AlbumViz

redesigning the digital photo album

contributor:
Nivay Anandarajah
Department of Mechanical Engineering
University of California, Berkeley
nivaya@berkeley.edu

instructor:
Maneesh Agrawala

CS294-10 · Visualization · Final Report · December 2008

Abstract

This paper discusses AlbumViz, a tool used to visualize and navigate personal photo collections. AlbumViz was developed in a user-centered study investigating the issues of visual information seeking. By studying how the common user identifies and groups his or her own photos, metrics for efficient design were established. This tool was implemented to create a holistic navigation anchored by the preferred method of search "time + keyword." By appropriately using tag clouds, and image clustering, data was encoded in an effective manner.

Keywords

Visual Information, Space Management, Photo Presentation and Seeking, User Study, Tag Cloud

Introduction/Motivation

Advancements in digital photo technology have allowed users to take an increasing number of photos. As the quantity of photos taken has risen, the means of organization has stayed relatively the same: the photo album.

Unfortunately, the structure of traditional photo albums does little to aid navigation. Finding photos can be an overwhelming experience. Typically photos are kept in a folder hierarchy with arbitrary ordinal labels. Within these folders, viewing speed is limited by the number of thumbnails one can scan at a time. This methodology only works provided the users are familiar with their albums. As data storage gets cheaper, these organizational problems will only grow.

The goal of the AlbumViz study is to investigate a user-centered solution to this problem. The tool aims to minimize the time to find a photo while allowing interesting patterns to emerge through organization and layout. The tool was scoped to work with photo collections of the common user. In effect, this would minimize the reliance on data preparation.

Related Work

There currently exists a wide array of existing interfaces that attempt to visualize existing photos. The one used most often is windows explorer shown in Figure 1.

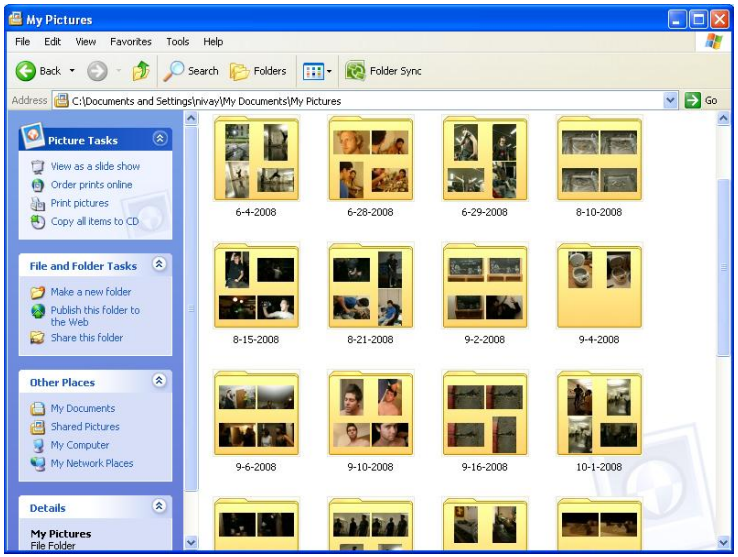

Figure 1. Windows explorer photo organization

This common methodology relies entirely on user input to provide a system of organization for one's photos. An example of more advanced photo album interfaces would be Picasa's web album shown in Figure 2.

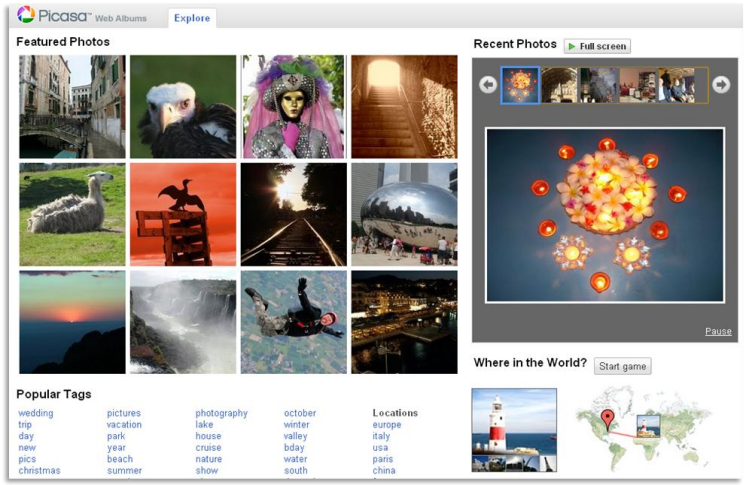

Figure 2. Picasa's Web Album Interface

This interface makes use of varying thumbnail size and metadata inputted by the user. The metadata is shown as a list of popular tags, or accessed through a straight keyword search. These interfaces give insight on how to use metadata and dynamic scaling, yet do not instruct how to navigate one's own photo collection besides direct keyword search.

Aside from presenting photos, there has been interesting research in preparing data for navigation. Professor Marc Davis at UC Berkeley worked on the Mobile Media Metadata project. By using cell phones and a series of sensors, they were able to automatically tag photos at the time of capture with geographic and context data. This metadata was shared through social exchanges and then used to navigate through personal shots. This project shows how metadata can be used in various ways to organize and share data. Professor James Wang at PSU developed an automatic linguistic editing of pictures that tagged photos in real time based on statistical modeling, learning and wavelet transforms. Automatically tagging pictures based on modeling could help visualize an album without the use of data preparation. Mark Sanderson and Paul Clough at the University of Sheffield developed ImageClef, a robust context based image retrieval system. ImageClef's context retrieval system algorithm could be useful for developing image clustering. Although data preparation is not within the scope of this study, these studies give an understanding of how to effectively deal with existing metadata or a lack of metadata.

Methods

With a problem so intimately tied to the user's personal life, it was first crucial to perform a user needs and usability test. The goal of the user test was to determine how individuals identify, group, and navigate their own photos. The methods used were interviews and digital photo album data analysis.

The key findings were users fall into two categories: those that don't put any effort into organizing their photos and those that do. The processes are as follows:

> *Without organization:* (1) import photos into folders organized by date (2) search conducted by estimating date followed by visual traversal
> *With organization:* (1) organize folders into event folders such as "Tahoe Trip 08' "; individual folders organized roughly by date (2) search conducted by starting with date approximation and finding specific event followed by visual traversal

With the advent of online photo albums, users were provided a different framework to interact with their photos. It was found, however, that organization and search of similar photos is still conducted similar to offline. The addition of metadata, and "similar" or "related" tags are used primarily to explore the photos of others. Upon questioning, it was found that people just did not feel comfortable enough in their tagging to perform a tag search for their own album. Direct search was too haphazard and they were prone to miss the photo they were looking for. They would rather navigate the album by approximating date and identifying event folders.

The user study showed that finding photos is not as much of a search form as a holistic navigation. It would be in the user's best interest to stay in the preferred method of "time + event keyword."

Using this as metrics for the design, layout techniques were explored to best encode time and keyword information.

It was chosen to use Flickr's web albums as the data set to build the tool around. Flickr is the one site where people are already dutifully in adding tags to their photos. By using existing metadata on flickr, the tool will gather a small level of data to create an effective visualization while not requiring additional data for preparation. The metadata available on Flickr for each photo includes: a tag (keyword) list, date added, date taken, sets a photo belongs to, and some geodata.

The layout was built around the framework of approximating the date the photo was taken followed by scanning through events. The date added metatag was the best correlation to defining the date parameter. The tag (keyword parameter) was the best correlation to defining the event parameter.

Instead of having a search form, it was important to show context around the input. It was decided that an effective method in presenting the data would be in date intervals with popular tags within each. The photos taken were scanned for each interval of time. The number of times a tag appears (it's popularity of use) within that time period is then correlated to its size in the display. This provided a simple yet effective method in correlating keyword text to photos. By ordering the tags by size, people immediately get the sense of what they were taking photos of most. By clicking on these tags for these intervals, the user could immediately be brought to the photos taken during this interval for this tag – similar to finding an event folder within a certain time period.

Following first round of user feedback testing, it was found people had difficulty first navigating their photos using pure text. Despite this being the least overwhelming method, representative images were preferred. Showing one context based clustered image for each tag was effective in reducing clutter. By allowing users to click on different tags, and see different representative images, a level of interaction was added to the visualization. By clicking on the representative image, users would be navigated to the

rest of the subset of photos represented by this tag and date interval.

The prototype was implemented using Adobe Flex and Actionscript. The Flickr API was used to retrieve user names, search for photos within a specified date, and retrieve photo tags, To create a list of popular tags, photos were searched and scanned in batches. The count of the occurrence of tags was kept. For the representative image, a random image was found. Unfortunately, content based image clustering was not implemented at this stage due to difficulty in working with the Flickr API. Implementation went smoothly overall except for running into problems with server lag. Due to Flickr taking time to return photos, many parameters such as maximum number of photos scanned had to be limited.

Results

The final tool was designed to run as a Flash file on the web for all Flickr users to access. Users start by inputting their Flickr user name, the start date, the length of the intervals they want to search for, and the number of intervals they want to search. They are given a screen showing all popular tags within each interval, and a representative image for each top tag.

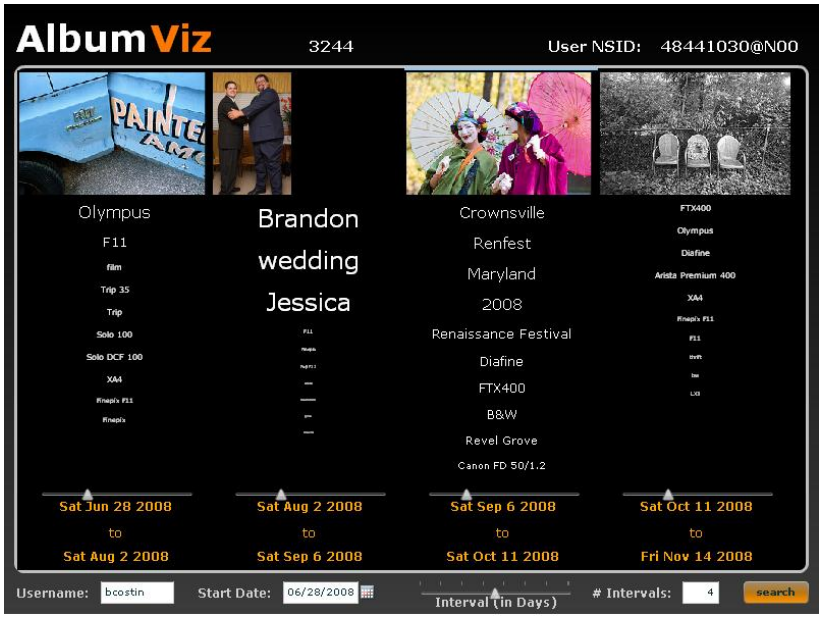

Figure 3. Screenshot of AlbumViz search

As Figure 3 evidences, large events are often highlighted by popularity. Users can then click on tags to change the representative image. Once they find the “time interval +

tag" subset they are looking for, they can click on the representative image to see the rest of the images that belong to that subset.

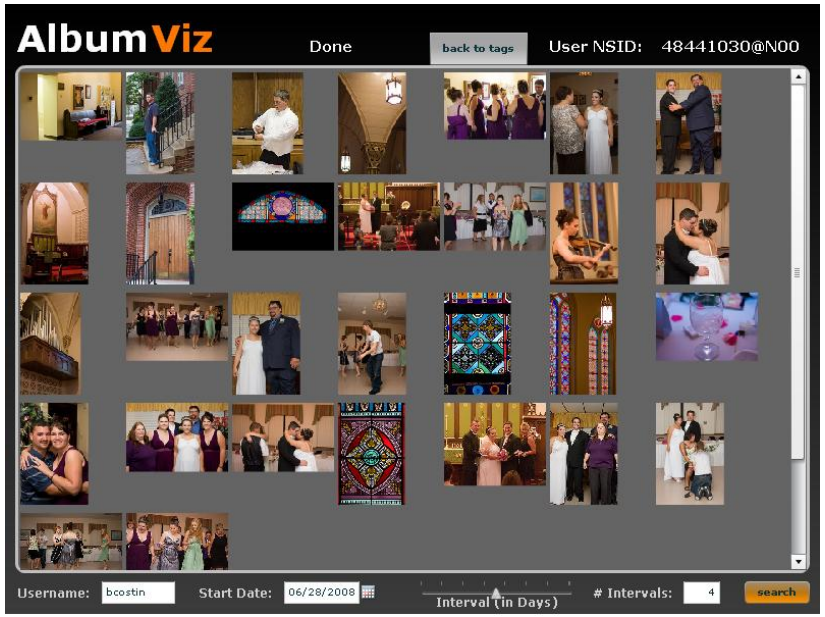

Figure 4. Screenshot of AlbumViz results display

These images can finally be clicked on for an enlarged view and information regarding the Flickr image title.

AlbumViz tested rather favorably. People enjoyed the display of key tags over time. As a proof of concept, people found this as a novel approach to a difficult problem. However people wanted more indication on how to use the interface at first glance. Most users needed an explanation as to exactly what is being displayed and how to interact appropriately.

Discussion

AlbumViz was developed in a user-centered study investigating the issues of visual information seeking. A number of contributions were made during this approach to a complex problem.

By performing a needs and usability assessment, natural habits of the user were identified and used as design metrics. It was found that finding photos is not as much of a search form as a holistic navigation. People's preferred method of navigating their photo set is "time + event keyword." By staying sensitive to these issues, the amount of data presented can be reduced to make visual information seeking a less overwhelming experience.

The next step was to create a visualization around this "time + event" framework. By using size to encode the popularity of tag and x-position to encode time intervals, a type of navigation histogram was created. The idea of relying on the text of associated metadata to navigate photos was rated as "surprisingly effective" in navigating visual information. This allowed users to gather concrete information without having to guess what a picture was about. The inclusion of having a representative image provided a strong correlation between text to images. A single content clustered image allows the user to get a sense of the style of photos without getting overwhelming by the whole range of a subset. This project provided a unique use of content based image characterization as a presentation tool as opposed to a search tool.

By using this array of design techniques, AlbumViz attempted to tackle the problem of the finite level attention people are able to give to a database of visual information. Where traditional paper photo albums were designed for the experience of slowly flipping through, digital photo albums have an entirely new context of use – navigation, and search. By using cognitive visualization techniques and interaction, AlbumViz provides a new way of addressing logistical issues of digital photo albums while maintaining the satisfaction of traditional albums.

Future Work

In order to have a fully working prototype, content based image clustering must be implemented. This would be the best way to choose a representative image without the need of instructive metadata. Additionally, to have a smoothly working prototype, key data should be loaded and cached ahead of time. It was found the effectiveness of the interaction started to fade due to long server lag. Once a final prototype is built, further user testing could be conducted.

It was found that the layout of the main AlbumViz search interface provides affordances as how to operate. Most users had to be explained what a certain encoding

means or what image corresponds to what text. Implementing these affordances could vastly improve the usability of the tool. The layout should be redesigned to instantly know what text corresponds to what images. This could be implemented through animation and reordering of the text and images.

Additionally, there are still some dangling encodings. The y-position of the tags is simply a redundant encoding of tag popularity from greatest to least. This was left as is to prevent the interface from being overwhelming. If an appropriate encoding was decided for this y-position encoding, the visualization could be stronger.

The ability to build a visualization tool for photo data is entirely limited by the quality of the metadata. AlbumViz would not be effective if the user did not input comprehensive tag data. An interesting problem to investigate would be how to facilitate data preparation for the sake of navigation. Metadata could be added automatically by the hardware at time of exposure, manually by the user afterwards, or methodically using statistical data analysis. This complex problem could have any number of solutions.

References

Ahlberg, Christopher and Ben Schneiderman. *Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays.* University of Maryland: College Park. Accessed October 2008. ftp://ftp.cs.umd.edu/pub/hcil/Reports-Abstracts-Bibliography/3131html/3131.html

Davis, Marc. *Mobile Media Metadata.* University of California Berkeley. Accessed October 2008. http://garage.sims.berkeley.edu/research.cfm#MMM

Sanderson, Mark and Paul Clough. *ImageCLEF.* University of Sheffield. Accessed October 2008. http://www.imageclef.org/

Schittly, Stephane. Create a Flickr Viewer Using Flex. Accessed October 2008.

http://www.thetechlabs.com/interfaces/create-a-flickr-viewer-using-flex/

Tufter, Edward. The Visual Display of Quantitative Information. 2nd Edition. Graphics Press, 2001.

Wang, James. *Automatic Indexing of Pictures – Real Time*, University of Pennsylvania. Accessed October 2008. http://wang.ist.psu.edu/IMAGE/