A Document Browser Based on a Book-Style Interface with Augmented Reality

[1]Masatoshi NISHIMURA, [2]Tsuneo KAGAWA, [3]Hiroaki NISHINO, [4]Kouichi UTSUMIYA

Department of Computer Science and Intelligent Systems, Oita University
700 Dannoharu, Oita 870-1192, Japan.
{[1]v12e3021, [2]t-kagawa, [3]hn, [4]utsumiya}@oita-u.ac.jp
[4]*currently with Institute for Hypernetwork Society, Oita, Japan.*

***Abstract*— We propose a document browsing method based on a book-style interface usable in a normal office environment. Augmented reality (AR) technology is applied to realize an easy-to-use interface. In the proposed method, documents and multi-media contents are projected on a real booklet including an AR marker printed on each page. Users can feel that they are reading a book consisting of document files, presentation files, movies, and images. Furthermore, we utilize the motion sensor Kinect to recognize the users' gestures. The detected gestures can be used to activate some functions such as picking a page up and placing it in a real work space. In this paper, we describe the proposed book-style interface, its implementation method, and evaluate and discuss the effectiveness of the proposed method.**

Keywords- Augmented Reality; book-style interface; gesture interaction; graphical user interface.

I. INTRODUCTION

Rapid progresses of graphics computing and networking performance allow users to easily acquire any information anywhere at any time. Additionally, they can comfortably browse and edit multiple contents consisting of texts, images, sounds, and movies in a multi-window environment in an office or at home. Whereas this multi-window based multitasking environment enhances the convenience of various work, it also confuses the users and requires them to put extra efforts on non-essential tasks such as frequent switching operations between the windows. This is the same situation as wasting time on finding a target document and identifying an important part in the document on a chaotic desk. Connecting multiple displays and printing and lining up related documents on the desk cannot improve the situation. Therefore, a method for collectively managing all contents and providing easy-to-use manipulation functions is required.

As a traditional way of managing and searching information in a specific field without using computers, we commonly collect and file paper documents. Although it is inefficient compared with a high-speed multimedia processing function by computers, browsing information with prints has the following advantages.

- Readability: If a user cannot set a clear index or keyword for searching information, he/she can examine each document page-by-page. Because such examination tends to be a lengthy task, it encourages the user for better understanding each document's principal contents.
- Intuitive operation: The user can roughly point a target page by using the thickness, weight or texture of a stack of papers as a clue. Such muscle memory used in reading a book enables the user to adequately flip pages and intuitively find the desired information.
- Easy customization: The user can easily change the order of print files and select his/her preferred documents.

As shown in Figure 1, we propose a document browsing method designed for managing multimedia contents with texts, images, and movies having the above mentioned advantages of the paper files. The proposed method uses a book-style interface for flexibly displaying multimedia documents with AR (Augmented Reality). The interface is a real booklet and each page has a different AR marker printed on it. The user wears an HMD (head mounted display) with an attached web camera for browsing the documents displayed by AR. When the user spreads a pair of pages, the web camera captures two AR markers printed on the two-sided pages, and then the system superimposes a specific document page on each marker. It allows the user to pick some pages in the booklet and place them in his/her surroundings as shown in Figure 1. We adopt the Kinect motion sensor for easily selecting and placing a set of pages in the file. The proposed method enables the users to spontaneously browse, search, select, and reuse his/her desired information in his/her desktop work space with natural motions used in reading books.

II. RELATED WORK

There are various trials for effectively assisting document browsing and information search. Many of these trials devise display method in a monitor screen such as 3D desktop environments. Other approaches apply digital book technology usable on various mobile terminals.

A. 3D Desktop Environment

A 3D desktop environment extends a standard 2D desktop GUI into a 3D counterpart for spatially laying out multitask windows and efficiently switching among them. BumpTop [1] presents a desktop environment in a 3D virtual space and lightens an operator's burden for managing multiple overlapped windows. It uses a mouse in the same way with the standard 2D desktop, but manipulating the 3D desktop with the 2D I/O device is a difficult task. The recent

978-0-7695-4992-7/13 $26.00 © 2013 IEEE
DOI 10.1109/CISIS.2013.133

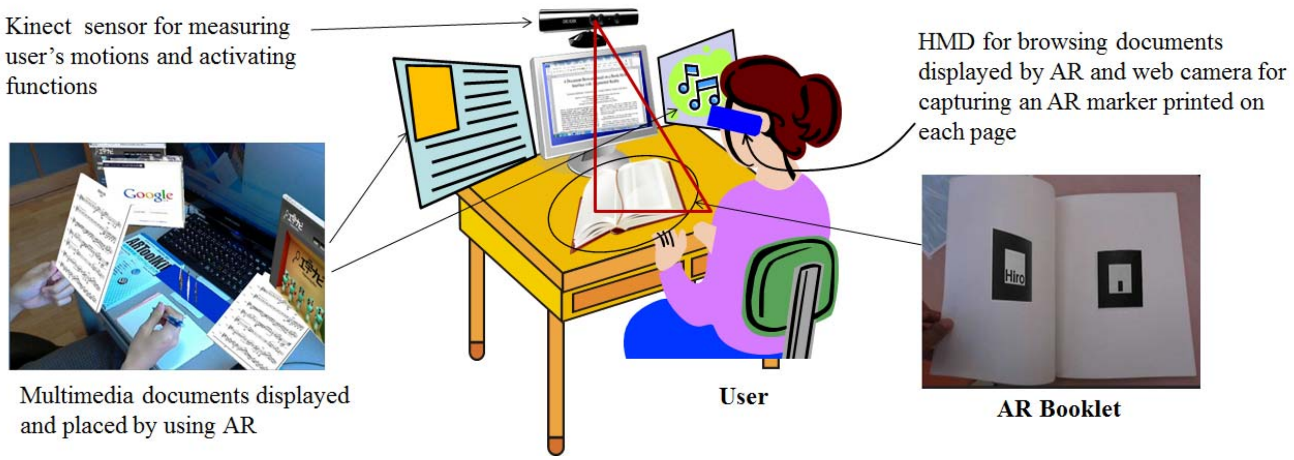


Figure 1. A proposed document browsing environment based on a book-style interface using AR and a motion sensor.

approaches, therefore, incorporate various devices such as 3D joysticks, touch screens, and motion sensors for easily controlling the 3D spatial operations. In our laboratory, a 3D desktop environment operable by using a haptic device was designed and developed [2]. Figure 2 shows a captured screen of the 3D desktop interface designed and developed in our laboratory. It allows a user to spatially operate various objects arranged in a 3D virtual space. When the user touches an icon or window, the system presents a reaction force enabling him/her to stably grab, move, and place it anywhere in the space. The screen size, however, is a limiting factor for efficient manipulations in the case of browsing many documents at the same time.

Figure 2. An example 3 dimensional desktop screen.

B. *Digital Book*

Various mobile devices suitable for browsing documents such as tablet PCs and e-book terminals recently become available. Browsing the large number of documents on the lightweight small devices turns to be a common practice. The digital book technology forms an international market and many researchers draw attention to its potential use for new human computer interaction (HCI) method. There are some comparative investigations between digital and paper books on the impact on sentence comprehension and memory in reading [3][4]. These studies concluded the paper book is better for understanding and memorize the book contents than the digital book. A reason is the digital book prohibits readers from remembering the book contents because of its simple and very convenient usage style. The results of the "digital vs. paper" comparisons, however, depend on various factors such as users' familiarity and experiences with the devices, and user-friendliness on page flipping and navigation functions. Based on these studies, implementing paperless environments with effective data management method will be a key issue for cost reduction and efficiency improvement in near-future office environments.

III. Basic Technology for Implementation

In this section, we summarize basic technologies adopted in implementing the prototype system elaborated in the next section.

A. *Augmented Reality*

The proposed system implements document display function using marker-based AR. The system captures a pair of markers printed on double pages, measuring their positions and angles in the working space from the shape and size of the captured marker images, and superimposing the contents of files and images as matched with the detected markers' positions and angles. Figure 3 shows an example case for displaying and projecting some document pages in a real working space using AR with the marker detection function. The user can naturally observe the presented documents through a body-worn HMD device. We use AR Toolkit [5] for implementing the document browsing environment in the real 3D space as shown in the figure.

B. *Real World Oriented Interface for Document Browsing*

Real world oriented interface adopts the routinely used activities of everyday life for implementing a natural HCI method. Borrowing the book reading activities for managing

Figure 3. Documents displayed and projected in a real working space using AR.

and browsing multimedia data is a good example for realizing an intuitive HCI method.

There are some approaches for implementing a natural "book reading" interface usable on various mobile terminals. Providing an intuitive "page flip" interface is a key design issue. Preceding studies adopted sensor devices for detecting the bending angle and shearing force occurred in flipping papers and fed them back to document output on the display [6][7][8]. Many of these trials, however, only focus on the flip operation interface part and lack a view on the smooth integration of the interface and the browsing parts. Flip interface [9] is a trial for attacking this problem by integrating a flip sensor with an iPad terminal. Because users commonly operate tablets by their fingertip, their fine and thin operations tend to provoke malfunctions. The proposed system resolves the problem by using real papers for presenting the realistic feeling of the page flip operation.

C. *Gesture Recognition by Motion Sensor*

The proposed system employs the Kinect motion sensor for supporting gesture-based document manipulations in the 3D work space. Kinect is a motion sensing input device developed and marketed by Microsoft. It supports an RGB camera, depth sensor, and multi-array microphone, providing a cost effective spatial sensing solution for many application systems. In 2012, Microsoft released the Kinect SDK for Windows to accelerate the use of the device in many application fields. It also features embedded proprietary software for recognizing gesture, face, and voice. We mainly use it for recognizing the shape and position of a user's hand, and applying some natural gestures for controlling the system. Figure 4 shows an example depth image captured in the middle of tracking a user's hand movements.

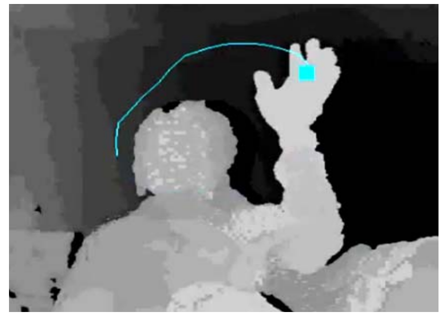


Figure 4. An example depth image for tracking a user's hand movements.

IV. PROTOTYPE SYSTEM IMPLEMENTATION

A. *Functional Overview*

Figure 5 shows the functional overview of the prototype system. As shown in Figure 2, the user operates the system by flipping the AR booklet at hand. A web camera attached to an HMD worn by the user captures a pair of AR markers printed in the current two pages of the booklet. Then, the system measures the angle and position of the booklet pages from the markers, and draws the associated document pages according to the detected angle and position in the work space. The user observes the document pages through the HMD. The system tracks the user's hand motions by the Kinect sensor set at front of the user for controlling the following operations:

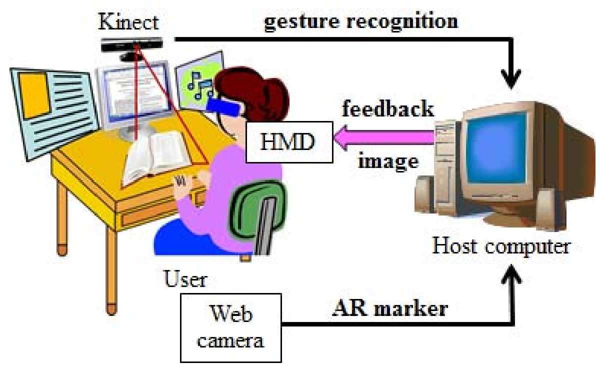


Figure 5. Functional overview of the prototype system.

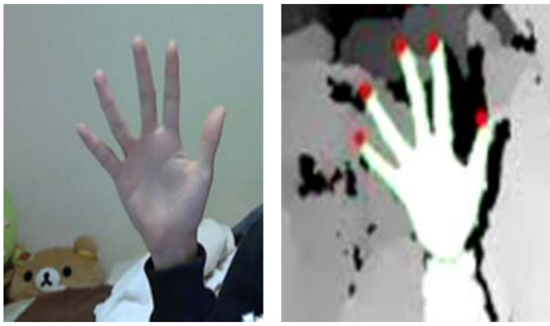
- Select a specific document page,
- Activate the browsing of a selected page, and
- Extract a page from the booklet and place it in the 3D work space.

The system maps all multimedia documents stored in the system's root folder and enables the user to browse them by turning the pages one by one.

B. *Gesture Interaction for Document Browsing*

The system allows the user for browsing multimedia documents by flicking through the booklet or carefully reading page by page. To realize such a natural reading environment, the system provides a function for browsing various documents by using simple gestures. The Kinect sensor tracks the user's hand movements and detects the shape and fingertips of the tracked hand in real time as shown in Figure 6. The prototype system supports the following interaction methods in addition to the page flip function.

- One-handed "grab" and "release" operations: When the user finds a page he/she prefers or needs to do some tasks, he/she can pick up and place it in the work space as shown



(a) RGB image (b) depth image

Figure 6. Example images for detecting the shape of an open hand.

in Figure 7. When he/she laps his/her hand over a specific page and makes a tight fist, the system activates the grab mode. Then, the user can freely move the grabbed page by moving his/her hand in the work space. When the user opens his/her hand, the page is released and stuck at the location. The stuck page can be grabbed and relocated many times. This 3D drag-and-drop operation enables the user to easily set out many documents in a space-efficient way. The grab motion also is used for selecting a specific page to change display mode as described in Section 4.C.

- Bimanual "pull" and "push" operations: A document placed in the work space by using the above mentioned grab-and-release operations can be scaled. As shown in Figure 8, the user can easily scale up and down by pinching at both ends of the target page, and pulling and pushing it for the scaling. The user can adjust the character size and image resolution included in the document page by using this scaling operation.

C. Data Structure for Browsing Documents

Figure 9 shows the internal data structure for browsing documents. As illustrated in Figure 9(a), the system uses the Window's hierarchical folder structure. All pages consisting of a document needs to be numbered and stored in a folder and all document folders are stored in the system root folder. When the user activates the system, it shows all documents in the preview mode as shown in Figure 9(b). It illustrates how the structured document data is displayed on the AR booklet in the preview mode. The system displays each document data on a booklet page and the reduced-size images of first five pages are laid out for previewing the document. The user can look at each document page by page and select a specific document (page) by grabbing it for browsing the contents of the document. Then the system changes its display mode to the contents browsing and presents a page image on each booklet page in numerical order as shown in Figure 9(c). Because the prototype can only display each page as an image file, the user needs to convert a document to a set of images beforehand. We are currently working on the rendering part for accepting other various formats like multimedia text, sound, and movies.

The system allows the user to collect various documents required for performing a specific project or a task and

Figure 7. Placing document pages in 3D work space.

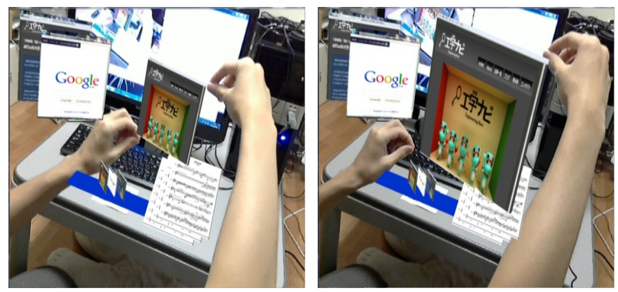


Figure 8. Bimanual scaling operation.

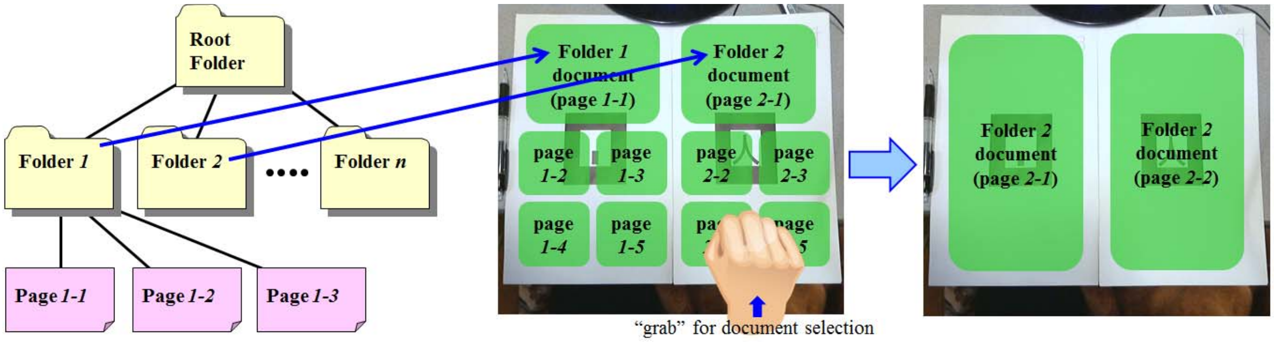
organize a booklet including all desirable information. Let us assume a case for making a new presentation file by incorporating carefully selected data among diverse information. The user can customize a booklet consisting of all gathered data prepared for making the file, browsing them page by page for finding appropriate information, picking important pages and lining them up in space for elaborating, and incorporating some parts in the new file. He/she can suspend the system by just closing the booklet.

V. Preliminary Experiment

A. Experimental Task

We conducted a preliminary experiment for evaluating the prototype system in terms of the viewability for AR-based document browsing and the operability for gesture-based document manipulations. We employed ten subjects and asked them for browsing an AR marker booklet consisting of fifty pages through a worn HMD. Each page includes multiple contents such as texts, images, and illustrations. We assigned the subjects to search blank pages with a printed number from one to four, and pick and place them in the work space as shown in Figure 10. We set a wall for placing the picked pages to clearly indicate the target locations in the positioning task.

We asked the subjects to answer the questionnaire after they completed the task. The questionnaire consists of the following four items and the subjects rate each item in ten ranks from one (very bad) to ten (very good).



(a) layered structure of documents (b) document preview mode (c) contents browsing mode

Figure 9. System's data structure for managing documents and displaying them on the AR marker booklet.

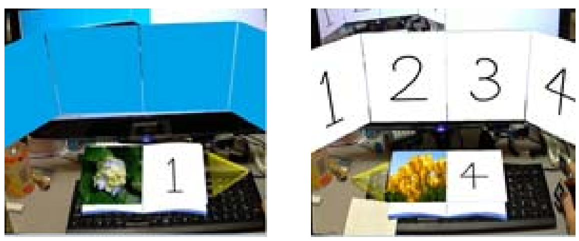


Figure10. Experimental task for searching pages with a printed number (left: start status, right: end status).

Item 1: Viewability of documents displayed by AR interface.

Item 2: Easiness for selecting a document.

Item 3: Easiness for moving a document.

Item 4: Easiness for placing a document.

B. Experimental Results

All subjects completed the task within five minutes and the average execution time is 101sec. Table 1 shows the results evaluated by the subjects for four items as described in Section 5.A.

Table 1. Evaluation results.

		Subjects										Ave.
		A	B	C	D	E	F	G	H	I	J	
Ev. Item	1	7	5	8	7	7	6	5	9	8	9	7.1
	2	7	8	6	5	3	6	7	5	5	6	5.8
	3	8	7	8	6	7	5	4	7	6	7	6.5
	4	7	6	9	4	6	7	6	9	5	8	6.5

Ev. Item: the number of evaluation item, Ave.: average rate value

Because the evaluation item 1 (document viewability) scores the highest average rate among the four questions, we confirmed the AR-based document drawing achieved a good quality for browsing operations. We used the OpenGL's texture mapping function for drawing documents in the work space. It can render the document image with enough clarity even if the booklet page size is different with the original image size. The user sometimes needed to resize the document with the scaling operation as shown in Figure 8 in only cases in which he/she is viewing the text document with many small-sized characters as shown in Figure 11.

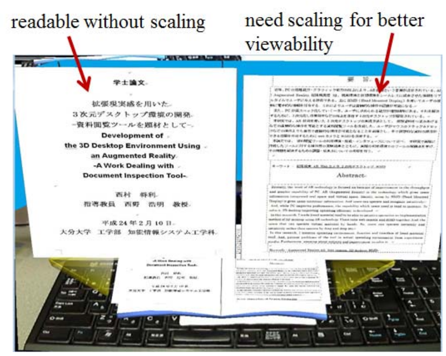


Figure11. Viewing text documents with many small-sized characters.

As for the easiness for document manipulations inquired in items 2 (select), 3 (move) and 4 (place), we confirmed the system achieved an acceptable operability level with a small number of errors. Because the current prototype system sets no limit on the range of document movements in the work space, a processed document sometimes occludes other documents. This problem frequently happened in the document scaling operation and complicated the operation environment. The stability of AR marker tracking was another issue in deteriorating the operability. Because the marker tracking requires for a full view of the tracked marker all the time, riffling the booklet sometimes interrupted smooth document rendering. We are now working on these issues to make the operation environment more stable and easier to use.

C. Discussions

Let us compare the system with a real paper book and a digital book. The real book is used in many scenes on a routine basis at schools, offices, and homes. It excels in familiarity and intuitiveness for many users. It also is better suited for poring through the whole information and imprinting key parts on the users' memory. It, however, is subject to a space constraint of real papers. In contrast, the digital book predominates in mobility and easiness to use. It also provides a powerful search engine. Its small screen size, however, tends to be a fundamental problems for preventing the user from quickly overviewing the contents of a document. In spite of some research achievements available, riffling through the document still is a difficult task.

The proposed method is considered to be a solution uniting the both technologies' advantages and features though it is weak in mobile usage.

VI. Conclusions

We proposed a new document browsing method enabling users to easily manipulate multimedia data by flipping physical booklet pages. We adopted AR with motion sensing technologies for implementing the browsing system. The implemented system allows the users to process documents with their simple gestures such as picking a specific page up, moving and lining it up in a 3D work space, and scaling it. The system unites the advantages of digital books with real book's operability. We confirmed the proposed system provides a customizable personal workspace with various kinds of multimedia data through a preliminary experiment.

We would like to consider and design new features such as going over a specific part in a page and reusing it in a new document for improving the functionality of the system. Conducting further experiments with a practical collection of multimedia data is another prime task in the near future.

Acknowledgment

A part of this work was supported by JSPS KAKENHI Grant Numbers 24500148 and 24603018.

References

[1] Agarawala, A. and Balakrishnan, R., "Keepin' It Real: Pushing the Desktop Metaphor with Physics, Piles and the Pen," *Proc. of the SIGCHI Conf. on Human Factors in Computing System (CHI'06)*, pp.1283-1292, 2006.

[2] Ouchi, Y., Nishino, H., Kagawa, T., and Utsumiya, K., "A Tangible 3D Desktop Environment with Force Feedback," *Journal of Mobile Multimedia*, Vol.8, No.2, pp.114-131, 2012.

[3] Kobayashi, R. and Ikeuchi, J., "Effects on Text Understanding and Memory by Types of Display Media: Comparison between E-book Readers and Papers," *IPSJ SIG Technical Reports*, 2012-HCI-147(29), pp.1-7, 2012. (in Japanese)

[4] Shibata, H., Takano, K., and Omura, K., "Can Electronic Reading Devices Replace Paper? Experiments to Evaluate Electronic Reading Devices," *Fuji Xerox Technical Reports*, No.21, pp.98-109, 2012. (in Japanese)

[5] Kato, H. and Billinghurst, M., "Marker Tracking and HMD Calibration for a Video-Based Augmented Reality Conferencing System," *Proc. of the 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR 99)*, pp.85-94, 1999.

[6] Watanabe, J. and Mochizuki, A., Bendable Device for Browsing Content for Use as Flexible Display Interface, *IPSJ Journal*, Vol.49, No.12, pp.3899-3907, 2008. (in Japanese)

[7] Mitsunaga, N., Yonezawa, T., and Tajika, T., "TABA(sheaf) on a Sheet: Intuitive Fliping Interface toward Electoronic Paper," *Proc. IPSJ Interaction Symposium 2008*, pp.39-40, 2008. (in Japanese)

[8] Holman, D., Vertegaal, R., Altosaar, M., Troje, N., and Johns, D., "PaperWindows: Interaction Techniques for Digital Paper," *Proc. of the SIGCHI Conf. on Human Factors in Computing System (CHI'05)*, pp.591-599, 2005.

[9] Izawa, K., Suzuki, N., Akabane, T., Yamakawa, H., Maruyama, J., Aikawa, T., Kubomoto, R., Shibayama, F., Takenaka, H., and Kobayashi, S., "A Proposal of New Interactions with Directly-manipulable `Mekuri' Interface," *Proc. IPSJ Interaction Symposium 2011*, pp.123-130, 2011. (in Japanese)