Capture, Annotate, Browse, Find, Share: Novel Interfaces for Personal Photo Management

Hyunmo Kang
Human-Computer Interaction Laboratory,
University of Maryland Institute for Advanced Computer Studies, College Park, MD

Benjamin B. Bederson
Human-Computer Interaction Laboratory,
Computer Science Department, University of Maryland
Institute for Advanced Computer Studies, College Park, MD

Bongwon Suh
Palo Alto Research Center, Palo Alto,CA

The vision of ubiquitous digital photos has arrived. Yet, despite their broad popularity, significant shortfalls remain in the tools used to manage them. We believe that with a bit more creativity and effort, the photo industry can solve many of these problems, offering tools which better support accurate, rapid, and safe shared annotations with comfortable and efficient browsing and search. In this article, we review a number of projects of ours and others on interfaces for photo management. We describe the problems that we see in existing tools and our vision for improving them.

1. INTRODUCTION

The days when people debated the relative merits of film versus digital imagery now seem almost quaint. And with hindsight, film seems destined to have been a chapter in history along with LP vinyl records—a temporary physical analog recording medium. Although some aficionados may still prefer some qualities of those dust-gathering mediums, the advantages of digital media have become clear. The rapid and inexpensive ability to edit, annotate, search, share, and access has brought digital media to ubiquity.

And yet, with all this promise, shortfalls remain in the overall user experience. How many of us have replaced old unlooked-at shoeboxes of prints with unlooked-at digital archives of image files? How much has our ability to find a particular set of

---

photos really improved (i.e., can you find those photos of a visiting uncle with your sister when they were children?) How do we record stories from our parents describing family photos? How do we make sure those stories stay with the photo in question and get distributed to all copies of that photo within the family? Perhaps of most concern, how do we ensure that these annotations stand the test of time and remain accessible as computers, file formats, recording mediums, and software change?

These changes are all happening within the context of human behavior, which does not change so rapidly. People still like immediate gratification: take pictures rapidly, print them, and share in social settings. Some people spend a lot of effort creating photo albums or "scrapbooking." Of course, many do not. Understanding which behaviors are fundamental and which are side effects of current technology is crucial, because this understanding can and should influence where researchers spend their effort.

We explore these issues and more with much of the intellectual motivation coming from Ben Shneiderman, our close colleague who has pushed for a deeper understanding of and better support of photo management for more than 10 years. His personal photo archives document the field of human–computer interaction (HCI) going back to its beginning. He regularly shares those photos with great enthusiasm to visitors, motivating and exciting all of us—largely because of the care and consistency he has applied to annotating and organizing his photos. He regularly pulls up old photos of lab visitors showing everyone what they worked on 5 or 15 years ago (and of course, showing what they looked like too!). His early exploration of tools to support photo management (with co-author Kang) led to PhotoFinder (Kang & Shneiderman, 2000), and the ensuing PhotoMesa tools (Bederson, 2001; Bederson, Shneiderman, & Wattenberg, 2002). His personal interest helped inspire the authors of this article as well as other lab members to pursue the development of approaches and software to improve all of our user experiences when managing photos.

This, of course, all happened during a time of tremendous commercial activity in this area. There are wildly popular photo sharing Web sites, such as Flickr (http://www.flickr.com; Yahoo!), Picasa Web Albums (http://picasaweb.google.com; Google), Snapfish (http://www.snapfish.com; HP), Shutterfly (http://www.shutterfly.com; Shutterfly), and PhotoBucket (http://www.photobucket.com; Photobucket), as well as equally well-used desktop photo applications such as Google Picasa (http://www.picasa.com), Adobe PhotoShop Album (http://www.adobe.com/products/photoshopalbum/starter.html), and ACDSee (http://www.acdsee.com). The two approaches (desktop application and Web site) are interesting to look at because they each offer distinct advantages to users. For example, Web sites are available anywhere and facilitate sharing, whereas desktop applications are faster, support higher resolution photos more easily, provide local ownership of photos, and offer richer interaction capabilities. It is interesting that each approach is gaining characteristics of the other. Web applications begin to offer dynamic, interactive content rather than static html pages through AJAX and Flash technologies. In addition, they often include plug-ins to ease uploading or improve performance, and some offer APIs to enable desktop applications to access their data directly. At the same time, many desktop applications are offering Web capabilities such as sharing.

Yet even with this commercial activity, the full potential of personal photo management has not been reached. There is the opportunity for richer annotation interfaces, automated content analysis, improved sharing, and more creative organizational strategies. Our hope is that more photos end up with better metadata, enabling faster, easier, and more accurate and enjoyable retrieval and use.

In this article, we look at some of the key activities and behavior patterns of personal photo users and examine how innovative user interfaces have the potential to enhance users' power, satisfaction, and control in managing and exploring their images. Starting with a close look at annotation, we examine how a combination of manual and automated techniques can improve how people associate metadata with photos. We then look at how the resulting richer metadata can enable better interfaces for searching and browsing photos. Finally, we end with a discussion about the importance of sharing photos and how new interfaces enable that.

2. GUIs FOR ANNOTATION

An essential question is, How valuable is photo metadata? Our own assessment of user needs (Shneiderman & Kang, 2000) coupled with reports from other researchers (Chalfen, 1987; Naaman, 2005; Rodden & Wood, 2003), and our personal experience come together on this. They indicate that the photo metadata such as dates, locations, and content of the photos (especially people's names) play a crucial role in management and retrieval of personal photo collections.

However, in spite of the high utility of the photo metadata, the usability of software for acquiring and managing the metadata has been poor. The manual photo annotation techniques typically supported by photo software are time-consuming, tedious, and error prone, and users typically put little effort into annotating their photos. In fact, the industry attitude tends to be that because users do not annotate their photos very much, it is not necessary to spend much energy adding good support for it. However the success of keyword labeling, so-called *tagging*, systems such as del.icio.us (http://www.del.icio.us) and Flickr hints that users indeed want to make annotations when reasonable utility and usability are supported. We believe that the photo annotation has enough utility for some users and it is the usability of software that needs to be improved.

In this section, a few innovative approaches are presented to show how interaction and graphical user interface (GUI) design can improve the usability of the photo annotation process when they are based on careful understanding of users' behavior and usage pattern. In addition, we explain how those designs have been evolving over time to support a broader range of annotation tasks by combining the accessible technologies with the analysis of users and their needs.

2.1. Advanced Manual Annotation: Direct Annotation

Because annotations based on automated image content analysis are still limited, we developed an advanced manual annotation mechanism that can significantly reduce users' annotation workload under certain circumstances. From the

observations of personal photo annotations (Shneiderman & Kang, 2000), we found that there were three interesting characteristics that could be useful for our interaction design.

- Personal photo libraries often contain many images of the same people at different events. In the libraries we looked at, we typically found 100 to 200 identifiable people in several thousand photos. Furthermore, the library has a highly skewed distribution with immediate family members and close friends appearing very frequently.
- Textual search often doesn't work reliably because of inconsistency in names with misspellings or variants (e.g., Bill, Billy, William).
- Lists of names of people appearing in photos are often difficult to associate with individuals, especially in group shots. Textual captions often indicate left-to-right ordering in front and back rows, or give even more specific identification of who is where. However this approach is tedious and error prone.

Based on these observations, we collaborated with Ben Shneiderman to develop the concept of *direct annotation* (U.S. Patent #7010751), which is a technique using selectable, draggable labels that can be placed directly on the photo (Shneiderman & Kang, 2000). Similar interfaces have also appeared recently on Flickr and MySpace. Users can select from a scrolling or pop-up list and drag by mouse or touch screen. This applies direct manipulation principles (Shneiderman, 1983, 2005) that avoid the use of a keyboard, except to enter a name the first time it appears. The name labels can be moved or hidden, and their presence is recorded in the database or in the header of an image file in a resolution-independent manner. The relative location of the target is stored based on an origin in the upper left-hand corner of the photo with the point in the range $(0, 0) – (1.0, 1.0)$ corresponding to the full image. This approach not only associates a name with a position on the photo but also ensures that each name is always spelled the same way.

This simple rapid process also enables users to annotate at any time. They can add annotations when they first see their photos on the screen, when they review them and make selections, or when they are showing them to others. This design, which supports continuous annotation encouraged users to do more annotation, especially in collaborative situations such as with PhotoFinder Kiosk (Kules, Kang, Plaisant, Rose, & Shneiderman, 2004; Shneiderman et al., 2002). In a public setting, visitors of PhotoFinder Kiosk added 1,335 name labels using direct annotation while adding 399 captions using traditional type-in method.

The direct annotation mechanism was revised later so that it enables users to define their own categories such as activities, events, locations, objects in a photo in a hierarchical way (Figure 1) in addition to person names. A few alternative and complementary direct annotation mechanisms such as split menu annotation, hotkey annotation, and pop-up menu annotation have also been designed and developed to accelerate the annotation process. These are described in more detail in the following subsection.

A pilot experiment was conducted to see if the direct annotation method improved the annotation process in terms of annotation time and users' subjective preference compared with the traditional caption method or the click-and-type
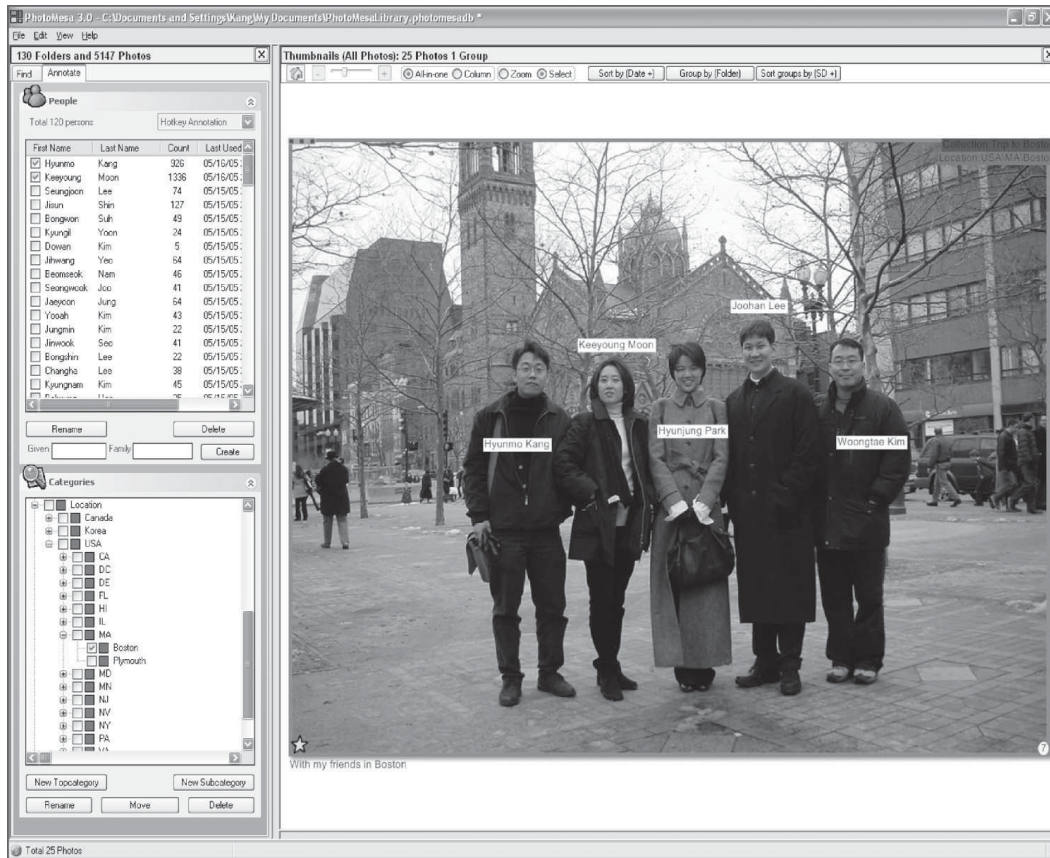
FIGURE 1   The revised direct annotation mechanisms as implemented in PhotoMesa, the successor to PhotoFinder. *Note.* Users can add a caption under a photo, or add a particular attribute such as favorite (a yellow star on the bottom left of photo) or hidden. In addition, labels can be dragged from the list of people (on the top left) or from the user-defined category tree (on the bottom left). A label can be directly placed on the photo to represent where the individuals or objects are located within the photo.

method (Goto, Jung, Ma, & McCaslin, 2000; Jung, 2000). Forty-eight volunteers participated in a within-subject design, whereby each participant attempted an annotation task (20 names in five photos) on each system. The direct annotation method was significantly preferred, but no significant difference was found in the mean annotation time. A more extensive user study might help identify under what circumstances (e.g., number of total labels, average number of people in a picture, pattern of people's appearance in a personal photo collection, etc.) different annotation mechanisms work better than others in terms of completion time, error rate, users' satisfaction, or confidence.

2.2. Enhanced Direct Annotation: Bulk Annotation

The direct annotation mechanism was enhanced through a series of design cycles to support more efficient and productive annotation. Perhaps the most notable

DOCKET ALARM

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.