Context-based Image Annotation & Retrieval for Mobile Device

Ila Naresh Patil[#1]

M.Tech CSE [Pursuing]

Infotech Education Society, Bhopal
ilanaresh@rediffmail.com

--------------------------- ◆ ---------------------------

Abstract-In this paper, a novel method of context-based image annotation & retrieval on mobile device is presented by using context information. The method helps in retrieving image files in content-based searching methodologies. Specially, for mobile devices, manual annotation is expensive & time consuming task. Here, mobile contexts are used, such as photo captured context and social network context. In this paper, we considered these contexts and get useful semantic content of the photo files. The outcome of this is treated as annotation suggestions. Our annotations include time, location, event, persons etc.

Index Terms-Context, image, annotation, fusion, filtering, tagging, storage, retrieval

1. INTRODUCTION

There are two dimensions along which the different incentives are placed for tagging photographs [4]. The first dimension, "sociality," relates to whether the tag's intended usage is by the individual who took and uploaded the photo or by others, including friends/family and strangers. The second dimension, "function," refers to a tag's intended uses. The "function" dimension focuses on the motivation for adding tags. The social category of the "sociality" dimension is where tagging incentives that are not part of traditional personal media management are introduced.

The area of content annotation has been an important research area in the field of human computer interaction. It is possible to motivate users to annotate content. Point-of-capture annotation (e.g., on the mobile device) can encourage the addition of tags. Tag suggestions and other methods of assisting mobile annotation proved to have broader implications that just assistance in text-entry. Additionally, the suggestions can encourage users to tag their photos and give them assistance for how best to annotate.

Image annotation is a very important step in the process of CBIR (content-based image retrieval). It broadly considered both features of the image visual and text messages, which can improve the accuracy of the content-based image retrieval and make image search system more accurate when getting target image. To perform effectively, organization, searching and browsing scale images resources, various image retrieval technologies emerged, and received widespread concern.

Even though of the limitation of computing and interaction ability, there are some advantages on annotation with mobile device. Nowadays, mobile devices are context-aware. With the context; we can assign some annotations to the image. Digital photo files contains metadata themselves (e.g.EXIF). This metadata include the date and time when the photo is taken, and sometimes even the GPS location of the photo, all of which can be very useful to annotate.

The photo sharing on mobile phone can be used [7]. The system gives annotation suggestions based on web server computing and user interactivity. The main focuses are on collections of personal photos and exploit the contextual information naturally implied by the associated GPS and time metadata [8]. The annotation of photos is done based on collections rather than individual. The framework to annotate images using personal and social network contexts is used [9].

2. RELATED WORK

Image annotation & retrieval has been a very focused area of research nowadays. In image annotation & retrieval, text annotation plays an important role. Efficient tagging of photos has been an active research at the present time. In [1] the authors gave idea & a survey on context awareness. In [2], the authors focused on the various context aware computing applications. In [3], the authors focused on the method of image annotation and retrieval on mobile device. In [4], the motivation of annotation is studied. It explains the many ways to be taken into account while tagging

their photos. In [5] the authors gave a study report on image annotation techniques available

1. Making use of Textual Information.
2. Manual Annotation
3. Image Annotation Based On Ontology
4. Semi-automatic annotations on mobile photo using contextual information. In [6], a background photo annotation technique on mobile device is introduced by author. The spatial, temporal and public contexts of a photo are offered for automatic annotation. In [9], the author has used the model in which two types of annotation proposals are taken into consideration. It can be either for a particular user context or for a community network.

3. PHOTO CAPTURED CONTEXT

Whenever we take a photo, the camera will record the time. Generally, the time is recorded in the photo metadata (EXIF: Exchangeable image file format).
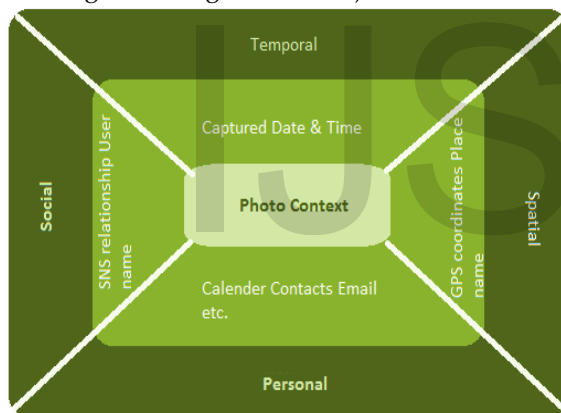

Fig 3.1 Photo Context

Similarly, [5] if the mobile device is prepared with a GPS receiver or other positioning system, it is location-aware. With the help of the GPS module, the GPS coordinates of the place where the photo is taken can also be recorded in EXIF of photo. Also when a photo is captured, date/time and GPS will be recorded. Therefore, we can get temporal and spatial context from photo metadata. Along with this data there is also a field in EXIF called "artist" which will record the photographer. Thus the photographer can also be recorded with some setting on the camera software. The figure 3.1 shows the photo context.

3.1 Personal Context

As mobile is a personal device, we can get additional user context information from mobile device. Calendar is a common service in mobile device for user making their timetable. So, when we take a picture at some time which is within a planned event, the picture is probably related to the event. Similarly, we can get more contexts from email, and even contacts. Event is the best semantic contents of a photo. When we can get event automatically, the annotation can become easier. When an event is planned, the photo taken at that time, it is probably the event is associated to the photo. Thus, it is logical to use the scheduled event as a suggestion of event annotation. In Image retrieval, annotation of person in a picture is very helpful. While it is impossible to automatically   identify person in the photo, we should manually annotate it. However, the contacts in the mobile device can be taken as candidates for annotations.

3.2 Social Network Context

As photo sharing is an important SNS (Social networking service) activity, social context is available in the social network based photo sharing system. The friendship between users is an implicit annotation of image. This annotation is added on the server side. Social relationship is very important for a photo sharing system. Here are several different ways to add friends in system:

1) Add by Name:
We can request somebody we know to be our friend by inputting his user name.

2) Add by Photo:
While browsing photos uploaded by other people if you want to make friends with the owner, you can send a friend request.

3) Import from Email Contacts: We can add friends by various different ways. Many emails service providers provide a service to retrieve contacts of your email. When a user wants to import friends, he has to input his email and password, the system gets contacts from email. Then the system compares these contacts with   users in our system, and recommends those users who are the user's contacts. The user can choose them to send friend requests.

4. CONTEXT ACQUISITION AND SENSING

Context acquisition is the most basic level of context awareness. In broad-spectrum, there are three ways to obtain context in the following

4.1 Sensed context: Environment information and user physical information can be acquired by physical or software sensors. And the user interaction habit and interactive historical record are gained mostly by logical sensor from the host.

4.2 Derived context: This kind of contextual information can be computed on the fly.

4.3 Context explicitly provided: For example, user's preferences can be obtained when they are explicitly communicated to the requesting application.

5. CONTEXT MODELING AND REPRESENTATION

In order to depict the user surroundings and situation, the system uses context modeling and representation in the context awareness computing. According to the current situation, the unified context modeling can be divided into two levels: Different context use the same data structure to express mode, such as Key-Value Models, Markup Scheme Models, etc. It also can support the semantic unity, such as ontology Based Models, Graphical Models, etc.

6. CONTEXT FILTERING AND FUSION

The raw data that we get after sensing is distinct, unstable and imprecise and it is necessary to efficiently use the context by filter and match conditions. The main purpose to filter is extracting simple semantic information from the sensor data, and protecting off the information which system does not care. The other one is filtering specific context information because of privacy.

To structure valuable and clear semantic context, it is required to develop the internal relation between context elements, and make a

fusion according to the specific application. There are varieties of ways in which the system can obtain the context information including physical sensor and logical sensor etc. but the information from the different variety of sources probably is distasteful, sometimes even contradictory. Therefore, context fusion is the request of maximum gain from the inconsistent information in the application based on certain knowledge and rules in order to avoid improper decisions by the system. After semantic relations fusion, context elements have more directly influence and effect for higher application and convert into higher level usable by applications, which is the vital context information.

7. CONTEXT STORAGE AND RETRIEVAL

The data collected during complete process i.e. raw context data as well as filtered and fused context information can be stored into some places for further retrieval by user to retrieve later. Context should be well structured into various data structures, such as table, object, tree, graphic, etc. Also, context storage is necessary to get history records in order to make processed context information out. The architecture for context storage can be either centralized or distributed.

From the perspective of using context information, context application generally can be divided into two kinds. For one thing, it focuses on physical context information, and the most common example is location based server. For the other thing, it can use the user context to the server, such as in the mobile computing the system can get the user preference and the user behavior information to provide with personalized recommendation, personalized retrieval information. Context information can be used for application more humanization and more convenient. Especially in the pervasive computing field, due to the users' demand and the rapidly changing environment, using context information is more value.

8. ANNOTATION PROCESS

As due to the limitation of input on mobile device, the operations of annotation process are reduced. For the

limitation computing ability of mobile phone, annotation algorithm has to be as simple as possible. The server based annotation suggestion system is not used in case the unstable wireless network environment. To give annotations to an image, follow these processes:

1. Metadata Analysis: Get metadata from the image, including time, GPS and artist etc;
2. Getting Personal Context: Use the time and GPS information to analyze useful information in personal utilities such as calendar, contacts and email, and the synthesized results are presented as annotation suggestions;
3. Tagging: Annotate image based on these suggestions and add other tags manually (Such as emotion).
4. Uploading: Upload the photo and the annotations to the server. The server stores the photo and the annotations and creates multidimensional indices for the photos.
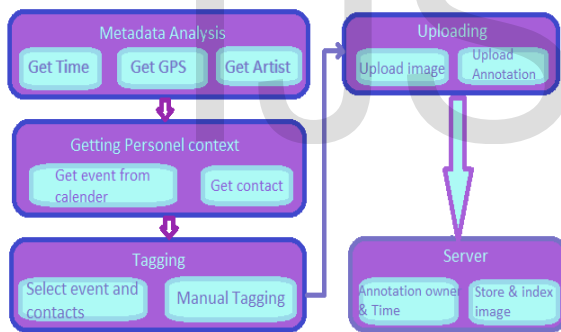

Figure 8.1 Annotation Process

When we select a photo to upload, the system reads date/time, GPS and photographer (artist) from the EXIF segment of the photo file. If it is failed in getting time, user can give the time manually. After the time is definite, it looks for the event which is programmed at that time in the calendar through some API (application programming interface) The search results are listed as event annotation suggestions.

After these processes, user can select the suggested annotations, or manually correct them. At the same time user can add some other annotation. Finally, we can set the range of the photo to be shared, which can be set to all, friends or owner.

1. Time Index: It is way to create index image on time, as people usually manage their images by time sequence.
2. Location Index: The GPS coordinates of images are disordered if they are directly display on the map. We use k-means algorithm to cluster the photos into several classes for different map zoom level.
3. Event Index: Event is an important semantic annotation for image, so search via event is common way.
4. User Index: Photos are uploaded by different people. People usually want to see their friend's photo.
5. Emotion Index: People may like a class of photos with come emotion. So one more index is created as emotion index.
6. Relevance Index: Some times, we want to retrieve images relevant to one image considering all aspect of the photo.
7. In the equation, a and b are two images, $St (a, b)$ is the time similarity of the two images, $Sl (a, b)$ is the location similarity of the location, $Se (a, b)$ is the event similarity of the event, and $wt + wl + we = 1$, $S(a,b)$ is the similarity of image a and image b.

$$S(a\ b) = wt\ St(a, b) + wl\ Sl(a, b) + we\ Se(a, b)$$

$$\dots\dots\dots\dots\dots\dots\dots(1)$$

8.1 Retrieval Model

Two kinds of photo browsing model are used: single dimension display and complex display query. These retrieval models are based on the indices we create.

In single dimension display model, we will treat an "album" as our clustered logical unit. With the annotations, we cluster our photos into different albums in different dimension such as time, location, event, SNS relationship and emotion.

In time dimension, we provide scalable granularity time range display. The albums are grouped by captured time of photo.
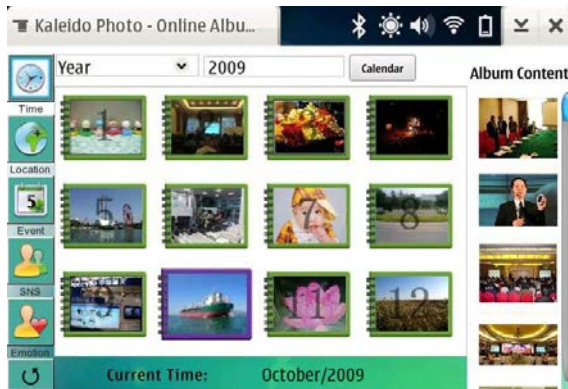
Figure 8.2 Album

Fig. 8.2 shows the 12 albums which sorted by 12 months in 2009. When we open an album, the photos in the album will be displayed on the right side of the screen. The photos can scroll in both direction and the photos are cyclically organized when they get the end, the first one appears again. We can browse each photo by scrolling the photo up or down by finger or stylus touch and we can select the right photo to have big photo view.

9. CONCLUSION

This is a semi-automatic way to add annotation to image by using contextual information of the mobile device. And it also provides a multidimensional way of image retrieval. This method reduces much work of manually input to simplify mobile image annotation. It provides good experience on image retrieval. Thus, it makes mobile photo sharing much easier.

10. REFERENCES

[1] Wei Liu , Xue Li, Daoli Huang "A Survey on Context Awareness" Computer Science and Service System (CSSS), International Conference on,29 June 2011 IEEE

[2] B. N. Schilit, N. Adams and R. Want, "Context-Aware Computing Applications",IEEE workshop on Mobile Computing Systems and Applications" December 8-9-1994.

[3] Shuangrong Xia, Xiangyang Gong, Wendong Wang,Ye Tia "Context-Aware Image Annotation and Retrieval on Mobile Device" 2010 IEEE.

[4] M. Ames, M. Naaman. "Why We Tag: Motivations for Annotation". In proc. CHI 2007, ACM Press (2007), 971-980.

[5] Reena Paranjape and Anita Shinde, "A Study of Recommender system Techniques" in IJECA [0975-888] volume 47-June2012.

[6] Lixing Jiang Jin Hou, Zeng Chen, and Dangsheng Zhang "Automatic Image Annotation Based on Decision Tree Machine Learning" 2009 IEEE.

[7] M. Davis, N. V. House, J. Towle, S. King, S. Ahern, C. Burgener, Perkel, M. Finn, V.Viswanathan, M. Rothenberg. "MMM2: Mobile Media Metadata for Media Sharing", Ext. Abstracts CHI 2005, ACM Press (2005), 1335-1338

[8] L. Cao, J Luo, H. Kautz, T.S. Huang. Image Annotation within the Context of Personal Photo Connections Using Hierarchical Event and Scene Models". In IEEE Multimedia 2009 11(2), 208-219.

[9] B. Shevade, H. Sundaram, L. Xie. "Modeling Personal and Social Network Context for Event Annotation in Images". In JCDL 2007, ACM press(2007).