Image Retrieval on Mobile Devices

Adrian Hub
University of Bamberg
D-96052, Bamberg, Germany
adrian.hub@uni-bamberg.de

Abstract

With the growing number of mobile devices and the access possibility to thousands of images from these devices, the users call for efficient image search techniques for mobile devices. Desktop paradigms cannot be used with the smaller screen sizes, hence it is needful to offer alternative searching and browsing strategies, which are adapted for mobile devices. In this paper we describe our ideas how image retrieval on mobile devices can be accomplished.

1 Introduction

The amount of small mobile devices, which we use in our every day life, grows constantly. There are cell phones, smart phones, tablet PCs, netbooks and so on. Most of them are equipped with more or less powerful cameras and all of them offer enough storage capacity to take and store a lot of photos on the device itself.

Furthermore the Internet is easily accessible via broadband connections offering access to an unlimited number of images. There are image search services like Google images[1], photo communities like Flickr[2] or Picasa[3] and social networks like Facebook[4]. All of them let the user search for thousands of images and the latter ones let him especially browse through the photos of friends and colleagues.

Due to the close integration of social networks, the access to the Internet and the huge storage capacity on mobile devices, a lot of own and external (from the web) photos are stored on the device or at least accessible. Unfortunately, most mobile devices are usually not designed to manage thousands of photos concerning the small screen size and limited control possibilities. The typical thumbnail view is inapplicable, as if the device should give a good overview with thumbnails, they had to be tiny making recognition of images very difficult or if the recognition with bigger thumbnails could be good, there can be placed only few on the small screen of the device regarding their typical screen sizes less than 4 inches.

Another drawback could be the lack of a keyboard, which makes tagging and searching for tags quite hard. As our measurements with Flickr crawls showed, a lot of images remain untagged (about 40% in this 'classical environment'), so tagging on mobile devices will probably be even less used. One more reason why image retrieval on mobile devices is a challenging topic.

Hence, we presented a hybrid system, called Picadomo [Hub *et al.*, 2009], that makes use of Hierarchical Faceted Search (HFS) [Hearst *et al.*, 2002] for the purpose of image retrieval and is adapted for the small screen size of mobile devices. It is based on our visual faceted search for desktop PCs VisualFlamenco [Müller *et al.*, 2008] and combines a tag-based search, search techniques based on EXIF data as well as Content-Based Image Retrieval (e.g. low-level visual features) to generate a good browsing experience and help the user for finding desired images on mobile devices (see Fig. 1).

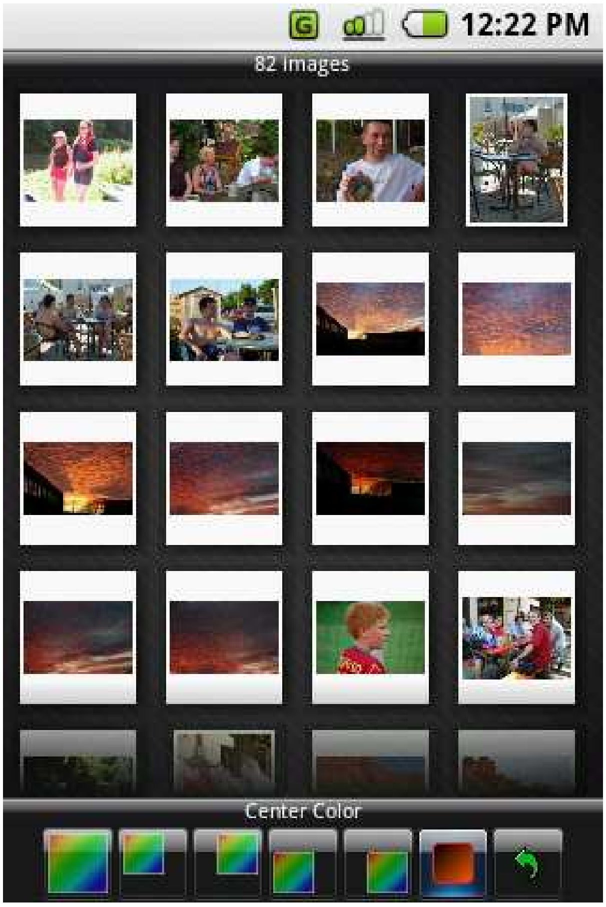

Figure 1: Screenshot of our Prototype

We now want to improve our prototype for mobile devices in different ways. First we want to address more image sources. Picadomo could handle only images stored on

[1] http://www.google.com/imghp
[2] http://www.flickr.com/
[3] http://picasaweb.google.com/
[4] http://www.facebook.com/

the internal memory card, due to the fact, that the feature extraction takes some time and cannot be done on the fly. Therefore it is essential to use server and/or peer-to-peer networks to broaden the searchable image sources.

Second, the facets used to search the images can be improved further. Our user experiments in [Hub *et al.*, 2009] showed, which facets were used often and which facets could be neglected. But it often depends on the user and/or the dataset, e.g. EXIF data like the time or place the picture was taken may be only interesting for own content, but not for images from the Internet.

Third, the search user interface can be enhanced in many ways. There are a lot of techniques in text retrieval, regarding the query specification, presentation of results, query reformulation, personalization and visualization, just to name a few. Some of them can be adopted for image retrieval as well, e.g. keyword-in-context (KWIC) views present extracted query terms along with other kinds of information (such as document title) about a specific search result. Similar to this, a selected color facet can be shown as overlay on image search results.

The paper is organized as follows: In the next section we briefly describe related work, in Section 3 we discuss different image sources along with their problems and methods of resolutions. Section 4 describes the two main search strategies the user typically is faced with. In Section 5 we discuss some user interface enhancements before we summarize our work in Section 6.

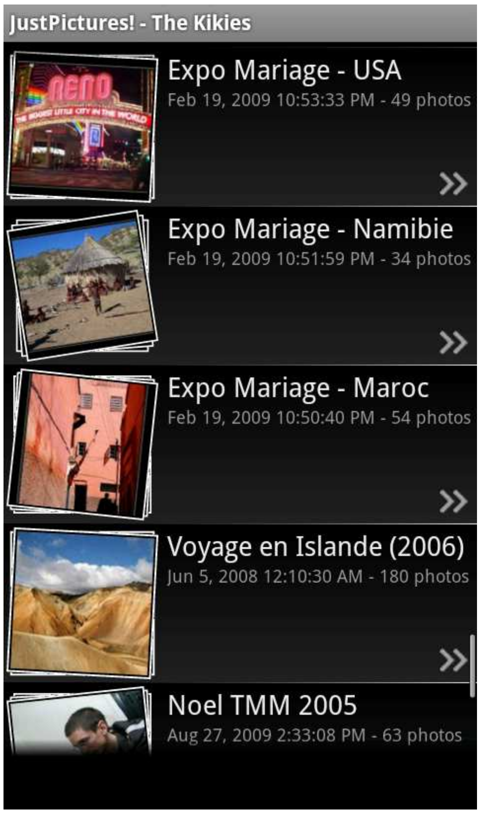

Figure 2: JustPictures in album view

2 Related Work

Most of the existing applications on mobile devices browse images just by folder, some let the user assign tags to images or classify them in albums and few let the user browse multiple photo albums from Facebook, Picasa, Flickr and your memory card.

The *JustPictures* [Quillard, 2010] application allows the user to browse the above mentioned photo albums on the web, it shows EXIF data, it automatically notifies the user of album updates, can handle authentication of users to access private albums and many more (see Figure 2 for the album view of this application).

But the user has still to know where to look for an image, as there is no search interface. JustPictures shows EXIF data for photos, but there is no possibility to browse the photos by one selected EXIF feature.

Another way to organize personal photo collections on mobile devices is using time information as main ordering criterion for visualization and interaction [Harada *et al.*, 2004]. But time information should not be the only aspect, since browsing and searching by facets can offer much more possibilities.

A multi-faceted image search and browsing system, named *Scenique* [Bartolini, 2009], allows the user to manage photo collections by using both visual features and tags, possibly organized into multiple dimensions (see Figure 3). But this solution is not designed for mobile devices, as it was made for desktop-PCs and therefore requires a screen with higher resolution.

Within this section, we only describe some of the existing applications that deal with the administration of image collections on mobile devices. Mor Naaman et al. give a more detailed overview in [Naaman *et al.*, 2008].

3 Feature Database

Our approach relies on extracted features, that are stored in some kind of database and are accessible through the mobile device. Obviously, every image needs to be processed with our feature extractor to make it findable on the mobile device. This is no problem for images on the internal memory, as their features can just be stored locally on the mobile device, after they have been extracted. But this could get a little bit more challenging for external image sources, like the web or social network sites. Therefore we need to differentiate between local images, images from friends (e.g. Facebook or Flickr) and external sources (e.g. Google images or any web page).

3.1 Server Database

If we want to use our faceted search for searching images on the Internet, the best solution would probably be a server, that stores the facet data. The server can crawl the Internet and extract the features for given images. The database could then contain all facet data along with a resized version of the image and an URL, where the original image can be found. The mobile device obtains this information from the server to offer the results to the user.

Another possibility could be the cooperation with photo album websites. Whenever users upload their images, the features could be extracted and offered via an interface to the mobile devices, to make the images easily searchable.

Obviously, both solution require server power, for extracting and storing the facets. To avoid these costs a peer-to-peer (P2P) solution may come in handy.

3.2 Peer-to-Peer Database

Another approach to store the facet data is to share it in a peer-to-peer environment. All users of the application can

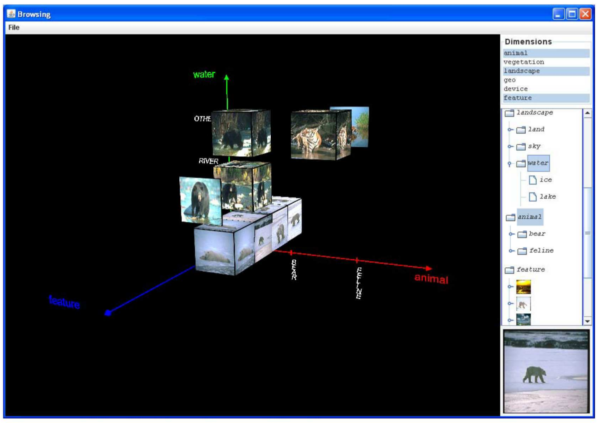

Figure 3: Scenique browsing interface

define their local photos as private, public or only viewable for friends. Depending on these settings and the relationship between the users they are able to browse public albums or photos of their friends using a P2P connection.

The process of feature extraction and the storage of facet data happens all on the mobile device. If a user starts the search for an image, the desired feature set will be exchanged using the P2P-network. Matching image results will then be forwarded to the searching user.

A JXTA-based implementation of a similar scenario was presented in [Müller *et al.*, 2007].

3.3 Local Database

If a user wants to use our faceted browsing application for browsing through her friend's photos, but her friends don't use the application or even don't have a mobile device, there is another possible solution. The user could simply mark some albums of her friends in our application as favorite albums, to indicate, that she wants to make these albums searchable via faceted browsing. Our application then processes these albums in the background, extracting the image features and stores them on the local memory. That way the user can search her own and the favorite albums of her friends with our faceted browsing application.

4 Search Strategies

When searching for images, there are generally two different use cases. First, recover images that the user has already seen before, e.g. photos taken with the camera of the mobile device, his own web albums or an image of a web album of any friend. Second, discover new images that the user hasen't seen before, e.g. images from the Internet or recently added photos of any web album of a friend. The search strategies for the two use cases may vary in detail.

4.1 Recover

Obviously the easier task is to recover an already seen image. The user may have a rough imagination of the image content, so we can use content-based image retrieval. As an example, the dominant color of a (region of an) image can help finding the image again. Other possibilities are the contrast of the image, texture or orientational features.

Any known metadata of an image can help finding it. EXIF features like the date or place the photo was taken, the camera model or if the photo was taken with or without flash allow simple browsing and searching. These facets can be used best for the user's own photos or for professional users.

If the user has his photos tagged, a tag based search can offer another possibility to search within photos. Due to the fact, that a lot of images remain untagged, this may be useful for only a small number of images.

Another option for searching images can be based on GPS data. If GPS data are available for a given set of photos, these photos can be placed on a map according to the place the photos were taken. In doing so, the user can easily search for photos of famous places via zooming and scrolling in a world map.

4.2 Discover

As a matter of principle, all approaches for recover images can also be used to discover unseen images. The expectation of the user regarding the resulting images is not as detailed given in this case. Hence, it depends on the individual use case, if it is an easier task to search for unseen images or not.

In addition the search for unseen images should support a key word query based on user input, as used for example on Google images. The result set of a key word query can then be further searched using our facets.

5 Search Interface

Regarding the search interface we want to describe in short how we could enhance the user interface and the presentation of results.

5.1 User interface

The search interface may be the most important part for searching and browsing through images. For a detailed illustration of our interface, see the prototype in [Hub *et al.*, 2009]. According to the user's feedback some improvements can be made. A lot of users asked for a timeline view to arrange the images chronologically. With individual time ranges this could indeed be a good way to search images, e.g. if the user wants to see the images of her last summer holiday, she can adjust the time range from August to September and only photos taken during that time will be shown.

Furthermore, our user experiments showed, that the number of color facets can be reduced. The fine grained color shades of our prototype were not used very often. We believe that ten to twelve main color shades are enough (see color picker at Google images). Other unused facets like the contrast for global an local regions can be reduced to only global contrast features.

Most of modern mobile devices are equipped with motion sensors, so it could be useful to use this for navigational purpose. We could imagine to scroll through a result list with pitching the mobile device. If the user shakes the device, the search could be reseted or the last step could be undone.

Alternatively multitouch gestures can be used for scrolling, zooming and navigation through search history.

5.2 Presentation of Search Results

Another very important aspect when searching for images is the presentation of search results. A simple list view is used most of the time, as usual with web search engines. Regarding text retrieval, theses lists can be enhanced with various improvements. There are summary information for every hit, query term highlighting, sparklines, preview of document content and many more (see [Hearst, 2009] for more information).

Some of the well-established techniques for text retrieval can be adjusted and adapted for image retrieval. For example, the global dominant color of an image can be represented with a colored frame around the image dyed in the corresponding color. This can also be used as overlay for some regions of an image, to indicate the selected color feature similar to query term highlighting.

The context of an image can also easily displayed. If the result image comes from a webpage or an web album, the previous and next image can be presented as thumbnail along with the result image. This gives the user a quick and small overview of the result and she can recognize, if there are more similar images in one continuous photo stream or if it is a collection of totally different images, what could come in handy for image search on web pages.

A totally different way to present the search result could be the arrangement of the result set as an three dimensional image globe. The images can be arranged on a virtual globe according to their dominant color using the HSV color model [Zhang *et al.*, 1999]. The north pole is for light images with high values and the south pole for dark images with low values. The hues are grouped around the equator and all images of the result set are positioned where they fit best. The saturation can be neglected in this model. The navigation could be done easily by pitching the mobile device, to let the image globe roll. For big result sets, the user can zoom in to show more or bigger images.

6 Conclusion

In this paper we described some ideas how image retrieval on mobile devices can be improved. Our first prototype should be enhanced with the presented features. We believe that this is a useful approach for finding images from webpages, online albums and internal storage.

Besides the realization of mentioned features in our prototype the search for partners could be very helpful, to include the faceted image search into existing photo sharing communities.

References

[Bartolini, 2009] Ilaria Bartolini. A mulit-faceted browsing interface for digital photo collections. *CBMI*, 7:237–242, 2009.

[Harada *et al.*, 2004] Susumu Harada, Mor Naaman, Yee Jiun Song, QianYing Wang, and Andreas Paepcke. Lost in memories: interacting with photo collections on pdas. In *JCDL '04: Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries*, pages 325–333, New York, NY, USA, 2004. ACM.

[Hearst *et al.*, 2002] Marti Hearst, Ame Elliott, Jennifer English, Rashmi Sinha, Kirsten Swearingen, and Ka-Ping Yee. Finding the Flow in Web Site Search. *Commun. ACM*, 45(9):42–49, 2002.

[Hearst, 2009] Marti A. Hearst. *Search User Interfaces*. Cambridge Univeristy Press, 2009.

[Hub *et al.*, 2009] Adrian Hub, Daniel Blank, Wolfgang Müller, and Andreas Henrich. Picadomo: Faceted image retrieval for mobile devices. *CBMI*, 7:249–254, 2009.

[Müller *et al.*, 2007] Wolfgang Müller, Soufyane El Allali, Daniel Blank, Andreas Henrich, and Thomas Lauterbach. Hunt the Cluster: A Scalable, Interactive Time Bayesian Image Browser for P2P Networks. *Multimedia Workshops, International Symposium on*, 0:317–322, 2007.

[Müller *et al.*, 2008] Wolfgang Müller, Markus Zech, Andreas Henrich, and Daniel Blank. VisualFlamenco: Dependable, Interactive Image Browsing Based on Visual Properties. CBMI International Workshop on Content-Based Multimedia Indexing, 2008.

[Naaman *et al.*, 2008] Mor Naaman, Rahul Nair, and Vlad Kaplun. Photos on the go: a mobile application case study. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 1739–1748, New York, NY, USA, 2008. ACM.

[Quillard, 2010] Gregory Quillard. Justpictures. http://code.google.com/p/justpictures/, 2010.

[Zhang *et al.*, 1999] L. Zhang, F. Lin, and B. Zhang. A CBIR Method Based on Color-Spatial Feature. *IEEE Region 10 Annual International Conference*, pages 166–169, 1999.