Metadata and Organizational Structures in Personal Photograph Digital Libraries

Sally Jo Cunningham and Masood Masoodian

Computer Science Department, University of Waikato, Hamilton, New Zealand
{sallyjo,masood}@cs.waikato.ac.nz

Abstract. We examine the 'native' metadata and organizational structures that individuals create for their personal photo digital libraries, by analyzing the behavior of photo collectors as recorded in 37 autoethnographies and ethnographies. The findings confirm several common assumptions about how people organize their photos that have been the basis for features in earlier photo digital libraries—that photos are commonly organized by time, event, and location, and that collection owners create very little metadata manually. We discuss alternate sources of metadata that arise as a consequence of sharing photos, and consider additional features for photo digital libraries that may be useful in supporting searching and browsing of personal collections.

Keywords: Personal digital libraries, photo metadata, user studies.

1 Introduction

Digital camera users take more photos than did film camera owners: digital cameras free users from the expense of developing film and creating prints, and offer instant feedback on each shot. Digital photos can be inexpensively stored as well—on CDROM, DVD, hard drives, and online archives. We can afford to create personal photo archives that document our lives, our hobbies, and indeed any passing interest.

Locating an existing photo in our collection and maintaining a sense of the collection's contents, however, may not be straightforward. Which PC did we store Uncle Bob's photo on—or was it on the hard drive that failed? How can we easily search our meagerly labeled photo directories to find my uni graduation shots? And this photo of a party: is this an event I attended, or did someone else send it to me?

A number of novel searching, browsing, and collection management features have been prototyped to address one or more of the above situations (Section 2). In this paper, we explore the ways that people currently organize their digital photo collections. Our work is based on a large-scale ethnographic investigation of personal photo management (37 participants, over 150 pages of observational summaries; Section 3). We examine analyze of the metadata and photo collection organization schemes created by participants, and suggest additional metadata that might be useful to these users (Section 4). We conclude (Section 5) by identifying photo management behavior that can suggest additional software support for personal photo collection digital libraries.

2   Previous Work

Earlier work on personal photo digital libraries has focused on using a commonsense understanding of how and when people take photos to inform the design of novel features specific to personal photo collections. For example, the insights that photos are taken in a time-linear order and that people tend to take several snapshots of the same event or person in a brief period of time, have led to experimental systems using photo timestamps as a basis for browsing structures [7], using time and image content to automatically cluster photos into 'events' [2], or using GPS location information to identify geographically proximate photos [11]. Since these are personal photos, the collection owner is expected to recognize the images—and so browsing, rather than searching, is anticipated to be the primary means of locating specific photos in the collection. A number of systems have been developed to enhance the layout and selection of representative thumbnails for browsing (eg, [7]). The insight that an individual tends to take multiple photos of the same people and places over time can be exploited by using time and location metadata in previous, annotated photos to develop metadata for new, un-captioned photos [9].

However, personal photo collections can rapidly become too large for browsing to be practicable as the main access method. Digital cameras enable their owners to take more photos than was feasible with film cameras, since the cost of film and processing are eliminated, and printing is optional. Content-based searching relies on image processing and pattern recognition to retrieve images based on similarity to a query specification (often in the form of a sample image); query features are usually limited to color, shape, texture, and spatial relations between image components [13]. These features rely on the user's ability to describe the contents of the desired image—and so content-based searching is most useful for locating the photos that could also be recognized while browsing.

Annotation tools based on manually created metadata are less commonly presented in the digital libraries research literature—possibly because of another assumption about personal photo-taking habits, that photos are rarely labeled by their owner. This assumption is confirmed by recent studies ([12], [4], [8]).

Public image collections usually rely on text-based rather than content-based searching, where the search metadata may be manually created or (more commonly) leveraged from text associated with the image (for example, a figure title, or document contents proximate to the image). Image management tools created for large scale, public photo libraries may be expected to be less relevant to the needs of the individual, managing a personal collection. The users of a public collection are relatively unfamiliar with its contents, whereas the photographer is the creator and curator of the individual collection—and so the user needs and requirements for the personal collection will be different from those of the general-purpose photo libraries.

3   Data Collection

Data for this study was gathered through a project assigned to undergraduate students in an upper level university human-computer interaction course. The students' goal

was to design and prototype a shared, online photo collection—essentially a digital library of personal photographs. The students based their designs on ethnographic investigations into the photo taking and sharing habits of themselves and at least one friend. These investigations were summarized, and these summaries are analyzed in this present paper. In total, the students conducted 18 personal ethnographies and 19 observations/interviews of another person, and the summaries come to over 150 pages (Table 1). To preserve anonymity, each participant is referred to with a randomly assigned letter of the alphabet.

The students first performed a 'personal ethnography' or autoethnography [3], examining their own photograph collections and photo-taking behavior. The students then performed a similar ethnographic observation and interview of a friend. In these investigations, the students describe their (or their informant's) photo collection's contents, when and under what circumstances they take new photos, how those photos are organized, metadata associated with the photos, and the ways in which photos are exchanged (e.g., by showing off photo albums, posting photos online, or sending photos via mobile phone). 'Organization' of a photo or collection was construed as broadly as possible, so as to capture as many aspects as possible for that behavior; in these investigations organization included creating physical and digital albums, storing photos on CD, tossing print photos in a drawer, and so forth. In this work, we focus on the behaviors associated with managing digital photo collections, rather than the now-historic management techniques for print photos.

An overview of photo organization behaviors, as described in the ethnographies and autoethnographies, is presented in Section 4. Grounded theory methods [14] were used to analyze the students' summaries of their interviews and observations. We analyze the summarized descriptions of photo taking, sharing, and storage/organization as reported in the ethnographies, rather than the students' own analyses and suggestions for photo digital library features for their projects.

Table 1. Gender and nationality of participants

	Male	**Female**	**National Origin**	**Count**
	22	15	NZ	17
			China	16
			Iran	2
			Korea	2

4 'Native' Metadata and Organizational Structures

In this section, we describe the types of metadata and organizational structures that are reported in the ethnographies.

Fig. 1. (a) Raw ingredients. (b) 'Steamed chicken on balsam pear'.

4.1   Group Labels and Individual Photo Labels

The most striking feature of the manually created photo metadata is that it is primarily applied to groups of photos, rather than to individual photos. People tend to take multiple shots, rather than individual snaps, and these groups of photos are downloaded and stored together. The folder that a group of photos are stored in is given a descriptive name (for example, "Wellington"). The filename of individual photos may be left at the default (eg, "Image12.JPG", "P80080003.JPG"), or the file may be renamed to be slightly more evocative (eg, "View 1", "View 2", etc. for photos of scenery taken on a trip; Participant I).

Many of the folder labels are 'events' in the sense used by Cooper et al [2]: a group of photos that are temporally clustered together (and that are temporally distant from other clusters), and that depict the same activity or location. A temporal cluster can be relatively tightly grouped (eg, "Wedding") or can include extended periods (eg, "Wellington trip"). Occasionally hierarchies of folders will be created (for example, to group together different days of a lengthy trip). Hierarchies very rarely are more than two folders deep.

Folders that are not event-related may simply be a group of photos that are downloaded from the camera at the same time (and so represent a number of un-labeled groups). These 'miscellaneous' folders might not be manually labeled, and may simply retain the default label of the photo management system. For these photos, the photo timestamps or the folder creation date may be the main access point.

Location-related labels may be records of a specific trip (and so represent an event, as well as a specific place), or may be (possibly temporally distant) images of a place that holds significance to the collection owner. Examples of location labels are "Home town" or "XXX University".

More rarely, a folder may be labeled with a theme that represents a hobby or special interest of the collection owner. These can be deeply idiosyncratic. Participant E's informant, for example, is a hobbyist cook; she takes photos of the raw ingredients and the final dish when she creates a meal "she could be proud of" (Figure 1a, b), with the cooking – related photos stored together across different dates.

Affect, or the emotional impact of a photo, also appears as a category for organizing photos. Only two emotion–related categories appear in the ethnographies:

sentimental (associated with close family members or romantically themed); and humorous or quirky. Sentimental photos are usually grouped with other photos associated with the individual (eg, "my mother"). Humorous photos may be part of a themed collection: for example, Participant K's informant "takes photos of 'engrish' – amusing mistakes in the use of the English language by foreign language users". Funny photos are also frequently reported as being shared with others by email or SMS, accompanied by a brief description of the comical aspect of the image ("hey look at this hilarious picture of my cat sleeping on my dog"; Participant M).

4.2 Time

Participants almost uniformly reported that they appreciated the association of a timestamp with a photo, and used the timestamps to browse along their personal timeline of activities. Without a timestamp, it can be difficult to distinguish between similar events (eg, is this a photo of last week's party, or last month's?). It can be annoying, however, to have the timestamp appear on the photo itself ("because it can make a really good picture look bad with the red/white date in the bottom of the screen"; Participant Q). The preference is for having the timestamp—and indeed, all metadata—viewable separately from the photo, so that the image is not spoiled.

A timestamp indicates when a photo was snapped. A second aspect of time is duration—how long it's been since that photo was taken. While raw timestamps are useful in sorting photos for browsing, an indication of duration may be useful in appreciating an individual photo. Participant G, for example, notes that for him the timestamp "serves dual functions, first of all, it tells me when I took this picture [of his mother] and secondly, the most of all, it also reminds me how long I have been away from home or my dear loving mother".

4.3 Verbosity of Metadata

Photo labels tend to be brief—a word or phrase, perhaps simply a few characters in a filename, rather than an exhaustive description of the people, places, events and so forth appearing in the image. Exceptions occur, with some photos given much more extensive metadata. Some photos may be selected to form a sort of visual, personal diary; these selected images are given more detailed captions ("I want to know exactly which Christmas that was, what club that concert was in, and so on"; Participant K). Photos that are intended to be shared are more likely to have manual descriptions, and these descriptions tend to be more detailed than those of photos retained for personal viewing only. The intention is to provide an explanation and interpretation of the photo for the viewer, ("if the photo shows people who were doing something and it isn't clear for the person who will see that photo, I add a brief explanation" (Participant L); "so that other people who happen to look at my photos know who the people in them are, or what we're doing, or where we are" (Participant Q)). At times this more detailed metadata can be critical: Participant K's informant, for example, posted on her website a photo in which it appears that several people are beating up one of their friends, and so "in the description, she explained that the fight was only a pretend one".

Photos shared via email or SMS usually have an accompanying message that explains the significance of the image to the photo taker. Even very brief messages can contain a surprising amount of contextual information. For example, the SMS message accompanying a photo of shelves of books ("This is one spot of the Waikato University's library! I read very hard there!" Participant T) includes the location (*Waikato University, library*) and the activity (*read*). Email messages tend to have much lengthier, and much richer, descriptions of attached photos.

While participants recognized that they infrequently recorded manual metadata on print photos, they also felt that they recorded longer captions on physical prints. Likely explanations for these phenomena are that captions on physical prints are usually created prior to archiving the print in a formal album, or to giving the print to someone else. The ease of writing on the back of prints—keeping the metadata bound to the photo, but not spoiling the image—may also be a factor.

Why is manually created metadata for digital photos usually so limited—or non-existent? One reason, it seems, is that photo takers tend to over-estimate their ability to remember the context surrounding their snapping a picture: a typical explanation for a lack of captioning is that "I did not note down when and where I took the pictures since I thought that it would not be necessary and I could remember" (Participant I). Several of the participants were shocked to realize while creating their autoethnographies that their memories, even of relatively recent events, were fallible:

> As I have just looked through my photo collection I am aware that I cannot remember what is happening in many of the photos. I also have no idea when most of the photos were taken, I can only guess—because of the chronological nature of the organisation of the photos and how old that people look in the photos. (Participant J)

Another impediment to the manual creation of metadata is the disjunct between creating a photo and adding it to the photo collection; a photo cannot be captioned until it is transferred from the digital camera to the photo storage system, and at that point the relatively large number of photos being transferred makes creating metadata seem a daunting task. The exception is the labeling of photos taken on a mobile; the user is generally taking only one or a few photos at a time, the mobile has a built-in keyboard or keypad for recording metadata, and the point of creating the photo is often to send it to someone else (with a brief message that can serve as a caption).

And, of course, it can simply be difficult to motivate oneself to put effort into organizing and providing metadata for a photo collection, when the payoff for this effort will be in increased ease locating particular photos in the distant future ("Also, I am too lazy to write down some information about those photos beside them [T]"). This behavior is hardly restricted to managing photo collections; most users have poorly maintained, badly organized, and cryptically labeled file storage structures on their personal computers [10]. An effective photo metadata scheme, then, would make it easy to create labels at the most convenient time for the user—perhaps at the time the photo is being taken—and would also be forgiving enough to automatically create metadata (for example, by image content analysis).

Fig. 2. (a) Photo taken in Guild Wars, 'as a tourist' at Perdition Rock (Participant K). (b) 'An action photo' of Participant K at Augery Rock, in Guild Wars.

4.4   Technical Metadata

Serious photographers may wish to have a record of the technical details of their photos: camera model, exposure time, environmental and lighting conditions, etc. These details may be held for personal use, or may be associated with photos that are intended for distribution, as they may be useful in critiquing the photos (Participant B notes that his informant believes that "Sharing photos [and their associated technical metadata] can improve his photography skill"). Note that some of these details could be automatically harvested from the camera, as a shot is being taken.

Photo filesize can be useful in deciding how to transfer a photo, or group of photos, to another person—can they be sent as an attachment, or will they break the mailer? Can an image be sent via PXT? Is my online photo account reaching its size limits?

Photo resolution is useful in selecting photos for printing or display as computer 'wallpaper'. Users may not be technically savvy enough to predict what resolution is appropriate for different uses. This understanding isn't critical for selecting wallpaper—the user can easily experiment with different background images—but it can be expensive and disappointing to print A3 copies of low-resolution photos.

Participants reported occasional use of Photoshop or other photo processing software to eliminate minor imperfections (for example, removing red eye), or to improve the appearance or impact of an image (by cropping, altering contrast, or removing people or objects extraneous to the focus of the photo). Serious photographers would want to document their processing of the raw photos.

'Photoshopping' is also used to radically alter photos, generally for creative or humorous effect. Sepia tinting is particularly popular, to make photos appear old, and converting to black and white makes an image appear more 'artistic' (Participant C). Swapping heads on bodies, including celebrities in photos of oneself and friends, or adding objects ("'classic example is you can stamp someone's cheeks with lipstick marks"; Participant D) are stock techniques in the creating of comic photos. Rarely, an image may be 'photoshopped' for more serious reasons; Participant T regretted having missed his graduation ceremony, and so he added himself to his copy of the official photograph of the graduating group. While many of these alterations are so outrageous or conspicuous as to be easily detected by the viewer, subtle changes to

content can be difficult to spot. Again, a record of modifications would be appropriate, for example to aid in maintaining an accurate record of one's activities.

Some photos may be of entirely virtual composition. Online communities offer the opportunity for online experiences that are compelling enough to capture; the 'photos' can be snapped from a first person point of view (Figure 2a) or third person (Figure 2b). These images can be shared within the game or other role-playing environment, and can also be mixed in with digital photo collections in the 'real' world ("I even used both of the images as my desktop wallpapers"; Participant K).

4.5   Heterogeneous Sources and Heterogeneous Storage Destinations

Although the photo collections described in the ethnographies were 'personal' (in the sense of belonging to one individual), one striking aspect of the collections was that the owner did not necessarily take the photos. Photos are frequently shared [1]— people create photo CDs as gifts, email photos of family events to relatives too far away to attend, and use mobile phones to snap and send quirky shots to friends. This situation suggests that browsing may not be sufficient for locating photos taken by others, since the collection owner will be less familiar with these images.

The participants reported that their digital collections tend to be distributed: some photos are on CDs or DVDs, the majority on a hard drive in a photo organizer, a few on their mobile, others in an online archive (perhaps for sharing with overseas friends and relatives), still more photos held within their email system as attachments. Finding a particular photo, then, involves first remembering how and where it is stored, and then recalling the particulars of that system's interface.

But a distributed (and duplicated) photo collection also has can be an advantage: Participant W's informant describes an exceptionally bad experience: '…where their harddrive suffered a melt down. Nothing was recoverable without paying an exorbient [sic] fee to professionals, subsequently many photos were lost. They now backup the images onto CD from time to time to safeguard against a reoccurrence." A physically distributed photo collection is at present a practical necessity; storing a photo collection across several media or storage systems provides a backup in case of hardware failure. Additionally, most participants recognized that storing photos in more than one software system allows the collection owner to separate out photos for personal viewing only from photos that are share-able with others (for example, by placing some photos in an online archive).

Given these arguments for the continuance of distributed collections across multiple photo organization systems, a meta-organizing system would be welcome. One such prototype, developed at Hewlett-Packard Laboratories, supports searching of photos across both local and remote computers, and maintains links between multiple copies of a photo [5].

5   Conclusions

Earlier research on personal photo collection digital libraries is based on assumptions about how people 'naturally' organize their photo collections: by time, by event, and by location. This paper provides evidence that these assumptions are valid; the

students and their informants did indeed report that significant portions of their personal photo collections were organized along these lines.

For digital photo collections, these organizing principles are frequently enforced by the file storage structures. Participants tend to label groups of digital photos, rather than individual photos. This behavior suggests that the most practical tools for locating photos within a collection will be based on browsing, rather than searching; the emphasis should be in identifying the group(s) most likely to include the desired image, and then aiding the user in efficiently browsing them.

It may be possible to automatically harvest some types of useful metadata. The timestamp for a photo allows the user to temporally order the collection for browsing. Different ways of recording the timestamp—for example, as the length of time since the photo was taken—may be helpful in appreciating a specific photo. Photo enthusiasts may desire technical details on camera settings and post-processing. Coupling a digital camera with a GPS could help the user identify photos of a single location that are scattered through time [11], although using GPS data to create human-readable location labels is still challenging. When we share photos by email or SMS, the email or SMS itself is a potentially rich source of metadata. These messages could supply keywords for searching, and the structure of the message itself can also be useful (for example, to record the sender and receiver of the photo).

The highest quality (in the sense of the most personally meaningful) metadata is that which the photo collection owner manually creates. Unfortunately, our participants rarely added more than a few terms to an aggregation of photos, primarily as a folder label. Very few individual photos were annotated. This behavior is not a sign of atypical laziness in our participants; other studies have noted a similar reluctance to formally organize and annotate their photos ([4], [12]). Given these behaviors, it appears that an annotation tool will be used only if it requires very little effort on the part of the user—if, for example, the user can apply metadata to variably sized groups of photos, rather than having to label each photo individually.

Participants recognized that they needed descriptive metadata for their photos both to locate specific images and to appreciate the photos whose contexts were being lost to imperfect human memory. Since photos may be given to others, there is also a concern that the recipients may not store descriptions of photos. But participants are reluctant to record captions visibly on the photo, as this detracts from the image. Ideally, the metadata would be separate from, but bound to, a photo, so that the metadata would 'follow' the photo as the photo is copied or passed along to others ("'captions should be allowed to stick with the picture" [S]).

When photos are grouped into labeled folders, the clustering principles are broad, and primarily based around events ("Wedding", "Birthday"), places ("Wellington"), or people ("Mother", "My Girlfriend"). But a classification scheme based solely on a photo's location limits the access points unnecessarily; a photo of a girlfriend in Wellington could logically fit into two folders, but will be physically stored in only one. The ability to easily (and we must stress "easily", as very few people appear to possess the discipline or desire to consistently catalog their photos) create virtual folders would allow the user to place a given image into multiple browsing categories.

Creating additional virtual folders (or more conventionally, 'albums') may be a better metaphor than providing textual tags, as albums are visually structured to

suggest browsing rather than searching—and browsing groups of related photos is by far the more frequent activity for a personal photo collection.

In the short term, the paucity of metadata does not pose a problem for managing photo collection. We can rely on our memory of the approximate time the photo was taken to locate it, and we are also confident that we can remember the significant details of people, places, and activities depicted. In the long term, however, our memory inevitably fails us—and we end up with a hard drive full of unidentifiable images. At that point, we are grateful for any metadata that helps us recall the context of our photos, to recover our chronicle of our life. Any tool that encourages us to record photo metadata, or that automatically collects metadata, will be appreciated.

References

1. Adams, A., Cunningham, S.J., Masoodian, M.: Sharing, Privacy and Trust Issues for Photo Collections. Working Paper 01/2007, Computer Science, Waikato Uni. (2007)
2. Cooper, M., Foote, J., Girgesohn, A., Wilcox, L.: Temporal event clustering for digital photo collections. ACM Transactions On Multimedia Computing, Communications, and Applications 1/3 (2005)
3. Cunningham, S.J., Jones, M.: Autoethnography: a tool for practice and education. In: Procs. of the 6th New Zealand Int. Conf. on Computer-Human Interaction, pp. 1–8 (2005)
4. Frohlich, D., Kuchinsky, A., Pering, C., Don, A., Ariss, S.: Requirements for Photoware. In: Proc. CSCW 2002, pp. 166–175
5. Gargi, U., Deng, Y., Tretter, D.R.: Managing and Searching Personal Photo Collections. HP Labs Technical Report HPL-2002-67, p. 9 (2002)
6. Glaser, B., Strauss, A.: The Discovery of Grounded Theory, Aldine, Chicago IL (1967)
7. Graham, A., Garcia-Molina, H., Paepcke, A., Winograd, T.: Image and cultural digital libraries: Time as essence for photo browsing through personal digital libraries. In: Proceedings of the ACM/IEEE Joint Conference on Digital Libraries 2002, pp. 326–335 (2002)
8. Mills, T., Pye, D., Sinclair, D., Wood, K.: ShoeBox: A Digital Photo Management System. AT&T Laboratories, Cambridge (2000)
9. Naamen, M., Yeh, R.B., Garcia-Molena, H., Paepcke, A.: Leveraging context to resolve identity in photo albums. In: Proceedings of the ACM/IEEE Joint Conference on Digital Libraries 2005, pp. 178–187 (2005)
10. Nardi, B.A.: A small matter of programming: perspectives on end user programming. MIT Press, Cambridge, Mass (1993)
11. O'Hare, N., Gurrin, C., Lee, H., Murphy, N., Smeaton, A.F., Jones, G.J.F.: My Digital Photos: Where and When? In: Proceedings of Multimedia 2005, pp. 261–262 (2005)
12. Rodden, K., Wood, K.R.: How do people manage their digital photographs? In: Proceedings of CHI 2003, pp. 409–416 (2003)
13. Smeulders, A.W.M.: Content-Base Image Retrieval at the End of the Early Years. IEEE Transactions on Pattern Analysis and Machine Intelligence 22/12 (December 2000)
14. Strauss, A., Corbin, J.: Basics of Qualitative Research: Grounded Theory Procedures and Techniques. Sage, Newbury Park (1990)