PERSONAL ARCHIVING AND RETRIEVING IMAGE SYSTEM (PARIS)

Pei-Jeng Kuo, Terumasa Aoki and Hiroshi Yasuda
University of Tokyo, Research Center of Advanced Science 4-6-1 Komaba, Meguro, Tokyo, Japan

Keywords: Ontology, Databases and Information Retrieval, MPEG-7, Spatial-Temporal, Digital Library Designs l, metadata, Semantic Web, semi-automatic annotation.

Abstract: In previous publications, we have proposed the DDDC architecture which annotates multimedia data with twelve main attributes regarding its semantic representation. In addition, we also proposed machine-understandable "Spatial and Temporal Based Ontologies" representation for the above DDDC semantics description to enable semi-automatic annotation process. PARIS (Personal Archiving and Retrieving Image System) is an experiment personal photograph library, which includes more than 80,000 of consumer photographs accumulated within a duration of approximately five years, metadata based on our proposed MPEG-7 annotation architecture, Dozen Dimensional Digital Content (DDDC), and a relational database structure. In this paper, we explain our proposed spatial and temporal coordinated Ontologies constructed based on personal related photograph preferences. As personal digital photograph libraries have specific characteristics and are particularly Spatial and Temporal associated, we envision various novel retrieval possibilities at semantic level can be developed based on the proposal described in this paper.

1 INTRODUCTION

With the proliferation of image devices such as digital video cameras, digital cameras, and mobile phone cameras, an increasing number of users are building up their private multimedia repositories. At the same time, the fast growing of hard drive capacity and digital image device resolution also results in an ever increasing size of personal multimedia collections. Currently, people who regularly take digital photographs, from general users, amateur photographers to professional photographers, are more or less accumulating their own photograph collections one way or another. We can see the trend that the number of photographs taken would soon exceed the critical amount for being simply manageable.

While immeasurable amount of multimedia information is accumulating in digital archives, from mobile terminals, on the web, in broadcast data stream and in personal and professional database, the value of the information depends on how easily we can manage, find, retrieve, and filter it.

Because of the pervasive of consumer imaging devices, building personal digital photograph libraries became an increasingly interested domain. Personal digital photograph collections have specific characteristics compare to general purpose image databases. Hence, annotation architecture specially designed for that plays an important role in building an interoperatable data repository for future indexing, browsing and retrieving purposes. We propose a MPEG-7 based multimedia content description architecture, Dozen Dimensional Digital Content (DDDC), which annotates multimedia data with twelve main attributes regarding its semantic representation. In addition, we also proposed a machine-understandable "Spatial and Temporal Based Ontology" representation for the above DDDC semantics description to enable semi-automatic annotation process.

The special address on integrated spatial and temporal features in terms of consumer photograph database embraces different approaches compare to traditional general purpose image database retrieval researches. PARIS (Personal Archiving and Retrieving Image System) is an experiment personal photograph library constructed according to our proposed semantic annotation structure.

Kuo P., Aoki T. and Yasuda H. (2007).
PERSONAL ARCHIVING AND RETRIEVING IMAGE SYSTEM (PARIS).
In *Proceedings of the Third International Conference on Web Information Systems and Technologies - Web Interfaces and Applications*, pages 347-350
DOI: 10.5220/0001289403470350

2 PROBLEMS WITH PERSONAL IMAGE ARCHIVING

While an increasing amount of people are building their online photo albums with the aid of off the shelf digital album tools as well as web album hosting sites, an effective and semantic way of retrieving context relevant images from the large repository of personal digital archives has yet appeared.

Previously, only commercial photograph database can reach this scale. Hence little research effort has been focused on large scale personal photograph collections. At current stage, general users tend to give simple text annotations based on folders of photograph categorized by related subject, event or date. But we can see a trend that the number of digital photographs taken would soon exceed the critical amount for being simply manageable.

Without constant organizing efforts, allocating a single photograph such as: "A photograph of mine that was taken around 5 years ago with three other friends in a beautiful coffee shop near the bank of Seine River at Paris during my 10 day summer vacation in Europe" would soon become a difficult task similar to look for old time photos stacked on shoeboxes. And it is almost impossible for any individual to annotate each of their photographs manually. (Platt , J. et al., 2002 ; Graham, A et al., 2002)

Most people make more photographs while they visit some new locations or during special events. As a result, the spatial and temporal attributes of personal digital photographs could contain some very relevant context information which was not addressed by most general purpose image archive researches.

The burst structure within collections of personal photographs tends to be recursive, while small bursts exist within big bursts as shown in Figure 1. And this recursive structure can be represented as a cluster tree, where photographs are stored only at the leaf nodes. (Graham, A et al., 2002)

In previous researches (Platt , J. et al., 2002 ; Graham, A et al., 2002; Stent, A. et al., 2001), semi-automatic event segmentation based on the recorded time tags made possible by most recent image devices were enabled. While our research emphasizes the importance of an integrated approach utilizing spatial information in addition to the ready-to-retrieve temporal information, we have designed metadata description architecture, DDDC (Dozen Dimensional Digital Content) (Kuo, P., Aoki, T. and Yasuda, H., 2004), extended from MPEG-7 multimedia description schema for annotating personal digital assets.

We also proposed the concept of "Spatial and Temporal Based Ontology" (Kuo, P., Aoki, T. and Yasuda, H., 2004), constructed based on the special pattern of personal photograph collections as we argue that time and location are two most important attributes in terms of personal photograph retrieval.

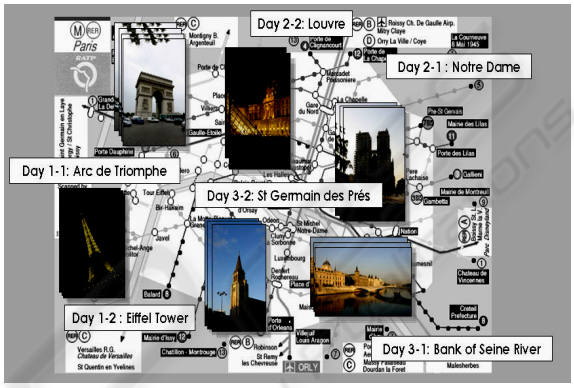

Figure 1: Personal Digital Photograph Clusters Tree Structure.

3 SPATIAL AND TEMPORAL COORDINATED ANNOTATION

Digital camera has become popular in the past few years, which means, a majority of people have only digital photograph collections accumulated within a limited time span of a few years. If we envision a continuous lifetime digital photograph archiving process, one might recalls, for example, his or her several trips to Paris a few years ago and one trip to Tokyo during Christmas season, but could not clearly memorize the exact year or dates.

Rodden, K., and Wood, K. talked about the two most important features of an efficient, reliable and well-designed system for managing personal photo-graphs are: automatically sorting photos in chronological order, and displaying a large number of thumbnails at once. While people are familiar with their own photographs, laborious and detailed keywords annotating are not specifically motivated for most people.

In addition to time, location has been argued as one of the strongest memory clues when people are recalling past events. In analyzing our prototype database raw data, we also find that most folder names contain words related to geographic

information such as “Asakusa”, “Yokohama”, “Boston”, NYC”, “Fukoka”, and “Paris”, as well as temporal information such “France National Day”, “Christmas”, or “ Winter”. All of the folder names start with temporal information and 673 out of the 711 folders contain geographical related words, which is about 93% in percentage.

While our prototype system only contains images from a single person, we also conducted oral interviews with morn than 10 active digital image users to ensure an unbiased presumption towards common user naming convention. The interview results show that all of the interviewees store their digital photographs in different folders roughly named by location, event, and time period, for example, “A trip to Europe_June_2004”, “Birthday Party with Nico_20040125”, etc.

According to the above discussion, we see a strong association between the image context and its respective spatial and temporal clues for personal digital photograph collections. With the aid of recent GPS receiving devices, it is also possible to embed geographic information in GPS format onto digital image data. Although the recorded device GPS information does not guarantee the exact location of subjects appeared in the photographs as we had discussed in our previous publications, utilize GPS information as a retrieving attribute greatly reduce the human effort compare to manual input.

Figure 2 and 3 shows examples of our proposed spatial coordinated Ontology and Temporal Coordinated Ontology. In Figure 2, we can see that from year 2001 to 2006, groups of photographs were taken at major period of times, such as: the first few days of a year, the season of cherry blossoming, days around the kichijouji autumn festival, the season of foliage, etc..

While those groups of photographs do not exist in every single year, most of the groups happened more than twice or three times over the past six year. This is around 33%~50% in ratio. The repeating pattern of similar topics or subjects taken around similar time in a year reveals the photographer’s personal history and preference on photographing topics.

In Figure 3, temporal coordinated events which are important to the specific photographer can be found in the burst photograph structure as we discussed in previous section.

In PARIS system, we annotate our archived photographs with sets of Spatial and Temporal Coordinated Ontologies.

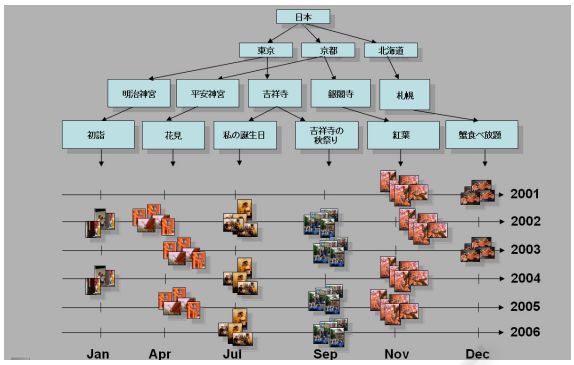

Figure 2: Example of Proposed Temporal Coordinated Ontology.

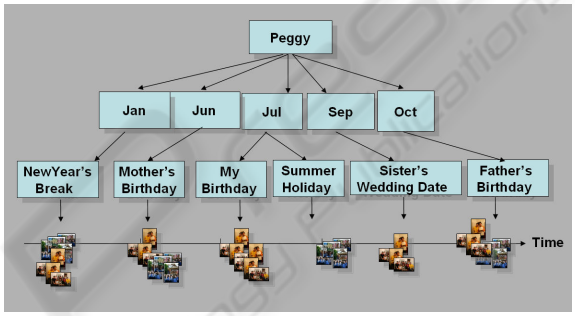

Figure 3: Example of Proposed Temporal Coordinated Ontology.

4 THE SYSTEM

Extended from the StructuredAnnotation Basic Tool of MPEG-7 Multimedia Description Schemes (MDS), we propose a semantic description tool of multimedia content. The proposed content description tool annotates multimedia data with twelve main attributes regarding its semantic representation. And the twelve attributes include answers of who, what, when, where, why and how (5W1H) the digital content was produced as well as the respective direction, distance and duration (3D) information. We also define digital multimedia contents including image, video and music embedded with the proposed semantic attributes as Dozen Dimensional Digital Content (DDDC). Due to limited space, detailed explanation and example codes can be found in our previous publications.

PARIS (Personal Archiving and Retrieving Image System) provides an image annotation methodology based on our proposed MPEG-7 annotation architecture, Dozen Dimensional Digital Content (DDDC). In annotating process, we also utilize a proposed Spatial and Temporal Ontology

(STO) designed based on the general characteristic of personal photograph collections.

To minimize the frustration of tedious manual inputs, we propose to build location specific "Domain Ontology" for popular tourist stops such as the city of Paris, Tokyo and New York based on their respective spatial and temporal attributes. Figure 4 shows the prototype structure of PARIS system.

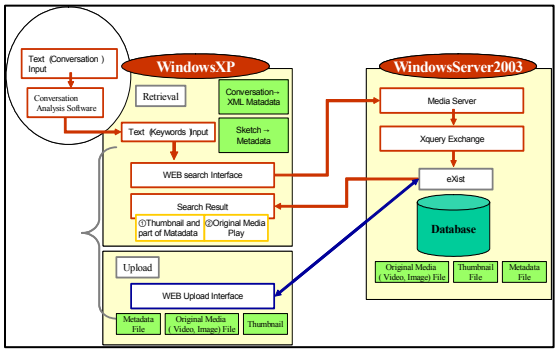

Figure 4: PARIS System Prototype Structure.

5 CONCLUSION AND FUTURE WORKS

In PARIS system, we try to construct a semi-automatic spatial and temporal coordinated personal photograph database with the following steps:

1. Construct common annotation architecture for building personal digital photograph libraries with our proposed "Dozen Dimensional Digital Content (DDDC)" architecture.

2. Construct a machine-understandable "Spatial and Temporal Based Ontology" with our proposed "Spatial and Temporal Based Ontology (STO)" representation for the above DDDC semantic description thus give a common background towards personal photograph annotation.

3. Enable semi-automatic annotation process according to previous annotated photographs. When recurrent patterns within collections of personal photographs can be found, a semi-automatic annotation process is possible base on similar spatial and temporal features.

As personal digital photograph libraries have specific characteristics and are particularly Spatial and Temporal associated, we envision various novel retrieval possibilities at semantic level can be developed in our future works.

REFERENCES

Platt , J. et al., 2002. PhotoTOC. Microfort TR.

ISO/IEC 15938-5, 2001. Multimedia Content Description Interface – Part 5, version 1.

Kuo, P., Aoki, T. and Yasuda, H. 2003. Semi-Automatic MPEG-7 Metadata Generation of Mobile Images with Spatial and Temporal Information in Content-Based Im-age Retrieval. Pro. SoftCOM 2003.

Kuo, P., Aoki, T. and Yasuda, H. 2004. MPEG-7 Based Dozen Dimensional Digital Content Architecture for Semantic Image Retrieval Services. Proc. EEE-04.

Soo, V. et al., 2003. Automated Semantic Annotation and Retrieval Based on Sharable Ontology and Case-based Learning Techniques. Proc , JCDL'2003.

Schreiber, T. et al., 2001. Ontology-Based Photo Annotation. IEEE Intelligent Systems.

Graham, A et al., 2002. Time as Essence for Photo Browsing Through Personal Digital Libraries. Proc.,JCDL'02.

Hyvönen, E. et al., 2002. Ontology-Based Image Retrieval. Proc. of XML Finland 02.

Stent, A. et al., 2001. Using Event Segmentation to Improve Indexing of Consumer Photographs. Proc. ACM SIGIR'01.

Jaimes, A. et al., 2003. Semi-Automatic, Data Driven Construction of Multimedia Ontologies. Proc.ICME 03.

Noy, N. et al., 2001. Ontology Development 101. SMI TR, 01.

Hunter, J., 2003. Enhancing the Semantic Interoperability of Multimedia Through a Core Ontology. IEEE Transac-tions on Circuits and Systems for Video Technology.

Kuo, P., Aoki, T. and Yasuda, H. 2004. Building Personal Digital Photograph Libraries: An Approach with Ontology-Based MPEG-7 Dozen Dimensional Digital Content. Proc. Computer Graphics International-2004.

Rodden, K., and Wood, K., 2003. How do People Manage Their Digital Photographs? ACM Conference on Human Factors in Computing Systems.