Information-Retrieval Behaviors and Content Browsing Interfaces

A Doctoral Dissertation Graduate School of Systems and Information Engineering University of Tsukuba

> March 2009 Jun-ichiro Watanabe

Abstract

We live in an age when content-rich information can be widely enjoyed thanks to the development of information technologies. However, as the amount of accessible information increases and information-retrieval appliances come equipped with more and more functions, the complexity of user interfaces increases, reducing accessibility to desired contents. For the digital world to be truly enjoyable, there must be computer-human interaction techniques that make accessing desired information not only efficient but also comfortable. This thesis describes two interfaces we developed for reaching this goal.

We first analyzed the relationships among the various factors that affect the way information-retrieval appliances operate and the various information-retrieval behaviors people use. The understanding gained of the mechanisms people use to retrieve information was then used to design two user interfaces. Much of the research on user interfaces, such as input devices and GUIs operated directly by the user, has focused on improving the interface itself. The improvements have been done with only a tacit understanding of the performance needed to achieve optimal usage. As result, there are many cases in which the actual operation differs from that best suited for the target usage. This gap between design and use is one of the factors in the creation of user interfaces that have less ease of use due to increased complexity.

The way a user operates an interface is affected by various factors, and the same user may use the same interface various ways depending on the situation. We focused on how the degree of human memorization and the features of the targeted information affect information-retrieval behaviors. For example, someone searching for a favorite photo in a folder may be able to find it immediately without straying because he/she remembers the folder's contents and the photo's whereabouts well. In contrast, someone searching for a friend's favorite photo in an unfamiliar folder will likely take longer to find it and may stray in the information space. This illustrates that memorization of the target contents affects the way a person navigates to find them.

In a user experiment, we identified a specific relationship between memorization and information-retrieval behavior. We found that the overall relative positions of the contents, the spatial order of the contents, and their grouping into discrete time periods (e.g., photos in a series) affect the degree of memorization and thus the retrieval behavior. It should be possible to develop an algorithm for adapting an interface to the way the user operates the interface, which reflects the user's intention. And it should be possible to design an interface for accessing desired information efficiently and comfortably due to not only the external design but also following the natural behaviors of people.

Using the understanding we gained from our user experiment, we developed a user interface that makes use of the fact that the overall relative position of the contents affects how people remember the contents. This is consistent with the general observation that people are usually better at memorizing and managing information spatially. Therefore, the first interface is based on the concept of grasping information presented spatially. It is based on the metaphor of a book and is designed for browsing contents presented in one dimension in the same manner as leafing through the pages of a book.

Just like pages in an actual book are presented in one dimension and are leafed through by someone browsing the book, our book metaphor device enables someone to leaf through digital contents in the same manner and with the same tangible feeling as reading a book. Real-world-oriented interfaces for accessing digital information are generally designed to mimic the way people handle familiar physical objects. In particular, tangible user interfaces enable people to access the digital world by physically operating real objects. However, though many of these interfaces bring a fresh sense to the access of digital contents compared to WIMP-style interfaces (windows, icons, menus, pointing devices), not many studies have focused on interfaces that provide the exact same feeling as that generated by operating real objects. Our book as a metaphor interface differs from other techniques using the book metaphor because it is based on the identification of factors that generate the pleasant feeling of leafing through pages and the incorporation of these factors as hardware, not software.

We also developed a time-oriented navigation interface for browsing digital contents. It is based on the use of time as the key to navigating content-rich information. The digital contents are presented temporally rather than spatially. The key feature is the use of an analog clock as a metaphor for handling time. The contents corresponding to the time indicated by the clock are shown, and the time can be freely set forward or back by simply rotating the arms of the clock. We use TV programs and related information as target contents because they are typically time-ordered and densely arranged along a time axis. They are identified by day and time of broadcast and station name or channel number. People usually remember or search for recorded programs roughly by date and time. Our browsing interface can be efficiently used to navigate a rich collection of data comprising recorded programs and information on future programs by operating the analog clock to move freely along the time axis.

Independently, these two user interfaces are an improvement over existing interfaces and will enable efficient and comfortable access to desired information. Together, however, along with a better understanding of the mechanisms people use to retrieve information, they will enable a truly natural and highly interactive computing environment for contentrich information browsing.

Acknowledgments

I am deeply grateful to Professor Jiro Tanaka, my thesis supervisor, for his many valuable suggestions, precise directions, and kind encouragement. I also thank Assistant Professor Shin Takahashi, also my thesis supervisor, for his useful suggestions and kind advice. I also greatly thank Associate Professor Kazuo Misue and Assistant Professor Buntarou Shizuki for their kind advice and suggestions regarding this research. I would like to thank Professor Yukio Fukui and Professor Hideaki Kuzuoka, the other members of my thesis committee, for their many useful comments on this research. I am also grateful to all the members of the IPLAB, Interactive Programming Laboratory, University of Tsukuba, for giving me many opportunities to discuss my research with them.

I would like to thank Youichi Horry for his encouragement and many valuable suggestions. I am grateful to all the members of the HHIL, Hitachi Human Interaction Laboratory, for their kind cooperation and encouragement.

Finally, I extend my sincere gratitude to my parents, Yoshiko and Ken-ichiro Watanabe, and my wife, Yoko Watanabe, for their constant encouragement and support.

Contents

1	INT	INTRODUCTION 1				
	1.1	Backg	round	1		
	1.2	Goal a	and Approach	2		
		1.2.1	Information-Retrieval Behavior	2		
		1.2.2	Sensibility-Oriented Interface	3		
		1.2.3	Time-Oriented Navigation	3		
	1.3	Overv	iew of Thesis	4		
2	Ana	lysis o	f Information-Retrieval Behaviors	7		
	2.1	Motiva	ation	7		
	2.2	Relate	d Work	9		
	2.3	Hypot	hesis	11		
		2.3.1	Degree of Memorization and Operation Patterns $\ldots \ldots \ldots$	11		
		2.3.2	Photo Type, Ease of Memorization, and Operation Pattern $\ . \ . \ .$	12		
	2.4	User S	tudy	13		
	2.5	Result	s	16		
		2.5.1	Classification of Scrolling Patterns	16		
		2.5.2	Degree of Memorization of Photo Itself and Operation Patterns	19		
		2.5.3	Degree of Memorization of Photo's Position and Operation Patterns	21		
		2.5.4	Types of Photos and Operation Patterns	24		
	2.6	Discus	sion	27		
		2.6.1	Effects of Type of Photo and Degree of Memorization	27		
		2.6.2	Effects of Incorrect Memorization and Position Confusion $\ . \ . \ .$	28		
		2.6.3	Effect of Arrangement Order	28		
		2.6.4	Possible Algorithms for User Interface Design	30		
	2.7	Remai	ks	30		
3	ΑE	Bendab	le Device for Browsing Contents Using Metaphor of Leafing			
	\mathbf{Thr}	ough l	Pages	31		
3.1 Sensibility-Oriented Interface						

CONTENTS

	3.2	Related Work			
	3.3 Basic Idea		Idea	42	
		3.3.1	Factors in Page Turning	43	
		3.3.2	Dynamics for Thin Plastic Sheet	45	
		3.3.3	Manner of Reading a Book	47	
	3.4	Imple	mentation	48	
		3.4.1	Prototype 1: Single-Display Structure with a Thin Plastic Sheet	49	
		3.4.2	Prototype 2: Dual-Display Structure with Two Thin Plastic Sheets .	52	
		3.4.3	Prototype 3: Dual-Display Structure Operated by Bending and		
			Rubbing	55	
	3.5	Applie	cations	59	
		3.5.1	Photo Browser	59	
		3.5.2	Leafing Through Wikipedia	61	
		3.5.3	CoverFlow Application	63	
		3.5.4	Switching between Web Browser Tabs	64	
		3.5.5	Following Links	64	
		3.5.6	Fast-Forwarding and Rewinding	67	
	3.6	Evalua	ation	67	
		3.6.1	Comparing Performance with Conventional Devices	67	
		3.6.2	Difference in Device Structure	68	
		3.6.3	Degree of Bending and Shearing Force	68	
		3.6.4	Finger-Bookmark Function	69	
	3.7	Discus	ssion	69	
4	Tim	ne-Orie	ented Information Browsing Interface	73	
	4.1	Inform	nation Navigation	73	
		4.1.1	Space-Oriented Navigation	73	
		4.1.2	Keyword Search	74	
		4.1.3	Time-Oriented Navigation	74	
		4.1.4	Goal and Approaches	75	
	4.2	Advar	nces in Technology and Changes in Viewing Styles	75	
		4.2.1	Technological Advances	75	
		4.2.2	Multi-Channel Continuous Recording	76	
	4.3	Recor	ding System	77	
	4.4	Imple	mentation	78	
		4.4.1	Multi-stream GUI	78	
		4.4.2	Views on Time Axis	81	
		4.4.3	Comparison with Conventional GUIs	82	
		4.4.4	Calendar View	85	

		4.4.5	Category View		 	 					85
	4.5	Evalua	ion		 	 					87
		4.5.1	Laboratory Findings		 	 					87
		4.5.2	Exhibition Feedback		 	 					89
	4.6	Related	Work		 	 					90
	4.7	Discuss	on		 	 					94
5	Con	clusion									97
	5.1	Retriev	al Behaviors and User Interfa	ices	 	 					97
	5.2	Summa	ry		 	 					98
	5.3	Future	Directions		 	 					101
	5.4	Conclu	ling Remarks		 	 					102

List of Figures

2.1	Examples of scrolling through photos.	8
2.2	Example plot of scroll track	12
2.3	Photo browser for experiment.	13
2.4	Inquiries.	14
2.5	Types of photos and examples	15
2.6	Typical scrolling patterns for keyboard operation	17
2.7	Typical scrolling patterns for mouse operation.	18
2.8	Time for retrieving	20
2.9	Details of operation.	20
2.10	Time for retrieving	22
2.11	Details of operation.	23
2.12	Memory score	24
2.13	Photo types and retrieval time	25
2.14	Details of operation.	26
2.15	Typical patterns of operations.	27
2.16	Effects of incorrect memorization and position confusion.	29
3.1	Factor 1: bending degree.	43
3.2	Factor 2: shearing force	44
3.3	Measured bending degree and shearing force	45
3.4	Bending thin plastic sheet	46
3.5	Prototypes for flexible display	46
3.6	Voltage measured with bend sensor for book and thin plastic sheets	47
3.7	Prototype 1: single-display structure with thin plastic sheet	49
3.8	Photograph of bend sensor.	49
3.9	Scrolling photos horizontally by bending sheet device.	50
3.10	Implementation of vibration and sound	51
3.11	Dynamics for hand acting as fulcrum.	51
3.12	Prototype 2: dual-display structure with two thin plastic sheets	52

LIST OF FIGURES

3.13	Two modes of operation	53
3.14	Leafing through Wikipedia page.	54
3.15	Prototype 3: dual-display structure operated by bending and rubbing	56
3.16	Detecting shearing force	57
3.17	Speed of page turning	58
3.18	Photo browser.	59
3.19	Finger-bookmarking function.	60
3.20	Zooming function.	61
3.21	Leafing through Wikipedia.	62
3.22	Key-event generator.	62
3.23	Operating CoverFlow interface	63
3.24	Volume control	63
3.25	Switching between tabs	65
3.26	Scrolling up and down	65
3.27	Following links.	66
3.28	Experimental results	68
3.29	Effect of features of target contents on retrieval behaviors	69
3.30	Projecting images on sheet surface	71
4.1	Multi-channel continuous recording system.	77
4.2	Main and schematic view of Multi-stream GUI.	78
4.3	Remote control	79
4.4	Views on time axis.	80
4.5	Multi-screen and full-screen modes	81
4.6	Conventional GUI on commercial HDD recorder.	83
4.7	Proposed GUI.	83
4.8	Calendar view function.	84
4.9	Category view function.	86
4.10	Exhibition at CES 2006	89

List of Tables

3.1	Operations for reading book and for browsing photos using conventional	
	information terminal device	48

Chapter 1

INTRODUCTION

1.1 Background

Accessing the digital world by using various information appliances is a common daily activity. Computers are no longer expensive and are commonly found in offices, in homes, and elsewhere. It is now difficult to find a person who does not own a mobile phone (or two). Multimedia devices such as portable game and music players have also spread widely, and their owners can enjoy the contents anywhere at anytime. Thus, as predicted by Weiser [140], a ubiquitous computing environment has been realized. We basically live in a world where everyone in a modern society can enjoy the digital world. However, "enjoyment" may not be the appropriate word to use. Dissatisfaction and complaints regarding the user interfaces of digital appliances continue in spite of the cool, brand-new designs.

One reason for the dissatisfaction may be that most conventional user interfaces are simply extensions of WIMP GUIs (windows, icons, menus, and pointing devices). WIMPstyle GUIs greatly expanded the number of users who can access digital information due to their ease of learning, ease of use, and consistency in look and feel for each application, compared to the command line interfaces found on the first-generation PCs [139]. However, as the amount of information has increased and the applications have become more complex, the harder it has become to learn how to use the interfaces fully. Moreover, the need to manipulate the controls of a WIMP GUI disrupts the process. For example, a user first has to grab the slider of a scrollbar to scroll through the contents, and opening up a menu hides some of the contents on the screen.

Various post-WIMP interface techniques have been proposed. One approach is to use a real-world-oriented interface in which a physical object is used as a metaphor for accessing digital information in a manner that is similar to a real world action. Particularly promising are tangible and graspable user interfaces [67]. They enable the operator to access digital information by directly grasping a real object and manipulating it. However, many of the interfaces studied use only the way a physical object is handled as a metaphor. They do not recreate the unique feelings generated when actually handling the object. As a result, some of them are not easy or comfortable to use though they may create a new feeling.

Another reason for the dissatisfaction is probably the gap between the expected experience of using an interface and the actual experience. While improvements in the user interface itself as an input device are eagerly promoted, scientific analysis of how the interfaces are actually used under various conditions is lacking. As a result, the gap between the interface designers' intentions and the actual usage causes some uneasiness for users.

1.2 Goal and Approach

The goal of this research was to develop user interfaces that can be used to access desired information efficiently and comfortably. To reach this goal, we took a two-step approach.

The first step was to gain a better understanding of the mechanisms people use to retrieve information. The objective was to establish an algorithm that takes these mechanisms into consideration when designing interfaces. We found that the overall relative positions of the contents affects how people remember the contents and the retrieval strategy used.

The second step was to develop interfaces based on the understanding gained in the first step. We developed two interfaces based on the concept of grasping information arranged spatially. One is a book metaphor device for browsing content presented in one dimension in the same manner as leafing through the pages of a book. The other is a time-oriented information navigation technique for browsing contents arranged along a time axis.

1.2.1 Information-Retrieval Behavior

We focused on how the degree of memorization and the features of the targeted information affect information-retrieval behavior. For example, searching for photos stored on a hard disk drive or other media is a common search activity, and finding some types of photos is often easier than finding other types. Why is this? We conducted a user study to clarify the fundamental mechanisms used by people to retrieve information. We found that the pattern of operation differed with the degree of memorization and the type of targeted content and that the differences can be interpreted as differences in retrieval strategies. These findings will contribute to the computer-human interaction research field. In particular, they will improve the understanding of the mechanisms people use to retrieve information. This should lead to the development of browsing interfaces that can dynamically and appropriately assess user intentions and retrieval situations.

1.2.2 Sensibility-Oriented Interface

We developed an interface device called "Bookisheet" that is based on the book metaphor. It is bendable and used to operate digital contents in the same manner and with the same tangible feedback as leafing through the pages in a book. It can be described as a sensibility-oriented interface. Books, the traditional information medium, not only offer the better legibility of paper compared with that of a computer screen but also several additional affordances [126, 95]. For example, a paper interface provides quick and flexible navigation [126]. Users can grasp the quantity and contents by leafing through the pages of a book or other document. The pages can be marked up or annotated. Taking notes regarding the contents of a book while reading it helps in understanding the contents. We focused on one affordance in particular—the pleasant feeling one gets when leafing through the pages of a book. A reader turning the pages can feel their elasticity and the friction they create as they pass under the finger or thumb.

The book is the classic medium that makes use of the human ability to grasp information that is presented spatially. For example, the order of the pages or the positional relationship between pages gives one the ability to accurately guess the location of the desired page. The thickness helps one to grasp the total size of the contents. It also helps one to guess the position of the desired page and to directly open to a page close to it. This ability to grasp spatial position or relationship is not inherent in the conventional interfaces used to access digital data.

Our Bookisheet device reproduces the pleasant feeling of leafing through the pages in a book and uses it as tangible feedback during the browsing of digital contents. To achieve this, we identified the factors that generate the feeling and implemented them in a hardware device for digital information operations such as scrolling. It demonstrates the feasibility of using this approach in next-generation user interfaces for information appliances, e.g., flexible displays.

1.2.3 Time-Oriented Navigation

The time-oriented information navigation technique is a unique way to grasp the whereabouts of digital data spatially. Digital data is often stored in a computer in a hierarchical structure, and the complexity of the structure increases with the amount of information. One approach to managing digital data spatially [3, 94] is to reproduce the way objects are typically arranged on a person's desk. For example, files are piled by type of information at different positions on a desktop-like GUI to enable the user to grasp their whereabouts. However, as the amount of information increases, scalability soon becomes a problem [141] and accessing desired information becomes more difficult. Another approach is keyword search, like that used to search for information using the Internet. However, information not related to the input keywords is generally not found, and some information cannot easily be presented using a combination of keywords, e.g., videos and sounds.

Time is a powerful cue in retrieving content-rich information. In the real world, memories of when an event, such as a trip, took place are often used to remember things that happened immediately before or after the event. In the digital world, we sometimes manage information by using the date and/or time as a filename for version management. Rearranging files or received e-mail by their dates of last modification or time of receipt often helps one to retrieve a desired file or e-mail more efficiently. The techniques proposed for managing digital information by using time as a key [119, 68, 97, 51, 106] were mainly designed for reproducing the past states of a file on a computer by using the file history (modification, deletion, etc.). Regardless of the name or contents of the file, by treating all the data in the computer as time-ordered data, one can more easily recall the relationship between the information at a particular time and information close to that time on the basis of association. Studies suggest that memories of information location on a time axis can be effectively used for information browsing.

We developed such an interface for retrieving information about recorded and onair TV programs, and their related information. Information regarding TV programs is typical time-ordered data densely arranged along a time axis from past to future. We use an analog clock metaphor for setting the time freely while letting the user feel the sense of time. The user can easily move from the past to the future by simply rotating the arm of an analog clock.

1.3 Overview of Thesis

In this thesis, the analysis of the behaviors people use to retrieve information is first presented. The two user interfaces are then described.

In Chapter 2, we describe the analysis of the relationship between memorization and retrieval behavior. We describe the user experiment we conducted, show the results, and discuss them. Findings such as that the degree of memorization of the contents themselves and their overall relative positions affect the way an information-retrieval device is operated and methods for developing an interface on the basis of these findings are described.

In Chapter 3, the book metaphor interface is described. Observations of handling an actual book and identification of the factors that generate the pleasant feeling of leafing through the pages of a book are described. Then, three prototype devices featuring these factors are introduced. Several applications for such a device and a simple evaluation are also presented.

In Chapter 4, an interface for navigating content-rich information using time as the key is described. We show that information arranged along a time axis can be efficiently navigated. In this study, we chose TV programs and related information as target contents to be navigated because they are typical densely arranged time-ordered data. The TV program recording system we used to collect the target data is first explained, and then the navigation interface using the analog clock metaphor is introduced.

Finally, Chapter 5 summarizes this thesis and our conclusions and discusses several future directions.

Chapter 2

Analysis of Information-Retrieval Behaviors

2.1 Motivation

What is the relationship between memorization of information and the behavior used to retrieve that information? Searching for photos stored on a hard disk drive or other media is a common activity. Chances are that it is easier to find some types of photos than others. To determine the reason for this, we conducted a user study to clarify the mechanisms people use to retrieve information. We found that the operational patterns differed with the degree of memorization and the types of target photos. In particular, we found that the overall relative positions of target contents and/or the order of the arrangement affect memorization. The difference in operational patterns can be interpreted as a difference in retrieval strategies.

Due to the rapid progress in information technologies, we live in a society where we can enjoy a rich variety of digital information. However, the complexity of the interfaces between digital information and potential users has increased, and accessibility to the information they require has been reduced. Even though a great deal of research on user interfaces has been done, and many excellent results have thus far been obtained, there have not been that many studies on the mechanisms people use to retrieve information. If we could understand them more clearly, novel and instinctive user interfaces based on these mechanisms could be attained.

In our user study, we focused on scrolling because this is one of the basic methods for retrieving digital information. It is used to move contents, such as photos arranged in one or two dimensions, along the dimension axis so that contents outside the window are brought into the window. The operation is executed by dragging the slider of the scrollbar, pressing arrow keys on the keyboard, or rotating the wheel on the mouse. The



(b) Searching for unfamiliar photo.

Figure 2.1: Examples of scrolling through photos.

speed of scrolling is controlled to some extent by the user adjusting the speed at which the slider is dragged, the duration for which the keys are pressed, or the speed at which the wheel is rotated.

We focused on how the degree of human memorization and the features of the targeted information affect the information-retrieval behavior. For example, someone searching for a favorite photo in a folder may be able to find it immediately without straying because he/she remembers the folder's contents and the photo's whereabouts well. In contrast, someone searching for a friend's favorite photo in an unfamiliar folder will likely take longer to find it and may stray in the information space (Figure 2.1). This illustrates that the scrolling pattern used may be affected by the target information.

In our experiment, we asked the participants to retrieve photos by scrolling, and we measured the time it took to find the target photo, recorded the identification numbers of the photos viewed (the "scrolling position") for every time unit, and derived the velocity from the changes in position. We also asked them to complete a questionnaire that enabled us to estimate how well they remembered the target photos, not only their features but also their position in the arrangement. We found that the pattern of scrolling, which consisted of time-ordered data reflecting position and velocity, varied with the degree of memorization. The degree of memorization was affected by the features of the photos, so the patterns also varied with the type of photo. Our results revealed that the strategies used to retrieve information depend on the degree of memorization of the target contents, especially the relative positions of the contents. They also indicated that the order of the contents generates a unique atmosphere for the entire contents. This indicates that the spatial arrangement plays an important role in memorization. This interpretation is consistent with the general observation that people have strong spatial memory abilities.

Our study contributes to the computer-human interaction (CHI) field in that the results will enable the design of user interfaces on the basis of an understanding of the mechanisms people use to retrieve information. Many of the studies on user interfaces have focused on the parts the user directly operates, such as the GUI. In contrast, we focused on the mechanisms underlying the retrieval behaviors so that we could design not only the outer layer but also the depth. That is, once we understand the fundamental parts of the mechanisms, we can develop an algorithm for a user interface that adapts to the various operation behaviors people use, which will enable the user to interact with digital contents more comfortably using natural behaviors.

The rest of this chapter is organized as follows. After briefly introducing related work, we describe the motivating hypothesis in Section 2.3. Then we describe the user study we conducted in Section 2.4, and present some of the results in Section 2.5. In Section 2.6 we discuss the results, and close with brief remarks.

2.2 Related Work

A good Web page provides information such as links and images at appropriate positions on the page to enable users to follow the "scent" of their goal, i.e., the information of interest to them. If a user loses the scent, he/she may stop browsing and begin again from another page. Several cognitive psychological studies have investigated how people use Web browsers and their information-retrieval behaviors [20, 25, 31, 32, 84, 114, 115]. These studies have shown that a user's retrieval behavior is driven in part by the way information is presented. That is, the user interface design affects the retrieval behavior. The results of these studies have been used to develop tools for designing and evaluating Web pages.

An investigation of the relationship between scrolling distance and the required precision of scrolling [56] revealed that Fitts' Law models the scrolling behaviors well, though it is usually used to evaluate the performance of pointing devices. This investigation of how conventional scrolling techniques are actually used resulted in a paradigm that can be used for designing new browsing and retrieval techniques.

Various methods have been proposed for the scrolling operation. Igarashi and Hinckley proposed a scrolling technique for browsing a large amount of content using a zoom function [63]. The pseudo speed of scrolling is kept constant by automatically zooming in and out in accordance with the speed of operation. They found that common problems such as losing the position and information when scrolling too fast can be avoided by using this technique. This technique utilizes the scrolling speed to dynamically change the content presentation. They defined a function that determines the relationship between the scrolling operation used and the content presentation. However, it does not take the relationship between the types of content and scrolling operation into consideration. The zoom level is simply a function of the scrolling speed.

Appert et al. proposed a technique for automatically adjusting the zoom level in

accordance with the user's operation, not the operation speed. Their OrthoZoom Scroller [6] controls the zoom level by moving the pointer in the direction perpendicular to the scrolling direction. As the pointer approaches the scrollbar, the contents are presented with lower and lower precision. As the pointer moves away from the scroll bar, the contents are presented in higher and higher precision. They found that users could find the target contents faster with their technique than with an automatic zooming technique.

The way items in a long list are thinned out while scrolling is an important and effective technique for browsing large amounts of information. In Masui's proposed technique for navigating text-list items [98], the items to be thinned out when zooming are determined on the basis of the degree of interest values assigned to the items. The higher the value of an item, the longer it remains on the list during zooming.

Kumar et al. proposed using eye gaze to control scrolling [83]. They focused on the finding that scrolling is strongly coupled with a user's ability to catch information using his/her eyes. For example, the placement of a document being read on a screen can be maintained even when the page up or page down key is pressed by detecting the eye-gaze point and using it to limit the scrolling edge.

For supporting user operations, some studies have suggested using the users' operation tendencies to deduce their intentions. Asano et al. used the direction and peak speed of pointer movement to deduce the target and automatically scroll to it [11]. Kobayashi et al. proposed a technique for operating a cascaded menu [73]. The direction of pointer movement is mapped to the direction of the cascaded-menu items, and the user does not have to actually point to an item to open it. If the pointer is slightly moved to the right, for example, the system deduces that the user intends to open sub-menus, and the sub-menus to the right of the main menus appear.

Sun et al. proposed a scrolling technique that uses both zooming and page flipping [132]. Their technique scrolls the contents when the slider on the scrollbar is moved slowly and flips them page by page when it is moved quickly. The transition between scroll mode and page flipping mode is seamless and depends on the preset threshold speed. They reported that, in page flipping mode, navigation making use of spatial memory, a strong human ability, is possible. They also tested a function for backtracking during page flipping in cases when the desired page was skipped, but they found that the function caused operational complexity and was relatively ineffective.

Ishak and Feiner described a scrolling method that depends on the content characteristics [66]. The speed of scrolling and zooming automatically changes in accordance with the context of the content. For example, if a document with two columns is being read, the scrolling operation is supported by a function for automatically jumping from the bottom of the left column to the top of the right column.

Improving the scrollbar has been another topic of interest regarding scrolling. In one

study, a rubber band metaphor was used to control the scroll speed [96]. The speed changed with the distance between the pointer and the scrollbar slider. When the user drags the slider, the speed of scrolling is the same as that of the slider. When the user drags somewhere else in the scrollbar area, the speed of scrolling is higher the greater the distance between the mouse pointer and the slider. This enables more precise pointing to the desired contents because the closer the target content comes to the screen, the lower the scrolling speed. This avoids the problem with conventional scrolling in which the content presentation changes rapidly with small movements of the slider.

The Alphaslider [4] works in a similar manner. The scroll speed depends on which part of the slider is dragged. For example, if the upper part of the slider is dragged, the scroll speed is high and the contents are switched roughly (200 items per unit of slider movement). If the lower part is dragged, the scroll speed is low and the contents are switched one by one.

Using a tilting device for scrolling and other operations has also been proposed. Rekimoto proposed the use of device tilting as an interaction technique for small mobile devices [118]. Menu selection, map browsing, and 3D object viewing are performed by tilting the device, and all can be done with one hand. Poupyrev et al. investigated the use of tilting combined with tactile feedback produced by device vibration [117]. They found that the tactile feedback is an effective way to let the user perceive the scroll speed or realize that the target contents are approaching.

Placement of this Study

The pioneering studies on "information scent" and "information foraging" investigated the effects of Web page design on information retrieval. Our study aims to clarify the mechanisms people use to retrieve information to enable us to develop better user interfaces. Most previous studies focused on the functions or design of the user interface itself, not on the effect the target contents have on using the interface. We focused on how the contents affect a person's use of an interface. We found that memory is a big factor affecting scrolling behaviors. And rather than define a rule or function between operations and content presentation, we aimed to reflect the user's intention or the situation in the content presentation. We clarified the mechanisms people use to retrieve information and explored the possibility of using them as a basis for novel computer-human interaction techniques.

2.3 Hypothesis

2.3.1 Degree of Memorization and Operation Patterns

In this study, we investigated the relationship between memory and retrieval behavior by conducting a user study in which users retrieved target photos using scrolling. The



Figure 2.2: Example plot of scroll track.

scrolling speed was controlled by the user and was not constant during the retrieval process. The identification numbers on the photos scrolled through were recorded, and the number of photos scrolled through per time unit was used as a measure of the scrolling speed. The speed varied with the operation pattern, for example, how long keys were pressed or the speed at which the slider was moved. We hypothesized that the shape of the scrolling pattern (Figure 2.2) depends on the user's memorization of the target contents and that the patterns can be categorized on the basis of their shapes.

The time it takes to find a target photo in a folder and the psychological load are affected by how well the searcher remembers the contents of the folder and their order. This means that the scrolling patterns for well-memorized photos should differ from those for poorly memorized ones. The memorization can be affected by both the features of the contents, that is, the contents themselves and their overall relative positions. We investigated the relationship between the degree of memorization of the contents themselves and their relative positions, and the patterns of operations used to retrieve them.

2.3.2 Photo Type, Ease of Memorization, and Operation Pattern

Though a computer cannot directly calculate the degree of a person's memorization, there is a relationship between the degree of memorization and the features of the contents that a computer can calculate. There are characteristic differences between photos that tend to be well memorized and those that are not. A series of photos with the same theme taken on nearly the same date or at the same time can usually be easily distinguished in a folder of photos. Photos with strong features such as tone or composition can also usually be easily distinguished. Thus, in this study, we categorized the target photos into three types before conducting the user experiment: "series," "impressive," and "featureless." We investigated the relationship between photo type and ease of memorization and between photo type and operation pattern.



Figure 2.3: Photo browser for experiment.

2.4 User Study

We conducted a user study to collect data on scrolling patterns, which are plots of timeordered data reflecting scrolling position (photo ID) and velocity. Using this data and answers to a questionnaire, we analyzed the information-retrieval strategies used to search for target photos by scrolling, a fundamental method of operation.

The participants were asked to find a target photo from among 200 photos using two types of operation.

- Operation 1: Use only left and right arrow keys on keyboard to scroll.
- Operation 2: Use only mouse movements to drag scrollbar slider.

Eight people (two women and six men, 25 to 42 years old) familiar with computer operation participated. Photo browsing software was run on a desktop PC (Epson, Intel Core 2 Duo, 2.2 GHz, 1 GB of RAM, 256 MB of graphics memory) with a 24-inch display (Dell). The photos were presented in one dimension horizontally across the middle line of the display (Figure 2.3). A fixed cursor was presented at the center of the display, and three photos were shown at once. The target photo was shown in the upper-center area (Figure 2.3 (a)). The upper-left area (correct picture area) was for showing the results of searching. It was colored in gray until the target photo was found. When the photo was found, it was displayed in that area (Figure 2.3 (b)). A scrollbar was presented at the bottom of the display. The participants scrolled through the photos by dragging the slider (red circle) using the mouse (in operation 2). They could also scroll by pressing the left and right arrow keys on the keyboard (in operation 1), and the slider moved in accordance with how the keys were operated. In both cases, the slider indicated the position of the photo currently located at the cursor position.

Q1: Did you remember the photo itself?						
1: Yes 2: No						
Q2:Did you remember its po	osition?					
1: Yes	(Accurately)					
2: Yeah	(Almost)					
3: No, but remembered	(Remembered during the operation)					
4: No	(Not at all)					
Q3: Did you find it where you expected it to be? 1: Yes (exactly) 2: No (different)						



There were five steps in the experimental process.

- 1. We gave the participants (one at time) 3 min to memorize the features of a total of 200 photos and their order.
- 2. The participant then pressed the Enter or Tab key to start searching. The target photo was displayed, and the timer started. The participant scrolled by pressing keys (in operation 1) or dragging the scrollbar slider (in operation 2) to find the target.
- 3. The participant pressed the Enter or Tab key again when the target photo was apparently found. If the photo was the target one, the task was accomplished and the timer stopped. If not, the left-upper area remained gray, a beep was sounded, and he/she resumed searching.
- 4. After the participant found the target photo, he/she answered three questions on a questionnaire.
- 5. Each participant repeated Steps (1) to (4) for 12 photos, once using keyboard operation and once using mouse operation, i.e., 24 tasks in total.

Each task was started and stopped when the participant pressed the Enter or Tab key. Use of the Tab key avoided the loss of time when mouse operation was used because the mouse could be operated with the right hand and the Tab key could be pressed with the left hand. Since the arrow keys and Enter key are close together, not much time would be lost when switching from pressing the arrow keys to pressing the Enter key with the right hand.



Featureless: photos not in series and with no impressive features.

Figure 2.5: Types of photos and examples.

Each time the participants finished a task, they were asked to write their answers to three questions (Figure 2.4). The first question ("Did you remember the photo itself?"; "yes" or "no") was used to investigate the effect remembering the photo itself had on the operation pattern. The second question ("Did you remember its position?"; "yes (accurately)," "yeah (mostly)," "no (but remembered during the operation)," or "no (not at all)") was used to investigate the effect remembering the photo's position had on the operation pattern. The difference between Q1 and Q2 was "position." Someone may remember the photo itself and the position as well. Others may remember the photo itself but not the position. The operation strategies may be different between these two. The third question ("Did you find it where you expected it to be?"; "yes (exactly)" or "no (different)") was used to confirm the accuracy of their memory. From the results of the last two questions, we defined a "memory score," which represented how memorable the photo was.

The potential targets were 200 photos taken with a digital camera. The same 24 of these photos were used as retrieval targets, and they were presented in the same order to all participants. The 24 photos were categorized into three types (Figure 2.5). The "series" type included photos that had been taken in close succession and had the same theme. The "impressive" type included photos that had strong, easily remembered features. They included photos that were striking in some way, such as photos with strong tones, an interesting composition, or a strange object. The "featureless" type included photos that were not in a particular series and had no strong features. For example, a photo between one series and another series could be of this type.

In the experiment, we recorded the identification numbers of the photos located at the fixed cursor position during each time unit (100 ms). We then derived the velocity from the changes in position. The entire time it took to find the target photo was also measured. When keyboard operation was used, the key-pressing events (keys up or down) were also recorded.

2.5 Results

2.5.1 Classification of Scrolling Patterns

We gathered scrolling-pattern data, i.e., the ID numbers of the photos scrolled through (i.e., the scrolling position) and the derived velocity of scrolling for the eight participants for both the keyboard and mouse operations. Some of the patterns had a similar shape even though they were for different target photos or were for different participants. This indicates that the scrolling patterns can be classified using several typical patterns and their combinations. We used four typical patterns for each type of operation to classify all the patterns. As shown in Figures 2.6 and 2.7, the patterns comprised two plots: position (ID) vs. time and velocity vs. time. The original velocity curves tended to zigzag quite a bit due to the way they were derived, so we smoothed them by using a simple moving average. The patterns in the figures are for actual data obtained from the user study.

Scrolling Patterns for Keyboard Operation (Figure 2.6)

Approach The user continued pressing (key down) an arrow key until the target photo appeared. He/she then stopped pressing and started to tap the key to scroll slowly until the target photo was reached. That is, he/she approached the target at high speed by continuously pressing a key and then slowed down. The slope of the position plot is initially steep, and then it becomes gentle near the target; the velocity plot forms a trapezoid.

Pass-by The user continued pressing an arrow key until the target photo had been passed. He/she then stopped pressing and started to tap the key for moving in the opposite direction to scroll slowly back to the target photo. That is, he/she approached the target at high speed, passed the target, stopped suddenly, and returned to the photo. The slope of the position plot is initially steep, and then it becomes gentle with opposite inclination; the velocity plot forms a trapezoid.

Tap The user scrolled by continuously tapping an arrow key. That is, he/she operated slowly and certainly by tapping a key. The position plot remains fairly steady, and the velocity plot has a very gentle slope.

Alternate The user alternated between continuously pressing an arrow key and tapping an arrow key. That is, he/she periodically repeated high- and low-speed moving. The position plot has steps, and the velocity plot has spikes at semi-regular intervals.

16



Figure 2.6: Typical scrolling patterns for keyboard operation.



Figure 2.7: Typical scrolling patterns for mouse operation.

Scrolling Patterns for Mouse Operation (Figure 2.7)

Approach The user initially dragged the slider quickly over a long distance, scrolling through numerous photos until the neighborhood of the target photo was reached. He/she then slowly searched through the neighboring photos until reaching the target. That is, he/she approached the target at high speed in one stroke and then carefully adjusted the position. The position plot has a very steep, almost vertical, slope, and the velocity plot has one sharp peak near the beginning and then remains flat. This is similar to the "approach" and "pass-by" patterns in keyboard operation.

Jump The user alternated between suddenly and rapidly moving the slider over a long distance and moving it very slowly until the target was reached. That is, he/she randomly changed the base position for retrieval by jumping long distances at once. Both the position plot and the velocity plot have discrete sharp peaks. This is a characteristic pattern of mouse operation and is not found in keyboard operation.

Move-slowly The user moved the slider slowly and continuously until reaching the target. That is, he/she operated slowly and certainly by dragging the slider. The average speed of scrolling was very low. The position plot remains fairly steady, and the velocity plot has a very gentle slope.

Alternate The user alternated between moving the slider over long and over short distances (or stopping). That is, he/she periodically repeated high- and low-speed moving. The position plot is stepped, and the velocity plot has spikes at semi-regular intervals.

We then defined representations for all the scrolling patterns. For example, if the scrolling pattern was simply "approach", (approach, pass-by, tap, alternate) was represented as (1,0,0,0). If the pattern included both "pass-by" and "tap," it was represented as (0,1,1,0). The former representation is interpreted as "approach 100%" (i.e., the rate of use was 100%); the latter is interpreted as "pass-by 50% and tap 50%" (i.e., the rates of use for both patterns was about 50%).

2.5.2 Degree of Memorization of Photo Itself and Operation Patterns

In the questionnaire (Figure 2.4), we asked the participants whether they remembered the target photo itself. Using the answers to Q1 and the scrolling pattern representations, we identified the relationship between the degree of memorization of the photo itself and the operation pattern. As shown in Figure 2.8, finding unremembered photos took longer, and the difference in retrieval times is consistent with it being more difficult to find unremembered photos than to find remembered ones.

Answer to Q1 and Keyboard Operation

As shown in Figure 2.9 (a), the participants who answered "yes" to Q1 for keyboard operation used the "approach" and "pass-by" scrolling patterns at a combined rate of



Figure 2.8: Time for retrieving.



Q1: Did you remember the photo itself?

(b) Mouse operation.

Figure 2.9: Details of operation.

Approach = Jump = Move-slowly = Alternate

about 50%. Those who answered "no" used them at a combined rate of about 25%. This is consistent with the idea that a user who remembers the target photo will tend to scroll quickly because he/she can catch a rough impression of it even when the photos are scrolled rapidly. Additionally, a user who remembers the position of the photo can move toward it without straying. A user who does not remember the photo has no clues for finding it and will thus tend to scroll through the photos more slowly, with more dependence on visual feedback.

Answer to Q1 and Mouse Operation

As shown in Figure 2.9 (b), the participants who answered "yes" to Q1 for mouse operation used the "approach" scrolling pattern at a rate of about 45%, while those who answered "no" used the "jump" pattern at a rate of about 35%. That is, someone who remembers the target photo can approach the target with one long movement. Someone who does not basically scrolls slowly, and, if he/she cannot find the target, he/she changes the base retrieval position by scrolling a long distance in one stroke. That is, they jump to a new position and start searching again slowly.

2.5.3 Degree of Memorization of Photo's Position and Operation Patterns

We have seen that differences in remembering the target photo caused significant differences in operation patterns and retrieval times. Next, we focus on the effect of remembering the target photo's position rather than the photo itself because this information could prove useful in finding the photo. The scrolling patterns for "remembering the photo but not the position" should differ from those for "remembering the photo and the position as well." We thus analyzed the representations for Q2 ("Did you remember its position?") for those participants who answered "yes" to Q1.

Figure 2.10 shows the retrieval time for each pattern. The participants who did not remember the position of the target photo (Q2: "no") even though they remembered the photo itself (Q1: "yes") took more than twice the time to find the photo than those who gave one of the other three answers.

Characteristic Tendencies for Keyboard Operation

As shown in Figure 2.11 (a), the rate of using the "approach" scrolling pattern for keyboard operation was significantly reduced when the degree of memorization was lower. That is, the more accurately the participant remembered the position, the easier it was for him/her to recognize when he/she was close to the target. He/she was able to move directly and rapidly toward the target because he/she knew where it was. As he/she approached the target, he/she slowed down to be able to stop directly at the target. Interestingly, when the answer to Q2 was "yes" i.e., the user remembered the position exactly, the "tap"



Figure 2.10: Time for retrieving.

scrolling pattern was not used. The basic strategy was to keep pressing the arrow key and moving quickly until the neighbors of the target were reached.

The characteristics for users who answered "yeah (mostly)" or "no, but remembered" to Q2 were similar. In particular, the percentages for the "pass-by" pattern were high. This was because it was more difficult for them to recognize when they were close to the target because they approximately rather than accurately remembered the position. They could find the photo visually because they knew the photo itself, so, when they found it, they immediately stopped moving and went back to the target.

When they did not remember the position (Q2: "no"), the percentage for the "alternate" pattern was more than 50%, and the two "keep-on pressing" patterns ("approach" and "pass-by") had the smallest percentages. This can be interpreted to mean that these participants used a "probabilistic" search strategy. That is, by periodically changing the "base" position of retrieval, they hoped to more quickly approach and reach the target.

Characteristic Tendencies for Mouse Operation

As shown in Figure 2.11 (b), when the participants remembered the position of the target exactly (Q2: "yes") or almost exactly ("yeah"), the rate of using the "approach" pattern for mouse operation was close to 50%. The "jump" pattern was virtually unused, especially for "yes." This means that, when the participants knew the position of the target, they moved toward it without hesitation. If they did not initially remember the position but remembered it during the operation ("no, but remembered"), the rate of using the "approach" pattern was lower, and that of using the "jump" pattern was higher. The "jump" pattern was also used by those who answered "yeah". As evident in the figure, the lower the degree of remembering the position, the higher the rate of using the "jump"
2.5. RESULTS



Q2: Did you remember its position?

Figure 2.11: Details of operation.

and "move-slowly" patterns. However, when the participants did not remember the position at all (Q2: "no"), the "jump" pattern was not used, and the rate of "move-slowly" was close to 50%. This can be interpreted to mean that, when a participant roughly remembered the position, he/she used the strategy of frequently changing the base position of retrieval, aiming to accidentally and probabilistically find a location near the target. And when they did not remember the position at all, they used the strategy of slowly searching from one end of the photo list one-by-one.

As we have seen, the operation patterns varied with the degree of how well the target's position was remembered. In keyboard operation, when the participants remembered the position accurately, the typical pattern used was "approach"; when they remembered the position approximately, it was "pass-by"; and when their memory was poor, it was "alternate." In mouse operation, when they remembered the position exactly, the typical

⁽b) Mouse operation.



Figure 2.12: Memory score.

pattern used was "approach"; when they remembered the position approximately, the "jump" pattern had a larger rate; and when their memory was poor, "move-slowly" was dominant.

In short, there is a relationship between the degree of memorization and the operation patterns used. Next, we discuss the relationship between the degree of memorization and the features of the photos. That is, what kinds of photos are easy to memorize and what kinds of photos are difficult? To evaluate this ease of memorization, we defined a measure.

2.5.4 Types of Photos and Operation Patterns

Memory Score

We defined "memory score" (MS) for evaluating the ease with which the photos were memorized. It was calculated for the photos for which the answer to Q1 was "yes." MSmakes use of the answers for Q2 and Q3 and is defined as;

$$MS_{i,j} = \begin{cases} 3 \times q_{1,i,j} + 2 \times q_{2,i,j} + 1 \times q_{3,i,j} + 0 \times q_{4,i,j} & \text{(if answer to Q3 was "yes")} \\ 0 & \text{(if answer to Q3 was "no")} \end{cases}$$

where *i* stands for participants and *j* stands for photos. The $q_{1,i,j}, q_{2,i,j}, q_{3,i,j}$, and $q_{4,i,j}$ are the answers to Q2 ("yes," "yeah," "no, but remembered" and "no") and took a value of one if they corresponded and 0 if they did not. The more accurately the participants remembered the position of the target photo, the higher the score that was awarded. If the answer to Q3, which was to confirm how accurately they remembered the position after finding the target photo, was "no," a score of 0 was given because their memorization was not accurate. The higher the MS, the easier it was for participant *i* to memorize photo *j*.

The average MS for each photo type we used in the experiment (Figure 2.5) is shown in Figure 2.12. The "series" photos tended to have higher scores, and the "featureless"



Figure 2.13: Photo types and retrieval time.

ones tended to have lower scores. The results show a correlation between photo type and MS, i.e., ease of memorization.

We have seen that the operation patterns varied with the degree with which the photo itself and its position were remembered. Since the degree of remembrance is affected by the photo type, there must be a correlation between photo type and patterns used. This is supported by the finding that it took longer to find featureless photos (Figure 2.13).

Photo Type and Keyboard Operation

As shown in Figure 2.14 (a), when searching for "series" photos, the participants used the "approach" scrolling pattern at a higher rate than for the other two types of photos. This can be explained in the sense that, when searching for a "series" photo, it was easier for a participant to recognize when he/she was in the neighborhood of the target photo even when they approached it quickly because the similar photos in the series were easily recognized. It was surprising to find that the "alternate" pattern had the highest rate $(\sim 40\%)$ for "series" photos. This could be because, when the participant did not remember the position of the series, he/she searched for the series of photos that included the target at a medium scrolling speed (not quickly by continually pressing the key, and not slowly by using the "tap" pattern). This would be more likely to happen when the series containing the target photo contained only a few photos, and the target photo itself was "featureless." When searching for "impressive" photos, the participants used the "passby" pattern at a rate of about 60%. This was because the "impressive" photos could be memorized more accurately, enabling the participants to move directly and quickly toward the target photo. However, since the target was not in a series, it was difficult to recognize when they were near the target so that they could slow down. They passed the target and then went back. The "tap" pattern was used more often to search for "featureless" photos. The participants tended to search for the target by scrolling slowly from one end



Photo Types and Patterns

Figure 2.14: Details of operation.

of the photo list photo-by-photo because they had few clues for recognizing the photo.

Photo Type and Mouse Operation

As shown in Figure 2.14 (b), the mouse operations used to search for "series" and "impressive" photos were mostly the same. In both cases, the "approach" pattern, i.e., moving the slider rapidly over a long distance toward the target, had a rate of about 50%. The difference is that the rates for "move-slowly" and "alternate" were reversed between "series" and "impressive." For "series," the rates were about 13 and 24%. For "impressive," they were about 26 and 15%. This indicates that, when the participants did not know the whereabouts of the target photo accurately and searched for a series of photos, they tended to use "alternate," so the average scrolling velocity was medium. When they searched for a single impressive photo without knowing its whereabouts, they tended to use "move-slowly," so the average velocity was low. For "featureless" photos, the rates for "jump" and "alternate" were higher (~21 and ~32%). The search strategy was to increase the



Figure 2.15: Typical patterns of operations.

probability of finding a neighbor of the target by changing the retrieval base periodically or by suddenly beginning to move the slider.

2.6 Discussion

2.6.1 Effects of Type of Photo and Degree of Memorization

We have seen that different scrolling patterns were used depending on the degree of memorization of the target photos and of their positions in an arrangement. There was also a relationship between the type of photo and the degree of memorization. This relationship was used to define a rule combining the patterns of operations and the type of target content. The differences in the scrolling patterns can be interpreted as difference in information-retrieval strategies.

In our user study, we found some typical patterns of operations (Figure 2.15). The "approach" and "pass-by" patterns observed for keyboard operation were consistent with the "approach" pattern observed for mouse operation. They were generally used to search for well-memorized photos that were either in a "series" or "impressive" and were not used much to search for poorly memorized "featureless" photos. For "series" photos, the searcher tended to slow down and stop at the photo. For "impressive" photos, the searcher tended to go past the target and return to it. The "tap" and "move-slowly" patterns observed in keyboard and mouse operations, respectively, were used to search for unfamiliar photos (poorly memorized and/or "featureless").

The interesting and unexpected patterns we found were "alternate," observed for both

keyboard and mouse operations, and "jump" for mouse operation. The "alternate" pattern was used for retrieving poorly memorized photos, especially in keyboard operation. The "jump" pattern was correspondingly used in mouse operation. It tended to be used to search for poorly memorized, "featureless" photos. With the "alternate" and "jump" patterns, the participants used a similar strategy for finding unfamiliar photos. They changed the base position of retrieval, aiming to accidentally and probabilistically come close to the target. The difference between the two is that with "alternate," the participants changed their base position periodically and generally continued scrolling in the same direction. With "jump," they changed the base position suddenly and did not necessarily continue scrolling in the same direction, so that the position (photo ID) plot zigzagged (see Figure 2.15).

2.6.2 Effects of Incorrect Memorization and Position Confusion

There were several cases in our user study in which the participant remembered an incorrect position or was confused about the position and made a comment about it on the questionnaire.

- 1. "I quickly searched forward and backward when the target was not where I thought it was."
- 2. "I was confused by the positions of similar photos."

In the first case, the participant first moved quickly to the incorrectly remembered position and then moved gradually while changing directions. As shown in Figure 2.16 (a), the position (photo ID) plot zigzagged a bit and was uneven, and the velocity plot was uneven and had regular spikes.

In the second case, when the participant reached the incorrect position, he/she stopped scrolling for a moment. Apparently, he/she was momentarily confused by the similarity between the target photo and the one at which he/she stopped (here, a small girl with a pink hat). The similarity could be used to form a "series" of photos, but here it simply was a trap because the photos were ordered without taking their similarities into consideration. As shown in Figure 2.16 (b), the position (photo ID) plot was like the combination of "pass-by" and the velocity plot consisted of multiple (in this case two) trapezoids.

2.6.3 Effect of Arrangement Order

We have seen that different information-retrieval strategies were used for retrieving different types of photos and for different degrees of memorization. The correlation between the degree of memorization and the scrolling patterns used seems to be especially significant. We received several useful comments regarding photo memorization. For example,



Figure 2.16: Effects of incorrect memorization and position confusion.

- 1. "I memorized the photos on the basis of the clothes the people wore or the season."
- 2. "I knew that the target photo was not in the latter half of the photo list, even though I did not remember the photo itself. The target photo showed some houses, and although I did not remember any pictures with houses, I did recall that there were no photos with houses in the latter half of the photo list, only ones with beautiful beach scenes."

The first 100 of the 200 photos used in the experiment were taken in May, and the other 100 were taken in October. People with short-sleeved shirts or trees with green foliage were included in the photos taken in May, reflecting the season becoming warm. Likewise, people with long-sleeved shirts or trees with foliage that had turned red were included in the photos taken in October. Comment 1 indicates that the participant had memorized photos by using a mental model of clothes and seasons. That is, a sense of the season of the photos was mapped onto the flow of time for all 200 photos and was used to estimate the position of the target photo. Comment 2 indicates that the position of the photo could be deduced even when the participant did not remember the photo itself. The participant searched for the photo by comparing it with the order or atmosphere of all the photos to estimate its position. These comments indicate that we make use of a mental model reflecting the atmosphere of all the photos, focusing on their features and order, rather than memorizing each photo exactly. If the order of photos is changed, the atmosphere generated by the whole collection of photos is also changed. This can change the scrolling patterns or time it takes to find the target photo. In future work, we intend to analyze the correlation between a change in photo order and the change in scrolling patterns.

The comments we received reflect the human ability to grasp, manage, and memorize objects and information spatially. For example, we can locate desired object, say a document, on a messy desk because we grasp its whereabouts in our brain. Comment 1 indicates that we memorize the information on a time axis using time as a key. Comment 2 indicates that we can more easily recall contents that are memorized spatially. We hypothesize that not only by classifying the scrolling patterns but also by using how contents are spatially and temporally arranged, we can achieve a better understanding of how people retrieve information.

2.6.4 Possible Algorithms for User Interface Design

Our ultimate goal is to understand the mechanisms people use to retrieve information and to establish a method for designing user interfaces based on that understanding. In the study described in this chapter, we investigated fundamentals of these mechanisms. We showed that it is possible to extract a principle of information retrieval that is based on memorization and to develop an algorithm for adapting user interfaces to human's behaviors. For example, an algorithm could be constructed that causes only photos in a series with high memorization potential to be presented with emphasis when a user scrolls through the photos rapidly and that causes featureless ones with low memorization potential to be presented with emphasis when the user scrolls slowly. Though this example is rough and the details must be worked out, if such algorithms were established, we could design interaction methods based on the natural behaviors of people rather than simply defining functions that treat the relationships between the various factors and the search behaviors in the same way.

2.7 Remarks

We have demonstrated that the scrolling patterns used for retrieving photos differ significantly depending on how well the searcher memorizes them and their positions in the arrangement. We also demonstrated that a photo's characteristics affect the degree to which it is remembered. The difference in scrolling patterns can be interpreted as a difference in strategies for retrieving information. It could be possible for applications to determine a user's intentions or situation by dynamically identifying the operation patterns used and to automatically display appropriate information at the right time. Such an interface, based on the mechanisms used by people to retrieve information, should provide more natural interaction between people and computers.

Chapter 3

A Bendable Device for Browsing Contents Using Metaphor of Leafing Through Pages

3.1 Sensibility-Oriented Interface

In the digital world, there is a rich variety and large amount of information. Thanks to the various types of information terminals such as desktop PCs and mobile devices, we can easily retrieve this information. As many studies have shown [126, 95], digital technologies have many advantages; for example, information can be updated regularly, related pieces of information can be linked to each other, and information can be browsed remotely.

However, to enjoy the contents in the digital world, we have to use special interfaces designed for information appliances. Most user interfaces are based on the WIMP-style GUIs (windows, icons, menus, pointing devices) widely used on desktop computers. Therefore, users who are familiar with computers can usually access this information without much trouble. Unfortunately, these conventional interfaces are not necessarily convenient for all users, and there are many cases in which the user must endure some unpleasantness while navigating digital contents. For example, in the home, a person can store many recordings of TV programs, digital camera photos, videos taken with a home video camera, and downloaded audio tracks on large-capacity hard disk drives in HDD recorders or PCs. Storing, retrieving, and enjoying these contents requires the operation of a remote control, a graphical user interface, a keyboard, a mouse, etc. However, the accessibility to the desired contents and the pleasantness of operation are affected by the quality of the user interface. For mobile devices, navigating contents is more difficult even for users who are familiar with computers. Because of the restrictions on screen and device size, user interfaces for desktop appliances are not appropriate for mobile devices. Even though mobile devices can be operated using buttons, a touch panel, and/or a stylus, these interfaces are still too troublesome for many users, especially ones who are unfamiliar with using such devices. Users may have to press a series of buttons to reach the contents they want, and they can easily lose their way while doing so.

Various post-WIMP interface techniques have been proposed. One approach is to use a real-world-oriented interface in which a physical object is used as a metaphor for accessing digital information in a manner that is similar to a real world action. Particularly promising are tangible user interfaces [67]. They enable the operator to access digital information by directly grasping a real object and manipulating it. However, many of the interfaces studied use only the way a physical object is handled as a metaphor. They do not recreate the unique feelings generated when actually handling the object. As a result, some of them are not easy or comfortable to use though they may create a new feeling.

We think it is important to take the unique feelings into consideration when designing user interfaces for accessing the digital world using natural behaviors. We can make not simply "tangible" user interfaces but "sensibility-oriented" ones once we deeply understand the essential qualities of these unique feelings and apply them to designing interfaces. In this study, we did not simply mimic the physical objects in the development of content browsing interfaces but also endeavored to give the user the ability to actually feel the operation.

Books, the traditional information medium, not only offer the better legibility of paper compared with a computer screen but also several advantages, as Sellen et al. pointed out [126, 95, 111, 1]. For example, quick and flexible navigation is one of the affordances of a paper interface. Users can grasp the quantity and contents of documents by leafing through the pages of a book or a stack of documents. Marking up documents or annotating the pages of a book is another affordance of a paper interface. Taking notes regarding the contents of a book while reading it helps in understanding the contents. We focused on one affordance in particular, the pleasant feeling created by leafing through the pages of a book. The physical tangibility of the pages in a book and the simplicity of the page turning generate a pleasant feeling for most people. As a reader turns the pages, he/she can feel their elasticity and the friction they create as they pass under his/her finger or thumb.

People tend to be good at recalling object locations memorized spatially, and we make use of this ability when we read a book. For example, the order of the pages or the positional relationship between pages gives a sense of guessing where the desired page is. The thickness helps the reader to grasp the total size of the contents. It also helps the reader to guess the position of the desired page and to directly open to a page close to it. It is difficult to grasp the spatial position of digital data or the spatial relationship between them using the conventional user interfaces on computers and information appliances. For example, someone following the links on Web pages may find it difficult to rediscover (reopen) the pages found along the way or to organize several interesting pages that were found.

We thus designed an interface based on the metaphor of turning pages in a book. It enables users to interact with a computer without using a mouse or keyboard, which were designed for computer operations. We applied this intuitive user interface to digital content browsing. Using our Bookisheet interface, which consists of two thin plastic sheets, bend sensors, and pressure sensors, a user can easily scroll digital contents such as photos or Web pages by bending one side of the sheet or the other, and rubbing along the edges. This action provides a tangible and pleasant sense, like turning pages in a book. Bookisheet can be used not only as an interface for conventional information terminals but also as one for flexible displays.

Flexible displays (for example, [116, 41]) are particularly attractive for applying the book metaphor interface. The technologies supporting flexible displays have been well studied and have rapidly advanced [122, 82, 60, 108] though they have not yet become widespread. Nevertheless, they will likely gain wide acceptance, so investigating novel interaction techniques is worthwhile. Our Bookisheet interface can be applied immediately when flexible displays become practical.

The rest of this chapter is organized as follows. Section 3.2 describes related research in several areas. The basic idea of our research is described in Section 3.3, and developed prototypes are explained in Section 3.4. After introducing several applications for our device in Section 3.5, we present the results of our evaluation in Section 3.6. In Section 3.7, we discuss several aspects of this research.

3.2 Related Work

In this section, we introduce research related to our work from several aspects. First, research that compares the digital and physical worlds is summarized. Studies on paper interfaces motivated us to explore the possibility of using a book as a metaphor for digital content browsing. Next, studies on user interfaces using the interaction manner of handling paper or books are described. Then, studies on linking a physical book with digital information are introduced. Since flexible displays are within the scope of our research, the technological progress in flexible displays is summarized and several studies on user interfaces using flexible displays are described. Studies on content viewer applications using the book metaphor, i.e., e-book reader devices, are also summarized. Finally, the placement of our research in relationship to these related studies is described.

Digital vs. Real

Real-world-oriented user interfaces aimed at accessing the digital world using interactions

similar to handling real objects have been well studied. Ishii et al. described a tangible user interface [67], and many user interfaces using physical objects as a controller in the digital world have been proposed. These studies made use of the affordances [48] of physical objects for intuitively accessing digital information. "Affordance" means an aspect of an object which makes it obvious how the object is to be used, and Norman pointed out its importance in designing objects used in our daily lives [110].

The paper interface has often been used as a typical physical object and compared with digital technologies. Sellen and Harper observed and analyzed how paper documents are handled in offices to determine why paper is still being so much despite the advent of paperless technologies [126]. They identified several affordances of paper and digital technologies.

Affordances of Paper

- Quick, flexible navigation through and around documents.
- Reading across more than one document at once.
- Marking up document while reading.
- Interweaving reading and writing.

Affordances of Digital Technologies

- Storing and accessing large amount of information.
- Displaying multimedia documents.
- Fast full-text searching.
- Quick links to related materials.
- Dynamically modifying or updating contents.

According to the results of their analysis, the main advantage of the paper interface is its free navigation and ease of being written upon. By leafing through the pages of a book or a stack of documents, a reader can easily grasp the total size of the contents, get an overview of them, and determine the position of interesting information. By annotating the pages, sometimes referencing multiple documents at the same time, the reader can better understand the contents. The disadvantages of the paper interface include the lack of remote access, the difficulty of updating the contents, and the impossibility of presenting movies.

In contrast, the advantage of digital technologies is the ability to handle large amount of information. By using the Internet, a user can access a large, virtually infinite, amount of information. Multimedia information including text, photos, movies, and music can be found. The disadvantages of digital technologies include the limitation of the methods and area of operation, and interruption of on-going work due to manipulating the widgets of the user interface. A user has to, for example, use a mouse to drag the slider on a scrollbar in order to browse the non-displayed contents or pull down menus and click on labels to execute other functions.

Marshall et al. investigated how people read magazines and found that "lightweight navigation" and "multiple-page turning" are the big differences between reading them on paper and on digital information appliances such as e-book readers [95]. Lightweight navigation represents a user's unconscious behavior while reading a physical magazine such as folding pages to hide a distracting page or part of the article, digressing from the flow of the text, and looking away from the advertisements. It is difficult to perform lightweight navigation on digital devices—a user has to zoom to part of the document consciously and then read the contents in order, linearly. With a physical magazine, the reader can turn multiple pages rapidly or turn several at once, helping him/her grasp the total number of pages and the relative positions of pages of interest. He/she can then easily find desired articles without using the table of contents. With a digital device, the reader can browse the contents rapidly as well, but cannot easily grasp the overview of them. Marshall et al. argued that it is important to apply not only the advantages of digital technologies such as hypertext but also the pleasant feelings created by lightweight navigation when designing user interfaces for digital appliances.

Terrenghi et al. investigated the differences between the behaviors on a digitally enhanced tabletop interface and those on a normal physical desk [135]. Their motivation was to determine whether imitating the real world is best for developing interfaces for ubiquitous computing. In their user study, they compared the operations supported by a multi-touch input tabletop display and by a physical desk. The tasks were retrieving and rearranging pieces of a puzzle or photographs, which were printed on paper-like cards or as digital images of the same size and presented on the tabletop display. They measured the time for completion. They also recorded movies of how the users operated and analyzed the movies. They found that the users tended to use only their dominant hand to operate the digital multi-touch display. Even when a user used both hands, he/she moved them symmetrically. In contrast, they tended to use both hands asymmetrically when handling the printed photos on the physical desk. Their findings demonstrate that a digital device that imitates a real-world object is not always used in the same manner as the actual object.

Using Affordances of Paper Interfaces

Fishkin, Harrison et al. investigated interfaces for mobile devices such as PDAs and e-book readers using the metaphor of turning pages or cards [45, 54]. They simulated the turning of pages by having the user press sensors at the upper corners of the frame of a device. They also explored the use of tilting to scroll menus, using the metaphor of turning cards

in a binder. They also proposed an interaction technique to annotate digital documents in a manner similar to writing notes in the margins of a book. They generated the margin by shifting the displayed digital document toward the grabbed side of the device—when the user grabbed the left side of the device, the document was scrolled leftward and a margin was generated on the right. They made several interesting discoveries.

- The pressure placed on the upper edge must be small for turning pages comfortably.
- The sound effect of page turning was turned off by most users because they interpreted it as noise.
- The tilting operation for scrolling must be ended by another gesture such as squeezing, not by returning the device to the normal (default) tilting position to prevent passing-by the target.

They used a "pressure strip" to measure the pressure along the width of the display in order to recreate the effect of the thickness of a book. When the user touched the left edge of the strip, the first page was presented; when he/she touched the right edge, the final page was presented; and when he/she touched the middle, a page near the middle of the book was presented. In this way it was possible to not only turning pages one by one but also to turn pages in chunks.

Schilit et al. investigated the enhancement of the reading experience with digital devices by using the annotating affordance of paper, e.g., underlining, highlighting, and writing comments on the document [123]. With their XLibris device, operations such as retrieving information, filtering, and sorting can be done by linking annotated information made with a pen tablet. They investigated ways of using both the paper interface's affordance of freely making annotation and the digital technologies' affordance of linking information freely.

Chen et al. examined interaction techniques for electronic books with dual displays [30]. The operations of two small displays joined together are mapped to those of an actual book, and the pages are flipped by fanning one of the two displays toward the other. They attached an accelerometer to each display, which are hinged together, to distinguish three modes: side-by-side (hold with pages opened), back-to-back (one page is folded behind the other), and detached (two displays are detached). For example, the document size is changed by switching from side-by-side to back-to-back mode, and pages are turned by flipping the device in back-to-back mode or fanning one display in side-by-side mode. This research is an example of applying lightweight navigation [95] to e-book readers. The operations for one page were simulated well, but those for multiple pages were not.

Siio and Tsujita proposed an interaction technique for handheld devices such as PDAs that is based on the metaphor of a paperweight [129]. When we write something on paper

in the real world, the paper slips if there is no weight on it. In their technique, operations on content by stylus differ depending on the position of one's palm. For example, a photo is rotated if one's palm is on the device, but it is scrolled if the palm is in the air.

Studies on using the affordances of a book as a bridging interface between the real and digital worlds have been done. Because a book is a traditional information medium familiar to most everyone, there have been studies on the use of paper as an interface for people unfamiliar with computers [38, 29]. For example, Davidoff et al. proposed a book-like device for supporting people who are unfamiliar with computers to enable them to send e-mails [38]. Each step in sending an e-mail is mapped to a page, and a user sends an e-mail by simply turning the pages and following the instructions on each page, just like reading a book page-by-page. Another method proposed for supporting people who are unfamiliar with computers enables them to execute computer tasks by scanning pages with bar-codes corresponding to each step in a task [29].

Linking Paper and Digital Information

Another approach to bridging the digital and real words goes beyond simply using the paper or book metaphor to actually linking the two worlds.

There have been studies on the use of a pen device, another familiar object, together with physical paper as an interface. As digital pen technologies have advanced, interaction techniques for their application have become more sophisticated [107, 127, 88]. Using these techniques, a user can, for example, manipulate a digital file such as a PowerPoint file by annotating physical printouts of the file with a pen. The user can also carry around actual paper documents linked to digital data.

Nelson et al. described a system that creates a PowerPoint presentation by scanning printed out slides [107]. Steps such as selecting appropriate slides, changing their order, and adding notation can be done in the same manner as handling a stack of cards. Signer et al. described a system that runs a PowerPoint presentation using printed out slides and a pen [127]. Operations, such as displaying or printing out slides, can be done simply using the pen to check the field next to the corresponding instruction, "display" or "print out", printed on a card. Liao et al. investigated a technique for browsing annotations written on paper on a digital document viewer [88]. A user can view the annotations in the margins on various pieces of papers as one digital document.

Several studies have look into enhancing the reading experience in the real world by using digital technologies. Various methods have been developed that make it possible to browse not only the contents of a book but also related digital information while maintaining the tangible feeling of handling a book. There have been studies on projecting related digital information such as text and images around the actual book or on the page margins [145, 52, 77, 74, 18].

In the system proposed by Wu et al., pages are identified by using a camera set above

the book, which is placed on a table [145]. The digital information is updated and modified to match the contents of the page being read. This ability to update information is one of the advantages of digital technologies and one of the disadvantages of a book. Grasset et al. developed an augmented reality (AR) technique for reading markers printed on the pages of a book for identifying the page number and accessing linked digital information appropriate for each page [52]. They used a separate handheld display instead of projecting digital information onto the margins of the book. A similar technique was used by Kobayashi et al. [77, 74] in their EnhancedDesk system. The page number is represented by a 2D marker and is identified with a camera. The user's finger tip is also identified by the camera on the basis of the extracted skin color. Billinghurst et al. described a technique for transiting between reality and virtual reality using a physical book and AR [18]. When book pages are viewed through a handheld display, they are overlaid with 3D virtual images.

IconSticker [128] links icons on a computer to stickers in the real world. A sticker with a particular identifying mark can be printed out by simply dragging the corresponding icon to a location on the computer screen labeled "entrance to the real world." By attaching the printed sticker to a physical object, a user can manage the digital data exactly like managing the object. For example, digital documents can be managed by placing corresponding stickers on physically printed documents and filing or stacking the printed documents. Just like he/she would handle and manage printed documents stacked on a desk or filed in a cabinet, the user can handle and manage the digital data. He/she can also memorize the data using spatial memory, which is a human attribute. To access particular digital data, the user simply scans the corresponding sticker.

Other techniques for bridging paper and digital information using a pen have been proposed. For example, PaperLink [9] presents digital information related to the information on the page of a document by scanning the words or marks written or made with a pen by using a small camera attached to the pen. Other examples include accessing digital data by manipulating physical documents written using conductive ink [89] and recording information both digitally and physically at the same time by using a graphic tablet [91].

Several research groups have explored ways of enhancing the reading experience by incorporating digital technologies. For example, Back et al. described a system for a museum setting that uses sound to enhance the reading experience [13, 14]. Sounds related to each page are played as the pages of an actual book are turned. The page number is identified by reading RFID tags placed on each page by using a tag reader embedded in the book spine. Since an actual book is used, the sound is added while maintaining the tactile feeling of paper. Conductive ink can also be used to identify the page numbers [101].

Inagawa et al. described a technique for presenting digital information related to the

page opened [64]. A markup language called eBookML is used to describe the relationship between the page number and linked digital information. The page number is identified by using a bookmark-type swan switch or by using a book cover that can detect the gesture of page pressing.

Interfaces for Flexible Displays

Before flexible displays can become widely used, improvements must be made in the performance of the materials used to make them (organic electro-luminescence, liquid crystal film, etc.) and of the thin film transistors that drive the display materials. Basic research has been conducted on the development of highly efficient luminous materials, the optimization of display structures, and the improvement of panel flexibility. There has also been evaluation of the reliability and application potential of flexible displays [122, 82, 60, 108]. For example, a flexible display with a speaker function has been presented [108]. This function is critical for making flexible displays portable, one of their key advantages. One of the problems with flexible displays has been their slow rewriting, but a controller with partial rewriting ability was recently presented that improves rewriting performance [125].

Several groups have investigated flexible display interfaces. For example, Schwesig et al. developed a bendable device and explored various ways it could be operated for menu selection, map zooming, and text input [124]. They implemented a function for blending a street map and an aerial photograph of the same area by changing the degree of overlap between them by bending the device. However, image scrolling is done by operating a touch sensor set behind the device, not by bending. Moreover, users have to learn new modes of operation; for example, to display the system menu, they must bend the device upward twice.

Matsumoto et al. presented a conceptual model for a small bendable display using the metaphor of a Post-It note for use as an interface for managing movie contents [99]. In their model, the contents stored in a computer are copied by attaching a small flexible display called a "post-bit" to the display, and fast-forward and rewind operations are done by bending the display.

Holman et al. suggested interaction techniques similar in manner to handling a piece of paper [57]. Eight actions for handling a piece of paper (hold, collocate, collate, flip, rub, staple, point with one hand, and point with two hands) are used for operating digital contents. For example, flipping is used for scrolling or page switching, and rubbing is used for copying information from computer to paper or from paper to paper. Digital information is projected onto the paper surface, the position of which is detected using multiple VICON cameras [104] and infra-red reflecting markings. However, the metaphor used is that of handling one piece of paper, not that of handling a book with multiple pages.

Lee et al. described a method for projecting images on foldable surfaces [85]. As with Paper Windows, the surface is tracked by recognizing the position of infrared light-emitting diodes (LEDs) embedded in the surface. A PixArt camera in a Wii remote, which contains a hardware blob for tracking four IR LEDs, is used for the detection. They succeeded in projecting images onto various types of surfaces such as a fan surface.

Digital Document Viewer Software

A number of studies have been conducted on interactive digital document viewing using the book metaphor, and many products have been commercialized. The information is presented as a 3D computer graphics book, and the operations used are the same as those for an actual book. The point of interest has been how the virtual book imitates the real one.

As an example of public use, the "Turning the Pages" document browsing system used in British libraries [22] presents virtual books on a touch panel display, and the pages are turned by rubbing the surface of the display. Commercial software includes PDF document viewers and photo browsers [105, 2]. They effectively present contents in the limited size of a screen [24] and can be used to annotate the pages in a displayed book [58]. A graphical image of a book is displayed on the screen, and photos can be attached to the pages. A user turns the pages by dragging the edge of the page, by clicking the corner of the display, and/or by pressing the right- or left-arrow key. However, the load imposed on the user exceeds that imposed by an actual book. Although the advantages of digital technologies such as zooming or jumping to a desired page immediately can be useful, there is no software that generates a feeling that is similar to that of leafing through the pages of an actual book. For example, the speed of page turning animation is too slow. The page turning is done by mouse or keyboard, so the feeling of operation is completely different compared with that of handling an actual book. While techniques for making the operation feel more natural have been studied, they are still insufficient [36, 37].

One topic of interest has been using the thickness of a book to directly open to a desired page, and related studies have tried to make use of the human ability to spatially memorize the locations of objects. Matsushita et al. developed a digital information viewer called "BookWindow" [112, 80, 8] in which digital contents are presented as pages of a virtual book. Users can browse the contents by turning the pages, can annotate the pages, and can bookmark pages of interest. The thickness of the book is graphically indicated, so the user can easily grasp the total size of the contents and the relative position of the currently opened page. Clicking a point on the thickness indicator opens the corresponding page, so the user can turn to a page close to the desired page or turn pages in chunks. Pages can be turned one by one by clicking on the displayed page, not on the thickness indicator. Functions are provided that combine the advantages of digital technologies and of actual books. They include jumping to pages that contain a keyword found using a full text search, copying favorite pages and arranging them around the virtual book, underlining text using a mouse, bookmarking pages, jumping to pages by clicking their bookmarks, and playing movies on the pages. In a user study comparing the retrieval performance of BookWindow with that of the "less" command, they asked the participants to turn to a chapter in a document for which they had read the chapter titles once before the experiment. The retrieval time using BookWindow was shorter than that using the "less" command, which they attributed to the use by BookWindow of the human ability to managing and memorize objects spatially. Moreover, the users reported that the action of page turning gave them a pleasant and peaceful feeling.

A similar digital information viewer was described by Yagawa et al. [147], who also aimed at using the thickness of a book to efficiently access desired information. In their book-like Digital Album viewer, the pages are tuned by clicking a "thumb button" field representing the thickness of the book. The speed of page turning depends on where the thumb button is clicked—the higher up the clicked position, the faster the page turning. They also asserted that people are generally good at managing and memorizing information spatially.

The book metaphor can be used for browsing information on the Internet as well as locally stored content. Card et al. noted that Web pages viewed while following the links to a destination page are often "lost." They proposed representing Web pages like pages in a book [26]. They implemented a Web browser that arranges such pages in a virtual 3D space along with the links followed to enable the user to access them more quickly. Ichimura et al. proposed enabling a user to leaf through a set of nodes retrieved from a hypermedia document in a manner similar to leafing through the pages in a book [62]. A similar technique was proposed by Miyazawa et al. [103].

E-Book Readers

While some e-book readers, especially designed devices for reading electronic publications, have reached the market [131, 100, 47, 5], they are not as widely spread as electronic dictionaries. This may be because they are not as portable because they need a display sufficiently large for comfortable legibility, and, as a result, they are not as light or small as an actual paperback book. In contrast, electronic dictionaries are smaller and lighter than actual dictionaries. Moreover, the user interfaces on e-book readers are based on button and menu operations, which are not used when reading an actual book. In other words, e-book readers are more like typical digital devices than books, with their familiar interaction operations. Once their portability and ease of operation are improved by, for example, the use of flexible displays and novel interaction techniques, e-book readers are more likely to become widely accepted by consumers.

Placement of this Study

In our research, we explored the possibility of making use of the pleasant feeling generated by leafing through pages in an actual book to browse digital contents pleasantly. As we have described, there has been much research into using a book as a metaphor for browsing interfaces. Some researchers focused on imitating the appearance of a book and representing it as an information viewer on the screen. However, the operations are much different than those used for an actual book. Others focused on operating existing information appliances using novel input methods such as pressure, tilt, or pen input, and they tried to recreate the manner of handling a book from the hardware point of view. However, because only a part of the appliance is replaced with sensors, the overall feeling is unlike that generated by handling a book.

Our research is the first attempt to use a book metaphor for browsing contents that provides tangible feedback, i.e., using hardware that creates the same pleasant feeling of leafing through the pages of a book. We did not simply mimic the physical objects but also endeavored to give the user the ability to actually feel the operation. Our inspiration came from a study on a pointing device called "Soap" [16], which provides a tactile sensation of spinning a wet bar of soap in one's hand. Moreover, we propose applying this approach to flexible displays, which are expected to become wide spread in the near future. In terms of flexible hardware, the approaches used for Gummi [124] and Paper Windows [57] are similar to ours. However, the former does not use the book metaphor, and the user must learn operations specific to the device. The latter does not handle data in the same manner as physically handling multiple pieces of paper such as a book or a stack of documents. In our approach, we use a book, which everyone knows how to operate, as a metaphor and recreate the pleasant feeling generated by leafing through the pages of a book.

3.3 Basic Idea

Our objective was to develop an interface that enables a user to access the digital world comfortably. To reach this goal, we explored the possibility of using a real object as a metaphor, with particular attention to a book. Using a book as a metaphor reduces the psychological load of operation as well as creates the pleasant feeling of leafing through the pages in a book. We first investigated the factors that generate the pleasant feeling, and then used them to create a unique interface that produces the essential manner of handling a book.



Figure 3.1: Factor 1: bending degree.

3.3.1 Factors in Page Turning

Exactly how are pages turned? It is a mystery, and we still have no exact answer. However, our observations of people turning the pages of a book revealed several interesting things.

- When a page is turned, the side of the page opposite the direction in which the page is turned is bent over. For example, if a page is turned from right to left, the right side of the page is bent over.
- If the degree of bending is large, the pressure of the thumb pressing the page determines whether the page is actually turned. If the pressure is strong, it creates much friction between the thumb and page, and the page is not turned.
- If the degree of bending is small, the page is not turned even if the thumb does not press the page.
- The speed of page turning depends on the degree of bending. If the degree of bending is small, the pages are turned slowly, one by one. If the degree of bending is large, the pages are turned quickly.
- The thumb shifts as pages are turned.



Figure 3.2: Factor 2: shearing force.

From these observations we identified two factors that play key roles in the dynamics of page turning. One is the bending of the pages, and the other is the shifting of the thumb along the edge of the pages.

Factor 1: Bending of the pages When a book is held, the fingers holding it from behind play the role of a fulcrum, and the pages are distorted. When the pages are turned, they are bent more. As a result, they have elastic energy E_1 or E_2 (Figure 3.1, $E_1 < E_2$) and tend to go upward. The thumb has to press with pressure F_{thumb1} or F_{thumb2} (Figure 3.1, $F_{thumb1} < F_{thumb2}$) on the page to balance the elastic force F_1 or F_2 (Figure 3.1, $F_1 < F_2$) and the weight of the pages. A page is turned when the elastic force is greater than the thumb pressure. The larger the degree of bending, the larger the elastic energy and the faster the speed of page turning.

Factor 2: Shearing force by shifting thumb The position of the thumb shifts as pages are turned (Figure 3.2). By shifting the thumb, the reader reduces the pressure on the page under the thumb, and the page turns. To leaf through pages smoothly, the reader has to add pressure toward the direction of thumb moving, what we call "shearing force." That is, the reader has to rub the edges of the pages, otherwise multiple pages are turned at once in a block unit (10 pages at once, for example), and no "leafing through" happens, and no pleasant feeling of page turning is generated. By rubbing the edges of the pages and controlling the amount of shearing force, the reader can freely control the speed of page turning. Shifting the thumb quickly causes the pages to turn quickly. It also creates a feeling of friction because pressure is added in the direction of the shearing force.



Figure 3.3: Measured bending degree and shearing force.

The elasticity caused by bending the pages and the friction caused by the shearing force make leafing through possible, so these two factors together generate the unique pleasant feeling. To investigate these two factors, we measured the degree of bending and the shearing force by attaching a bend sensor to a page and a pressure sensor to a reader's thumb. Figure 3.3 shows the values measured under four conditions; pages held open, pages leafed through slowly and quickly, and page turning stopped. When the pages were turned slowly, they were bent slightly. When they were turned quickly, they were bent more. The shearing force showed a similar tendency. The faster the pages were turned, the larger the pressure that was added in the shearing force direction. This means the page edges were rubbed more strongly when the pages were turned more quickly.

From these results, we identified bending degree and shearing force as the two main factors in page turning and generating a pleasant feeling. We then considered ways of applying them.

3.3.2 Dynamics for Thin Plastic Sheet

The same dynamics are found if we bend a thin plastic sheet (Figure 3.4), like one used for a flexible display. An appropriate force $(F'_{thumb1} \text{ or } F'_{thumb2}, F'_{thumb1} < F'_{thumb2})$ must be applied to the surface by the thumb to balance the elastic force $(F'_1 \text{ or } F'_2, F'_1 < F'_2)$, and elastic energy $(E'_1 \text{ or } E'_2, E'_1 < E'_2)$ is generated depending on the degree of bending. Therefore, bending a sheet should generate a feeling similar to that of leafing through the pages in book.



Figure 3.4: Bending thin plastic sheet.

We made a preliminary study of the characteristics of bending a book and of bending a thin plastic sheet. We attached a bend sensor to (1) a book, (2) two thin plastic sheets connected together, and (3) a thin plastic sheet. The plastic sheets, (2) and (3), correspond to two prototypes we developed for dual- and single-display flexible displays (the structures are illustrated in Figure 3.5 and will be described in the next section). Changing the degree of bending changed the voltage measured using the bend sensor (the larger the degree of bending, the smaller the measured voltage).



Figure 3.5: Prototypes for flexible display.



Figure 3.6: Voltage measured with bend sensor for book and thin plastic sheets.

As shown in Figure 3.6, for the book, the measured voltage decreased when pages were turned. The higher the turning speed, the lower the measured voltage. In other words, the degree of bending is larger when pages are turned more quickly. Similar characteristics were observed when the plastic sheets were bent. Though there was a difference between the single- and dual-sheet structures in terms of the degree of bending (the dual-sheet one was easier to bend), the tendencies in the measured values were the same. This indicates that bending a thin plastic sheet generates a tangible sense similar to that of turning pages. Therefore, to implement the book metaphor interface, we used sheet bending to implement factor 1 and created a mechanism for implementing factor 2.

3.3.3 Manner of Reading a Book

While there are differences in the way people read a book, there are several basic operations. We compared these operations with those for browsing photos using a conventional information terminal device (Table 3.1).

We focused on the two operations in particular that create a pleasant feeling when leafing through the pages of a book— operations 4 and 5 in Table 3.1. The physical tangibility of the pages in a book and the simplicity of the page turning generate the pleasant feeling. Operation 2 in Table 3.1, "adjust point for starting retrieval," is exemplified by a reader who would like to start searching around 3/4 of the way into the book after being told to turn to page 75 of a book with 100 pages. This operation is unique to

Operation	Reading a book	Browsing photos
1	Open/close cover.	Turn power on/off.
2	Adjust point for starting retrieval.	Specify a photo.
3	Turn pages one by one.	View photos one by one by pressing a
		button.
4	Leaf through pages slowly.	Continue pressing button.
		Use mouse wheel.
5	Leaf through pages quickly.	Use scroll bar.
6	Place book face down.	Pause slide show.
7	Insert a bookmark.	Bookmark a slide.
8	Turn pages while keeping a finger	
	between pages at a point of interest.	_
9		Zoom in/out.

Table 3.1: Operations for reading book and for browsing photos using conventional information terminal device.

reading a book though a "specify page number directly" function is supported by several conventional information devices. Operation 8 in Table 3.1, "turn pages while keeping a finger between pages at a point of interest," is also unique to reading a book. We applied operations 2, 4, 5, and 8 to the browsing of digital contents. Moreover, we aimed to not simply duplicate the interaction manner of handling an actual book but also to add some advantages of digital technologies, such as zooming (operation 9, Table 3.1).

The main challenge in developing a hardware interface based on the book metaphor was to recreate the pleasant feeling generated by leafing through the pages of a book. Another challenge was to implement the unique and essential aspects of handling a book, including making use of its thickness. We did not simply try to faithfully implement the manner of handling a book. We also explored the possibility of creating a unique manner of handling digital data using an interface based on the book metaphor.

3.4 Implementation

Using the basic ideas discussed above, we developed three prototype bendable sheet devices: a single-display structure and two dual-display structures. Prototypes 1 and 2 were designed for testing the first factor of page turning dynamics, that is, page/sheet bending. The browsing operation is controlled by changing the degree of sheet bending. Prototype 3 was designed not only for testing sheet bending but also for testing the second factor, that is, shearing force. We also implemented additional functions that enable operations unique to reading a book, e.g., finger-bookmarking and starting point adjustment.



(a) Appearance.

(b) Components.

Figure 3.7: Prototype 1: single-display structure with thin plastic sheet.

3.4.1 Prototype 1: Single-Display Structure with a Thin Plastic Sheet

Structure

The single-display structure has bending characteristics similar to those of reading a book though the degree of bending required tended to be less than that required for a book or the dual-display structures. Given the likelihood of future products using flexible displays, single-display structures can reasonably be expected. We thus developed a singledisplay version of our Bookisheet interface. The appearance of this first prototype and its components are shown in Figure 3.7. It consists of a thin plastic sheet (1 mm thick), two bend sensors, two speakers, and two micro switches. The analog voltage values from the sensors and switches are sent to a micro controller, and the digitized value is output to a PC.

Degree of Bending Detection

To detect the degree of bending, we used bend sensors (FLX-01, Jameco Electronics [71], Figure 3.8). The resistance of each was 10k ohm when the sensor was flat and increased as it was bent, reaching 40k ohm at 90° bending. We used two bend sensors, one on the right side of the plastic sheet and the other on the left side. The resistance of one



Figure 3.8: Photograph of bend sensor.



Figure 3.9: Scrolling photos horizontally by bending sheet device.

of the sensors increased when one side of the sheet was bent while that of the other side remained unchanged, so the side of the sheet that was bent could be determined, and that information could be used to control the scrolling or to switch the contents.

Basic Operation

We developed a photo browser application to be operated by the prototype. We arranged 200 photos horizontally, and they were scrolled when the sheet device was bent. For example, if the right side of the device was bent, the photos were scrolled from right to left (Figure 3.9). We know heuristically that the speed of turning pages increases as the degree of bending is increased. This was confirmed by the results of our simple measurement (Figure 3.6). Since we have not yet quantified this relationship, we implemented two options as an initial step: (1) increase speed of scrolling linearly or (2) exponentially with degree of bending.

Vibration and Sound

A person reading a book can feel the vibrations created by the pages passing under his/her fingers and hear the sound created by the pages as they are turned. Although the vibration and sound are not usually experienced consciously, they are important feedback for controlling the speed of page turning.

We thus attached small speakers (Figure 3.10) to the right and left sides of the sheet. A brief sound is emitted when a photo is scrolled so that the user senses vibration of the sheet surface and sound when bending the device. That is, he/she can control the speed of scrolling naturally by tangibly recognizing the degree of bending and speed of scrolling.

Other Functions

We also implemented a function for turning pages one by one (operation 3 in Table 3.1) and one for turning pages while keeping a finger between pages of interest (operation 8 in Table 3.1). These functions are operated by using micro switches and are also implemented on our third prototype. Their evaluation is described below, in the description for our

3.4. IMPLEMENTATION



(c) Output brief sound when photo is scrolled.

Figure 3.10: Implementation of vibration and sound.

third prototype.

Evaluation

An initial user evaluation of our first prototype in a laboratory setting revealed some interesting aspects of this device.

Control of Scrolling Speed Positive comments were made regarding the free and easy control of the scrolling speed by changing the degree of bending. All of the participants (six men and two women) preferred increasing the scrolling speed exponentially against the degree of bending. They also recognized the value of sound and vibration as feedback, but commented that the sound should be carefully adjusted to avoid being taken as noise. **Feeling in Hand not Doing the Bending** Most participants reported feeling elasticity



Figure 3.11: Dynamics for hand acting as fulcrum.



Figure 3.12: Prototype 2: dual-display structure with two thin plastic sheets.

in the hand doing the bending similar to that when turning pages in a book. However, the elasticity did not feel similar in the other hand, the one simply holding the sheet. This was because that hand had to press the surface downward and work as a fulcrum to prevent the entire sheet from lifting up (Figure 3.11).

Using these findings, we developed a second prototype with a dual-display structure that should generate a feeling of tangibility more similar to that of a book.

3.4.2 Prototype 2: Dual-Display Structure with Two Thin Plastic Sheets

Structure

The second prototype has a dual-display structure, like that of an actual book, and bending characteristics similar to those of reading a book, as shown in Figure 3.6. It was developed to not only reproduce the tangible feeling of an actual book but also to make use of the thickness of a book for efficient content browsing. We focused on operation 2 in Table 3.1, "open directly to a page that is close to the desired page". Because this operation must be done when the book is about to be opened, we created an explicit open-close structure.

As shown in Figure 3.12, this prototype has two thin plastic sheets that are connected, with a bend sensor attached to each sheet. The analog voltage value from the sensors, which changes with the degree of bending, is sent to a micro controller, and a digitized value is output to a PC. An application running on the PC, such as a dictionary page viewer, works in conjunction with the input sensor values. A light-dependent resistor (LDR) is attached to one sheet, about 1/4 of the page width from the center (hinge). It detects the brightness, which differs between the open and closed states, enabling detection of open and close.

Modes of Operation and Testing

The prototype has two modes of operation. The first mode is for indicating where to start



(a) Mode 1: Indicate starting point

Figure 3.13: Two modes of operation.

the retrieval operation. It corresponds to using the marks along the edges of the pages in a dictionary to find the pages with the words starting with the corresponding letter (Figure 3.13 (a), right). It is enabled when the two plastic sheets are close together (Figure 3.13 (a), left). The potential starting points are pre-set, for example, at the beginning of each section in a dictionary. If the user bends the sheet while keeping the "root" of the two sheets close together, the index pointer jumps incrementally from A to Z depending on the degree of bending.

The second mode is for turning pages normally. It enables the user to turn the pages one by one by bending the left or right sheet. It is enabled when the two sheets are far from each other (Figure 3.13 (b)). If the user bends the right sheet, the pages are turned from right to left. They move one by one continuously, and the speed is adjusted by changing the degree of bending. The larger the degree of bending, the faster the page turning.

The two modes are distinguished by whether the two sheets are close together or not. Whether they are close together or not is detected by the LDR, which detects the brightness. When the sheets are close together, less light is detected. When they are not,



Figure 3.14: Leafing through Wikipedia page.

more light is detected. Mode 1 is when only the tip of the sheet (around one-third of the sheet width) is bent and the area where the LDR is attached remains dark. Mode 2 is when the sheets are opened and the area around the LDR is bright (above a threshold).

To test our second prototype, we arranged 10,000 captured images of Wikipedia pages (Japanese version) in alphabetical order. Index characters (ten Japanese kana symbols) were positioned down the left side (see Figure 3.14). The currently selected index was indicated in red. A slide bar at the bottom of the screen showed the relative position of the currently displayed page. In mode 1, the user selected the target index by bending sheets while keeping them close together. After selecting the index, the user then smoothly transferred to mode 2 by separating the sheets widely to turn pages continuously.

Evaluation

Testing was done using five participants (two female, three male). In the real world, a book with 10,000 pages is very rare. We instructed the participants to leaf through the pages of this large book and find indicated words (five different words for each participant).

All of them found the indicated words within a reasonable amount of time (average time was 68.7 sec). All of them used modes 1 and 2 effectively. For example, for indicated word "Eye," they consistently first jumped to "F" using mode 1 and then scrolled back to "E" using mode 2. This is because the indicated word came at the end of index group "E." The metaphor of opening to a page by referring to the thickness of the book seemed to work well for mode 1. The participants said that they liked the feeling of leafing through the pages by bending the sheets, which is mode 2.

3.4.3 Prototype 3: Dual-Display Structure Operated by Bending and Rubbing

In our first and second prototypes, we focused on using one of the two main factors in page turning—bending pages. Evaluation of both prototypes demonstrated that bending a plastic sheet generates tangible feedback caused by the elasticity of the sheet, and this feedback enables the user to operate digital contents intuitively. For example, the scrolling speed can be easily and freely controlled on the basis of the feeling of elasticity corresponding to the degree of bending.

In our third prototype, we focused on the second factor of page turning, rubbing the edges of the pages, together with bending.

Structure

Figure 3.15 shows the components and appearance of prototype 3. The third prototype is basically the same as the second—two thin plastic sheets jointed together with a bend sensor attached under each one. In addition, it is equipped with a solid edge piece along the outside portion of each sheet, and a pressure sensor to detect the shearing force caused by thumb rubbing is embedded in each piece. A micro switch for implementing optional functions is also embedded in each edge piece. These sensors and switches communicate with the PC using Bluetooth, and the power needed to drive the micro controller and Bluetooth antenna is supplied by a small 6V battery. The sensors, switches, and battery are hidden in a spine, so prototype 3 is more like a book than a sheet. This device can wirelessly operate an application running on a PC.

Degree of Bending Detection

As in the first two prototypes, the bend sensors attached to the sheets generate an analog value depending on the degree of bending. As in the second prototype, we used two bend sensors, one on each sheet. The resistance of one of the sensors increased when one of the sheets was bent while that of the other remained unchanged, so the sheet that was bent could be determined, and that information could be used to control the scrolling or to switch the contents.

Detecting Shearing Force

To detect the direction of the shearing force created by the thumb rubbing the edge of the sheet, we installed an L-shaped plate in each edge piece so that its surface was flush with the edge surface, as illustrated in Figure 3.16. Two pressure sensors connected in parallel were attached to each plate for detecting the pressure and direction of the shearing force. The shearing force applied to any part of the plates, upper, middle, or lower portions, can thereby be detected. As shown in Figure 3.16, the L-shaped plate moves slightly towards the outside, that is, the direction of the shearing force. If there is no pressure in the direction of the shearing force (no thumb rubbing), the plate moves back to its original



<Front side>

<Backside>



<Appearance>

Figure 3.15: Prototype 3: dual-display structure operated by bending and rubbing.



Figure 3.16: Detecting shearing force.



Figure 3.17: Speed of page turning.

position because the rubber part of the pressure sensor has some thickness. Any vertical pressure on the L-shaped plates is not detected.

Speed of Page Turning

As we found in the evaluation of prototype 2, the users preferred that the speed of page turning vary exponentially with the degree of bending. That is, the larger the degree of bending, the faster the page turning. The difference between prototypes 2 and 3 is the use of factor 2, shearing force.

Figure 3.17 shows how the speed of page turning was determined. The speed of page turning, v, is expressed as $v = c_1 p \times (\exp(c_2 b) - 1)$, where p and b represent the shearing force and bending degree, and c_1 and c_2 are constants. Basically, the stronger the bending, the faster the turning exponentially. And the stronger the shearing force, the faster the turning linearly because the numbers of pages turned in a second should be proportional to the distance the thumb moves along the edge in one second. In Figure 3.17, if p = 0, v = 0, no matter how large the bending degree. As shearing force increases, a high speed of page turning can be obtained with the same bending degree. If shearing force is kept constant, the larger the bending degree the faster the page turning. Combining bending and shearing force enables pages to be turned exactly like in a book.

In our implementation, we tuned the speed carefully by giving c_1 and c_2 appropriate values so that the page turning produced a natural feeling. We set the maximum speed on the basis of observations of leafing through pages in a book and legibility on the display. Using a high-speed camera, we filmed how pages are turned in a book. We found that the number of pages turned in one second varied with how bending and shearing force were used together and that there is a maximum speed for keeping the pleasant feeling of page turning. On a screen, if the contents are scrolled too quickly, the user cannot judge the direction of scrolling. Given these considerations, we set the maximum speed depending on the application. The refresh rate of the display device should also be considered when


(a) Bending and rubbing.

(b) Scrolling photos.



setting the speed for an actual implementation.

3.5 Applications

We created several typical applications for use with our Bookisheet device. Basically, any type of digital content managed in units of pages can be operated using our device.

3.5.1 Photo Browser

We created a photo browser application that faithfully reproduces the tangible pleasant feeling of leafing through the pages of a book. Scrolling of the photos is controlled by adjusting the degree of sheet bending and the shearing force created by rubbing the edges.

Figure 3.18 (a) shows a picture of a user bending the right page of the device and rubbing the L-shaped plate embedded in the right edge outward with his thumb at the same time. These actions correspond to leafing through the pages in a book from right to left. The photos, which are horizontally presented in one dimension on the display of a PC (Figure 3.18 (b)), are scrolled leftward. Performing the same actions on the left sheet causes the photos to scroll rightward. The speed of page turning, scrolling in this case, is determined by the combination of the degree of bending and the amount of shearing force caused by the thumb rubbing.

Finger-Bookmarking

As a reader turns the pages of a book, he/she may find an interesting page unexpectedly and insert a forefinger or thumb into the book at that page to mark it while continuing to turn the pages (Figure 3.19 (a)). After browsing the remaining pages, he/she can easily return to the page of interest. Although conventional information devices are often equipped with a bookmark function, the user must often stop the current task, such as



(a) Insert finger at page of interest (actual book).



(b) Finger-bookmark function (our device).

Figure 3.19: Finger-bookmarking function.

viewing a slideshow, and call up a bookmark list to register the page of interest. In contrast, using a "finger-bookmark" does not interrupt the operation. This is a unique characteristic of reading a book.

We implemented a finger-bookmarking function by placing a micro switch in each edge piece near the forefinger locations. Pressing the switch on the side opposite the bending bookmarks the opened page. For example, the user might find an interesting photo while scrolling through the photos leftward by bending the right side of the sheet device. He/she presses the left micro switch to bookmark the photo (photo 24, Figure 3.19 (b), left). While continuing to press the switch, he/she continues scrolling through the remaining photos by bending the right side of the sheet (Figure 3.19 (b), middle). After browsing the photos, he/she stops bending the sheet and quickly returns to the bookmarked photo by simply releasing the left micro switch (Figure 3.19 (b), right).

Using this finger-bookmarking function, a user can bookmark contents of interest at anytime without interrupting the scrolling. Compared with the bookmark functions on conventional information terminal devices, which require a registering step or calling bookmark list step, our finger-bookmarking function follows the natural behaviors and should thus be easier to use.

Turning Pages One-by-One

When reading an actual book, the reader not only leafs through multiple pages but also

3.5. APPLICATIONS



(c) View of photo browser.

Figure 3.20: Zooming function.

turn pages one-by-one to better determine the contents of each page (operation 3, Table 3.1). In our Bookisheet device, we implemented this function by using micro switches. If the user presses the right switch, for example, the photos are scrolled one photo right to left.

Interactions that are not Similar in Manner to Handling an Actual Book

One approach to developing an interface for handling digital contents by using the book metaphor is to do so faithfully. The photo scrolling operation we described above requires both bending and rubbing a sheet at the same time, and these actions faithfully replicate the dynamics of turning pages in a book. As a result, a user scrolling photos experiences a tangible sense of leafing through the pages in a book. Finger-bookmarking and turning pages one-by-one faithfully replicate actions associated with reading a book.

Another approach is to take advantage of the technologies involved with handling digital data and thereby enhance the manner of handling a book. For example, zooming is not possible in a physical book but is possible for digital photos displayed on a computer screen (operation 9, Table 3.1). We implemented a zooming function in our photo browser application by using operations on the L-shaped plates and micro switches. To zoom in, the user rubs both L-shaped plates outwards at the same time, as shown in Figure 3.20 (a). To zoom out, the user presses both micro switches at the same time (Figure 3.20 (b)).

3.5.2 Leafing Through Wikipedia

Browsing a rich collection of digital contents in the same manner as leafing through actual pages is generally a novel experience. The largest dictionary in the real world contains



Figure 3.21: Leafing through Wikipedia.

at most ~ 3000 pages and is at most 15 cm thick. Leafing through the pages of such a large dictionary is not a pleasant experience—the dictionary must be placed on a desk, and the reader turns pages one-by-one or grabs a section of pages (100 pages or so) and leafs through them.

The Wikipedia online dictionary contains much more information than a physical dictionary and is used in a much different way. It is safe to say that many users of Wikipedia do so as a result of using a search engine such as Google. Though Wikipedia has an A-Z index, and users can search for entries by using it, the amount of information in Wikipedia is simply too large for the index to be a practical means of entry. In contrast, information is ordered alphabetically in a physical dictionary, so readers can guess the location of a desired entry by considering the thickness or index of the dictionary. As discussed in Section 3.2, research has shown that people are good at managing information spatially, and the index for a dictionary is suitable for this ability.

We created an application for browsing 10,000 captured Wikipedia pages arranged alphabetically (Figure 3.21) by using our Bookisheet device. In the real world, turning 10,000 pages would be a virtually impossible task, but doing so is no problem with this application. The operation is the same as that of the photo browser application, i.e., bending a sheet and rubbing an edge. Turning pages alphabetically arranged gives the



Figure 3.22: Key-event generator.

3.5. APPLICATIONS



Figure 3.23: Operating CoverFlow interface.

feeling of handling an actual book and helps the user grasp the position of the desired page relative to the current page.

3.5.3 CoverFlow Application

Most conventional applications can be operated using shortcut keys. We therefore created a template application that receives signals from the Bookisheet device sensors and generates key events for operating such applications (Figure 3.22). With this "key-event generator" application, many conventional applications can be operated using our Bookisheet device. We applied it to two conventional applications.



(b) Volume DOWN.

Figure 3.24: Volume control.

CoverFlow is an application from Apple Inc. [7] that presents CD jacket images on the graphical user interface of their iTunes music player. The images are arranged in one dimension horizontally, and a user can scroll through the images with nice animation by pressing keys on the keyboard, dragging the mouse pointer, or rotating the mouse wheel. With the key-event generator application, we could use our device to operate CoverFlow (Figure 3.23). The CD jacket images could be scrolled in the same manner as leafing through the pages of a book. Moreover, we could use it to control the volume in iTunes. The volume is controlled using up and down keys on the GUI. We defined rubbing of the L-shaped plate on the right side with pages opened, i.e., no bending, as "volume up" and rubbing of the L-shaped plate on the left side as "volume down" (Figure 3.24).

3.5.4 Switching between Web Browser Tabs

The latest Web browsers, Mozilla Firefox 2.0, Microsoft Internet Explorer 7.0, etc., feature "tabbed" browsing. Switching between tabs is done by mouse clicking or pressing a combination of keys on the keyboard. With the key-event generator application, we could use our device to switch between tabs (Figure 3.25) in the same manner as leafing through the pages in a book and with the same feeling.

Many Web pages are too long for their contents to be displayed on one screen, so a reader has to scroll up and down to read them. Scrolling is done by dragging the scrollbar slider or pressing the "page up" and "page down" keys. We defined rubbing the L-shaped plate on the right side with pages opened, i.e., no bending, as "scroll down" and rubbing of the plate on the left side as "scroll up" (Figure 3.26).

3.5.5 Following Links

Our Bookisheet device met the challenge presented by the first four applications we developed, i.e., browsing a rich but finite collection of contents, such as photos and music stored in a computer and bookmarked or captured Web pages. The next challenge was to browse an infinite amount of information.

The application we developed to test its ability to browse an infinite amount of information loads all the pages linked to the parent page in background (Figure 3.27). In other words, it creates tabs dynamically. The links are obtained using one of the standard control functions found in Microsoft .NET Framework 2.0, WebBrowser Active X Control. The user can then leaf through the linked pages. If a linked page remains open for a certain time, e.g., 30 seconds, it becomes the new parent page, and its linked pages are loaded and can be leafed through. It is thus possible to follow an infinite number of links by turning pages.

3.5. APPLICATIONS



Figure 3.25: Switching between tabs.



(b) Scroll DOWN.

Figure 3.26: Scrolling up and down.



Figure 3.27: Following links.

3.5.6 Fast-Forwarding and Rewinding

The final application was the fast-forwarding and rewinding of animations or movies. Drawing cartoon figures in the corners of a stack of pages and animating them by leafing through the pages is an interesting real-world example. We developed a simple way of rapidly displaying a series of digitally drawn cartoon figures using our device. We also developed a way of fast-forwarding or rewinding movies using our device. Controlling movies is a natural application of our technique for leafing through digital contents because movies are the ultimate example of serially arranged data.

3.6 Evaluation

3.6.1 Comparing Performance with Conventional Devices

Even though Bookisheet generates a nice feeling during operation, it wouldn't be much to write home about if the performance was worse than that of more conventional devices. We thus tested its performance against those of four conventional devices for a retrieval task.

Task and Method

We evaluated the usability of our first and third prototypes by comparing their operation with that of four conventional devices: a book, the right- and left-arrow keys on a PC keyboard, a mouse wheel, and a touch panel. Eight people (two female, six male, 25-42 years old who often use computers and various types of information terminals) were given a task (search for a specific photo) designed to evaluate the ease of image retrieval. There were 200 photos, each participant executed each task five times, and the time to execute each task was measured.

Prior to the experiment, the participants viewed all 200 photos at least once, so they were familiar with their contents and order. They were also given three minutes to practice using the Bookisheet device. The experiment started with the display of the target photo and the start of the timer. Prototype 1 is a single-display structure and consists of a single sheet, and only bending is used for scrolling photos. Prototype 3 is a dual-display structure and consists of two sheets joined together. Both bending degree and shearing force can be used for controlling scrolling. Therefore, for the third prototype, we measured the retrieval time for two cases; scroll photos by bending only and by both bending and rubbing along the edge.

Results

The average time it took each participant to complete the task and the average time for all participants are plotted in Figure 3.28 for each device. Bookisheet prototypes 1 and 3 had the same level of performance for retrieval as the book, the right- and leftarrow keys, and the mouse wheel. In the results for prototype 3, we can see that using



Figure 3.28: Experimental results.

a combination of bending and shearing force tends to show better results compared to using only bending. Comparing the results for prototypes 1 and 3, that is, the singleand dual-display structures, we can see the better results obtained with the dual-display structure.

3.6.2 Difference in Device Structure

We also collected comments from the participants. For prototype 1, the single-display structure, all of them had positive comments regarding using bending as a means of scrolling as it enabled them to control the speed of scrolling tangibly. All of them preferred using the exponential relationship rather than the linear one for the degree of bending and the speed of scrolling. However, all of them preferred the dual-display structure, that is, prototype 3. As we described briefly in sub-section 3.4.1, the dynamics of the hand holding a single sheet is different from that of holding a book. The third prototype, the dual-display structure, had exactly the same structure as a book, and the bending operation generated feelings similar to those when handling a book.

3.6.3 Degree of Bending and Shearing Force

For prototype 3, we found that with only bending operation, the user found it difficult to stop scrolling precisely at the target position. He/she had to stop bending consciously and keep the two sides of the pages flat to stop scrolling. As a result, he/she often passedby the target. With both bending and rubbing operation, the user could stop scrolling simply by stopping the rubbing, which immediately reduced the shearing force to zero.



Figure 3.29: Effect of features of target contents on retrieval behaviors.

This difference in the accuracy of stopping at the target caused the difference in retrieval times.

3.6.4 Finger-Bookmark Function

The finger-bookmarking function was explained to all the participants, and they were asked to execute it. We found that all of them were able to execute the function even though it requires three simultaneous operations: button pressing, sheet bending, and edge rubbing. Finger-bookmarking has not been implemented on conventional information terminal devices, and the participants found it very useful.

From our experimental results, we can say that our Bookisheet interface has the same level of performance for retrieval and paging as conventional devices (books, arrow keys, and mouse wheels). Bookisheet is an intuitive, tangible interface based on the metaphor of turning pages in a book. All eight participants in our evaluation were positive about its tangibility and the use of the book metaphor.

3.7 Discussion

Effect of Features of Targeted Contents

While there were no big differences in device performance, except for the touch panel, the features of the targeted contents did affect performance. In other words, the time it took

to find the specified photo depended on how well the participant remembered the overall contents. In the 200 photos, some photos were in a series with a theme (e.g., birthday party) and others were not. Some photos had a strong feature (e.g., a bright pink hue) that would tend to remain in memory and others did not. The target photos for each participant included at least one of these four types of photos; type 1 included photos that were in a series and had a strong feature; type 2 included ones that were in a series but had no strong features; type 3 included ones that were striking in some way but not in a series; and type 4 included photos that were not in a series and had no strong features.

Figure 3.29 shows the results broken down by the four types of photos for all the devices except the touch panel, which had the worst scores. The times were shorter for the photos in a series because the participants could generally recall the position of the series in the 200 photos and begin scrolling in the correct direction immediately. The times were higher for the photos not in a series because the participants had less help in recalling them. The times were shorter for the photos with a strong feature whether they were in a series or not. This finding is consistent with the results presented in Chapter 2 and confirms that content type and memorization affect retrieval behaviors.

Use of Book Metaphor for Accessing Digital World

Our aim was to develop a digital content browsing interface using the book metaphor that gives the user the same pleasant feeling obtained by leafing through the pages in a book. From observations of how people handle books and measurement of the forces they exert on the pages of a book, we identified two factors that play key roles in page turning dynamics: the bending and rubbing of pages.

Our first and second prototypes made use of only factor 1, bending, and the third prototype made use of the shearing force caused by rubbing along the edge as well. Through our development efforts and user study, we found that a dual-display interface device is better at generating a feeling of leafing through the pages in a book than a single-display one. Combining bending with the shearing force created by the thumb rubbing along the page edge is better at generating the feeling compared with only bending. This is the first attempt at using the book metaphor in hardware for accessing digital contents so as to obtain the unique pleasant feeling that comes from handling a book. We believe this approach, identifying the factors in creating the pleasant feeling that comes from handling a familiar object and applying them, will lead to user interfaces that will enable easier and more comfortable access to the digital world.

Book Thickness

In our second prototype, we implemented a function for reproducing the effect of the thickness of a book, enabling the user to open directly to a page close to the desired page by bending the sheet an appropriate amount. Thickness plays an important role in the handling of an actual book. It gives the reader an idea of the total size of the contents

3.7. DISCUSSION



Figure 3.30: Projecting images on sheet surface.

and helps the reader judge where to open the book so as to be close to the desired page. Grasping the total size of the contents is difficult to achieve for digital content using most conventional devices. One approach could be to use a touch sensor to judge the position of a page relative to the entire volume. The sensor could be used for implementing a function for adjusting the opening page instead of using a light-dependent resistor.

In our third prototype, we developed a device with some thickness, so it is more like a book than a sheet. We noticed that this thickness tended to give the user peace of mind. This was not observed for prototypes 1 and 2, which have thin-sheet structures.

Other Ways of Implementation

In this work, we separated the display from the interface device. Currently, it is difficult to obtain really bendable displays. Therefore, we tested projecting images, e.g., Wikipedia pages, on the sheet (Figure 3.30). The viewers' impressions differed from those when they were displayed on a computer screen. The technical issue is causing the projected images to follow the surface of the Bookisheet interface. Several techniques, for example, using multiple infrared reflective markers and cameras [57, 85], can be used to do this. We tested the use of ARToolKit [10] for detecting the position and orientation of the surface of the interface. Though the implementation is still in progress, we believe this will improve the quality of our interface. We could attach a small solid display to the bendable sheet, but this would restrict the degree of freedom of bending (only the very edge part would be bendable; the device would be too heavy for handheld operation, etc.). Nevertheless, attaching bendable parts to a conventional small display as an input device is worth studying. We would like to develop such a device and compare its usability with that of Bookisheet.

Chapter 4

Time-Oriented Information Browsing Interface

4.1 Information Navigation

In the previous chapter, we described an interface based on the book metaphor for browsing digital contents. This device is aimed at giving a person browsing digital information the tangible sense of leafing through the pages in an actual book. A book is an information medium that makes use of a strong human ability—managing and memorizing information spatially. Spatial information, such as the order of pages and the thickness of the book can be used as clues for navigation.

This chapter describes an interface that aims at giving a person the ability to browse information using time as a key. Although people are generally good at memorizing the locations of objects spatially in the real world, it is not easy to memorize the location of information spatially in the digital world. A hierarchical structure of directories is often used in a computer to manage information. However, navigation efficiency is reduced as the amount of information and the complexity of the hierarchy increase.

Most digital data have temporal information such as the time a file was modified. As in the real world, this temporal information can be a strong cue for retrieving information. We developed an interface that can be used to access large amounts of information regarding television, such as recorded programs and movies, on-air programs, and text information about future programs. These are typical types of data that are ordered temporally. They are densely arranged on the time axis.

4.1.1 Space-Oriented Navigation

While people are generally good at retrieving things stored spatially, for example, finding a particular object from amongst stacks of objects on an apparently messy desk, it is difficult to grasp the whereabouts of digital data spatially. As mentioned, using a hierarchical structure of directories to organize digital data in a computer becomes problematic as the structure becomes more complex and the information multiplies. Moreover, a hierarchical approach to organizing information is not suitable for every information appliance, particularly consumer products like TVs, because often times the user is unfamiliar with computers.

Several groups have conducted studies aimed at organizing digital data spatially like we do in the physical world [94, 3]. One approach has been to reproduce the arrangement of objects on a physical desk on the computer. Stacks of files are arranged at spatially different positions on a desktop-like GUI as groups of information. However, scalability is a problem with this approach [141]. It becomes more difficult to access desired information from the stacks as the amount of information increases.

4.1.2 Keyword Search

People often use key words to search for information using the Internet. For example, Google [49] and Flicker [46] present search results related to input keywords from the rich, and what some would say almost infinite, amount of information on the Internet. By selecting appropriate keywords, a user can get a list of information well matched to the desired information. However, the user is unlikely to find unexpected information not related to the keywords, and it is difficult to search for information that cannot be expressed precisely by a combination of keywords.

4.1.3 Time-Oriented Navigation

Time is a powerful cue in retrieving content-rich information. In real life, people often use the memory of the date and time of an event such as a trip to remember things that happened before or after the event. In the digital world, people sometimes organize information by using the date and/or time as a filename for version management. Rearranging files or received e-mails by their modified date or received time often helps in finding a desired file or e-mail efficiently.

Several groups have proposed techniques for managing digital information by using time as a key [119, 68, 97, 51, 92, 93, 81]. Their main focus was on reproducing the past state of works on a computer by using their histories, such as file modification or deletion dates. Regardless of the file names and contents, by treating all the data as time-ordered data, a user can remember the relationships between information at a particular time or remember information close to the time by association. Studies suggest that memories of the whereabouts of information along the time axis can be effectively used for information navigation.

4.1.4 Goal and Approaches

The goal of this research was to develop a browsing interface using time as a key for retrieving information efficiently. To reach this goal, we first focused on broadcast television programs and their guides (program guides) as target contents because they are typical types of information strictly managed by time. In particular, we did not simply use many videos but assumed a situation in which all the channels are continuously recorded. Next we developed a system that can record multiple channels continuously. We then developed a GUI that uses an analog clock as a metaphor for time-oriented information retrieving. The clock represented on the GUI is operated using a dial on a remote control. User can freely move along the time axis because the arm of the analog clock rotates in the same direction as dial rotation.

The rest of this chapter is organized as follows. Section 4.2 briefly summarizes the advances in television technology and the corresponding changes in TV viewing styles. Section 4.3 describes the system we developed for recording multiple channels continuously. Section 4.4 explains the GUI we developed that is based on an analog clock as a metaphor for browsing contents using time as a key. The evaluation results are presented and discussed in Section 4.5, related research is introduced in Section 4.6, and the key points are summarized in Section 4.7.

4.2 Advances in Technology and Changes in Viewing Styles

TV programs, a typical time-ordered type of data, are broadcast by TV stations in accordance with a schedule, which is generally published in newspapers and magazines and shown electronically on digital TV and the Internet. A decade ago, TV viewers typically checked one of these program guides to determine the time and channel of programs they wanted to watch or record. However, recent advances in information technologies for TV broadcasting and related products have brought about dramatic changes in viewing styles. In this section, we describe the main advances and predict how viewing styles will change in response to them.

4.2.1 Technological Advances

Large Capacity Hard Disk Drives

A decade ago, a person had to either watch his/her favorite programs at the time they were broadcast or carefully set the recording times and channels on his/her video recorder in order to record the programs. Nowadays, the style of watching/recording TV programs is changing dramatically with the spread of hard disk drive (HDD) recorders. Most commercial HDD recorders can store several hundred gigabytes of data, enabling users to record many TV programs; for example, most can store over 100 hours of high-definition

broadcast programs. Moreover, the programs to be recorded are set by simply selecting the desired programs from the displayed program guide, and the recorder can even be set to record them on a daily or weekly basis. This makes it much easier for users to record not only their favorite programs but also programs that they may enjoy viewing.

Moreover, as standards for home network systems [17], such as that of the Digital Living Network Alliance, DLNA [39], have been consolidated, makers have begun to market products that can be networked. This will enable the networking of several large-capacity HDD devices in the home—an HDD recorder in the living room, PCs in the children's rooms, laptops on the work desks (and even in the bathroom), and so on. Therefore, the total capacity of the linked HDDs in a home could be of the terabyte (TB) order.

Increased Number of Channels

As the forms of TV program distribution have become diversified, the number of channels a user can watch has increased. In addition to traditional terrestrial analog television broadcasts, programs are now available from terrestrial digital television broadcasts, satellite broadcasts, and cable providers, and more than 100 channels are typically available. Moreover, video sharing services such as YouTube [148] are widely used as sources of video contents. With these services, a user searches for a program of interest by using keywords rather than using a program guide.

Advanced Functionalities

In line with the social and technological trend of "broadcast-telecommunication linkage," some TVs now come equipped with functions for browsing the Internet [43, 150]. For example, a viewer watching a program about a historical event could simultaneously go on-line and retrieve information about the event.

Cable boxes with multiple TV tuners are now available. This enables a user to watch and/or record multiple programs simultaneously, eliminating the problem of deciding which program to watch and/or record. The same capability will be made available by the networking of all the TV tuner devices in the home.

4.2.2 Multi-Channel Continuous Recording

These advances will lead to the continuous recording of multiple channels becoming normal. The users will then have to search the contents of programs previously recorded and of programs currently being broadcast and information about programs yet to be broadcast, which is typically found in an electronic program guide (EPG). As a result, the amount and variety of contents that users can watch will become much greater, and the access path to the target contents will become more complex. The complexity of operating existing commercial interfaces will thus be increased. Interfaces that help users more easily operate digital devices will be necessary to meet a strong consumer demand. In



Figure 4.1: Multi-channel continuous recording system.

response to this need, we have developed a GUI that enables users to access a rich variety of contents in an intuitive and easy way.

4.3 Recording System

To experience the situation of multi-channel continuous recording, we built an experimental system for continuously recording eight channels simultaneously. The system configuration and specifications are shown in Figure 4.1. Eight TV tuners that can simultaneously encode eight TV signals [65] are connected in a cascade manner and connected to a PC through a USB 2.0 connection. We developed a GUI and two software libraries ("DSAccessDll.lib" and "EPGDll.lib") for accessing program and EPG data, respectively, recorded on the HDD. TV signals are received by the antenna, and the EPGs are obtained from the Internet.

The specifications of the PC and the recorded images are listed in Figure 4.1. Given these specifications, the system can record eight channels continuously for eight days. Furthermore, by managing the space on the HDD by using scripts, we can record eight channels for many days. The scripts delete older recordings to create space for newer ones. The EPG data covers the following week for all eight channels; it is managed using the



Figure 4.2: Main and schematic view of Multi-stream GUI.

Microsoft Office Access database management software.

The recorded programs, in MPEG2 format, are stored in folders labeled with the date, time, and station, for example, "10/Dec/2008/12:30/NHK-TV." Each file can hold up to 24 hours of recorded data. Using a more manageable length of 1 hour per file, we get 24 files per channel, 192 files for all eight channels each day, and 1536 files for the entire eight-day period. Similarly, the program information obtained from the EPG for each recorded program, such as "10/Dec/2008/12:30/NHK-TV/Daily News/Newsreader name/Brief explanation of program" is managed using the database management software.

Two important keys for searching for a particular program from this large amount of data are broadcast time and station name. Both keys can be found in the name of the folder containing the MPEG file in which the program is recorded and in the text information managed by the database management software. To find a particular scene in a program requires careful viewing of the file from the beginning. We thus designed a user interface for browsing these large amounts of data effectively. We utilized the fact that all these data are strongly related to time. They are densely arranged along the time axis from the past to the future.

4.4 Implementation

4.4.1 Multi-stream GUI

We developed a GUI named Multi-stream GUI that can simultaneously handle multiple data inputs at any point on the time axis. Figure 4.2 (a) shows an overview of it. An analog clock for setting the time is placed at the center, and eight windows for the eight

4.4. IMPLEMENTATION



Dial: rotate arms of clock. Colorkeys: change minutes unit of arms movement per unit rotation. Arrowkeys: move cursor. Selectkey: transfer between multi- and full screen.

Figure 4.3: Remote control.

channels are arranged around the clock. The software was developed using Visual C++. NET 2003 and Microsoft DirectX SDK 9.0C (December 2004); DSAccessDll.lib is used for handling the program data, and EPGDll.lib is used for handling the electronic program guide (EPG) data (see Figure 4.1).

Figure 4.2 (b) shows a schematic view of the interface. In addition to the analog clock and eight windows, the main components are a cursor with small triangles indicating possible cursor movement directions, a ticker area for presenting program information in text, a field showing the current time, operation guidance controls, and a view mode area.

Figure 4.3 shows the remote control we developed for operating the Multi-stream GUI. It was made by reconstructing a conventional commercial remote control. A dial switch instead of the regular arrow keys is embedded in the upper portion. This switch can be used as four regular direction arrow keys, a select key, and a dial. Dial rotation moves the arms of the clock and produces a clicking sensation. Eight clicks are felt for each 360° rotation (rotation unit is 45°). The infrared signal emitted by the remote control is picked up by a receiver connected to the PC through a USB connection, and software running on the PC translates the signal into key press events. The assignment of key events and buttons on the remote control is programmable. The other buttons, such as the blue, red, green, and yellow ones, can be programmed as well. We programmed the four colored buttons to set the time interval for arm movement per unit rotation (1, 5, 10, or 30 min per 45° rotation).

The arms of the analog clock are rotated by turning the dial on the remote control, and the cursor is moved by operating the arrow keys. The cursor can be placed on any of the eight windows, the analog clock, and the view mode area. Usually, one of the eight windows is selected using the cursor, and sound is heard from only this window. If a program is presented in this window, related information is shown in the ticker area.



Figure 4.4: Views on time axis.



Figure 4.5: Multi-screen and full-screen modes.

4.4.2 Views on Time Axis

A key feature of this GUI is using the analog-clock metaphor for setting the time. Users can move the clock arms by rotating the dial on the remote control. Programs or EPG data corresponding to the time on the clock are shown in each window. Here we explain the operations for the three display modes.

Current View

The application starts up in "current view." The clock displays the current time (Figure 4.4, top), and eight "on-air" TV programs can be viewed in the eight windows. Image data for recorded and on-air programs and/or text data for EPG information are spread around the virtual space defined by the time and channel axes. The information on a surface cutting a point on the time axis is presented in the windows, and currently on-air programs are shown in each window. If the user selects one by pressing the select key, the program in that window is shown full screen and the sound for that program is played (Figure 4.5). The programs, not still images, are shown in the other windows (which is possible because there are eight tuners). The arms of the clock are shown in white, and the background color is brown for AM and gray for PM to enable the user to immediately recognize the time setting of the GUI.

Past View

If the dial is rotated leftward, the arms on the clock are wound backwards, the view switches to "past," and programs recorded on the HDD can be watched. The surface (yellow) cuts the time axis at a point corresponding to recorded programs (pink), as shown in Figure 4.4 (middle). The application searches for folders labeled with the corresponding time and finds the files with the programs recorded for the eight channels at that time. It then sets the start time for each program to match the time indicated on the clock. For example, if the clock indicates 10:15 (Figure 4.4, middle) the application searches for the eight folders labeled with a corresponding time, such as 10:00. It then sets the start time for each program accordingly, such as 15 minutes from the beginning of the file. The arms of the clock are shown in orange, and the background colors are the same as in current view.

Future View

If the dial is rotated rightward, the arms of the clock are advanced, and the view switches to "future." The surface (yellow) cuts the time axis at a point corresponding to EPG information (blue), and EPG text information is displayed (Figure 4.4, bottom). The application searches for the EPG information in the database with the corresponding time and finds the items for the programs to be shown at that time for the eight channels.

A speaking text function was implemented specifically for this view. The text information is converted using text-to-speech (TTS) technology. The GUI speaks the text using a TTS engine implementing Microsoft's speech application programming interface (Speech API). Speech synthesis technology could be a good way of implementing the "10-foot user interface" supposedly needed for the interfaces of digital devices.

Both text information and the current on-air programs are presented in the subwindows. This is because it is likely that users will switch to future view to obtain information on the programs that will air immediately after the current on-air programs. That is, the user may want to continue watching a current on-air program while checking information on the immediately following programs. The arms of the clock are yellow, and the background colors are the same as in current view.

Setting Time Interval for Arm Movement

As described above, the display view is selected by rotating the dial to move the arms of the clock, and dial rotation produces a clicking sensation. Eight clicks are felt for each 360° rotation (rotation unit is 45°). The arms of the clock move in 1-, 5-, 10-, or 30-minute intervals for each click depending on the setting. The interval is set by pressing the corresponding colored button on the remote (Figure 4.3). The default setting is a 5-minute interval (blue button).

4.4.3 Comparison with Conventional GUIs

The conventional GUIs found on commercial HDD recorders generally do not display the lists of recorded contents, current on-air programs, and future program information in a unified manner. Moreover, to switch between these three modes, different buttons on the remote control must be pressed. As shown in Figure 4.6, the conventional displays are usually vastly different. The list of recorded programs may contain a thumbnail image and descriptive text for each program, such as the title, station name, broadcast time, and performers. The cursor is moved using the four direction arrow keys on the remote, and only the program for which the thumbnail is selected is played. In contrast, the EPG view



Figure 4.6: Conventional GUI on commercial HDD recorder.



Figure 4.7: Proposed GUI.

shows the time axis vertically and the channel axis horizontally. In both views, the cursor is moved in a 2D plane, vertically or horizontally. This 2D scrolling operation forces the user to read text information, such as program titles, rather than feel the time element and use it as a key for retrieval. Switching between the three display modes is not seamless, and this complexity and lack of seamlessness can make the user operations unpleasant.

On the other hand, in our Multi-stream GUI, the basic designs of the recorded program, on-air program, and EPG information views are the same (Figure 4.7). Additionally, there is no need to press different buttons such as "EPG" or "Playlist" to access them. The user simply rotates the dial to seamlessly switch between them. There is no need to scroll to an area not shown on the display, and scrolling in one dimension, the time axis, enables the user to navigate the contents using time as the key, not text or keywords. The operation of the control dial is reflected in all eight windows. This means that not only can all eight channels be watched at the same time, but all the channels can be fast-forwarded or rewound simultaneously without switching the target channel.



(a) Calendar view.





(b) Cursor has been moved onto the clock.



(c) Pop-up menu appears.



(e) Calendar area is active.

(d) "AM/PM" area is active.



(f) Rightmost area is active.

Figure 4.8: Calendar view function.

4.4.4 Calendar View

It is possible to set the time or search near the current time setting by rotating the dial to turn the arms of the clock. However, if the target day is some distance from the current time setting, setting the time simply by rotating the dial is not efficient. We therefore implemented another function called "calendar view" as a pop-up menu for setting the time and date more easily in such cases (Figure 4.8 (a)). The user can set the day, AM/PM, and time. "Return to current time," "one week before," and "one week later" can also be set. Using this function, users can easily search for programs broadcast some time ago, or for information related to programs that will be broadcast some time in the future.

To use this function, the user moves the cursor onto the analog clock (Figure 4.8 (b)) and presses the select key to open the pop-up menu (Figure 4.8 (c)). This menu is semitransparent, so the program being shown can still be seen. The pop-up menu is separated into four areas: analog clock (for time setting) in the center, AM/PM and date (calendar) on the left, and one week jumping or returning to current time on the right. One of these four areas is selected by pressing the arrow keys. Once the area is selected, the values in the area can be selected by rotating the dial. For example, if the current date and time are 10 December 2008, 10:30, and the user wants to jump to 5 December 2008, 20:45, to watch a recorded program, he/she opens the pop-up menu and selects the clock area (Figure 4.8 (c)). He/she rotates the arms of the clock leftward to set the time at 08:45. Next, he/she presses the left key once to activate the AM/PM area (Figure 4.8 (d)) and switches from AM to PM to change the set time from 08:45 to 20:45. Then he/she presses the left key again to activate the calendar area (Figure 4.8 (e)). The calendar rotates as the user rotates the dial. Once the desired date, 5 December, is reached, the user presses the select key, the pop-up menu disappears, and the programs recorded at that date and time are displayed in the eight windows.

4.4.5 Category View

Our Multi-stream GUI is designed for browsing TV-related information including recorded programs, on-air programs, and EPG information using time as a key for retrieval. The information presented in the eight windows corresponds to a surface cutting a point on the time axis.

Filtering information on the basis of certain keywords is demanded by many users, not for browsing, but for retrieving. Filtering is also demanded in terms of gathering a particular set of programs and selecting one to be watched. We thus developed another function called "category view." A screen image of it is shown in Figure 4.9 (a), and a schematic illustration of it is shown in Figure 4.9 (e). It differs somewhat from the channel view. First, the windows are not arranged in rows but in a spiral. Second, each window



(c) Cursor has been moved to upper-left (d) Pull-down menu has been opened. corner.



() 0 /

Figure 4.9: Category view function.

does not correspond to a channel. Instead, the windows represent programs in selected categories, so each window has no fixed channel as in channel view (see Figure 4.2 (b) or 4.8). Third, the cursor is fixed at the upper-left window.

To switch from channel view to category view, the user moves the cursor to the upperleft window (Figure 4.9 (c)) and presses the up arrow key to move the cursor onto the view mode area to open the pull down menu (Figure 4.9 (d)). From the pull down menu, the user selects the desired category by operating the arrow keys. The view then changes from channel view to category view (Figure 4.9 (a),(e)). For example, if the keyword "movie" is input, TV programs in the movie genre are shown in the windows. A movie broadcast at the set time is shown in the uppermost window (top-left), and movies shown previously are arranged behind it in a rightward spiral. For example, if the user enters the title of a movie program that was broadcast at 17:00, the program is shown in the top-left window, another movie program that was broadcast at 12:00 is shown in the window to the right, and so on. If the arms of the clock are rotated leftward, the movie next in line moves to the uppermost window, and the one previously in that window disappears. The analog clock indicates the starting time of the newly selected program, which was broadcast before the previous one. The images of the older programs that are retrieved one by one with this process appear to gush out from the center (behind the clock) in an animated manner as the arm is rotated leftward. If the arms of the clock are rotated rightward, the image of the newer movie apparently hidden to the left of the cursor appears. If the newer information is from an EPG, the text programming guide for the selected category appears. The images appear to be sucked into the center in an animated manner as the arms are rotated rightward.

4.5 Evaluation

We have been using the developed multi-channel continuous recording system and GUI since 2005 in our laboratory. We have also shown the GUI at several exhibitions including *CEATEC 2005* in Tokyo and *CES 2006* in Las Vegas and have received much feedback (Figure 4.10). Here we describe some of our laboratory findings and exhibition feedback.

4.5.1 Laboratory Findings

Laboratory tests on the GUI revealed some interesting aspects.

Time-oriented Browsing One of the most impressive and interesting comments from users in our laboratory is something like,

"I did not know that such an interesting TV program was being broadcast by that station at that time!?" This comment reveals that users tended to discover unexpectedly interesting programs at unexpected time in unselected windows, that is, windows that were not showing the initial target channel. It may also indicate that some users simply rotated the dial on the remote without any target program in mind. They just moved forward and/or backward along the time axis, checking out the programs displayed in each window. During this "killing time" operation, they tended to discover interesting programs unexpectedly.

Another comment from laboratory users was something like,

"I wanted to watch the program broadcast around 10 o'clock last night from the beginning because I missed the first part."

The user simply had to rotate the arms of the clock backward and set the time to around the time the program was aired. He/she then set the time to that at the beginning of the program. This shows that a user can retrieve a desired program by simply roughly remembering the time it was shown. The time setting operations can be easily done using the analog clock metaphor. This is because everyone, from child to grandparent, knows how to use an analog clock. The operation of rotating the dial is very simple and intuitive, and users can easily change the setting at any time. Furthermore, the operation of rotating the dial is reflected in all eight windows. This means users can handle eight streams simultaneously and view eight channels at the same time.

Effects of Filtering by Categories We found that the category view function is very useful for searching for contents in a proactive manner. Filtering by category, title, artist, etc. shortens the path to the target contents. Navigation along the time axis is possible in this view as well, and the time line, though represented as a spiral, was sensed well by the users. Moreover, presenting multiple images of same the program at once on the time line was useful for searching for multiple episodes of the same program, and gave user a satisfaction of collecting his/her favorite programs.

Changes in Television Viewing Styles We found that using multi-channel continuous recording changes a person's television viewing style.

First, "pre-set recording" is no longer necessary because all the programs are automatically stored in the HDD. As a result, no one says, "Shoot, I forgot to record an important program!" any more. The task becomes finding the desired program from content-rich information, and this is the task for which our Multi-stream GUI was designed.

Second, we found that a group of people could more easily discuss topics related to TV programs because, when someone mentioned an interesting program he/she had watched, the others could use our interface to find and watch it. As a result, a person could find interesting programs in a genre or with performers he/she had never watched before they were recommended by others. This non-automatic but user-initiated recommendation method is well suited for today's world of content-rich television.

4.5. EVALUATION



Figure 4.10: Exhibition at CES 2006.

4.5.2 Exhibition Feedback

The GUI we developed was originally designed for television viewing, but its ability to handle multiple streams of time-ordered data simultaneously make it potentially applicable to other types of applications, as revealed by some of the comments we received at the exhibitions.

Surveillance A number of the comments mentioned using it not only as a TV interface but also as an interface for surveillance camera monitoring systems. Such systems typically use multiple cameras and record continuously, and our GUI is well suited for monitoring multiple video inputs.

Medical Diagnosis Several comments noted that various types of patient records, such as x-ray photographs of certain body parts, should be managed as time-ordered data to enable the doctor to grasp the changes over time. This would be an excellent application for our time-based GUI.

File Management Other comments pointed out that other types of media, such as home video, digital photos, and MP3 files, are stored on HDDs. Since they are also stored in files with time stamps, our GUI could manage them as well.

Highlight Scene Playback Several models of HDD recorders and TVs now come equipped with a function for detecting the highlight scenes, such as a homerun in a baseball game. Many of the people making comments were familiar with this "digest" function and mentioned that showing thumbnails of detected highlight scenes in a time-ordered manner, as in our "category view," would help them to find a desired scene.

4.6 Related Work

This section first describes research on information-retrieval techniques using time as a key. Next, background technologies related to user interfaces on television are described. Finally, the placement of this research is discussed.

Time-Oriented Information Browsing

Several studies have investigated the use of information-retrieval techniques in which the state of the work environment on a computer at a particular time is reproduced to help the user associate the target object with memories from around that time.

Rekimoto et al. described a time-centric user interface called "Time-Machine Computing" [119] that can be used to manage the state of the computer desktop by time. It aims to put information in order not only spatially but also temporally by managing the histories of operations such as creating and deleting files, reproducing the desktop state at a particular time, and visualizing the time information. The user specifies the time for the reproduction, and information sources, such as files that were on the desktop at that time, are presented where they were. They used various techniques to give the user a sense of the time, such as using semi-transparent menus and coloring the displayed icons so as to show the time order.

Ishii et al. investigated many computer-human interaction techniques that use familiar physical objects. Their "ambientROOM" system aims at helping people notice digital information unconsciously [68]. In this system, people use physical objects as controllers to retrieve information. For example, a physical analog clock is used to reproduce past information or access future information. If the arms of the clock are rotated leftward, to the past, the amount of e-mail received at the indicated time is represented by the sound of rain or the brightness in the room. During the operation of moving to the past or the future, the current time is represented as a shadow of the arms projected on the surface of the clock.

Masui et al. made use of the fact that, when people try to remember something, they generally use vague association rather than indicate a specific retrieval condition such as a keyword [97]. That is, even if someone cannot remember the desired information directly, he/she can usually retrieve the information by making use of the relationship between it and related information. The related information is remembered step by step until the target information is finally reached. They applied this to the retrieval of information on a computer. The degree of relatedness is determined by the contents of the files, the structure of the directories, keywords, and time information. Files with a high degree of relatedness are presented around the watched file. For example, files and other kinds of information sources created at about the same time as the file are presented around it. In this way, information is retrieved by using time association as a retrieval key.

4.6. RELATED WORK

Graham et al. developed a calendar browser for organizing and presenting photos by making use of the time stamps on the photos [51]. They compared its operation with that of a conventional browser based on scrolling retrieval and found that the calendar browser was faster for finding target photos. They implemented a function in their calendar browser for automatically clustering photos with close time stamps.

Some studies have investigated the visualization of time-ordered data using 3D expression. For example, the "perspective wall" proposed by Mackinlay [92, 93] presents a schedule or time-ordered files in a perspective view. The "dynamic timeline" interactive photo browser presents photos chronologically using semi-transparency or zooming technique [81].

Mynatt et al. also made use of the fact that time is a powerful cue for retrieving content-rich information. They developed a whiteboard system that can reproduce text previously written on a whiteboard [106].

Background Technologies Related to TV User Interfaces

Server-Type Broadcasting In terms of recording large amounts of data related to TV programs, a technology called "server-type broadcasting" has been studied [12, 75]. It assumes that there is a large-capacity server available in the home. The broadcaster sends program information, such as title, broadcast date and time, and performers, as metadata. A user makes use of this metadata to decide which programs stored on the home server to watch. Our proposed browsing technique based on the analog clock metaphor is used actively by the user to find and watch contents on a large-capacity local hard disk drive, while server-type broadcasting aims to let users watch stored contents rather passively.

Speech Technologies Speech technology is a strong candidate for achieving a humancentric user interface and has been applied as an operation method for television [50, 27, 143]. The biggest advantage of speech technologies is that they could enable the user to operate a device using spoken commands rather than a complex GUI or remote control. With it, virtually anyone would be able to operate a device without having to learn how to use an interface. However, in practice, its application to television is problematic due to limitations on the speech recognition rate and the inherent problem of recognizing synthesized speech in a television environment.

Cultural difference may affect the use of speech technologies. Tan et al. compared the acceptance of using speech commands between users in the U.S. and in Japan [134]. Users compared the use of speech commands with 100% recognition by the "Wizard of WOZ" method to remote control operation. They found that users in the U.S. could easily use the speech commands and preferred doing so while users in Japan had more difficulty and preferred using a remote control.

Another approach to operating a television using natural dialogue is to use an avatar as an agent that receives the commands and operates the device. The idea is to reduce the operational complexity by letting an avatar perform the tasks and thereby improve the attractiveness of the interface [102, 15].

Scene Detection As the amount of watchable video content increases, the need for effective ways of browsing such content has grown, and technologies such as ones for detecting specific scenes have been developed. For example, Li et al. investigated how users watch videos using the various functions made available to them and found that the ones most used were those for time compression, pause removal, and navigation using shot boundaries [87].

Detecting particular scenes efficiently is another area that has attracted much attention. Of particular interest is detecting highlight scenes in sports programs [133, 86, 59, 78, 146, 35], and a function for detecting them is now available in some commercial TVs and HDD recorders. Highlight scenes, such as a goal in a football match, can be detected by detecting changes in the audio and/or video signals or by recognizing superimposed text information. Color or edge information of the objects in the video can be used to detect specific objects in the video. For example, a technique was developed that detects only players in the scenes of a football match [138].

Several methods for presenting detected scenes have been developed and evaluated. For example, Boreczky et al. developed a method for presenting detected key frames in comic book style [21]. Christel et al. developed an interface that arranges thumbnails of video scenes spatially on a map or temporally on a time line [34]. It recognizes the contents of the video automatically using audio, text, images, or facial images of people and arranges the thumbnails in the related area on the map. It also arranges video contents filtered by keywords along the time line to make it easier to retrieve them using time as a key. Another study investigated the relationship between the display rate, the number of key frames, and user perception when detected key frames are presented as a slideshow [137].

Other studies investigated ways of navigating programs using active searching by the user rather than using automatic system recommendations. Wittenburg et al. developed a technique for retrieving desired scenes that presents thumbnails taken from a video at fixed time intervals, which greatly reduces the use of fast-forwarding and rewinding [142]. Testing showed that a user could find the indicated scene by remote control operation more accurately with this technique, but not more quickly.

A similar technique was developed by Drucker et al. It takes thumbnails from a video at fixed time intervals and superimposes them on the display to help the user find a desired scene by fast-forwarding and rewinding [40]. The thumbnails are arranged horizontally at the bottom of the display, and the one selected is shown a bit larger. The time interval is adjustable and ranges from 10 seconds to 8 minutes. Testing revealed that, though this technique takes a bit longer and requires more button pressing on the remote to find the target than the user interfaces on TiVo [136], the users reported a higher level of satisfaction with the proposed technique.

Personalization and Program Recommendation As the number of TV channels and the amount of video content increases, letting the system chose what to show has been another topic of interest. Rather than watch channels that they select, users watch preset programs or ones recommended by the system.

Many program recommendation systems using the user's profile have been developed [69, 70, 130, 33, 61]. Isobe et al. identified four styles of watching TV by observing representative viewers [69]. They then developed a system that presents recommended programs in one of four styles, which the user can specify depending on the degree of program selectable they want: "watch preset channel," "watch program recommended by system," "select one from list of recommended programs," and "retrieve program." Zimmerman et al. developed a system that improves the reliability of recommendations [149]. It guides users to new programs by presenting recommendations based on a conversational process with the user.

Remote Controls Devices that control GUIs enable users to browse TV content conveniently. Most conventional remote controls have so many buttons that users are often unsure of how to use a function. Various remote control devices have thus been developed with reduced complexity.

One approach is the universal remote, which can be used to control the TV as well as various other devices in the home. Conventional handheld devices, such as PDAs and tablet PCs, could be used as a universal remote [121, 109, 19]. The graphical user interface displayed on the handheld device would change automatically depending on the target device. A particular benefit of using a conventional device with a display is that the menu graphics would be displayed on the handheld device while the video contents would be displayed on the main display without being obscured by a graphical overlay. Kohtake et al. developed a technique for using one device to control the operation of several devices connected through the Internet [76]. A camera in the control device recognizes the 2D marker attached to each device to be controlled, and data can be copied or moved from one device to another by "pointing" at the appropriate devices. Preset recording can be initiated by pointing at a recording device.

Creating new hardware is another approach. Komine et al. embedded a trackball in a remote control to reduce the number of buttons and optimized the GUI for the remote [79]. User testing revealed that users looked at the remote control less frequently to confirm the operation than they did with a conventional remote. Ferscha developed a tangible cube-shaped remote control small enough be held in the hand [44]. Rolling or rotating the cube controls TV operations. Paper and pen based input techniques have also been applied to remote control. Hess proposed using a paper-based interface and text recognition for changing channels, controlling the volume, and inputting text [55].

Placement of this Study

In this research, we proposed using time as a key for retrieving or browsing contents and developed an interface for searching a large volume of TV-related data, which is typical type of time-ordered data. In the GUI we developed, we used the metaphor of an analog clock to enable the user to intuitively and effectively recognize the time. We also developed a remote control with a dial switch to operate the analog clock on the GUI. By simply turning the dial, a user can rotate the arms of the clock and seamlessly move along the time axis from the past to the future (and back again) to browse the contents. Related research on time-centric information browsing interfaces mainly focused on reproducing past states of a computer and the data on it at that time. Compared to these studies, our interface treats the past, present, and future equally, and the target data, i.e., data related to TV programs, is densely arranged along the time axis.

Our proposed interface has a high affinity for existing and developing TV-related technologies such as scene detection, program recommendation, and speech technologies. Combining these technologies with our interaction technique featuring time as the key for browsing should result in an intuitive and effective information browsing interface.

4.7 Discussion

We have developed an interface that handles recorded TV programs and program information using the metaphor of an analog clock. We use an analog clock as a metaphor because TV programs and the information about them in an electronic program guide (EPG) are typical types of time-ordered information, so using time as a key for retrieving a desired program or browsing along the time axis should work well. As described in the previous section, users could handle the interface easily; that is, they could operate the clock intuitively. The time-oriented browsing method generates positive effects such as discovering unexpected interesting programs and retrieving a desired program by roughly remembering the date and time it was broadcast. The analog clock metaphor was used to give users the sense of moving freely along the time axis. We think that time-ordered data such as that for TV programs and related information can be suitably managed using the proposed interface.

Time is a powerful cue for remembering an event in daily life. We created a situation in which all TV programs broadcast on eight channels were continuously recorded. In this situation, the data collected for the recorded programs, on-air programs, and EPG were densely arranged along the time axis. Users tend to roughly remember the date and time when they watched a particular program, so they can usually navigate to that program later using that memory. Our GUI supports such time-oriented navigation by giving the user the sense of time through the use of an analog clock, which can be operated freely to
move along the time axis.

Multi-channel continuous recording is one possible near-future television viewing style, and our Multi-stream GUI was designed especially for that situation. In addition, our timeoriented interface is well suited for various movie-content viewing situations and related technologies are rapidly advancing. In our study, we focused on contents broadcast by TV stations as the data to be managed using our interface. In addition to such contents, shared videos on the Internet can be accessed. The contents submitted to video sharing services have temporal information as well, including the time they were uploaded and/or downloaded. This information can be used for retrieving such content using our timeoriented GUI, together with other retrieval techniques such as keyword search.

The number of video channels available for viewing continues to increase. One of the most frequently asked questions at the exhibitions was the ability of the GUI to handle more than eight channels. From personal experience, we think eight is the maximum number of channels that can be watched at one time. The problem is finding a way to select, from the hundred or more channels available, which ones to watch. One approach is to combine our GUI with a program recommendation system. Besides program recommendation systems, there are many related technologies that could be used with our interface to improve the user's experience. Since a person using our interface to search for contents must be relatively active, combining it with technologies that the person can use passively, such as program recommendation and highlight scene detection, should work well.

Finally, in this era of "broadcast-telecommunication linkage", TV viewing styles are changing day by day [43, 150]. In addition to the programs broadcast by TV stations, there is a rich supply of video contents on the Internet. With the spread of Internet video delivery services, TV viewing styles have changed from passive browsing to active searching [23]. Moreover, interactive television and digital television have been introduced, and communicating with other users through the TV is possible [90, 28, 53]. Various user interfaces and designs for them have been eagerly studied [120, 72, 42, 144, 113]. Our future and continuous challenge is to explore ways of accessing these rich collections of digital contents effectively and comfortably. As described above, time is a strong cue for intuitively managing a rich collection of temporally ordered contents. Management of contents on the Internet, such as YouTube videos, is mainly done using keyword search. However, for navigating contents on the Internet, other methods must be developed that make navigation more efficient. We believe that our proposed interface using time as a key can be widely applied in the era of broadcast-telecommunication linkage.

Chapter 5

Conclusion

5.1 Retrieval Behaviors and User Interfaces

We have covered three topics: analysis of information-retrieval behaviors, use of a book metaphor interface, and use of a time-oriented navigation interface for television. We first investigated how people operate user interfaces depending on the degree of memorization of target contents. We found that there are four typical operation patterns for keyboard operations and four for mouse operations. In particular, the relative position of the target content or the entire order of arrangement affects the memorization, rather than the content itself. These findings indicate that the relative position and order of arrangement play important roles in information-retrieval behaviors, that is, how people operate user interfaces.

The human ability to memorize information spatially is well known in terms of memorizing the location of information by using knowledge of relative position. A book is a typical physical information medium for which we use this ability. For example, we can easily guess the whereabouts of a particular page in a book. Our book metaphor interface handles contents arranged in one dimension, and users browse them by leafing through the pages. We think that there must be typical patterns of leafing through pages, like the four patterns we found for keyboard and for mouse operations. We also think that, once these typical patterns are identified, we can develop an algorithm that represents contents appropriately depending on the pattern used. For example, if the pattern is quick leaving, familiar contents can be shown with emphasized representation, and if the pattern is slow leaving, unfamiliar contents can be emphasized. The results of our experiment on finding photos using our developed interface showed that the retrieval time tends to depend on the photo type. This indicates that the leafing pattern is affected by at least the target contents type and that there must be other factors such as memorization of relative position and order of arrangement.

Similarly, typical patterns of operating an analog clock were identified for the time-

oriented interface using the analog clock metaphor. This interface handles contents arranged on a time axis in one dimension, and users browse them by moving along the time axis freely. The whereabouts of content on the time axis can be guessed by referencing absolute temporal information instead of relative position. How far the target is from the current position on the time axis is expected to affect how users operate the interface.

Independently, these two user interfaces are an improvement over existing interfaces and enable efficient and comfortable access to desired information. Together, however, along with a better understanding of the mechanisms people use to retrieve information, they will enable a truly natural and highly interactive computing environment for contentrich information browsing.

5.2 Summary

This thesis described two interfaces for accessing desired information efficiently and comfortably. We first analyzed how people retrieve information and used the findings to establish a basis for designing improved interfaces (Chapter 2). There has been no previous research focusing on the effect of various factors such as the user's ability to recall information and the features of the target contents on how a user operates the interface for an information appliance. We analyzed the relationship between the degree of memorization and the patterns of user operations. Using the understanding gained, we developed an improved user interface for browsing digital contents (Chapter 3). It is a bendable device based on the book metaphor that recreates the factors that generate the pleasant feeling that comes from leafing through the pages in a book. This is a pioneering attempt at applying the book metaphor to hardware. We also developed a time-oriented navigation interface that uses time as the key to retrieval (Chapter 4). It is based on the analog clock metaphor and can be used to efficiently manage time-ordered data such as that for TV programs and related information.

This thesis makes three particularly important contributions.

Clarification of Effect of Memorization on Information-Retrieval Strategies (Chapter 2)

The results of our user experiment clarified the effects of memorization on the operation patterns of a user interface. The aim was to better understand the mechanisms of information retrieval and apply them to the design of user interfaces for information appliances. Such an understanding should make it possible to create an algorithm that adjusts the interface to the user's operation behaviors.

We first investigated the behaviors used to retrieve photos. Though we did not investigate all the situations needed to fully understand retrieval behaviors, we believe our results do contribute to understanding the fundamental mechanisms. We made use of the

5.2. SUMMARY

fact that the degree of memorization affects the way people operate information appliances. We investigated the relationship between scroll operation patterns and (1) memory of photo contents, (2) memory of photo positions, and (3) type of photo, which can affect memorization. We plotted the scrolling patterns, i.e., position (photo ID) and velocity against time. Our hypothesis was that the shapes of these patterns could be classified on the basis of the degree of memorization and the type of photo.

The results of our user experiment showed that particular operation patterns tended to appear depending on the degree of memorization. We found that the larger the degree, the higher the percentage of particular operation patterns, and the lower the degree, the higher the percentage of other patterns. Though the system is unable to judge the degree of memorization, the degree can be affected by the features of the photos. We classified the photos into three types, "series," "impressive," and "featureless," and measured the ease of memorization for each type. The operation patterns tended to uniquely correspond to a particular type. The features of the target contents that affect the degree of memorization can be interpreted by a computer.

Given these findings, we can construct an algorithm that identifies the features of the intended target contents on the basis of the way the user operates. For example, if the user operates the widgets of the user interface, such as the arrow keys and scrollbar rapidly, the system can be instructed, for example, to present only photos with a particular feature or ones that have been frequently accessed because the user's operation behavior can be interpreted to mean that he/she tends to search for familiar photos. On the other hand, if the user operates the widgets slowly (carefully) or irregularly, the behavior can be interpreted to mean that he/she tends to search for unfamiliar photos. In this case, the system can be instructed to present photos with "weak features" or ones with a low rate of being accessed, for example. Such an algorithm can be easily and quickly applied to any photo searching or browsing software system.

Our study contributes to the field of computer-human interaction because it will enable the design of user interfaces based on an understanding of the mechanisms people use to retrieve information. Many of the studies on user interfaces have focused on those components that the user operates directly and have resulted in, for example, novel input techniques and intuitive GUI designs. We focused instead on the underlying mechanisms of retrieval behaviors so that we could design not only the outer layer of user interfaces but also the depth.

Identification and Application of Factors that Generate Pleasant Feeling (Chapter 3)

To realize a user interface that can be used to access content-rich information both efficiently and comfortably, we developed a device for browsing digital contents that provides a tangible sense of leafing through the pages of a book. We first investigated the factors that generate the pleasant feeling. While the dynamics of page turning remain a mystery, we have identified the two main factors. One is the elasticity created by the bending of the pages, and the other is the shearing force created by the rubbing of the edges of the pages with the thumb. Using this knowledge, we designed a bendable device for handling digital contents using the book metaphor.

While real-world-oriented user interfaces have been studied well, many of them use objects in the real world simply as a metaphor to reduce the psychological load of operation. For example, though there has been much research on the use of the book metaphor for information browsing, it has been mostly implemented as an interface in software. We think our effort is the first attempt to implement it in hardware. Doing so not only reduces the psychological load of operations but also gives tangible feedback similar to that of handling an actual book. In other words, our interface is not simply physically oriented and tangible, it is also sensibility oriented.

In terms of memorization ability and information-retrieval behaviors (Chapter 2), our developed device makes use of the ability people have to spatially manage and memorize information, enabling it to handle content-rich information. Our device handles digital contents arranged in one dimension and in page units, in the same way an actual book presents information as pages. The order in which information is arranged must play an important role in the way people memorize it and navigate to find it. Arranging information in one dimension, rather than in a hierarchical structure, and then accessing the information using a device such as ours should reduce the complexity of navigation and enable more intuitive interaction.

Our developed device demonstrates the feasibility of applying our findings to flexible displays in general, which are expected to become widely used in the near future. Current flexible displays are not well suited for use in our bendable device because they are technologically immature and limited in capability, e.g., the refresh rates tend to be too low. However, our proposed technique can be applied to them once these problems are resolved. We think investigating and developing interaction techniques for applying flexible displays is worthwhile because flexible displays will certainly change the trend in the development of information appliances once they are popularized. We believe the interaction technique used in our developed interface will play a key role in appliances using flexible displays.

Development of Time-Oriented Information Navigation Technique Using Analog Clock Metaphor (Chapter 4)

An information navigation technique using time as a key has been proposed and applied as an interface for television-related data, which is a typical type of time-ordered data. Using the analog clock metaphor, which everyone is familiar with, enables a user to freely navigate information arranged along the time axis.

Time is a powerful cue for associating memories in daily life. While related research

has used time as the key to navigating information, most of it has focused on reproducing the past states on a computer. Our interface handles not only past information, but also current present and future information densely arranged along the time axis in the same manner.

The metaphor we used to give the user an intuitive sense of time is an analog clock, an object everyone knows well. The ease of understanding the relationship between rotating the arms and setting the time enables users to easily and freely navigate along the time axis. The consistency between turning the dial on the remote control and the rotation of the clock arms on the GUI produces a strong feeling of fast-forwarding or rewinding time and thereby moving freely along the time axis.

One of the big effects of navigating information by moving along the time axis is encountering unexpected information at an unexpected time. This was frequently demonstrated in the user evaluations. We believe the ease of time handling produced a sense of free and easy movement along the time axis, and this led to the discovery of unexpected information at an unexpected time. Such discovery is not possible with keyword search or a conventional thumbnail- and text-based 2D scroll interface.

In terms of an interface for browsing TV programs, our study contributes in the respect that it predicted and realized the situation of multi-channel continuous recording. Though styles of TV viewing have been changing day by day due to technological changes related to the trend towards "broadcast-telecommunication linkage," our predicted situation will likely be one of the future video content browsing styles. We believe that our interface has high affinity for other technologies and possible TV watching situations.

5.3 Future Directions

One of our future challenges is to establish a design framework for a user interface based on an understanding of the mechanisms people use to retrieve information. We investigated the relationship between memorization and information-retrieval behaviors. We are planning to analyze other factors such as eyesight, computer familiarization, display size, and refresh rate. Once these relationships are clarified, we should be able to create an algorithm for the developed interface that adjusts its operations to the user's retrieval behaviors and/or predicts the user's intention on the basis of how they operate the interface. Moreover, as we have described here, the features of the target content affect the retrieval behaviors. We plan to investigate the relationship between not only photos but also other contents such as text or sound, and these behaviors. We also plan to develop an algorithm for our developed user interface that adapts its operations to the user's behaviors.

Another challenge is to identify the factors that generate a pleasant feeling when operating familiar physical objects other than a book, which we featured in this thesis. Not only by using a familiar object as a metaphor for reducing the psychological load of operation but also by identifying the core factors of the feeling and applying them to a browsing interface, we can develop an interface that a person can use to access information comfortably. We plan to establish a basis for designing a user interface that makes use of the factors that generate a pleasant feeling when a physical object is operated even if there is no fresh outer design or novel functions. We think that such a "sensibility-oriented" interface would open up new horizons in the computer-human interaction field, even if the interaction appears normal and casual at first glance.

5.4 Concluding Remarks

In spite of the great efforts made to improve user interfaces and thereby facilitate access to the digital world, removing the obstacles between the real and digital world completely is still a challenge. In this thesis, we emphasized the need to focus on understanding the fundamental mechanisms of information retrieval and of how pleasant feelings are generated when we handle physical familiar objects so as to enable more natural and free access to digital contents. Our ongoing challenge is to develop user interfaces based on the scientific and quantitative analysis of the relationship between a person's surroundings and how his/her behaviors are stimulated by them to realize natural interaction with digital contents.

Bibliography

- Adler, A., Gujar, A., Harrison, B.L., O'Hara, K., and Sellen, A. A Diary Study of Work-Related Reading: Design Implications for Digital Reading Devices, *Proceed*ings of CHI '98, pp. 241-248 (1998).
- [2] Adobe Digital Editions, http://www.adobe.com/products/digitaleditions/.
- [3] Agarawala, A., and Balakrishnan, R. Keepin' It Real: Pushing the Desktop Metaphor with Physics, Piles and the Pen, *Proceedings of CHI '06*, pp. 1283-1292 (2006).
- [4] Ahlberg, C., and Shneiderman, B. The Alphaslider: A Compact and Rapid Selector, Proceedings of CHI '94, pp. 365-371 (1994).
- [5] Amazon.com, Inc., *Kindle*, http://www.amazon.com/.
- [6] Appert, C., and Fekete, J-D. OrthoZoom Scroller: 1D Multi-Scale Navigation, Proceedings of CHI '06, pp. 21-30 (2006).
- [7] Apple iTunes, http://www.apple.com/itunes/overview/.
- [8] Arai, K., Yokoyama, T., and Matsushita, Y. A Window System with Leafing Through Mode: BookWindow, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 291-292 (1992).
- [9] Arai, T., Aust, D., and Hudson, S.E. PaperLink: A Technique for Hyperlinking from Real Paper to Electronic Content, *Proceedings of CHI '97*, pp. 327-334 (1997).
- [10] ARToolKit, http://www.hitl.washington.edu/artoolkit/.
- [11] Asano, T., Sharlin, E., Kitamura, Y., Takashima, K., and Kishino F. Predictive Interaction Using the Delphian Desktop, *Proceedings of UIST '05*, pp. 133-141 (2005).
- [12] Baba, A., Nishimoto, Y., Ishikawa, K., Nakamura, H., Yoshimura, T., and Kurioka, T. A Study on Metadata Technology for Broadcasting System Based on Home Servers, *IEICE Technical Report*, Vol.104, No.279, pp. 11-16 (2004, in Japanese).

- [13] Back, M., Cohen, J., Gold, R., Harrison, S., and Minneman, S. Listen Reader: An Electronically Augmented Paper-Based Book, *Proceedings of CHI '01*, pp. 23-29 (2001).
- [14] Back, M.J., and Cohen, J. Page Detection Using Embedded Tags, Proceedings of UIST '00, pp. 159-160 (2000).
- [15] Barrett, B.H. A Digital Library Architecture for Interactive Television, IEEE International Conference on Systems, Man, and Cybernetics, 'Computational Cybernetics and Simulation', Vol.3, pp. 2380-2385 (1997).
- [16] Baudisch, P., Sinclair, M., and Wilson, A. Soap: A Pointing Device that Works in Mid-Air, *Proceedings of UIST '06*, pp. 43-46 (2006).
- [17] Bell, G., and Gemmell, J. A Call for the Home Media Network, Communications of the ACM, Vol.45, Issue 7, pp. 71-75 (2002).
- [18] Billinghurst, M., Kato, H., and Poupyrev I. MagicBook: Transitioning Between Reality and Virtuality, CHI '01 extended abstracts on Human factors in computing systems, pp. 25-26 (2001).
- [19] Bing, J., Dubreuil, J., Espanol, J., Julia, L., Lee, M., Loyer, M., and Serghine, M. MiTV: Rethinking Interactive TV, *Proceedings of International Conference on Virtual Systems and Multimedia*, pp. 365-369 (2001).
- [20] Blackmon, M.H., Polson, P.G., Kitajima, M., and Lewis, C. Cognitive Walkthrough for the Web, *Proceedings of CHI '02*, pp. 463-470 (2002).
- [21] Boreczky, J., Girgensohn, A., Golovchinsky, G., and Uchihashi S. An Interactive Comic Book Presentation for Exploring Video, *Proceedings of CHI '00*, pp. 185-192 (2000).
- [22] British Library: Turning the Pages, http://www.bl.uk/collections/treasures/ about.html.
- [23] Brown, B., and Barkhuus, L. The Television Will be Revolutionized: Effects of PVRs and Filesharing on Television Watching, *Proceedings of CHI '06*, pp. 663-666 (2006).
- [24] Card, S.K., Hong, L., Mackinlay, J.D., and Chi, E.H. 3Book: A Scalable 3D Virtual Book, *Proceedings of CHI '04*, pp. 1095-1098 (2004).

- [25] Card, S.K., Pirolli, P., Wege, M.V.D., Morrison, J.B., Reeder, R.W., Schraedley, P.K., and Boshart, J. Information Scent as a Driver of Web Behavior Graphs: Results of a Protocol Analysis Method for Web Usability, *Proceedings of CHI '01*, pp. 498-505 (2001).
- [26] Card, S.K., Robertson, G.G., and York, W. The WebBook and the Web Forager: An Information Workspace for the World-Wide Web, *Proceedings of CHI '96*, pp. 111-117 (1996).
- [27] Carmichael, A., Petrie, H., Hamilton, F., and Freeman, J. The Vista Project: Broadening Access to Digital TV Electronic Programme Guides, *PsychNology, Journal*, Vol.1, No.3, pp. 229-241 (2003).
- [28] Cesar, P., Chorianopoulos, K., and Jensen, J.F. Social Television and User Interaction, *Computers in Entertainment (CIE)*, Vol.6, Issue 1, Article No.4 (2008).
- [29] Chand, A., and Dey, A.K. Jadoo: A Paper User Interface for Users Unfamiliar with Computers, CHI '06 extended abstracts on Human factors in computing systems, pp. 1625-1630 (2006).
- [30] Chen, N., Guimbretiere, F., Dixon, M., Lewis, C., and Agrawala, M. Navigation Techniques for Dual-Display E-Book Readers, *Proceedings of CHI '08*, pp. 1779-1788 (2008).
- [31] Chi, E.H., Pirolli, P., Chen, K., and Pitkow, J. Using Information Scent to Model User Information Needs and Actions and the Web, *Proceedings of CHI '01*, pp. 490-497 (2001).
- [32] Chi, E.H., Rosien, A., Supattanasiri, G., Williams, A., Royer, C., Chow, C., Robles, E., Dalal, B., Chen, J., and Cousins, S. The Bloodhound Project: Automating Discovery of Web Usability Issues Using the InfoScentTM Simulator, *Proceedings of CHI '03*, pp. 505-512 (2003).
- [33] Chorianopoulos, K., Lekakos, G., and Spinellis, D. Intelligent User Interfaces in the Living Room: Usability Design for Personalized Television Applications, *Proceedings* of IUI '03, pp. 230-232 (2003).
- [34] Christel, M.G., Hauptmann, A.G., Wactlar, H.D., and Ng. T.D. Collages as Dynamic Summaries for News Video, *Proceedings of MULTIMEDIA '02*, pp. 561-569 (2002).
- [35] Christel, M.G., Smith, M.A., Taylor, C.R., and Winkler, D.B. Evolving Video Skims into Useful Multimedia Abstractions, *Proceedings of CHI '98*, pp. 171-178 (1998).

- [36] Chu, Y-C., Bainbridge, D., Jones, M., and Witten, I.H. Realistic Books: A Bizarre Homage to an Obsolete Medium?, *Proceedings of JCDL '04*, pp. 78-86 (2004).
- [37] Chu, Y-C., Witten, I.H., Lobb, R., and Bainbridge, D. How to Turn the Page, Proceedings of JCDL '03, pp. 186-188 (2003).
- [38] Davidoff, S., Bloomberg, C., Li, I.A.R., Mankoff, J., and Fussell, S.R. The Book as User Interface: Lowering the Entry Cost to Email for Elders, *CHI '05 extended abstracts on Human factors in computing systems*, pp. 1331-1334 (2005).
- [39] DLNA, http://www.dlna.org/home.
- [40] Drucker, S.M., Glatzer, A., Mar, S.D., and Wong, C. SmartSkip: Consumer Level Browsing and Skipping of Digital Video Content, *Proceedings of CHI '02*, pp. 219-226 (2002).
- [41] E-INK Corp., http://www.eink.com/.
- [42] Eronen, L., and Vuorimaa, P. User Interfaces for Digital Television: A Navigator Case Study, *Proceedings of AVI '00*, pp. 276-279 (2000).
- [43] Ferretti, S., Roccetti, M., and Andrich, J. Living the TV Revolution: Unite MHP to the Web or Face IDTV Irrelevance!, *Proceedings of WWW '06*, pp. 899-900 (2006).
- [44] Ferscha, A., Vogl, S, Emsenhuber, B., and Wally, B. Physical Shortcuts for Media Remote Controls, Proceedings of the 2nd international conference on INtelligent TEchnologies for interactive entertainment (INTETAIN), Article No.9 (2008).
- [45] Fishkin, K.P., Gujar, A., Harrison, B.L., Moran, T.P., and Want, R. Embodied User Interfaces for Really Direct Manipulation, *Communication of the ACM*, Vol.43, No.9, pp. 74-80 (2000).
- [46] Flicker, http://www.flickr.com/.
- [47] FUJITSU FRONTECH Ltd., FLEPia, http://www.frontech.fujitsu.com/ services/products/paper/flepia.
- [48] Gibson, J.J. The Theory of Affordances, R. Shaw & J. Bransford (Eds.), Perceiving, Acting, and Knowing: Toward an Ecological Psychology, pp. 67-82 (1977).
- [49] Google, http://www.google.co.jp/.
- [50] Goto, J., Kim, Y-B., Miyazaki, M., Komine, K., and Uratani, N. A Spoken Dialogue Interface for TV Operations Based on Data Collected by Using WOZ Method, *Proceedings of ACL '03*, pp. 101-104 (2003).

- [51] Graham, A., Garcia-Molina, H., Paepcke, A., and Winograd, T. Time as Essence for Photo Browsing Through Personal Digital Libraries, *Proceedings of JCDL '02*, pp. 326-335 (2002).
- [52] Grasset, R., Duenser, A., Seichter, H., and Billinghurst, M. The Mixed Reality Book: A New Multimedia Reading Experience, *CHI '07 extended abstracts on Human factors in computing systems*, pp. 1953-1958 (2007).
- [53] Harboe, G., Metcalf, C.J., Bentley, F., Tullio, J., Massey, N., and Romano, G. Ambient Social TV: Drawing People into a Shared Experience, *Proceedings of CHI* '08, pp. 1-10 (2008).
- [54] Harrison, B.L., Fishkin, K.P., Gujar, A., Mochon, C., and Want, R. Squeeze Me, Hold Me, Tilt Me! An Exploration of Manipulative User Interfaces, *Proceedings of CHI* '98, pp. 17-24 (1998).
- [55] Hess, J., Kustermann, G., and Pipek, V. pRemote: A User Customizable Remote Control, CHI '08 extended abstracts on Human factors in computing systems, pp. 3279-3284 (2008).
- [56] Hinckley, K., Cutrell, E., Bathiche, S., and Muss, T. Quantitative Analysis of Scrolling Techniques, *Proceedings of CHI '02*, pp. 65-72 (2002).
- [57] Holman, D., Vertegaal, R., Altosaar, M., Troje, N., and Johns, D. Paper Windows: Interaction Techniques for Digital Paper, *Proceedings of CHI '05*, pp. 591-599 (2005).
- [58] Hong, L., Chi, E.H., and Card, S.K. Annotating 3D Electronic Books, CHI '05 extended abstracts on Human factors in computing systems, pp. 1463-1466 (2005).
- [59] Hua, W., Han, M., and Gong, Y. Baseball Scene Classification Using Multimedia Features, *Proceedings of IEEE International Conference on Multimedia and Expo* (*ICME*), Vol.1, pp. 821-824 (2002).
- [60] Huang, T-C., Tseng, H-Y., Kung, C-P., and Cheng, K-T. Reliability Analysis for Flexible Electronics: Case Study of Integrated a-Si:H TFT Scan Driver, *Proceedings* of DAC '07, pp. 966-969 (2007).
- [61] Huet, B., Jiten, J., and Merialdo, B. Personalization of Hyperlinked Video in Interactive Television, *Proceedings of IEEE International Conference on Multimedia and Expo (ICME)*, pp. 462-465 (2005).
- [62] Ichimura, S., and Matsushita, Y. Another Dimension to Hypermedia Access, Proceedings of HYPERTEXT '93, pp. 63-72 (1993).

- [63] Igarashi, T., and Hinckley, K. Speed-Dependent Automatic Zooming for Browsing Large Documents, *Proceedings of UIST '00*, pp. 139-148 (2000).
- [64] Inagawa, N., and Fujinami, K. Making Reading Experience Rich with Augmented Book Cover and Bookmark, *Proceedings of APCHI '08*, pp. 157-166 (2008).
- [65] I-O DATA, http://www.iodata.jp/product/tv/.
- [66] Ishak, E.W., and Feiner, S.K. Content-Aware Scrolling, Proceedings of UIST '06, pp. 155-158 (2006).
- [67] Ishii, H., and Ullmer, B. Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms, *Proceedings of CHI '97*, pp. 234-241 (1997).
- [68] Ishii, H., Wisneski, C., Brave, S., Dahley, A., Gorbet, M., Ullmer, B., and Yarin, P. ambientROOM: Integrating Ambient Media with Architectural Space, CHI '98 conference summary on Human factors in computing systems, pp. 173-174 (1998).
- [69] Isobe, T., Fujiwara, M., Kaneta, H., Morita, T., and Uratani, N. Development of a TV Reception Navigation System Personalized with Viewing Habits, *IEEE Transactions on Consumer Electronics*, Vol.51, Issue 2, pp. 665-674 (2005).
- [70] Isobe, T., Fujiwara, M., Kaneta, H., Uratani, N., and Morita, T. Development and Features of a TV Navigation System, *IEEE Transactions on Consumer Electronics*, Vol.49, Issue 4, pp. 1035-1042 (2003).
- [71] Jameco Electronics, http://www.jameco.com/.
- [72] Karsvall, A. Personality Preferences in Graphical Interface Design, Proceedings of the second Nordic conference on Human-computer interaction (NordiCHI), pp. 217-218 (2002).
- [73] Kobayashi, M., and Igarashi, T. Considering the Direction of Cursor Movement for Efficient Traversal of Cascading Menus, *Proceedings of UIST '03*, pp. 91-94 (2003).
- [74] Kobayashi, M., and Koike, H. EnhancedDesk: Integrating Paper Documents and Digital Documents, *Proceedings of APCHI '98*, pp. 57-62 (1998).
- [75] Kohda, T., and Kawarasaki, M. A Study on Server Based Broadcasting Systems using P2P Technologies, *IEICE technical report*, *Information networks*, Vol.107, No.98, pp. 1-6 (2007, in Japanese).
- [76] Kohtake, N., Rekimoto, J., and Anzai, Y. InfoPoint: A Device that Provides a Uniform User Interface to Allow Appliances to Work Together over a Network, *Personal and Ubiquitous Computing*, Vol.5, Issue 4, pp. 264-274 (2001).

- [77] Koike, H., Sato, Y., Kobayashi, Y., Tobita, H., and Kobayashi, M. Interactive Textbook and Interactive Venn Diagram: Natural and Intuitive Interfaces on Augmented Desk System, *Proceedings of CHI '00*, pp. 121-128 (2000).
- [78] Kolekar, M.H., and Sengupta, S. Event-Importance Based Customized and Automatic Cricket Highlight Generation, *Proceedings of IEEE International Conference* on Multimedia and Expo (ICME), pp. 1617-1620 (2006).
- [79] Komine, K., Sawahata, Y., Uratani, N., Yoshida, Y., and Inoue, T. Evaluation of a Prototype Remote Control for Digital Broadcasting Receivers by Using Semantic Differential Method, *IEEE Transactions on Consumer Electronics*, Vol.53, Issue 2, pp. 561-568 (2007).
- [80] Kudo, M., Okada, K., and Matsushita, Y. Flexible User Interface for Accessing Information Spatially: BookWindow, *IPSJ SIG Notes* (HCI), vol.48, No.3, pp. 9-16 (1993, in Japanese).
- [81] Kullberg, R.L. Dynamic Timelines: Visualizing the History of Photography, CHI '96, Conference Companion, pp. 386-387, (1996).
- [82] Kumaki, D., Ando, S., Shimono, S., Yamashita, Y., Umeda, T., and Tokito, S. Significant Improvement of Electron Mobility in Organic Thin-Film Transistors Based on Thiazolothiazole Derivative by Employing Self-Assembled Monolayer, *Applied Physics Letters*, Vol.90, No.5, pp. 053506 (2007).
- [83] Kumar, M., and Winograd, T. Gaze-Enhanced Scrolling Techniques, Proceedings of UIST '07, pp. 213-216 (2007).
- [84] Lawrance, J., Bellamy, R., Burnett, M., and Rector, K. Using Information Scent to Model the Dynamic Foraging Behavior of Programmers in Maintenance Tasks, *Proceedings of CHI '08*, pp. 1323-1332 (2008).
- [85] Lee, J.C., Hudson, S.E., and Tse, E. Foldable Interactive Displays, Proceedings of UIST '08, pp. 287-290 (2008).
- [86] Li, B., and Sezan, I. Semantic Sports Video Analysis: Approaches and New Applications, *Proceedings of International Conference on Image Processing (ICIP)*, Vol.1, pp. I-17-20 (2003).
- [87] Li, F.C., Gupta, A., Sanocki, E., He, L-w., and Rui, Y. Browsing Digital Video, Proceedings of CHI '00, pp. 169-176 (2000).
- [88] Liao, C., Guimbretiere, F., and Hinckley, K. PapierCraft: A Command System for Interactive Paper, *Proceedings of UIST '05*, pp. 241-244 (2005).

- [89] Luff, P., Heath, C., Norrie, M, Signer, B., and Herdman, P. Only Touching the Surface: Creating Affinities Between Digital Content and Paper, *Proceedings of CSCW* '04, pp. 523-532 (2004).
- [90] Luyten, K., Thys, K., Huypens, S., and Coninx, K. Telebuddies on the Move: Social Stitching to Enhance the Networked Gaming Experience, *Proceedings of NetGames* '06, Article No.18 (2006).
- [91] Mackay, W.E., Pothier, G., Letondal, C., Boegh, K., and Sorensen, H.E. The Missing Link: Augmenting Biology Laboratory Notebooks, *Proceedings of UIST '02*, pp. 41-50 (2002).
- [92] Mackinlay, J.D., Robertson, G.G., and Card. S.K. The Perspective Wall: Detail and Context Smoothly Integrated, *Proceedings of CHI '91*, pp. 173-179 (1991).
- [93] Mackinlay, J.D., Robertson, G.G., and DeLine, R. Developing Calendar Visualizers for the Information Visualizer, *Proceedings of UIST '94*, pp. 109-118 (1994).
- [94] Mander, R., Salomon, G., and Wong, Y.Y. A 'Pile' Metaphor for Supporting Casual Organization of Information, *Proceedings of CHI '92*, pp. 627-634 (1992).
- [95] Marshall, C.C., and Bly, S. Turning the Page on Navigation, Proceedings of JCDL '05, pp. 225-234 (2005).
- [96] Masui, T., Kashiwagi, K., and Borden, G.R.IV. Elastic Graphical Interfaces to Precise Data Manipulation, *Proceedings of CHI '95*, pp. 143-144 (1995).
- [97] Masui, T., Tsukada, K., and Takabayashi, S. Information Navigation by Neighbor Hopping, Workshop on Interactive Systems and Software XI (2003, in Japanese).
- [98] Masui, T. LensBar Visualization for Browsing and Filtering Large Lists of Data, Proceedings of IEEE Symposium on Information Visualization, pp. 113-120 (1998).
- [99] Matsumoto, T., Dunnigan, T., and Back, M. Post-Bit: Embodied Video Contents on Tiny Stickies, *Proceedings of MULTIMEDIA* '05, pp. 263-264 (2005).
- [100] Matsushita Electric Industrial Co., Ltd., Σ Book, https://www.sigmabook.jp/.
- [101] May, K.W. Conductive Ink Based Page Detection for Linking Digital and Physical Pages, CHI '01 extended abstracts on Human factors in computing systems, pp. 275-276 (2001).
- [102] Meuleman, P., Heister, A., Kohar, H., and Tedd, D. Double Agent—Presentation and Filtering Agents for a Digital Television Recording System, CHI '98 conference summary on Human factors in computing systems, pp. 3-4 (1998).

- [103] Miyazawa, M., Kinoshita, K., Kobayashi, M., Yokoyama, T., and Matsushita, Y. An Electronic Book: APTBook, Proceedings of the IFIP TC13 Third International Conference on Human-Computer Interaction, pp. 513-519 (1990).
- [104] Motion Capture Systems from Vicon, http://www.vicon.com/.
- [105] My Page View, http://www.page-view.jp/english/index.php.
- [106] Mynatt, E.D., Igarashi, T., Edwards, W.K., and LaMarca, A. Flatland: New Dimensions in Office Whiteboards, *Proceedings of CHI '99*, pp. 346-353 (1999).
- [107] Nelson, L., Ichimura, S., Pedersen, E.R., and Adams, L. Palette: A Paper Interface for Giving Presentations, *Proceedings of CHI '99*, pp. 354-361 (1999).
- [108] NHK STRL Open House 2007, Blazing a Trial to the Future: Loudspeaker for Flexible Displays, http://www.nhk.or.jp/strl/open2007/en/tenji/t29.html.
- [109] Nichols, J., Myers, B.A., and Rothrock B. UNIFORM: Automatically Generating Consistent Remote Control User Interfaces, *Proceedings of CHI '06*, pp. 611-620 (2006).
- [110] Norman, D.A. The Psychology of Everyday Things, Basic Books, ISBN: 0465067093 (1988).
- [111] O'Hara, K., and Sellen, A. A Comparison of Reading Paper and On-Line Documents, Proceedings of CHI '97, pp. 335-342 (1997).
- [112] Okada, K., and Matsushita, Y. Beyond the Book Media: BookWindow, Transactions of Information Processing Society of Japan, Vol.35, No.3, pp. 468-477 (1994, in Japanese).
- [113] Peng, C., Lugmayr, A., and Vuorimaa, P. A Digital Television Navigator, Multimedia Tools and Applications, Vol.17, Issue 1, pp. 121-141 (2002).
- [114] Pirolli, P., and Crad, S.K. Information Foraging, Psychology Review 106, 4, pp. 643-675 (1999).
- [115] Pirolli, P. Computational Models of Information Scent-Following in a Very Large Browsable Text Collection, *Proceedings of CHI '97*, pp. 3-10 (1997).
- [116] Plastic Logic Ltd., http://www.plasticlogic.com/.
- [117] Poupyrev, I., Maruyama, S., and Rekimoto, J. Ambient Touch: Designing Tactile Interfaces for Handheld Devices, *Proceedings of UIST '02*, pp. 51-60 (2002).

- [118] Rekimoto, J. Tilting Operations for Small Screen Interfaces, Proceedings of UIST '96, pp. 167-168 (1996).
- [119] Rekimoto, J. Time-Machine Computing: A Time-Centric Approach for the Information Environment, Proceedings of UIST '99, pp. 45-54 (1999).
- [120] Rice, M., and Alm, N. Designing New Interfaces for Digital Interactive Television Usable by Older Adults, *Computers in Entertainment (CIE)*, Vol.6, Issue 1, Article No.6 (2008).
- [121] Sanguinetti, A., Haga, H., Funakoshi, A., Yoshida, A., and Matsumoto C. FReCon: A Fluid Remote Controller for a FReely Connected World in a Ubiquitous Environment, *Personal and Ubiquitous Computing*, Vol.7, Issue 3, pp. 163-168 (2003).
- [122] Sato, H., Fujikake, H., Murashige, T., Kikuchi, H., Kurita, T., and Sato, F. A4-sized Flexible Ferroelectric Liquid Crystal Displays with Micro Color Filters, *Journal of Society for Information Display*, vol.13, no.6, pp. 461-468 (2005).
- [123] Schilit, B.N., Golovchinsky, G., and Price, M.N. Beyond Paper: Supporting Active Reading with Free Form Digital Ink Annotations, *Proceedings of CHI '98*, pp. 249-256 (1998).
- [124] Schwesig, C., Poupyrev, I., and Mori, E. Gummi: A Bendable Computer, Proceedings of CHI '04, pp. 263-270 (2004).
- [125] SEIKO EPSON Corp., http://www.epson.jp/device/semicon/product/lcd_ controllers/.
- [126] Sellen, A.J., and Harper, R.H.R. The Myth of the Paperless Office, *MIT Press*, ISBN:026269283X (2003).
- [127] Signer, B., and Norrie, M.C. PaperPoint: A Paper-Based Presentation and Interactive Paper Prototyping Tool, *Proceedings of TEI '07*, pp. 57-64 (2007).
- [128] Siio, I., and Mima, Y. IconSticker: Information Management Using Paper Icons, Japan Society for Software Science and Technology, Vol.16, No.6, pp. 24-32 (1999, in Japanese).
- [129] Siio, I., and Tsujita, H. Mobile Interaction Using Paperweight Metaphor, Proceedings of UIST '06, pp. 111-114 (2006).
- [130] Smyth, B., and Cotter, P. A Personalized Television Listings Service, Communications of the ACM, Vol.43, Issue 8, pp. 107-111 (2000).
- [131] Sony Corp., LIBRIe, http://www.sony.jp/products/Consumer/LIBRIE/.

- [132] Sun, L, and Guimbretiere, F. Flipper: A New Method of Digital Document Navigation, CHI '05 extended abstracts on Human factors in computing systems, pp. 2001-2004 (2005).
- [133] Takahashi, Y., Nitta, N., and Babaguchi, N. Video Summarization for Large Sports Video Archives, *IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1170-1173 (2005).
- [134] Tan, G., Brave, S., Nass, C., and Takechi, M. Effects of Voice vs. Remote on U.S. and Japanese User Satisfaction with Interactive HDTV Systems, CHI '03 extended abstracts on Human factors in computing systems, pp. 714-715 (2003).
- [135] Terrenghi, L., Kirk, D., Sellen, A., and Izadi, S. Affordances for Manipulation of Physical versus Digital Media on Interactive Surfaces, *Proceedings of CHI '07*, pp. 1157-1166 (2007).
- [136] TiVo, http://www.tivo.com/.
- [137] Tse, T., Marchionini, G., Ding, W., Slaughter, L., and Komlodi, A. Dynamic Key Frame Presentation Techniques for Augmenting Video Browsing, *Proceedings of AVI* '98, pp. 185-194 (1998).
- [138] Utsumi, O., Miura, K., Ide, I., Sakai, S., and Tanaka, H. An Object Detection Method for Describing Soccer Games from Video, *Proceedings of IEEE International Conference on Multimedia and Expo (ICME)*, Vol.1, pp. 45-48 (2002).
- [139] van Dam, A. Post-WIMP User Interfaces, Communications of the ACM, Vol.40, No.2, pp. 63-67 (1997).
- [140] Weiser, M. Computer for the 21st Century, Scientific American, 265 (3), pp. 94-104 (1991).
- [141] Whittaker, S., and Hirschberg, J. The Character, Value, and Management of Personal Paper Archives, ACM Transactions on Computer-Human Interaction (TOCHI), Vol.8, Issue 2, pp. 150-170 (2001).
- [142] Wittenburg, K., Forlines, C., Lanning, T., Esenther, A., Harada, S., and Miyachi, T. Rapid Serial Visual Presentation Techniques for Consumer Digital Video Devices, *Proceedings of UIST '03*, pp. 115-124 (2003).
- [143] Wittenburg, K., Lanning, T., Schwenke, D., Shubin, H., and Vetro, A. The Prospects for Unrestricted Speech Input for TV Content Search, *Proceedings of AVI '06*, pp. 352-359 (2006).

- [144] Wood-Bradley, G., and Blashki, K. Interactive Television and the Role of Affective Usability, *Proceedings of International Conference on Multi-Media Modeling*, pp. 485-489 (2006).
- [145] Wu, C-S(A)., Robinson, S.J., and Mazalek, A. Turning a Page on the Digital Annotation of Physical Books, *Proceedings of TEI '08*, pp. 109-116 (2008).
- [146] Xiong, Z., Radhakrishnan, R., and Divakaran, A. Generation of Sports Highlights Using Motion Activity in Combination with a Common Audio Feature Extraction Framework, *Proceedings of International Conference on Image Processing (ICIP)*, Vol.1, pp. I-5-8 (2003).
- [147] Yagawa, Y., Iwai, N., Yanagi, K., Kojima, K. and Matsumoto, K. The Digital Album: A Personal Filetainment System, *Proceedings of MULTIMEDIA*, *IEEE Computer Society Press*, pp. 433-439 (1996).
- [148] YouTube, http://www.youtube.com/.
- [149] Zimmerman, J., and Kurapati, K. Exposing Profiles to Build Trust in a Recommender, CHI '02 extended abstracts on Human factors in computing systems, pp. 608-609 (2002).
- [150] Zimmerman, J., Dimitrova, N., Agnihotri, L., Janevski, A., and Nikolovska, L. MyInfo: A Personal News Interface, CHI '03 extended abstracts on Human factors in computing systems, pp. 898-899 (2003).