS was developed in close collaboration with Steelcase Innovation Management Office (IMO) and the manufacturing organization at Steelcase's Kentwood West Factory (KWW). We specifically focused on the c:scape design system office furniture production line, and the station we decided to augment is shown in Figures 7-1 and 7-2.

The MARS used the LuminAR projector, a compact and self-contained projected augmented reality interface (See Chapter Three). The device was coupled to a database backend server. The MARS system was designed to promote the adoption of lean manufacturing principles within a production operation. The general hypothesis in this dissertation is that the use of projected AR interfaces in assembly tasks can increase worker productivity, reduce errors, and increase worker satisfaction. In addition, our system can also be used to collect detailed metrics that are critical to facilitate the continual process improvement that lean manufacturing seeks. We devised a study to evaluate this dissertation's goal and to pilot the system in a real-world production environment. From the data gathered in this study, we measured the following quantitative metrics:

1. reduction of manufacturing errors;
2. increased operator productivity (in terms of build time and/or completed units per rotation);
3. increased operator knowledge and proficiency in the standard work protocols; and
4. increased build operation consistency and adherence to standard work.

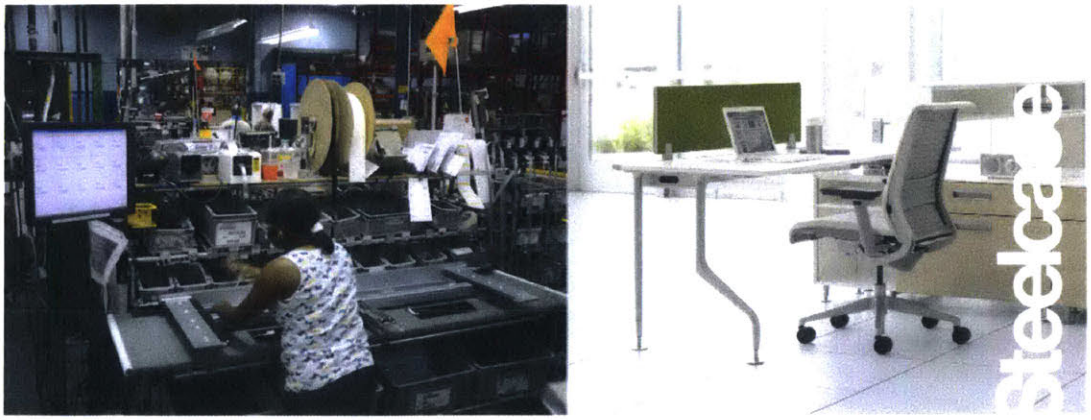


Figure 7-2 The original state of the c:scape station at Steelcase Kentwood West factory

Along with the quantitative analysis we performed, we conducted extensive interviews and feedback sessions with production line operators and with other related stakeholders in the KWW manufacturing organization.

Henderson's and Feiner's research widely explored the effectiveness of Head Worn Display-based AR interfaces in procedural assembly tasks and to provide Computer-Aided Instruction (CAI) to assembly workers, Their results demonstrated the usefulness of AR as a tool to guide assembly tasks **[182]**. However, the domain of projected SAR interfaces that support lean manufacturing applications is relatively unexplored. There are few empirical results that stem from actual deployments of SAR interfaces in real-world, live production-line environments.

Our prototype augments workpieces and provides task instructions. It provides direct user interface for the assembly worker while logging all operations. Work order information is pulled from a database server that is integrated with the factory manufacturing planning system. The initial analysis of the results of the production trial shows that users exhibited a short learning curve, adjusting well to the AR interface without significant alteration to user operation or impacting work performance. It also proved an efficient data-gathering tool, collecting worker performance metrics unattainable, by current factory methods. Qualitative feedback indicates that assembly workers and other factory stakeholders overwhelmingly acknowledged that the usability benefits of the system were a highly reliable means to enforce standard work procedures and to measure in high granularity actual work performance. We also found that the system can serve as an effective on-the-job training tool.

Figure 7-3 Applying Spatial Augmented Reality to Facilitate In-Situ Support for Automotive Spot Welding Inspection Jianlong Zhou et al., 2011

Extension to the MARS system capabilities were developed in collaboration with the Advanced Manufacturing Initiative of Lockheed Martin Corporate Engineering Division. Lockheed Martin is an American global aerospace, defense, security, and advanced technology company. One of Lockheed Martin's flagship projects is the Lockheed Martin F-35 Lightning II. The current process for manufacturing the F-35 requires over 70,000 hours of touch labor per aircraft. This labor comprises a variety of tasks, a significant portion of which involves the installation of fasteners, brackets, boots, and other components in a particular order according to precise specifications.

Despite the many technologies successfully inserted into production, many portions of final assembly resort to legacy workflows, with mechanics relying on scattered product, procedure, safety, and requirements data to carry out their jobs. In addition, while Lockheed Martin has many procedures and checks in place to minimize human error and its impact on production, it is inevitable that manufacturing technicians will occasionally make mistakes. Developing a system that improves the workflow by utilizing modern technologies could increase mechanic efficiency while simultaneously reducing human error.

Lockheed Martin also has also successfully integrated related technologies such as optical projection and non-contact metrology into F-35 production. Our work integrates many useful technologies into a powerful and flexible system applicable to a broad range of production environments. Within Lockheed Martin Aerospace division, and specifically the F-35 program, an ideal work environment exists that could benefit from the enhanced capabilities associated with using a system such as MARS. Detailed touch labor procedures are found throughout the F-35 final assembly process, and MARS can provide a more effective way for manufacturing technicians to interact with the information required to perform the required tasks.

Because the current system utilizes many off-the-shelf technologies, the total hardware cost for the base system is less than $2,000. While there is an initial upfront cost associated with developing the projected interface, development will emphasize minimizing that cost by utilizing commonalities across production tasks and through the development of software tools that will automatically generate MARS workflows based on existing work instructions. By keeping system costs low, there will be greater opportunity for applications that can benefit from the technology. Together with Lockheed Martin, we have created a mobile version of the MARS system—mounted on a mobile pedestal. The mobile unit is designed to fit in the complex production line of the F-35, allowing the unit to be placed ad-hoc according to assembly needs.

Across its use cases, MARS is designed to provide the following benefits:

- Improve the flow of information to mechanics, which will improve efficiency and reduce potential errors. This includes providing a step-by-step walkthrough of the process and tracking task progress for digital buy-off of successful completion.

- Utilize sensing capability to do real time quality assurance and provide feedback to mechanics throughout the assembly process by detecting errors and preventing rework.
- Collect highly detailed data about workflow timing, which can be used to implement lean production processes and optimize worker assignments to boost efficiency. This data will also enable statistical analysis of the entire manufacturing process more effectively than do current methods.
- Serve as a training tool for expediting the time required for new mechanics to become proficient at a specific task
- Strengthen the digital thread by providing a mechanism to collect manufacturing information for a specific aircraft that will follow the product all the way through sustainment.

7.4 Related Work

Early AR manufacturing research focused on Head Mounted Display (HMD) or Heads Up Display (HUD) based system. This approach typically involves overlaying instructions in the direct view of the operator. A work area is registered in a 3D real-world coordinate system in relation to a workpiece. The operator's head position is also continuously tracked and correlated.

Figure 7-4 KARMA - Knowledge-based augmented reality. Steven Feiner et al., 1993 **[183]**.

This early work further developed and highlighted the potential for AR to dramatically improve manual assembly operations **[183]**. Later work from Neumann and Majoros suggested a framework for six types of overlay representations for AR information **[184]**. They explored how AR application flow can be described in the ARML markup language. ARML is an early example of an architecture that addresses the challenges of authoring an AR application.

SAR interfaces for manufacturing were also widely explored in **[185]**-**[187]** . However, the specific domain of projected SAR interfaces that support Lean Manufacturing applications is unexplored. There are very limited empirical results that stem from actual out-of-lab deployments of SAR interfaces in real-world, live production line environments. Most recently Fite-Georgel attempted to answer the question "Is there a reality in Industrial Augmented Reality?" [188]. He provides a comprehensive survey and taxonomy of recent research findings and notes that only two projects have broken out of laboratories into actual use in the field [189], [190].


Figure 7-5 Augmented Reality in the Psychomotor Phase of a Procedural Task Steven J. Henderson and Steven Feiner 2011.

AR has also been found to be an effective training tool [191]-[193]. Wiedenmaier conducted an experiment that measured the performance of novice assembly workers, comparing an AR condition versus a 3D manual print. The results suggest that the AR condition yielded a lower learning curve and cognitive work load, resulting in faster task completion times with less human error [194]. Webel et al. also provided excellent design guidelines in recent work [195]. AR technologies have also been used to augment physical tools in the manufacturing environments. For example, Echtler et al. introduced an intelligent augmented welding gun to the automotive industry which shoots studs with high precision in vehicles [189]. Early results indicated significant time improvements over the traditional stud welding process. Many additional industrial use cases are covered by the excellent surveys [196], [197].

7.5 Research Questions

We focused on the following research questions to advance the empirical findings as those relate to the real-world AR interface deployment to a working production line.

> *1. Can projected AR interfaces be employed to reduce operator errors and increase quality of worker output?*

This question relates to one of LM's core principles of just-in-time information. The assumption is that humans, no matter how proficient in a certain task, are still continuously taxed by memory recall operations. In the context of assembly tasks, and specifically tasks with a high degree of customization, yet high degree of similarity (as the tasks at the c:scape line where we installed MARS/LuminAR), the potential for errors is very high. Given our observations of actual assembly operations and based on previous work [196], [197], it is our hypothesis that AR interfaces augmenting a physical workspace could help operators recall the operations they need to perform at a given time. No less important, the introduction of sensing-based augmentation that takes into account the production step, and forces the user to adhere to standard work practices, can dramatically help with operator error reduction.

> *2. Can projected AR interfaces be used without severely impacting production throughput? Can they help eliminate unnecessary paper or context switching screen-based interactions during the assembly process?*

Our assumption is that the use of augmented interfaces eliminates unnecessary paper or screen-based interactions during the build process. Such interactions can increase overall build time as they introduce operator context switch between the build operation activities and the screen or paper instructions. In this study, we did not perform a controlled experiment to verify this assumption. Instead, we chose to focus on showing that an AR interface immediately adjacent to the work context will not introduce significant overhead to the production throughput when compared to the current screen and paper baseline available to us.

> *3. Can projected AR interfaces can be used to effectively teach standard work processes and reinforce standard work knowledge?*

One of LM's most powerful concepts is the Kaizen principle. Kaizen refers to a work philosophy that focuses on continuous improvement [198]. These questions explore how AR interfaces can support the use of standardized work procedures. This is important as standard work procedures provide a baseline for improvements.

Compared to paper or computer-based procedures, this question asserts that the use of augmented interfaces would help with the reinforcement of standard work procedures and the teaching of complex assembly tasks.

Wiedenmaier's work has demonstrated that AR based assembly instructions, synchronized with the real environment and the object of instruction, would be more effective compared to an instructional video or a paper procedure [194]. Those are typically delivered in a classroom environment, decoupled from the object of instruction.

A controlled experiment to test this assumption must involve other methods that make for a better comparison case, for example, the use of on-the-job-training procedures. We observed that many of the senior operators on the c:scape line have issues adhering to standard work—it became relevant to explore opportunities to use augmented interfaces to change work habits for a population of the workforce that is less likely to change its so called "old work habits." This hypothesis also has strong grounding in the literature; there has been extensive research into the use of AR as a training tool for manufacturing [191], [192], [199], [200].

7.6 MARS

MARS began at a workshop on augmented reality systems for manufacturing that was held with the representatives of Steelcase Inc., a workspace furniture manufacturer. This workshop was followed by a visit to several Steelcase factories. For development of MARS, we chose the manufacturing line for Steelcase's c:scape office tables. The c:scape assembly line is comprised of several stations at which different components of the table are assembled. The undercarriage assembly station was selected for augmentation after review because the workpiece remains mostly stationary, the station has a horizontal interaction surface, and it was appropriately sized for projection.

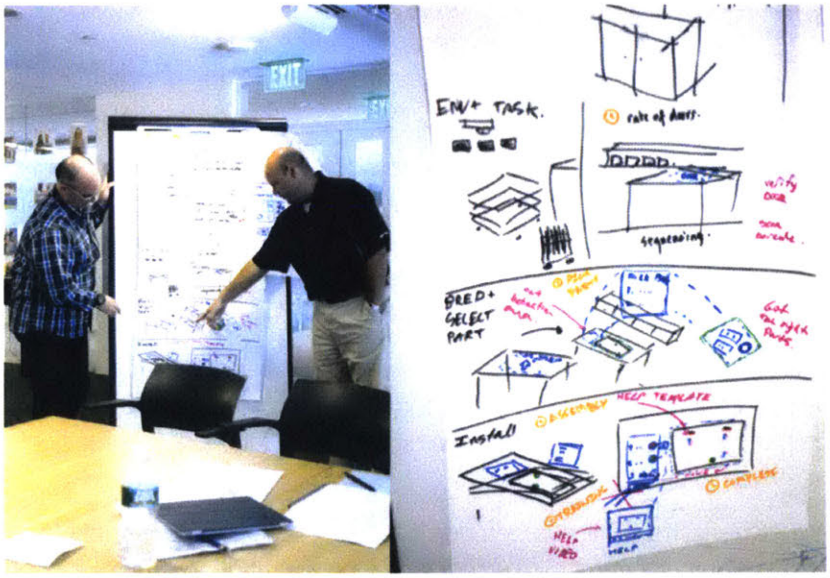


Figure 7-6 Augmented Manufacturing workshop at the MIT Media Lab

7.6.1 System Design

Our design was motivated by two key principles: First, in terms of user interaction, we wanted to introduce the AR experience with minimal overhead for the users and focus on modeling the existing behavior and assembly procedures. Second, we strove to create a simple and inexpensive hardware design that would be easy to integrate into the existing production line.

To design and test the system, we replicated in lab the factory workbench, making the transition to the factory easy. Any additions or features added to the workstation were tested using materials that would be present in the factory. Thus, when we deployed the system for the trial, it was simply a matter of duplicating the workstation.

The c:scape assembly procedures were studied and incorporated into the system. Particular steps were monitored to ensure they were completed. Information relevant to users and observers was projected in key places. A system was devised to display errors, warnings, and information in such a way that was instantly distinguishable by the user and anyone in the vicinity. Current methods of calling for help or requesting additional parts were also implemented in the system. We believe that by modeling the current work processes and involving the different stakeholders early on, users experienced a quick learning curve when introduced to the actual system. We will discuss this further in the results section.

Although much of the workstation and existing work process was maintained, enhancements were made to the system to gather information. Current production processes did not allow for time-to-completion to be measured without having someone monitoring the process with a stop watch. Our system passively measures the overall build time and has the ability to measure the time to complete each individual step in the process. To complement the physical system, we designed a backend database to allow for easy manipulation of information displayed to the user. Specifically, the backend interface allowed for factory news, the editing of standard work procedures, and various LM notifications. The database also logs valuable user-performance metrics such as build times, number of help calls, break times, and quality control images for each completed build. With the exception of the quality control images, this information was not currently logged and could only be obtained through active human-based means, versus our system's passive automatic data-collection facility.

Steelcase manufacturing stakeholders at all levels—from the plant manager to line workers—provided crucial feedback at key development milestones. At different stages of the design process, the system was demoed for employees, ensuring the direction taken was relevant to factory work procedures. This approach assured that the final design answered the specific needs of bench-top assembly operations. It also promoted a sense of ownership and acceptance of the system, which was very helpful as we proceeded to integrate the system into the actual production line.

7.6.2 System Implementation

In this section, we describe the implementation details of the MARS system.

7.6.3 MARS Hardware

MARS was implemented using the LuminAR projector hardware we described in detail in Chapter 5. Figure 7-7 shows the hardware mounted above the workbench.

Figure 7-7 MARS test station at Steelcase's Kentwood West factory.

7.6.4 Software

While there has been substantial research into Augmented Reality (AR) interfaces, it is still difficult to build applications that leverage these technologies, due to the lack of development tools. The MARS device runs the LuminAR LuXor runtime stack in combination with the Lens software stack (see chapter four).

LuXor is a software framework designed to support interactive projected AR scenarios. LuXor comprises a web-based application framework and a set of core services that provide hardware drivers, application runtime logic, event management, computer vision services, and projected GUI utilities.

The MARS front-end application was developed using the Lens software framework we described in detail in Chapter Four. Lens is LuXor's Javascript API that allows developers to use standard web development tools to build projected AR applications. This eases the development process and opens up this interaction modality to all levels of developers.

Augmentation and Vision Capabilities

At present, our system's computer vision capabilities support:

- Simple one-time calibration of projector and camera.
- Blob-based hand and fingers detection.
- Multi-touch interaction based on browser touch events.
- Basic object identification and tracking based on fiducial or image markers.
- Capturing and cropping snapshots from a camera based on the position of the web page's HTML elements.

As discussed above, Lens also provides applications with a number of pre-built interface widgets and components specifically designed around the constraints of projected AR interfaces. By providing high-level abstraction that is decoupled from the underlying augmented reality platform, Lens provides a developer-friendly environment that makes it fast and easy to develop augmented reality applications.

MARS Frontend Application

The MARS application pulls data from a production work order database and augments the assembly bench with assembly information accordingly. In some specific steps, the application augments the workpiece itself. In addition to the actual augmentation, the application has communication interfaces that allow the operator to send text messages asking for help or calling for more parts.

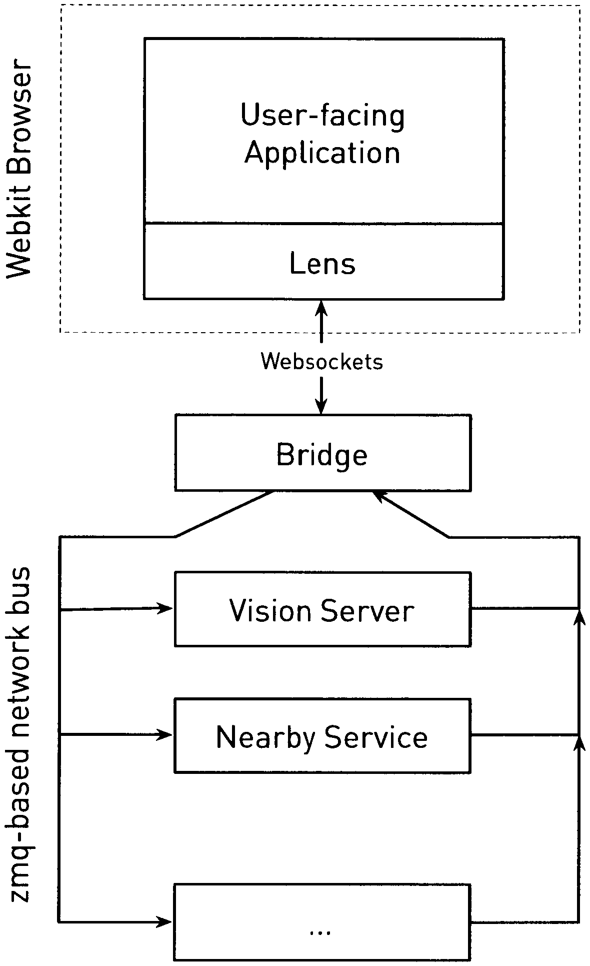


Figure 7-8 The MARS software stack is based on the LuXor and Lens frameworks

SteelData Backend Server

The MARS backend server, SteelData, is a Ruby-on-Rails web application. SteelData provides production engineers with the interface to input manufacturing data to the system, such as standard work information, operator information, factory news and more. SteelData also provides means to import and query work-order data. Most importantly, SteelData exposed a visual dashboard that summarized the vast information collected by MARS. This provides production line operators, their team leaders and the factory management with visibility to performance metrics, and work progress. It also allowed them to perform basic comparison and analysis operations on the data.

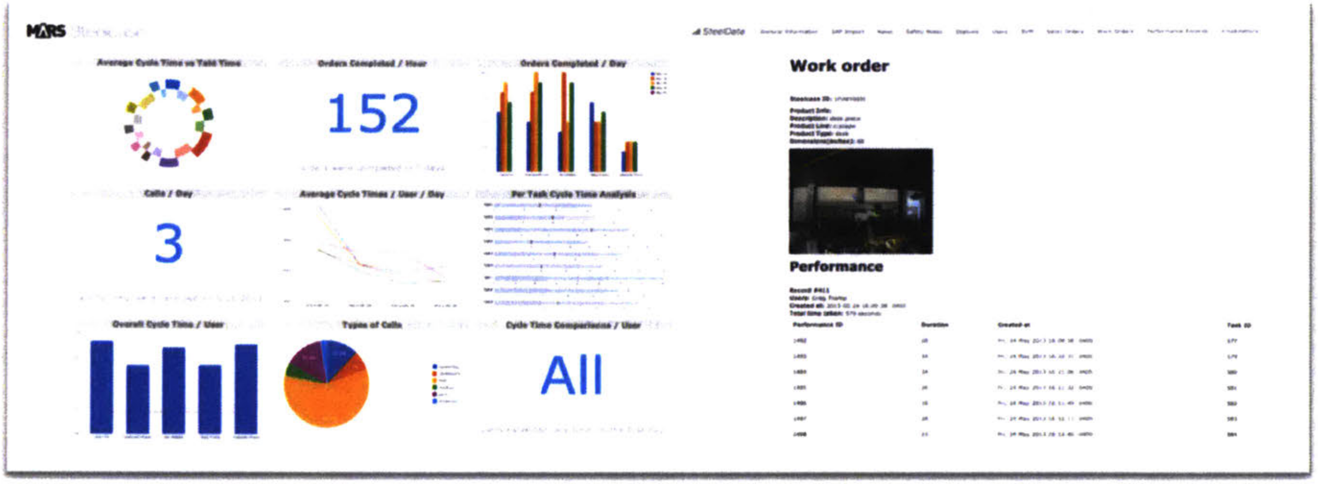


Figure 7-9 SteelData provided a web based dashboard for the MARS system

7.6.5 c:scape Desk Station

c:scape is a high-end office furniture line from Steelcase. It served as a development test bed for the MARS system. The MARS unit was installed on a custom-made rig. In addition, the original assembly station was modified in three minor ways:

- The existing factory workstation has magnetic reed switches underneath the work surface to measure the length of the table workpiece. Our system integrates with these sensors via serial communication to an existing sensor board. This had no effect on the user and was unnoticeable.
- The workbench length was slightly extended by eighteen inches to provide a larger free projection surface.
- A mounting rig was installed to support the system and camera. This addition had no effect on the the work area.

Figure 7-10 show the integrated projector camera device and its mounting to the production workstation. The device was mounted approximately 1.8 meters above the production workstation, giving a projection area of 1.00 by 0.55 meters, covering the left half of the production table and workpiece.

The system was also connected to two additional external peripherals for the manufacturing application:

1. A handheld laser barcode scanner was connected to the device to integrate with the factory's existing paper-based workflows. The existing workflow relies on a work-order sheet

of paper that serves as an identifier for a single workpiece to be assembled. The sheet is attached to the workpiece as it proceeds through the factory. At each assembly station, a barcode identifier on the work-order sheet is scanned into the factory's production management database. Worker ID badges also have unique barcodes which are scanned into the system.

2. An external IP camera captures the quality-control (QC) photograph during the last step of the assembly task. The external camera provided a higher-resolution QC image than available using the color camera onboard the device. Additionally, the external camera allowed for the QC image to be taken directly above the workpiece.

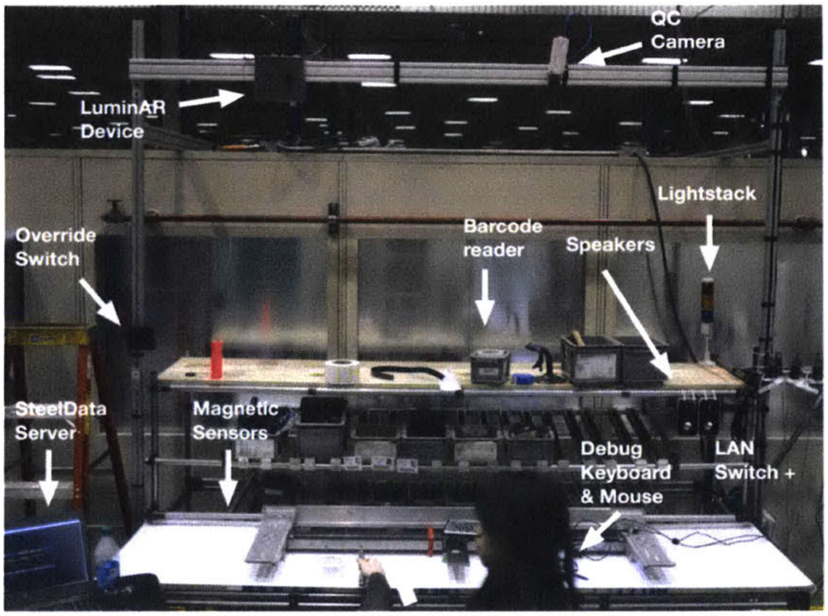


Figure 7-10 The c:scape station with a MARS device mounted to the production workstation onsite at Steelcase Kentwood West factory

7.7 User Studies - Production Line Trial

To assess if the MARS system can be used as an effective real-world AR interface for assembly operations and as training tool to teach and reinforce standard work, we designed a five-day-long production trial at the Steelcase Kentwood West factory (KWW).

The study, "Effects of Augmented Reality Interface on Performance of Specified Tasks[1]" was designed to measure the effects of using projected augmented reality interfaces to perform repetitive tasks, by measuring the change in performance of subjects conducting the same tasks with and without the presence of our system. The purpose of this study was to show how the disadvantages of prolonged repetitive tasks can be mediated with a more interactive system.

7.7.1 Study Preparation

Prior to the study, in a separate visit to the KWW factory, employees were exposed to the system and were given a demonstration and tutorial on the use in a real assembly scenario of the c:scape product. The tutorial gave potential subjects an extensive overview of the system.

Figure 7-11 MARS Demo and Tutorial at Steelcase Kentwood West Factory

Study subjects were then walked through the progressive flow of the user interface system as well as the features. Instruction was given on how to activate each feature, such as a call for help or call for parts, and how to deactivate them. Subjects were shown ways to avoid and fix error messages, as well as what actions may result in an error.

---

[1] Effects of Augmented Reality Interface on Performance of Specified Tasks (MIT COUHES #1302005532)

Finally, each subject was shown what information was gathered by the system and how this information was kept in the database. Following the tutorial, each subject had a clear introduction of the system's parts, interface, and database.

As part of the tutorial, we studied how subjects interacted with the system as they were asked to go through a build of a table—without physically constructing it—using our system. As they progressed through this task, comments and actions were observed and noted. This approach informed our production trial experiment. We found that generally operators could complete a build with little to no assistance from the moderators.

7.7.2 Experimental Setup

The production line experiment was specifically designed to study both operator and system performance. We were interested in evaluating the adjustment of operators to the use of the MARS interface over time. Another focus area was the verification of our manufacturing data collection approach.

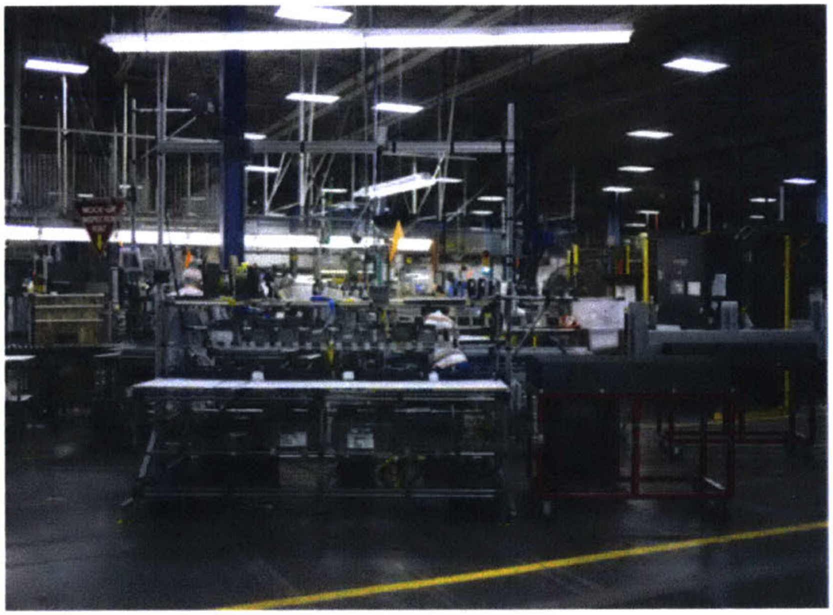


Figure 7-12 The MARS c:scape system on the actual production line floor

The purpose of our study was to (1) evaluate how well operators follow the current standard work procedures; and (2) measure and analyze the impact introduced by the system on time-to-complete assembly operation. Our assumption was that the study results would shed light on the standard work operations time allocations (Cycle

Time) in use. Cycle time is defined to be the measured time it takes to complete the set of discrete assembly operations in a defined station. Cycle time is currently obtained by time studies that are subjective and suffer from human errors, and are hence inaccurate. We also wanted to prove that the MARS system can actually be used as a data collection tool that can yield improved task time completion.

The trial group included five assembly operator test subjects. Four subjects were male (ages 35-45). The fifth subject was female (age: 22). The four male subjects had previous experience with the c:scape production line. The female subject was a temporary worker training on the production line with no prior knowledge of either the work process or the MARS system. The experiment took place on the actual production line stream. During a work shift, the five test subjects rotated as they would usually do on the c:scape line, spending approximately two hours each day performing assembly work on the experimental augmented station.

To support the real-world scenario, we integrated SteelData to Steelcase's backend manufacturing planning system. Each night, current production work-orders were loaded into the system to be used in the experiment. The c:scape station used as a host for the trial was equipped with video cameras viewing both the overall experiment and detailed interaction with the AR interface. The experiment ran with little moderation by the research team, allowing operators to interact directly with the system and carry on as they would normally perform the assembly operations. Finally, we conducted daily interviews to assess the subjective experiences with the MARS system. We also held extended post-trial interviews that extended the qualitative feedback on the use and potential impact of the system with both operators and the factory management staff.

7.7.3 Methodology

Due to the nature of running an experiment in a live production line, it was difficult to allocate a formal control group that would help solidify our findings. Cognizant of this fact, we opted to use as the control for performance in this experiment the known and well-established standard work Steelcase has been using. While it is not a formal control, we believe it serves as a reasonable baseline for our experimental results.

The c:scape undercarriage component included 20 well-defined and separate assembly steps. To establish the baseline further, prior to the actual trial assembly sessions, we presented subjects with a randomly ordered list of standard work steps and we asked them to arrange them according to the correct assembly order. This approach established their perceived personal knowledge of the standard work required for assembly. This also proved a useful method to assess standard work knowledge over the duration of the experiment.

7.8 Results: Qualitative Analysis and Discussion

System Performance

The first and most significant experimental result is that during the production trial using the MARS system, the Steelcase KWW operators were able to complete 152 c:scape undercarriage units, their planned quota for product built for the duration of the experiment. This indicates that the system at least maintained the production level required from the this specific c:scape line, and that in fact the MARS system did not slow down the planned production.

Operator Performance vs. Cycle Time

We found that operators quickly adjusted over time to the AR assembly process and that MARS did not significantly impede cycle performance.

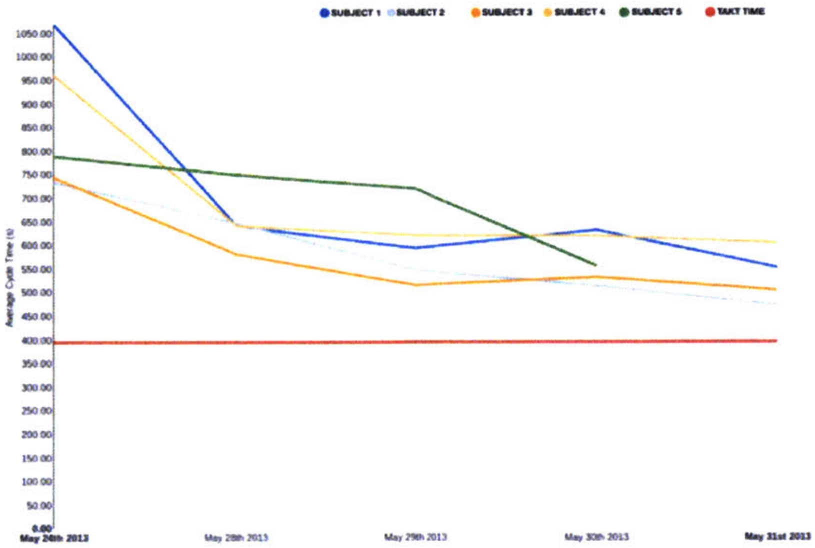


Figure 7-13 The average assembly cycle time during the production trial

Figure 7-13 shows the average build cycle time during the study, showing a clear drop in time as the study progresses. The average cycle time of the operators at the end of the fifth day of the experiment was 538 seconds compared to 853 seconds on the first day. This performance should also be compared to the established 395 seconds planned cycle time. Our findings indicate that the system introduced an overhead of an average of 143 seconds to the total assembly time.

The additional time can potentially be explained by simple observation that the system indeed requires the workers to perform additional confirmation steps. However, there is also indication in our results that the actual assigned cycle time for the different steps in the assembly process are incorrect. Figure 7-14 provides a summary of this trend via a comparison of per-task operator performance variances.

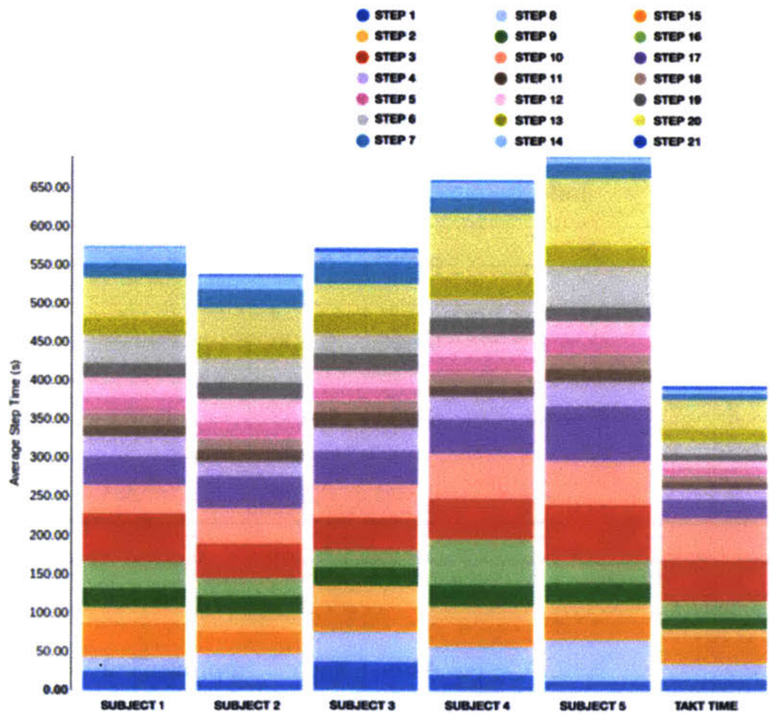


Figure 7-14 Per task operator performance variances

SteelData, the system's backend server dashboard and analytics application, emerged as a helpful tool in identifying cross-operator variations. Figure 7-14 also provides a glimpse into the step-specific variances we discovered among different operators. Such variances provide clues as to which steps have been previously assigned an inaccurate takt time value. The data also reveals clearly that different operators excel at different steps.

Training and Enforcing Standard Works

After five days working with the system, we asked the study subjects to complete a post-completion survey and order the standard work steps, similar to the pre-study survey. We deduced that the actual knowledge in standard work procedures has improved, on average, by approximately 17%. Table 7-1 summarizes the improvement in standard work; we have used the Kendall rank correlation coefficient to measure the magnitude of improvement. The analysis of these results provides an indication that the system helped with standard work training and reinforcement.

Table 7-1 Scores are normalized Kendall's-Tau statistic. Participants were asked to correctly order standard work steps; The statistic compares participant answers against the correct ordering. The score ranges between 0–1, with 1 being perfect score.

Participant	C:Scape Experience	Before	After	Improvement
1 (temp)	none	0.545	1.000	**0.455**
2	5 years	0.727	0.822	**0.095**
3	1 year	0.600	0.600	(none)
4	5 years	0.564	0.867	**0.303**
5	2 years	0.800	0.800	(none)

	Before	After	Improvement
Average	0.647	0.818	**0.171**

It is also important to address the lack of formal control in this study and the implication on the results presented below. Two key factors made it impossible to perform a formal control study alongside the experiment we preformed. First, like many of the manufacturing operations we visited, Steelcase did not keep record for previous operation performance and, secondly, it was not practically possible to run an parallel experiment at the same time we had allocated to complete the experiment. That said, we believe there is an inherent control in the production quota allocated to our work bench during the experiment. That quota represents the regular system Steelcase is using in operations. The fact that MARS did not slow down the workers and that the production quota was met provides in my opinion real-world validation of our results per a known baseline of production quota.

We can also make the following observations based on this analysis:

- Participant #1 (a temp worker) has clearly benefited from the system as a training tool.
- Participants #2 and #5, both five-year veterans of the c:scape line, did not have perfect knowledge of standard work procedures as one might expect. Both exhibited improvement in their knowledge, indicating the system also performed as a reinforcement tool.

User Experience and Feedback

The user feedback can be divided into two categories: feedback given by assembly operators and feedback given by observers.

Operators were those that participated in the trial user studies and were able to interact with the system directly. Observers were those given demonstrations in large groups. These groups were usually comprised of workers in managerial positions. Table 7-2 summarizes some of the surveys we performed to assess operator satisfaction, all the responses to the question were scaled between 1-7.

Overall assembly operators had positive experiences with MARS. Several users expressed that they liked the projection aspect of the system. Information usually not present to users on the line could be projected in an efficient way on the work surface. One user explicitly noted that our system was better than the current method of using a touchscreen.

This opinion was directly influenced by the ability to have information projected on the workpiece rather than having to look up and interact with a screen. The screen interaction is a process that is often misused by operators, as they are currently able to simply approve production steps after they were actually performed.

All of the users complained about the sensitivity and responsiveness of the projected buttons. This could be addressed by giving clearer instruction, to how to activate a button, since this process was not intuitive to the user. However, over time, each user was able to find a technique that worked. We also performed observations to assess the efficacy the system button press module, and found that the overall success rate was 95%.

Table 7-2 Summary of overall participant subjective feedback and satisfaction post the MARS production trial (see full data in Appendix B)

Question	Range: 1	Range: 7	Mean	Std. Err
How was your experience pressing "augmented" buttons compared to using a touch screen?	Prefer Touch Screen	Prefer Augmented Buttons	4.8	0.489
Did the addition of touchable buttons distract you while attempting to perform tasks?	It was distracting	It was Helpful	6.0	0.632
Did you find the buttons easy to activate?	No	Yes	5.8	0.583
How seamless was the transition between the current system to the MARS system	Not seamless at all	Seamless	3.4	0.871
How often were you aware of your current cycle time	I ignored it	I was always aware	5.6	0.678
How did your performance feedback after each build affect your performance?	No affect	Influenced my performance	3.2	0.969
How often did you finish a build within the average cycle time?	Rarely	Always	2.2	1.2
How comfortable are you using the MARS system	Uncomfortable	Very Comfortable	6.0	0.774

Button press accuracy was assessed by recording the hit and miss rates for each participant during the production trial. Since the study's completion, we have improved our multi-touch detection capabilities relying on depth data, and as a result, button press accuracy improved. Please refer to Appendix A for additional quantitative user study data. We included the raw data and additional analysis we created. We plan to continue and assess responsiveness and ease of use in future studies.

Overall, younger users (age=20-35, mostly temporary and less veteran workers on the c:scape line) were more receptive to the system than their older co-workers. We observed and heard from the veteran workers that they had ingrained habits ("muscle memory") that allowed them to achieve fast assembly times. Some indicated this as the reason they were reluctant to change to a method that strictly enforced standard work, as our system introduced. Since MARS is designed to be as nonintrusive as possible, we are confident that even senior workers will be able to achieve the performance metrics they are accustomed to after a short adjustment period. This is evident as four out of five of the test subjects are indeed considered seniors and have been working as senior operators many years. Figure 7-15 shows that overall, participants' perception of performance and experience improved during the study. This phenomenon is also visible when we examine Figure 7-13, the male subjects (1-4) are considered senior, and their learning curves are very similar. This indicates a relatively close adjustment period to the system.

The most common positive aspect of the system as expressed by observers was the amount of data passively collected. Our system could collect data that was either not collected before or that took manpower and time to collect

Another common feature to which observers reacted positively was the ability to use the system as a training tool. One participant noted: "I think it's a great tool to have; when you are training especially you can't make any mistakes. As far as a training tool, you can't beat it." This may lead to changing training processes by introducing training directly on the production line with the actual workpiece. This would supplement and possibly change the standard classroom-based training currently in use. This will also ensure that training is done in a consistent fashion for all workers, as the system removes the subjective human-based performance evaluation factors. Training is also more effective, since users are able to interact with the workpiece while being given instruction. Management also expressed interest in incorporating information they felt critical and important to the user into the system after seeing the value of having the information projected on the surface to each user individually

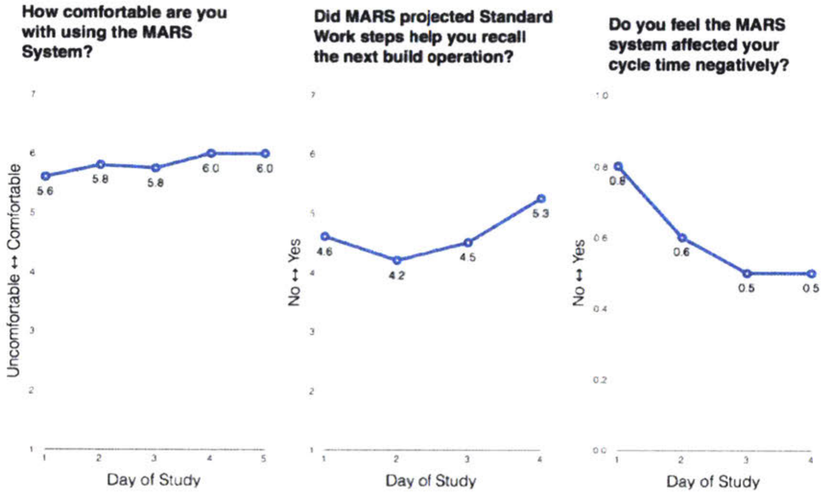


Figure 7-15 Subjective MARS user experience feedback

While most of the comments given by observers were positive, a few comments suggested improvements we are currently implementing. One observer commented on the white surface not being the best choice due to dirt that it will attract by being touched constantly. This can easily be solved by finding a more durable surface that does not affect the quality of the projection. There were concerns about the brightness and readability of the content that was projected. Our projector had to compete with the extreme lighting conditions of the factory, affecting the perceived brightness of the projected material.

The readability of the content was mostly affected by the surface we chose. Light dispersion into the material causes blurriness, therefore making the content less legible than when compared to standard displays or paper. To address the dirt and readability issues, we are searching for materials that are durable and do not disperse light as much as the current material.

7.9 Discussions and Contributions

This chapter described the design, implementation, and evaluation of the MARS system for augmented manufacturing. We have provided extensive details of a real-world evaluation of our experimental system in a factory environment. This work promotes the adoption of AR as a technology that could greatly benefit manufacturing environments. We found that the MARS system required a relatively small adjustment to physical facility and to the worker routine, while providing immense potential benefits to the manufacturing operation.

Our approach delivered a simple-to-use and easy-to-integrate AR system to the factory floor. We enabled extensive data collection that was previously impractical to obtain. Our results show that users quickly adjusted to MARS with impact on assembly performance time. However, the overhead introduced did not actually impact the overall planned work in the duration of study. The MARS system enforced consistent use of standard work procedures among study participants. We have shown that on-the-job training can also benefit from augmentation. The systematic collection of manufacturing metrics proved to be an important tool for refining the lean manufacturing process. We hope that this work will open the door for future real-world AR systems that will use the core results we present here to create solutions for manufacturing, as well as for additional domains.

8 Augmented Education Case Study

During the course of this dissertation, work I became a dad for the first time. Naturally, this event had a great impact on how I think about my work; it also served as the main motivation to explore interfaces for augmented education. The core idea was to use augmented reality interfaces to teach students various concepts in science.

Educational technology and the general availability of computers in classroom environments have made simulation and visual aids common tools for teaching abstract concepts. Simulations help students develop intuition as to how natural phenomena—which can be typically abstract and complex—behave and are composed. More recently, education technology has made a leap to online platforms such as Coursersa [201], and Khan Academy [202], helping students connect to each other and amplify learning using the vast information available on the internet.

It was my hope that we can use projected augmented reality interfaces to provide an interactive learning experience that takes the traditional simulation to the domain of Augmented Simulation. Augmented simulations (AS) exist in the real world, yet are amplified by overlaid digital information that allows a student to perceive the experience while the natural phenomena occur. Such simulations can help understand complex concepts and serve as a superior education too. It is important to note that AS simulations follow a long history of real-world physics demonstrations. Physics teachers rely on such demonstrations to explain complex topics such as motions, energy, electricity and more. Sprott published an excellent source book for physics teachers [203], that have numerous examples of such real-world simulations.

This work included quite a few early explorations, with a focus on the Enlight project that I will describe in this section. I share this work with collaborators Tal Achituv and Yi Hui Saw, who published our results and discussion in her thesis [204]. Our work was also performed in collaboration and with the support and guidance from Dr. Katherine McKnight, who at the time directed the Center for Educator Learning & Effectiveness at Pearson. In addition, we are supported by two leading physics education experts: Prof. John Belcher of the MIT Physics Department [205], is heavily involved in the effort to change introductory physics education at MIT into an interactive format. He led the TEAL (Technology Enabled Active Learning) classroom project that introduced several education technology innovations [205], [206]. He is especially interested in visualizations of electromagnetism, because it is a hard concept for students to grasp. We also collaborated with Prof. Eric Mazur who is Area Dean of Applied Physics at Harvard University [207]. Prof. Mazur is considered an expert in the domain of Interactive Teaching, and he has pioneered and published several seminal books and papers in the domain of Peer Instruction [208]-[210]. In addition, we also collaborated with Andrew Kim and his team from the WorkSpace Futures group at Steelcase. I am thankful for all the generous support we have received as team. It allowed us to explore this domain in a deep way and show how projected augmented reality can be put to use to advance education.

8.1 Motivation

The idea that humans learn from their physical environment is likely older than our written history is. However, modern education over the years has become predominantly ruled by "learning by instruction" over "learning by doing." A long chain of modern philosophers, epistemologists, and education researchers that span Locke, Rousseau, Itard, Pestalozzi, Froebel, Montessori, Piaget, and Dewey represent this development of human knowledge. Zuckerman provides an excellent review of the sources and impact of tangibility and learning in his PhD dissertation [211]. In our era, a vast body of knowledge on instruction and learning using technology and interaction emerged as computers proliferated to everyday life.

The constructivist theory [212] promotes the concept of "Learning by doing." It asserts that learning, or the construction of knowledge, occurs on the mind when one performs an interaction between an abstract idea and the experience it holds [213]-[215]. Papert's seminal research has evolved the idea further, introducing the use of tangible objects as part of the learning process. This approach has been used to create a myriad of simulation-based learning techniques [216]. Naturally, and with the evolution of computers as educational tools, virtual simulation and games have emerged as the key technology that helps support learning-by-doing approaches. Price provides an excellent review of the research in the field [217].

Active learning is an instructional approach to education that engages students during the learning process. It requires students to perform activities and think about the actions. The active learning model emphasizes the learning skill as opposed to traditional raw transmission and consumption of information [218]-[220]. Students engaged in active learning classes are expected to work in teams, share ideas, and collaboratively solve problems. This is generally referred to as peer instruction [221]. In the TEAL classroom [222], Prof. Belcher created several educational simulations and the instructional environment that promoted student to explore such active learning methodologies. He focused on teaching complex phenomena in electricity and magnetism. Simulations were made available to students using PC applications [223]. On the other hand, Prof. Mazur at Harvard uses a collaborative approach to teach undergraduate physics. Teams of students collaborate during the term on engineering projects and distill physics knowledge in the process. Both serve as great examples for real-world active learning success.

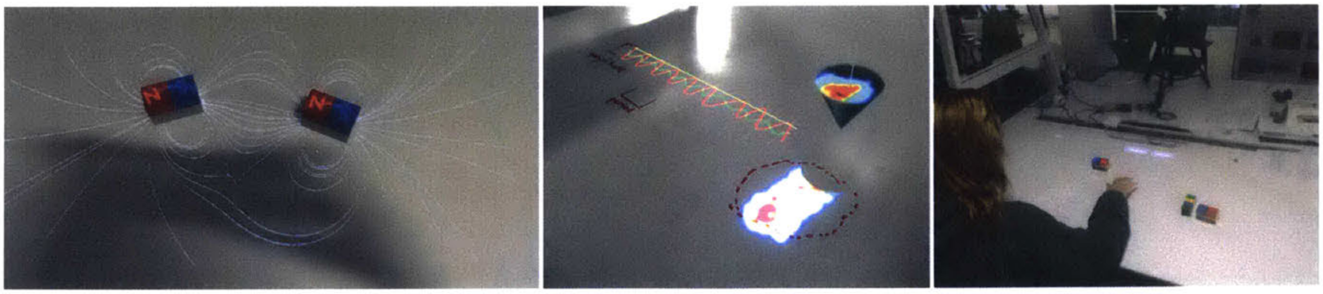
Both peer instruction and active learning served as important design inspirations and requirements in the design our system. To truly advance traditional computer simulations, we needed to create environments that allow students to freely explore physical phenomena while enhancing it with digital information. Specifically, this approach emphasizes the subjective and qualitative experience perceived by the senses over a qualitative experience reliant upon information consumption, rigor, and intellectual abstraction capabilities.

8.2 Augmented Simulations

The parameters of any given simulation are set at a point in time, defining the constraints of the simulation. The projected AR interfaces this dissertation presents attempt to blur the boundaries of real-world phenomena and augmentation that adapt per the natural phenomena, and as such, become much more open ended compared with traditional computer simulation. I believe that Augmented Simulations (AS) are, in fact, more customizable and adaptable to the user who can alter the simulation with real-world objects. My work on the Physics AS presented in the Enlight project brought about this concept.

Augmented simulation (see Figure 8-1) is particularly suited for sciences education that requires deep understanding of underlying abstract concepts. Such concepts can come in the forms of complicated models, formulas and equations that provide a simplistic representation to complex phenomena. Some of these phenomena are not always easy to perceive with simple human senses. For example, physical forces may act on an object, yet they are invisible. A biological mechanism, for example, can be microscopic hence impossible to perceive with the naked eye. As a result, it is hard to teach those concepts using simple instruction. Traditional simulation has come to fill this gap, and it has been shown that visual aids and help with the understanding of complex scientific concepts [224], [225]. Augmented Simulation holds the promise to take the traditional simulation to point where students can use their senses to not only understand the natural phenomena in situ, but also to develop intuitions based on the additional layers of information that annotate the real world with real time context.

Augmented Simulations enable us to make the invisible visible. The emergence of high-performance computing and sensing capabilities creates a new possibility. Imagine having the learner explore the science concepts in the physical world, just like in lab experiments, but with the experience enhanced with the digital simulations previously available only on screens. This is made possible with object tracking and image projection onto the objects themselves. In other words, augmented reality interfaces have the potential to create direct personal experiences described by Papert with tangible objects [226]. By painting on the physical objects and making the phenomena that are invisible to the naked eye visible—like field lines emerging from a magnet or microscopic views of the skin—learners can gain direct experiential intuitions on how the natural world functions. Combined with good instructional design, augmented reality can provide a learning environment for scientific knowledge that engages all the learner senses.



(a) (b) (c)

Figure 8-1 (a) Enlight's magnetic field lines augmented simulation activity; (b) Augmented pendulum simulation visualizes the motion of the pendulum and graphs the X, Y, Z motion components as the system tracks the pendulum; (c) The Enlight system supports augmented education content creation, allowing teachers to create lesson plans. here we show a student taking the magnetic fields augmented lesson activity.

One of the largest efforts to reform how introductory physics courses were taught in the United States was pioneered by the MIT Physical Science Study Committee (PSSC) [227]. Educators found that textbooks were ineffective in stimulating student interest in physics and failed to teach students how to think like a physicist. After the Soviet Union sent Sputnik into orbit, science education leadership in the United States feared that American schools lagged behind in science, and efforts at the PSSC were expanded. The PSSC designed many teaching materials aimed to improve understanding and engagement, as opposed to memorization. These included films, teacher guides, experimental apparatus, and an accompanying textbook for all the materials. For phenomena that were more complex, specialized equipment was built, and the explanations for them were recorded on film. The effort's impact was felt worldwide, and by the 1960s, over 20% of all American high school physics teachers were involved in the project. One notable example of this effort was captured in the famous Richard Feynman lectures [228] that were widely successful promoting physics education to a wide audience.

Despite their clear advantages, computer simulations have definite drawbacks. In principle, a screen-based computer interface experience introduces an attention context switch between a viewer and the physical phenomena. Accordingly, computer simulations are naturally virtual and therefore lack tangible feedback. Where possible, such feedback is key to developing deep understanding of abstract concepts and reduce potential learning [229], [230]. Finally, simulation is finite and programmed, while our physical world is not. The behavior change of a physical object in an experimental environment is governed by natural physical laws (e.g., gravity, magnetism) that always operate.

8.3 Related Work

The emergence of the modern computer brought about new efforts to increase student engagement in physics courses. At MIT, the TEAL (Technology-Enhanced-Active-Learning) project [222] aimed to further reduce the mismatch between traditional teaching methods and how students actually learn. Instead of the traditional lecture format, students were given shorter lectures interspersed with discussion questions designed to emphasize correcting common misconceptions. Students also conducted experiments in class and were given animated simulations to visualize the concepts taught (see Figure 8-2).

At Harvard, Professor Eric Mazur designed his class, Applied Physics 50, to feature team projects throughout the course. Students use their new knowledge to reverse-engineer musical instruments, build Rube Goldberg machines to complete specific tasks, or even design circuits for secure safes. The idea was that students will need to learn the concepts to be able to successfully complete the projects **[209]**. Mazur also developed a correlative site, Learning Catalytics, which allowed students to provide feedback and solve concept questions on their mobile devices. This new approach showed the largest gain on the force concept inventory (a measurement of students' understanding of basic concepts of any Harvard physics course taught in the six years), and class attendance surged to 97% in the fall of 2012 **[231]**.

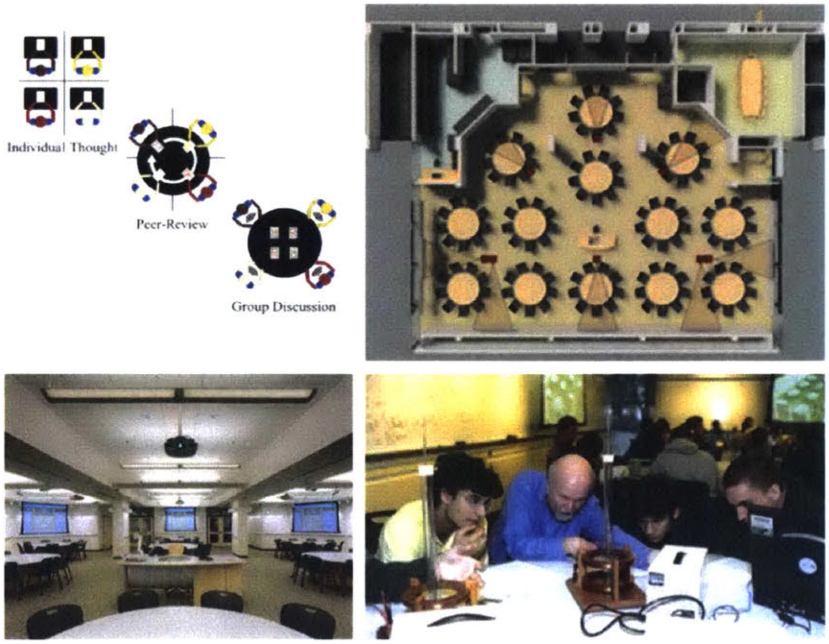


Figure 8-2 MIT TEAL – Technology Enhanced Active Learning Class Room

Augmented reality displays can take many forms. One class of displays utilizes a secondary screen such as a tablet or a laptop to provide the additional graphical information to the user. One study at the University of Washington applied this form of AR to an application for teaching chemical structures **[232]**. Users would hold fiducial markers in front of a webcam, and the laptop would render the structures associated with the specified markers. When this interface was compared with physical models, students who had cited a preference for the physical models indicated that they liked being able to hold, rotate, and count the atoms on the models. Blanca et al. conducted an experiment that utilized tablets as the secondary screen for overlaying visualizations on fiducial markers [233]. The researchers developed an application that allowed students to construct circuits using marked boxes. Current flows in circuits would then be shown when a tablet was held over the boxes. Similarly, learners can also view electromagnetic field lines using the interface. When compared with a basic web application, their research found that students learned significantly better on the AR interface, based on pre-test and post-test results. For additional interesting work in this domain see Wojciechowski [234] and Billinghurst [235].

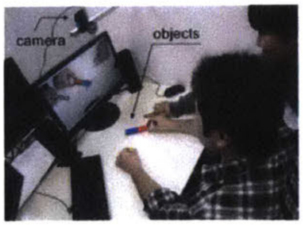


Figure 8-3 Real-Time Visualization System of Magnetic Field Utilizing Augmented Reality Technology for Education Shinya Matsutomo et al., 2011

Projected AR systems create a new possibility over what secondary screens can provide for education applications. They allow students to use their hands freely without needing to hold any device, while digital information is projected directly on the physical world. The TinkerLamp [236] combines projection and a paper-based interface for teaching geometry and fractions. The learner can move pieces of marked paper that represent a specific shape or number around and under the Tinker Lamp. The projection would then show additional information that helps the user answer questions on a problem sheet that is also marked. There are several additional examples of projected AR systems used for education, please refer for to Bower et al. for an excellent recently published review [237].

Figure 8-4 Augmenting Magnetic Field Lines for School Experiments, Florian Mannuß et al., 2011

Head-mounted displays (HMD) also served as a great tool to teach complex scientific phenomena. Shelton et al. used HMD to teach earth-sun relationships in 3D space. Kaufmann et al. found that HMDs can help improve the spatial comprehension of 3D geometry [238].

More recently, with the emergence of advanced mixed reality headsets, new immersive educational experiences are appearing. One example is HoloStudy [239] that makes use of holograms to teach astrophysics.

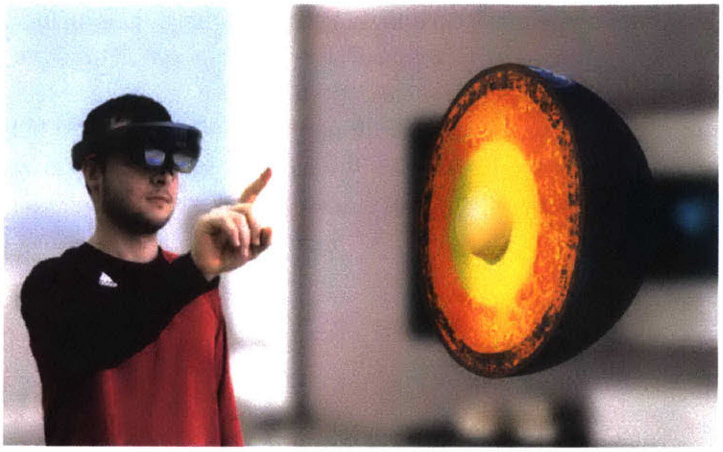


Figure 8-5 HoloStudy offers immersive educational experience using Microsoft's HoloLens.

8.4 Research Questions

We focused on the following research questions to advance the empirical findings as these relate to the real-world AR interface deployment to an educational setting, such as a lab or an interactive classroom. The primary goal of the augmented education experiments was to measure the impact of augmentation on the learning process. Specifically, in this part of this dissertation work, I address the following questions:

1. *How might projected digital information superimposed on the physical world assist with abstract concepts' learning activities?*

Specifically, we are interested in assessing how a student's understanding of high level abstract concepts changes because of using the system.

2. *To what extent do students perceive the use of the system as engaging?*

It is important to evaluate how students perceive the user experience the system provides, with respect to the lesson's content we selected to run the experiment on. Later in this chapter, we describe in more detail the specific lesson content we focused on - the principles and mechanics of magnetic field forces.

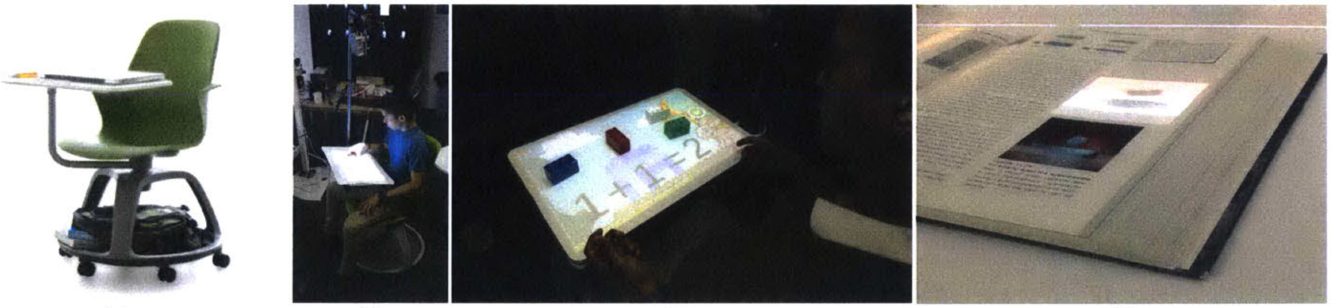
3. *Can we deduce clear advantages to the projected augmented-tactile learning vs. comparable virtual representations?*

Any user interface evaluation should be comparative. Our main comparison baseline in this research is the non-augmented state as students experienced it in the classroom. The focus of our evaluation centered on understanding what students believe is the most helpful educational aid for them and why. Additionally, we are curious to understand the student's learning curve when interacting with a new type of interface.

8.5 Design Explorations

8.5.1 Node Chair

Our first exploration into augmented education began with reimagining the iconic Node chair from Steelcase **[240]** (see Figure 8-6(a)). The Node chair is designed to support active and shareable learning environments. Active learning refers to a set of educational instruction models that shift the responsibility of learning from the instructor (teachers) to the learners (students) **[219]**.



(a) (b) (c) (d)

Figure 8-6 (a) Steelcase Node chair; (b) Hacking the Node chair, we added a specialized LuminAR device that augments the work surface; (c) augmented counting application; (d) augmented chemistry text book application;

One of the key elements of active learning is the reconfiguration and dynamism of the learning space. The physical space should support the pedagogies at work in the classroom. It was therefore almost natural for us to choose the Node to host our LuminAR system. By collaborating with Steelcase, we modified the Node chair to explore how projected AR interfaces could be incorporated in the classroom. We integrated the device in a way that retains all of the original flexibility the chair offered to encourage collaboration. The projected interface allowed for immersive experiences while encouraging multiple students to interact with the projected augmented reality technology.

We created two demo applications to explore the affordances of the chair after we enabled it with a projected interface.

8.5.2 Augmented Count Application

The first application was designed as an early age (ages 4–5) counting learning experience. In this simple application, we illustrate how the technology can be used to teach students how to count, using physical LEGO blocks. It shows the possibilities for students to learn naturally by playing with objects in the physical world and using them to interact with information in the digital world, opening possibilities for new experiences. The Count game is a simple Lens application that tracks objects that are placed on the interaction surface. It displays a colorful and playful playing board that presents the used with a simple arithmetic problem. We used LEGO Duplo [241] web interfaces and blocks as transitional objects. But the system can also work with arbitrary objects. The objects are detected by the LuXoR vision engine's contours module. The gameplay progresses by presenting the users stages that teach the basic operations such as addition, subtraction and multiplication. We installed the system on the Steelcase Node chair (see Figure 8-6)

Figure 8-7 Count application user interface with physical object interaction. On left, the Count application is showing a simple count step. The mug is placed in the white projected boundary, it is detected, and the count is incremented. On the right, the user is presented with a simple addition statement.

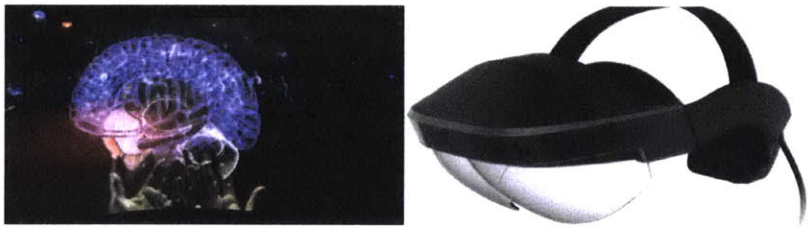
8.5.3 Augmented Textbook Application

The second application was a physical textbook augmentation, targeted at early college education. Using LuminAR's capability to detect specific objects, we created an augmentation scenario for a college-level chemistry book. Once the book is detected, supplemental educational content becomes available to the reader.

There are several examples of augmenting text books with AR content [242] using standard see-through mobile AR methods [243], [244] .

More recently, we can point to great examples of using mixed reality headsets such as Meta to create immersive educational content that fundamentally stretch the bounds of the traditional text book. Radke provided several example in his recent TED talk [244].

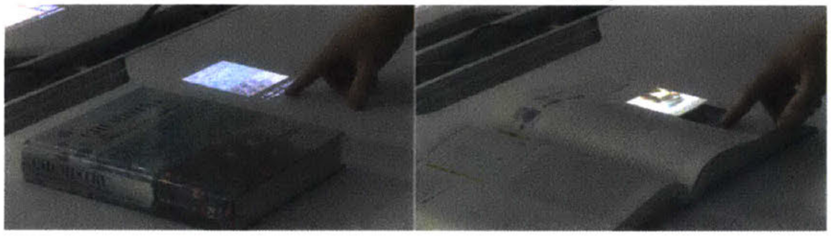
In our case, we wanted to use the classic book, and augmented it with light. This is not new and was explored well by others [245], [246]. We used this as an exercise to use the components of the LuminAR hardware and software to quickly prototype such an application. We installed the Augmented Book application on the Verb table that we used for the white:scape project we described in Chapter Six.



(a) (b)

Figure 8-8 (a) Meta's brain image demo uses volumetric rendering (b) Meta 2 Headset (source: vbandi.net, Meta)

We briefly describe the application scenario (see figure 8-9 below). We used a chemistry text book. The cover of the book serves as a marker for the system enabling the book detection. As soon as the book is detected, LuminAR can display relevant information about the book itself. In our example, the system prompts the user about AR demo content that is included with the book on certain pages.



(a) (b)

Figure 8-9 (a) The book is recognized and displays pointers to AR content within the book (b) when the user is flipping to the correct page a video augmentation is triggered.

For example, we had a "super conductor levitation demo on page 439"; when the user flips to the correct page, the system retrieves the simulation demo from the database and projects the content in a location relative to the printed text and pictures in the book. Hence the video supplements the text and helps visualize and explain abstract concepts that are hard to understand.

8.6 Enlight

Enlight is an end-to-end system for employing AR in education. We envision the system to complement existing approaches in science instruction. The system allows an instructor to incorporate AR applications with existing learning materials and to provide easy deployment of the application to the learners. One initial question was to identify use cases that benefit from the enhancement of augmented reality. In the current iteration of the system, we focused our development and subsequent pilot studies on science education, specifically physics. Some previous research on using AR in science education made use of handheld see-through AR displays and stereo video see-through head-mounted displays [247], [248]. This section will describe Enlight's system design and implementation.

8.6.1 System Design

In physics education, virtual simulations have allowed us to show and explain phenomena that are otherwise invisible to the naked eye. However, experiments with analog devices still play an important role. They allow us to verify theories and discover ideas through experiments that are not constrained by software. What if we could combine the best of both worlds? We achieve that by building our applications on a projected augmented reality system. By projecting onto physical objects, we can paint invisible phenomena. With our system, we have built "physical playgrounds"—simulations that are projected onto the physical world and respond to detected objects in the physical-augmented space. Thus, we can draw virtual field lines on real magnets, track and provide history on the location of a pendulum, or even build circuits with both physical and virtual components.

The key design principle of Enlight was to provide just-in-time information to the learner, allowing her to freely explore a physical phenomenon in the real world, and interacting with projected digital information at the same time. Enlight was designed with two key users in mind: teacher and student. We summarize the design goals below.

8.6.2 Open-Ended Augmented Simulation Space

The focus on physics AS dictated designing an environment where a physics rule engine can produce augmented reality content dynamically per the sensed state of the real world and make it available to the student who experiences the simulation. Standard computer simulations are naturally limited by the content of virtual simulation objects that were programmed into the simulation. Augmented simulations are open ended and can theoretically work with an unlimited number of real world objects. However, AS are also limited by the ability of system to detect such objects as well as user actions, and supply them to physics rule based engine to support a full augmentation feedback loop.

8.6.3 Collaborative Learning Space

Projected interfaces are naturally collaborative. They can be experienced by multiple users at once. This fact lends them well for use as an environment for active learning [249], where students can solve problems together and teach each other [247], [248]. Research shows that active learning improves abstract concept understanding [218]-[220]; thus, our goal was to create an environment that enriches collaboration to support such active learning processes.

8.6.4 Dynamic Content, Live Data

A great active learning experience is dependent on the teacher's ability to create content that will engage students and naturally compel them to work on a problem together. Enlight is designed to make it easy for teachers to create multimedia, dynamic learning experiences. In addition to making it possible to create engaging content, is also important to measure the efficacy of those experiences. Enlight was also designed to utilize web-based interfaces to collect user interaction data and provide real-time analytics to the teacher during an active learning session.

8.7 System Implementation

8.7.1 Classroom Furniture

In this section, we describe the key technical concepts of the Enlight system. The Enlight system was installed on a Steelcase Verb table [250]. The Verb classroom furniture line was designed to support a reconfigurable classroom environment.

We collaborated with the Workspaces future group to create two prototypes: In the first prototype the system was directly situated on steel construction that was added to the Verb tables that were adapted to carry the LuminAR hardware.

The second variant used a special monitor arm that carried the LuminAR hardware and also adding mount for a tablet. The furniture and hardware setup supported the classroom mock-up experiments well.

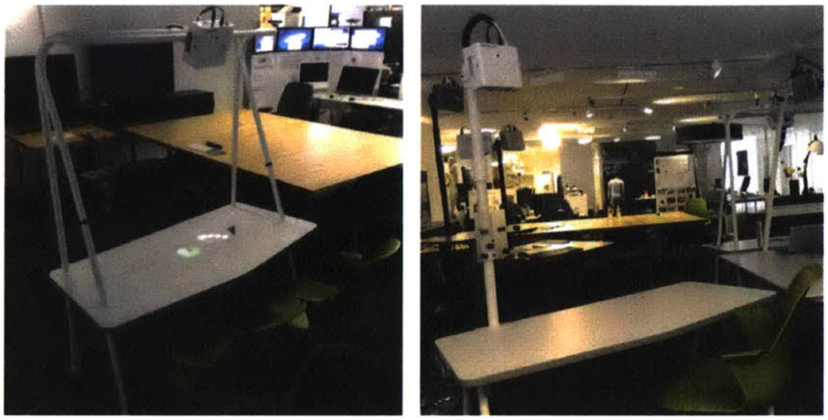


Figure 8-10 Steelcase Verb table installed with two variants of the Enlight system; On the left we show the arc configuration, on the right we show the single mounting pole configuration.

8.7.2 Building Blocks for Active Learning

The key elements of the system consisted of:

- Augmented simulation lessons - a sequence of learning steps, typically focused on single concepts. A step is very similar to a slide in online presentation tools, yet the content was designed to interact with the physical object in the augmented simulation scenario.
- Lessons could contain various multimedia materials, such as videos, images, or other dynamic GUI elements
- Lessons could contain activities or quiz questions that would be presented to the student at the right moment per the instructional plan implemented by the teacher.
- A special type of activity we called "playgrounds," which supported an open-ended exploration environment, where the student could freely interact in the state of augmented simulation.

In the next section, we describe the specifics of the AS playgrounds we created to explore usability and prepare for the overall evaluation of the system.

8.7.3 Augmented Simulation Playgrounds

The AS playgrounds were designed to explore different physics education concepts and provided early validation of our design direction.

Figure 8-11 Enlight's application launcher (left) and Playgrounds menu (right)

We implemented the playgrounds using the Lens software framework discussed in Chapter Four. We summarize the specifics of each playground we implemented.

Vectors

- Designed to teach basics of Euclidean vector space definitions and operators.
- Devised as a game scenario using navigation of a boat in river as a key metaphor.
- In each stage, the force and direction of the vector was determined by a user touching points on the table to set the boat's direction and navigate towards the target goal.
- Interaction was simple and used a touch-screen-like mechanism.

Pendulums

- Designed to teach basic mechanics of a pendulum's harmonic motion, using an historical heat map that revealed the invisible path of the pendulum as it travels.
- The pendulum weight could be any object, and have different cord lengths. We were constrained by the system's physical dimension and the detection range of the sensor
- Time series visualization of the 3D (X,Y,Z) coordinates that represented the pendulum travel location were graphed alongside the real time augmentation of the pendulum

- This AS provided haptic feedback, while students could speed up and slow down the pendulum motion and immediately see the results in the visualization and the time series graph

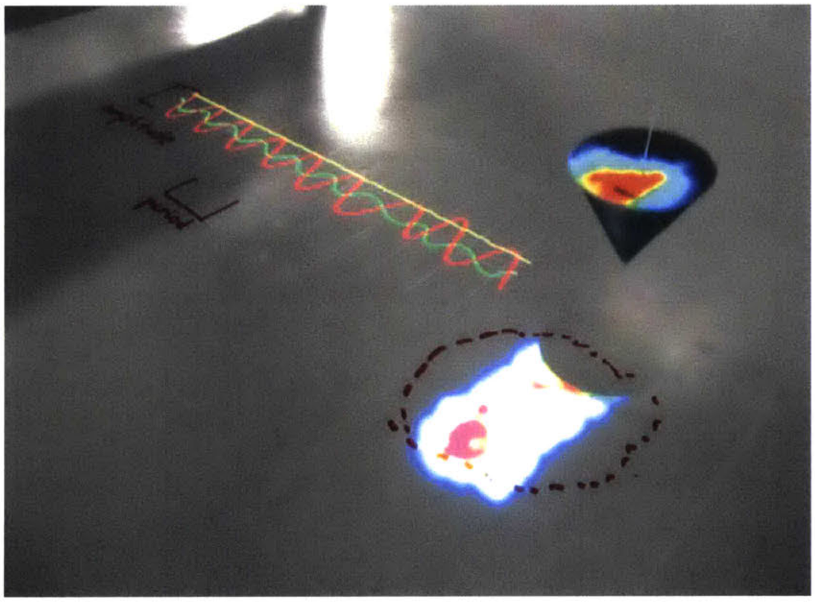


Figure 8-12 Pendulum Playground

Vector Fields

- Designed to teach the key concepts of magnetic vector fields.
- The system was extended to detect the location and type of magnetic dipoles and monopoles.
- The system would track and simulate vector fields according to the physical state detected.

Magnetic Fields

- Extended the core idea of the Vector Fields AS to visually represent a magnetic field.
- Adopted Prof. Belcher's physics teaching approach **[205], [206]** we incorporated real magnets to this AS. Magnets were color coded red and blue and would be tracked by the system as a dipole.

- We extended the system to also detect the contours of a magnet block, allowing us to generate a visualization of the magnetic fields and project while tracking the object.
- Using color-based tracking of the location of the dipole, we could calculate the moment vector based on orientation and the contour. Saw's research included a good review of the mathematical approach we took in her master's thesis [204].
- The user experience allowed students to perceive and predict the motion of the magnets as a function of the visualized magnetic field.
- We used a compass to enhance the experience; when students were moving the compass on the virtual magnetic fields lines, the compass needles would naturally (and physically!) adhere perfectly to the simulated projected lines.

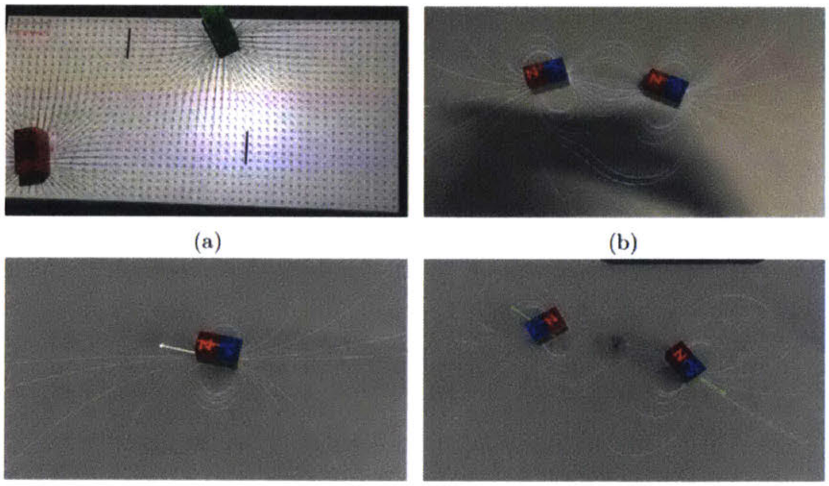

(a) (b) (c) (d)

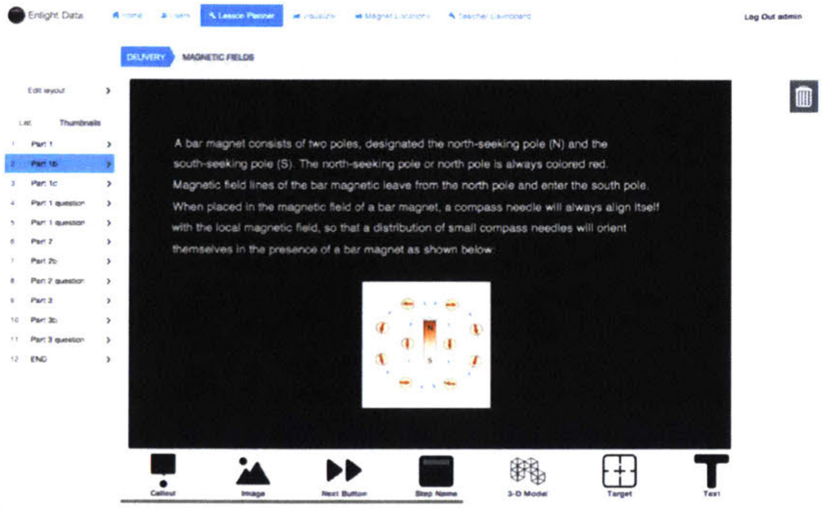
Figure 8-13 (a) vector fields playground first iteration (b) vector fields playground second integration with color support (c) a projected yellow arrow indicates the dipole moment vector. (d) projected green demonstrates the force vectors on the magnetic blocks. The compass needle accurately matches the projected magnetic field lines.

8.7.4 Interfaces

Lesson Builder

The lesson builder is a web application for teachers to create AS content for the Enlight system. It follows a simple online presentation user-experience scheme. The teacher can use a drag-and-drop interface to compose a multimedia experience. The tool also allowed the teacher to define tracked objects using either a marker or markerless approach.

Teachers could create applicative flow based on simple conditions and triggers created to advance the student through the designed steps in the lesson; the system requires no programming experience of the teacher.



A bar magnet consists of two poles, designated the north-seeking pole (N) and the south-seeking pole (S). The north-seeking pole or north pole is always colored red. Magnetic field lines of the bar magnetic leave from the north pole and enter the south pole. When placed in the magnetic field of a bar magnet, a compass needle will always align itself with the local magnetic field, so that a distribution of small compass needles will orient themselves in the presence of a bar magnet as shown below:

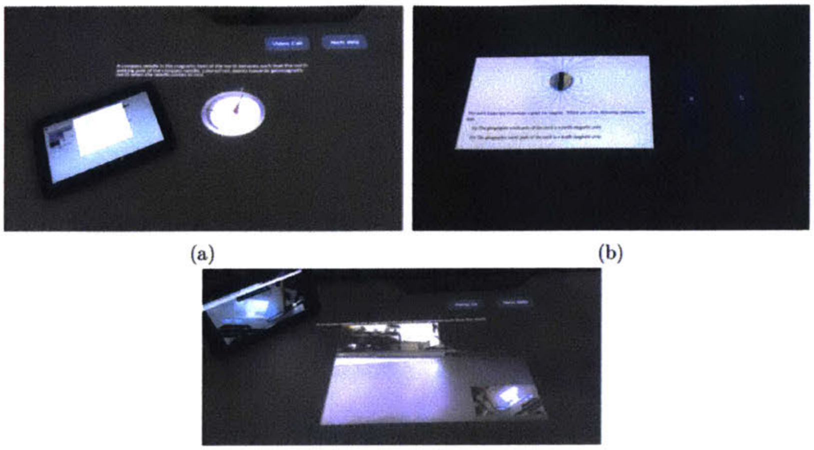
Figure 8-14 Enlight's Lesson Builder application. Users can drag widgets from the bottom bar to the art board to create desired screen composition and applicative behavior.

Lesson Player

The lesson player was implemented using the LuminAR Lens framework. It was specifically adapted to load the content from the lesson builder. The player handled all required client/server transactions and provided real-time feedback to the backend databases. The real-time approach to data implemented in Lens made it easy to provide teachers with feedback as students were progressing through the lesson plans. The player recorded data such as time spent on lesson steps or time to answer quiz questions. We also incorporated video conferencing capabilities using WebRTC [251] to facilitate student-teacher communication in an online setting.

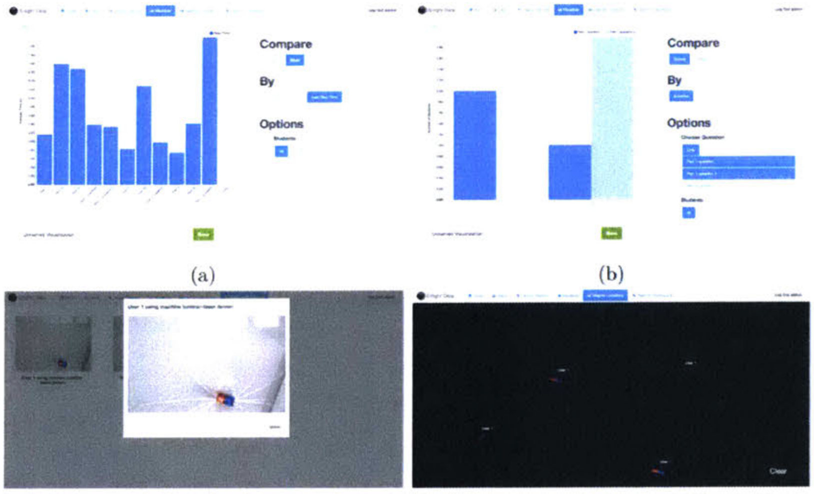
Classroom Dashboard

The classroom dashboard provided live information on the lesson progress, visible to both student and teacher. It was designed to explore the full classroom scenario where multiple LuminAR systems might be deployed in a classroom environment, giving the participants the means to understand with a simple glance the progress of multiple activities taking place in the classroom. In another scenario, the teacher would use the dashboard to manage the requested communication she would get from participants in their activities.


(a) (b)

(c)

Figure 8-15 Enlight's Lesson Player; (a) the player receives real time updates between the tablet and the lesson builder (b) Player showing a quiz application (c) Enlight App with two-way video using WebRTC.



(a) (b)

(c) (d)

Figure 8-16 Enlight's classroom dashboard: (a), (b) shows class room analytics. Teachers can graph the amount of time students spend per activity (c) and (d) shows snapshots from the different student players are available for introspection by teacher and peers

8.8 User Studies - Classroom Test Trial

In this section, we describe the formal user study we devised to evaluate the Enlight system and its applicability as an educational tool. The study was titled Effects of Augmented Reality in the Learning of Concepts in Physics.

8.8.1 Evaluation Methodology

Our evaluation intent was to measure the effects of using a projected augmented reality interface to teach and learn concepts in physics, by measuring the change in performance of subjects completing assessment tests. Students also participated in think-aloud self-assessments as they worked through an exercise worksheet similar to a typical in-class lab. The Enlight system provides them with the augmented simulations and tools they need to answer the questions in the activity worksheet. The purpose of the study is to show how learning can possibly be improved with a more interactive system that encourages tangible learning, as we described in the research questions section earlier in this chapter.

8.8.2 Study Population

The study participants (n=30) were recruited from Prof. Belcher's MIT undergraduate Physics class (MIT Course 8.02). The class covers basic physics concepts in electricity and magnetism. The class is taught in the TEAL format.

8.8.3 Inherent Control Baseline

We intentionally selected the TEAL classroom population as they have been exposed to active learning aided with multimedia content. During the term, the students were exposed to visualizations of physical phenomena, such as electricity and magnetism using Java applets and YouTube videos. This fact served as an inherent control baseline that helped compare our augmented simulation methods with the more traditional computer simulation approach.

8.8.4 Study Experimental Procedure

The experimental procedure was devised in collaboration with and under the guidance of Dr. Katherine McKnight, who directs the Center for Educator Learning & Effectiveness at Pearson. Dr. McKnight was instrumental in providing us with the correct mental model to conceive an educational study. Prof. John Belcher provided the essential subject matter physics context and served as our pedagogical advisor.

The key components of our experimental procedure were:

- Conceptual test—subject matter pre-test and post-test.
- Online survey—to gauge qualitative subjective feedback.
- Think-aloud session with the students—used to collect the student feedback while engaged in the activity.
- Data analysis of all the information the Enlight system captured during the experiment lessons.

Our study focused on the magnetic fields subject matter, as described above. Prof. John Belcher designed a worksheet to complement the magnetic field playground. The worksheet consisted of four sections, each introducing a different concept on dipoles. We provided minimal instruction to students as to how to use the system, since we wanted to maximize the learning curve effect we wanted to measure.

The activity was intentionally designed to be open-ended. We placed no emphasis on actually explaining the subject matter formal definitions and equations to the students. This encouraged conceptual exploration by the student participants using the system.

8.8.5 Study Protocol

We conducted the study at the MIT Media Lab in three different locations. In each location we installed a full Enlight setup. Groups of three participants were given a short introduction, worksheet, and magnets, then were directed to separate location to conduct the study. We had a researcher-observer in each location alongside the participant.

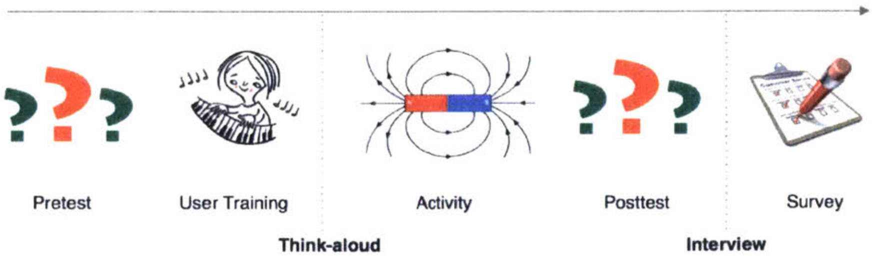
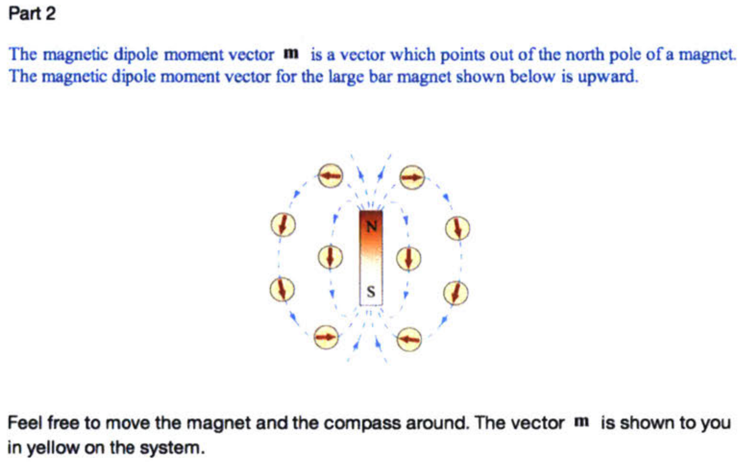


Figure 8-17 Enlight Study Protocol

Pre-test and Post-test

We used the Enlight system to perform the pre-test and the post-test quiz. We recorded the time spent by each participant on the pre-test and the post-test questions. All answers were stored in the Enlight system's database.



Part 2

The magnetic dipole moment vector **m** is a vector which points out of the north pole of a magnet. The magnetic dipole moment vector for the large bar magnet shown below is upward.

Feel free to move the magnet and the compass around. The vector m is shown to you in yellow on the system.

Figure 8-18 Example activity from worksheet

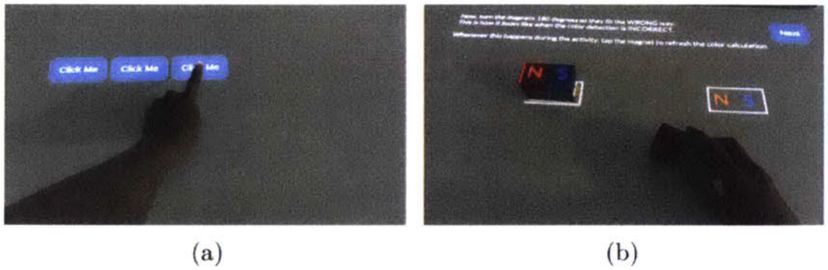
User Training

The participants were generally tech-savvy, and more so as they were all students in the TEAL class; however, most of them had never interacted with a projected reality system before. We decided to provide users with minimal training, provided in the form of a lesson.

The training lesson prompted the participants to "move an object" or "click a button here." We believe this training was essential to reduce noise from misuse of the system due to a lack of basic understanding of how AR interfaces work. However, we don't believe this training had a serious impact on our effort to measure the learning curve of using a projected augmented-reality interface.

Think-aloud session

The classic think-aloud technique defines the "productive thinking" approach teaching researches to speak to participants so they can further understand their thought process as they complete an activity [252]-[254]. In usability studies, researchers typically expand the method to include probing using neutral cues to encourage participants to share more without introducing bias into the experimental results [255]. To capture the think-aloud sessions, the experiment was video recorded and later transcribed. Participants were not given a time limit to complete the activity, and the researchers were careful not to rush them. The research staff was trained to provide a unified set of prompts to facilitate the think-aloud session.



(a) (b)

Figure 8-19 The Enlight user training app helped familiarize first-time users with the basic use of the system

We coded all the direct quotes from the participants into common recurring themes, generating a cognitive model of the participant pool. This was done following the Someren at el. [252] procedures representing and tagging actual events in the experiment. The annotations were conducted by the staff of Dr. Katherine McKnight at George Mason University, who were introduced to the project at the analysis phase. This step provided an additional layer of independent review of the transcripts compared to our results. It serves as an additional layer of objective analysis of the results we present here.

Online Survey

Once the study was concluded, we performed an online survey. The survey focused on the usability aspect of the Enlight system. We also asked the participants to provide their perception of how helpful the system was to achieve the educational goal upon which we focused. The survey questions consisted of multiple choice, ranges, and freeform questions. Participants were given an Amazon gift card as a reward for completing the study and the online survey.

Interviews

To supplement the think-aloud sessions, we conducted secondary interviews with participants. The questions asked were repeats of questions first introduced in the initial interview, designed to verify several concepts at the end of the experiment. The interview questions covered user experience with the system, as well as addressing the subject-matter knowledge of magnetic fields.

8.8.6 Results Analysis and Discussion

In this section, we present our experimental results and discuss our analysis. We can dissect our findings into three distinct categories:

- **Cognitive model**—comparing the impact of the system on participant results, while relating the findings to student perception of their own performance. In other words, we want to understand how the participants perceived the augmented simulation impact on their learning process.
- **User experience**—analyzing the usability of the system per the student observations during the experiment.
- **Student preferences**—collecting and analyzing feedback based on survey responses and think-aloud sessions to distill key recurring themes informing what users like and dislike about the system.

Cognitive Model

The cognitive model helps us understand the impact of augmented simulation on the students' conceptual performance understanding the topic of magnetic fields during the experiment.

We posed specific questions as we began building the cognitive model:

1. What is the change in student performance between the pre-test and post-test?
2. Which of the students demonstrate improvement and why?
3. Which of students did not show improvement and why?

To support the pre-test and post-test score comparison, we first needed to define the concept of learning gain.

Learning Gain

Learning gain (g), is defined as a measure of relative improvement between a user's pre-test score compared to their post-test score. We present this as a ratio of success percentages between the pre-test and the post-test:

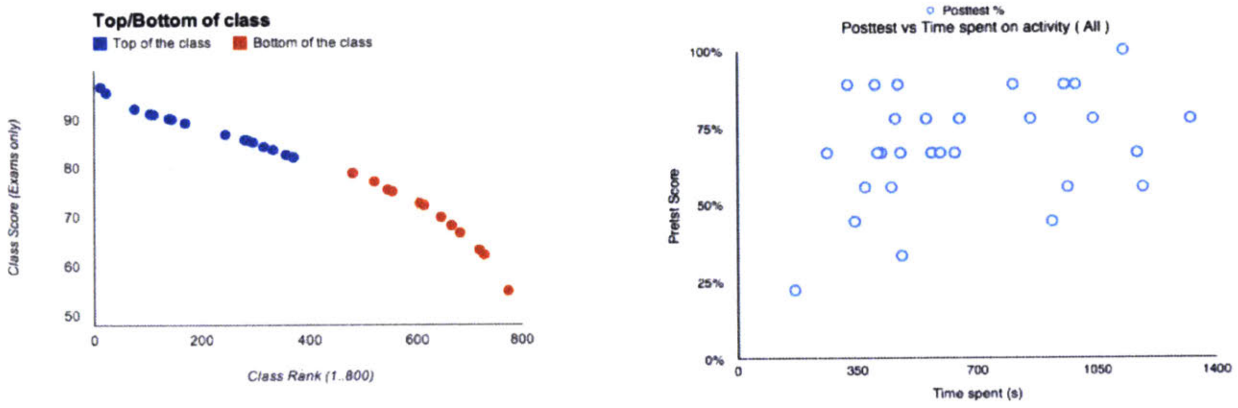
$$\boxed{g = \frac{Posttest(\%) - Pretest(\%)}{100 - Pretest(\%)}}$$

Figure 8-20 Learning Gain definition

Class Rank

Next, we needed to define the notion of class rank. Class rank is simply the distribution of grades in the Spring 2014 semester MIT 8.02 course, ranging from 1 (top-rank) to 800 (bottom-rank) for all students in the course. The actual rank is calculated based on the test scores only. Figure 8-21(a) shows the distribution of study participants compared to the overall 8.02 course population of 800 students.

When we compare the pre-test and post-test results across the study population (n=30), we can further understand the variance in learning gains across the group as a function of the larger population of the entire 800 students in the class. It provides a good foundation and baseline to assess the participants' level of conceptual knowledge in the magnetic fields subject matter. This metric is useful to put the pre-test and post-test results in context.



(a)

(b)

Figure 8-21 Distribution of participants from 8.02. (b) Post-test scores vs Time spent on activity: Time spent on activity refers to the amount of time spent on the activity portion of the study where students were exposed to the magnetic fields playground

If we limit the analysis to the test score (Figure 8-21(b)), it is hard to find correlation between the pre-test and post-test scores or class rank across the board. It is also hard to indicate a strong correlation between the amount of time spent in training to use the system

This trend is also visible from Figure 8-23. We cannot say the same about the student who belongs to the top 50 of the class (400 < rank < 800). This may suggest that students' level of conceptual understanding and prior knowledge of the physical phenomena, may impact their disposition towards the augmented simulation, thus impacting the overall impact. To put it simply, if the students already have a deep understanding of the abstract concept, the augmented simulation does not add much value to their learning experience. This may explain why strong students spent less time.

Time Spent vs Improvement Correlation

With the definition of Learning Gain and class rank, as well the initial observations above, we can now compare and check the correlation between the time spent on the augmented simulation and the rank position in the class. Using the Pearson correlation coefficient **[256]** and the Spearman correlation coefficient **[257]** we can compare learning gain per each participant and the amount of time spent on the activity. Table 8-1 shows the results for both groups: students from the top 50% of their class and students from the bottom 50% of their class.

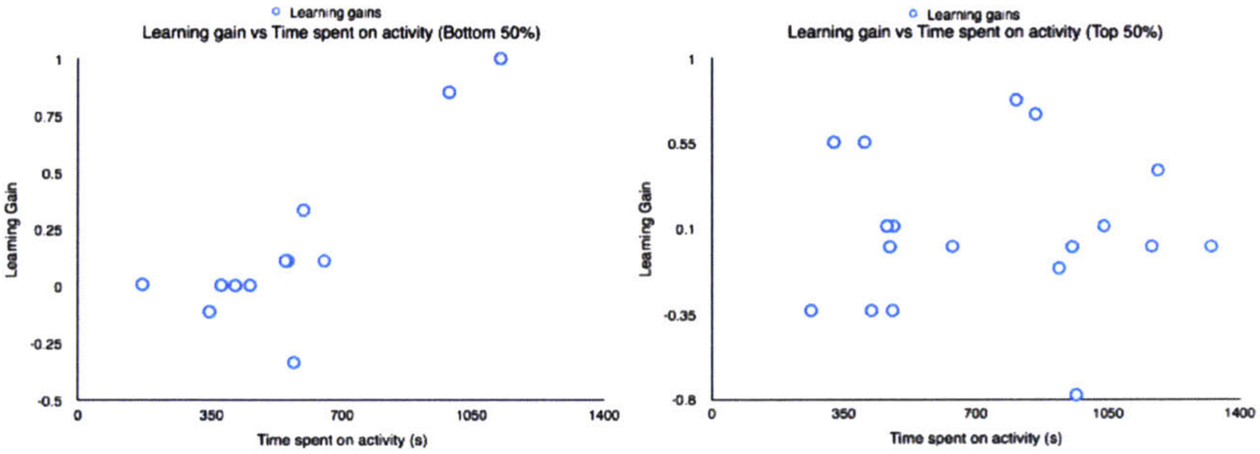
The analysis of these results provided by Saw **[204]**, shows that preliminary comparisons from our study indicate significant strong positive correlation (0.8465 ∫ 0.00195) between time spent on the augmented simulation activity and the resulting learning gains for students from the bottom 50% of the class, more so than for the top 50%. We also believe that there is a high (r = 0.8465) and significant (p < 0.005) correlation between time spent on training with the system and the resulting learning gains for the bottom 50% group. We used both Pearson and Spearman as a sanity check due to our small sample size and the results are consistent as shown in Table 8-1.

Table 8-1 Correlation values between time spent on activity and learning gain g

Measures and Values	Bottom 50% ($n = 12$)	Top 50% ($n = 18$)
Pearson correlation coefficient, r	0.8465	0.0195
$p-$value from Pearson r	0.002009	0.934966
Coefficient of determination, R^2	0.7166	0.0004
Spearman correlation coefficient, r	0.69138	0.01868
$p-$value from Spearman r	0.01276	0.94135

In Figure 8-24, we show the time spent on augmented simulation activity vs learning gain g. It reinforces our prior assertion that students who are already higher performers are subject to a ceiling effect that limits their gain from our augmented simulation and very likely from other learning tools as they have already mastered the concept. The interesting and meaningful discovery we made is that the lower-rank students were performing at the same level as the students with higher rank in our conceptual tests.

Since our work was centered on the design and testing of student knowledge gain using augmented simulation, our research did not handle the pedagogical validity of the actual questions posed to the student to assess their actual level of knowledge.



(a) Plot for the lower 50% of the class.

(b) Plot for the top 50% of the class.

Figure 8-24 Time spent on activity vs learning gain g plots separated by class rank.

In this research, we relied our results on the deep pedagogical expertise Prof. Belcher provided us. Using his well test of pedagogical approach gives us confidence that the data collected is properly grounded as we measure the effect of the augmented simulation on the student's learning gain.

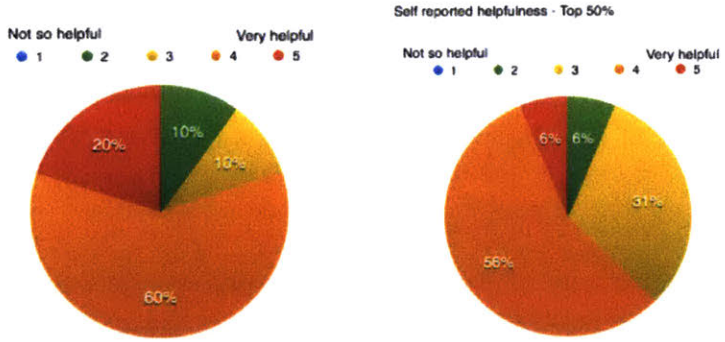
Self-Reported Helpfulness

When examining the online survey result for self-reported helpfulness we find consist results. 87% of the participants answered the survey, indicating a range between '1' (Not so helpful) and '5' (Very helpful) to the question whether the augmented simulation was helpful to understand the concept of magnetic fields.

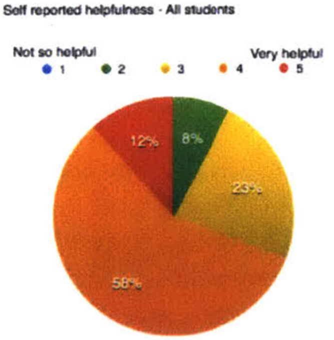
20% of the lower performing students found the augmented simulation very helpful (5) compared to only 6% for the higher performing students. This trend is consistent when we extrapolate across the survey range.

Overall, the results suggest that augmented simulation has the potential to be helpful to those students that do less well with the current method of traditional computer simulation. This finding is also consistent with the pedagogical method research that suggests that students prefer a more "hands on" or active approach to learning by doing as opposed the "abstract, written" methods. For further reading and expansion on this body of knowledge, we refer the reader to Kolb **[258]**, **[259]**, Gregorc **[260]**, and Hawk **[261]**, who have written extensively on the topic.

As mentioned above, the TEAL classroom method of instruction provided a natural baseline or an inherent control baseline to compare the traditional technology-driven education method with the proposed augmented simulation. The finding of the cognitive model reported here can certainly motivate future exploration to expand the finding to crystalize what specifically helps students understand abstract concepts when presented with augmented simulation.



(a) Plot for the lower 50% of the class. (b) Plot for the top 50% of the class.



(c) Plot for all students of the class.

Figure 8-25 Self-reported helpfulness of object augmentation on learning

User Experience

Our user experience evaluation was designed to validate Enlight as a viable projected augmented reality interface situated in a real-world classroom environment.

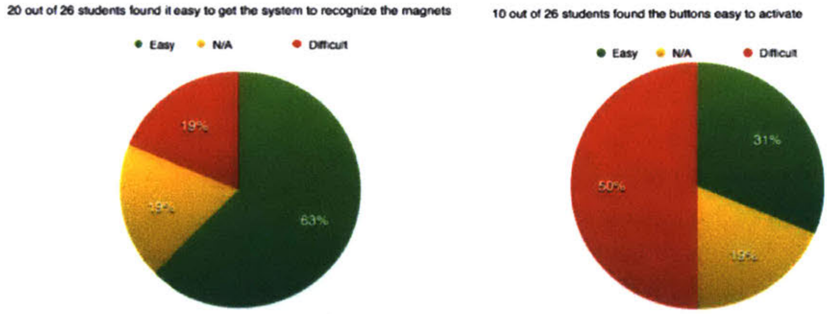
To that end, we focused the measurements on two key usability factors:

1. Perceived ease of use of the system
2. Ease of use for first-time users of projected augmented reality interface.

Our online survey presented participant with the following questions:

1. "Did you find the buttons easy to activate?" (Yes/No)
2. "Did you find getting the system to recognize the magnets easy?" (Yes/No)

The results are reported in Figure 8-26 below:



(a) Object recognition.

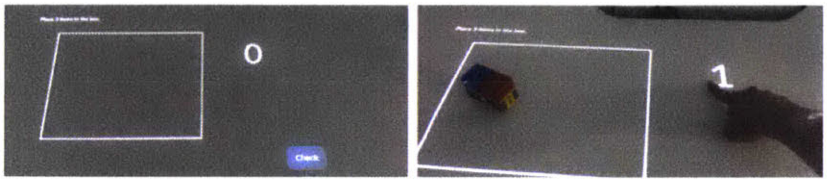
(b) Projected buttons.

Figure 8-26 Self-reported ease of use of the interface.

63% of the participants reported it was easy for them to get the system to recognize the simulation objects (magnets). However, the projected buttons were perceived as difficult to operate. This result is likely exposing the drawbacks of the current implementation sensing capabilities.

Another interesting finding exposed the effect of the common use of touch screens on the initial learning of an augmented reality interface.

The Enlight training application prompted users to place three physical objects in a projected box (Figure 8-27). The surprising phenomena we found was that most of the participants attempted to drag projected virtual objects to the screen when instructed to put an object in the virtual projection area. In other words, the participants completely ignored the real world and assumed the interface and the instructions expected them to drag objects as they do normally with their smartphone screen.



(a) (b)

Figure 8-27 User training: The step that requests the participant to place three items in the white box projected. (left) The behavior demonstrated by some participants attempting to drag a virtual object into the box. (right)

(a) (b)

Figure 8-28 User behavior with virtual and physical objects. (left) User behavior for participants who attempted to drag on screen objects. (right)

After analyzing the video footage from the experiment, we found that 60% of the participants exhibited this behavior. Out of this set, 50% of the participants eventually recognized their error and reached to place physical objects in the projected box. This result yields that over 60% of the study population were able—without the intervention of the researchers—to adapt to the physical-digital experience provided by the Enlight projected augmented reality interface. This suggests that AR interfaces in general, and specifically projected ones, are indeed intuitive and become a viable modality if developed further. In parallel, this finding also unearthed the strong bias effect that prevailing user interfaces have on users as they encounter new experiences.

Student Preferences

In the final section of the results discussion is centered on student preferences. We rely in our analysis on the work of Dr. Katherine McKnight's staff at George Mason University, who annotated the transcriptions of the experiment's video recording. Their work provided an objective frame of reference to the abundant qualitative data we collected. In the sections that follow we will focus on common themes we discovered.

Annotation Methodology

The annotations were divided to groups of five overlapping transcripts. The annotators reviewed the raw data and decided on the following common high-level themes:

- Conceptual understandings of electromagnetism as observed during demo.
- Conceptual misunderstandings as observed during demo.
- Usability of system (Affirmative).
- Usability of system (Negative).
- Students call for an app of demo (asking for help).

The themes above break down to several more specific sub-themes. In the annotations analysis process, we only account for the absence or presence of comments and not the accurate counts of comments by the users, as some participants have repeated a comment with the same notion or feedback multiple times. The final data set was checked to ensure consensus between the two annotators.

Physical Objects, Projected Interfaces

Participants certainly enjoyed the overall experience of the projected augmented reality user interface. Over 60% made positive statements on interacting with the physical objects and the Enlight interface.

> ***"There is something so much more fun about playing with physical legos than with using a mouse to control little points on a screen."***

> ***I think this is really similar, but it is never quite the same moving around on the screen. I found it easier to just move things around then doing something on a screen. At least with a tablet, versus a computer screen, you are moving things around, but it is not the same as picking it up and moving it."***

The participants also responded to the inclusion of real, functioning objects (e.g., magnets, compass). Nine students made positive comments on the haptic feedback provided by the magnets, and 15 students responded positively to the use of a functioning compass.

> ***"I guess this is more engaging because it allows you to do more hands-on type activities. Also because these are actual magnets, i.e. if you bring them close together, you can actually feel force, you can feel them repelling or attracting. That is good because you see the arrows that tell you they are going to attract and as you bring them closer, you feel them going to attract. That is something you cannot get on a computer, i.e. that these two things are definitely pushing each other apart."***

> ***"There is just a sense that it is a real thing and that you are seeing, where the projection is showing what is underlying the nature of it."***

> ***"Because it is real and you are actually doing a lot of the stuff we do in physics they kind of just tell you about, and you believe them. When you get to do it with your hands, it very much reinforces the fact that these are real forces, that dipoles are not some abstract thing, vectors and everything, the cross product - that they are things in nature that we have discovered, not just made up."***

> ***"Using the compass with the line is a good idea. It kind of helps you get a more accurate depiction of what is going on...because you can see it falling along the contour lines and actually orienting itself, which kind of adds a real thing to the concept."***

The participants also appreciated the projected information included as part of the augmented simulation. Almost half (47%) asserted that the simulated projected magnetic lines helped them understand the invisible, unseen aspect of this real-world, physical phenomena.

> ***"It was just cool to see real field lines from magnets. It is kind of an abstract concept, and seeing it right in front of you makes it more clear."***

> ***"I think that some of the problems with E&M versus classical mechanics is that you see classical mechanics as it happens. E&M you cannot see it, so, being able to put a magnet on the table and then all the lines—that was useful. It***

> ***was also a lot of problems that other digital interfaces have is that it kind of all goes in one direction so you are still manipulating digital objects—there is no real physical analog. So the idea of having physical systems is good, especially having a physical compass, instead of a virtual compass, that points...a physical compass seems more real. I can make a digital compass point whatever with a few lines and a GL, but the physical compass tied kind of tied the concepts better."***

> ***"This kind of system would be more useful where there is actual physical phenomena that you are modeling...where it is almost entirely conceptual, because (xx unclear from audio) you have an actual object that you are working with the digital interface so you kind of need a class where you have actual constructs."***

> ***"It was cool to see exactly the force where it is pointing. The magnet still does the same thing along the lines, but to see that the green line is showing that the force...It was nice, later, having the force vectors also on top, so you could see how they all related to each other...."***

Errors, Challenges, Drawbacks

At the time of the experiment the Enlight system was a mature prototype. However, as with any nascent technology, it suffered from bugs and limitations. We had to deal with several issues that impacted the usability of the system and our research results. The sources of errors we dealt with stemmed from:

- Failures of our detection system.
- Failures to calibrate and/or recalibrate the system pro-cam setup.
- Responsiveness of our GUI in comparison to state of the art touch screens.

> ***"Something that is bugging me is that whenever I move it, it disappears for several moments until I move my hand."***

> ***"I wish it was a little more responsive, but I guess that is because I am spoiled: I am use to using the touch screen on my iPhone or on computers."***

> ***"The sensitivity is a little bit. Sometimes I have to twice— I don't know if that is going to get in the way."***

> ***"It would be useful if there was another arrow showing the torque on the dipole at the same time, but, they are kind of separate ideas so maybe that would be something you could add on at the end, as a Part 5, where you could show the torque and force together."***

> ***"I have one thing that I am noticing is that sometimes the hyper field lines are broken up so they are not connected for me. I guess it would help to have them connected so I can try to move the compass around field lines to be able to trace it exactly on the dipole."***

> ***"I think it could be helpful if you could move them and see how it (force) is changing as you move them."***

> ***"Something that may help is being able to recognize whether it is a hand or an object because sometimes when I put my hand it would think it was a magnet but it was just my hand."***

> ***"it looks to me, that the calibration of the fields is just slightly askew. It still gets the point across, it's just not perfect—it's hard to get it perfect. It makes sense though."***

Online Survey versus Think-aloud

Overall, the annotation results seem to be consistent with the results of the online survey. The common themes we distilled from the think-aloud protocol, and described above, were also amplified by the results of the online survey.

In the online survey participants were asked to comment on their preferences with relation to the augmented reality simulation provided by the system. Specifically, we asked two important open ended questions:

1. "Would you prefer this system over visualizations on tablets/laptops/desktops?"
2. "What will make the system more enjoyable?"

For the purpose of this analysis, the annotators classified a positive response only if the response included a specific 'Yes'. All other responses were sub-classified (Figure 8-29). Overall, more than 50% prefer the augmented simulation. However, many of the students who prefer the augmented experience require it to dramatically improve. This is understood given the nascent state of the technology we tested.

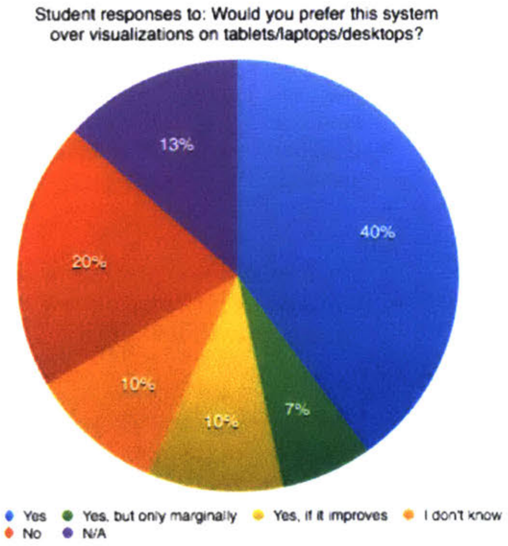


Figure 8-29 Student preferences based on survey responses: interface comparison.

Our annotators classified several specific areas that require technical and interface improvements to address the second question. Figure 8-30 shows the aggregated results. We also include examples of the suggestions we got from participants below.

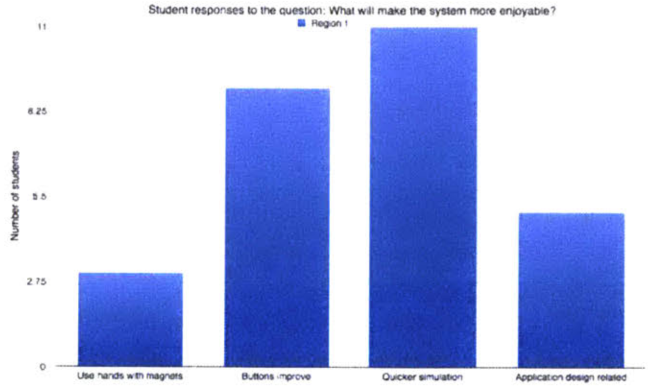


Figure 8-30 Student preferences based on survey responses: areas for improvement.

"Use of hands with magnets" (3 students) refers to responses that wished for the ability to keep their hands on top of the magnets while the visualization for the fields lines is updated. With the current implementation, users have to remove their hands from the magnets for approximately 1 second before the fields lines are calculated and shown.

"Buttons improve" (9 students) refers to responses that wish for projected buttons that are more responsive and predictable.

"Quicker simulation" (11 students) refers to responses that wish for a more responsive simulation. This is easily rectifiable with the use of a more powerful computer than the current single board computer on the LuminAR system.

"Application design related" (5 students) refers to responses that suggest improvements that are related to the application design like "the color of the arrows projected should be red" or "I'd like to be able to go back to a previous part" etc.

8.9 Discussion and Contributions

Enlight was the final project in the LuminAR family of projects. It provided a great opportunity to integrate all the knowledge we have accumulated over the years trying to build viable projector demented reality interfaces. Through this project we have demonstrated the real-world application of AR interfaces in science education. We have demonstrated how the future of active learning experiences might look like using augmented simulations.

We created hardware, software, and interaction techniques, as well as content that paves the way for continued work that can help transform the educational medium down the road, and help prepare for *reality* when augmented reality is widely available. We did so using contemporary classroom furniture and undergraduate level physics education contents from one of the best teachers on the planet. These choices made us produce much more than a lab experiment.

We also tested and analyzed the efficacy of our solution with a cohort of students who provided inputs as we were testing our hypothesis. Though preliminary, our qualitative and quantitative results suggest that augmented simulation may provide enhanced means to teach students who require alternative learning experiences to reach the same level of performance as their high-performance peers.

Finally, we have shown that users who never experienced augmented reality before can very quickly adapt to this new type of interface. Users of the system could naturally combine physical objects that are familiar and common in the physical world, and combine them with routine use of superimposed digital information, using a device with which they had no experience. This combination of digital-physical experience is an exciting path for us to follow.

9 Conclusion and Future Work

This dissertation focused on the design, engineering, and interaction techniques for a compact and projected augmented reality interface. LuminAR challenges existing computer-interface paradigms and form factors, offering a novel AR user experience that combines digital interaction in a physical workspace.

This chapter provides a high-level summary of the shared learning we collected as we deployed projected AR interface prototypes in real-world scenarios. The learning synthesizes field work with formal experiments. As a result, we can suggest a unique generalized perspective, that can inform new application domains we did not yet explore. Next, I outline the major technical trends that I believe will govern the domain in the decade to come, attempting to distill a vision for the near-term evolution of projected AR interfaces.

We conclude this chapter with concrete research directions that can promote the field further, and a summary of the direct contribution of this research.

9.1 Learning from Real World Deployments

It has been an extraordinary journey to try and advance the state of the art of projected AR interfaces. In the course of this research I explored six distinct application domains and watched numerous users play around with projected interfaces. In this section, I will summarize some of my own learnings from these multiple observations, focusing on what would make projected interfaces a tangible part of our interactive future:

9.1.1 Robust Interfaces

Users have very low tolerance to clunky interfaces. We have all had the experience of trying a new consumer device and having a bad experience. Robust interfaces are those that reach a certain level of technology and usability, interfaces that can provide a consistent, learnable and pleasant user experience. Simply put, the interface works as promised. Projected AR interfaces did not yet reach this level of robustness and maturity. We have discovered again and again that deploying such interfaces is not a simple task. It requires understanding of the target augmentation environment and light conditions, as well as making inventories of objects to be recognized and mapping the specific user interactions required. My early intuitions were that for projected AR to be a viable medium, it would require what we take for granted when interacting with PC or smartphone applications.

Specifically, hardware should be installed in minutes, with minimal interaction from the user. In fact, the user should be shielded from any calibration or custom setup routines. That is not the case today. This is the reason why I pursued the design of new hardware form factors that work out of the box.

9.1.2 Light is Magic

Once a projected AR is installed properly in a physical space, people cannot ignore it, at least not now, as it is not as common as public displays. People respond to light. Light is playful and inviting. I have personally observed numerous users become fascinated with the idea and the experience interacting with light.

Humans have a peripheral vision that is tuned very well to processing light events. Boyce describes the human vison systems in his great book Human Factors in Lighting **[262]**. Humans respond well to light. We all understand the idea of illumination, shadows and projections. Therefore, when photons become interactive and provide information, putting novelty aside, most people I have seen become extremely curious and begin to play with the interface.

One concrete example of that was the Boids particles simulation **[263]** application (see Figure 9-1). Boids is a flocking behavior simulation. Typically, the input to this simulation is a click on a computer screen. In our version, the Boids flock to the object placed under the LuminAR projector or to the user hands in a mid-air gesture.

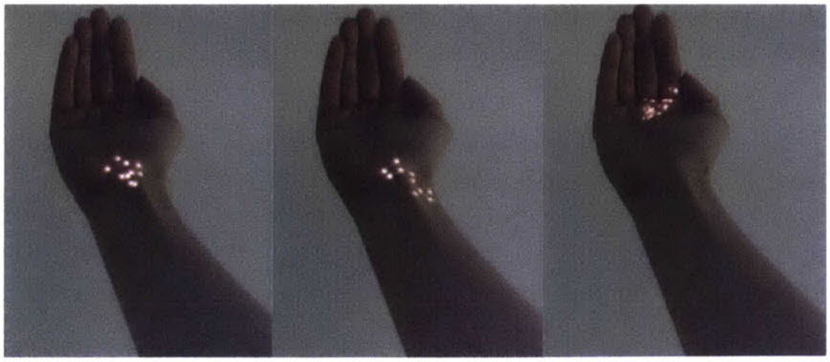


Figure 9-1 Boids application. The emergent behavior of the Boids flock is following the user's hand in a mid-air gesture.

The hands can freely move in 3D under the LuminAR projector. This application demonstrates a playful and instantly learnable light-based experience. My prediction is that this 'magical' property of light-based interface can contribute to the mass adoption of such interfaces.

9.1.3 Learnable Interfaces

Much like the Boids experience, we have found that new users who are exposed for the first time to projected AR interfaces can learn to use the interface very quickly. Our findings are reported in the formal experiments in the Manufacturing and Education domains in Chapters Seven and Eight, respectively.

In the past decade, smartphones have become a life necessity. This fact made touch screen interactions entirely common and expected by users. The results we have seen in our augmented simulation work seem to indicate that touch screen interaction makes any projected interface easier to learn, as the user expects to touch the interface and for interaction to occur. This combination would make projected AR interfaces highly learnable.

9.1.4 Complex Augmentation

This insight led me to focus on the idea behind Dynamic Multi Touch presented in Chapter Five. My work extends classic multi-touch interactions to the physical domain, where objects and users co-exist in a 3D space.

However, we did not yet handle a key requirement to support real-time, geometrically registered 3D texture mapping with an adaptive background model compensation. This is required to support projection on complex surfaces in arbitrary environments. Once implemented, it will support the painting of objects with light directly while tracking the objects and the users. I discuss concrete future research path that addresses this approach in the next section.

To fully support such complex augmentations, there is an increasing need to perform a myriad of computer vision-based tasks such as object detection and tracking, gesture detection as well as automatic and self-healing calibration routines. Latency can kill any interface, and in particular any computer vision-based interface. It is therefore clear to me that one of the key technologies that would help any AR interface to support real-time interaction is hardware-accelerated computer vision. This trend is already a reality, a prominent example of that is the CEVA-XM4 computer vision hardware created by CEVA to support natural interaction, AR and various other embedded machine vision applications **[264]**.

9.1.5 Augmented Guidance

Finally, this works demonstrated the findings of previous AR researchers **[182]**, **[183]**, **[189]**, **[265]** proving that guidance provided to humans using projected AR has meaningful, positive impact on learning. This was demonstrated both in the manufacturing and education domains we heavily explored. In the manufacturing domain, we measured the increase in the work instructions knowledge (see Chapter Seven). In the education work, we saw students respond positively to augmented simulation and found correlation between the time they spent using the augmented simulation and theory learning gain (see Chapter Eight). I am convinced that our findings can span to other application domains.

9.2 The Future of Projected AR Interfaces

This thesis is written in 2017, a decade after the introduction of the iPhone, the canonical smartphone. In this period of time, consumer electronics productizing cycles have grown shorter. We know how to design and ship hardware products very fast. The proliferation of smartphones has also reduced the cost and simplified the supply chain required to design new hardware. This phenomenon also impacted the AR product development. In Chapter 2, we covered some of the new product introduced by major consumer electronics companies.

9.2.1 Common AR Bulb

However, we are still far from the vision we described in this thesis, of a standard, affordable and highly functional AR bulb that is available as a common interface, naturally installed in our existing architectural spaces. My collaborator and good friend, Yoav Reches, captured this vision in the rendering he created in Figure 9-2.

I believe that the inflection point will occur when such a common AR bulb is available. Currently, the existing products are good as proxies for the real impactful form factor. The bulb is a natural form factor to carry a projected interface. This idea is a repeating theme in the past 30 years of research, as we discussed in detail in Chapter 2. The lack of simple form factors and authoring tools inhibit the mass adoption of augmented reality as a mainstream user interface modality.

Figure 9-2 LuminAR Bulb Concept (Source: Yoav Reches)

The simplicity of the bulb as a form factor can drive users and developers to invent new types of applications that currently cannot be built or distributed. However, just having a common form factor is not enough, it has to be coupled with authoring tools that allow everyone to develop content. I describe a future path for such tools in the section that follows.

9.2.2 A World with More Displays

Projected AR interfaces are not going to replace our traditional active flat displays. We are a display-centric society. In the past few years, more flat panel displays have been sold compared to desktop systems, and interactive panel sales have surpassed one million annual units **[266]**. This trend is not slowing down. Display technologies are moving fast with the introduction of OLED and transparent displays. It is safe to predict that screens will become ever more integrated into our electronics and architectural spaces. This raises the question how projected AR interfaces will interplay in this massive display-driven future. I think the answer is quite simple. Projected AR is just another type of display that has very specific capabilities. With time, the projection engines will catch up to certain performance specs of the active displays, however active displays, will forever have pixels contained in a rigid form. This will be the key driver for adoption. There are use cases that are simply unavailable for standard display. We covered the advantages of light-based interfaces in Chapter One. As a recap, the main difference we are discussing is the ability to place pixels on physical objects in real time and by that extend functionally of physical objects with digital interfaces.

I believe that when AR bulbs become common, they will use the same internet-driven computing substrate and as a result they will become integrated to spaces that are rich with displays. In many ways, the internet has already become the de-facto operating system that synchronizes all of our various displays, so this is another somewhat safe prediction.

9.2.3 Augmented Reality Interface Integration

As we discussed in Chapters One and Two, projected AR is following the evolution of wearable AR interfaces. Currently, wearable AR and VR interfaces are maturing faster than projected AR interfaces. Other forms of handheld AR and contact-lens driven AR are also under development. It is logical that projected AR interfaces will interplay with those interfaces as they become available.

There are many possible scenarios we can imagine. I am particularly interested in the shared nature of projected AR when compared to the wearable AR interfaces, which is an inherently private interface. The combination of the two can create an environment for interesting collaboration. Of course, the main advantage of the wearable HMD AR interface is the ability to use head-tracking to provide an immersive 3D object augmentation, an effect that is impossible to achieve with a simple projected AR interface. It is plausible that the combination can be a very powerful shared 3D experience.

It is also clear to me that other forms of natural interfaces will eventually become integrated with projected AR. Devices like Alexa [267] and Google Home [268] provide easy means to use natural language as interface, and they can be easily integrated today to any AR interface with a microphone. My view is that with time, voice-driven interaction modality will become a vital and expected component of any computing experience, and especially if it is an architecturally embedded computing experience, as the projected AR interfaces this work presents.

9.3 Research Directions

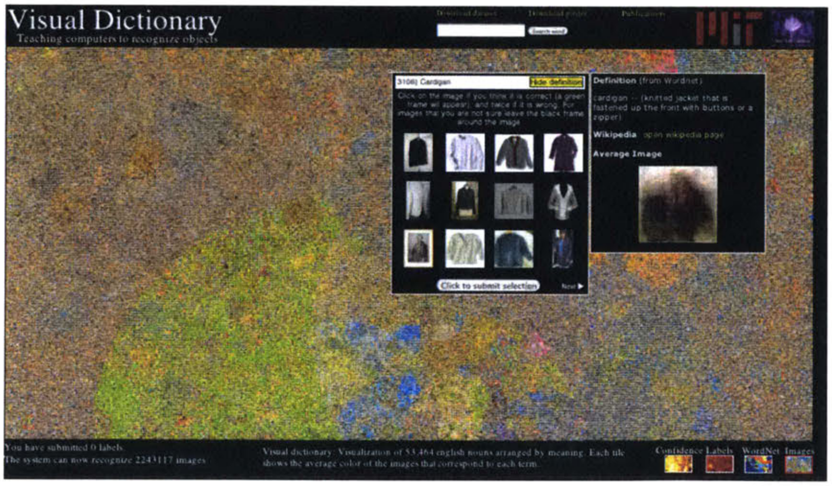
In the first two sections of this chapter, I presented the common learnings from our real-world deployments of projected AR interfaces and the major future technology trends that will impact the evolution of such interfaces. In this section, I propose three concrete future research directions that can help promote this evolution.

9.3.1 Object Augmentation and Manipulation

For projected augmented reality interfaces to become widely adopted, it is necessary to enable them with robust capabilities for object detection and tracking. Creating a framework for general-purpose object recognition is difficult and very much an open problem within the computer vision research communities. From an interaction point of view, such capabilities are important as they enable the design of user scenarios that include object augmentation.

Kinetic interfaces can potentially assist with solving aspects of this problem, utilizing the ability to dynamically position a sensor's point of view to improve object detection and registration into an interactive scene. To achieve that, computer vision engines should be extended with software interfaces that can make perspective change requests in real time to the motor control system. In doing so, and considering the already existing and known geometric model, the system may gain superior object recognition and tracking capabilities. Therefore, future systems could also benefit from the kinetic motion capabilities to dynamically augment objects with projected information.

From a hardware perspective, it is also possible to better integrate the sensor inputs in the bulb, namely combining depth sensor frames with standard camera frames to improve detection. In addition, automatic compensation for ambient lighting conditions should be added and used in real-time fashion while the system is moving from one location to another.



Click on the image if you think it is correct (a green frame will appear), and twice if it is wrong. For images that you are not sure leave the black frame around the image

cardigan -- (knitted jacket that is fastened up the front with buttons or a zipper)

You have submitted 0 labels.
The system can now recognize 2243117 images

Visual dictionary: Visualization of 53,464 english nouns arranged by meaning. Each tile shows the average color of the images that correspond to each term.

Figure 9-3 MIT CSAIL Visual Dictionary by Antonio Torralba et al (Source: CSAIL website)

Finally, this approach should also consider the advances made in the field of computer vision cloud-based datasets. Projects like Visual

Dictionary **[269]** propose new mechanisms to use billions of images to solve the general problem of recognizing all different classes of objects in the real world. This approach, combined with an interface like LuminAR, can provide a huge leap for projected augmented reality, providing AR interfaces with the ability to create close to real-time, cloud-based, visual based queries. The result of such queries will allow AR interfaces to truly respond to actions performed in the physical world by users.

9.3.2 Opportunistic Projection Surface Detection

Kinetic interfaces would be even more useful if they were enhanced with surface detection capabilities. During user interaction, such future systems would be able to adjust to the targeted projected area. This approach has advantages for adapting the system to a changing physical environment. This can be beneficial for desk workspaces that are usually cluttered. We call this approach "opportunistic location detection."

The algorithmic basis for implementing such features already exists. Known techniques like PTAM **[270]** can be used to efficiently identify surface candidates for projection. The real challenges lie in the interaction design required to create a valid user experience. Users' fault tolerance for kinematic systems is very embryonic, and perhaps Hoffman's work on fluency can serve as a starting point **[58]**.

9.3.3 Augmented Reality Interface Design Tools

If we follow the trajectory of Moore's law, it is clear that at some point, projection and sensing hardware capabilities integrated with powerful microcontrollers will enable designers to build rich augmented reality interfaces.

In Chapters Six, Seven and Eight, we have shown the power of a generic web-based AR authoring tool, that allowed us to create many applications in different domains using the same system. Here we highlight three additional important aspects and exemplify them. First, the bar for content creation should be low. In our case, Lens (see Chapter Four) only assumes basic web-development skills. In Figure 9-4, I present Shou Yan's mixed reality bookshelf project, built with LuminAR and Lens. Shou is a great designer and at the time we met she had very basic web development skills. She quickly picked up the Lens APIs and created a neat experience. Shou did not need to have any knowledge about the underlying AR or computer vision algorithms that actually enabled her application. That is a principle that all future AR content creation tools should adhere to, and strive to implement.

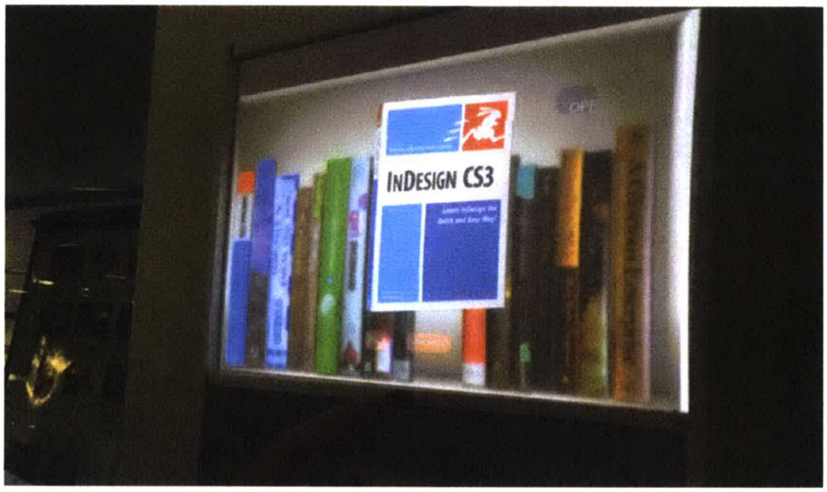


Figure 9-4 Shou Yan's interactive mixed-reality bookshelf

Second, any content framework must be extensible and allow developers to improve it via API. Our platform, Lens, was designed as such. For example, Walter Menéndez created an extension library to Lens that allowed adding interactive shadows to objects (see Figure 9-5). This library used the core AR and computer vision services provided by Lens (e.g. object and color detection).

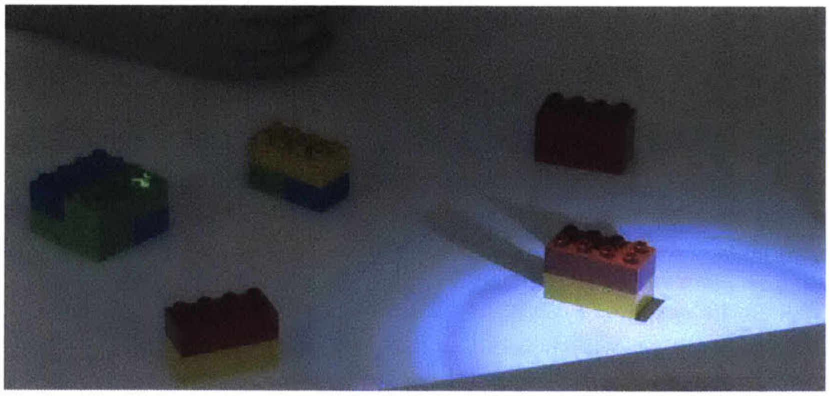


Figure 9-5 Walter Menéndez's interactive simulated shadow Lens extension library

He also used standard Javascript user interface libraries to create a new API that allowed to attach a simulated shadow to objects detected by LuminAR. The significance for AR content authoring is critical. All future AR content systems must be open to leverage the functionality-rich web-based application development environments provide.

Finally, current user interface design technologies have clear drawbacks when applied to the design of augmented reality interfaces. Current Digital Content Creation (DCCs) tools and Integrated Development Environments (IDEs) are completely geared toward the development of display-centric interfaces.

It is not hard to imagine how we can redesign such tools so they fit new interaction paradigms specific to augmented reality. I believe the gap is greater when it comes to the support of dynamic projected content and kinematic systems. Such tools should support the concept of Just-in-Time-and-Place Interfaces and augmented reality affordances. For example, development tools today are unable to model complex application flows that take into account projection location, size, object detection, and tracking. The underlying computational models for such systems have been explored before, for example Roy [57].

I also believe that future AR application development environments will use authoring-by-demonstration techniques, allowing users to simple show the projected AR system a new object, and by using voice command and gestures will create a new experience.

9.4 Contributions

LuminAR is inspired by more than 30 years of active research in the domains of interactive spaces, augmented reality, computer vision, and personal robotics. We covered the key works in these domains in Chapter Two. The research approach for the work presented here integrated these domains with a design-driven development process. This approach motivated several prototype iterations described in Chapter Three. The family of LuminAR devices developed in the course of this dissertation work represents the result of this exploration.

The LuminAR Bulb, LuminAR Lamp, and Spotlight represent original hardware designs for a compact and kinesthetic interactive projector-camera system. Product design and industrial design played key roles and informed the development of LuminAR prototypes. The result pushes the boundaries of embedding computation in everyday objects, making interfaces to digital information truly embedded in our environment. This is particularly relevant to the LuminAR Bulb, which can be simply screwed into a standard Edison socket, or to the LuminAR Lamp which can replace an existing Anglepoise lamp on your desk.

We also demonstrated that LuminAR devices are practical for real-world scenarios. The applications we developed provided a glimpse into the future of projected reality interfaces in the personal desktop workspace and retail domains. Even though our demos were domain specific, we can already see merit in exploring LuminAR applications for new domains such as education, medical applications, command centers, and many more.

The interaction techniques we described in Chapter Five provide a glimpse into a future of kinetic interfaces and contribute directly to the evolution of natural augmented reality interfaces. In this work, we propose new interaction concepts enabled by LuminAR, namely Dynamic Multi-Touch and Just-in-Time-and-Place Interactions. Our initial results and feedback, as captured in the demos of the applications we created, show the potential of such interfaces, but leave an open door for additional work that will explore further augmented reality as a mainstream user interface modality. Specifically, kinematic interfaces enabled with kinetic I/O are still very much in the future and relatively unexplored. I believe that 'interfaces that can move' hold a great potential for the development of the field of human-computer interaction, but require further investigation and an abundance of design efforts. It is also important to conduct further formal evaluation of the interaction techniques we proposed.

Finally, through this work we have demonstrated the viability of projected augmented reality interfaces. We have deployed such systems and applications in real-world scenarios and environments, across the domains of retail, collaboration education and manufacturing. We have also exposed a great number of users in non-formal and formal settings, and shown how quickly those users are able to use our system, suggesting that projected interfaces are close to becoming a common reality.

Through our work, we have also shown that developers can easily learn to develop projected augmented reality applications using our software framework without any deep prior knowledge of AR or computer vision algorithms.

Even though we have made great advances in the current state of the art projector-sensor systems and their accompanying software frameworks, we are clearly just scratching the surface of what such interfaces can become in the future. Moreover, I believe that adding light-emitting devices as a first-class interface medium will become commonplace in our digital-physical reality. It is my belief and hope that the work done here can inspire and promote this future.

10 Bibliography

[1] L. F. Baum, The Master Key: An Electrical Fairy Tale Founded Upon the Mysteries of Electricity and the Optimism of Its Devotees. It was Written for Boys, But Others May Read it. Bowen-Merrill Company, 1901.

[2] M. Weiser, "The Computer for the 21st Century," *Scientific American*, vol. 265, pp. 94–104, 1991.

[3] I. E. Sutherland, "A head-mounted three dimensional display," December 9-11, 1968, fall joint computer conference, part I, New York, New York, USA, 1968, pp. 757–764.

[4] I. E. Sutherland, "The Ultimate Display," Proceedings of the IFIP Congress, 1965, pp. 506–508.

[5] M. W. Krueger, T. Gionfriddo, and K. Hinrichsen, "VIDEOPLACE -- an artificial reality," *SIGCHI Bull.*, vol. 16, pp. 35–40, Apr. 1985.

[6] Feng Zhou, H. B.-L. Duh, and M. Billinghurst, "Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR," Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR), 2008, pp. 193–202.

[7] T. P. Caudell and D. W. Mizell, "Augmented reality: an application of heads-up display technology to manual manufacturing processes," Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences, 2004, vol. ii, pp. 659–669 vol.2.

[8] P. Milgram and F. Kishino, "A taxonomy of mixed reality visual displays," *IEICE TRANSACTIONS on Information and Systems*, vol. 77, no. 12, pp. 1321–1329, 1994.

[9] Ronald T. Azuma. 1997. A survey of augmented reality. *Presence: Teleoper. Virtual Environ.* 6, 4 (August 1997), 355-385.

[10] Oliver Bimber and Ramesh Raskar. 2005. *Spatial Augmented Reality: Merging Real and Virtual Worlds*. A. K. Peters, Ltd., Natick, MA, USA.

[11] J. Patten, H. Ishii, J. Hines, and G. Pangaro, “Sensetable,” the SIGCHI conference, New York, New York, USA, 2001, pp. 253–260.

[12] O. Cakmakci and J. Rolland, “Head-Worn Displays: A Review,” *J. Display Technol.*, vol. 2, no. 3, pp. 199–216, Sep. 2006.

[13] N. Zerkin, “Is Google Glass an Augmented Reality Device?.” [Online]. Available: http://blog.integratedrealities.com/?p=261. [Accessed: 22-Jul-2017].

[14] Microsoft, “Microsoft - HoloLens.” [Online]. Available: https://www.microsoft.com/en-us/hololens. [Accessed: 22-Jul-2017].

[15] Meta, “Meta | Augmented Reality.” [Online]. Available: https://www.metavision.com/. [Accessed: 29-Jul-2017].

[16] Ramesh Raskar, Greg Welch, and Henry Fuchs. 1999. Spatially augmented reality. In *Proceedings of the international workshop on Augmented reality : placing artificial objects in real scenes: placing artificial objects in real scenes* (IWAR '98), Reinhold Behringer, Gudrun Klinker, and David W. Mizell (Eds.). A. K. Peters, Ltd., Natick, MA, USA, 63-72.

[17] E. Kruijff, J. E. Swan, and S. Feiner, “Perceptual issues in augmented reality revisited,” 2010 9th IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 2010, pp. 3–12.

[18] Apple, “ARKit.” [Online]. Available: https://developer.apple.com/documentation/arkit. [Accessed: 28-Jul-2017].

[19] PTC, “Vuforia.” [Online]. Available: https://developer.vuforia.com/. [Accessed: 29-Jul-2017].

[20] Google, “Google Glass .” [Online]. Available: https://x.company/glass/. [Accessed: 22-Jul-2017].

[21] P. Mistry and P. Maes, “SixthSense: a wearable gestural interface,” ACM SIGGRAPH ASIA 2009 Sketches, New York, NY, USA, 2009, pp. 11:1–11:1.

[22] W. Newman and P. Wellner, “A desk supporting computer-based interaction with paper documents,” the SIGCHI conference, New York, New York, USA, 1992, pp. 587–592.

[23] Pierre Wellner. 1993. Interacting with paper on the DigitalDesk. *Commun. ACM* 36, 7 (July 1993), 87-96.

[24] A. D. Edgar, "Apparatus and method for remote displaying and sensing of information using shadow parallax," US4468694, 1984.

[25] P. Wellner, "The DigitalDesk Calculator: Tangible Manipulation on a Desk Top Display," Proceedings of the 4th Annual ACM Symposium on User Interface Software and Technology, New York, NY, USA, 1991, pp. 27–33.

[26] R. Raskar, J. van Baar, P. Beardsley, T. Willwacher, S. Rao, and C. Forlines, "iLamps," ACM SIGGRAPH 2003 Papers, New York, New York, USA, 2003, pp. 809–818.

[27] AccuVein, "AV400 Vein Viewing System | AccuVein."

[28] *christiemed.com*. [Online]. Available: https://www.christiemed.com. [Accessed: 20-Aug-2017].

[29] P. M. Central, "Projection Mapping Central Community," Dec. 2013. [Online]. Available: http://projection-mapping.org/. [Accessed: 20-Aug-2017].

[30] B. Sparrow and L. Chatman, "Social cognition in the Internet age: Same as it ever was?," *Psychological Inquiry*, vol. 24, no. 4, pp. 273–292, 2013.

[31] B. Sparrow and L. Chatman, "We're Not Burning Down the House: Synthesizing Pre-Internet, Current Findings, and Future Research on Social Cognition and Being Online," *Psychological Inquiry*, vol. 24, no. 4, pp. 349–355, 2013.

[32] M. Weiser, "Hot topics-ubiquitous computing," *Computer*, vol. 26, no. 10, pp. 71–72, Oct. 1993.

[33] D. A. Norman and S. W. Draper, *User Centered System Design; New Perspectives on Human-Computer Interaction*. Hillsdale, NJ, USA: L. Erlbaum Associates Inc., 1986.

[34] D. A. Norman, *The Psychology of Everyday Things*. Basic Books, 1988.

[35] Y. Rogers, K. Connelly, L. Tedesco, W. Hazlewood, A. Kurtz, R. E. Hall, J. Hursey, and T. Toscos, "Why It's Worth the Hassle: The Value of In-Situ Studies When Designing Ubicomp," in *UbiComp 2007: Ubiquitous Computing*, vol. 4717, no. 20, Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 336–353.

[36] Y. Rogers, "Interaction design gone wild," *interactions*, vol. 18, no. 4, pp. 58–62, Jul. 2011.

[37] A. Crabtree, A. Chamberlain, R. E. Grinter, M. Jones, T. Rodden, and Y. Rogers, Eds., "Introduction to the Special Issue of 'The Turn to The Wild'," *ACM Trans. Comput.-Hum. Interact.*, vol. 20, no. 3, pp. 1–4, Jul. 2013.

[38] E. Hutchins, *Cognition in the Wild*. Cambridge, MA: MIT Press, 1995.

[39] P. Marshall, R. Morris, Y. Rogers, S. Kreitmayer, and M. Davies, "Rethinking 'multi-user'," the 2011 annual conference, New York, New York, USA, 2011, pp. 3033–3042.

[40] J. S. Underkoffler, "The I/O bulb and the luminous room," Massachusetts Institute of Technology, Cambridge MA, Jan. 1999.

[41] N. Linder, "LuminAR : a compact and kinetic projected augmented reality interface," Massachusetts Institute of Technology, Cambridge MA, 2011.

[42] N. Linder and P. Maes, "The design evolution of LuminAR," the 2012 ACM annual conference extended abstracts, New York, New York, USA, 2012, pp. 1435–1436.

[43] N. Linder, P. Maes, "Kinetic input/output," US Patent 8751049, Jun. 2014.

[44] N. Linder and P. Maes, "LuminAR: portable robotic augmented reality interface design and prototype," Adjunct proceedings of the 23nd annual ACM symposium on User interface software and technology, New York, NY, USA, 2010, pp. 395–396.

[45] T. Berners-Lee, R. Cailliau, and J.-F. Groff, "The world-wide web," Conference proceedings on 3rd joint European networking conference, New York, NY, USA, 1992, pp. 454–459.

[46] D. C. Engelbart and W. K. English, "A Research Center for Augmenting Human Intellect," vol. 0, p. 395, 1968.

[47] Bush, Vannevar (1945). "As We May Think," Atlantic Monthly 176 (July 1945) pp. 101-108.

[48] H. Ishii and B. Ullmer, "Tangible bits: towards seamless interfaces between people, bits and atoms," Proceedings of

the SIGCHI conference on Human factors in computing systems, New York, NY, USA, 1997, pp. 234–241.

[49] Ben Shneiderman and Pattie Maes. 1997. Direct manipulation vs. interface agents. *interactions* 4, 6 (November 1997), 42-61.

[50] Ronald Azuma, Yohan Baillot, Reinhold Behringer, Steven Feiner, Simon Julier, and Blair MacIntyre. 2001. Recent Advances in Augmented Reality. *IEEE Comput. Graph. Appl.* 21, 6 (November 2001), 34-47.

[51] Mario Martínez Zarzuela, Francisco J. Díaz Pernas, Leire Barroso Martínez, David González Ortega, Miriam Antón Rodríguez, Mobile Serious Game Using Augmented Reality for Supporting Children's Learning About Animals, Procedia Computer Science, Volume 25, 2013, Pages 375-381, ISSN 1877-0509.

[52] E. Redondo, D. Fonseca, A. Sánchez, and I. Navarro, “New Strategies Using Handheld Augmented Reality and Mobile Learning-teaching Methodologies, in Architecture and Building Engineering Degrees,” vol. 25, pp. 52–61, 2013.

[53] A. Bin Tomi and D. R. A. Rambli, “An Interactive Mobile Augmented Reality Magical Playbook: Learning Number with the Thirsty Crow,” vol. 25, pp. 123–130, 2013.

[54] T. P. Caudell and D. W. Mizell, “Augmented reality: An application of heads-up display technology to manual manufacturing processes,” System Sciences, 1992. Proceedings of the Twenty-Fifth Hawaii International Conference on, 1992, vol. 2, pp. 659–669.

[55] Wolfgang Friedrich. 2002. ARVIKA " Augmented Reality for Development, Production and Service. In *Proceedings of the 1st International Symposium on Mixed and Augmented Reality* (ISMAR '02). IEEE Computer Society, Washington, DC, USA, 3-.

[56] Van Krevelen, Rick & Poelman, Ronald. (2010). A Survey of Augmented Reality Technologies, Applications and Limitations. International Journal of Virtual Reality (ISSN 1081-1451). 9. 1.

[57] D. K. Roy, “Learning Words from Sights and Sounds: A Computational Model,” Massachusetts Institute of Technology, Cambridge, MA, USA, 1999.

[58] G. Hoffman and C. Breazeal, “Achieving fluency through perceptual-symbol practice in human-robot collaboration,” New York, NY, USA, 2008, pp. 1–8.

[59] Jun Rekimoto and Masanori Saitoh. 1999. Augmented surfaces: a spatially continuous work space for hybrid computing environments. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (CHI '99). ACM, New York, NY, USA, 378-385.

[60] H. Koike, Y. Sato, and Y. Kobayashi, “Integrating paper and digital information on EnhancedDesk: a method for realtime finger tracking on an augmented desk system,” *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 8, no. 4, pp. 307–322, Dec. 2001.

[61] J. Underkoffler, B. Ullmer, and H. Ishii, “Emancipated pixels,” the 26th annual conference, New York, New York, USA, 1999, pp. 385–392.

[62] J. Underkoffler and H. Ishii, “Illuminating light,” the SIGCHI conference, New York, New York, USA, 1998, pp. 542–549.

[63] J. Underkoffler and H. Ishii, “Urp,” the SIGCHI conference, New York, New York, USA, 1999, pp. 386–393.

[64] C. S. Pinhanez, “The Everywhere Displays Projector: A Device to Create Ubiquitous Graphical Interfaces,” Proceedings of the 3rd international conference on Ubiquitous Computing, London, UK, UK, 2001, pp. 315–331.

[65] Noi Sukaviriya, Rick Kjeldsen, Claudio Pinhanez, Lijun Tang, Anthony Levas, Gopal Pingali, and Mark Podlaseck. 2004. A portable system for anywhere interactions. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '04). ACM, New York, NY, USA, 789-790.

[66] M. Ashdown and Y. Sato, "Steerable Projector Calibration," *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops*, San Diego, CA, USA, 2005, pp. 98-98.

[67] Andrew D. Wilson. 2005. PlayAnywhere: a compact interactive tabletop projection-vision system. In *Proceedings of the 18th annual ACM symposium on User interface software and technology* (UIST '05). ACM, New York, NY, USA, 83-92.

[68] F. Kaplan, S. Do-Lenh, and P. Dillenbourg, “Docklamp: a portable projector-camera system.”, 2nd IEEE TableTop workshop, Nwport, RHODE ISLAND, USA, October 10-12 2007.

[69] Microsoft, “*Mobile Surface*” [Online]. Available: https://www.microsoft.com/en-us/research/project/mobile-surface/. [Accessed: 10-Jul-2011].

[70] Andrew D. Wilson and Hrvoje Benko. 2010. Combining multiple depth cameras and projectors for interactions on, above and between surfaces. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology* (UIST '10). ACM, New York, NY, USA, 273-282.

[71] R. Ziola, M. Kellar, and K. Inkpen, “DeskJockey: Exploiting Passive Surfaces to Display Peripheral Information,” in *Human-Computer Interaction -- INTERACT 2007: 11th IFIP TC 13 International Conference, Rio de Janeiro, Brazil, September 10-14, 2007, Proceedings, Part I*, C. Baranauskas, P. Palanque, J. Abascal, and S. D. J. Barbosa, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 447–460.

[72] S. K. Kane, D. Avrahami, J. O. Wobbrock, B. Harrison, A. D. Rea, M. Philipose, and A. LaMarca, “Bonfire,” the 22nd annual ACM symposium *on User interface software and technology* (UIST '09). ACM, New York, NY, USA, 129-138.

[73] Brygg Ullmer and Hiroshi Ishii. 1997. The metaDESK: models and prototypes for tangible user interfaces. In *Proceedings of the 10th annual ACM symposium on User interface software and technology* (UIST '97). ACM, New York, NY, USA, 223-232.

[74] Daniel Wigdor, Hao Jiang, Clifton Forlines, Michelle Borkin, and Chia Shen. 2009. WeSpace: the design development and deployment of a walk-up and share multi-surface visual collaboration system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '09). ACM, New York, NY, USA, 1237-1246.

[75] Chia Shen, “Multi-User Interface and Interactions on Direct-Touch Horizontal Surfaces: Collaborative Tabletop Research at MERL,” First IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP '06), Washington, DC, USA, 2006, pp. 53–54.

[76] S. Hunter, "MemTable-Contextual Memory in Group Workspaces," *MIT*, Jan. 2009.

[77] Z. Ryder, G. Shweta, L. Nate, F. James, and H. Beverly, "OASIS: Examining a Framework for Interacting with General-Purpose Object Recognition," *Intel Labs Seattle Technical Report, 2010*

[78] B. Tognazzini, "The Starfire video prototype project: a case history," Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence, New York, NY, USA, 1994, pp. 99–105.

[79] Julien Letessier and François Bérard. 2004. Visual tracking of bare fingers for interactive surfaces. In *Proceedings of the 17th annual ACM symposium on User interface software and technology* (UIST '04). ACM, New York, NY, USA, 119-122.

[80] Shahzad Malik and Joe Laszlo. 2004. Visual touchpad: a two-handed gestural input device. In *Proceedings of the 6th international conference on Multimodal interfaces* (ICMI '04). ACM, New York, NY, USA, 289-296.

[81] J. Segen and S. Kumar, "Shadow gestures: 3D hand pose estimation using a single camera," Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1999, vol. 1, pp. 479–485.

[82] J. Rekimoto, "SmartSkin," the SIGCHI conference, New York, New York, USA, 2002, pp. 113–120.

[83] Jefferson Y. Han. 2005. Low-cost multi-touch sensing through frustrated total internal reflection. In *Proceedings of the 18th annual ACM symposium on User interface software and technology* (UIST '05). ACM, New York, NY, USA, 115-118.

[84] Paul Dietz and Darren Leigh. 2001. DiamondTouch: a multi-user touch technology. In *Proceedings of the 14th annual ACM symposium on User interface software and technology* (UIST '01). ACM, New York, NY, USA, 219-226.

[85] B. Buxton, "Multi-Touch Systems that I Have Known and Loved." [Online]. Available: http://www.billbuxton.com/multitouchOverview.html. [Accessed: 10-Jul-2011].

[86] Jun Rekimoto, Masaaki Oka, Nobuyuki Matsushita, and Hideki Koike. 1998. HoloWall: interactive digital surfaces.

In *ACM SIGGRAPH 98 Conference abstracts and applications* (SIGGRAPH '98). ACM, New York, NY, USA, 108-.

[87] Z. Zhang, Y. Wu, Y. Shan, and S. Shafer, "Visual panel: virtual mouse, keyboard and 3D controller with an ordinary piece of paper," *PUI '01*, pp. 1–8, Nov. 2001.

[88] A. Agarwal, S. Izadi, M. Chandraker, and A. Blake, "High Precision Multi-touch Sensing on Surfaces using Overhead Cameras," Second Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP'07), 2007, pp. 197–200.

[89] Alex Butler, Shahram Izadi, and Steve Hodges. 2008. SideSight: multi-"touch" interaction around small devices. In *Proceedings of the 21st annual ACM symposium on User interface software and technology* (UIST '08). ACM, New York, NY, USA, 201-204.

[90] Andrew D. Wilson. 2004. TouchLight: an imaging touch screen and display for gesture-based interaction. In *Proceedings of the 6th international conference on Multimodal interfaces* (ICMI '04). ACM, New York, NY, USA, 69-76.

[91] Chris Harrison and Scott E. Hudson. 2008. Scratch input: creating large, inexpensive, unpowered and mobile finger input surfaces. In *Proceedings of the 21st annual ACM symposium on User interface software and technology* (UIST '08). ACM, New York, NY, USA, 205-208.

[92] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '09). ACM, New York, NY, USA, 1083-1092.

[93] M. Wu and R. Balakrishnan, "Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays," the 16th annual ACM symposium, New York, New York, USA, 2003, pp. 193–202.

[94] Andrew D. Wilson. 2010. Using a depth camera as a touch sensor. In *ACM International Conference on Interactive Tabletops and Surfaces* (ITS '10). ACM, New York, NY, USA, 69-72.

[95] oblong, “oblong industries, inc.,” *oblong.com*. [Online]. Available: http://oblong.com/. [Accessed: 28-Jul-2011].

[96] Pranav Mistry, Pattie Maes, and Liyan Chang. 2009. WUW - wear Ur world: a wearable gestural interface. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '09). ACM, New York, NY, USA, 4111-4116.

[97] Johannes Schöning, Michael Rohs, Sven Kratz, Markus Löchtefeld, and Antonio Krüger. 2009. Map torchlight: a mobile augmented reality camera projector unit. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '09). ACM, New York, NY, USA, 3841-3846.

[98] Xiang Cao and Ravin Balakrishnan. 2006. Interacting with dynamically defined information spaces using a handheld projector and a pen. In *Proceedings of the 19th annual ACM symposium on User interface software and technology* (UIST '06). ACM, New York, NY, USA, 225-234.

[99] Xiang Cao, Clifton Forlines, and Ravin Balakrishnan. 2007. Multi-user interaction using handheld projectors. In *Proceedings of the 20th annual ACM symposium on User interface software and technology* (UIST '07). ACM, New York, NY, USA, 43-52.

[100] N. Sakata, T. Konishi, and S. Nishida, “Mobile Interfaces Using Body Worn Projector and Camera,” in *Internet of Things, Smart Spaces, and Next Generation Networking*, vol. 5622, no. 4, S. Andreev, S. Balandin, and Y. Koucheryavy, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 106–113.

[101] L. B. Optics, “Light Touch - instantly turns any flat surface into a touch screen.” [Online]. Available: lightblueoptics.com. [Accessed: 12-Jul-2017].

[102] Canesta, “Canesta,” *en.wikipedia.org*. [Online]. Available: https://en.wikipedia.org/wiki/Canesta. [Accessed: 20-Aug-2017].

[103] Sony, “Xperia® Touch Official Website - Sony Mobile (United States).”

[104] R. Xiao, C. Harrison, and S. E. Hudson, “WorldKit: Rapid and Easy Creation of Ad-hoc Interactive Applications on Everyday

Surfaces," CHI '13 Extended Abstracts on Human Factors in Computing Systems, New York, NY, USA, 2013, pp. 2889–2890.

[105] R. Xiao, S. Hudson, and C. Harrison, "Supporting Responsive Cohabitation Between Virtual Interfaces and Physical Objects on Everyday Surfaces," *Proc. ACM Hum.-Comput. Interact.*, vol. 1, no. 1, pp. 12:1–12:17, Jun. 2017.

[106] Sony, "AIBO," *support.sony-europe.com*. [Online]. Available: http://support.sony-europe.com/aibo/. [Accessed: 24-May-2011].

[107] InnovoLabs, "Pleo," *pleoworld.com*. [Online]. Available: http://www.pleoworld.com/Home.aspx. [Accessed: 24-May-2011].

[108] iRobot, "iRobot," *irobot.com*. [Online]. Available: http://www.irobot.com/. [Accessed: 24-May-2011].

[109] *wowwee.com*. [Online]. Available: http://www.wowwee.com.

[110] I. Automata, "Autom," *intuitiveautomata.com*. [Online]. Available: http://intuitiveautomata.com/. [Accessed: 24-May-2011].

[111] S. A. Green, J. G. Chase, X. Chen, and M. Billinghurst, "Evaluating the Augmented Reality Human-Robot Collaboration System," 2008 15th International Conference on Mechatronics and Machine Vision in Practice (M2VIP), 2008, pp. 521–526.

[112] S. A. Green, M. Billinghurst, X. Chen, and J. G. Chase, "Human-Robot Collaboration: A Literature Review and Augmented Reality Approach in Design," *International Journal of Advanced Robotic Systems*, vol. 5, no. 1, p. 1, Mar. 2008.

[113] G. Hoffman, "Ensemble : fluency and embodiment for robots acting with humans," *dspace.mit.edu*. [Online]. Available: http://dspace.mit.edu/handle/1721.1/41705. [Accessed: 16-Mar-2011].

[114] S. Andre and L. Markus, " Outerspace." [Online]. Available: outerspace-robot.com. [Accessed: 12-Jul-2017].

[115] G. Pingali, C. Pinhanez, A. Levas, R. Kjeldsen, M. Podlaseck, H. Chen, and N. Sukaviriya, "Steerable Interfaces for Pervasive Computing Spaces," Proceedings of the First IEEE

International Conference on Pervasive Computing and Communications, Washington, DC, USA, 2003, pp. 315–315.

[116] Intel, "Intel® M2M Reference Design for Smart Services Development based on the Intel® AtomTM Processor," *download.intel.com*. [Online]. Available: http://download.intel.com/embedded/applications/M2M/324391.pdf. [Accessed: 10-Jul-2011].

[117] ALA, "Types of Light Sources and Light Bulbs | American Lighting Association," *americanlightingassoc.com*. [Online]. Available: http://www.americanlightingassoc.com/Lighting-Fundamentals/Light-Sources-Light-Bulbs.aspx. [Accessed: 11-Jul-2011].

[118] Ivan Poupyrev, Tatsushi Nashida, and Makoto Okabe. 2007. Actuation and tangible user interfaces: the Vaucanson duck, robots, and shape displays. In *Proceedings of the 1st international conference on Tangible and embedded interaction* (TEI '07). ACM, New York, NY, USA, 205-212.

[119] Wikipedia, "Balanced-arm lamp,". [Online]. Available: http://en.wikipedia.org/wiki/Balanced-arm_lamp. [Accessed: 15-Mar-2011].

[120] M. J. French and M. B. Widden, "The spring-and-lever balancing mechanism, George Carwardine and the Anglepoise lamp," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 214, no. 3, pp. 501–508, Jan. 2006.

[121] N. Linder, "LabCAST - The MIT Media Lab Video Podcast » Archive » #47 LuminAR," *labcast.media.mit.edu*. [Online]. Available: http://labcast.media.mit.edu/?p=129. [Accessed: 28-Jul-2011].

[122] ROBOTIS, ":::ROBOTIS:::," *robotis.com*. [Online]. Available: http://www.robotis.com/xe/dynamixel_en. [Accessed: 28-Jul-2011].

[123] Ramesh Raskar, Greg Welch, Matt Cutts, Adam Lake, Lev Stesin, and Henry Fuchs. 1998. The office of the future: a unified approach to image-based modeling and spatially immersive displays. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques* (SIGGRAPH '98). ACM, New York, NY, USA, 179-188.

[124] Apple, “iOS (Apple),” *en.wikipedia.org*. [Online]. Available: http://en.wikipedia.org/wiki/IOS_(Apple). [Accessed: 28-Jul-2011].

[125] Google, “Android,” *en.wikipedia.org*. [Online]. Available: http://en.wikipedia.org/wiki/Android_(operating_system). [Accessed: 28-Jul-2011].

[126] W. Garage, “ROS,” *ros.org*. [Online]. Available: http://www.ros.org/wiki/. [Accessed: 28-Jul-2011].

[127] Google, “protobuf - Protocol Buffers - Google's data interchange format,” *code.google.com*. [Online]. Available: http://code.google.com/p/protobuf/. [Accessed: 28-Jul-2011].

[128] H. Matthew, “A Motor Control Framework for Many-Axis Interactive Robots,” *M.Eng. Thesis Department of Electrical Engineering and Computer Science.*, Massachusetts Institute of Technology, May 2003.

[129] opencv, “OpenCV,” *opencv.willowgarage.com*. [Online]. Available: http://opencv.willowgarage.com/wiki/. [Accessed: 28-Jul-2011].

[130] reacTIVision, “reacTIVision,” *reactivision.sourceforge.net*. [Online]. Available: http://reactivision.sourceforge.net/. [Accessed: 28-Jul-2011].

[131] PMDTEC, “PMDTec: Reference Design,” *pmdtec.com*. [Online]. Available: http://www.pmdtec.com/products-services/reference-design/. [Accessed: 28-Jul-2011].

[132] metaio, “Metaio SDK.”

[133] A. H. Behzadan, B. W. Timm, and V. R. Kamat, “General-purpose modular hardware and software framework for mobile outdoor augmented reality applications in engineering,” *Advanced Engineering Informatics*, vol. 22, no. 1, pp. 90–105, Jan. 2008.

[134] M. Bauer, B. Bruegge, G. Klinker, A. MacWilliams, T. Reicher, S. Riss, C. Sandor, and M. Wagner, “Design of a Component-Based Augmented Reality Framework,” Proceedings of the International Symposium on Augmented Reality (ISAR), 2001, pp. 45–54.

[135] R. Ballagas, M. Ringel, M. Stone, and J. Borchers, “iStuff: A Physical User Interface Toolkit for Ubiquitous Computing

Environments," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA, 2003, pp. 537–544.

[136] Daqri, "Daqri - Smart Glasses." [Online]. Available: https://daqri.com/products/smart-glasses/. [Accessed: 22-Jul-2017].

[137] Daqri, "Daqri - Smart Helmet." [Online]. Available: https://daqri.com/products/smart-helmet/. [Accessed: 29-Jul-2017].

[138] J. Hardy and J. Alexander, "Toolkit support for interactive projected displays," MUM '12: Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia, 2012.

[139] S. Ahn, H. Ko, and S. Feiner, "Webizing mobile AR contents," 2013 IEEE Virtual Reality (VR), 2013, pp. 131–132.

[140] R. Xiao, C. Harrison, and S. E. Hudson, "WorldKit," the SIGCHI conference, New York, New York, USA, 2013, pp. 879–888.

[141] M. Weigel, S. Boring, J. Steimle, N. Marquardt, S. Greenberg, and A. Tang, "ProjectorKit: Easing Rapid Prototyping of Interactive Applications for Mobile Projectors," Proceedings of the 15th International Conference on Human-computer Interaction with Mobile Devices and Services, New York, NY, USA, 2013, pp. 247–250.

[142] M. Brubeck, S. Moon, and D. Schepers, "Touch events version 1 ," 2013.

[143] Designmodo, "Designmodo. Flat UI Free web user interface kit." [Online]. Available: http://www.designmodo.org/. [Accessed: 22-Jul-2017].

[144] N. Linder, R. Kubat, B. Weissmann, and P. Maes, "Lens: A javascript sdk for building web-based augmented reality applications on the Luminar platform," EECScon.

[145] I. Z. W. Group, "Zero Configuration Networking (Zeroconf)." [Online]. Available: http://www.zeroconf.org/. [Accessed: 22-Jul-2017].

[146] Donald A. Norman. 2010. Natural user interfaces are not natural. *interactions* 17, 3 (May 2010), 6-10.

[147] Christian von Hardenberg and François Bérard. 2001. Bare-hand human-computer interaction. In *Proceedings of the 2001 workshop on Perceptive user interfaces* (PUI '01). ACM, New York, NY, USA, 1-8.

[148] Microsoft, "Kinect for Windows SDK from Microsoft Research," *research.microsoft.com*. [Online]. Available: http://research.microsoft.com/en-us/um/redmond/projects/kinectsdk/. [Accessed: 28-Jul-2011].

[149] OpenNI, "Introducing OpenNI," *openni.org*. [Online]. Available: http://www.openni.org/. [Accessed: 28-Jul-2011].

[150] NUI, "NUI Group - Natural User Interface Group," *nuigroup.com*. [Online]. Available: http://nuigroup.com/go/. [Accessed: 28-Jul-2011].

[151] A. Parkes, I. Poupyrev, and H. Ishii, "Designing kinetic interactions for organic user interfaces," *Communications of the ACM*, vol. 51, no. 6, pp. 58–65, Jun. 2008.

[152] H. Raffle, A. Parkes, and H. Ishii, "Topobo: a constructive assembly system with kinetic memory," *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems*, Apr. 2004.

[153] D. Leithinger and H. Ishii, "Relief: a scalable actuated shape display," Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction, 2010, pp. 221–222.

[154] I. Poupyrev, T. Nashida, S. Maruyama, J. Rekimoto, and Y. Yamaji, "Lumen: interactive visual and shape display for calm computing," ACM SIGGRAPH 2004 Emerging technologies, 2004, p. 17.

[155] J. K. Salisbury and M. A. Srinivasan, "Phantom-based haptic interaction with virtual objects," *IEEE Comput. Grap. Appl.*, vol. 17, no. 5, pp. 6–10, Sep. 1997.

[156] William Buxton, Ralph Hill, and Peter Rowley. 1985. Issues and techniques in touch-sensitive tablet input. In *Proceedings of the 12th annual conference on Computer graphics and interactive techniques* (SIGGRAPH '85). ACM, New York, NY, USA, 215-224.

[157] W. Westerman, “Hand Tracking, Finger Identification, and Chordic Manipulation on a Multi-Touch Surface,” University of Delaware, 1999.

[158] M. Hirsch, D. Lanman, H. Holtzman, and R. Raskar, “BiDi screen,” ACM SIGGRAPH ASIA 2009 Art Gallery & Emerging Technologies: Adaptation, New York, New York, USA, 2009, pp. 62–62.

[159] B. J. Rhodes and P. Maes, “Just-in-time Information Retrieval Agents,” *IBM Syst. J.*, vol. 39, no. 3, pp. 685–704, Jul. 2000.

[160] Jun Rekimoto. 1997. Pick-and-drop: a direct manipulation technique for multiple computer environments. In *Proceedings of the 10th annual ACM symposium on User interface software and technology* (UIST '97). ACM, New York, NY, USA, 31-39.

[161] A. Butz, T. Hollerer, S. Feiner, B. MacIntyre, and C. Beshers, “Enveloping users and computers in a collaborative 3D augmented reality,” Augmented Reality, 1999. (IWAR '99) Proceedings. 2nd IEEE and ACM International Workshop on, 1999, pp. 35–44.

[162] Evernote, “Remember Everything | Evernote Corporation,” *evernote.com*. [Online]. Available: http://www.evernote.com/. [Accessed: 28-Jul-2011].

[163] Intel, “The Responsive Store,” *edc.intel.com*. [Online]. Available: http://edc.intel.com/Video-Player.aspx?id=4158. [Accessed: 17-Jun-2011].

[164] Stacey D. Scott, Karen D. Grant, and Regan L. Mandryk. 2003. System guidelines for co-located, collaborative work on a tabletop display. In *Proceedings of the eighth conference on European Conference on Computer Supported Cooperative Work* (ECSCW'03), Kari Kuutti, Eija Helena Karsten, Geraldine Fitzpatrick, Paul Dourish, and Kjeld Schmidt (Eds.). Kluwer Academic Publishers, Norwell, MA, USA, 159-178.

[165] Pranav Mistry, Suranga Nanayakkara, and Pattie Maes. 2011. Touch and copy, touch and paste. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '11). ACM, New York, NY, USA, 1095-1098.

[166] “Swyp.” [Online]. Available: https://github.com/swyp/swyp. [Accessed: 05-Jul-2017].

[167] Richard Buchanan. Wicked problems in design thinking. Design Issues, 8(2):5–21, 1992.

[168] GRID, "How to Grid." [Online]. Available: https://vimeo.com/69590221. [Accessed: 05-Jul-2017].

[169] Steelcase, "Campfire Big Lamp | Turnstone Furniture." [Online]. Available: https://www.steelcase.com/products/lighting/campfire-big-lamp/. [Accessed: 05-Jul-2017].

[170] P. André, M. L. Wilson, A. Russell, D. A. Smith, A. Owens, and M. C. schraefel, *Continuum: designing timelines for hierarchies, relationships and scale*. New York, NY, USA: ACM, 2007, pp. 101–110.

[171] R. Dachselt and M. Weiland, "TimeZoom," CHI '06 extended abstracts, New York, New York, USA, 2006, pp. 682–687.

[172] Chia Shen, Frédéric D. Vernier, Clifton Forlines, and Meredith Ringel. 2004. DiamondSpin: an extensible toolkit for around-the-table interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '04). ACM, New York, NY, USA, 167-174.

[173] O. Hilliges, D. Baur, and A. Butz, "Photohelix: Browsing, Sorting and Sharing Digital Photo Collections," Horizontal Interactive Human-Computer Systems, 2007. TABLETOP '07. Second Annual IEEE International Workshop on, 2007, pp. 87–94.

[174] Intel and N. Linder, "LuminAR Product Counter," *download.intel.com*. [Online]. Available: http://download.intel.com/pressroom/kits/embedded/pdfs/MIT_LuminAR_Counter_Overview.pdf. [Accessed: 28-Jul-2011].

[175] T. P. Caudell and D. W. Mizell, "Augmented reality: an application of heads-up display technology to manual manufacturing processes," 2004, vol. ii, pp. 659–669 vol.2.

[176] J. L. Nevins and D. E. Whitney, "Assembly research," *Automatica*, vol. 16, no. 6, pp. 595–613, Nov. 1980.

[177] Seymour, W. D. *Industrial skills*. London: Pitman, 1966.

[178] J. P. Womack, D. T. Jones, and D. Roos, *Machine that changed the world*. Simon and Schuster, 1990.

[179] M. Holweg, "The genealogy of lean production," vol. 25, no. 2, pp. 420–437, 2007.

[180] F. Capozzi and M. Sacco, “Exploiting Augmented Reality in Lean Manufacturing: Opportunities and Challenges,” in *Advances in Sustainable and Competitive Manufacturing Systems*, no. 110, Heidelberg: Springer International Publishing, 2013, pp. 1361–1373.

[181] Steelcase, “c:scape Office Workstations & Desk Systems - Steelcase.” [Online]. Available: https://www.steelcase.com/products/desk-systems/c-scape-4/. [Accessed: 12-Jul-2017].

[182] S. J. Henderson and S. K. Feiner, “Augmented reality in the psychomotor phase of a procedural task,” 2011.

[183] Steven Feiner, Blair Macintyre, and Dorée Seligmann. 1993. Knowledge-based augmented reality. *Commun. ACM* 36, 7 (July 1993), 53-62.

[184] U. Neumann and A. Majoros. 1998. Cognitive, Performance, and Systems Issues for Augmented Reality Applications in Manufacturing and Maintenance. In *Proceedings of the Virtual Reality Annual International Symposium* (VRAIS '98). IEEE Computer Society, Washington, DC, USA, 4.

[185] M. Sugi *et al.*, "Motion Control of Self-Moving Trays for Human Supporting Production Cell “ Attentive Workbench ”," *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, Barcelona, Spain, 2005, pp. 4080-4085.

[186] B. Schwerdtfeger and G. Klinker, “Supporting order picking with Augmented Reality,” 2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR), Washington, DC, USA, 2008, pp. 91–94.

[187] Jianlong Zhou, Ivan Lee, Bruce Thomas, Roland Menassa, Anthony Farrant, and Andrew Sansome. 2011. Applying spatial augmented reality to facilitate in-situ support for automotive spot welding inspection. In Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry (VRCAI '11). ACM, New York, NY, USA, 195-200.

[188] P. Fite-Georgel, “Is there a reality in Industrial Augmented Reality?,” 2011 IEEE International Symposium on Mixed and Augmented Reality, Washington, DC, USA, 2011, pp. 201–210.

[189] F. Echtler, F. Sturm, K. Kindermann, G. Klinker, J. Stilla, J. Trilk, and H. Najafi, “The Intelligent Welding Gun: Augmented Reality for Experimental Vehicle Construction,” in *Virtual and Augmented Reality Applications in Manufacturing*, no. 17, London: Springer London, 2004, pp. 333–360.

[190] P. F. Georgel, P. Schroeder, and N. Navab, “Navigation Tools for Viewing Augmented CAD Models,” *IEEE Comput. Grap. Appl.*, vol. 29, no. 6, pp. 65–73, Nov. 2009.

[191] A. C. Boud, D. J. Haniff, C. Baber, and S. J. Steiner, “Virtual reality and augmented reality as a training tool for assembly tasks,” 1999 IEEE International Conference on Information Visualization, 1999, pp. 32–36.

[192] Hou, Lei & Wang, Xiangyu. (2013). A study on the benefits of augmented reality in retaining working memory in assembly tasks: A focus on differences in gender. Automation in Construction. 32. 38–45.

[193] Arthur Tang, Charles Owen, Frank Biocca, and Weimin Mou. 2003. Comparative effectiveness of augmented reality in object assembly. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '03). ACM, New York, NY, USA, 73-80.

[194] S. Wiedenmaier, O. Oehme, L. Schmidt and H. Luczak, "Augmented reality (AR) for assembly processes - an experimental evaluation," *Proceedings IEEE and ACM International Symposium on Augmented Reality*, New York, NY, 2001, pp. 185-186.

[195] Sabine Webel, Ulrich Bockholt, and Jens Keil. 2011. Design criteria for AR-based training of maintenance and assembly tasks. In *Proceedings of the 2011 international conference on Virtual and mixed reality: new trends - Volume Part I*, Randall Shumaker (Ed.), Vol. Part I. Springer-Verlag, Berlin, Heidelberg, 123-132.

[196] S. K. Ong, M. L. Yuan, and A. Y. C. Nee, “Augmented reality applications in manufacturing: a survey,” *International Journal of Production Research*, vol. 46, no. 10, pp. 2707–2742, May 2008.

[197] J. Novak-Marcincin, J. Barna, M. Janak, and L. Novakova-Marcincinova, “Augmented Reality Aided Manufacturing,” *Procedia Computer Science*, vol. 25, pp. 23–31, 2013.

[198] Ángel Medinilla. 2014. *Agile Kaizen: Managing Continuous Improvement Far Beyond Retrospectives*. Springer Publishing Company, Incorporated.

[199] A. Y. C. Nee, S. K. Ong, G. Chryssolouris, and D. Mourtzis, "Augmented reality applications in design and manufacturing," *CIRP Annals - Manufacturing Technology*, vol. 61, no. 2, pp. 657–679, Jan. 2012.

[200] X. Wang and P. S. Dunston, "Design, strategies, and issues towards an augmented reality-based construction training platform," *Journal of information technology in construction (ITcon)*, vol. 12, no. 25, pp. 363–380, 2007.

[201] Wikipedia, "Coursera." [Online]. Available: https://en.wikipedia.org/wiki/Coursera. [Accessed: 12-Jul-2017].

[202] Wikipedia, "Khan Academy." [Online]. Available: https://en.wikipedia.org/wiki/Khan_Academy. [Accessed: 12-Jul-2017].

[203] J. C. Sprott, *Physics Demonstrations: A Sourcebook for Teachers of Physics*. University of Wisconsin Press, 2015.

[204] Y. Saw, *Enlight : a projected augmented reality approach to science education*. Massachusetts Institute of Technology, 2014.

[205] J. Belcher, *Professor John Belcher website*. [Online]. Available: http://web.mit.edu/jbelcher/www/. [Accessed: 12-Jul-2017].

[206] J.W. Belcher, "Studio Physics at MIT", *MIT Physics Ann.*, pp. 58-64, 2001.

[207] "Professor Eric Mazur." [Online]. Available: http://mazur.harvard.edu/eric_mazur.php. [Accessed: 12-Jul-2017].

[208] E. Mazur, "Peer Instruction: A User's Manual," vol. 67, no. 4, pp. 359–360, 1999.

[209] E. Mazur, "Peer instruction: Getting students to think in class," *AIP Conference Proceedings*, vol. 399, no. 1, pp. 981–988, 1997.

[210] C. H. Crouch and E. Mazur, "Peer Instruction: Ten years of experience and results," *American Journal of Physics*, vol. 69, no. 9, pp. 970–977, Sep. 2001.

[211] O. Zuckerman, "Flowness+ FlowBlocks: uncovering the dynamics of everyday life through playful modeling. ," MIT, Cambridge, MA.

[212] E. V. Glasersfeld, "The Construction of Knowledge, Contributions to Conceptual Semantics.," 1987.

[213] M. Montessori, *The Secret of Childhood*. Orient Blackswan, 1978.

[214] J. Piaget, *Principles of Genetic Epistemology*. Routledge, 2013.

[215] S. Turkle, S. Papert, Massachusetts Institute of Technology. Epistemology & Learning Research Group, *Epistemological Pluralism*. 1990.

[216] S. Papert, *The children's machine: Rethinking school in the age of the computer*. Basic books, 1994.

[217] Price, S. Tangibles: Technologies and interaction for learning, in Sage Handbook of Digital Technology Research, Chapter 20, 2013.

[218] M. W. Keyser, "Active learning and cooperative learning: understanding the difference and using both styles effectively," *Research Strategies*, vol. 17, no. 1, pp. 35–44, 2000.

[219] D. W. Johnson, R. T. Johnson, and K. A. Smith, *Active Learning: Cooperation in the College Classroom*. Interaction Book Company, 1991.

[220] C. C. Bonwell and J. A. Eison, *Active Learning: Creating Excitement in the Classroom. 1991 ASHE-ERIC Higher Education Reports.* ERIC, 1991.

[221] David P. Bunde, Joe Hummel, and Jaime Spacco. 2015. Experiences adopting peer instruction. *J. Comput. Sci. Coll.* 31, 1 (October 2015), 106-107.

[222] R. S. Shieh, "The impact of Technology-Enabled Active Learning (TEAL) implementation on student learning and teachers' teaching in a high school context," *Computers & Education*, vol. 59, no. 2, pp. 206–214, Sep. 2012.

[223] J Y.Dori and J. Belcher. How does technology-enabled active learning affect under-graduate students' understanding of electromagnetism concepts? The Journal of Learning Sciences, pages 243–279, 2005.

[224] B. Victor, “Simulation as a Practical Tool,” Oct. 2009. [Online]. Available: http://worrydream.com/SimulationAsAPracticalTool/. [Accessed: 22-Jul-2017].

[225] C. Chae and K. Ko, “Introduction of Physics Simulation in Augmented Reality,” International Symposium on Ubiquitous Virtual Reality ISUVR, 2008, pp. 37–40.

[226] Seymour Papert. 1980. *Mindstorms: Children, Computers, and Powerful Ideas*. Basic Books, Inc., New York, NY, USA.

[227] U. Haber-Shaim, J. H. Dodge, R. Gardner, E. A. Shore, Physics Textbook Review Committee, “PSSC (Physical Science Study Committee) Physics,” *The Physics Teacher*, vol. 37, no. 5, pp. 294–295, May 1999.

[228] R. P. Feynman, R. B. Leighton, and M. L. Sands, *The Feynman Lectures on Physics*, no. 1. Addison-Wesley, 1963.

[229] V. Colella, R. Borovoy, and M. Resnick, “Participatory simulations: Using computational objects to learn about dynamic systems,” CHI 98 Conference Summary on Human Factors in Computing Systems, 1998, pp. 9–10.

[230] Lana Trey and Samia Khan. 2008. How science students can learn about unobservable phenomena using computer-based analogies. *Comput. Educ.* 51, 2 (September 2008), 519-529.

[231] C. Perry, “In AP 50, students own their education,” Sep. 2013. [Online]. Available: https://www.seas.harvard.edu/news/2013/09/in-ap-50-students-own-their-education. [Accessed: 22-Jul-2017].

[232] Yu-Chien Chen. 2006. A study of comparing the use of augmented reality and physical models in chemistry education. In *Proceedings of the 2006 ACM international conference on Virtual reality continuum and its applications* (VRCIA '06). ACM, New York, NY, USA, 369-372.

[233] María Blanca Ibáñez, Ángela Di Serio, Diego Villarán, and Carlos Delgado Kloos. 2014. Experimenting with electromagnetism using augmented reality: Impact on flow student experience and educational effectiveness. *Comput. Educ.* 71 (February 2014), 1-13.

[234] Rafał Wojciechowski and Wojciech Cellary. 2013. Evaluation of learners' attitude toward learning in ARIES augmented

reality environments. *Comput. Educ.* 68 (October 2013), 570-585.

[235] Mark Billinghurst and Andreas Duenser. 2012. Augmented Reality in the Classroom. *Computer* 45, 7 (July 2012), 56-63.

[236] Son Do-Lenh, Patrick Jermann, Amanda Legge, Guillaume Zufferey, and Pierre Dillenbourg. 2012. TinkerLamp 2.0: designing and evaluating orchestration technologies for the classroom. In *Proceedings of the 7th European conference on Technology Enhanced Learning* (EC-TEL'12), Andrew Ravenscroft, Stefanie Lindstaedt, Carlos Delgado Kloos, and Davinia Hernández-Leo (Eds.). Springer-Verlag, Berlin, Heidelberg, 65-78.

[237] M. Bower, C. Howe, N. McCredie, A. Robinson and D. Grover, "Augmented reality in Education — Cases, places, and potentials," *2013 IEEE 63rd Annual Conference International Council for Education Media (ICEM)*, Singapore, 2013, pp. 1-11.

[238] H. Kaufmann, "Collaborative Augmented Reality in Education," in *Imagina Conference 2003*, Monaco: Monaco Mediax, 2003.

[239] HoloStudy, "HoloStudy." [Online]. Available: http://www.holo.study/. [Accessed: 22-Jul-2017].

[240] Steelcase, "Steelcase - Node: Seating for the active learning classroom." [Online]. Available: https://www.steelcase.com/eu-en/products/collaborative-chairs/node/. [Accessed: 22-Jul-2017].

[241] "LEGO Duplo." [Online]. Available: http://www.lego.com/en-us/duplo. [Accessed: 12-Jul-2017].

[242] A. Dunser, L. Walker, H. Horner, and D. Bentall, "Creating Interactive Physics Education Books with Augmented Reality," Proceedings of the 24th Australian Computer-Human Interaction Conference, New York, NY, USA, 2012, pp. 107–114.

[243] S. Uras, D. Ardu, G. Paddeu, and M. Deriu, "Do Not Judge an Interactive Book by Its Cover: A Field Research," Proceedings of the 10th International Conference on Advances in Mobile Computing & Multimedia, New York, NY, USA, 2012, pp. 17–20.

[244] F. Radke, *How Augmented Reality Will Change Education Completely*. TEDxGateway. [Online]. Available: https://www.youtube.com/watch?v=5AjxGqzqQ54. [Accessed: 12-Jul-2017].

[245] R. E. L. Grasset, A. Dunser, and M. Billinghurst, "Edutainment with a Mixed Reality Book: A Visually Augmented Illustrative Childrens' Book," Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology, New York, NY, USA, 2008, pp. 292–295.

[246] H.-T. Jeong, D.-W. Lee, G.-S. Heo, and C.-H. Lee, "Live Book: A mixed reality book using a projection system," 2012 IEEE International Conference on Consumer Electronics (ICCE), 2012, pp. 680–681.

[247] F. Mannus, J. Rubel, C. Wagner, F. Bingel, and A. Hinkenjann, "Augmenting magnetic field lines for school experiments," 2011 IEEE International Symposium on Mixed and Augmented Reality, 2011, pp. 263–264.

[248] A. Gillet, M. Sanner, D. Stoffler, and A. Olson, "Tangible Augmented Interfaces for Structural Molecular Biology," *IEEE Comput. Grap. Appl.*, vol. 25, no. 2, pp. 13–17, Mar. 2005.

[249] M. H. Towns and E. R. Grant, "'i believe i will go out of this class actually knowing something': Cooperative learning activities in physical chemistry," *Journal of Research in Science Teaching*, vol. 34, no. 8, pp. 819–835, 1997.

[250] Steelcase, "Steelcase - verb." [Online]. Available: https://www.steelcase.com/eu-en/products/meeting-tables/verb/. [Accessed: 12-Jul-2017].

[251] WebRTC, *WebRTC (Web Real-Time Communication)*. [Online]. Available: https://webrtc.org/. [Accessed: 12-Jul-2017].

[252] M. W. van Someren, Y. F. Barnard, and J. Sandberg, *The Think Aloud Method: A Practical Guide to Modelling Cognitive Processes*. Academic Press, 1994.

[253] C. Lewis, *Using the "thinking Aloud" Method in Cognitive Interface Design*. IBM T.J. Watson Research Center, 1982.

[254] C. J. Johnstone, N. A. Bottsford-Miller, and S. J. Thompson, "Using the Think Aloud Method (Cognitive Labs) to Evaluate Test Design for Students with Disabilities and English Language Learners. Technical Report 44.," 2006.

[255] K. A. Ericsson and H. A. Simon, "Verbal reports as data.," *Psychological review*, vol. 87, no. 3, p. 215, 1980.

[256] K. Pearson, "Note on Regression and Inheritance in the Case of Two Parents," *Proceedings of the Royal Society of London*, vol. 58, no. 347, pp. 240–242, 1895.

[257] C. Spearman, "The proof and measurement of association between two things," *International Journal of Epidemiology*, vol. 39, no. 5, pp. 1137–1150, 2010.

[258] D. A. Kolb, *Experiential Learning: Experience as the Source of Learning and Development*. Prentice-Hall, 1984.

[259] A. Y. Kolb and D. A. Kolb, "Learning styles and learning spaces: Enhancing experiential learning in higher education," *Academy of management learning \& education*, vol. 4, no. 2, pp. 193–212, 2005.

[260] A. F. Gregorc and H. B. Ward, "A New Definition for Individual," *NASSP Bulletin*, vol. 61, no. 406, pp. 20–26, 1977.

[261] T. F. Hawk and A. J. Shah, "Using Learning Style Instruments to Enhance Student Learning," *Decision Sciences Journal of Innovative Education*, vol. 5, no. 1, pp. 1–19, 2007.

[262] P. R. Boyce, *Human Factors in Lighting, Third Edition*. Taylor & Francis, 2014.

[263] C. W. Reynolds, "Flocks, Herds and Schools: A Distributed Behavioral Model," Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques, New York, NY, USA, 1987, pp. 25–34.

[264] CEVA, "ceva-xm4," *ceva-dsp.com*. [Online]. Available: http://www.ceva-dsp.com/product/ceva-xm4/. [Accessed: 27-Aug-2017].

[265] Jianlong Zhou, Ivan Lee, Bruce Thomas, Roland Menassa, Anthony Farrant, and Andrew Sansome. 2011. Applying spatial augmented reality to facilitate in-situ support for automotive spot welding inspection. In *Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry* (VRCAI '11). ACM, New York, NY, USA, 195-200.

[266] M. Richtel, "For Multitaskers, Multiple Monitors Improve Office Efficiency," *New York Times*, 07-Feb-2012.

[267] Amazon, "alexa," *developer.amazon.com.* [Online]. Available: https://developer.amazon.com/alexa. [Accessed: 27-Aug-2017].

[268] Google, "Google Home." [Online]. Available: https://madeby.google.com/home/. [Accessed: 27-Aug-2017].

[269] A. Torralba, R. Fergus, and W. T. Freeman, "80 Million Tiny Images: A Large Data Set for Nonparametric Object and Scene Recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 11, pp. 1958–1970.

[270] G. Klein and D. Murray, "Parallel Tracking and Mapping for Small AR Workspaces," 2007 6th IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 2007, pp. 1–10.

Appendix A: LuminAR Technical Specifications

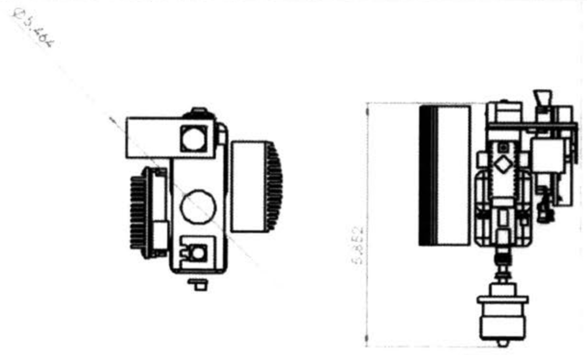
This appendix includes mechanical specifications for the various LuminAR systems developed in the course of this thesis work.

LuminAR Bulb Integration

Weight (w/out M2M Embedded computer)	460g
Diameter	150mm
Actual rotation range (degrees)	45°
Length	190mm

LuminAR Bulb Final

Weight	550g
Diameter	130mm
Actual rotation range (degrees)	270°
Length	170 mm



LuminAR Spotlight BD

Overall length	1450mm
Belt & pulley mechanism length	1200mm
Pivot servo actual motion range (deg)	120°
Weight	3700g
Depth - to explain the mounting requirements	100 mm in the wall/ 18mm above +Bulb

LuminAR Optimus

Base diameter	304.8 mm (12 in)
Base rotation actual rotation (degrees)	90°
Hard stops	~90°, driven by cable length
Length of shoulder link	355.6mm (14in)
Length of elbow link	355.6mm (14in)
Interesting mechanical elements	All linkages are "doubled," servos directly attached to arms
Weight (estimated)	4000g
Elbow joint range (degrees)	100°
Shoulder joint range	90°

(degrees)	
Wrist joint range (degrees)	100° (actual range 65° because of weight + spring)

LuminAR Aluminum

Base diameter	254mm (10in)
Base rotation actual rotation (degrees)	360° / 90° with stops

Hard stops	Cable management
Length of shoulder link	355.6mm
Length of elbow link	355.6mm
Interesting mechanical elements	Elbow spring is parallel to the base and attaches to the elbow with a long string, which goes through a pulley
Weight (estimated)	5000g
Elbow joint range (degrees)	~90°
Shoulder joint range (degrees)	~90°
Wrist joint range (degrees)	~70°

LuminAR Tipsy

Base circumference	241.3 mm (9.5 in)
Base rotation actual	90°

rotation (degrees)	
Hard stops	Cable management
Length of shoulder link	355.6mm
Length of elbow link	355.6mm
Interesting mechanical elements	Fork; servomotors are located on the base and connect trough hard linkages to the arms; Dynamixel servos
Weight (estimated)	4300 g (+ bulb)
Elbow joint range (degrees)	~90°
Shoulder joint range (degrees)	~90°
Wrist joint range (degrees)	~70°

LuminAR SilverJack / BlackJack

Base circumference	241.3mm (9.5 in)
Base rotation actual rotation (degrees)	300°
Hard stops	Fork
Length of shoulder link	355.6 mm (14 in)
Length of elbow link	355.6 mm (14 in)
Interesting mechanical	The base is divided in two parts (lower steel

elements	for more mass); aesthetically designed fork and arms
Weight (estimated)	4300g (+ Bulb)
Elbow joint range (degrees)	90°
Shoulder joint range (deg)	93°
Wrist joint range (deg)	70

LuminAR Retail

Base circumference	241.3 mm (9.5 in)
Base rotation actual rotation (degrees)	300°
Hard stops	Lower arm / base cover
Length of shoulder link	355.6mm (14in)
Length of elbow link	355.6 mm (14 in)
Interesting mechanical elements	The base is embedded in the counter (86 mm); wrist has its own 4 bar linkage
	5000g

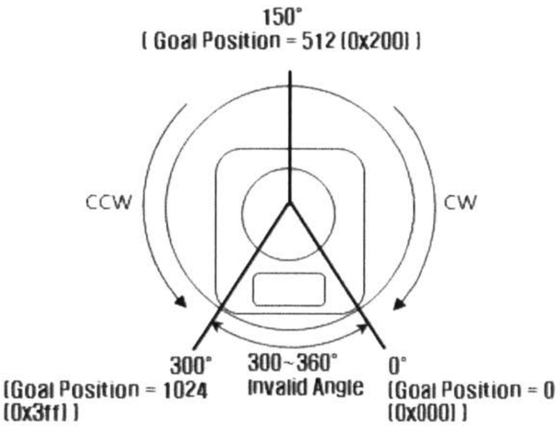
Weight (estimated)	
Elbow joint range (degrees)	75°
Shoulder joint range (degrees)	74°
Wrist joint range (degrees)	15°

LuminAR BlackJack / SilverJack S/W Limits (in encoder values)

Joint	Min	Max
Base	100	950
Shoulder	350	600
Elbow	370	635
Wrist	50	250

To convert Dynamixel DX-117 encoder values to degrees:

See more at: http://support.robotis.com/en, Product Information / Dynamixel / DX Series / DX-117



128

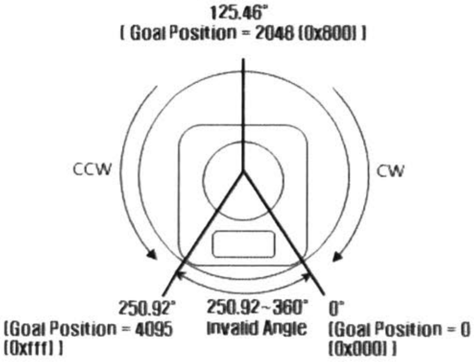
LuminAR Retail S/W Limits (in encoder values)

Joint	Min	Max
Base (DX-117)	100	950
Shoulder (EX-106+)	1651	2672
Elbow (EX-106+)	1306	2302
Wrist (DX-117)	400	600

To convert Dynamixel DX-117 encoder values to degrees - see above

To convert Dynamixel EX-106+ encoder values to degrees:

See more at: http://support.robotis.com/en, Product Information / Dynamixel / EX Series / EX-106+



Appendix B: MARS Quantitative User Study Data

MARS Button Press Data

Day 2 (5/28/2013)			
Trial	**Hits**	**Misses**	**Percentage**
1	22	3	88.00%
2	20	5	80.00%
3	23	2	92.00%
4	25	0	100.00%
5	23	2	92.00%
6	23	2	92.00%
7			0.00%
8			0.00%
9			0.00%
10			
		Average	90.67%

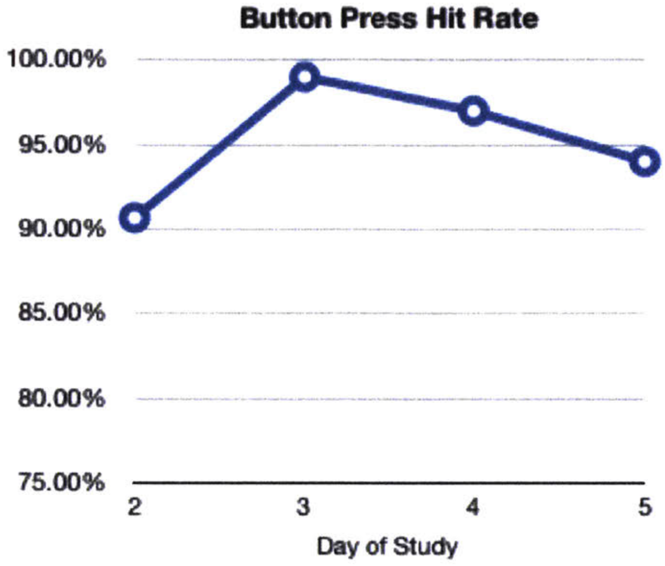
Day 3 (5/29/2013)			
Trial	**Hits**	**Misses**	**Percentage**
1	25	0	100.00%
2	25	0	100.00%
3	25	0	100.00%
4	24	1	96.00%
5			0.00%
6			0.00%
7			0.00%
8			0.00%
9			0.00%
		Average	99.00%

Day 4 (5/30/2012)			
Trial	**Hits**	**Misses**	**Percentage**
1	25	0	100.00%
2	25	0	100.00%
3	24	1	96.00%
4	23	2	92.00%
			0.00%

			0.00%
			0.00%
			0.00%
			0.00%
			0.00%
		Average	97.00%

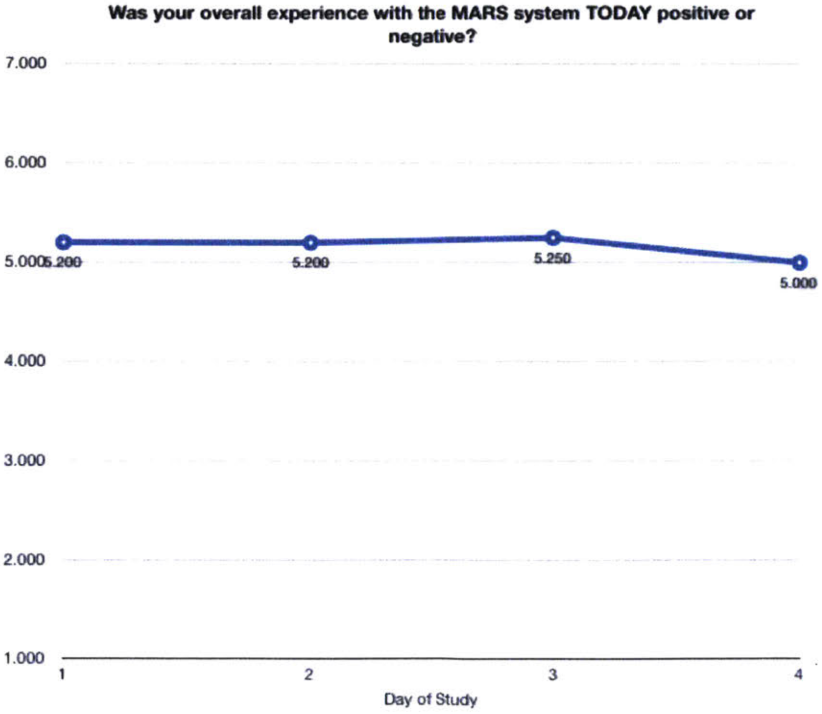
Day 5 (5/31/2012)

Trial	Hits	Misses	Percentage
1	21	4	84.00%
2	25	0	100.00%
3	25	0	100.00%
4	23	2	92.00%
			0.00%
			0.00%
			0.00%
			0.00%
			0.00%
			0.00%
		Average	94.00%

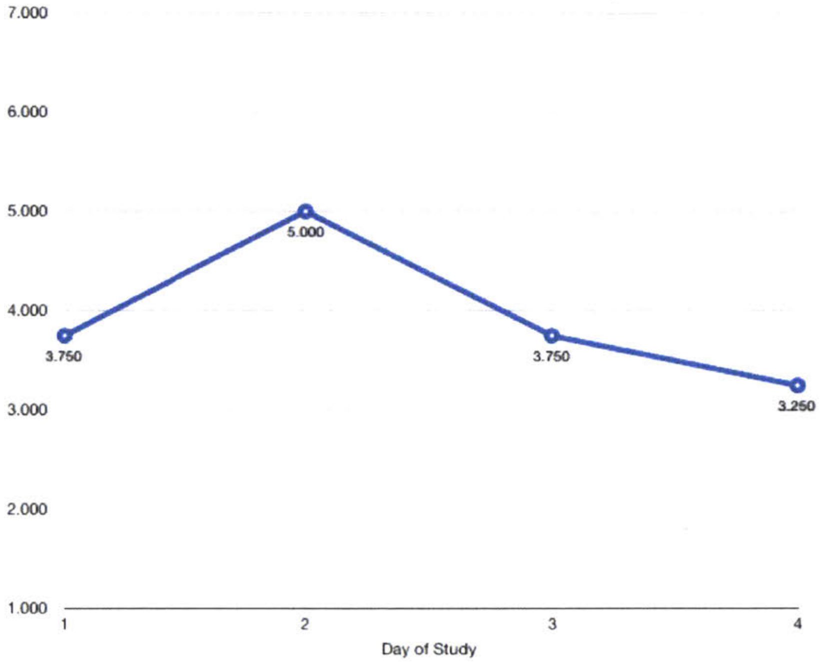


Day	2	3	4	5
Button Press Hit Rate	90.67%	99.00%	97.00%	94.00%

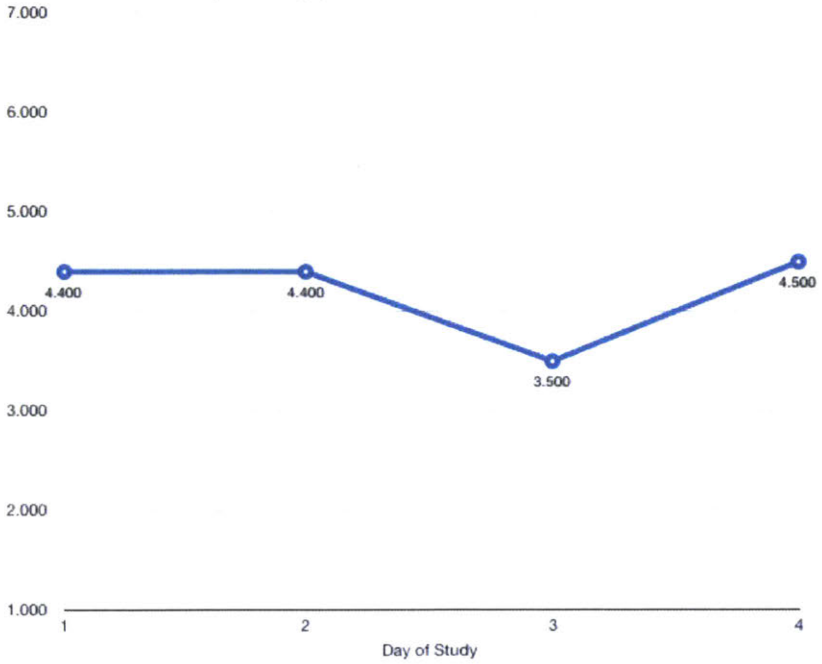
MARS User Survey Data



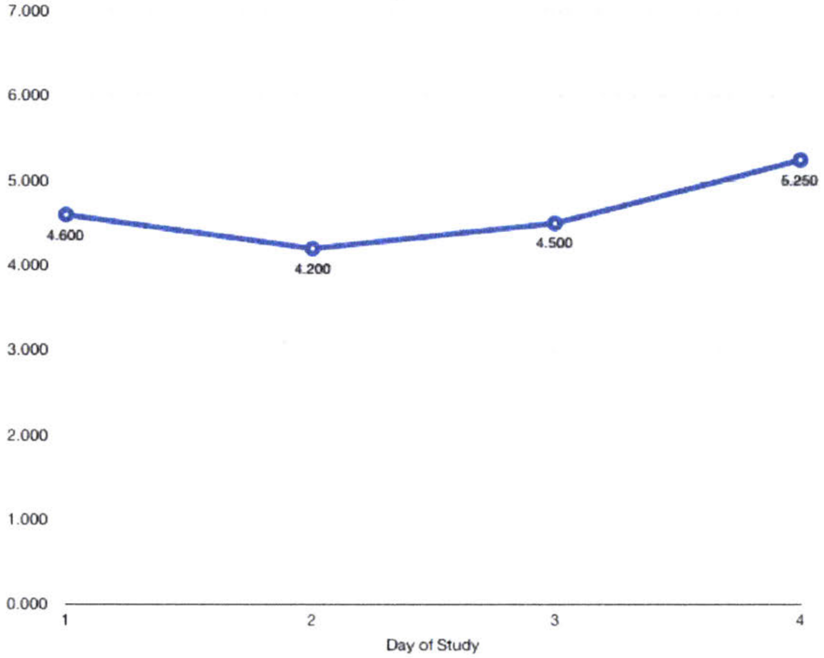
How did your cycle time performance compare to the previous shift?



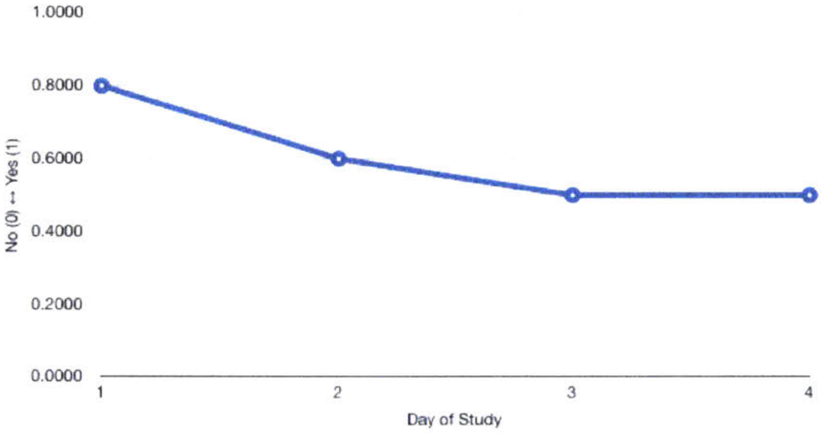
Did MARS system help you to learn and remember the standard work?

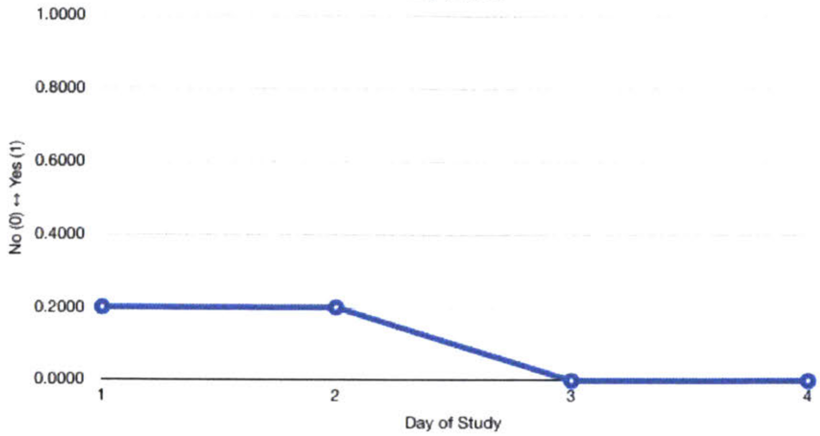


Did MARS projected Standard Work steps help you recall the next build operation?

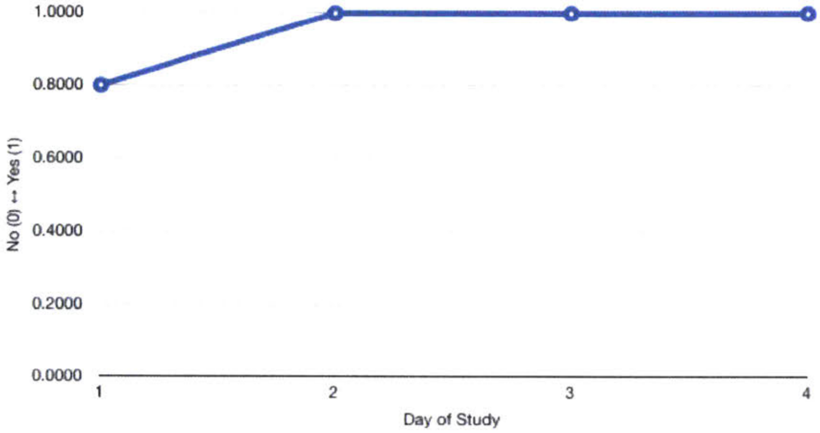


Do you feel the MARS system affected your cycle time negatively?

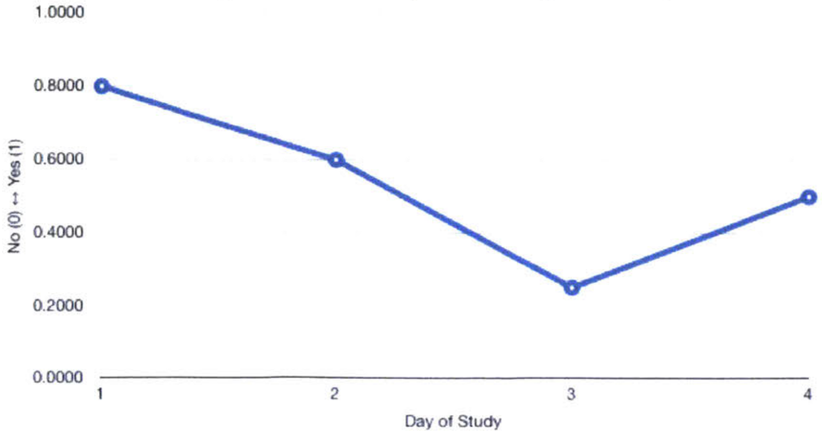




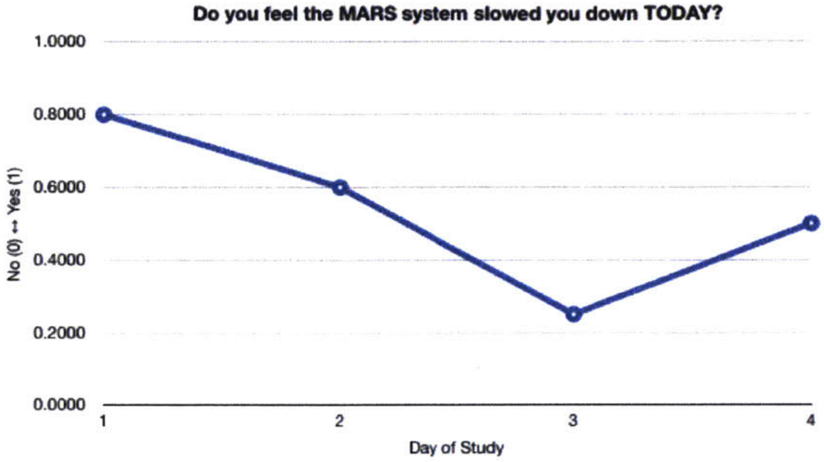
Did the MARS system negatively affect your ability to perform your build at any point?



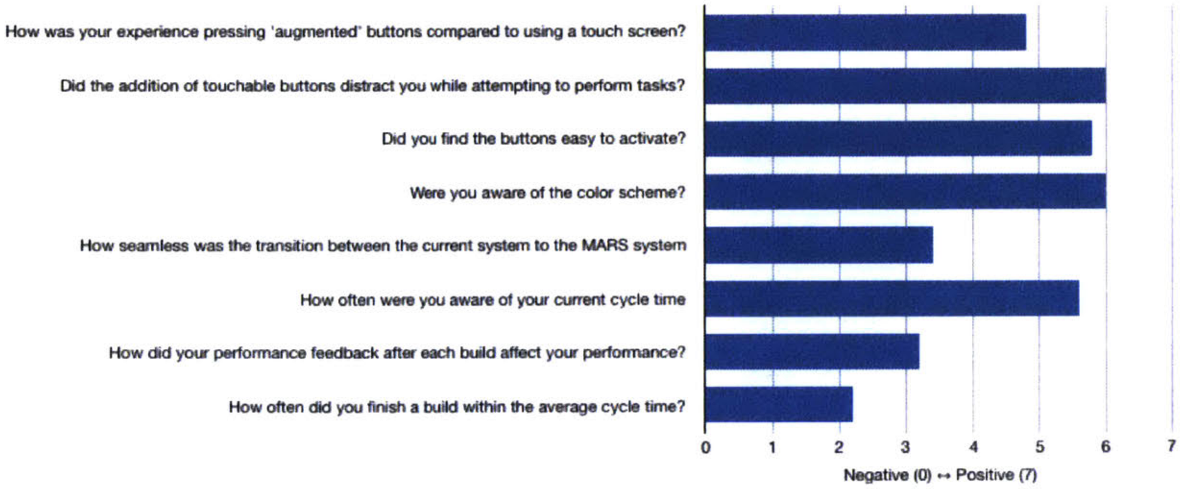
Do you feel you were able to perform your builds adequately?



Do you feel the MARS system slowed you down TODAY?



Do you feel the MARS system slowed you down TODAY?
1.0000
0.8000
0.6000
0.4000
0.2000
0.0000
No (0) ↔ Yes (1)
1
2
3
4
Day of Study



How was your experience pressing 'augmented' buttons compared to using a touch screen?
Did the addition of touchable buttons distract you while attempting to perform tasks?
Did you find the buttons easy to activate?
Were you aware of the color scheme?
How seamless was the transition between the current system to the MARS system
How often were you aware of your current cycle time
How did your performance feedback after each build affect your performance?
How often did you finish a build within the average cycle time?
0
1
2
3
4
5
6
7
Negative (0) ↔ Positive (7)

Question	Range: 0	Range: 7	Mean	Std. Error
How was your experience pressing "augmented" buttons compared to using a touch screen?	Prefer Touch Screen	Prefer Augmented Buttons	4.8	0.490
Did the addition of touchable buttons distract you while attempting to perform tasks?	It was distracting	It was Helpful	6.0	0.632
Did you find the buttons easy to activate?	No	Yes	5.8	0.583
Were you aware of the color scheme?	I was not aware	It was very clear	6.0	1.000
How seamless was the transition between the current system to the MARS system	Not seamless at all	Seamless	3.4	0.872
How often were you aware of your current cycle time	I ignored it	I was always aware	5.6	0.678
How did your performance feedback after each build affect your performance?	No affect	Influenced my performance	3.2	0.970
How often did you finish a build within the average cycle time?	Rarely	Always	2.2	1.200
How comfortable are you using the MARS system	Uncomfortable	Very Comfortable	6.0	0.775

(Figure 8-22). This suggests that that the amount of time spent training, which could be explained by the simple interest in the novelty of the interface, has little impact on the actual results.

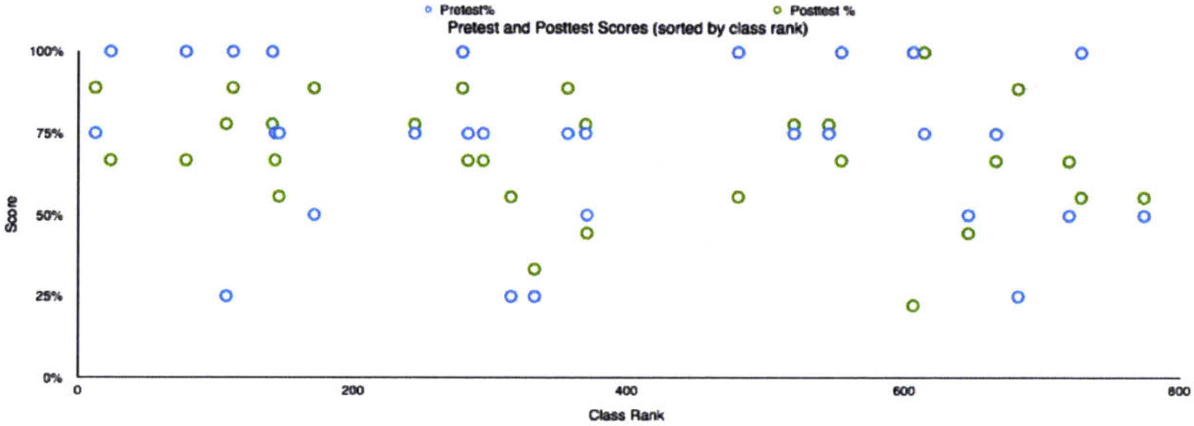
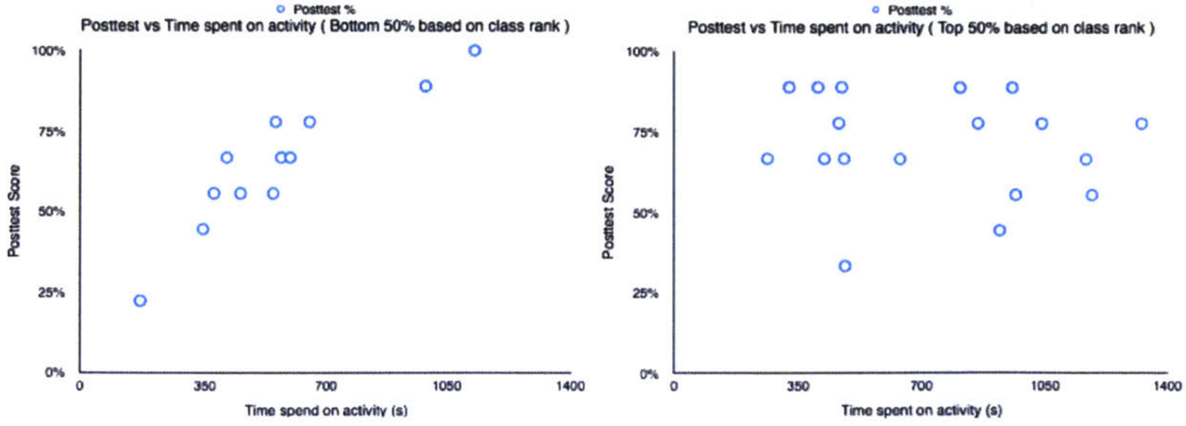


Figure 8-22 Pre-test and Post-test Scores: Each column in the graph represents the pre-test and post-test score of one student. The columns are sorted by the class rank of the student in 8.02.

However, when we focused the analysis on the top 50% of the class, where rank <= 400, we seem to identify a significantly stronger correlation between the post-test score and the time spent on the augmented simulation activity. This may indicate that students with lower rank may have struggled to grasp some of the concepts in the traditional way simulations were presented in class, and potentially had developed a better understanding when presented with the augmented state.



(a) Plot for the lower 50% of the class.

(b) Plot for the top 50% of the class.

Figure 8-23 Post-test scores vs Time spent on activity plots separated by class rank. (a) Plot for the lower 50% of the class. (b) Plot for the top 50% of the class.