

The Digital Album: A Personal File-tainment System

Yuichi Yagawa, Noriyuki Iwai, Kunihiro Yanagi, Keiji Kojima
Systems Development Laboratory, Hitachi, Ltd.
1099 Ohzenji, Asao, Kawasaki, Kanagawa 215 Japan
{yagawa, iwai, k-yanagi, k-kojima}@sdl.hitachi.co.jp

and Kiyoshi Matsumoto
Data Storage & Retrieval Systems Division, Hitachi, Ltd.
(Presently with Graduate School of Engineering, The University of Tokyo)

Abstract

We describe the Digital Album: a personal File-tainment system that allows users to enjoy multimedia data as well as file and retrieve them. This system incorporates an album metaphor which imitates a photo album and features pseudo-3D page-turning animation. Stories described in this album metaphor entertain users. We also adopt a multiview classification scheme in which each data item is registered to any nodes in multiple classification hierarchies. The system is suitable for filing and retrieving multimedia data visually and bibliographically. Moreover, users can enjoy creating their own albums based on the classification easily and dynamically. The album metaphor is used as a PhotoCD user interface in Hitachi's multimedia products.

1. Introduction

Recently, computers are not just thought of as a business or scientific tool but are being considered sources of entertainment, and the main interest of users has been shifting from efficiency to enjoyment. Multimedia systems are expected to satisfy these new requirements. For example, some multimedia titles published on CD-ROM specifically aim at entertainment and, as a result, they can delight and impress users. We enjoyed "Just Grandma and Me"[1] very much, and also "I Photograph to Remember"[2] is one of the most impressive multimedia titles we have seen.

At the same time, the need for personal multimedia database systems is gradually increasing because users can make personal use of much more multimedia data than ever before. One factor in this trend is the progress in hardware technologies, such as input, output, and storage devices; another is the standardization of multimedia data formats, such as JPEG (Joint Photographic Experts Group) and MPEG (Motion Picture Experts Group). Such standardization allows even novice users to use computers to handle digital photographs and video. Moreover, CD-ROM has come into wide use as a distribution method, and DVD (Digital

Video Disc) or the NII (National Information Infrastructure) may be available in the near future, making it much easier for users to access multimedia data.

However, most of the current multimedia database systems lack entertainment value, even though many users would like to have fun while using multimedia databases. For example, today's multimedia database products, which generally adopt a user interface that shows small thumbnail images in a window or that show each data item with detailed information on a card, are not suitable for describing stories or scenarios in a way that creates a feeling of pleasure. For personal use, a multimedia database system should be a balance of easy filing, efficient retrieval, and fun.

In this paper, we focus on a personal *File-tainment* system that lets users not only file and retrieve, but also enjoy multimedia data, such as pictures, video scenes, and so on. Because a real photo-album is a medium in which we can file, browse, and enjoy photographs in the real world, we call this system the Digital Album. Our approach is to imitate both the external appearance and the hidden structure of the real album as much as possible. In Sec. 2, we show the conceptual model of the Digital Album. In Sec. 3, we explain how the Digital Album imitates the external appearance of the real album (later, we define this as the album metaphor), and in Sec. 4, we describe the database design as the hidden structure of the Digital Album. Finally in Sec. 5, we discuss the user evaluation of the Digital Album and show other applications of the *File-tainment* system.

2. Conceptual Model of The Digital Album

One factor that increases the entertainment value of multimedia titles is a story (We use this word to mean the same as a scenario in this paper). For example, in "I Photograph to Remember", photographs are shown in a pre-defined sequence, which organizes the data into a meaningful overall story. As in "Just Grandma and Me", the more complicated the story is, the

more user interaction is possible.

Before we look at the Digital Album, we would like to discuss how a story is organized in a real album. Generally speaking, an album might be considered to consist of only visible elements, such as photographs and comments that accompany them. However, especially in a well-organized album, meaningful relationships between the photographs exist apart from the external appearance and people can easily understand these relationships.

For example, when people place photographs in an album, they arrange related photographs on the same pages. Even though most people dislike troublesome filing, they put photographs into an album at least in the order they were taken and this time sequence often produces a meaningful story. Also, when they look for photographs, they usually expect related photographs to be close to each other. When looking through the album, they enjoy not only individual photographs but also the overall feeling that is often due to the photographs' layout.

These examples show that there are internal semantics hidden within the external appearance and these semantics organize the story and the layout which we see in the album. The examples also indicate that the real photo album satisfies the definition of the *file-tainment* system explained in Sec. 1.

With this in mind, we designed the conceptual model of the Digital Album to consist of two parts (Fig. 1): one is the external appearance, which we call the album metaphor, and the other is the internal semantics, which we call the multiview classification[3]. These two parts are also related to each other. Here, we briefly explain the two parts.

The album metaphor is a user interface that imitates the look and feel of a real photo album in order to organize a story in the Digital Album and to entertain users. Moreover, the album metaphor is effective for visual filing and retrieving multimedia data. How this is achieved is discussed in Sec. 3.

The multiview classification scheme was first introduced in Hitachi's document management software, "Bibliotheca/IS (InfoShare)"[3], as a data model based on orthogonal classification hierarchies. In this model, each hierarchy represents a classification from one viewpoint and each data item can be registered into any nodes in multiple classification hierarchies. This means that users can classify or retrieve data from multiple viewpoints.

We adopted this technology to represent the internal semantics in the Digital Album because, generally speaking, filing is based on classification. Even in a real photo album, people classify and arrange photographs. Also, multimedia data contain detailed information not seen by users. For example, in the case of photographs, there is information about subjects, photographers, places where the photos were taken, and so on. So, this detailed information clearly has multiple categories. Moreover, once

multimedia data are classified into the multiview classification, users can enjoy creating their own albums and stories and going through them at any time (We call this feature as dynamic album creation). Users select nodes in the hierarchies and the appropriate multimedia data are retrieved. The data are arranged on the pages by ranking classification hierarchies. How this is achieved is explained in Sec. 4.

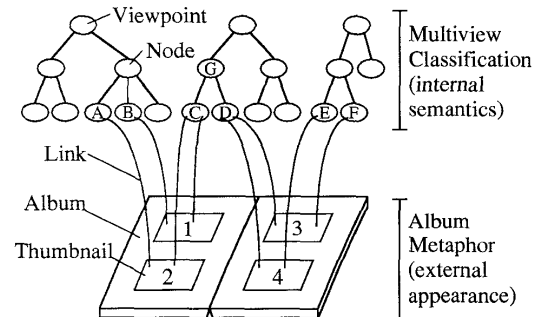


Fig. 1 Conceptual Model of Digital Album

Note that the internal semantics are described as classification hierarchies and that each data item may be connected to several nodes in different hierarchies.

3. Album Metaphor

As we indicated before, the album metaphor entertains users and helps users file and retrieve multimedia data visually. In this section, we explain how this is achieved.

3.1 Interface Design

The interface design of the album metaphor is quite simple and similar to a real album as Fig. 2 shows. The album interface consists of two pages spread open, page titles, two areas indicating the album's thickness (which we call the "Thumb Button"), no more than two thumbnail images on each page, thumbnail titles, and page numbers. Also, some pages contain Book Marks on themselves. The following paragraphs describe the characteristics of this user interface.

We designed the album interface to be attractive so users will enjoy seeing it; for example, it has well-proportioned spaces and a favorable background color on the pages. We also defined simple layout rules according to the design: the system resizes thumbnail images to fit the pre-defined spaces, puts them on pages in order, and arranges them on a page according to the balance of each thumbnail's size. For example, if two thumbnails on a page are vertically long, they are horizontally shifted, and elsewhere, they are vertically aligned, as Fig. 2 shows. This means that users are released from the troublesome task that they arrange thumbnails on pages for themselves.

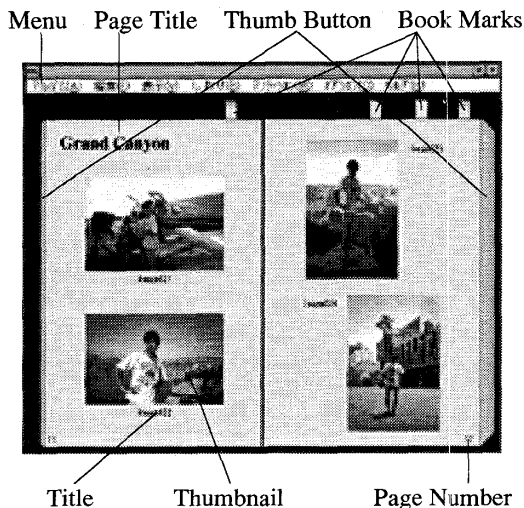


Fig. 2 Interface Design of Album Metaphor

Album design consists of two pages and four thumbnail images. The page size is 320x480 pixels and the thumbnail size is around 180x120 pixels, which is large enough for users to clearly see what the icon represents. Each page represents a meaningful scene because the thumbnails on the same page are related to each other.

Although each thumbnail image is small, it is large enough so that its contents can be enjoyed. Here, a thumbnail image is a reduced image of an original picture, a clipped image of a video scene, or a related image to a sound, such as a CD jacket. While having more thumbnail images on a page increases browsing efficiency, there is a trade-off between efficiency and image clarity. We considered clarity more important than efficiency, so we set the page size to 320x480 pixels and the thumbnail size to around 180x120 pixels, which is sufficient for users to clearly recognize the image content. Of course, the higher the display resolution will become in the future, the more thumbnail images can be displayed.

Also, since the thumbnails are encapsulated as objects with detailed information such as original data and attributes, users can easily access this information through the album interface. In the current version of the Digital Album, users see original data by clicking a thumbnail image, and users access to attributes by clicking a thumbnail's title.

The thumbnails on the two facing pages are often related, so the overall impression is of one meaningful scene. For example, a user may include pictures and video scenes from the same occasion on the same facing pages. A few years later, when the user views them again, the whole atmosphere of the facing pages may bring to mind that occasion and create powerful memories for the user.

A set of pages may also represent one meaningful scene.

Together, these pages make up a story. In this case, it is very important that users can feel the continuity of the story by turning pages. Also, the page-turning action should be as realistic as the design of album metaphor is. To represent continuity and realism, we implemented page-turning animation[4] and its operation methods, discussed in detail below.

3.2 Page-turning Animation

Visual browsing with page-turning animation is both enjoyable and effective for retrieving multimedia data. As we mentioned, the thumbnail layout in the album interface corresponds to relationships among the data; therefore, there is a kind of a story created in the album. Following the story by turning pages helps users find the data. Moreover, a previous study indicates that page-turning animation is useful during retrieval because people are good at memorizing objects spatially and page-turning animation stimulates this spatial memory[5].

The idea of page-turning animation was introduced in the "Media Room"[4], in which the system used the realistic animation to turn pages. Also, this idea was implemented in several systems such as the desktop environment[6], the user interface of the Digital Library[7], and so on. However, these systems required graphic accelerators to represent reality, and, as a result, the implementation costs were high as a personal system. Although, some other systems, such as "Book Window"[5], implemented page-turning animation as software only, the reality of the animation was not so favorable.

There is a trade-off between an animation's realism and its implementation cost. We implemented the Digital Album as personal computer software. To achieve rich realism while avoiding heavy transactions, we used an optical illusion to good account.

Principle

Fig. 3 illustrates the principle of our page-turning animation by showing a book from above. The system creates a page-turning illusion by mapping the pages onto a projection surface while the pages are being turned over. In this figure, θ is the page angle (Rad, $0 \leq \theta \leq \pi$) and α is the turning speed (Rad/sec), which users can change at any time.

The geometric definitions of the mapping functions are:

$$X = k \times x \times \cos(\theta) + (k - x \times \sin(\theta)) \quad (1)$$

$$Y = k \times y + (k - x \times \sin(\theta)) \quad (2)$$

where k is the distance between the light source and the projection surface, and (x, y) and (X, Y) are the coordinates of each pixel on a page in the case of being still and turning, respectively. For every

frame in the animation, these functions are calculated for all pixels on the turning page and the mapped image is also compounded with the images of still pages.

Although this page-turning illusion is quite realistic, it is very difficult for a personal computer to execute the projection transaction in real time. Even if the system has a pre-defined mapping table, the system wastes as much memory as all pixels of a page upon only the user interface. Therefore, we invented a new method that is executable on a personal computer in real time but still offers reasonable realism.

Assuming that the light source is put on the vanishing point in Fig. 3, we have the following formulas approximated to the formulas (1) and (2) respectively:

$$X = x \times \cos(\theta) \quad (3)$$

$$Y = y \quad (4)$$

because $k = k - x \times \sin(\theta)$ is given. This method uses orthographic projection and cuts the cost of vertical projection. Moreover, horizontal reduction is substituted for horizontal projection.

This technique reduces the processing burden, but the animation realism is not adequate, mainly because users seem to feel some incongruity since there is neither change nor movement in the vertical direction. Therefore, we modified the method so that the thumbnail images would move in the vertical direction to simulate projection in the vertical direction.

Algorithm

When the system detects a user's command to turn over the pages, it executes the following algorithm.

STEP 1: Set the turning speed and direction according to user's operation which is explained below.

STEP 2: Load thumbnails for the next or the previous page, depending on the turning direction.

STEP 3: Calculate the number of frames in one page-turning, which is in inverse proportion to the turning speed, and execute the following for each frame.

(a) Reduce the width of the page being turned according to the expression $PW = pw \times \cos(\theta)$.

(b) Execute the following functions for each thumbnail on the page being turned.

1) Reduce the thumbnail width according to the expression $IW = iw \times \cos(\theta)$.

2) Move the thumbnail in the horizontal direction according to the expression $X = x \times \cos(\theta)$.

3) Move the thumbnail in the vertical direction according to the current frame number.

(c) Show the page and thumbnails which were hidden

behind the page being turned.

STEP 4: Set the width of the Thumb Buttons. This function is also explained below.

STEP 5: Continue the functions from STEP 1 to 4 unless a user's command to stop turning pages is detected.

Here, each original page width is denoted by pw , the original thumbnail width by iw , and the thumbnail location by x and y . Each width of the page being turned is denoted by PW , the width of the turning thumbnail by IW , and its location by X and Y .

We estimated that this method can cut more than 98% of the steps compared to the original method.

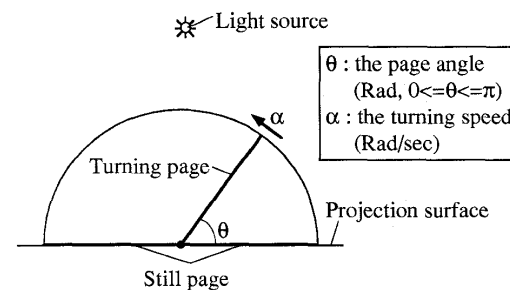


Fig. 3 Principle of Page-turning Animation

This shows a book above. The system creates a page-turning animation by projecting pages onto a surface as the page turns.

3.3 Operation Methods

Fig. 4 shows an operation method of the page-turning function. Users turn pages by clicking the Thumb Button on either side: clicking the right button turns to the next page and clicking the left button turns back to the previous page. The point at which the button is clicked determines the page-turning speed according to $\alpha = A \times n + N$. Here, A is the maximum speed, n is the distance from the clicking point to the page bottom, and N is the page height. If the cursor is moved up, the speed increases and if it is moved down, the speed decreases.

Also, users can jump to pre-defined pages by clicking Book Marks. In the current version of the system, one Book Mark corresponds to one page and the place of the Book Mark corresponds to the page number, so users can roughly guess which page they are going to jump. In Fig. 4, the horizontal points of the Book Marks are in proportion to the page numbers. About the way to put Book Marks, a user chooses pages and puts Book Marks, or the system automatically puts Book Marks to pages according to the classification when a user dynamically creates an album based on the multiview classification. Of course, the user can remove any Book Marks at any time. We also developed page jump operation by clicking a title in a table of contents. Here, the table of contents is shown in the menu or in another window.

Moreover, the width of the Thumb Buttons indicates the number of pages. The left Thumb Button corresponds to the current page number and the right Thumb Button corresponds to the number of remaining pages. When the user is viewing the first page, there is no left Thumb Button, except for the operation area. As the user turns pages forward, the left Thumb Button widens and the right Thumb Button narrows.

Album editing functions are also available from the menu. The functions are used to arrange thumbnail images on the pages, or, in other words, for visually filing multimedia data. Users can place, move, or delete thumbnails on any page or at any place they want by dragging and dropping the thumbnails. Cut-and-paste and copy-and-paste functions are also available. In this case, the system grids the thumbnails automatically to maintain the balance and beauty of the layout.

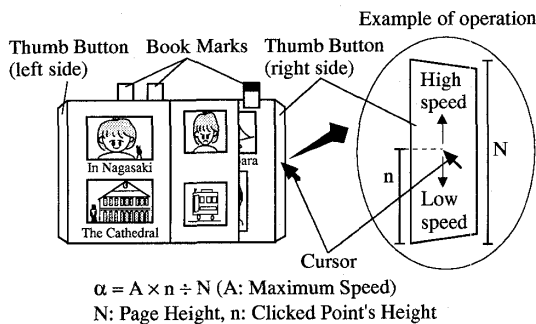


Fig. 4 Operation Method of Page-turning Function

4. Multiview Classification

In this section, we briefly explain how the multiview classification achieves easy filing and efficient retrieval, and we show the mechanism of the dynamic album creation which stimulates users' enjoyment.

4.1 Overview

The multiview classification is a data model that represents the internal semantics of the Digital Album. The main characteristics are as the following.

First, as Fig. 1 shows, the structure is based on orthogonal classification hierarchies and each hierarchy represents a classification from one viewpoint. Users can define and edit the hierarchies at any time.

Secondly, users can register a multimedia data item into any number of nodes in multiple hierarchies. For example, in Fig. 1, data 1 is connected to both node B and node C, each of which belongs to a different classification hierarchy. This means that users can classify multimedia data from multiple viewpoints.

Thirdly, in the hierarchies, each node has bi-directional links to multimedia data. When a node is selected, all multimedia data connected to the selected node as well as all nodes below the selected node are collected through an inheritance mechanism. Users can simultaneously select any number of nodes and the system will retrieve a subset of multimedia data collected from all selected nodes. In Fig. 1, when a user selects node A and node G, the result of the retrieval will be data 2. This search mechanism is also used for dynamic album creation explained in the following section.

Furthermore, we also developed the tool which visualize the multiview classification. This tool displays a list of viewpoints, a hierarchy of the selected viewpoint, and thumbnail images of multimedia data which the selected nodes contain together. Therefore, in order to retrieve multimedia data, users just select views and nodes on the display.

In the case of filing, thumbnail images of available data are displayed in another window, so users select appropriate thumbnail images, and drag and drop them to the selected nodes. Editing functions are also available for modifying views and hierarchies.

The Digital Album is based on an object-oriented database. Fig. 5 shows this object model. Classes are organized in hierarchical structure to implement the multiview classification. Each multimedia datum is managed as an object encapsulated with a title, a thumbnail image, an original datum and attributes. Also, an album consists of a title, a table of contents, and several pages, each of which also consists of a title and links to objects. When the album metaphor shows an album's contents, objects belong to each page are collected according to the page-turning. Moreover, we define the mapping rules from the database to the album metaphor, which enable users to create an album dynamically.

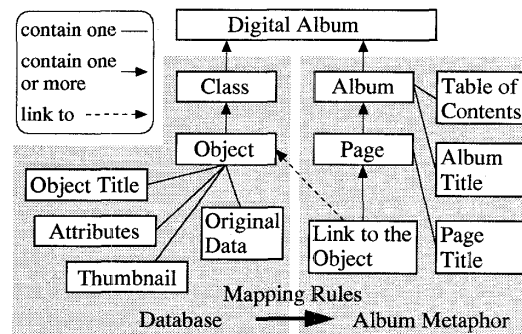


Fig. 5 Object Model of the Digital Album

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts



Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research



With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips



Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.