Variable Reality: Interacting with the Virtual Book

by

Hye Soo Yang

B.A. Media Arts and Sciences, Wellesley College (2013)

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning, in partial fulfillment of
the requirements for the degree of Master of Science

at the

MASSACHUSETTS INSTUTUTE OF TECHNOLOGY

June 2015

Signature redacted

Author____
Program in Media Arts and Sciences
May 8, 2015

Signature redacted

Certified by____________________________________
Chris Schmandt
Principal Research Scientist
MIT Media Lab
Thesis Supervisor

Signature redacted

Accepted by____________________________________
Patricia Maes
Alexander W. Dreyfoos Professor of Media Arts and Sciences
Academic Head
Program in Media Arts and Sciences

Variable Reality: Interacting with the Virtual Book

by
Hye Soo Yang

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
on May 8, 2015, in partial fulfillment of the requirements for
the degree of Master of Science.

Abstract

This thesis presents Variable Reality, a wearable augmented reality-based system focused on creating a unique on-the-go reading experience that combines the readily accessible nature of digital books with the favorable physical spatiality of a paper book. The two types of Variable Reality books are the *Augmented Book* and the *Virtual Book*. They differ in the way they are displayed as the former augments virtual pages onto an actual book whereas the latter virtually augments a 3-D book on the palm of the user's hand. Designed to take the physical form of a book, Variable Reality books make use of human cognitive ability in storing and retrieving information in a spatial manner. Easy-to-use hand gestures naturally associated with reading activity are integrated with the system to help bring an intuitive user experience in reading.

Thesis Supervisor: Chris Schmandt
Title: Principal Research Scientist, MIT Media Lab

Variable Reality: Interacting with the Virtual Book

by
Hye Soo Yang

The following people served as readers for this thesis:

Signature redacted

Thesis Reader ______________________________
Patricia Maes
Alexander W. Dreyfoos Professor of Media Arts and Sciences
Academic Head
Program in Media Arts and Sciences

Signature redacted

Thesis Reader ______________________________
Steven K. Feiner
Professor of Computer Science
Director, Computer Graphics and User Interfaces Lab
Columbia University

Acknowledgements

I would like to begin by thanking my advisor, Chris Schmandt, who has given me the opportunity to be a part of the Living Mobile Group (previously called the Speech + Mobility Group) since my undergraduate years. Your patience and openness to new ideas always helped me try many new and different things with confidence, which in return enabled myself to come up with an idea that I felt truly passionate about, eventually leading to this thesis work. I cannot thank you enough for making my career at the Media Lab so exciting and so wonderful.

Also, I am sincerely grateful to my thesis readers, Professor Pattie Maes and Professor Steven Feiner, who generously offered insightful advices and feedback on this thesis work.

I want to thank my colleagues, Misha Sra, Cindy Kao and Dhruv Jain for willingly making the time to discuss with me whenever I needed an external opinion. I truly appreciate your kindness, advice and friendship. Special thanks to Misha Sra for introducing me to Virtual Reality and your frequent advice that greatly helped on this research work.

I would also like to thank my mentors and friends, Andrea Colaço and Ahmed Kirmani for generously helping me and guiding me throughout my research career since my undergraduate years. Your support and thoughtful advice have always nurtured and encouraged me.

A big thanks to my friends and colleagues at the Media Lab, a truly amazing group of intelligent people who have done and will continue to produce incredible work. A big thank to my friends, Eugenia Kang, Ashley Kim and Jenny Choi for their unfathomable support and encouragement. Many thanks to Linda Peterson for her generous help in making sure that the students are in the right track and making the right choices. Special note of thanks to Tom Lutz for your patience and help in printing the 3D mounting piece for this project.

And finally, a very special thanks to my parents, my grandparents, my brother, my sister-in-law and my little sister for their endless sacrifice, love and support dedicated to me. You make me who I am today and your words constantly inspire me to improve as a person.

Thank you very much!

❤

Table of Contents

List of Figures

Variable Reality: Interacting with the Virtual Book

by
Hye Soo Yang

Figure: Variable Reality hardware setup and a physical book

1

Introduction

Since the advent of electronic books, the number of worldwide digital books has been increasing at an extremely fast rate, marking an aggregate sum of more than thirty million by 2013 [9]. These electronic reading devices, so-called eBooks, are compact and light in weight and provide the users the ability to retrieve any book from the digital library with just one tap on the screen. Recent advances in display technology enable realistic visual quality for reading that mimics the look of a real page. However, despite the convenient features available on these digital books, studies have shown that the majority of people still prefer the "look and feel," that is, the physicality of a paper bound book rather than pressing the 'next' button on electronic devices [23]. They enjoy the feel of holding a book, posing and flipping through the pages. The mere reading of dry texts on screens is insufficient to account for the entirety of reading experience. Further, the physicality of a book and the way its numbered pages are put together in a sequential order allow users to access pages in a non-linear manner. Accessing information in the book in a non-linear fashion promotes the human spatial cognitive ability, which enables users to rapidly store and retrieve information.

While many still prefer to read paper books as opposed to the digital ones given the positive reading experience it brings, the convenient and portable aspect of digital books cannot be underestimated, especially when the number of books is only going to grow; the ability to quickly sort and access books will become extremely critical in the future. This is the reason why the eBook industry is continuing to grow independent of the negative reviews and criticisms. This is also the reason for myself to own an electronic book despite my unquestionable predilection for paper bound books.

This thesis work is inspired by the need to combine the benefits of a digital book and a physical book. It presents Variable Reality, a novel Augmented Reality (AR) based interactive system focused on bringing intuitive on-the-go reading experiences to the users by enabling always-available access to and presentation of any digital book contents in a complete spatial form of a book. Variable Reality introduces two types of books: the *Augmented Book* and the *Virtual Book*. The *Augmented Book* incorporates computer vision based corner detection and tracking systems to detect the corners of a physical book, information of which is used to overlay paper book pages with digital pages using a Head-Mounted Display (HMD) device. As a result, any real book can instantly turn into a different book while still retaining the physical properties of the book, conforming to the legacy of the paper book metaphor. The *Virtual Book* is an entirely virtual 3-D book that extends the accessibility and mobility of reading materials. When in absence of a physical book, the user can retrieve this virtual book on the palm of a hand –a well-established pose of how one holds a book when reading. The weightless and portable nature of the *Virtual Book* makes it possible for users to comfortably hold larger physical books such as unabridged dictionaries or the World Book Encyclopedia. In both augmented and virtual book scenarios, natural hand gestures inherently associated with the reading activity are defined and used in various Variable Reality applications.

The core contributions of this thesis work include 1) the design and implementation of a system that enables a novel mobile reading experience that combines the physicality or the spatiality of a paper book with the utilities available on a digital one, 2) the design and implementation of software algorithms for detecting and tracking a book and a hand in multiple layers, 3) the implementation of easy-to-learn gestural interaction styles naturally associated with reading and, lastly, 4) the implementation of applications that demonstrate novel navigation techniques and manipulation of books.

1.1 Concept

The basic concept of Variable Reality is essentially based on the idea of enabling users to read a digital book anywhere while maintaining the physical quality of a real book. The Variable Reality system uses a lightweight and compact head-mounted display (HMD) device that is see-through. Two types of books designed for the system are the *Augmented Book* and the *Virtual Book*.

Figure 1-1: Variable Reality Augmented Book concept

In the *Augmented Book* scenario, the user can take any book from the real environment and turn it into a digital book, which can be seen through the HMD (see Fig. 1-1). The book becomes an augmentable surface for overlaying paper pages with digital information without the need of a visual marker for tracking. As a result, it enables the user to quickly switch between reading a paper book and reading a digital book with a single physical book. The title of a book is detected using a feature detector and the page number can be detected with the help of the OCR technology. This enables the user to store and retrieve digital data on a specific physical page. The *Augmented Book* also integrates natural hand gestures associated with reading activity in order to provide an intuitive reading experience. A fingertip is predominantly used as a mouse cursor but is used as a pen tip for annotations.

Figure 1-2: Variable Reality Virtual Book concept

Ideally, the *Virtual Book* makes it possible for users to hold a 3-D virtual book in the hand and read it anywhere. It is designed to bring extended mobility and accessibility to the digital library designed for on-the-go reading experience. The user does not need anything other than the hand to bring up a virtual book on the palm (see Fig. 1-2). Similar to the *Augmented Book*, it integrates the hand gesture recognition system, which enables users to efficiently manipulate the book.

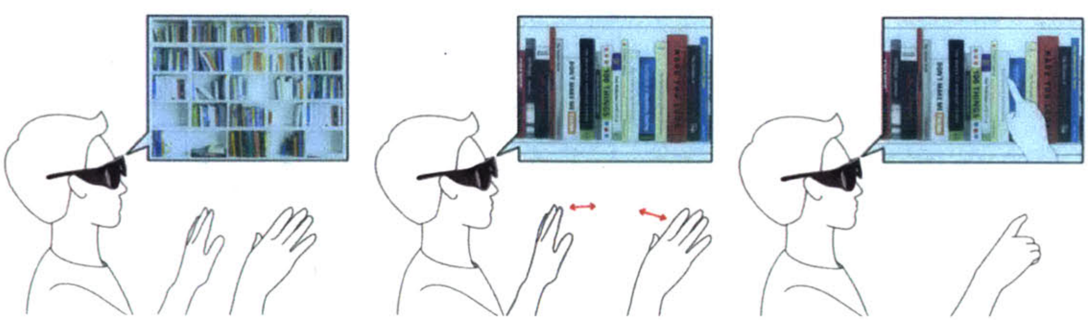

Figure 1-3: Variable Reality Virtual Bookshelf concept

Variable Reality further incorporates a three-dimensional digital bookshelf that closely resembles the look of a real bookshelf (see Fig. 1-3). Using a simple gesture, the user can display it, zoom into a portion of it and pick out a book to read. This book can be in the form of either the *Augmented Book* or the *Virtual Book.*

1.2 Thesis Overview

Chapter 2 covers the background and related work, including the concept of physical 'affordances' of an object and the physical metaphors particularly geared towards a paper book. It also touches upon the potentials of human spatial cognitive ability that, when used, can significantly enhance the overall reading experience. Next, it discusses several research fields that explore ways to bridge the gap between the physical and the digital world. The chapter ends by describing a number of book related projects.

In Chapter 3, the Variable Reality systems are introduced in more depth along with a variety of interaction styles designed for both the *Augmented Book* and the *Virtual Book* scenarios. It also includes a technical explanation of how these systems are implemented.

Chapter 4 presents various applications designed for the two Variable Reality books. Chapter 5 discusses the user study conducted to gauge the effectiveness of the system software and to gather user experience results. Limitations manifested in the process of building and testing the system are discussed in Chapter 6. It also talks about the attempts made to overcome the handicap and possible future directions of the project. The final chapter summarizes the thesis work and concludes the document.

2

Background and Related Work

Living in the era of high material productivity, people are constantly surrounded by numerous objects (e.g. a watch, a book, clothes etc.) that they use and interact with on daily basis. Often they are manufactured goods that have a specific purpose for use meant to provide convenience to the users. Over the years since the birth of consumer electronics, these everyday objects have been transforming into digital artifacts that enable utilities and functionalities that are traditionally not afforded by physical objects. An electronic book is one of them; it is portable and grants the users a quick and easy access to the massive digital library through a single device. Unsurprisingly, when designing such a device, the creators and interface designers greatly consider the physicality of the object the device is trying to replace because it can help create a better user interface.

This chapter begins by discussing the physicality of a book and the 'affordances' [25] it provides to the users. Variable Reality takes into account what it means to hold a physical book and flip through the paper pages for a reading experience. The chapter also touches upon the popular paper book metaphor that is analogous to the desktop metaphor and the need to incorporate the concept when designing a user interface. It includes an explanation dedicated to the human's exceptional cognitive ability in spatially storing and retrieving information. Several research fields that are related to integrating real objects and environment with the digital information are discussed next. Lastly, it presents book related projects from the past that demonstrate the effective exploitation of the paper book metaphor.

2.1 Book Affordances and Paper Book Metaphor

The term "affordance" was invented by JJ. Gibson [25] and refers to the performable properties provided by the object that the user can perceive through its forms and shapes. For example, a mug affords the user the ability to hold it using the handle. In this case, the mug affords holding (see Fig. 2-1 (left)). Often Gibson uses '-able' suffix to describe the performable actions on objects such as 'climb-on-able' and 'fall-off-able' for terrestrial surfaces. Similarly, paper books have their own affordances that they offer to the users. Using Gibson's suffix, books are hold-able, open-able, flip-able, turn-able, write-able and so on (see Fig. 2-1 (right)). These affordances and physical properties of a book are important factors that contribute to the user's overall reading experience.

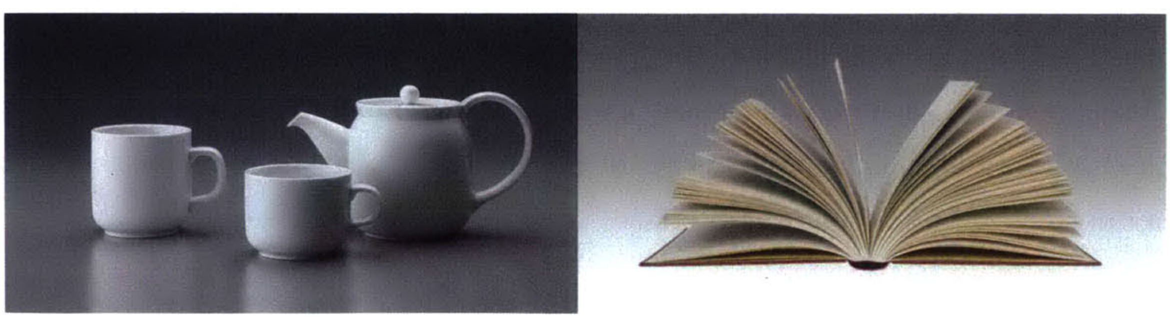

Figure 2-1: Object affordances; a mug (left) and a physical book (right)

The paper book metaphor is closely related to the affordances of a book in that it refers to the idea of embedding the physical properties and aspects of a book in an alternative system or device. Similar to how the desktop metaphor was employed by interface designers when designing a digital desktop workspace, today's digital books are an outcome of the inventor's efforts in adhering to the paper book metaphor. Wilson's work in [1] demonstrates the importance of the paper book metaphor and how each page should be regarded as a distinct "visual space." Complying with these metaphors can help engineers build an intuitive and natural user interface in which users can quickly and effortlessly learn how to interact based on long established behaviors associated with the object through practice [8]. Variable Reality systems are built based upon this physical metaphor and enables seamless user-interaction.

2.2 Human Cognitive Ability and Spatial Memory

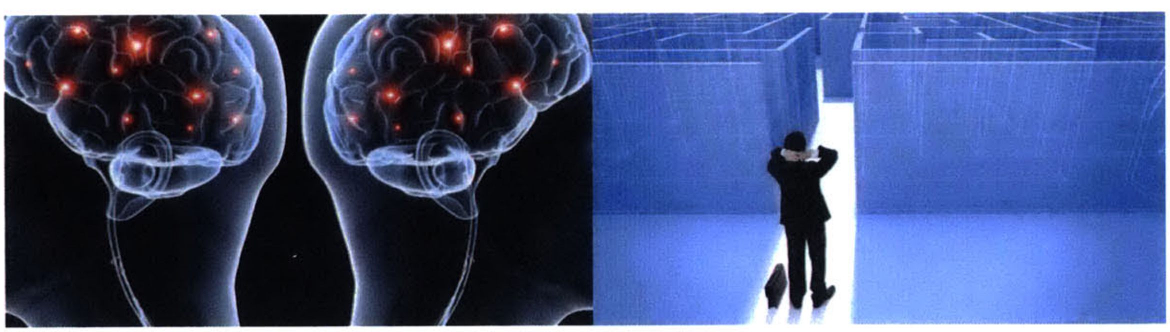

Figure 2-2: Human spatial memory

Spatial memory is one of many cognitive abilities that humans possess. It is critical for remembering information about a spatial setting relative to one's position and orientation. Studies have shown that the spatial cognitive ability enables rapid storage and retrieval of data in graphical user interfaces [26]. The way human location memory is formed and processed is essentially effortless whereas remembering and manipulating abstracted representation of information requires time and effort [30]. The desktop metaphor attempts to take advantage of this spatial memory by allowing users to organize information in a spatial manner. However, the 2-1/2-dimensional workspace on screen is still limited for users to perceive it as a three-dimensional space.

Figure 2-3: Spatial Data Management System (SDMS)

Unlike the spatially constrained user interfaces, the **Spatial Data Management System (SDMS)** by the Architecture Machine Group was built using an entire media room as an experimental setup that enabled users to spatially organize digital data and retrieve it efficiently [28] (see Fig. 2-3). The setup included a large back-projected display that simulated realistic two-dimensional space, generating an "information surround" [28]. Similarly, Robertson et al.'s **Data Mountain** [24] attempts to take advantage of the spatial ability by allowing the users to easily manipulate browser windows in a 3-D graphical user interface.

Figure 2-4: An eBook (left) and a literal manifestation of the paper book metaphor (right)

A physical book stores information just as desktop folders do but in a spatial and sequential order. This makes it possible for users to access information in a non-linear fashion, which facilitates search. Further, because the order flows from the front to the back, the user has a visual and haptic feedback on how far along the book one is on. As studies show, this "fixed layout" of paper pages triggers spatial memory, which, in turn, allows effective data retrieval [14]. It is also known to help students digest and remember materials better [6]. However, current electronic books are analogous to the 2-1/2-dimensional user interface in that their "perceived affordances" (e.g. location indication bar, clickable icons) [31] require additional effort and practice to process the information (see Fig. 2-4 (left)). Variable Reality is designed to make use of human spatial cognitive ability with a strong emphasis on the adherence to the paper book metaphor.

2.3 Bringing Together Physical and Virtual World

The need for a more user-friendly and intuitive user interface led to the emergence of a number of research fields intended for seamlessly converging the physical and the digital world. Augmented Reality (AR), Tangible User Interfaces (TUI) and Tangible Augmented Reality (Tangible AR) are among them just to name a few. Pioneers in each field have produced breakthrough work that not only served as an inspiration to Variable Reality but also as system enabling technologies.

2.3.1 Augmented Reality

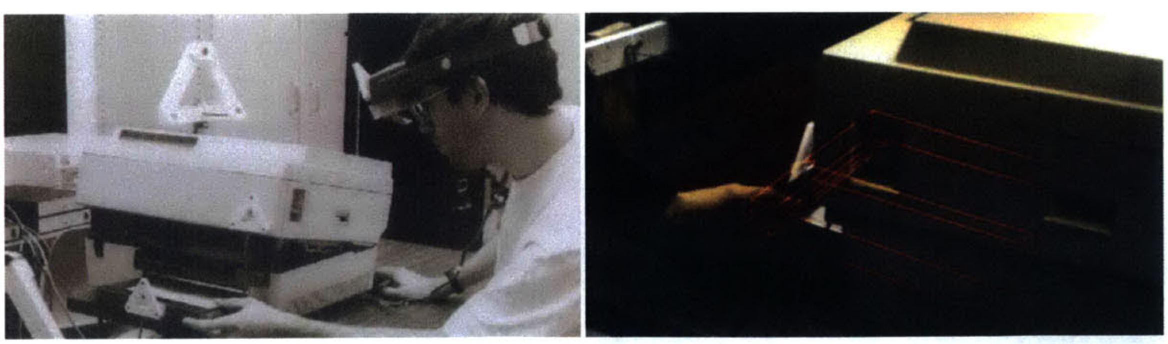

Figure 2-5: Knowledge-based Augmented Reality

Augmented Reality (AR) aims to combine the real scene view and a computer-generated scene by superimposing additional 3-D information over the elements in the real environment. It was first demonstrated in 1968 by Ivan Sutherland using a transparent head-mounted display (HMD), which allowed the user to see the combined view of the physical world and the digital imagery that was produced from miniature CRT displays [4]. 1992, AR pioneers such as Steven Feiner, Blair MacIntyre and Doree Seligmann developed **Knowledge-based Augmented Reality** [35], a breakthrough work that guided untrained people to load paper into the printer without the need of an instruction manual (see Fig. 2-5). In the later years, the AR toolkit provided by Hirokazu Kato enabled a stable tracking system based on a physical marker.

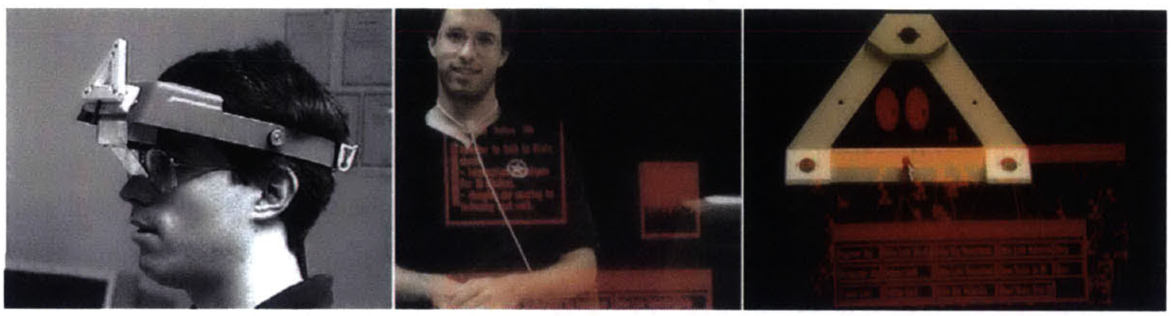

Figure 2-6: Windows on the World

See-through HMDs make it possible to easily augment (or overlay) digital contents onto the physical environment. In Feiner et al.'s **Windows on the World** system [27], the user can easily create links between different windows, which can be attached to real objects in an office (see Fig. 2-6). Objects can be tracked for the attached windows to stay with them.

Figure 2-7: Mobile Augmented Reality System (MARS)

The wearable nature of an HMD device makes it well suited for both indoor and outdoor mobile use cases. The **Mobile Augmented Reality System (MARS)** [36] built by Hollerer et al. enables users to manage spatially oriented information in the real environment (see Fig. 2-7). The real-time-kinematic GPS system integrated with the MARS system is capable of detecting the user's centimeter-accurate location and orientation. As a result, the user has the freedom to digitally annotate and leave real-time traces on the real environment on the go. Sinem Guven et al. developed MARS Authoring Tool designed for creating 3D hypermedia narratives using a wearable device that turns the user's surrounding environment into a narrative book [34].

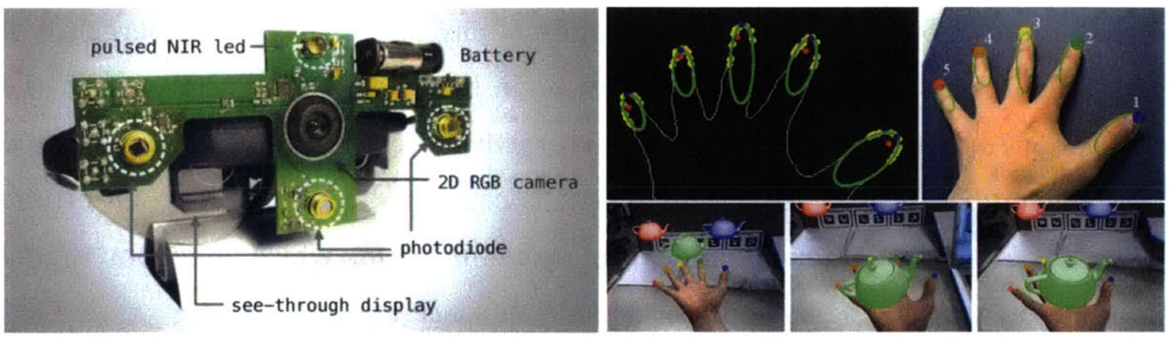

Figure 2-8: Mime sensor (left) and Handy AR (right)

While HMDs are great for seamlessly merging digital information with the physical space, its limitation lies in its input mechanism, which had traditionally been available through external controllers such as keypads, keyboards or mice. In an attempt to overcome this handicap with HMDs, Thad Starner et al. integrated RGB-camera-based hand tracking system to a networked wearable computer in which a fingertip served as the system's mouse cursor [29]. Similarly, **Mime** [37] developed by Andrea Colaco et al. manages to transport the entire spatial volume around the HMD device into an input generable space by using a clever algorithm to extract out millimeter-accurate 3-D hand position (see Fig. 2-8 (left)). **Handy AR** [32] by Taehee Lee et al. turns the palm of a hand into an augmentable surface on which 3-D objects can rest (see Fig. 2-8 (right)).

2.3.2 Tangible User Interfaces

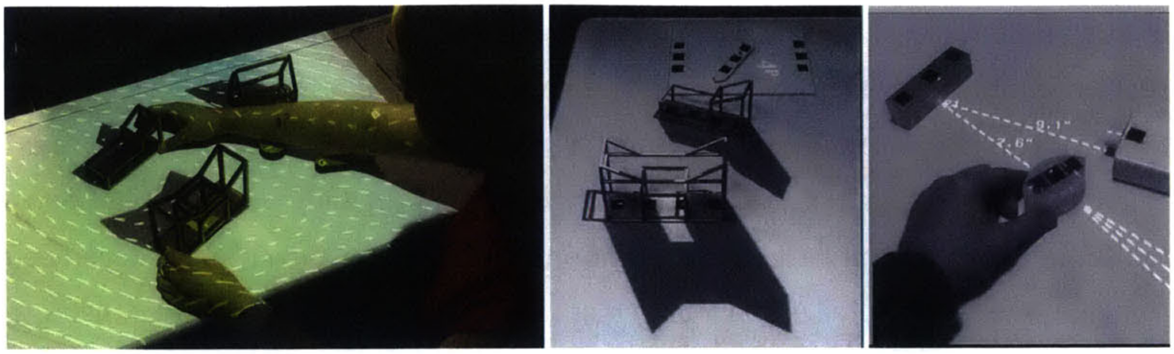

Figure 2-9: Urp

Tangible User Interface (TUI) is a user interface that allows the user to interact with digital information through the manipulation of physical objects. The emergence of TUI is dedicated to the pioneering work done by Hiroshi Ishii, who created the concept of *Tangible Bits*, which refers to interactive digital information manipulated by changing physical objects or surfaces it is connected to. Projects based on the framework of TUI largely employ projectors in setup, which augment digital information onto a designated physical model or space.

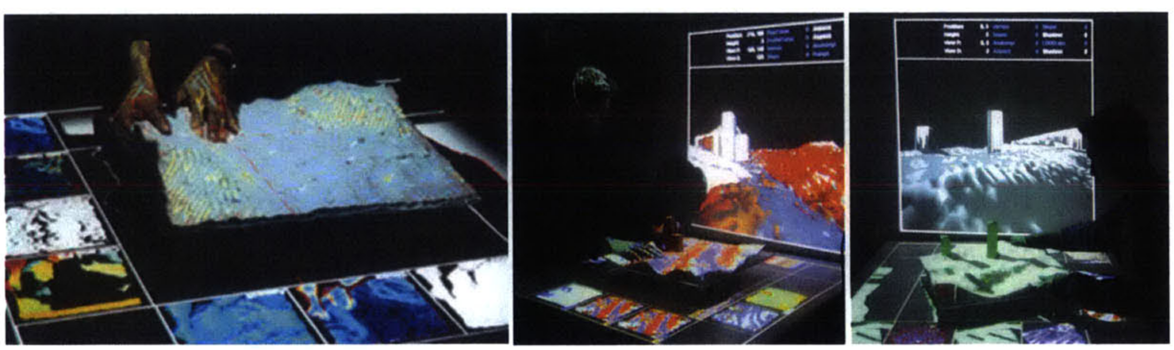

Figure 2-10: Illuminating Clay

Urp [38] is a system designed for urban planning, which uses physical architectural models placed on an ordinary tabletop and a projector. Upon the manipulation of these architectural models, the system computes the effect (e.g. the wind velocity between adjacent buildings) in real-time and redisplays the updated result (see Fig.

2-9). **Illuminating Clay** [40] similarly performs a real-time mathematical analysis of landscape models using a laser scanner. It is capable of capturing the surface deformation of the clay model placed on the (see Fig. 2-10). **I/O Brush** [39] likewise deploys an ordinary physical object, a brush, which can capture images, patterns and videos of the real world using an embedded camera inside the brush (see Fig. 2-11). Just as another paintbrush, it takes colors from the real objects and environment and paints them onto the digital canvas.

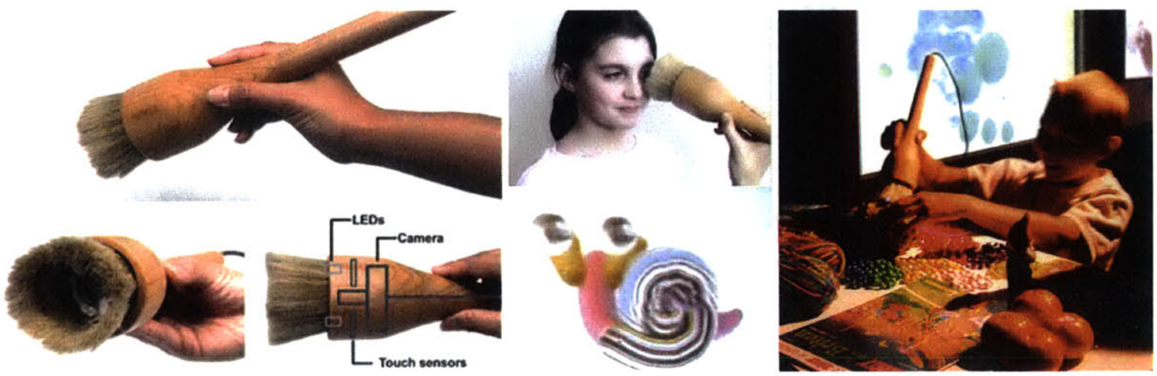

Figure 2-11: I/O Brush

In TUI, physical objects serve as the window to controlling digital contents and, therefore, the affordances inherent to the objects are present for user interactions. As a result, the user can easily adapt to the systems like the **I/O Brush** since one already has an understanding of how to paint with a brush from the real world practices.

2.3.3 Tangible Augmented Reality

While AR systems seamlessly integrate digital information with the real world, the physicality of the objects and their implied affordances that help create natural user interaction tend to have less emphasis. On the other hand, TUI systems provide the users with intuitive ways of physically interacting with the digital world but the experience is limited to the designated space and objects.

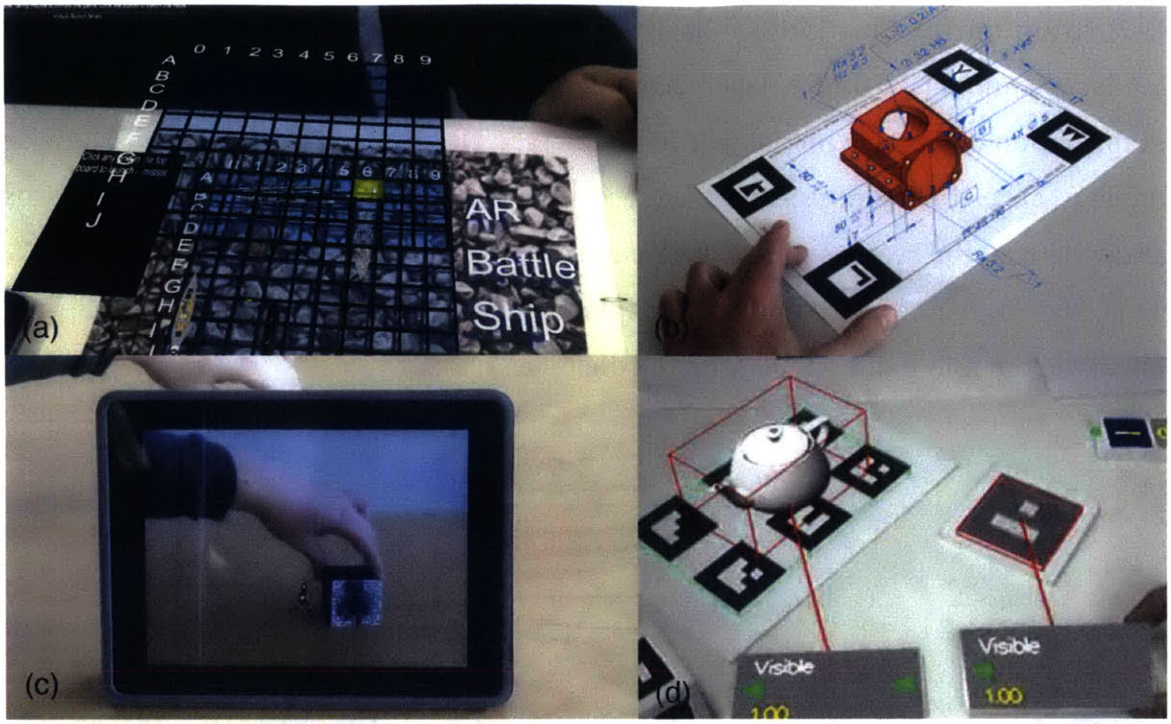

Figure 2-12: Examples of Tangible AR projects: (a) Tangible AR Battleship Game; (b) Tangible Interface for Augmented Engineering Data Management; (c) AvatAR; (d) Immersive Authoring of Tangible Augmented Reality Applications

Tangible Augmented Reality (Tangible AR) attempts to combine input methods prevalent in TUI with AR display techniques. The goal of Tangible AR is to link virtual contents or functions to the real-world objects and use these objects to physically manipulate the digital information. It was first introduced by Billinghurst et al. in 2008 as a concept of a metaphor. It emphasizes the need for a seamless interaction and display between the physical and the virtual world [7]. The design principles for Tangible AR interfaces state that 1) each virtual object is registered to a physical object and that 2) the user interacts with virtual objects by manipulating the corresponding tangible objects [7]. For example, the immersive authoring tool [33] and the augmented data management system [45] are Tangible AR based projects that enable the users to author and manage virtual information by directly manipulating of physical objects within the AR environment (see Fig. 2-12 (d, b)). In Tangible AR Battleship Game [43], a physical paper becomes a game interface with which the user can make inputs using a hand (see Fig. 2-12 (a)). **AvatAR** developed by Fernandez-Baena et al. is another project that uses a physical cube for controlling avatars interacting in the real environment [17] (see Fig. 2-12 (c)).

The concept of "physical metaphor" [5] and the paper book metaphor are closely related to Tangible AR as these metaphors refer to removing the *functional* and *cognitive seams* found in traditional AR systems [7]. Variable Reality systems stick with these Tangible AR design principles and use a physical book or a hand as the surface for augmenting (or registering) and manipulating digital book contents.

2.4 Related Work

Many book related projects had been produced in the recent past that are aimed at bringing better reading experience to the users. The paper book metaphor is often discussed and applied to the project systems. For example, the **Dual-display E-book Readers** [2] developed by Chen is an implementation that directly achieves the paper book metaphor; the hardware design mimics the dual-page characteristic of a physical book. More closely related projects to Variable Reality are presented next in four categories: augmented book, virtual book, mixed reality book and the rest.

2.4.1 Augmented Book

Using augmented reality markers to overlay virtual contents over the physical book pages has been frequently explored. Billinghurst's **MagicBook** [10] is built based on the idea of smoothly transporting users from the augmented book environment to the immersive Virtual Reality setting (see Fig. 2-13). **MagicBook** is a normal paper bound book some of whose pages have AR markers. The augmented virtual contents can be seen through a handheld display device. They follow the position and orientation of the marker. The user can seamlessly fly into an entirely virtual environment that is an extended version of the virtual contents shown on the page. With the **MagicBook**, the user can experience the story in the book in a highly immersive fashion. Furthermore, the **MagicBook** supports collaboration among people by allowing multiple users with a handheld display device to experience the story in a shared virtual setting.

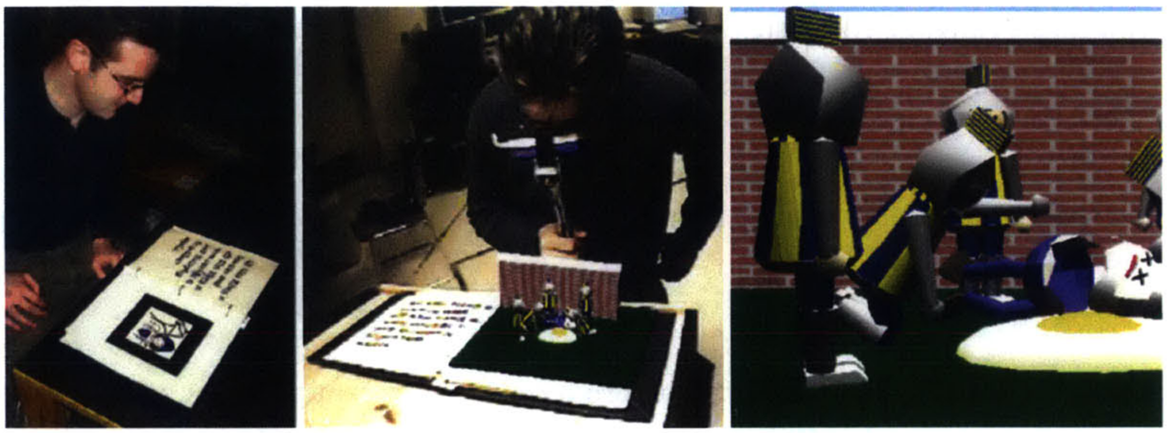

Figure 2-13: The MagicBook

WikiTUI [12] is an over-head projection based system that overlays digital information onto the pages of a physical book. The goal of the concept project is to bridge the gap between the physical and the digital world within the scope of reading. By using the book surface as the interface, the users can enjoy ordinary reading experience while having the capability to display additional useful information on the side. The setup requires an overhead projector directly looking down at the table on which a real book rests. The fingertip is used as a mouse cursor for the interaction. **WikiTUI** also supports collaborative annotation using the online wiki technology.

The **Haunted Book** [18] is an illustration book designed for artistic story telling. It uses computer vision techniques to perform marker-less AR tracking through the feature detection method. As a result, the illustrated animation augmented on the page has a natural feel that does not distract from the reading experience. The **Haunted Book** setup requires a lamp that has a built-in webcam, which looks down at the book for detecting book features to track. The animation is displayed on the computer screen placed right next to the book.

Augmented reality books also have been largely developed for educational purposes. The use of instructional 3-D contents in learning is known to help on students' understanding. **Live Solar System (LSS)** [11] is one example of educational augmented reality books designed to help students learn Astronomy. The user study conducted on the system has shown that **LSS** is easy to use and has a positive impact on student's overall understanding of the material. For complicated procedures such as surgical tasks, it is often difficult to train students due to limited resources. To alleviate such adverse learning environment, Welch et al. invented the **Immersive Electronic Books (IEBook)** [42] for surgical training. The goal of the project is geared towards providing surgeons the opportunity to witness and experience previously done surgical procedures in 3-D with the help of augmented reality technology.

2.4.2 Virtual Book

Figure 2-14: The Web Forager (left) and the WebBook (middle, right)

Organizing digital information in a form of a physical book has long been attempted particularly aimed at promoting spatial memory in order to enables rapid storage and fetching of information. In 1978, **Pages Without Paper** [22] developed by Schmandt introduced a virtual book display as a part of the **SDMS** [28]. In the system, high quality fonts were displayed in a book-like format on a mini-computer driven raster scan display and a realistic page turning animation was played upon the user's fingertip control. The **WebBook** [16] organizes a collection on web pages in a 3-D form of a book (see Fig. 2-14). A page turn is animated to create a realistic feel to the virtual system. One of the several methods designed to flip the **WebBook** pages is based on the relative distance of the mouse cursor from the edge of the book. This navigation technique allows the user to quickly scan through the contents. This facilitates a quick visual search through the web pages in the **WebBook**. The **Web Forager** [16] is a larger and more inclusive entity that contains multiple **WebBooks** in a three-dimensional virtual workspace. Similar to the **Data Mountain** system introduced in Chapter 2.2, the **Web Forager** ensures the effective organization and management of the **WebBooks**.

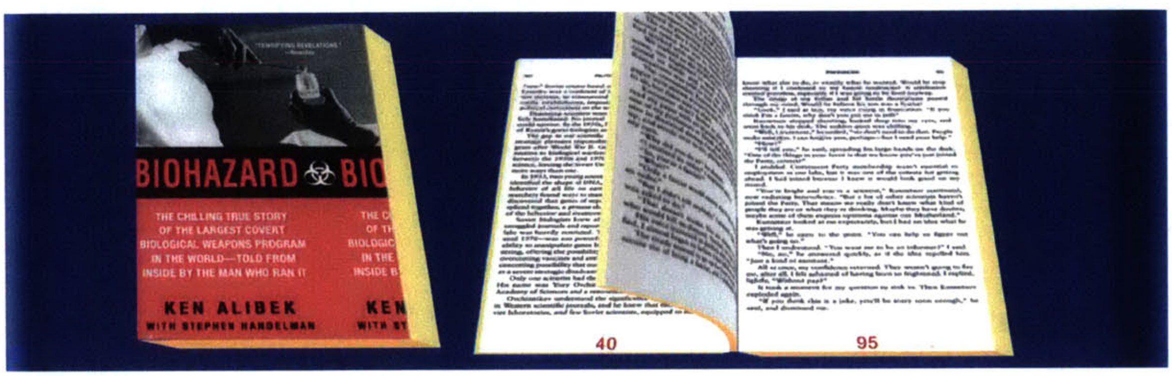

Figure 2-15: The 3Book

The **3Book** [21] is also an entirely virtual book that takes the complete physical form of a real book (see Fig. 2-15). While maintaining the appearance and the form of a physical book, **3Book** enables convenient functionalities traditionally unavailable on paper books such as smart indexing, automatic zooming, and cut-and-paste extraction tool just to name a few. To facilitate user interaction, it also provides the so-called fisheye visualization technique with which the user can use to focus on specific areas within the page. Another version of **3Book** development is the scalable **3Book** [19] that can represent any book independent of its length and size. This allows users to read any book with any size. It affords the users with the means for leaving bookmarks or highlights on a page, the mechanism for comparing multiple pages and the capability to excerpt a portion on a page out to an external editor.

2.4.3 Mixed Reality Book

Beyond what augmented books and virtual books offer, mixed reality books additionally incorporate auditory feedback that extends the scope of user experience in reading. Grasset et al. developed a unique children's illustration book [3] that offers an audio-visual interactive user experience. It provides animated virtual 3-D contents, interactive features and sound feedback that are associated with storytelling contents of a page (see Fig. 2-16).

Figure 2-16: Children's illustration book by Grasset et al.

The mixed reality book developed by Grasset et al. [41] is similar to the above-mentioned interactive book but with an improved sound system that brings spatial sound rendering to the system that allows the users to feel a part of the story presented in the book. Moreover, it provides the option for the user to take presence in the story as a live video character. Both the children's mixed reality edutainment book [3] and the interactive mixed reality books [41] require a handheld display device.

2.4.3 Others

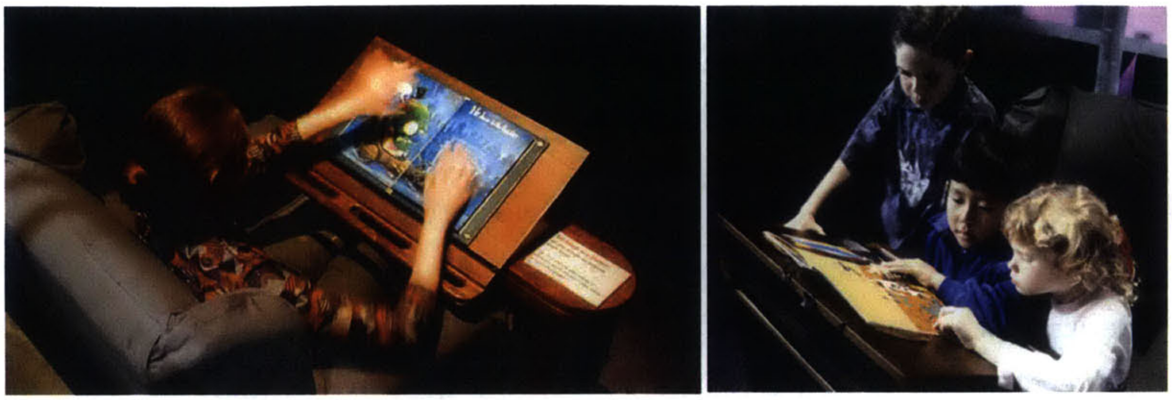

Figure 2-17: Listen Reader

There are several book related projects that are not necessarily based on either augmented or virtual reality but that nonetheless are effective systems that adhere to the paper book metaphor.

Listen Reader [20] is an electronically augmented paper-based book that combines rich quality of music and sound effects with unique physical location on pages in a real book (see Fig. 2-17). **Marginalia** [13] is a digital annotation system that extends the margins of a paper book on both sides (see Fig. 2-18). Using an electronic pen, the user can click on icons that are displayed on the electronic margin near the contents on the physical page. Individual pages are networked meaning that the sharing of digital media can easily be achieved.

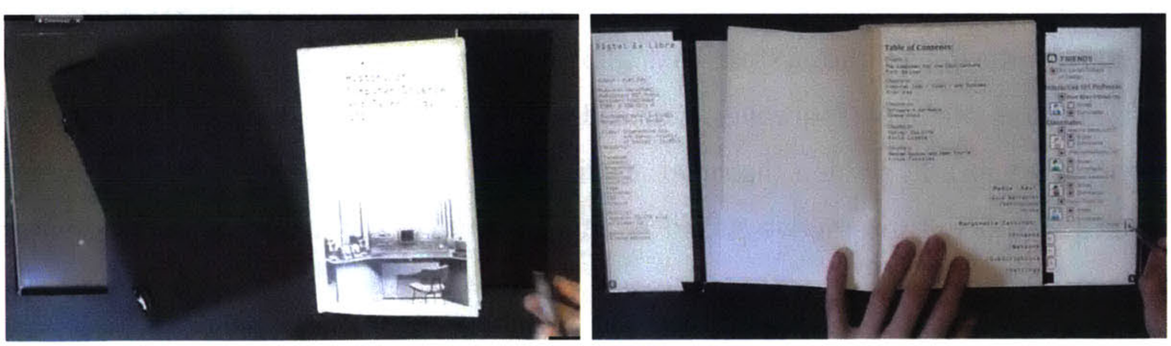

Figure 2-18: Marginalia

All of the book related projects presented in previous sections manages to create positive reading experience by combining the physical look and feel of a book with digital bits. Augmented books and mixed reality books are effective particularly in creating effortless user interaction by using physical surface of an ordinary book as the base for augmenting digital information. As a result, it eliminates the need for the user to newly learn how to interact with the system. However, the challenge lies in the fact that these unique augmentable pages require a visual marker or predefined features for stable detection and tracking. As a consequence, **MagicBook**, **LLS**, **IEBook** and the discussed mixed reality books [3, 41] need a specially designed book that has pages with unique markers that provide visual cues for displaying virtual contents in the correct orientation. Creating this custom-made book that is a couple hundred pages in length will be painful. Making a library of these books will not only be more painful but also be extremely inefficient, if not impossible. The **Haunted Book** performs marker-less tracking by detecting features in the illustration book, which also have to been predefined by the system designer. On the other hand, **WikiTUI** system requires an overhead projector in a fixed setting, which contradicts the portable nature of a book. Virtual books such as **WebBook** and **3Book** manage to create three-dimensional books that have the realistic appearance of a physical book. Users are provided with perceived spatial cues for their page location upon reading or for organizing a book in the 3-D repository, the **Web Forager**, for example. These visually perceived spatial cues on screen, however, are still limited to the conventional input mechanisms (e.g. clicking or dragging the mouse, typing on keyboard), which are distant from how one interacts with a physical book while reading.

Contrary to the marker based augmented book applications like the **MagicBook**, **LLS**, and **IEBook**, Variable Reality does not need markers or specific features to track because it uses the contour of the book shape as the marker. As a result, the user can seamlessly switch between reading a paper book and a digital book while maintaining the physicality of a real book. Further, it allows the users to hold an entirely virtual book on the hand, a long-established posture for reading. The users are able to interact with the 3-D book with quick and easy hand gestures that are associated with reading activity. Next two chapters introduce Variable Reality and its systems in more depth.

3

Variable Reality

Today's digital book technology enables realistic visual display for reading, ensures instant access to millions of books with just a few clicks and provides functionalities such as smart search that are not possible with traditional paper books. However, the majority of people still prefer to read from paper books. Anita Li reports in her article that studies have shown that a preference for physically turning pages over swiping "next" on screens may, in fact, help readers thoroughly digest what they read [44]. This conclusion was derived from Garland (a lecturer at the University of Leicester)'s study, the findings of which showed participants were much quicker in transforming information from *"remember"* to *"know"* when reading paper than when reading screens [44]. Further, Li includes quotes from Daniel Wigdor, a computer science professor at the University of Toronto, regarding this study outcome:

> *"The current e-reader design is all about removing the physicality, and it's with that physical interaction that you gain that experience and apply spatial*

memory. The basic thing we're trying to get to is active reading, where you're not just visually engaged with the task, but physically engaged with reading."

The comments from Wigdor imply that the significance of the physicality (or the "look and feel") of a book is closely tied to enabling a positive reading experience.

Variable Reality is an augmented reality-based interactive system aimed at enhancing reading experiences by allowing users to take advantage of human spatial memory by emulating paper books. By combining the physicality of a book with the functionalities of a digital book, it offers convenient features available in digital books while maintaining key "affordances" of real books. The system deploys markerless tracking, which enables the user to augment any chosen digital book over any paper book. To promote mobile use cases, the system provides two different types of books: the *Augmented Book* and the *Virtual Book.* A number of quick and natural user interaction styles are designed to enhance the overall experience and convenience of reading. This chapter introduces the Variable Reality system in depth, including the implementation techniques, hardware setup, interaction styles and early prototypes.

3.1 Introduction

Variable Reality is a wearable, augmented reality-based interactive system focused on bringing intuitive, on-the-go reading experiences to users by enabling always-available access to and presentation of any chosen digital book's contents in the complete spatial form of a book. With a special attention to mobile use, Variable Reality takes on two different forms of a book: the *Variable Reality Augmented Book* (or the *Augmented Book*) and the *Variable Reality Virtual Book* (or the *Virtual Book*). The *Augmented Book* is an augmented book designed so that the user can open up any real book and read a different book using the body of the physical book by digitally overlaying information. The *Virtual Book* is a virtually augmented book that allows the user to pull any chosen book into one's hands just as the person would hold a paper book. As a result, the user does not even need a book to read, making the system more mobile than a digital book. The system is also designed to detect a number of quick and easy hand gestures connected to reading activity.

In both versions, the system attempts to take advantage of the spatial memory. The *Augmented Book* directly achieves this metaphor by incorporating a physical book and, thereby, enabling interaction styles naturally associated with reading. Additionally, the augmented virtual page is animated, following the physical turn of a page in order to further enhance the user experience. The *Virtual Book* metaphorically complies with the concept by allowing users to hold a realistic 3-D book and pose just as one would hold a real book when reading. Similar to the *Augmented Book*, it animates the turning of a page when the page-turn gesture is detected. The dual design aspect of the Variable Reality system makes it possible for users to open multiple books in the interactive space around one's body which is a feature not available on a digital book. Additionally, the system incorporates a virtual bookshelf that contains the user's personal collection of books.

3.2 Implementation

Designed to support mobile user scenarios, Variable Reality uses a minimal setup for creating an augmented reality environment and for integrating natural hand gestures. The unique computer vision-based tracking technique for tracking a book is the key system-enabling factor. In the following subsections, the specifications for hardware, mechanical design and software processing implemented for the Variable Reality systems are explained. The interaction styles built based on these technologies are described next.

3.2.1 Hardware

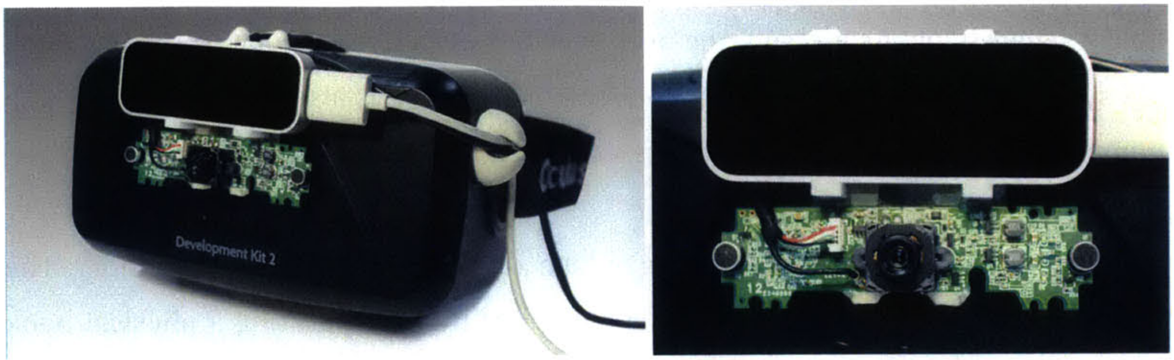

Figure 3-1: Variable Reality hardware setup

In order to create an immersive augmented reality environment with hand gestural controls, Variable Reality system uses Oculus Rift DK2 for display, a high-resolution webcam for capturing the real worldview and a Leap Motion Controller for acquiring 3-D hand positions (see Fig. 3-1).

Oculus Rift DK2

Figure 3-2: Oculus Rift DK2 and an example of the display (bottom right)

The Variable Reality system is built using Oculus Rift DK2 [52], which has a display resolution of 960 × 1080 per eye. It is one of the popular head-mounted display

(HMD) devices largely used for virtual reality applications and gaming. Its large-enough field of view allows users to experience the virtual world in an immersive manner (see Fig. 3-2).

Logitech HD Pro Webcam C920 Web Camera

For the Variable Reality system, a compact yet high-resolution webcam (Logitech HD Pro Webcam C920 Web Camera [53] (see Fig. 3-3 (left))) is mounted at the front of the HMD device at the same height as the user's eye level. Bringing the webcam's real-world scene view to the Oculus Rift displays can create an immersive, augmented reality environment. The webcam has a resolution of 1920 x 1080 and works at 30FPS. It includes an autofocusing feature that allows clear display of the page text and content by adjusting the lens focus on the book. The camera frame is also used for multiple purposes within the Variable Reality software pipeline (e.g. acquisition of a book or hand positional information).

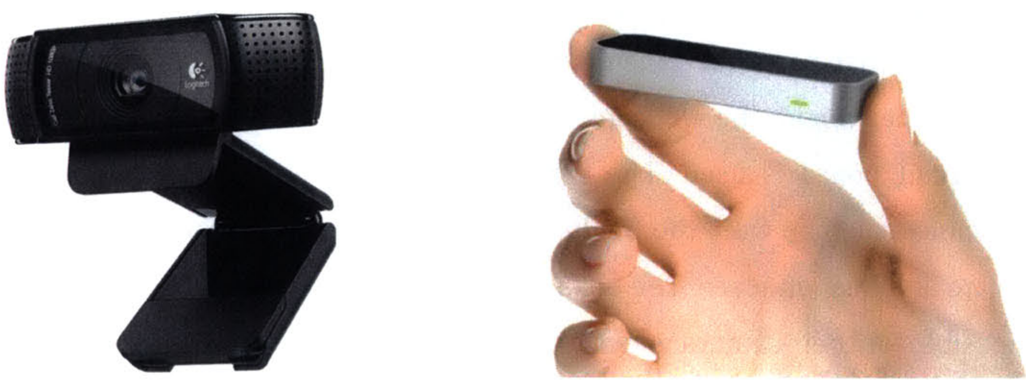

Figure 3-3: Logitech HD Pro C920 webcam (left) and the Leap Motion (right)

Leap Motion Controller

Fig. 3-3 (right) shows a Leap Motion Controller [54], which is mounted slightly above the webcam on the front of the HMD device. This 3-D sensor is mainly used to detect hand positions and finger joint information for hand gestural interaction. It is composed of two monochromatic IR cameras and three infrared LEDs. The video frames from the device's infrared light camera are occasionally used for book tracking when book tracking with the webcam fails due to a bright or cluttered background.

3.2.2 Mechanical Design

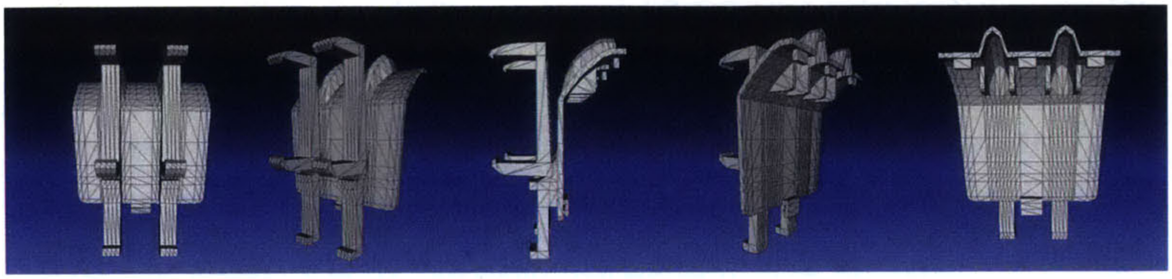

Figure 3-4: The 3-D mounting piece model

Because the system's setup requires the user to wear a device on the head, minimizing the additional weight of the device incurred from attaching the cameras was important. For this reason, a 3-D mounting piece created from a lightweight material was designed identical to the removable front cap of Oculus Rift but with some additional parts for placing the cameras (see Fig. 3-4). To most efficiently minimize the weight, the webcam's packaging has also been removed, and only the bare circuit board is used. The current Variable Reality setup has replaced the removable front cap from the Oculus Rift and replaced this mounting piece with the webcam and Leap in place. The Leap and webcam are placed as close to one another as possible for their camera views to overlap.

3.2.3 Software

In order for the Variable Reality system to reliably work in various different settings, unique tracking techniques were devised using the two cameras. Tracking systems can be divided into two main categories: *book-tracking systems* and *hand-tracking systems*. Each of these tracking systems can be further divided into two categories based on the camera device upon which it relies. Only one book-tracking system and one hand-tracking system can work at a time. The system is designed to automatically switch the tacking mode between "RGB" and "IR"; RGB is designated for the RGB webcam use and IR is designated for the Leap IR camera use. While both devices operate simultaneously and continuously, the tracking mode determines which camera frames are to be processed. In this way, the devices can supplement one another based on the environmental setting. For example, the IR camera can be used when RGB book tracking fails due to a white background. The benefits and disadvantages of each camera and the tracking systems are discussed in more detail below.

Book-Tracking Systems

RGB-Based Book Tracking

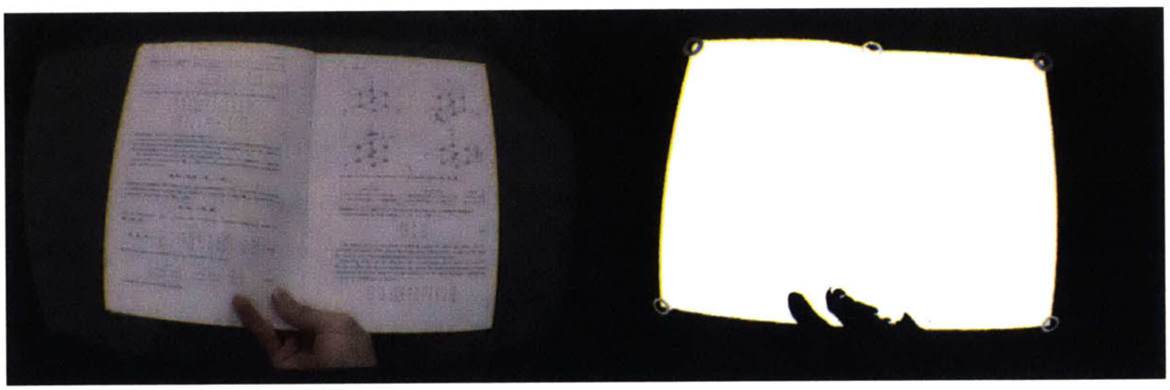

Figure 3-5: RGB camera view (left) and a thresholded image with tracked book corners (right)

Variable Reality book tracking primarily uses the RGB webcam feed (see Fig. 3-5 (left)) for the *Augmented Book*. It first looks for objects with bright color (as bright as ordinary white paper). If detected, it then turns the frame into a binary image by applying a threshold to filter out unnecessary objects. If the detected object mesh has a reasonable size for it to be a good book candidate, then it processes the Harris corner detection algorithm (available in OpenCV SDK [51]) on the frame to see if any corners exist.

> The *Harris corner detector* finds corners in a given image first by looking at the gradient of the image, which is a representation of the change in the image in both x and y direction. Since a 'corner' is a point in which two non-parallel

edges meet, it has a high gradient (or variation). Based on the scaling and threshold parameters, the detector gives a score to each detected point. The detector determines whether a detected point is a 'corner' based on the assigned score and outputs points with higher scores as detected 'corners'.

Normally, more than a few dozen corners get detected due to its sensitivity, and this calls for filtering. Upon filtering these corner data, the most-bottom-left corner of all the detected points acts as the most important corner point upon which detection of all the other book corners depends. The reason for using this corner as the anchor point lies in the fact that it is more easily detected than any of the other corners. It is also hard to loose once detected because users normally hold a book in the left hand and flip with the right hand, leaving this bottom-left corner intact.

If the object mesh is a true book, then this corner has to be the left bottom corner of the book. Based on this information, it looks for the most-bottom-right corner that does not deviate too much from the anchor point's y position. (When the user holds a book, the book is rather quite horizontal because otherwise it would make it hard for the user to read.) If such a corner exists and has a reasonable x distance away from the bottom left corner, then the system considers the most-bottom-right corner as the bottom right corner of the book.

Once these two bottom corners are in place, it eliminates all other points in the detected corner dataset whose y heights are smaller than 1/3 of the book height. This operation narrows down the candidates for the top two book corner points and the mid-ridge point. Based on x positions, the surviving corner points are then divided into three buckets: first one for possible top-left corner, second one for possible mid-ridge and the third one for possible top-right corner. Among the corner points in the first bucket, the system looks for a corner that is topmost but whose x position is within some distance from the x position of the detected bottom-left corner point. This significantly reduces jumping of the top-left book corner. The same is performed on the points in the third bucket to find the top-right book corner in respect to the detected bottom-right book corner. For finding the mid-ridge point, which marks the center of the book, the system uses the median point between the detected top two corners as the anchor point to filter out non-mid-ridge points. With a very small threshold for distance, the true mid-ridge point is detected.

Once all five points are detected as shown in Fig. 3-5 (right), the system records the distances between the bottom-right book corner and each one of the other corner points. As a result, even if one of the corners except the bottom-right corner jumps by a large amount (e.g. if a hand gets in the way and covers one of the book corners), the system is able to recover the position using the stored distances from the bottom-right corner. This enables robust tracking, especially when the user interacts with the pages. If the bottom-right corner jumps by a lot or is lost, then the system goes through the entire filtering algorithm in order to find the book corners again. Because the user holds the book in the left hand and flips pages using the right hand, the bottom-left corner is secured most of the times.

Additional algorithm is implemented in order to animate the flipping of the virtual page along with the physical page that is turning. Based on the detected five corners, the system segments out two possible regions where the flipping can start: the left half and right half of the book. When the user flips a physical page in the forward direction (meaning that the right side of the page gets flipped to the left side of the book) or in the backward direction, the corner of the page that is turning is the highest point of all the detected corner points for most of the times. Using this characteristic, the system, after a book has been detected, checks if there exists an additional corner point that is higher than the top two corner positions. If it detects one, it remembers on which side it first happened and tracks this point continuously until it gets too close to one of the two top book corners. It creates a third page behind the turning page as soon as the turning starts so that the user can see the next virtual page. The virtual page is augmented following this moving corner point of the page; this creates an illusion of a real page turn. If the corner of the turning page is detected first on the right side and then gets merges with the left top corner, then the system accepts that a forward page turn has been detected.

The downside of this book-tracking method is that it significantly depends on the background settings such as the color and light conditions. To supplement failure cases, Leap IR frames are used (instead of the RGB webcam frames), in which a similar image processing and filtering algorithm is applied to detect and track the physical book. More explanation is provided in the next part.

Leap IR-Based Book Tracking

Leap's IR camera-based book-tracking system (or "IR tracking" for short) is switched on when RGB book tracking fails to locate a book due to a bright or cluttered background. Because the three IR LEDs embedded in Leap illuminate the scene, objects closer to the IR cameras reflect back more IR lights. Furthermore, any object outside of the working range (approximately 25 to 600 millimeters) gets excluded automatically. For example, a white wall a few meters from the user is not visible to these cameras, whereas it will cause the RGB book-tracking system to fail. Using this advantage, the system takes the IR camera frame and successfully detects the book and its corners in a scene by going through similar image-processing and corner-data-filtering procedures described in the previous section (see Fig. 3-6).

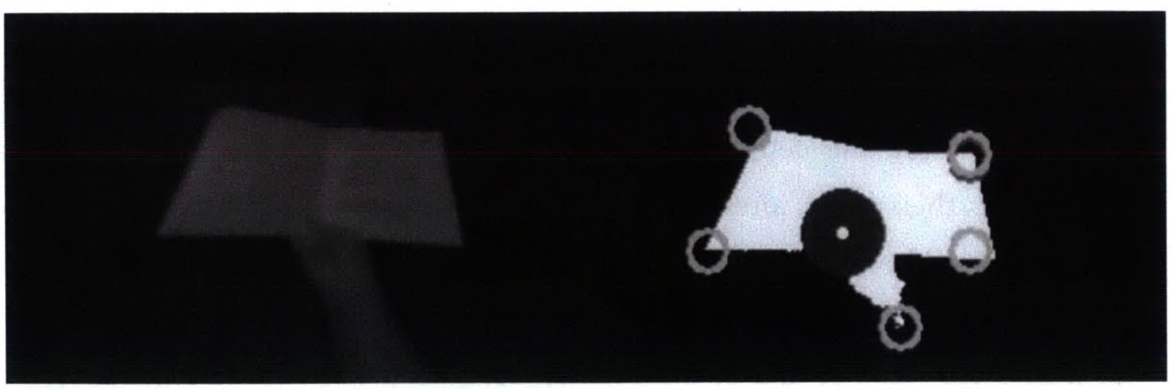

Figure 3-6: Leap camera view (left) and thresholded image with detected corners (right)

However, IR tracking has drawbacks inherent to the device hardware. While the fisheye lens built into the device helps the cameras get a larger field of view, it significantly distorts the image. Programmatically reversing this distortion (or warping) has an insignificant effect. Furthermore, the resolution of the IR cameras is only 320 x 120, which is about four times smaller than that of the webcam. For this reason, the book corners detected using the IR book-tracking system are not as accurate as when using the RGB book-tracking system because the low-resolution IR frames are scaled up about four times to match the resolution of the RGB camera. Nonetheless, it serves as a nice alternative to the RGB book-tracking system whenever it fails to produce reliable results.

Hand-Tracking Systems

Leap IR-Based Hand Tracking

The Leap IR-based hand-tracking system (or "IR hand tracking" for short) is predominantly used in the *Virtual Book* scenario, in which the user is capable of retrieving an entirely virtual book in one hand. The Leap Motion SDK [50] provides the methods this tracking system uses to acquire precise 3-D hand position and individual finger joint information (see Fig. 3-7). Based on this data, a number of motion-based gestures can be designed.

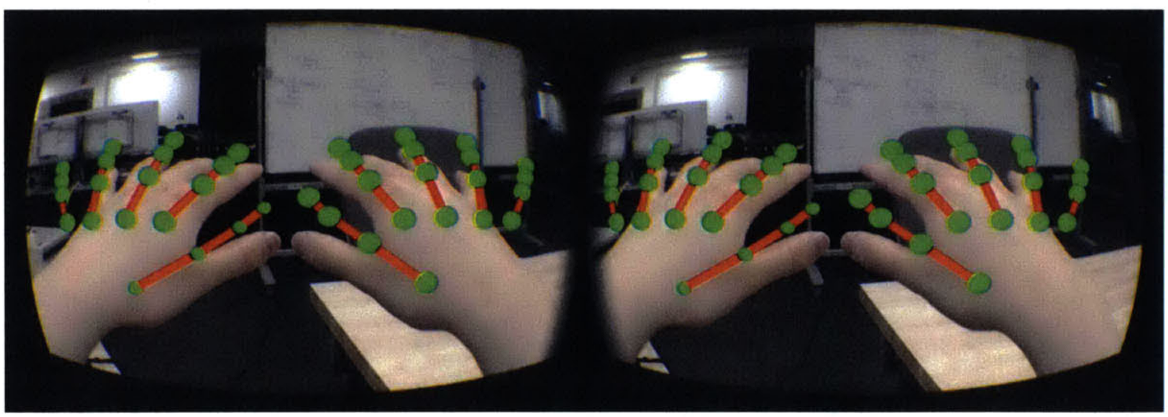

Figure 3-7: IR hand tracking results shown through Oculus Rift display

While the IR hand-tracking system is well suited for the *Virtual Book* use cases, it unfortunately fails to detect hands in the *Augmented Book* scenario. Because the user places the hand very near the book pages when interacting with it (e.g. annotation), the difference in z position (or the depth) is far too little for the IR camera to differentiate between the hand and the paper. Moreover, because the skin and paper have similar brightness in color, the monochromatic device gets confused when the user's hand overlaps with the paper pages. For this reason, I implemented another method for tracking the hand using the webcam as the alternative.

RGB-Based Hand Tracking

As soon as IR hand tracking fails to perform, the system changes the mode to RGB-based hand tracking (or "RGB hand tracking" for short). When the hand-tracking mode is on RGB, the system grabs the webcam frame and applies a number of computer vision techniques to recognize a hand. Before these techniques can be applied, however, the image has to go through a number of image-processing operations for the best result. Once an RGB frame is acquired, the system first looks for colors that are similar to a skin color. For stable color detection, HSV segmentation is used to separate the skin tone blobs from the rest of the image. HSV segmentation is more effective than regular RGB segmentation because it is independent of intensity or lighting. The HSV segmented image then gets turned into a binary image, which is then blurred to smooth out the shape of the detected object blobs.

Figure 3-8: RGB camera view (left) and the thresholded image of the hand mask (right)

At this point, it is ready to undergo more intricate image-processing operations. For detecting the hand mesh and the five fingers, the method first approximates the contour of the blob using OpenCV [51] and computes the convex hull of the shape. Next, it computes the interior angles of the hull corners. If the interior angle is small, meaning the shape is pointy, and if this interior angle is located within the top 1/5 of the convex hull, then it detects the angle as a fingertip (see Fig. 3-8).

The RGB hand-tracking system is primarily used in user interaction scenarios for the *Augmented Book*. It is true that a challenge in the system still remains in the fact that it is sensitive to color and light settings. If a book is made up of pages similar to the user's skin color, then the system will get confused. However, given that most book pages are white in background with occasional color images, it manages to work reliably with the majority of ordinary books.

As the Variable Reality tracking system is designed to support an on-the-go reading experience, it is built in multiple layers to enable stable tracking independent of the environment. Individual tracking system is backed up by the other in case its tracking fails. Automatic mode switching of the tracking system allows users to interact with Variable Reality books in a non-intermittent manner. Initially, the lag in the system due to intensive image processing operations deteriorated the quality of the user experience. This issue, however, has been resolved in the current implementation, in which each of the tracking systems run concurrently in separate threads [49].

3.3 Interaction Techniques

Using the detected hand and finger information acquired through the Variable Reality hand-tracking systems, I designed a number of gestural interaction techniques for the *Augmented Book* and the *Virtual Book*. The *Augmented Book* uses the RGB hand-tracking method while the *Virtual Book* employs the IR hand-tracking system.

3.3.1 *Augmented Book* Interaction Styles

Index Fingertip

The user's index fingertip is used for interaction just like a mouse cursor on screens (see Fig. 3-9 (a)). The index finger is used because the way one holds it up is similar to the hand shape when one points at something, which is a long-established way of conveying intention. Therefore, any motion-based hand gesture available for interacting with the *Augmented Book* is based on the movement of the index fingertip.

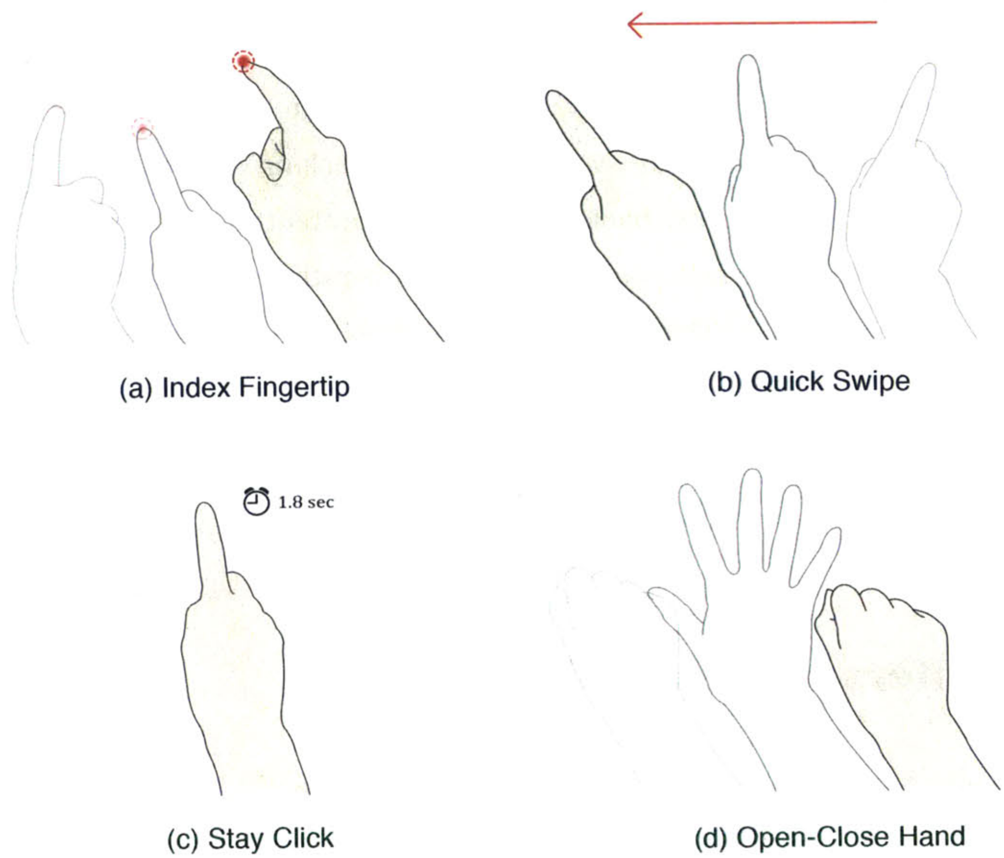

Figure 3-9: Variable Reality Augmented Book Interaction Styles

Open-Close Hand

By counting the number of detected fingers, a quick opening and closing of the hand motion is detected by observing the change in the number at a given time (see Fig. 3-9 (d)). Additionally, the location where the Open-Close Hand motion takes place can be used for manipulation.

Stay Click

Using the index fingertip, the user can click on icons and contents on a page by staying (or hovering) over them for 1.8 seconds (see Fig. 3-9 (c)). The current stay time is determined just so that it does not give the impression of a system lag but is adjustable.

Quick Swipe

The user can perform a quick swipe by quickly moving the hand from right to left while keeping the pointing gesture (see Fig. 3-9 (b)). This is a time-based interaction, which observes the index fingertip movement in a given time interval.

3.3.2 *Virtual Book* Interaction Styles

Open/Close Hand (Face Up)

The IR hand tracking-system conveniently provides information about all of the hand and fingertip information. It uses this data to determine whether a hand is open or closed. If the distance between the thumb and pinky is larger than a certain amount, then it detects it as open and vice versa (see Fig. 3-10 (a)). Recognizing whether the hand is facing up or down can also be useful. A *face up* state of the hand is true if the pinky is located on the right side of the middle finger.

Quick Open-Close Hand (Face Down)

In the *Face Down* state (*Face Down* refers to when the pinky is located on the left side of the middle finger), quickly closing and reopening the hand is called the *Quick Open-Close Hand* interaction style (see Fig. 3-10 (i)).

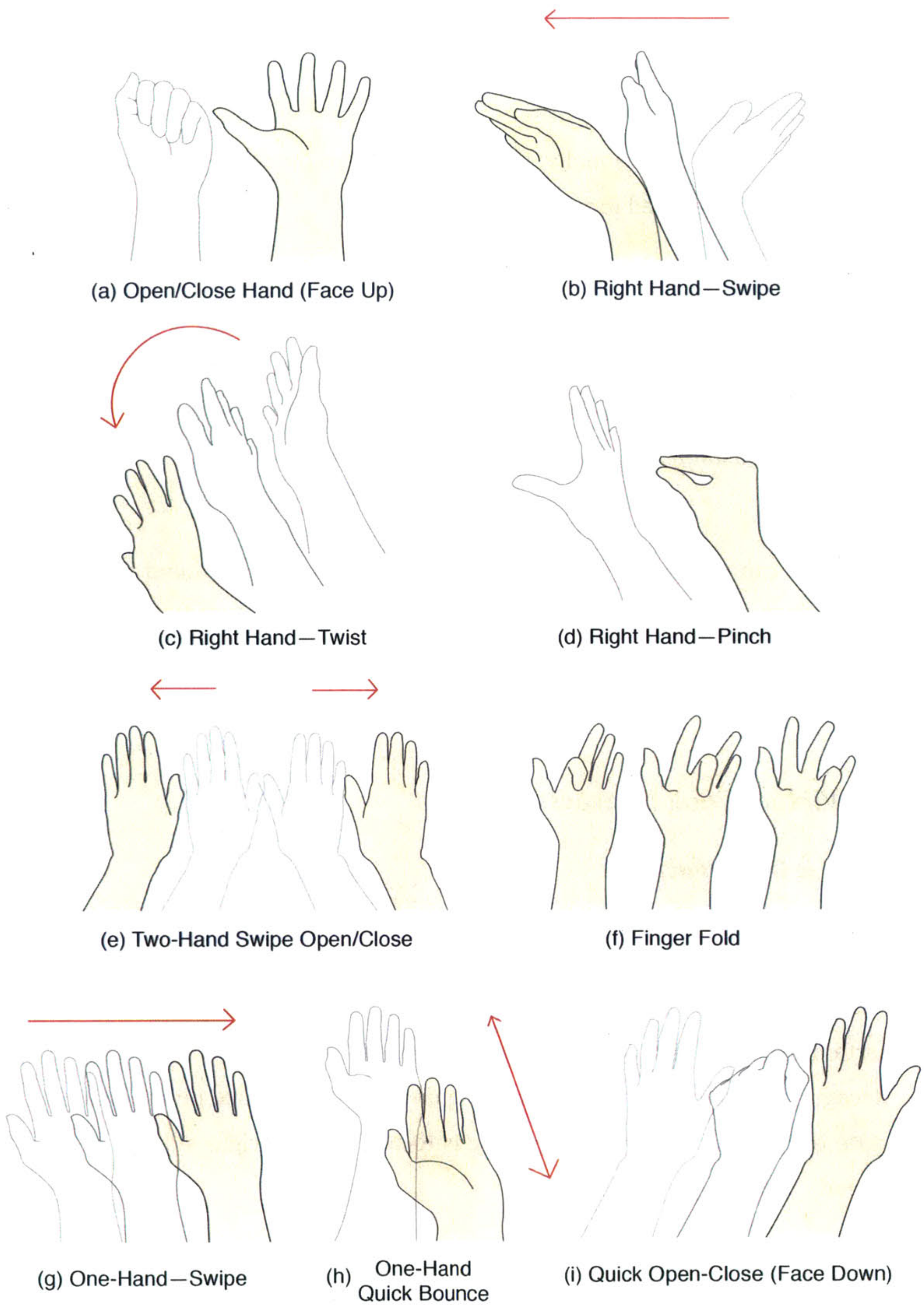

Figure 3-10: Variable Reality Virtual Book Interaction Styles

Right Hand Swipe, Twist, Pinch

While the user virtually holds a 3-D book in the left hand, the right hand is free to make gestures as an input mechanism. A few quick and simple motion-based hand gestures are defined: swipe, twist and pinch. A swipe gesture refers to a swift, right-to-left hand movement (see Fig. 3-10 (b)). In this interaction, the hand shape is insignificant. A twist gesture refers to a quick rotating motion of the hand around the middle finger using the wrist (see Fig. 3-10 (c)). A pinch gesture is similar to how one would pinch someone. It is activated when the distance between the index fingertip and the tip of the thumb suddenly decreases (see Fig. 3-10 (d)).

Two-Hand Swipe Open/Close

Unlike in the *Augmented Book* scenarios, the user has the freedom to interact with the *Virtual Book* system using both hands. A *Two-Hand Swipe Open* is detected when the user initially holds two hands near one another and suddenly brings them apart (see Fig. 3-10 (e)). In the reverse, a *Two-Hand Swipe Close* is detected when the user rapidly brings together two hands that were originally farther apart.

Finger Fold

By calculating the distance between each fingertip location and the palm point, the center of the hand, whether an individual finger is folded or not can be recognized. The *Finger Fold* function detects which of the five fingers is folded based on how close it is to the palm point (see Fig. 3-10 (f)).

One-Hand Interaction – Swipe; Quick Bounce

As opposed to holding the *Virtual Book* in the left hand and interacting with it using the right hand, the *One-Hand Interaction* style enables the user to use only the left hand for manipulation. A *Swipe* is a quick left hand movement from left to right (see Fig. 3-10 (g)). A *Quick Bounce* is a sudden down and up hand movement, slightly coming toward the user's body and then back out (see Fig. 3-10 (h)).

3.4 Early Prototypes

The current Variable Reality system is a newly developed prototype that has improved upon and enhanced the challenges and limitations experienced in the earlier prototypes. The previous setup (as shown in Fig. 3-11) consisted of a webcam and the Leap Motion Controller mounted on the front of Oculus Rift DK1 [52] with a display resolution of 1280 x 800. The webcam used was called Logitech C270 HD Webcam [55], which had a resolution of 640 x 480. The resolution of both the HMD device and the webcam were far too low for the user to clearly read text on paper pages. Hand tracking and gesture recognitions were performed using the Leap Motion Controller only. As a result, the first Variable Reality book, called the "AR Book," had to be consisted of black papers with AR markers on each page.

Figure 3-11: Earlier Variable Reality prototype

In the early prototypes, there was a clear need for improvement in both the system's hardware and software. 1) The resolution of the RGB camera had to be much higher to capture clearer images. 2) The HMD display resolution had to be high enough to display the captured high-resolution image so as to enable users to read texts on paper through the system. 3) The AR Book was a custom-made book, which hindered with readily available aspect of the Variable Reality system; further, its design, with AR markers at the center of each black page, was visually unpleasant.

Based on these observed room for improvements, I upgraded the hardware to more powerful devices such as the Oculus Rift DK2 and a webcam that has twice as much resolution as the previous one. I also devised and implemented book-tracking methods that turn any physical book into an AR Book without the need for any additional tracking aids (e.g. AR markers or features). As a result of the hardware and software design iteration, the current Variable Reality system allows users to effortlessly switch their reading from a paper book to a digital book using a single physical book. In mobile cases, users do not even need a real book because their hand essentially is the book.

4

Applications

Based on the stable book- and hand-tracking algorithms along with the recognition of various gestures designed for user interaction, I built several Variable Reality applications for both the *Augmented Book* and the *Virtual Book*. These applications import convenient navigation techniques available on digital books and yet maintain the physicality of a paper book. The goal of each application is to create an intuitive and effective reading experience for the users. Incorporating natural hand gestures associated with reading activity adds to the positive user experience in reading.

4.1 Variable Reality Augmented Book

Using the interaction styles designed for the *Augmented Book* (described in Chapter 3), I created several applications that enable users to perform quick navigation, link to external books, detect page numbers and display graphical elements associated with the page materials.

Page Flip

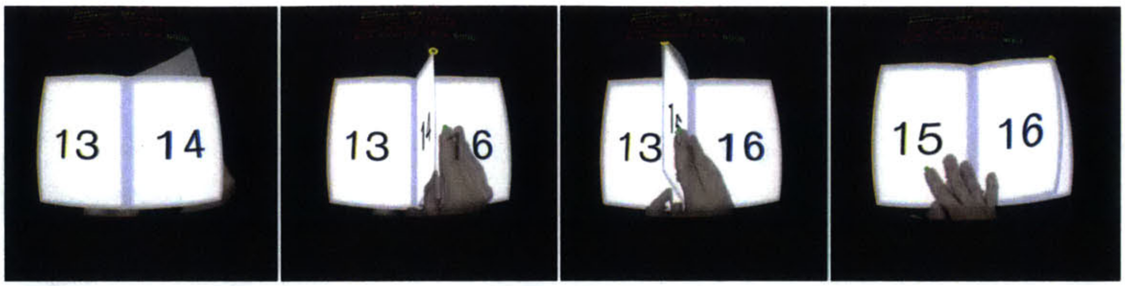

Figure 4-1: Page Flip in action

Page Flip realistically animates the turning of a page following the motion of the real page, which the user manipulates physically (see Fig. 4-1). To enable *Page Flip*, an additional image-processing algorithm works within the RGB book-tracking framework to detect the edge of the page that is turning (see p47, *RGB-Based Book Tracking*). The physical turning of a page that is replicated with the virtual page creates a successful illusion that the digital book is the physical book.

Quick Page Flip

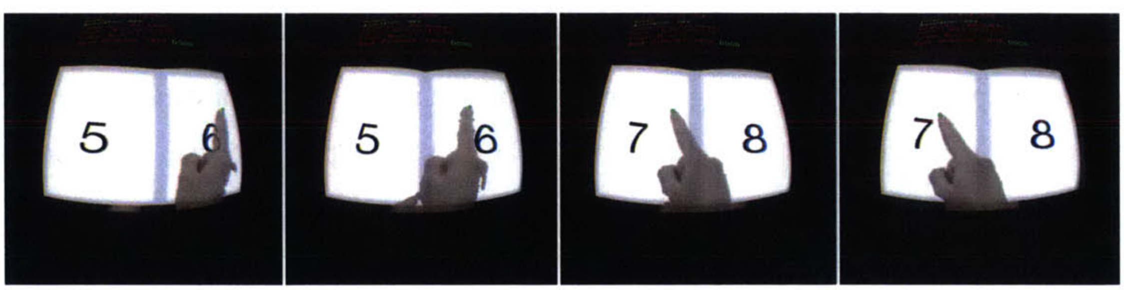

Figure 4-2: Quick Page Flip in action

Quick Page Flip provides users with a convenient navigation technique. While *Page Flip* ensures a realistic feel to user interaction, physically turning every page can be slow. *Quick Page Flip* incorporates the *Quick Swipe* interaction style to enable the user to quickly traverse the pages with a simple swipe (see Fig. 4-

2). When a *Quick Swipe* is detected, the virtual page on the book promptly changes into the next page without the need to physically turn the page. Using this application, the user can navigate through the pages much more rapidly than physically flipping through the pages.

Quick Preview

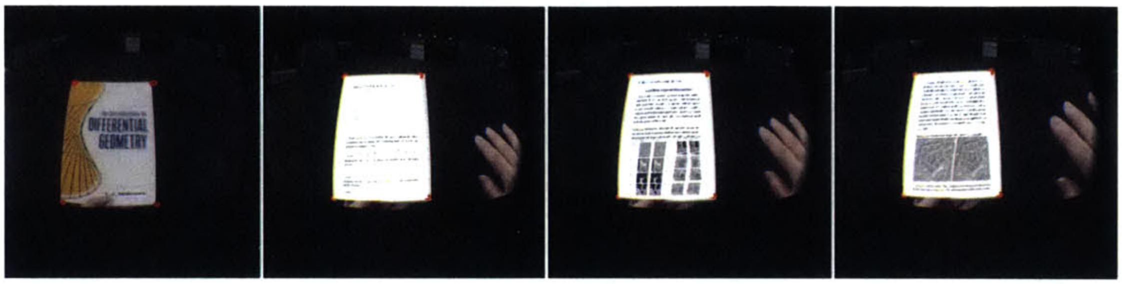

Figure 4-3: Quick Preview through a book

Often people want to have a sense of what the book is about but simply looking at the book cover does not give them much information about the book. *Quick Preview* enables users to quickly see through all of the book pages and materials without opening the book (see Fig. 4-3). Using a widely known computer vision-based feature detection algorithms, it detects the book and its title. The contents inside the book appear onto the front cover of the book and the user can traverse forward and backward using the right hand. Bringing the hand closer to one's body will bring the augmented page towards the beginning of the book and vice versa.

Table of Contents, Index

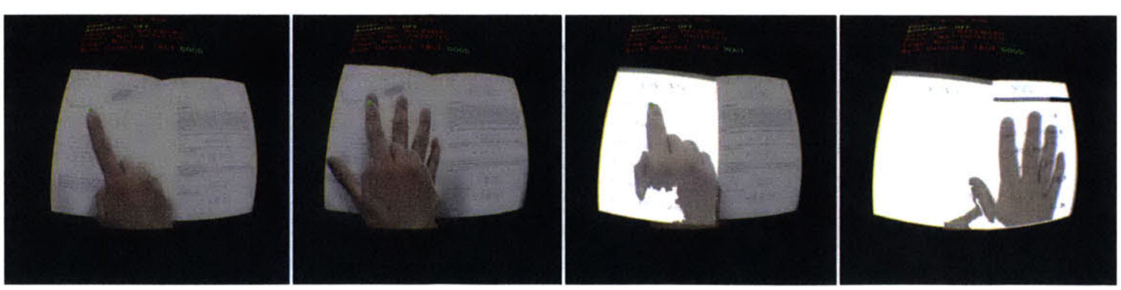

Figure 4-4: Open-Close Hand to display Table of Contents and Index

One of the useful features available on a digital book is the easily accessible table of contents and index. With a physical book, moving back and forth between the table of contents page and the index can be tiring. The *Augmented Book* employs the *Open-Close Hand* interaction style to allow users to display the table of contents and index information on the page they are on (see Fig. 4-4). If *Open-close Hand* is detected on the left side of the book, it replaces the left side page with the table of contents. If *Open-close Hand* is detected on the right side of the book, it replaces the right side page with the index.

Annotation

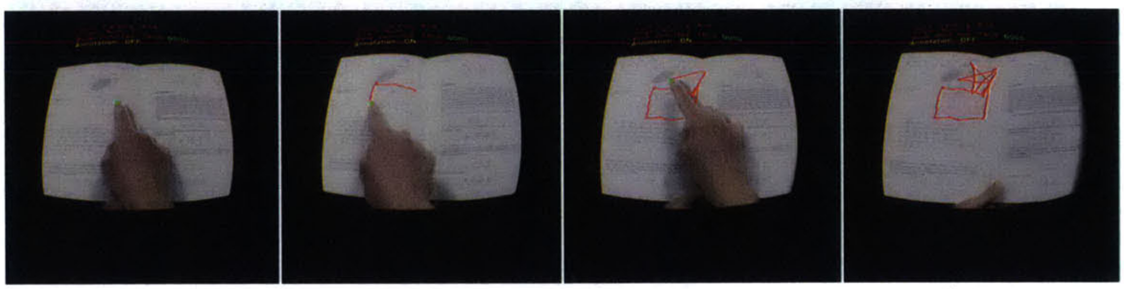

Figure 4-5: Free-form annotation

One of the downsides of a digital book is that it does not provide a way for users to input free-form annotations as they could with a paper book. While digital books do offer highlighting and inserting notes by typing, users do not have much freedom when it comes to drawing circles and putting a star mark next to important information. The *Annotation* application enables the user to digitally draw on and annotate a physical page (see Fig. 4-5); this can also be done on an augmented virtual page. The *Index Fingertip* is used as a pen tip; the user can *Stay Click* to start and end an annotation and the finished annotation stays on the page.

Virtual Tags

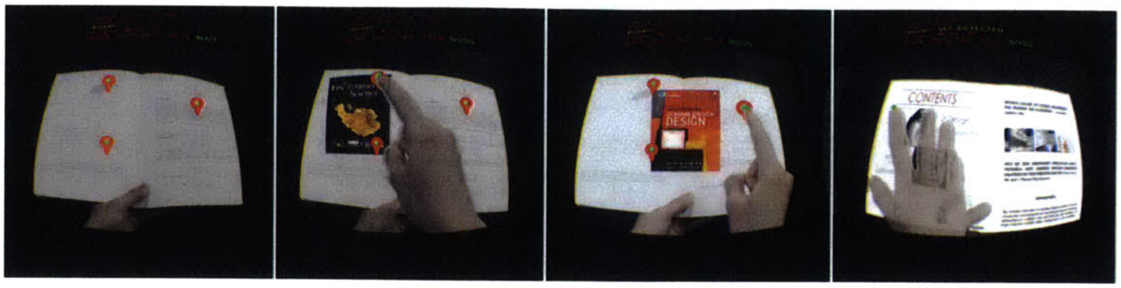

Figure 4-6: Click on a Virtual Tag to display the book

Virtual Tags application allows the user to link information on a physical page to external books and materials. It uses the *Stay Click* interaction style with which the user can click one of the *Virtual Tags* icons on a page and display a related page from an external book (see Fig. 4-6). In this way, the user can quickly acquire information from multiple books without having to physically open them and search for relevant pages.

OCR Page Detection

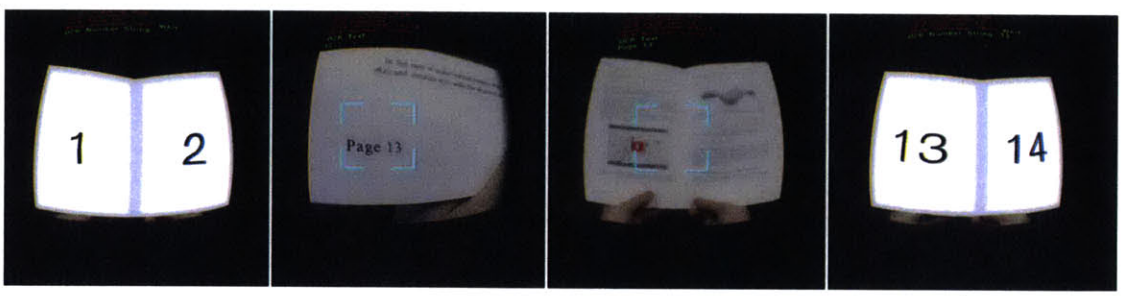

Figure 4-7: OCR to go to a specific page

For *OCR Page Detection* application, the Optical Character Recognition (OCR) engine detects a page number of a page (see Fig. 4-7). Due to the resolution of the camera, the user needs to bring the book closer to the camera and focus it on the number printed region. Hairlines are drawn in the view to indicate the working boundaries of OCR. When a page number is detected, the *Augmented Book* displays the corresponding virtual page of a designated book.

Hover-display Graphics

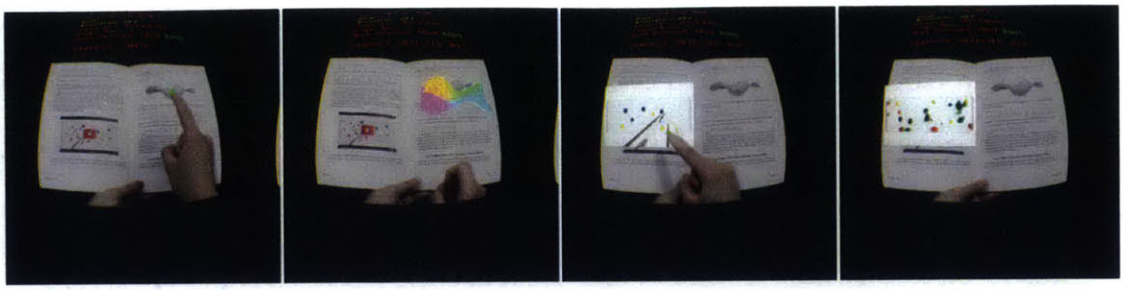

Figure 4-8: Click on multimedia contents to watch it live

Physical pages are affixed to the book in a sense that the contents on a page is neither going to move or change. Therefore, digital contents can be linked to materials on a specific page. Once a page number is detected on a physical page, pre-designated digital contents can be augmented. For example, the image of an intricate 3-D math graph shown on a page is quite difficult to grasp just by looking at the image taken from one perspective. In *Hover-display Graphics* application, the user can *Stay Click* to display an animated version of the graph that shows all sides of it more clearly. It can play other contents like videos that are impossible to play on paper pages (see Fig. 4-8).

4.2 Variable Reality Virtual Book

Designed for mobile user scenarios in which there is a lack of available physical books, *Virtual Book* applications make it possible for a user to easily find a virtual book from a personal digital library and hold it just as one would hold a paper book for reading. Using the interaction styles designed for the *Virtual Book*, I implemented a number of applications that take advantage of the weightlessness and digital aspects of the book.

Open/Close Book

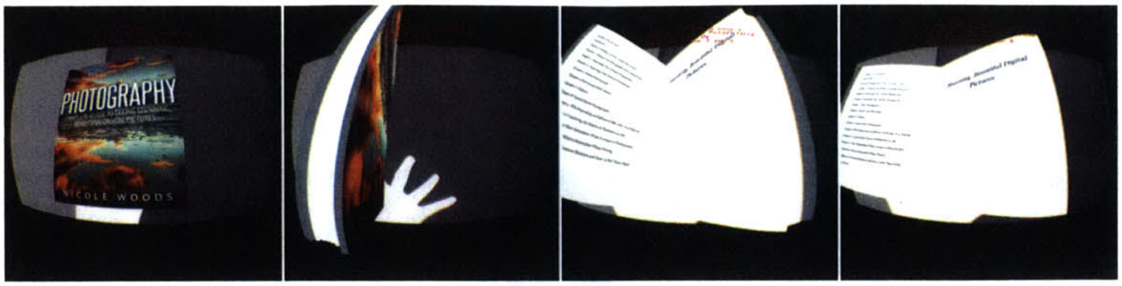

Figure 4-9: Open/Close the Virtual Book with hand gesture

In the *Face Up* state of the left hand, a user-selected book gets virtually augmented on top of the hand just as a real book would rest on the hand. The *Open/Close Book* application uses hand gestures designed in the *Open/Close Hand* interaction styles. The opening hand gesture opens the book to the last opened page (or to the first page if it is opened for the first time) while the closing hand gesture closes the book (see Fig. 4-9). The book follows the hand position, creating the illusion of the book resting on the palm.

Page Flip, Reverse Flip

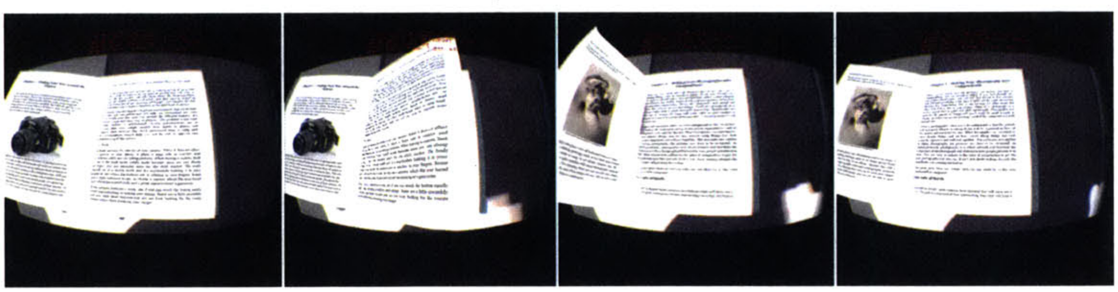

Figure 4-10: Flip pages with Right Hand Swipe

In the *Page Flip* application, the user can turn a page of the *Virtual Book* with the *Right Hand Swipe* interaction style (see Fig. 4-10). The turning of a virtual page is realistically animated to enhance the user's reading experience. By performing *Right Hand Twist*, the user can reverse the direction of the page turning, which activates the *Reverse Flip* application. In both applications, the *Right Hand Swipe* motion can have similar motion, but the turning of the page will happen according to the designated direction of a page turn.

Orientation

Figure 4-11: The manipulation of the book orientation

In the earlier prototype, the *Orientation* application was missing. As a result, the *Virtual Book* would follow the hand even when the hand's palm was facing sideways, which ordinarily would make the book fall due to gravity if it were real. Hence, the *Virtual Book* had an undeniable supernatural feel to it. The *Orientation* application in the current prototype makes the book follow the hand orientation (see Fig. 4-11). For example, if the fingers fold perpendicular to the palm, then the book is orientated to stand vertically. The user can control the orientation of the book in six degrees of freedom (6DoF).

Pinch to Scale

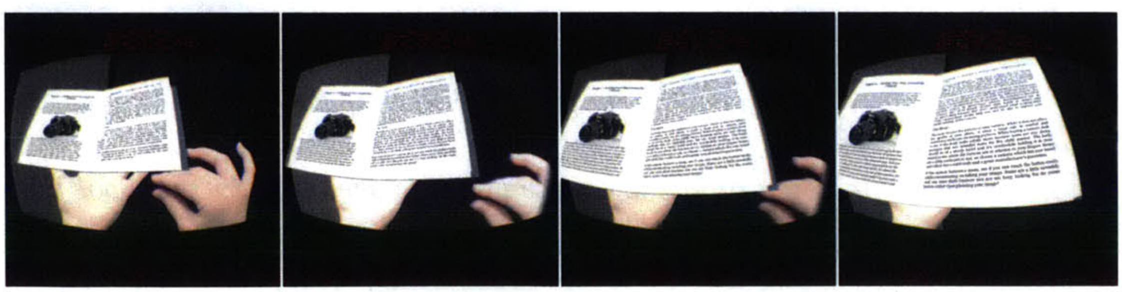

Figure 4-12: Pinch and drag to scale the book

The advantage with the *Virtual Book* as opposed to a physical book lies in the fact that its features can be manipulated digitally. *Pinch to Scale* application is designed to change the size of the book through the *Right Hand Pinch* interaction style. With a *Right Hand Pinch*, the user can pinch one corner of the book and drag it horizontally to increase or decrease its size (see Fig. 4-12). Releasing the pinch fixes the book size. As a consequence, the user changes the font size with respect to the new book size. A digital book adjusts the font size by simply increasing the letter size in the same screen space and pushing back all the following pages; for this reason, a page number on a digital book is meaningless. On the contrary, the way font size adjustment is managed in the *Pinch to Scale* application does not interfere with the spatial arrangement of the book and this aspect promotes the use of spatial memory.

Fold to Go to Location

The *Fold to Go to Location* application is built to make book navigation easier. The user can quickly open up the book to a page that is linked to each finger. Using the *Finger Fold* interaction style, the system determines which finger is folded and flips the pages until the book is on a corresponding page. Currently, the system automatically assigns a location to each finger based on the length of the book. For example, the index finger is assigned to the first page of the book while the pinky is assigned to its last page. The middle and the ring ringers are assigned to pages in the book that equally divide it.

One Hand Interaction – Flip, Lock-in-Air

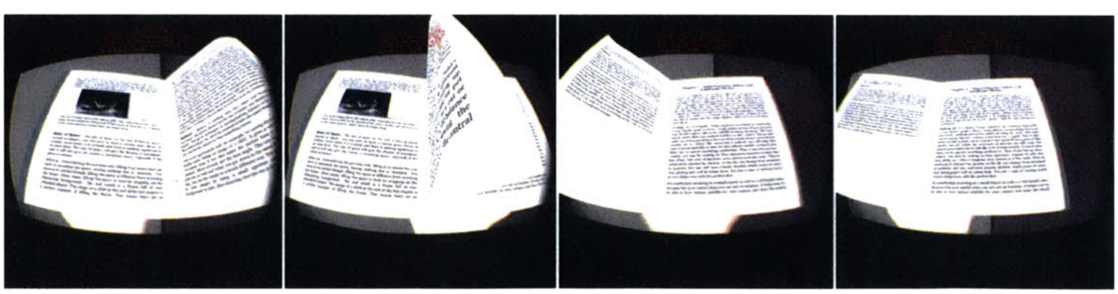

Figure 4-13: One Hand Interaction - Flip

If the user wants to drink a cup of coffee while reading, *One Hand Interaction* comes in handy. The user can hold the *Virtual Book* on one hand and perform a quick swipe similar to the *Right Hand Swipe* to turn the book page (see Fig. 4-13). *Lock-in-Air* interaction makes use of the *Quick Bounce* interaction style and fixes the book to the user's head. As a result, the user can read the book in any posture without the need to hold it in hand.

Virtual Bookshelf

Figure 4-14: Selecting a book from the Virtual Bookshelf

Today's digital libraries offered by digital books attempt to mimic the appearance of a real bookshelf. However, the interface is nonetheless limited to 2-1/2-dimensional space, which makes it hard to remember book locations. The *Virtual Bookshelf* application is built to offer the users a three-dimensional bookshelf that contains user's personal collection of *Virtual Books* (see Fig. 4-14). The user can bring up the *Virtual Bookshelf* anytime from anywhere with a simple *Two-Hand Swipe Open* hand gesture. The *Virtual Books* in the bookshelf are animated as the user hovers over them. Using a *Quick Open-Close Hand (Face Down)* gesture, the user can pick out a book to read. The *Virtual Bookshelf* allows the users to spatially organize and manage books and this enables rapid retrieval of a book in need.

4.3 Multiple Book Scenarios

The multiple book scenarios refer to using both the *Augmented Book* and the *Virtual Book* in a shared spatial setting. It is designed to bring together the "affordances" of a physical book and of a digital book. Digital books can instantly display any book from the digital library but cannot display multiple books simultaneously. With physical books, one can open multiple books at once but it takes time and manual labor. More importantly, one has access to only so many print books. With Variable Reality, the user can have multiple books open and available at once. While holding a physical book for the *Augmented Book* interactions, the user can pick out a book from the *Virtual Bookshelf* and read it by virtually holding it with the other hand (see Fig. 4-15).

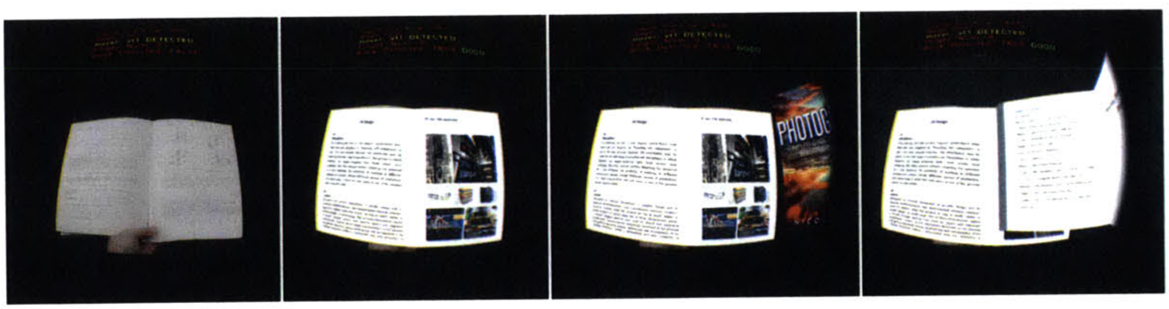

Figure 4-15: Multiple books –a physical book, the Augmented Book and the Virtual Book

5

Evaluation

Experiments were designed to observe the technical performance of the Variable Reality software and the user experience of using the *Augmented Book* and the *Virtual Book*. A special focus was on the gestural interaction styles that the participants used to manipulate the two books. Tests were conducted on two consecutive days between which I made some improvements on the system based on the feedback from the participants.

The recruited participants were asked to experiment with both the Augmented Book and the Virtual Book, one at a time. Half of the participants tried the *Augmented Book* first whereas the other half experienced it second. For the *Augmented Book* experiment, the participants were asked to use the *Virtual Tags* application. For the *Virtual Book* experiment, the participants were asked to try to open up one of the predefined *Virtual Books* and flip its page. Depending on how comfortable the participants became with the system, I had some of them experiment with additional interaction styles and applications. After completing these tasks, each participant filled out a questionnaire (Appendix A). This chapter discusses the testing process in more detail, technical difficulties encountered and the feedback gathered from the participants.

5.1 Test Setup

The test setup had a long rectangular table that was covered with a large black tablecloth. The table was placed against a wall that was also dark in color. This was an optimized setup for the *Augmented Book* experiments as the book is easier to detect in a darker background. The participants sat in front of the table, wore the Variable Reality hardware (Oculus Rift with a webcam and Leap mounted on it) and faced the wall. For the *Augmented Book* experiment, the participants had the freedom to choose between two physical books; one of them was smaller in size and much lighter while the other was slightly larger and heavier.

5.2 Participants

A total of 10 participants were recruited either in person or through an email. The group was comprised of 5 men and 5 women spanning 22-34 years of age. The participants were mainly MIT undergraduate and graduate students but one of the subjects was from a different organization and called him or herself "non tech-savvy." Nine of the participants were eBook users, four of whom indicated their frequent use. They particularly liked the portable and lightweight aspect of a digital book but disliked its unintuitive user interface and the interaction styles such as scrolling. Nine participants preferred to read a physical book over a digital book because they felt that it provided a more natural and intuitive reading experience. However, they expressed their frustration with the heavy weight and the inability to perform a quick scan or word search with a physical book. None of the participants have used the Variable Reality books before.

5.3 Experimental Procedure

The participants were asked to try reading with the *Augmented Book* and the *Virtual Book*. Half of the participants tried the *Augmented Book* first and the other half tried the *Virtual Book* first. The entire testing lasted for 15-20 minutes (8-10 minutes for each book experiment) depending on the participant. Before trying each book, the participants were given 3 minutes of instruction time during which I showed them a demo video of the book application that they will try. I also taught them the interaction styles involved with each of the book such as the *Quick Swipe* and *Open/Close Hand*. Depending on how comfortable the user felt using the system, I guided them to try different applications using different set of hand gestures.

The *Augmented Book* experiment was designed to have the participants try the *Virtual Tags* application; I decided to have the participants try this application because it included a number of major interaction styles. I guided each participant in advance to hold the book with the left hand and interact with the right hand once the book is detected. I mentioned that all four of the book corners have to be seen in order for the book to be detected. I also told them that they would have to give the system about 1.5 seconds after the book has been detected for it to start performing stable tracking. Half of the participants who experimented on the first day had verbal feedback of this state; I called out "wait" until the system was ready for the stable book tracking and then called out "good" for the participant to start interacting with the book using the right hand. Based on the feedback, I modified the system so that the participant had text feedback of this state in their Oculus Rift view, which made it easier for them to know when was a good time to start using the gestures. I instructed the participants that when the book is detected by the system, the virtual tags should appear on the book pages and that they should use their right hand *Index Fingertip* to hover one of them to perform *Stay Click*, which would bring up the virtual pages. I asked the participants to flip through the pages using the *Quick Swipe* interaction style. Depending on how comfortable the participant felt using the system and the gestures, I had some of them try the *Quick Open/Close* to bring up the index page on the right and the contents page on the left side of the physical book.

The tasks related to the *Virtual Book* experiment involved the *Open/Close Book* and *Flip Page* application. Before the experiment, I instructed the participants on how to perform the *Open/Close Hand* gesture for bringing up the book and opening it up, and the *Right Hand Swipe* gesture for flipping the virtual book pages. The system overlaid the physical hand with the 3-D hand and finger joint information acquired through the Leap Motion Controller; this was intended to give feedback to the user as to when the hand was detected and which hand it was. With the left hand closed, the Virtual Book was designed to show up on the palm of the left hand. With the *Open Book* gesture, the book was animated open. The *Virtual Book* displayed was predefined for all participants. I guided the participants to start interacting with the *Virtual Book* using the right hand once the book was open. I told them to make sure that the two hands do not overlap because the Leap Motion Controller would get confused and fail to detect the hand. I mentioned that their goal was to get comfortable with the system and the gestures as much as possible in the given time. Depending on how quickly a participant learned to adapt to the system and the gestures, I had him or her try the *Right Hand Twist* interaction style to reverse the flipping direction and also to try opening and closing the *Virtual Bookshelf* and pick out a book to read.

5.4 Results

Seven participants successfully completed the basic tasks for the *Augmented Book* experiment. Four participants felt comfortable enough to try different interaction styles such as quick *Open-Close Hand* to quickly switch back and forth between the pages with the index or table of contents. Two participants had a different skin color from the color the system was designed to detect. Thus, the hand detection algorithm failed and the participants could not perform either *Stay Click* to bring up the virtual pages or the *Quick Swipe* to flip through the pages. A participant failed to detect the book and, therefore, could not further experience the book.

Nine out of the ten participants successfully completed the basic tasks for the Virtual Book. Three participants felt comfortable enough with the system and went onto try different interaction styles and applications such as the *Right Hand Twist* and the *Virtual Bookshelf*. One participant was successful at opening up the Virtual book but failed to flip a page; the right hand was not being detected properly.

5.4.1 Augmented Book

User Experience

In overall, the participants who successfully completed the basic tasks reacted positively to the experience and expressed that the system was easy to use. One of the participants described using the *Augmented Book* as "natural like a paper book." The participants found the hand gestures easy to learn an intuitive. The majority of them found storing additional digital data on the physical pages in a book and the ability to access it by simply pointing at the virtual tags particularly useful. Participants found the navigation method extremely convenient as they could quickly scan through the book content and navigate between chapters without having to manually flip the pages, which would, they expressed, be tiring and slow. A few participants who were comfortable enough to try the *Table of Contents and Index* application found it extremely convenient to look up the index without having to open up the back of the page.

Technical Performance

All participants found it difficult to detect the book at first. They particularly found it frustrating to have to make sure that all four book corners be seen for the tracking to work. As a result, the participants had to stretch out their arms far enough to be able to see all four book corners; they described this pose as "tiring" and "hard."

Participants with shorter arm length had trouble detecting the book. Except for one, all participants succeeded in detecting the book after a few practices. Among them, six of the participants were able to stably detect the book. They expressed that 'once the book was found, tracking was fairly easy.'

Half of the participants found the hand tracking difficult, two of whose hands did not even get detected due to the different skin color. A few participants commented that the hand tracking slowed down the system, which took away from the experience. The participants with the skin tone that the system was designed to detect had less trouble in interacting with the system. The best result came out when I told individual participants that they could actually touch the physical page with their right hand. Often the participants thought that the right hand had to be interacting in the air in front of the book.

5.4.2 Virtual Book

User Experience

In general, the participants enjoyed interacting with the Virtual Book and found it "cool and interesting" and easy to use. One of the participants expressed that the 3-D form of a book "made it feel more natural (like a real book)." One other participant expressed that the *Virtual Book* was "more like a physical book" in that one could have a better visual sense of the physicality of a book, like which page one is on. A different tester described the reading experience "very similar to reading a paper book." All participants except for one, whose right hand could not be detected, found the flipping gesture very natural and easy to use. Many participants liked its weightless aspect, which would allow them to read a book on a bed without having to hold a physical book in the air, which gets tiring after a while. Three participants who became comfortable enough with the system along with the gestures further interacted with the *Virtual Bookshelf*, which they found extremely handy.

Technical Performance

All participants were able to successfully bring up the *Virtual Book* on the palm of the left hand and open it using the *Open/Close Hand* gesture. However, the participants generally had difficulty detecting the right hand. The Leap Motion controller often got confused between the left and the right hand and often detected the right hand as the left hand. As a result, the participants caught the *Virtual Book* with right hand when it was supposed to stay with the left hand. One of the participants failed to successfully detect the right hand in the given time. However, the rest of the participants, once the right hand was properly detected, were able to perform *Right Hand Swipe* for flipping the pages. One technical limitation that people found frustrating was the fact that the two hands could not overlap, in which case the Leap Motion Controller would lose the detected hands.

5.4.3 Interaction Styles

The interaction styles used in the Augmented Book experiment were *Index Fingertip, Stay Click* and the *Quick Swipe*. A few participants went on to try the *Quick Open/Close* gesture for bringing up the index or table of contents. The participants who were able to interact with the book using these gestures (because the system was able to detect their skin tone) found the interaction styles extremely natural, intuitive and easy to learn. One of the participants mentioned that it requires some time getting used to.

The interaction styles used in the *Virtual Book* experiment were mainly *Open/Close Hand* and *Right Hand Swipe*. A few participants felt comfortable enough with the system after a few practices and tried additional gestures such as the *Two Hand Swipe Open/Close* for the *Virtual Bookshelf* application and the *Right Hand Twist* to reverse the flipping direction. All of the participants who were able to properly detect the right hand (nine out of the ten participants) expressed that the gestures made sense and that the flipping gesture was like interacting with a physical book. They felt that the gestures were straightforward, and easy to learn and use.

5.4.4 Device

The participants found the Oculus Rift with the two additional devices too "heavy" and "bulky" to wear for a reading experience. Further, the participants had a different perception of depth between the real world and the augmented reality environment seen through the Oculus Rift display. As a result, the participants found it difficult to realize how far the book had to be placed in order for the book tracking to work reliably. The resolution of the webcam and the display resolution of the Oculus Rift, although both were better than the first prototype, were still not enough for the participants to clearly read small text in the books.

5.4.5 General Responses & Suggestions

Nine out of the ten participants said that they are likely to use the system again and to suggest it to their peer readers. The remaining participant expressed that he or she would use it again only if the device were lighter and more compact. The participants responded that they are more likely use the system on a plane or train because they do not produce quick motion as the subways do. A few participants mentioned that wearing the device in a public place may appear out of place until the HMD devices become widely used. A number of participants found the system particularly useful for thick and heavy textbooks and for linking among these books for a quick access.

The participants in general suggested the use of an alternative HMD device that is more lightweight and less bulky. They would like to use the system again but are hesitant to wear the heavy device on their head again. Further, they felt that the display was not clear enough for them to read smaller text in the book and suggested for a device with higher resolution. One of the participants suggested allowing the users to custom design their own flipping gesture because everyone has a different style of flipping pages.

5.5 Summary

The experiments show that the Variable Reality books provide a positive reading experience. It is clear that enabling the users to read a digital book using a physical book can create a natural and intuitive reading experience. It has also shown that readers enjoy reading an entirely virtual 3-D book; they are able to visually perceive it as a physical book. Quick and easy-to-learn hand gestures that are associated with reading activity add to the experience.

At the same time, there exists a clear need for improvement in both hardware and software. The limitations inherent in the devices such as the camera resolution or the bulkiness of the HMD device make it difficult for the users to feel comfortable while using the system. The tracking systems for both hand and the book require novice users to practice for a period of time; this implies that the tracking systems need to be more advanced in order to enable a natural and comfortable reading experience. These limitations and future work will be covered in the next chapter.

6

Limitations and Future Work

It is important that Variable Reality systems perform reliably in many different situations in order to enable intuitive mobile reading experience. There are, however, a few limitations that are inherent in the system that challenges achieving this goal. In this chapter, I discuss the limitations that hinder with the system and possible ways of overcoming them. Next, I present the future direction of the project.

6.1 Limitations

The most crucial enabling factors of the system are the book- and hand-tracking systems. If these systems fail to produce reliable results, none of the interaction styles or applications is possible. Currently, these tracking systems are sensitive to the light settings and background clutter. Further, the field of view of the cameras is either too narrow or too large for the system. This hampers the realistic experience the system is expected to produce.

Sensitivity to Light and Color

The color-based detection and algorithms used in RGB tracking make it very sensitive to colors and, as a consequence, helplessly fail when in the presence of competing targets. It is also sensitive to light settings because the detection works on a static color. Variable Reality systems thus require a relatively dark background to work robustly. To provide a truly mobile user experience, the system will need a tracking method independent of light and color settings. While such limitations in color-based object tracking have long been observed and studies for overcoming it are still an active field of research, algorithms presented in some past work can be applied to the system to increase the accuracy. For example, adaptive skin detection methods [15] and object detection methods can help eliminate false positives, using a classifier to detect the hand and book regions first.

The issue with the light setting is also present in IR tracking systems. Although the IR cameras of the Leap Motion controller have their own source of illumination, the three IR LEDs, they take on excessive external illumination when in sunlight, which creates noise and, thereby, significantly interferes with the tracking systems. To support outdoor user scenarios in bright sunlight, a 3-D sensor like Mime [37] that is uninfluenced by bright sunlight can be used as an alternative.

Field of View

To track a book, all four corners of the book have to be visible to the webcam. However, an opened book fills most of the camera's narrow view, leaving not much space for the book to move. The Oculus Rift creates an immersive environment by surrounding the user's eyes with a view that covers peripheral vision. An RGB camera that has a large enough field of view to capture the user's peripheral vision would make the tracking systems much more robust. In fact, using a see-through HMD device would quickly solve this problem. (The next section discusses implementing the system with a see-through HMD device.)

On the other hand, the Leap IR camera frames had too large of a field of view due to the fisheye lens built into the device to cover large input volume. Due to the warping and distortion inherent in the camera view, correctly mapping the tracked data to the user's real view was a big challenge. For example, moving the hand slightly to the left translated into a huge jump in the Leap space.

Book Sizes and Length

Because the field of view of the system is quite narrow, current implementation of the *Augmented Book* is limited to medium sized books (7.5 x 11 inches) or slightly smaller. If the book size is large, then the user would have to hold the book very far from the device in order to make sure that all four corners are seen. This limitation would cause unwanted distortion when augmenting a coffee table book contents onto a small paperback novel.

Furthermore, the interface of the *Augmented Book* is bound by the number of pages in a physical book. If a given digital book requires more number of pages than the number of pages in a physical book, then the user would either have to stop reading on the last page of the physical book or re-use one of the visited pages.

Bookshelf Capacity

The *Virtual Bookshelf* is limited in size and can hold up to 15-18 books depending on their thickness. However, current prototype does not reflect the initial idea behind the Virtual Bookshelf, which will be discussed in more detail in the next section.

6.2 Future Work

The work of the project can be extended in a number of ways that will create a more effective and user-friendly system for reading. Below, I present a number of possible future project ideas.

Collaboration

Reading is a social activity that encourages people to share thoughts with one another. By extending the Variable Reality systems to support a collaborative setting, the users can be more engaged in the reading experience. Instead of telling a friend to Google a book, one can simply drag and drop a *Virtual Tag* onto the friend's book.

Auto-Creation of Virtual Books

Currently, virtual book pages in the *Augmented Book* and the *Virtual Book* are manual creations from downloaded electronic versions of books. To support the quickly growing digital library, a system that can automatically download and create these virtual pages will enable access to millions of available books.

Multiple Books

The Multiple Books Scenario introduced in the previous chapter simultaneously uses the *Augmented Book* and the *Virtual Book*. However, the goal of the scenario is actually to implement multiple *Augmented Books* and *Virtual Books* in the space around the user's body. The user can lock a number of these books in the air and come back to them whenever needed. As a result, the entire spatial volume around the user's body becomes a user interface.

Additionally, this idea can expand to allow many more books in a shared spatial setting using a depth camera. For example, the surface of the table and its orientation can be recognized using the information. With this data, the user can put *Virtual Books* on a table without having to hold them.

Scalable Virtual Bookshelf

In the current prototype, the *Virtual Bookshelf* is not large enough to store more than a couple tens of books. In the future, the size of the *Virtual Bookshelf* should be scalable, meaning that the user should be able to quickly zoom-in to a portion of a large 3-D digital library and back out to see the entire collection again. The *Scalable*

Virtual Bookshelf could be thought of as a wall-sized 3-D bookshelf that is consisted of many of current *Virtual Bookshelves* nicely stacked up in the x and y direction.

Form Factor

Despite the immersive visual experience Oculus Rift offers, its bulky form factor makes it hard for users to wear it without feeling too uncomfortable. Furthermore, seeing the real world through a secondary display can create a disconnected experience. Enabling the system using see-through augmented glasses like Vuzix Start 1200 XLD [48] or the Meta One [46] (see Fig. 6-1) and HoloLens [47] (see Fig. 6-2) will be a worthwhile attempt to create a truly seamless user experience.

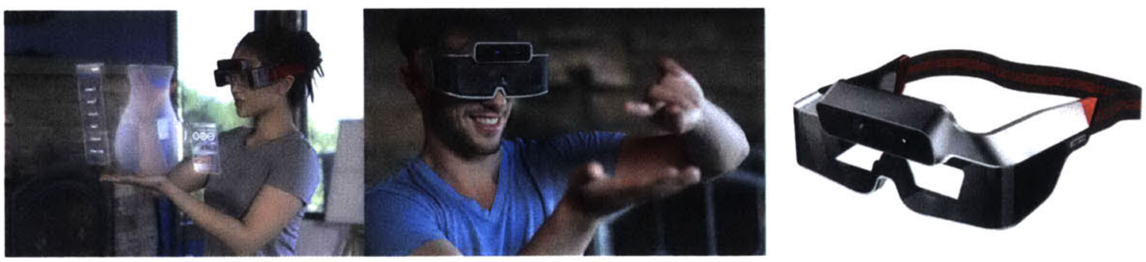

Figure 6-1: The Meta Spaceglasses

Figure 6-2: Microsoft's HoloLens

7

Conclusions

This thesis presented Variable Reality, an augmented reality-based book application designed for HMD displays that brings a unique on-the-go reading experience to the users. It attempts to combine the widely preferred physicality or spatiality of a paper book with the useful and portable nature of a digital one. It places great emphasis on the physical metaphor (or the paper book metaphor), a concept offering significant benefit to the user experience when applied appropriately. Two types of Variable Reality books were presented that are used distinctly based on the user scenarios: the *Augmented Book* and the *Virtual Book*.

The *Augmented Book* borrows the physical body of a real book. That is, it can instantly turn any paper bound book into a different book digitally. The book tracking systems implemented for Variable Reality stably tracks the book corners, using them to overlay paper pages with virtual pages. As a result, each page on a physical book can turn into a different page from another book. The physical turning of a page is replicated to the virtual page, which is animated along the moving edge of the real page. Several styles of natural hand gestures, easy to learn and use, are also part of the system and further facilitate intuitive reading.

The *Virtual Book* is a virtually augmented book that takes the complete 3-D form of a physical book. It adheres to the paper book metaphor through its realistic appearance of a book and gestural interaction styles similar to how one would interact with a physical book (e.g. posing, holding a book, flipping the pages). It targets a mobile reading user scenario in which there are no available physical books; it allows the user to access any book from a personal digital library from anywhere. It incorporates a number of useful navigation and manipulation techniques traditionally impossible with either paper books or digital books. Applications built for the *Virtual Book* manifest its practical and always-available nature.

The user study conducted on a group of participants demonstrated both the effectiveness and the drawbacks of the system in technical performance and in user experience. While both the Augmented Book and the Virtual Book received positive feedback on user experience, it was clear that the book and hand tracking systems had to be more robust in order to enable seamless interaction. Based on the feedback from the participants, limitations of the system are discussed as well as a number of possible future directions for the project.

Appendix A

User Evaluation Questionnaire

Background

- Name:
- Age:
- Gender: Female / Male
- Do you have a digital book or an eBook (on your smartphone or laptop)?
- Do you use it often?
- What aspects of it (a digital book) do you like and dislike?
- Do you prefer to read a physical book?
- What aspects of it (a physical book) do you like and dislike?

Augmented Book

- Did you find it easy or difficult to use?
- What aspects of the book did you find useful?
- What aspects of the book did you find frustrating?
- Did you find the gestures easy or difficult to learn?
- Did the gestures make sense to you?
- How would you compare this reading experience to reading a paper book?
- How would you compare this reading experience to reading an eBook?

Virtual Book

- Did you find it easy or difficult to use?
- What aspects of the book did you find useful?
- What aspects of the book did you find frustrating?
- Did you find the gestures easy or difficult to learn?
- Did the gestures make sense to you?
- How would you compare this reading experience to reading a paper book?

- How would you compare this reading experience to reading an eBook?

Combined

- Do you think you are likely to use the system again?
- If the Oculus Rift were not much more obtrusive than eyeglasses, would you be more likely to use it?
- Would you use it on the T (subway)? On a plane?
- Would you suggest the system to peer readers?
- Do you have any other comments?

Bibliography

[1] Wilson, Ruth. "The look and feel of an ebook: considerations in interface design." *Proceedings of the 2002 ACM symposium on Applied computing*. ACM, 2002.

[2] Chen, Nicholas, et al. "Navigation techniques for dual-display e-book readers."*Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2008.

[3] Grasset, Raphaël, Andreas Dünser, and Mark 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*. ACM, 2008.

[4] Sutherland, Ivan E. "A head-mounted three dimensional display." *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*. ACM, 1968.

[5] Grasset, Raphael, Andreas Dunser, and Mark Billinghurst. "The design of a mixed-reality book: Is it still a real book?." *Mixed and Augmented Reality, 2008. ISMAR 2008. 7th IEEE/ACM International Symposium on*. IEEE, 2008.

[6] Cheng, Kun-Hung, and Chin-Chung Tsai. "Affordances of augmented reality in science learning: suggestions for future research."*Journal of Science Education and Technology* 22.4 (2013): 449-462.

[7] Billinghurst, Mark, Hirokazu Kato, and Ivan Poupyrev. "Tangible augmented reality." *ACM SIGGRAPH ASIA* (2008): 7.

[8] Billinghurst, Mark, Hirokazu Kato, and Ivan Poupyrev. "Collaboration with tangible augmented reality interfaces." *HCI International*. Vol. 1. 2001.

[9] Darton, Robert. "The National Digital Public Library Is Launched!" *The New York Review of Books* (2013): n. pag. Web. 20, April 2015.

[10] Billinghurst, Mark, Hirokazu Kato, and Ivan Poupyrev. "MagicBook: transitioning between reality and virtuality." *CHI'01 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2001.

[11] Sin, Aw Kien, and Halimah Badioze Zaman. "Live Solar System (LSS): Evaluation of an Augmented Reality book-based educational tool." *Information Technology (ITSim), 2010 International Symposium in*. Vol. 1. IEEE, 2010.

[12] Wu, Chih-Sung Andy, Susan J. Robinson, and Ali Mazalek. "Turning a page on the digital annotation of physical books." *Proceedings of the 2nd international conference on Tangible and embedded interaction*. ACM, 2008.

[13] Becker, Chris R., *Marginalia: The Hybrid Textbook*. The Art College of Design: Media Design Program, 2010. Web. 27 Apr. 2015.

[14] MacFadyen, Heather. "The reader's devices: The affordances of ebook readers." (2011).

[15] Zhu, Hong-Min, and Chi-Man Pun. "An Adaptive Superpixel Based Hand Gesture Tracking and Recognition System." *The Scientific World Journal* 2014 (2014).

[16] Card, Stuart K., George G. Robertson, and William York. "The WebBook and the Web Forager: an information workspace for the World-Wide Web."*Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 1996.

[17] Fernández-Baena, Adso, La Salle-URL, and David Miralles. "Avatar: Tangible interaction and augmented reality in character animation." *on Interacting with Smart Objects* (2014): 28.

[18] Scherrer, Camille, et al. "The haunted book." *Proceedings of the 7th IEEE/ACM international Symposium on Mixed and Augmented Reality*. IEEE Computer Society, 2008.

[19] Card, Stuart K., et al. "3Book: a scalable 3D virtual book." *CHI'04 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2004.

[20] Back, Maribeth, et al. "Listen reader: an electronically augmented paper-based book." *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2001.

[21] Card, Stuart K., et al. "3Book: a 3D electronic smart book." *Proceedings of the working conference on advanced visual interfaces*. ACM, 2004.

[22] Schmandt, Chris M. "Pages Without Paper," Bachelors Thesis, Department of Computer Science, MIT, Cambridge, January 1979.

[23] Card, Stuart K., et al. "3Book: a 3D electronic smart book." *Proceedings of the working conference on advanced visual interfaces*. ACM, 2004.

[24] Robertson, George, et al. "Data mountain: using spatial memory for document management." *Proceedings of the 11th annual ACM symposium on User interface software and technology*. ACM, 1998.

[25] Gibson, James J., "The Theory of Affordances." n.d. Web. 20 Apr. 2015 < http://cs.brown.edu/courses/cs137/readings/Gibson-AFF.pdf>

[26] Cockburn, Andy, and Bruce McKenzie. "Evaluating the effectiveness of spatial memory in 2D and 3D physical and virtual environments." *Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM, 2002.

[27] Feiner, Steven, et al. "Windows on the world: 2D windows for 3D augmented reality." *Proceedings of the 6th annual ACM symposium on User interface software and technology.* ACM, 1993.

[28] Donelson, William C. "Spatial management of information." *ACM SIGGRAPH Computer Graphics.* Vol. 12. No. 3. ACM, 1978.

[29] Starner, Thad, et al. "Augmented reality through wearable computing." *Presence: Teleoperators and Virtual Environments* 6.4 (1997): 386-398.

[30] Jones, William P., and Susan T. Dumais. "The spatial metaphor for user interfaces: experimental tests of reference by location versus name." *ACM Transactions on Information Systems (TOIS)* 4.1 (1986): 42-63.

[31] Norman, Donald A. "Affordance, conventions, and design." *interactions* 6.3 (1999): 38-43.

[32] Lee, Taehee, and Tobias Hollerer. "Handy AR: Markerless inspection of augmented reality objects using fingertip tracking." *Wearable Computers, 2007 11th IEEE International Symposium on.* IEEE, 2007.

[33] Lee, Gun A., et al. "Immersive authoring of tangible augmented reality applications." *Proceedings of the 3rd IEEE/ACM international Symposium on Mixed and Augmented Reality.* IEEE Computer Society, 2004.

[34] Sinem, G., and Steven Feiner. "Authoring 3D hypermedia for wearable augmented and virtual reality." *2012 16th International Symposium on Wearable Computers.* IEEE Computer Society, 2003.

[35] Feiner, Steven, Blair Macintyre, and Dorée Seligmann. "Knowledge-based augmented reality." *Communications of the ACM* 36.7 (1993): 53-62.

[36] Höllerer, Tobias, et al. "Exploring MARS: developing indoor and outdoor user interfaces to a mobile augmented reality system." *Computers & Graphics* 23.6 (1999): 779-785.

[37] Colaço, Andrea, et al. "Mime: compact, low power 3d gesture sensing for interaction with head mounted displays." *Proceedings of the 26th annual ACM symposium on User interface software and technology*. ACM, 2013.

[38] Underkoffler, John, and Hiroshi Ishii. "Urp: a luminous-tangible workbench for urban planning and design." *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 1999.

[39] Ryokai, Kimiko, Stefan Marti, and Hiroshi Ishii. "I/O brush: drawing with everyday objects as ink." *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2004.

[40] Piper, Ben, Carlo Ratti, and Hiroshi Ishii. "Illuminating clay: a 3-D tangible interface for landscape analysis." *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2002.

[41] Grasset, Raphael, et al. "The mixed reality book: a new multimedia reading experience." *CHI'07 extended abstracts on Human factors in computing systems*. ACM, 2007.

[42] Lastra, Anselmo, et al. "Immersive electronic books for surgical training." *Simulation* 1.4 (1981): 29-68.

[43] Voleti, Kiran. "Tangible Augmented Reality Battleship Game." Web blog post. *Augmented Reality*. Real@Real.com, 21 Aug. 2013. Web. 1 May 2013.

[44] Li, Anita. "Tradition trumps tech: Textbooks better than e-books when it comes to learning." *LIFE* (2012): n. pag. Web. 20, April 2015.

[45] Fiorentino, Michele, Antonio E. Uva, and Giuseppe Monno. *Tangible interfaces for augmented engineering data management*. INTECH Open Access Publisher, 2010.

[46] Meta Spaceglasses *https://www.getameta.com/*

[47] Microsoft HoloLens *https://www.microsoft.com/microsoft-hololens/en-us*

[48] Vuzix Start 1200XLD glasses *http://www.vuzix.com/augmented-reality/products_star1200xld/*

[49] Boost C++ Libraries *http://www.boost.org/*

[50] Leap Motion SDK *https://developer.leapmotion.com/*

[51] OpenCV *http://opencv.org/*

[52] Oculus Rift *http://www.oculus.com/*

[53] Logitech HD Pro Webcam C920 *http://www.logitech.com/en-us/product/hd-pro-webcam-c920*

[54] Leap Motion controller *https://www.leapmotion.com/*

[55] Logitech HD Webcam C270 *http://www.logitech.com/en-us/product/hd-webcam-c270*