Department of
Information Engineering
and Computer Science
DISI

UNIVERSITY
OF TRENTO - Italy

DISI - Via Sommarive 14 - 38123 Povo - Trento (Italy)
http://www.disi.unitn.it

PHOTO ANNOTATIONS: PRESERVATION OF CONTENTS AND CONTEXTS

Tabin Hasan and Fausto Giunchiglia

May 2012

Technical Report # DISI-12-021

Submitted at APCHI2012 (http://apchi2012.org/).

Photo annotations: Preservation of contents and contexts

Tabin Hasan
University of Trento
Via Sommarive, 5 I-38123 POVO
tabin@disi.unitn.it

Fausto Giunchiglia
University of Trento
Via Sommarive, 5 I-38123 POVO
fausto@disi.unitn.it

"We try to grab pieces of our lives as they speed past us. Photographs freeze those pieces and help us remember how we were. We don't know these lost people but if you look around, you'll find someone just like them."

-Gene McSweeney, *Grey Water Photography, 06-04-2006*

ABSTRACT

Photo is taken as memory of life and lives are the units of history. The value of photo fades away when the content becomes unrecognizable and context is lost over time. In this digital era, it has become easy to build contextual information for each photo with the process of annotating metadata. This paper proposes some aspects and techniques for semi-automatic photo annotation addressing several missing considerations to build a complete context that can survives over time. We also described the necessary metadata that unlocks hundreds of contextual information and proposes a model for photo labeling. The semantic organization and retrieval of photo collection is powered by hierarchical structure of Light-weight Ontology.

Author Keywords

Photo annotation, Active Learning (AL), Semi-supervised Learning (SSL), content based image retrieval.

INTRODUCTION

With the extensive use of digital cameras some metadata are acquired automatically, especially the timestamp. In our study, we have identified two key metadata that reveals many secrets of an event at which photo was shot – the time and the geo-reference, and two set of key elements – subject and the objects in the content. Both time [1] and location burst in clusters then recess for a long period of time and their relationship is always reciprocal.

In a personal photo collection, time and location is the primary metadata on which the person's timeline story begins to unfold. We remember events by their context and manual annotation of contextual information is frustrating for general users. We need more metadata to be annotated to make the collection worth of a person's lifetime.

Given the time and location, we can retrieve event information from a public event calendar and store new private events whenever necessary. Weather information as well as day light status can be retrieved from nearest weather station. Based on these metadata we can build a context rich archive for managing personal photos that should speak a story like "Photos of *last summer* in *Florence* of a *rainy morning*".

Figure 1: Recognition of contents and context. [Photo source: Personal collection]

There are various challenges to retrieve content based metadata and needs to adapt some intuitive user interaction to label the contents and other context information. Among other content data, the detection and recognition of persons in the photo discussed detailed in later sections.

RELATED WORKS

Though some high-end cameras support for GIS data annotation, Google map service or *World Wide Media eXchange (WWMX)* are easy to use tool for annotation [6]. Time and geo-tag, by using external resources, can be used to retrieve a meaningful name for the locations, daylight conditions, weather conditions, public and private event information. Image annotation with Photocopain [9] has used camera metadata, GPS, calendar data and more, missing weather data. Mor Naaman *et al* extended this in PhotoCompas for management of personal photos [3]. In a user study, their result shows that these derived metadata are extremely effective clues for photo retrieval. A work has been done by Neil O'Hare *et al* in their MediAssist system where they have extended with Content Based Photo Indexing for human faces and buildings [4]. Bongwon and Benjamin proposed used event based clustering and cloth based person recognition [5].

Labeling with Computer game is a popular approach in object labeling in the photos [7]. In a Personal Photo Management Framework, Sama Cooray *et al* proposes a person identification method by their re-occurrence in the images [10]. Cha Zhang *et al* used Active Learning approach for CBIR [8].

Anastasia Krithara *et al* demonstrated a better learn-ability by combining Active Learning and SSL [11].

CONTEXT BASED METADATA ANNOTATION

Our proposed approach based on many real life scenarios and we will proceed from simple to complex one.

Fixing time

Digital cameras automatically add current time in file header. Usually a person adjust time of a camera only once after purchase and that is local. When the person travels to a different time zone, our solution to this problem is to calculate the offset from the geo-references between the locations and readjust the time. It can be done automatically by the application once geo-reference is annotated.

Tagging geo-reference

The proposed application may use Google Map or Microsoft MapPoint programmable interface to get the geographic coordinate. This extra work would not have been necessary if the camera would have been equipped with GPS. But geo-reference does bear any meaning to a user unless the location has a name. In many cases gazetteer like Alexandria [12] could be useful to retrieve geographic location names out of geo-reference. The drag and drop method works well for tagging geo-reference.

Event calendar

Our proposed system allows user either to retrieve event information from external sources or from his/her personal chronological data, e.g., anniversaries.

Weather data

Given the timestamp and location stamp, system can add all available weather information. For Trento the weather station is Paganella and the system queries necessary weather information and interprets with meaningful annotations such as cloudy or clear, snowy or rainy.

CONTENT LABELING

There are many theories and solutions exist for recognizing person or object in an image, but labeling them is a still very labor intensive. “Learning by Example” [13] considered to be an adaptable means for recognizing objects in the photos. As described in [10] BDF (Bayesian Discriminating Feature) method for face detection, and conjunction with Color Coherent Vector (CCL) and Color Correlogram for body patching proved effective. However, we consider AL method in conjunction other promising methods so far being described.

Active Learning cycle

We follow a little different AL process in selection of training data set. Instead of choosing a training set randomly and/or by users selection, we propose an approach by maximum content detection for training set. This set of photos becomes the primary dataset to be labeled first by user interaction.

Mattheiu Cord *et al* proposed an image retrieval strategy called RETIN that is organized in two-stage sequential process [14].

ORGANIZING COLLECTIONS

A classification scheme of Lightweight Ontology developed by Fausto Giunchiglia *et al* provides the state-of-the-art hierarchical content management framework for the application [2]. The whole collection of photos can be organized, retrieved and shared by the annotations and their semantic relations.

CONCLUSION

The initial proposal for the Personal Photo Management System amassed some possible scenarios to comprehend personal needs of preserving the memories of a person’s life. We tried to exploit all related technologies ranging from Human Language Technology to CBIR usable from a single application to best support the memory preservation and retrieval. We are still studying all other necessary attributes to make a real complete “Story of a Life” and as a consequence of this effort they will well be added in our future works. The work has only been partially implemented.

REFERENCES

1. Graham, A., Garcia-Molina, H., Paepcke, A. and Winograd, T. 2002. Time as essence for photo browsing through personal digital libraries. In JCDL, pages 326-335.
2. Giunchiglia, F., Marchese, M., Zaihrayeu, I. 2007. Encoding Classifications into Lightweight Ontologies.
3. Naaman, M., Harada, S., Wang, Q. Y., Garcia-Molina, H. and Paepcke, A. 2004. Context data in geo-referenced digital photo collections. In ACM Multimedia, pages 196-203.
4. O'Hare, N., Gurrin C., Jones, G. J. F., Lee H., O'Connor, N. E., and Smeaton, A. F. 2007. Using text search for personal photo collections with the mediassist system. In SAC, pages 880-881, 2007.
5. Suh, B. and Bederson, B. B. 2007. Semi-automatic photo annotation strategies using event based clustering and clothing based person recognition. Interacting with Computers, 19(4):524-544.
6. Toyama, K., Logan, R., and Roseway, A. 2003. Geographic location tags on digital images. In ACM Multimedia, pages 156-166.
7. Ahn, L. V., Liu, R., and Blum, M. 2006. Peekaboom: a game for locating objects in images. In CHI, pages 55-64.
8. Zhang, C. and Chen, T. 2002. An active learning framework for content-based information retrieval. IEEE Transactions on Multimedia, 4(2):260-268.
9. Mischa, M. Tu_eld, Harris, S., Dupplaw, D. P., Chakravarthy, A., Brewster, C., Gibbins, N, O'Hara, K., Ciravegna, F., Sleeman, D., Shadbolt, N. R. and Wilks, Y. 2006. Image annotation with PhotoCo-pain, Proceedings of the First International Workshop on Semantic Web Anno-tations for Multimedia (SWAMM) (Edinburgh).
10. Cooray, S., O’Connor, N. E., Gurrin C., Jones, G. J. F., O’Hare, N., and Smeaton, A. F. 2006. Identifying Person Re-occurences for Personal Photo Management Applications.
11. Krithara, A., Goutte, A., Amini, M. R. and Renders, J. M. 2007. Active, Semi-Supervised Learning for Textual Information Access, XEROX Research Centre Europe.
12. Hill, L. L., Frew, J., and Zheng, Q. 1999. Geographic names – the implementation of a gazetteer in a georeferenced digital library. CNRI D-Lib Magazine.
13. Sabera, E., Xu, Y. and Tekalp, A. M. 2005. Partial shape recognition by sub-matrix matching for partial matching guided image labeling.
14. Cord, M., Gosselin, P. H. and Philipp-Foliguet, S. 2006. Stochastic exploration and active learning for image retrieval.