Photo-triage: Rapidly annotating your digital photographs

Steven M. Drucker, Curtis Wong, Asta Roseway, Steve Glenner, Steve De Mar
Next Media Research Group

Abstract:

The Photo-triage application is meant to be an essential part of the digital photography lifestyle. It can fit as a component wherever photo management is done (shell, picture-it, media-center, etc.). The central idea is to facilitate rapid, convenient categorization of one's personal photos into at least the following categories: hidden/private, majority, highlights, best and/or representative. See figure 1. This application is meant to fill an empty niche in the usage of digital photos: that is, there's no easy way add metadata to photos to mark them for printing, for sharing, or for slideshows without creating separate versions of the photos and copying into separate folders. We propose a sorting metaphor that will add implicit metadata when one first goes through the photos.

Introduction:

The concept of photo-triage is inspired by the process that people used to go through when first receiving pictures back from the photo processor. People would go through a stack of photos and quickly sort pictures into different groups such as good, bad, private etc. This could be done in collaboration with others or as an individual process. People might iteratively go through the favorites and pull out the best of those.

As we've moved into the age of digital media, we have yet to have an interface that is as easy to use as the analog process for sorting photos. In fact, there have been repeated reports of people using a pad and paper next to the screen to write down the filenames of photos for subsequent printing or emailing. If we expect the computer to be a digital hub for managing one's media, it is a necessity that we have a satisfying and efficient process for sorting our photos.

The key idea behind Photo-triage is RAPID sorting and thus implicitly tagging the photos with metadata. This should be done in the context of looking at one's photos in a natural manner. Ideally, a slideshow or filmstrip view would suffice augmented by a few simple UI controls to permit rapid annotation of the image.

There are several different preferred interfaces for sorting through photos and we are currently investigating the tradeoffs in image size, number of images on the screen, and UI for sorting and annotating the photos. However, the basic requirements are all the same:
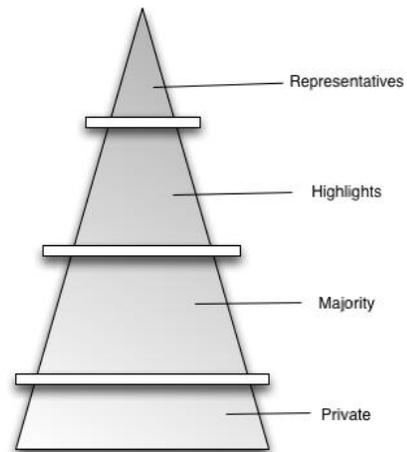

Figure 1: The Photo-triage pyramid.

Photo sorting needs:

1) <u>Part of a natural viewing experience:</u>
   The pictures need to be large enough to see details like faces, so thumbnail views are only marginally useful. It is best if we augment the slideshow or the filmstrip viewer.
2) <u>Preserve context for annotated (sorted) photos:</u>
   Photos should not be copied into separate folders, but kept with the photos around them.
3) <u>Rapid</u>
   People don't have the time or the desire to use awkward interfaces for minimal reward.
4) <u>Immediate feedback and reinforcement of sorting</u>
   File views of photos should immediately reflect results of sorting (ie added annotations).
5) <u>Enable simple comparisons of similar looking photos</u>
   Should have a mechanism for comparing between several, similar photos to select the best few.

After an image is sorted, we need to make use of this annotation in a variety of ways within our photo browsing applications to change the way that images are presented to the user. We can store the annotations either within the file (EXIF tags) , within WinFS, or using auxiliary files. Shell views can then adaptively show the best images, or all the images.

Basic Slide Show implementation

Our first vision for the Photo-triage application is primarily based on an augmented slide show interface. A stream of images is shown to the user with the currently considered image expanded the largest. See figure 1. As images roll out of the right side (can be customized to either side), they are separated based on time clusters. Images can easily be categorized as positive or negative (or skipped). This can be done via keyboard, remote, mouse, or even gestural interfaces. See figures 2 and 3.

Figure 1: A basic interface for Phototriage. We retain a slideshow like view, but add information about clustering, and rated documents.

Figures 2 and 3: Image classified as like or dislike.

A default behavior to cycle the images on to the next one can be added so that users do not need to interact at all with the interface creating a kind of slideshow. The image in focus slowly zooms larger to enable examining details until the person chooses which category to assign it to.

Several added features should be supported:

Undoing:
One of the first enhancements to the basic Slide Show Implementation would be the ability to take back the last classification. Multiple levels of undo should be supported (as well as redo). The image should animate back from the appropriate area on the screen where the image was classified and any indications on the image itself should be removed.

Clustering:
One common issue is when multiple snapshots are taken in immediate temporal proximity with each other. This is often in portraits or group shots where a single best image is desired. We should work out some way in the UI of indicating in advance that a number of similar shots are coming. See figure 4.

The spacing between the photos can give some indication of the difference in time between when the shots were taken.

In addition, there should be a way so that the user can simultaneously look at all the shots within a cluster and just pick the best shot.

Figure 4: Document clusters

Tablet Based Interface:
A tablet or a Mira-device can be a natural interface for browsing and classifying photos. While current tablets need a pen based input, if a hand can be used, the natural flow of the interface can be enhanced even more. See figure 5.

Figure 5(a): A Tablet based interface for phototriage. Figure 5(b): Using the fingers to sort items on the tablet.

Multiple, parallel presentations:
Looking at pictures is often a joint experience. This naturally lends itself to Media Center settings. Augmenting the slideshow from a 12 foot view along with a rich tablet interface presents an exciting possibility for digital media enjoyment in the future. Figure 6.

Figure 6: Letting more than one person sort through photos at the same time. The large display contains only the appropriate UI elements necessary for the task.

Simultaneous Comparison:
Sometimes a single image on the screen at any one time is best (maximizes the screen size of image allowing best observation of the details, and joint attention from many people on the same image). However, there are also cases where one might wish to have several images on the screen at the same time to allow for rapid comparisons. In general, people can sort through images quicker if they can see more than one at once. We need to figure out a balance between multiple images being shown simultaneously and maximizing screen real estate. One option is shown in figure 7:

Figure 7: Choosing the best shot from several similar pictures with auto zoom of the selected image

Subsequent Applications:
As we've observed with the MediaFrame project, it really helps to SHOW the benefits of annotation in some direct way. Other applications should make use of the simple classifications that have been put onto the data. The shell should allow users to group by triage category allowing users to easily find the best pictures that they've previously classified.

Once images have been classified, it should be easier to generate presentations that don't follow the typical linear format. People should be able to delve into areas where they want to see more information, or pivot on some piece of metadata to branch off in new directions. We envision an ongoing stream of events available for people to browse (figures 8, 9). We call this ongoing stream and system for viewing photos and other media a Lifespace which is another project that is currently in progress within the Next Media Research Group. Lifespaces might be authored via dynamic presentation templates which leverage the information provided during the triage process and whatever other metadata is available (time, gps location, etc.).
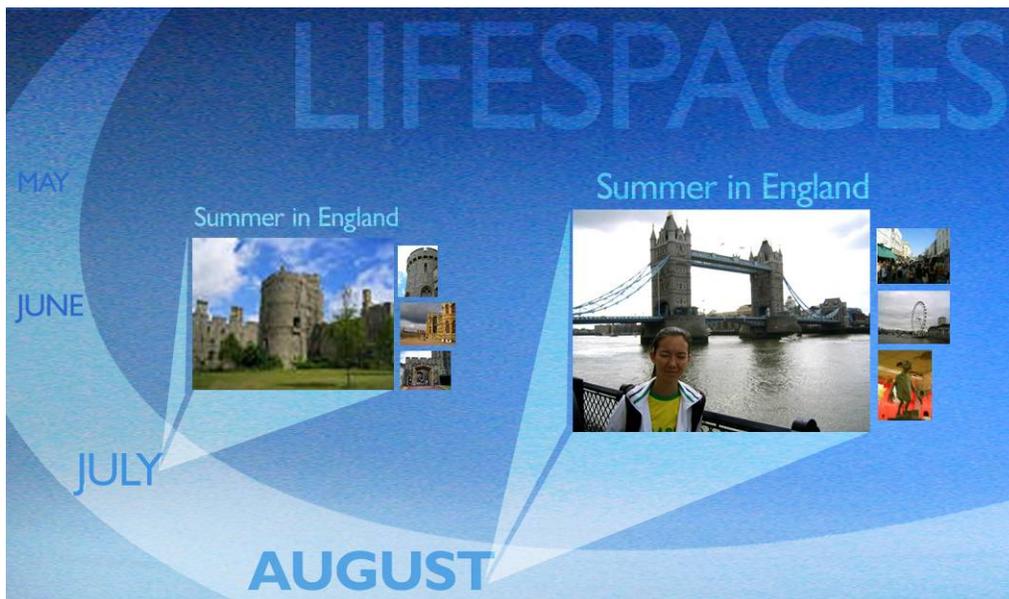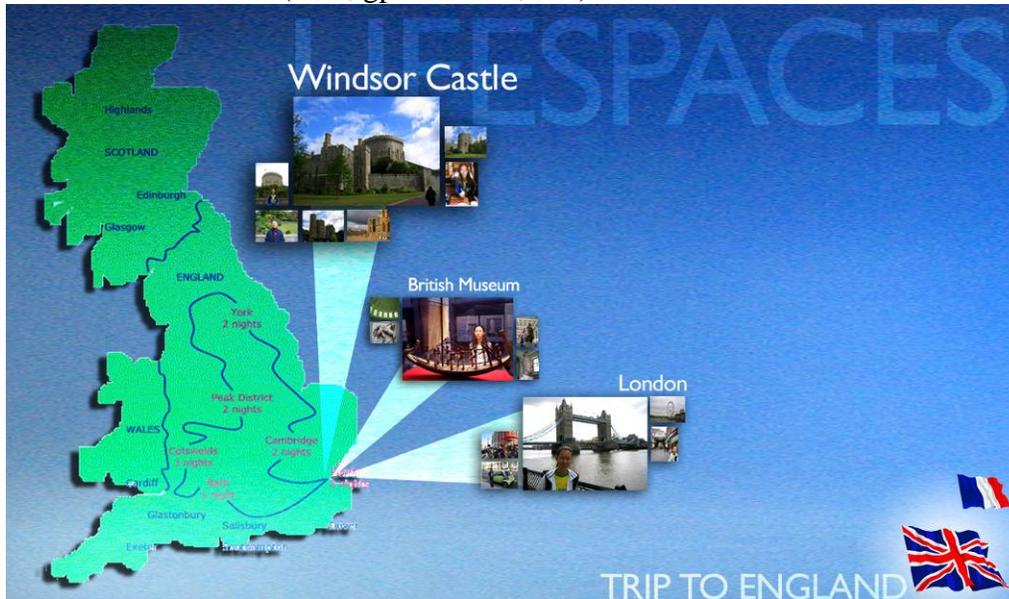

Figure 8: Two automatically generated presentations from metadata from the photo-triage experience and from other associated metadata (time of shot, location).

Figure 9: Parents viewing a Lifespace browsing experience from 12' created via photo-triage and other metadata with minimal authoring effort.

Summary:

We all know that applying metadata to our data is simultaneously incredibly beneficial, and incredibly painful. This basic conflict threatens to impoverish our experiences with digital photographs. Our attack on this problem is 2-fold. First, design interfaces and experiences that make annotating and classifying the photos easier and faster.

Our previous work on the MediaFrame Browser (http://mediaframe) used automatic clustering and automatic annotation tools to assist in manual annotation, while the current work on Photo-triage focuses on interfaces to facilitate rapid sorting.

At the other end of the experience, we must find ways to make it clear that annotating the data has a clear and obvious benefit to the author and the viewer. The hardware frame and the PocketPC experiences for the MediaFrame Browser and the Lifespaces project both work in this area.