Turning a page on the digital annotation of physical books

Chih-Sung (Andy) Wu, Susan J. Robinson and Ali Mazalek
Synaesthetic Media Lab, GVU Center
Georgia Institute of Technology, Atlanta, GA 30332-0280 USA
+1-404-352-7200
{andywu, srobinson, mazalek}@gatech.edu

ABSTRACT

The Graphical User Interface (GUI) has created an efficient work environment for many applications. However, when users are confined by keyboards and mice, they lose the ability to interact with the virtual world using habits from the real world. Our research examine how emerging modes of authorship, such as wikis, can be used to generate new possibilities for bringing atoms and bits together for digital annotation. Our goal is to combine the everyday habits in reading books with emerging digital possibilities.

In this paper, we present a prototype system called WikiTUI, which brings digital media to physical paper books. This system allows readers to access the digital world through fingertip interactions on books, and enables them to share information with other readers using wiki technology. WikiTUI not only bridges the gap between the digital and the physical worlds, but also facilitates multiple contributions to a reference base spanning across these worlds. We present user evaluations of the WikiTUI prototype and discuss research implications.

Author Keywords

Tangible user interface, augmented books, annotation, wiki, CSCW, hypermedia, electronic books, fingertip detection, Human-Computer Interaction, augmented reality, computer vision, paper-based user interface, gestural input.

ACM Classification Keywords

H.5.1 [**Multimedia Information Systems**]: *Artificial, augmented, and virtual realities*; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces---*input devices and strategies, interaction styles*; H.5.3 [**Group and Organization Interfaces**]: Collaborative computing; I.7.2 [**Document Preparation**]: Hypertext/hypermedia.

INTRODUCTION

The concept of creating interconnections between separate pieces of information originated in 1945 when Vannevar Bush proposed a theoretical machine called Memex to help scientists organize complex data [7]. Even though Bush's vision was not realized at that time, it inspired the creation of hypertext, so coined by Ted Nelson in 1965 [20]. The idea of hypermedia continues to thrive with the continuing expansion of networked media and personal computing. Nonetheless, the immense cyber world is typically only accessible using keyboards and mice. Traditional media such as physical papers, books and magazines benefit from digital publishing, but in their physical form remain separated from the digital world by a heavy reliance on the existing form factor of personal computing interfaces. The link between physical and digital texts and media has been missing since the birth of hypertext.

Contemporary electronic media provides a number of advantages over physical paper, such as greater ease of editing, reproduction, transmission and storage. Nevertheless, such media has not eliminated paper from our daily lives. Studies have determined that reading from the screen is generally slower than from paper. Moreover, working on paper is more efficient than working with a computer in certain circumstances [10, 22], such as reading activities associated with writing tasks. However, the lack of a linkage between physical and digital media forces users to create workarounds to connect knowledge on paper with bits. Some create digital copies of physical documents, while others attach printouts of digital files to physical documents. Such problems call for the development of a platform to link the use of paper and books to the digital domain through a new form of hypermedia.

Our WikiTUI concept explores the possibility of manipulating digital media using physical media. Similar ideas have been proposed before, but the novel perspective of WikiTUI is the addition of distributed collaborative annotation through online wiki services. Digital information in different formats (videos, Web links, images) can be retrieved from a wiki server using a book as a key, with the information being displayed by projector near or on the book. Reading *Lord of the Rings*, a reader could watch Hobbits and elves dance across the page and explore the white towers and castles of Minas Tirith with the touch of a fingertip. WikiTUI thus uses physical books coupled with projected visual patterns as tangible user interfaces that provide direct manipulation and collaborative authoring of digital media.

RELATED WORK

Coupling Atoms and Bits

Many researchers have tried to bridge the gap between the digital and physical worlds. In 1997, Ishii et al. [13] presented their vision of "Tangible User Interfaces" (TUIs) to enhance the connection between cyberspace and the physical environment. Our method in this project extends their vision by emphasizing the use of familiar physical objects, books, in a non-destructive way, rather than creating new or exotic interfaces and interaction metaphors to connect the two realms.

Among all TUI related projects created by the pioneers of tangible user interfaces, there are many applications utilizing cameras and projectors as input and output devices [30, 31]. Even though these projects make use of phidgets to interact with the digital world, the realization of coupling atoms with bits using projected light led us to brainstorm more on the applications of projected light patterns. Another domain that combines vision, graphics and HCI is Tangible Augmented Reality [14], which applies tangible user interface techniques to augmented reality environments. This work also informed our exploration.

Hypermedia

Hypermedia has been presented as the medium to store and distribute the world's entire literary production [21], as a common development space for programmers to work on heterogeneous software development environments [2], and as a tool for writers to engage in composing narratives [6]. Sinclair et al. proposed to combine the real and virtual worlds using commonplace and real world metaphors in an augmented reality setting to facilitate contextual linking in hypermedia [27]. In our application, we propose to connect to hypermedia through simple hand movements rather than mice and keyboards.

Human-Paper Interfaces

In his project Digital Desk [32], Wellner tried to merge the physical desktop and documents with their virtual counterparts by using digital projections on the desk surface and computer vision to capture user actions. Instead of projecting on a desktop and paper, WikiTUI uses books as the screen which receives projections. Moreover, WikiTUI connects to the Internet and makes use of the self-maintaining wiki community. Yet, we are indebted to Wellner's original idea of using fingers for certain tasks as an alternative to traditional pointing devices.

Another related project which informed our work is the EnhancedDesk designed by Koike et al. [16]. This work uses IR cameras to capture user motion as input. The researchers have also developed a method for real-time finger tracking [17]. In contrast to our work, their emphasis is focused strictly on the real-time tracking of fingertip interaction by the user at a desktop, whereas we attempt to apply their ideas to a broader application. More recently, Escritoire has brought the idea of collaborative editing to a projected display [3]. Their system has demonstrated positive results in user studies, showing that users can adapt themselves to the system without extra training.

Gestures as Inputs

Hardenberg et al. [11] achieved direct manipulation on a projected screen through fingertip interactions using a novel algorithm. In this research, a user uses her fingertip as a substitute of a mouse cursor to control the computer. This is similar to our idea of controlling digital media through the user's gestures. In addition to designing an algorithm to recognize fingertips, they also utilized their fingertip and gesture interaction design for manipulating content on projected screens.

Learning and recognizing gestures have become another mode of human-machine communication [24]. A number of scientists have developed models for the interpretation of gestures. Interpretation of gestures is critical to our system's success. A desired seamless interaction between the user and WikiTUI largely depends on a robust gesture recognition model. Many works have addressed the design of gesture interactions based on different models. For example, in the case of symbolic gesture analysis, motions of movements have been interpreted using the Hidden Markov Model (HMM) [18, 28, 33] or Finite State Machine [5, 8, 12].

WIKITUI DESIGN

The WikiTUI architecture is envisioned to sit on a specially instrumented desk that provides overhead projection onto the book and desktop surface as shown in Figure 1.

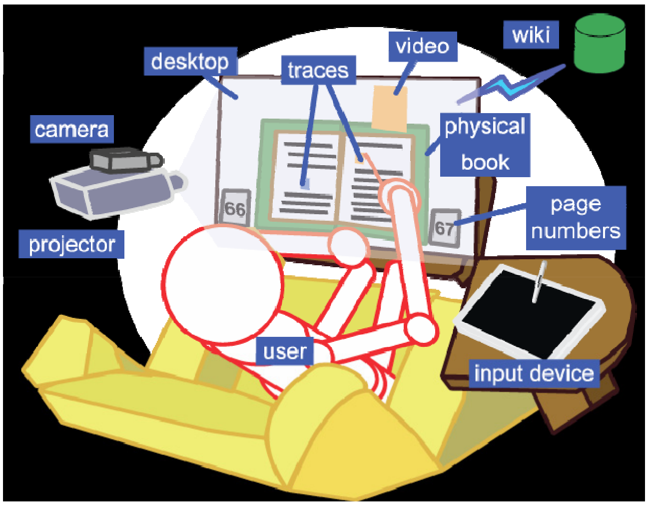

Figure 1. A camera captures the user's finger gestures on the desktop and book. These finger movements are used to access digital media. An input device is used to enter new annotations, which results in new traces being displayed.

An overhead projector projects digital multimedia onto a physical book, while an overhead camera records interactions between the user and the multimedia interface (see Figure 2). The digital multimedia elements are drawn from a wiki database, which stores the multimedia content and metadata, including an icon, or projected *trace* to be displayed at the edges of the book. The traces, which change page by page, cue the reader that multimedia content is available. A separate tablet or laptop interface could be used to enter new annotations. In this form, WikiTUI might be installed in a public library or a user's home. In the library setting, readers might retrieve data from rare books that are only available on site, with the ability to access, generate, and leave behind rich interpretations through digital annotation.

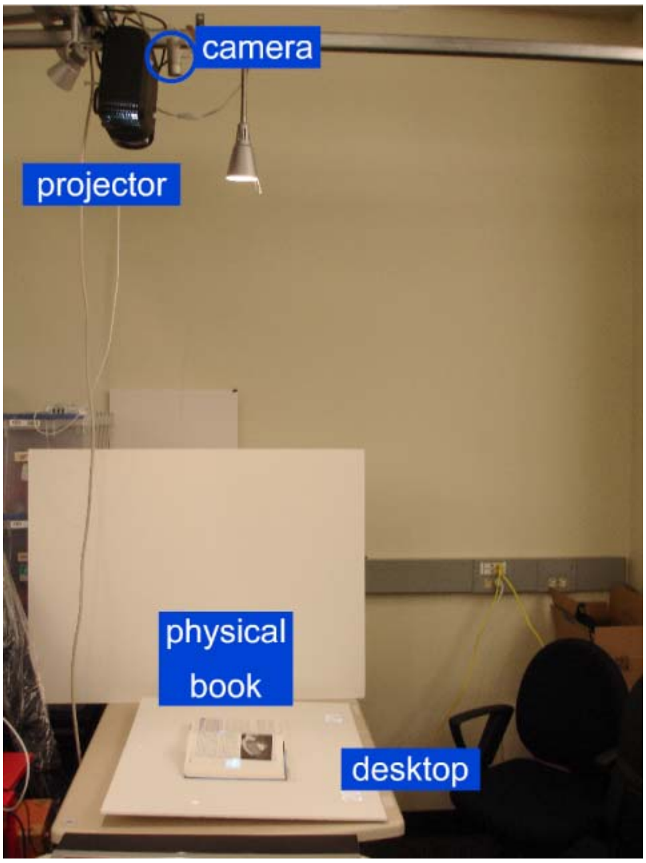

Figure 2. A side view of the WikiTUI prototype.

To simplify the alignment task for computer vision and content display, two rods lying horizontally at the bottom edge and vertically at the leftmost edge confine the book's position (see Figure 3).

We also envision a second form factor that could be used in areas where no special desk or projector system is available. In this case, the output would be provided not on the surface of the book itself, but on a separate display provided by a PDA. The PDA's built-in camera could be moved over the book's pages in order to retrieve and view the annotations on-screen.

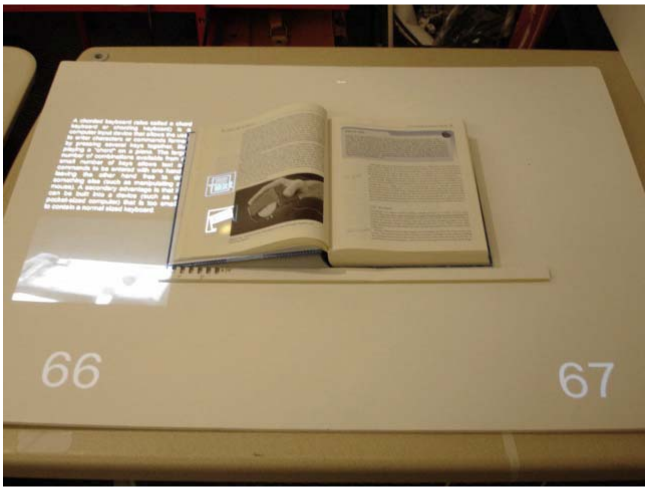

Figure 3. The book's bottom edge is aligned to the horizontal rod on the desktop. There are two *traces* projected on the book and their corresponding media showed at the desktop margin. One media object is a paragraph of text and the other is playing the so-called "mother of all demos" video by Douglas Engelbart in 1968 [9].

User Scenario

We envision WikiTUI for public libraries or home settings, in the form of an augmented desk that provides interaction via an overhead camera and projector. After selecting a book to read, a user carries her book and laptop to WikiTUI and sits down. The overhead camera detects the movement and turns on the projector. Next, the user scans her ID card and the barcode of the book at a specific scan zone projected onto the desktop. Once both the user and the book's identities have been recognized, WikiTUI displays the corresponding information. As she reads, traces indicating available digital content are projected onto the book's pages. The user may use fingertip movements on the book to view a video. Perhaps she finds a passage that reminds her of a video she saw online. She uses her nearby laptop to look up the video and she uploads it to the wiki server. The annotation is immediately projected onto the book and she drags it to the desired location on the page. Since the database is connected to Internet, other readers of this book can see her entered data as well. By moving her fingertip to the trace, she watches the video attached to the trace.

Implementation

In order to realize the scenario described above from a system perspective, we devised methods to implement the following functionality: authentication, page identification, gesture interaction, and annotation. Our approaches are described here. An overarching design goal was to enable a seamless transition for readers from their typical reading behavior to reading and retrieving multimedia content in essentially augmented books.

The WikiTUI software architecture includes a wiki server that stores the data and the user end interface. The wiki server stores data that helps deliver multimedia information shown on a book. This includes a book's ID, page number, and metadata such as a trace cue's position in a certain page, type of multimedia, contents of multimedia data and the creator's information. The data entered into WikiTUI is in two forms: multimedia content with associated metadata that is contributed manually by the user, and trace position data captured by the system, which require image recognition algorithms to interpret.

We implemented the system using Java with multithreaded programming. An isight Firewire camera right above the system captures the user's movements on the desktop. To map points on the image plane to the corresponding coordinates on the projection plane, we adopted the projector-camera homography method proposed by Sukthankar et al. [29]. With their proposed eight-point calibration method in addition to the alignment of our projector and camera (see the upper left in Figure 2), we were able to map the coordinates even they were on two different resolutions of planes.

Separating out skin tones within a video signal generated by a regular camera is a tough task due to a host of issues. Hardenberg et al. [11] achieved direct manipulation on a projected screen through fingertip interactions using a novel algorithm. Their method is especially suitable for a system with a projector and a regular camera such as WikiTUI. The fundamental idea of their method is that a finger is actually composed of a circle and a rectangle. When an algorithm to detect such shapes is repeatedly applied, which gradually adds different weights to the target image, the system can discriminate foreground hand shapes from background noise. During this process, we also applied morphological operations such as closing or opening to enhance hand shapes. Their algorithm can be used recognizing different fingers by calculating the width of each finger. Yet we implement a system with single fingertip interaction in our first prototype.

Authentication

Like a regular wiki service, WikiTUI requires each user's identity for data entry. For a WikiTUI system deployed in a public environment, this could be accomplished by scanning the user's ID card through either an RFID or bar code reader. For a home installed WikiTUI system, the home network should provide sufficient information about user identity, profile and preference settings to the system. Currently, users log into the system by entering user IDs using keyboard and mouse interaction.

In the future, a scanner could be used to read the barcode on the book's cover to identify the book, which has been a reliable method adopted by many bookstores. At the present time, users have to enter the ISBN manually. The system creates a new book record if the entered number cannot be located in the wiki database.

Gesture-based interactions

In our initial design, users move their fingertips to manipulate the traces left on a book. Basic interactions include creating a trace, moving it to a new position, enlarging it to a viewable size and turning pages in the virtual space. In this finger-paper interface, we envision that a pie menu could provide additional functionality.

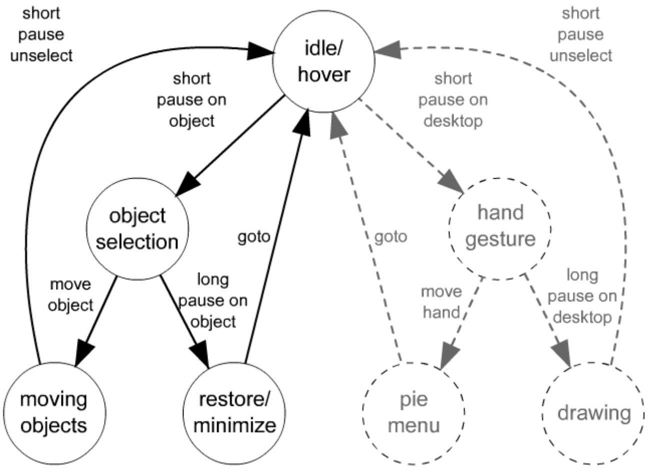

Figure 4. The state transition diagram of WikiTUI. Shapes with dotted lines represent future features of the system. All users start at the "idle/hover" state. When a user pauses her fingertip on top of a trace for 2 seconds, the system enters the "object selection" state. If she pauses on the object for another 2 seconds, then the associated digital medium will either start playing or become minimized depending on the system's status; the system then arrives the "restore/minimize" state. If she decides to move her fingertip instead of pausing longer, then the selected object is also moved; thus the system become in the "moving objects" state.

Possible finger interactions are shown above in the diagram in Figure 4 and are grouped into five categories:

- *Interactions based on action time*: A one-second pause on a "*trace*" maps to a GUI click (selection), while a longer pause on the target suggests a double click (open/play). A user stops her fingertip on top of a video trace and opens the associated video in Figure 5.
- *Interactions based on hand shapes*: A fist shape, namely a no-fingers-jut-out shape, represents a stop or abort. An index finger acts as a pointer, while a "V" sign (index finger combined with middle finger) stands for zooming.
- *Interactions based on drawing*: Drawing a triangle might restore an annotation to its initial state. Left and right arrows could be symbols for backward and forward movement through multimedia files.
- *Interactions based on movement*: Moving the "V" sign upward might zoom in the whole view, and vice versa. A hand wave outward might hide all traces on a page.
- *Interactions with two hands involved*: Placing two index fingers at the upper right and lower left corners of a trace and dragging in/out could resize the annotation.

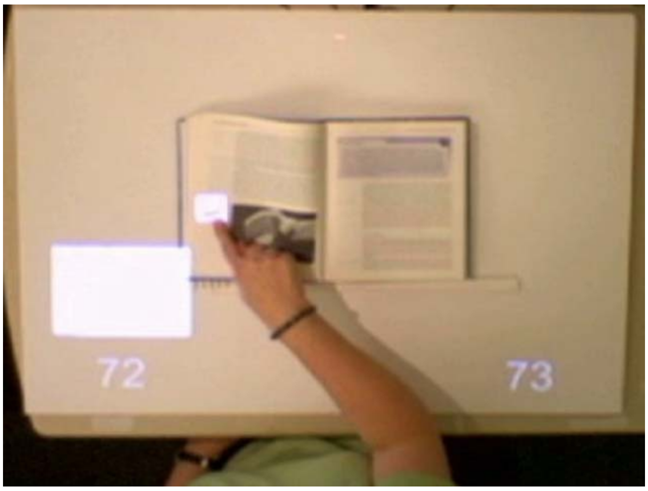

Figure 5. WikiTUI prototype showing a user interacting with *traces* on the book with her fingertip to access hypermedia.

Page identification

There has been work on determining the page numbers of a physical book [4, 19]. However, these involve the redesign of paper, which is not desired in our vision. An alternative approach is to use a camera to search for and identify the number on each page. However, in the WikiTUI prototype, rather than detecting page numbers from the physical book, the system projects page numbers corresponding to the location of digital annotations. The page numbers are projected at the two bottom corners of the desktop on either side of the book. User must manually match the page numbers for *digital annotations* to whatever page they are reading in the book using the projected page numbers. Thus the projected page numbers act as page turn buttons, and are used to keep the displayed annotations in sync with the physical book page as shown in Figure 6.

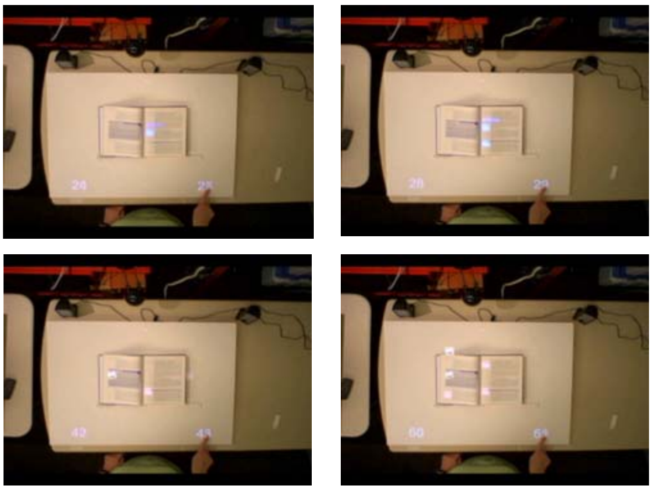

Figure 6. A user places her index finger on the projected page number at the bottom right to advance pages of digital annotations. She cycles between turning these digital pages and turning the physical pages. As she turns virtual pages, the projected contents also change, resulting in different projection patterns in the photos.

Annotation

Based on early observations and brainstorming with users, we believe that typing and handwriting are easy and reliable methods for text input. Since most annotations might be on-line media URLs or long text comments, both laptop and tablet computers would work well. We expect that users could carry their preferred input devices to WikiTUI. These devices would be used to create annotations through an on-line wiki interface, resulting in a new trace being projected onto the desk surface. The user would then drag the trace to the desired location, for example, next to a related phrase.

For the current prototype, users log on to a website to create their annotations. A user enters her username, the page to be matched to the digital annotation, and a link to on-line hypermedia. After that, she drags the trace to a new location.

EVALUATION AND DISCUSSION

We first conducted a formative evaluation to inform WikiTUI design. After building a semi-functional prototype, we then performed a user evaluation to inform the next iteration of the system.

Formative Evaluation

We observed readers in the Georgia Institute of Technology library, which integrates aspects of a computer center with its stacks storing books. Most students use computers in this space, significantly altering our notion of the typical "reader" in a library. This discovery lead us to believe that the audience of WikiTUI should be groups of readers that regularly use tools or technology to support the reading of physical books they are reading in common.

Based on our initial observations, we believe that reading behaviors change when the type of book changes. In other words, when reading a novel, a reader only reads; for a history book, he/she annotates or marks passages a lot; and when studying calculus, the reader may spend most of the time calculating.

User Evaluation

Next, we sought to discover if the system was usable, particularly if users found its operation seamless. We also were interested to see if users felt such a system would enhance their reading experience.

Even though a related research [4] has proved that fingertip tracking using regular cameras is feasible and the method of skin tone selection is adopted in many studies [1, 9, 15], our situation of introducing a physical book as part of the background image complicated our work. In our pilot test, we found that users tend to put their arms on the desktop and push the desktop. Moreover, the book's pages are not flat, and they contain different colors of figures and texts. Sometimes the reading and page turning were so fast that the system could not have time to establish new information for background subtraction. In brief, the image surface was not a flat plane when a user did not flatten the pages or put her arms on the desktop. In short, our system worked in the ideal environment but not in a practical condition.

Consequently, to obtain preliminary user feedback, we conducted user evaluations with our prototype using Wizard of Oz techniques. An experimenter sitting at a computer away from the system played the wizard behind the scene, while another experimenter conducted the user study. The wizard watched the user's movements from a screen and controlled the cursor, which is projected onto the desktop of WikiTUI to follow the user's fingertip. The software is connected to WikiTUI's wiki server and retrieved the information when necessary. We did not inform the subjects of this technique, but told them we were testing a semi-functional prototype.

We drew a convenience sample of users from an HCI introductory class to ensure that our users had in common a text, learning resources, and shared knowledge of a subject. The protocol called first for the administration of a 20-minute semi-structured interview regarding the user's current reading and annotation habits, in conjunction with looking at the readers' artifacts such as their textbook, class handouts, and notebooks. Next we familiarized the user with the system and then conducted a 20-minute cooperative think-aloud featuring four tasks, followed by a 10-minute debrief. The tasks users performed were to: read the book, change pages, use traces to play videos, and add a trace with rich media.

We confirmed our expectation that our readers read texts in a variety of locations: at desks in labs, on beds at home in the evening, and coffee shops. Our sample featured some students for whom English was a second language (ESL). These students said they learned mainly through the use of class handouts and annotated these rather than their books. Both ESL subjects wrote annotations in their native languages. All readers used the Internet while reading, with one ESL student noting that she used on-line dictionaries to understand words and hear them pronounced. When looking at pages related to the class readings, students did not have similar methods of saving links: some saved to browser "Favorites," one downloaded pages, and another created summaries with lines within. The greatest sharing of notes occurred around test taking with study partners.

The think-aloud revealed several consistent issues with the interface that need to be adjusted in future prototypes and systems. The projection of page numbers on both sides of the book presented difficulties for the users in trying to locate a specific page by pointing: they were required to go from one side to the other to cue the system. Users also consistently tried to "tap" their way to pages; pointing (leaving their finger motionless) until the system reached the page they wanted was not natural. Users readily accepted the use of input devices such as the keyboard and stylus pen.

In terms of reading, one user reported that the projected traces were distracting, particular text versus iconic traces. One reader noted that he would like the option of annotating the projected traces, for example, to note that the trace contained something worthwhile. All users mentioned that when playing videos they want to have full control of the video player (stop, rewind, play) versus no controls. When adding content, users wanted the flexibility to title the traces and content they contributed, versus a set format (author's name). One user mentioned the desire *to keep some things for myself and others to share.* Another user observed: *If lots of people add text and video, it could become confusing* [too many traces].

In the debrief, the general consensus was that the WikiTUI prototype was intriguing, but needed refinement. One mentioned that he would like to be able to virtually "highlight" text. Other themes included wanting elements (like page numbers and traces) to be near or on the book versus on the desktop and for the book to be near the reader's body.

RESEARCH IMPLICATIONS

Our user evaluation produced a number of potential areas for additional research that can inform future work on this project.

1. The confirmation by our users that they had a variety of different reading modes, across the physical and digital domains, and that all users used the Internet supports the vision of a WikiTUI system for linking materials across the digital and physical spaces. The affordances of WikiTUI could be particularly relevant to users in the context of reading/studying and sharing annotations within groups of readers.
2. However, the fact that people read books in different places suggests that a mobile version of WikiTUI, where media could be displayed on a camera-enabled mobile device rather than projected on a book at a fixed WikiTUI station, should be explored. In fact, similar usage of mobile devices, such cell phones' cameras as barcode scanners to retrieve online information has been quite popular in Japan [26].
3. Our good fortune in recruiting ESL students resulted in our recognizing that an augmented book could provide unique utility to those who may need translation support, particularly the match of spoken language with text. The suggestion that the effects of the traces may vary depending on the type of trace projected (text or image), particularly in terms of disrupting attention, warrants additional investigation into the design of the traces and their effects.
4. Although having page numbers on either side of the book mirrors the design of physical books, this feature presented difficulties for all users. This suggests that all controls, such as page numbers, and projected content should be placed in a consistent place on the desk, rather than being displayed near the associated phrases in the book (although the traces would still be projected

near related text.) The impulse for readers to "tap" to the next page might be translated into a discreet finger movement, such as a wiping motion (simulating page turning) to overcome problems with leaving the finger still.

5. Also, users brought up the problem of too many traces causing clutter or becoming distracting. Accordingly, methods for hiding certain traces or filtering the annotation material based on user preferences or queries would be important for a fully implemented system.
6. Finally, additional interview data and a review of existing literature regarding privacy is needed to determine if more users would also share the desire for a feature that allows for them to create content "for their eyes only" versus content to share. As a matter of fact, lack of preference settings in this prototype forces users to change their reading habits.

CONCLUSIONS

The introduction of windows, icons, menus and pointing devices in GUIs has provided a mechanism for users to interact with digital media applications. At the same time, it has also created a gap between the physical and digital worlds. In this paper, we have proposed the WikiTUI infrastructure to realize the goal of bridging these two realms seamlessly within the scope of reading. As such, we have adopted certain GUI approaches for inputting annotations, and have coupled these with natural reading interactions with physical books that are augmented with digital annotations. Yet the technology is not the main focus of this research. Our purpose is to extend the concept of coupled atoms and bits using a non-destructive method, and by so doing extend the use of paper books into new realms.

In our initial design and development, we have found that challenges in computer vision and image recognition still present problems for creating a truly robust system. Although our current prototype is still far from a fully functional system, our evaluation data supports the notion that WikiTUI could enhance the reading experience for certain types of users. For example, readers from different regions of the world might create their own culturally-relevant content in different languages, as our ESL users suggested. Such content could then be tailored to others sharing the same cultural backgrounds when they read the same books.

Further, it is also possible that writers or publishers might want to create some permanent traces associated with their printed books. For example, a magazine might add a trace of a video clip and embed it in the content. Readers could then watch the video when reading the magazine. Consequently, magazine contents could be more than just text and figures, and contents from online articles could be linked to the printed versions. Similar applications may affect the use of textbooks as well.

While such a vision is ambitious, the new constellation of computer vision, networked media, and wiki technologies available today brings new possibilities for coupling atoms and bits. Combined with the affordances of traditional media, these developments could enable us to turn a page on the digital annotation of physical books and other printed media.

ACKNOWLEDGEMENTS

The authors wish to thank the reviewers for their helpful comments and synlab members, who helped shape our ideas. They also thank all the participants in the user study. Moreover, they would like to thank Dr. Jim Foley and the Learning Technology Lab members at GVU Center for their inputs in the early stage of the project. Georgia Institute of Technology GVU Center and School of Literature, Communication & Culture have been continuously supporting synlab and its projects. The authors thank to the faculty, staff and students as well.

REFERENCES

1. Ahmad, S., A Usable Real-Time 3D Hand Tracker. *IEEE Asilomar Conf. 1994*, (1994), 1257-1261.
2. Anderson, K.M., Taylor, R.N., and Whitehead, E.J. Chimera: hypermedia for heterogeneous software development environments. *ACM Trans. Inf. Syst. 18, 3* ACM Press (2000), 211-245.
3. Ashdown, M. and Robinson, P. Escritoire: a personal projected display. *Multimedia, IEEE 12, 1*, IEEE CS Press (2005), 34-42.
4. Back, M.J. and Cohen, J. Page detection using embedded tags. *Proc. UIST 2000*, ACM Press (2000), 159-160.
5. Bobick, A.F., Wilson, A.D. A State-Based Approach to the Representation and Recognition of Gesture. *IEEE Transactions on Pattern Analysis and Machine Intelligence 19, 12*, (1997), 1325-1337.
6. Bolter, J.D., *Writing space: the computer, hypertext, and the history of writing*. Lawrence Erlbaum Associates, Inc., Mahwah, NJ, USA, 1991
7. Bush, V. As we may think. *Interactions* 3, 2 (1996), 35-46.
8. Davis, J. and Shah, M. Visual gesture recognition. *Vision, Image and Signal Processing 141, 2*, (1994), 101-106.
9. Doug Engelbart 1968 Demo. http://sloan.stanford.edu/MouseSite/1968Demo.html
10. Hansen, W.J. and Haas, C. Reading and writing with computers: a framework for explaining differences in performance. *Commun. ACM 31, 9*, ACM Press (1988), 1080-1089.
11. Hardenberg, C. and Bérard, F. Bare-hand human-computer interaction. *Proc. of the 2001 Workshop on PUI '01*, ACM Press (2001), 1-8.

12. Hong, P., Huang, T.S. and Turk, M. Gesture Modeling and Recognition Using Finite State Machines. *Proc. Int’l Conf. Automatic Face and Gesture Recognition 2000*, (2000), 410-415.
13. Ishii, H. and Ullmer, B. Tangible bits: towards seamless interfaces between people, bits and atoms. *Proc. CHI 1997*, ACM Press (1997), 234-241.
14. Kato, H., Billinghurst, M., Poupyrev, I., Imamoto, K. and Tachibana, K. Virtual object manipulation on a table-top AR environment. *Proc. ISAR'00* (2000), 111-119.
15. Kjeldsen, R. and Kender, J., Finding Skin in Color Images. *Proc. Int’l Conf. Automatic Face and Gesture Recognition 1996*, (1996), 312-317.
16. Koike, H. and Kobayashi, M. EnhancedDesk: Integrating Paper Documents and Digital Documents. *Proc. APCHI 1998*, ACM Press (1998), 57-62.
17. Koike, H., Sato, Y. and Kobayashi, Y. Integrating paper and digital information on EnhancedDesk: a method for realtime finger tracking on an augmented desk system. *TOCHI 8, 4*, ACM Press (2001), 307-322.
18. Martin, J., and Durand, J-B. Automatic Handwriting Gestures Recognition Using Hidden Markov Models. *Proc. Int’l Conf. Automatic Face and Gesture Recognition 2000*, (2000), 403-409.
19. May, K.W. Conductive ink based page detection for linking digital and physical pages. *Ext. Abstracts CHI 2001*, ACM Press (2001), 275-276.
20. Nelson, T.H. Complex information processing: A File Structure for the Complex, the Changing, and the Indeterminate. *Proc. of ACM National Conference*, ACM Press (1965), 84–100.
21. Nelson, T.H. *Literary Machines*. Swarthmore, PA, USA, 1981.
22. O'Hara, K. and Sellen, A. A comparison of reading paper and on-line documents. *Proc. CHI 1997*, ACM Press (1997), 335-342.
23. Oka, K., Sato, Y. and Koike, H. Real-Time Tracking of Multiple Fingertips and Gesture Recognition for Augmented Desk Interface Systems. *Proc. Int’l Conf. Automatic Face and Gesture Recognition 2002*, (2002), 429-434.
24. Pavlovic, V.I., Sharma, R., Huang, T.S. Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence 19, 7*, (1997), 677-695.
25. Quek, F.K.H., Mysliwiec, T. and Zhao, M. Finger Mouse: A Freehand Pointing Interface. *Proc. Int’l Workshop on Automatic Face and Gesture Recognition 1995*, (1995), 372-377.
26. Rao, B. and Minakakis, L. Evolution of mobile location-based services. *Commun. ACM 46, 12*, ACM Press (2003), 61-65.
27. Sinclair, P., Martinez, K., Millard, D.E., and Weal, M.J. Links in the palm of your hand: tangible hypermedia using augmented reality. *Proc. of the Thirteenth ACM Conference on Hypertext and Hypermedia*, ACM Press (2002), 127-136.
28. Starner, T. and Pentland, A. *Visual Recognition of American Sign Language Using Hidden Markov Models*. M.I.T. Media Laboratory, Perceptual Computing Section, Cambridge MA, USA, 1994.
29. Sukthankar, R., Stockton, R., and Mullin, Smarter Presentations: Exploiting Homography in Camera-Projector Systems. *Proc. Int'l Conf. Computer Vision '01*, (2001), 82-87.
30. Underkoffler, J. and Ishii, H. Illuminating light: an optical design tool with a luminous-tangible interface. *Proc. CHI 1998*, ACM Press (1998), 542-549.
31. Underkoffler, J., and Ishii, H. Urp: A Luminous-Tangible Workbench for Urban Planning and Design. *Proc. CHI 1999*, ACM Press (1999), 386-393.
32. Wellner, P. Interacting with paper on the DigitalDesk. *Commun. ACM 36, 7*, ACM Press (1993), 87-96.
33. Yamato, J., Ohya, J., and Ishii, K. Recognizing human action in time-sequential images using hidden Markov model. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition 1992*, (1992), 379--385.