J-065

MPEG-7 Based Personal Digital Photograph Library System

Pei-Jeng KUO   Terumasa AOKI   Hiroshi YASUDA

Yasuda-Aoki Laboratory, The University of Tokyo

Research Center for Advanced Science and Technology

4-6-1 Komaba, Meguroku, Tokyo, 153-8904 JAPAN
E-mail: {peggykuo, aoki, yasuda}@mpeg.rcast.u-tokyo.ac.jp

Abstract: **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.**

Keywords:

Ontology, Digital Image Database, MPEG-7, Spatial-Temporal, image retrieval, metadata, Semantic Web, semi-automatic annotation

1. INTRODUCTION

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.

Two approaches have been studied in the research community:

1. Content-Based Image Retrieval (CBIR)
2. Metadata-Based Image Retrieval

Currently, two problems hamper the development of "semantic" level image retrieval given the availability of carefully annotated external metadata :

1. There is lack of common annotation architecture for personal digital image library.

2. Annotations require domain knowledge.

We try to tackle the above two problems with the following steps:

1. Construct common annotation architecture for building personal digital photograph libraries –We proposed The "Dozen Dimensional Digital Content (DDDC)" architecture extended from MPEG-7 Multimedia Description Scheme.

2. Construct a machine-understandable "Spatial and Temporal Based Ontology" representation for the above DDDC semantic description to enable semi-automatic annotation process.

2. PERSONAL DIGITAL PHOTOGRAPH LIBRARIES

2.1 Burst Structure

Consumer photograph collections are different from most general purpose image databases. People tend to take personal photographs in bursts. Which means, a group of photos may be taken for a semantic related event, but a few, if any, photos may betaken until another significant event started. Figure 1 shows a time and location related burst tree structure in a demonstrative "Three Day Trip in Paris" personal photograph collection.

2.2 Proposed MPEG-7 Based Spatial and Temporal Retrieval

In addition to "time", the temporal element, we argue that "location", the spatial element, plays an equally important role as hint to the semantic context of personal photograph collections. One of the most interested topics for personal
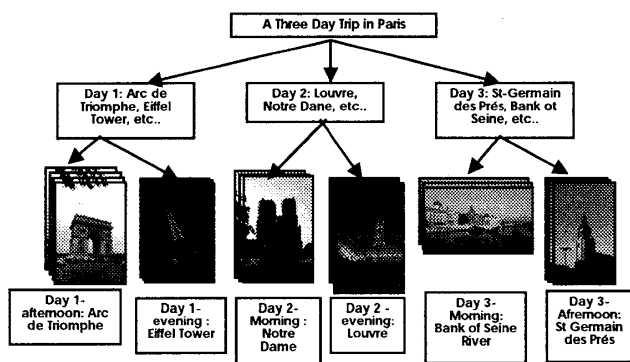
Figure 1 Personal Digital Photograph Clusters Tree Structure

photographs is tourist photographs. People tend to take a great number of photographs especially during their trips to a new place. To enable a spatial and temporal based personal digital photograph retrieval system, we adopted the MPEG-7 standard as the basis of our proposed annotation architecture, Dozen Dimensional Digital Content (DDDC), which will be explained in the next section.

3. DOZEN DIMENSIONAL DIGITAL CONTENT (DDDC) ARCHITECTURE

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. 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 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 [1].

4. SPATIAL AND TEMPORAL BASED ONTOLOGY

4.1 Ontology-Based Photograph Annotation

Given the above DDDC architecture, we provide a structured methodology to annotate most significant, if not explicit, semantic answers of personal digital photograph collection contexts. In [2], the idea of Ontology-Based Photo Annotation was described. An ontology is a formal, explicit specification of a domain. Typically, an ontology consists of
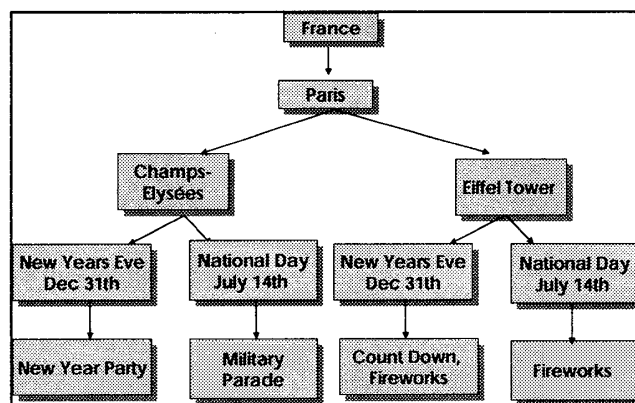

Figure 2 Figure 2 Concept of Proposed Spatial and Temporal Ontology

concepts, concept *properties*, and *relationships* between concepts. Ontology concepts are represented by terms, which can help the user in formulating the information needed, the query, and the answers.

4.2 Our Proposed Spatial and Temporal Ontology

Figure 2 illustrates an example of our proposed Spatial and Temporal Ontology built for the city of Paris. In building Spatial Ontology, we firstly separate Paris into several popular tourist districts such as "The Latin Quarter" and "St-Germain des Prés". Under each district, we again separate it into sub-districts or point of interests. The construction of Temporal Ontology requires more domain knowledge of the specific location. We suggest building up the location specific Temporal Ontology according to the photographer's personal interest and experience as well as with the aid of third party databases.

5. CONCLUSION

As personal digital photograph libraries have specific characteristics and are particularly Spatial and Temporal associated, we envision various novel browsing possibilities at semantic level can be developed based on the proposal described in this paper.

REFERENCES

[1] P. J Kuo, T. Aoki and H Yasuda, "MPEG-7 Based Dozen Dimensional Digial Content Architecture for Semantic Image Retrieval Services", *Proc. EEE-04*, Mar 04.

[2] A. T. Schreiber et al., "Ontology-Based Photo Annotation", *IEEE Intelligent Systems*, May 01.

[3] A. Graham et al., "Time as Essence for Photo Browsing Through Personal Digital Libraries", *Proc.,JCDL '02*, Jul 2002.