

J_056

SemiANNOTATE: A Semi-automatic Approach to Personal Photo Album Annotation Based on Public and Personal Information

Supheakmongkol SARIN[†] Toshinori NAGAHASHI[‡] Tadashi MIYOSAWA[‡] and Wataru KAMEYAMA[†]

[†]GITS, WASEDA University
mungkol@fuji.waseda.jp
wataru@waseda.jp

[‡]Corporate R & D Div., SEIKO EPSON Corporation
nagahashi.toshinori@exc.epson.co.jp
miyosawa.tadashi@exc.epson.co.jp

1 Introduction

Annotation is the core item to enable efficient access to content. This is also the case of personal photo collections. Yet none of today's real-world applications responds ideally to this concern. None of them provides practical annotation tools that would help end-users to easily annotate thousands of their photos. In this paper, we propose a novel approach, SemiANNOTATE, by using public and personal information to extract contextual metadata.

Presently, all digital cameras provide time information and most camera-phones can infer rough location from GPS or CellID information. It is promising that all cameras will eventually be enable with location information capturing system. Moreover most digital photos contain location data in the form of coordinate set stored in their EXIF header [2]. As the result, all the photos will be embedded by two main key informations in their EXIF header namely, timestamp and location. According to [1], time and location information are the main factors in human episodic memory. SemiANNOTATE, our proposed engine, semi-automatically annotates the photos by using the time-location pair as pivot to leverage contextual metadata from public and personal repositories. In other words, SemiANNOTATE automatically propose contextual metadata candidates to users to describe the photos without any effort from users. In addition, users may be asked for confirmation, correction and augmentation of those proposed metadata candidates to increase the accuracy.

2 Related Work

Some recent works [7, 8, 10] enable digital photos to be annotated with metadata of spatial context using exact location information. Nevertheless, none of these systems explores the use of public and personal information as means to extract contextual metadata. All of them are rather limited to location/time metadata or use social collaboration as means to gather metadata. To the best of our knowledge, our approach described in this paper is the first one ever to be presented in the research.

3 The SemiANNOTATE Approach

3.1 Our Concept

In our concept, all the best contextual metadata about the photos comes from either the users' personal information or the surrounding environment of the users. Therefore, in our system, we specify the personal and public information repositories as the principle re-

sources to look for key metadata. Furthermore because a part of the metadata comes from the users' own perception about the photo via their personal information, this could eliminate the semantic gap as users might use the same keywords or concept to search or browse their photo album later. Thus search and browse processes would be improved significantly.

3.2 Automatic Metadata Candidates Generation and Validation

We have produced the prototype of SemiANNOTATE. The schema of our annotation process is illustrated in Figure 1. The processing steps can be described as the following:

1. **Time and Location as Keys:** We assume that the camera is GPS enable. Therefore all the photos captured should come with time and location informations. These two key informations will be used to query the public and personal repositories via their index. It should be noted that, in our prototype implementation, we assume that the location information is provided.
2. **Indexing and Searching:** We consider personal emails, schedules and diary as personal information and MSN-Mainichi Daily News [6] as our public resource. We use Apache Lucene [5] as indexing and searching tool of the two main resources.
3. **Metadata Candidates Extraction:** Information Extraction (IE) is composed of a number of analysis processes. This might take long time to compute if there are many relevant sources. As shown in the *Metadata Candidate Extraction* step in Figure 1, the refinement is done by limiting to 3 most relevant sources. This helps us improve the speed performance and the quality of the metadata to be generated. Afterward, the 3 most relevant sources will go through 2 different analyzing processes:
 - In (A), we apply the IE engine- A Nearly-New Information Extraction system (ANNIE)- to the sources in order to compute the name entities (NE) [4]. In our case, ANNIE is used to extract *Who, Organization, When, Where* information. In the example, after analyzing the 3 sources, we could obtain some NE by category (*Who, Organization, Where* and *When*).
 - In (B), we perform the potential keywords extraction. To realize this process, we first use Lucene to index the refined relevant sources. Then we do the ranking by the statistic of the frequency of the terms provided by Lucene.

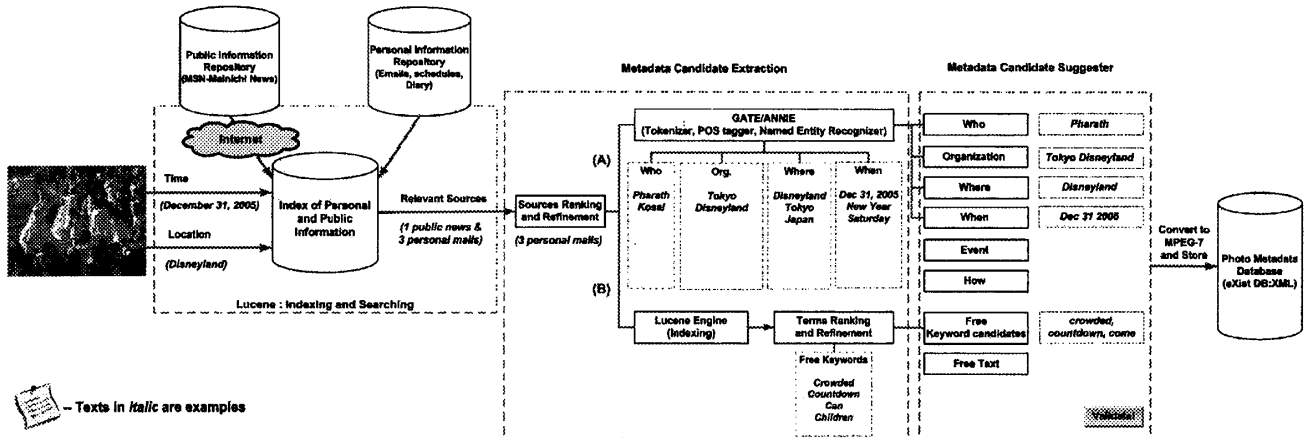


Figure 1: Processing steps in semi-automatic annotation of SemiANNOTATE, with example

4. **Metadata Candidates Suggester:** A subset of MPEG-7 has been selected as our metadata format. We extend from the StructuredAnnotation Basic Tool of MPEG-7 Multimedia Description Schemes (MDS) [9] by including the *Who*, *Organization*, *When*, *Where*, *Event*, *How* and *Free Keyword* fields. It aimed to be able to answer all the questions related to the photos and also to improve semantic integrity. All the metadata candidates (of the *Who*, *Organization*, *Where*, *When* and *Free Keyword* field) are presented to the user. For the *Event* field, default candidates are provided. Users can select one of them. For the other two fields (*How* and *Free Text*), users have choices to leave them empty or fill in manually. Users can make changes to the proposed metadata candidates if necessary.
5. **Validation and Metadata Storage:** Once users have validated the metadata, it will be converted to MPEG-7 format and will be sent to our metadata repository. We choose eXist [3] as our XML native database to store those metadata files.

4 Evaluations and Results

Experiments of our system were conducted. 5 users were recruited for the experiments. Each user contributes their photos and their emails and diary. We have a total of 112 photos. Users were asked to annotate each of their photos and fill in a 5-likert scale questionnaire. We evaluate the accuracy level of our proposed metadata as well as the satisfaction of SemiANNOTATE as an annotation tool by comparing with 3 real-world applications namely ACDSee Pro, Google Picasa and Adobe Photoshop Album.

Part (A) of Figure 2 shows accuracy level of our proposed metadata candidates. The result informs us that users find our proposed metadata candidates appropriate and acceptable. Part (B) depicts the satisfaction level of users by comparing our SemiANNOTATE engine with the annotation engine of the three others. The result shows that our proposed engine gains much satisfaction over the others. This concludes that our tool could help cope with the current annotation burden.

5 Conclusions

In this paper, SemiANNOTATE, a novel approach in generating contextual metadata for photos is pre-

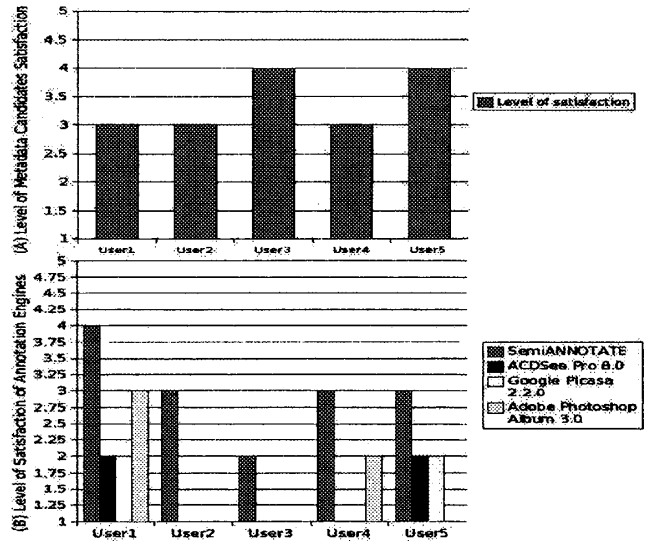


Figure 2: Satisfaction level results

sented. This proposed approach lessens much effort from the users. Users' verification on metadata is recommended by the system as an option to improve the accuracy. As a result, this eliminates the major effort users need to organize their large photo collection and enables better results for searching and/or browsing. Our preliminary experiments confirm the assumption and give us very encouraging results.

In the future, we would like to focus more on the IE part in order to achieve better metadata candidates.

References

- [1] C. C., O. M., and M. S. Using a time and location combination clustering model for browsing personal images. In *Proceedings of British HCI 2005, Edinburgh*, 2005.
- [2] Exif and related resources. <http://www.exif.org>.
- [3] Exist xml database. <http://exist.sourceforge.net>.
- [4] Gate and annie. <http://gate.ac.uk>.
- [5] Apache lucene. <http://lucene.apache.org>.
- [6] Mainichi news. <http://mdn.mainichi-msn.co.jp>.
- [7] M. Naaman, A. Paepcke, and H. Garcia-Molina. From where to what: Metadata sharing for digital photographs with geographic coordinates. In *10th CoopIS*, 2003.
- [8] M. Naaman, Y. J. Song, A. Paepcke, and H. Garcia-Molina. Automatic organization for digital photographs with geographic coordinates. In *Proceedings of the Fourth ACM/IEEE-CS JCDL*, 2004.
- [9] P. Salembier and J. Smith. *Overview of Multimedia Description Schemes and Schema Tools*. Addison-Wesley, 2001.
- [10] R. Sarvas, E. Herrarte, A. Wilhelm, and M. Davis. Metadata creation system for mobile images. In *Proceedings of the 2nd MobiSys*, pages 36-48. ACM Press, 2004.