

Geo-referenced Digital Photograph Library

Pei-Jeng Kuo, Terumasa Aoki and Hiroshi Yasuda
The University of Tokyo
Tokyo, Japan

Abstract

We have proposed a MPEG-7 based annotation architecture, Dozen Dimensional Digital Content (DDDC), specially designed for facilitating the managing, browsing and retrieving of personal digital photograph collections. In this paper, we explain the concept and structure of our MPEG-7 based geo-referenced digital photograph library prototype system, named PARIS (Personal Archiving and Retrieving Image System). In addition to utilizing the temporal information recorded directly from image devices, our approach integrates spatial coordinates to realize novel retrieving methodology with geo-referenced photograph collections.

Introduction

We argue that people make more photographs while they visit some new locations or during special events. Hence spatial and temporal attributes of personal digital photographs could contain the most relevant context information. While personal photograph collections have very different characteristics compare with traditional commercial stock image collections, which were normally used for image retrieval researches, organizing, archiving, and retrieving personal photograph collections also require different approaches.

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. This paper explains our approach in building up a geo-referenced photograph library system.

Geo-Referenced Annotation

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 enables geo-referenced multimedia data annotation 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). Figure 1 shows an example of

our proposed MPEG-7 DDDC annotation code. Due to limited space, detailed explanation and more example codes can be found in Ref. [1].

```
<TextAnnotation id="Ann1">
  <FreeTextAnnotation xml:lang="en">
    2001_07_04_People and Eiffel Tower in Paris on the
    National Day of France
  </FreeTextAnnotation>
  <StructuredAnnotation>
    <Who>
      <Name xml:lang="en">People
      </Name>
    </Who>
    <What>
      <Name xml:lang="en">
        Eiffel Tower
      </Name>
    </What>
    <Where>
      <Name xml:lang="en">
        Paris; France
      </Name>
    </Where>
    <When>
      <Name xml:lang="en">
        2001-07-14T20:17+01:00;
        National Day of France
      </Name>
    </When>
  </StructuredAnnotation>
</TextAnnotation>
```

Figure 1. MPEG-7 DDDC Metadata Example

Geo-Referenced Ontology

The above DDDC architecture provides a structured methodology to annotate most significant, if not explicit, semantic answers of personal digital photograph collection contexts. However, some of the DDDC annotations such as the free text part of who, where and what attributes still require manual inputs. In Ref. [3], several difficulties have been pointed out in terms of the annotation process. First, different annotator might use different terms to annotate the same concept. Second, the users who do not have specific domain knowledge might not be able to input the right keywords or natural language query for semantic image retrieval. And third, the manual annotation of a large amount of personal digital photograph collections, if not impossible, is a laborious task.

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. Illustrations in Fig. 2 and Fig. 3 present example of geo-referenced hierarchical ontology for the city of Paris. More information for our proposed geo-referenced ontology is explained in Ref. [2].

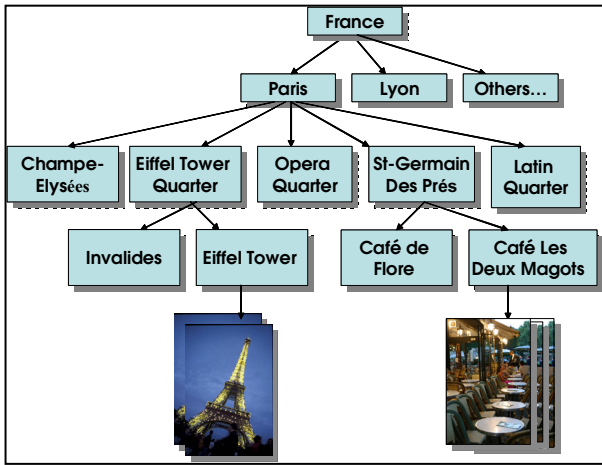


Figure 2. Concept of Proposed Spatial Ontology

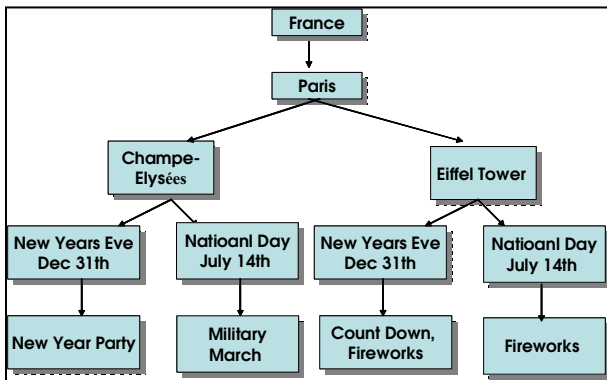


Figure 3. Concept of Proposed Spatial and Temporal Ontology

Conclusion

PARIS system is build in conjunction with a main project called AVR (Advanced Video Retrieval), funded by Japanese Government. The AVR project explores novel methodologies for multimedia content retrieval including video, audio and image. Within AVR project, the PARIS system targets on geo-referenced personal photograph collection retrieval in particular. Figure 4 illustrates the

structure of AVR testbed. The PARIS system contains around 80,000 photographs with a time span of around five years. At current stage, around 10,000 of them are geo-referenced and hence related spatial and temporal based annotations are generated based on our “Spatial and Temporal Ontology”.

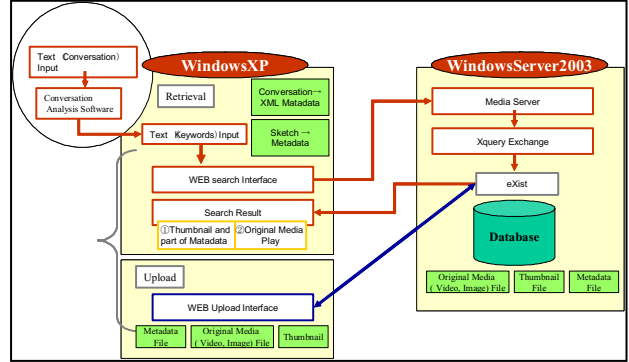


Figure 4. Testbed Structure used for constructing PARIS System

References

1. P. J Kuo, T. Aoki and H Yasuda, MPEG-7 Based Dozen Dimensional Digital Content Architecture for Semantic Image Retrieval Services, In Proc.IEEE EEE-04 ,Mar 04.
2. P. J Kuo, T. Aoki and H Yasuda, Building Personal Digital Photograph Libraries: An Approach with Ontology-Based MPEG-7 Dozen Dimensional Digital Content. In Proc. IEEE Computer Graphics International-2004 ,Jun 04.
3. A. T. Schreiber et al., “Ontology-Based Photo Annotation”, IEEE Intelligent Systems, May 01.

Biography

Pei-Jeng Kuo received her B.S. in Physics from National Taiwan University, Taiwan and M.E. from Massachusetts Institute of Technology, Cambridge, USA. She has interned with the Ericsson Eurolab Multimedia Application Research Department, French Telecom Creative Research Lab and NTT DoCoMo Wireless Laboratory. She is currently a Ph.D candidate in the University of Tokyo and her research topics include the indexing, archiving, delivering, and retrieving of multimedia contents with MPEG-7 technology. {peggykuo,aoki,yasuda}@mpeg.rcast.u-tokyo.ac.jp.