

論文の内容の要旨
Abstract of Dissertation

論文題目
Title of Dissertation

Continuous Personal Digital Photograph Archival and Retrieval with Spatial and Temporal Coordinated MPEG-7 Description Scheme

(時空間情報MPEG-7記述スキームを有する
個人向けデジタルフォトアーカイブならびに 検索技術)

氏 名 郭 正 佩
Name of Author

(本文) (Abstract)

A Picture is worth a 1,000 words. How about 1,000 pictures? Or 1,000,000 pictures?

While taking photographs with digital cameras has become easier than ever, any user can generate thousands of photographs conveniently and inexpensively. In addition to the widely available digital cameras, the mobile image services firstly introduced in Japan several years ago also soared in popularity. The introduction of compact sized digital cameras together with mobile phone cameras enables consumers to capture digital images almost anytime and anywhere. And as a result, we can envision the proliferation of personal digital photograph collections and image sharing facilitated by the rapidly increasing network bandwidth and the ever-increasing ubiquitous wired and wireless environment.

Think of the current camera user behavior especially in Japan, one normal consumer can easily take an average of 10 digital images per day. In fact, some of the younger generations even take pictures for each dish of their meals with mobile phone cameras. The result would be a 3,650 photograph collection in a year. Within 10 years, the collection may expand to around 30,000 to 40,000 images. And in 30 years, 100,000 photographs collection can be accumulated easily. While more enthusiastic users might take several thousands of photographs during a single 10 day trip, personal image databases which contain more than 100,000 photographs can be expected in the near future.

How can we deal with a collection of 100,000 photographs? It is almost impossible for any individual to annotate each photograph manually. However, without constant organizing efforts, how can we allocate something such as: "A photograph of mine that was taken around 5 year's 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"?

Currently, most people only have a collection of digital photographs accumulated within duration of around two to three years due to the short history of digital camera. A two to three year photograph collection might still be manageable with our memory; however, how about ten years later?

The rapid expansion of personal digital photographs would eventually make it impossible to retrieve in later days without a systematic indexing, categorizing, and browsing interface. 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. We plan to review, evaluate and discuss recent research efforts and directions towards semantic personal photograph database retrieval with special address on spatial and temporal-based approaches.

Several research efforts start to focus on the importance of location and time aspects of consumer photograph annotations in the past few years. Differ from traditional visual feature based image retrieval; metadata has been utilized in order to achieve higher level data interpretation and semantic retrieval. We proposed a semi-automatic metadata generation methodology with a MPEG-7 based annotation architecture, DDDC (Dozen Dimensional Digital Content), which enables spatial and temporal 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.

Spatial and temporal referenced digital photographs stamped with location information such as GPS data have become possible with the aid of GPS receiving devices, GPS software and increasingly sophisticated digital cameras. We devoted our effort on developing an experiment personal photograph library, which includes more than 80,000 of consumer photographs, metadata annotation based on our proposed MPEG-7 annotation architecture. In annotation process, we utilize our proposed Spatial and Temporal Ontologies (STO) designed based on the general characteristic of personal photograph collections, including special emphasis on photograph patterns induced from personal history, human relations, hobby, taste, and preferences.

We further elaborate on the evolvement of image retrieval research efforts from traditional stock image database towards personal photograph collections. And detailed explanation on

the proposed system structure and future directions towards building up long term personalized image retrieval applications can be found in full thesis.