Archiving of Personal Digital Photograph Collections with a MPEG-7 Based Geotag Related Annotation Methodology

Po-Yen Chen, Pei-Jeng Kuo, Department of Computer Science, National Chengchi University, Taipei, Taiwan

Abstract

Nowadays, everyone may easily possess one or more cameras. The innovation of digital camera technology is getting faster and faster, and its performance and capacity is constantly growing. With the pervasive of digital recording devices, personal digital data may show an exponential growth. However, it is difficult and time consuming to organize and relocate specific photograph collection within such a huge amount of digital data when there is lack of prebuilt descriptions.

In 2004, Kuo et al. [10] proposed a multimedia description schema system based on MPEG-7 technology called PARIS (Personal Archiving and Retrieving Image System). It was designed to integrate spatial and temporal information of multimedia content into a MPEG-7 based semantic description. With this description architecture, PARIS envisioned being able to continuously capture and archive personal experience with audio-visual recording data and utilize potential social networking annotations provided by third party services.

We carry out an ongoing project with the PARIS architecture using a semi-automatic crowdsourcing annotation possibility enabled by third party social networking services. While most current smartphone cameras are equipped with GPS recording features, the possibility of semi-automatic recording of personal recognizable text based spatial data is still unable. In this paper we explains our proposed methodology of adding text based location description into multimedia contents with the aid of current social networking services.

Introduction

"Go Shabu Shabu, tonight." -my friend posted this invitation on my Facebook wall. Then I headed out to the restaurant with both my compact digital camera and camera-equipped smartphone. While enjoying the dinner, I use my smartphone to "check-in" our currently location, and take photographs of our dining moments with both digital cameras. It's fairly easy to take hundreds of photographs during such events.

Normally, users tend to upload photographs to their computer hard drives after a few events, and leave those photographs into subfolders with their original serialized file names untouched. This thus poses great challenges when users are trying to identify or categorize their photographs in the future. The need of an efficient tool and methodology to annotate and organize personal photographs is essential. Otherwise, it could become increasingly difficult to retrieve specific photographs as personal multimedia collection accumulates. The problem of how to organize and annotate personal photographs has become a major issue in discussion in the area of personal digital assets management.

Previously, personal photograph retrieval is composed of two methods. Content-Based Image Retrieval (CBIR), which compares image features such as the color histogram, color layout or scale-invariant feature transforms (SIFT) descriptor from a set data to achieve automatic retrieval [1]. For example, current face recognition technology can achieve high recognition rates when detecting different people's faces [2]. Thus we can easily identify specific friends in large photograph collections. There are a number of successful commercial photo management softwares that support the above-mentioned functionalities, such as Apple's iPhoto and Aperture. However, topics that require higher precisions are yet to be implemented.

Another method is Metadata-Based Image Retrieval (MBIR), which focuses on adding text-based annotations to perform image retrieval tasks. Although this approach was very popular in the past, it faced many difficulties in processing unstructured text annotations and keyword matching. In [3–5], more and more semantic based metadata annotations with the aid of ontology were proposed to extend the design of structured annotation approach. Moreover, [6–8] proposed to use external metadata annotations when dealing with conceptually higher level of contents.

Based on the above-mentioned approach, Kuo proposed to utilize MPEG-7 standard in order to establish a metadata-based Personal Archiving and Retrieving Image System (PARIS) [9]. PARIS extended the MPEG-7 StructuredAnnotation DS of Multimedia Description Schemes (MDS) to annotate multimedia data with twelve semantic attributes. Those attributes include data concerning the who, what, when, where, why and how (5W1H) aspects as well as the respective direction, distance, and duration (3D) information of the related digital contents, and is referred to as Dozen Dimensional Digital Content (DDDC) [10–11]. A well-designed annotation scheme can enhance management and improve retrieval accuracy, thus PARIS architecture envisioned to achieve a better retrieving possibility. The popularity of modern mobile devices with GPS technology allows more DDDC metadata attributes to be automatically generated, and this gives DDDC more advantages over previous proposed schemes.

Nowadays, various social networking services provide check-in functionalities. Those check-in point of interest (POI) data could be very useful when we try to improve our previously proposed DDDC annotation architecture with integrated geotags. Currently, most POI information provided with social networking services did not capture the multi-leveled concept or geographical locations, thus make it very difficult to express complex location information. In this paper, we propose to establish a geotag related ontology, which used to describe the relationships between current available POIs. We try to initial a methodology, which improves our previously proposed spatial-based ontology. We also try to implement a revised system, which connects current available POIs with their geographical relationships.

Today, social network services are pervasive. It is relevantly easy to connect to the Internet with one's mobile device anytime, anywhere, and it has become popular to share one's location and activity with modern social network services. Popular social network services usually provide options for user to mark their current location with geospatial information. Through various location-based social network services, such as Gowalla, Foursquare, Facebook Place and so on, users can Check-In their location with the aid of latitude and longitude information received via the GPS sensor provided with modern mobile devices, and a recommended lists of POIs are popped out for a semi-automatic geotagging process. With the aid of recent location-based social network services, we hope to improve our DDDC architecture to achieve a better annotating possibility.

We try to integrate two social community resources into our revised system: (1) Gowalla, which consists of various POI data and (2) Facebook check-ins, which includes huge amount of personal geographical data. And we try to use these POI data to improve the previous proposed semi-automatic annotation system.

In this paper, we will explain an improved DDDC that integrates check-in information provided by social network services. The current design of geotagging with most social network services do not capture the multi-leveled concept of geographical locations, which make it difficult to express complex location relationships. Thus we also developed a Geotag Related Ontology. We will describe the improved DDDC scheme in Section 2, and explain how to build a Geotag Related Ontology in Section 3. Section 4 will cover how to produce the improved DDDC semi-automatically. In Section 5 we provide our conclusion and potential future works. This paper explains the possible usage of current geotaggings provided by modern location-based services, we envision to improve our proposed annotation methodology and to facilitate the management of long term personal photograph archives.

Building Integrated Dozen Dimensional Digital Content (iDDDC) with MPEG-7

***Figure 1**. Where:Geotag and What:Geotag attributes*

The DDDC scheme proposed by Kuo expands the StructuredAnnotation DS of MPEG-7 Multimedia Description Schemes (MDS), providing a machine readable, semi-automatic annotating method. While DDDC was proposed, social network services were not widely available and it was relatively difficult to obtain locational related information automatically. The possibility of using geotags provided by location-based social network services to facilitate personal photograph annotations was not available back then. We hope to propose an improved scheme that integrates check-in information from location-based social network services into DDDC, providing a novel geotagging methodology to simplify the archiving and retrieving of personal photograph collections with sematic annotations. The modified DDDC requires combining original Where attribute and its three sub-attributes with *Where:GeoPosition*, and adding geotag attribute to both Where and What attributes to form *Where:Geotag* and *What:Geotag*. (Fig.1) The description of the two attributes are explained below:

Where:Geotag

Describes the geographic location of the photographer. In DDDC no attribute was used to describe this, mainly relying on latitude and longitude data to automatically determine the position. However, we soon found this made location-based search queries difficult, since we could not know what location the machine assigned to each photo.

What:Geotag

Adding descriptive geotags (POI) to the places or sights that appear in the photo. For example, we can take a photograph of Taipei 101, the highest skyscraper in Taipei City, from either Taipei World Trade Center (0.1 km apart from Taipei 101) or Maokong (7.2km apart from Taipei 101). Although we can easily annotate the *Where:Geotag* attributes of each photographs as Taipei World Trade Center and Maokong according to the related GPS location of the photographing device (smartphone, for example), it might be of greater interest to the audience to know the exact geotag of the subject (Taipei 101, for example) appears on each photographs. We try to accommodate such situation switch an extra *What:Geotag* attribute (Taipei 101 in this case) in order to indicate the information of subjects appeared in a specific photograph. It can be useful when users try to retrieve photographs of geographic related subjects within the photograph collections.

The above changes maintain the original dozen dimensional setting of DDDC, and we named the modified DDDC configuration as Integrated Dozen Dimensional Digital Content (iDDDC). With the advance of emerging technologies, various resources provided by pervasive social network services became easier to obtain. We believe that the above-described changes provide a more flexible geo-information description methodology and would achieve a better semantic retrieval of personal photograph collections. In the next section, we will describe our concept of Geotag Related Ontology architecture.

Geotag Related Ontology

The above-described iDDDC integrates the concept of social network geotagging with our previous proposed DDDC architecture. The current design of geotagging with most social network services do not capture the multi-leveled concept of geographical locations, which make it difficult to express complex location relationships. If we wish to utilize the geotags provided by

recent social network services, an improved geotagging structure would be crucial. Thus, we have designed a Geotag Related Ontology as shown in Fig.2, which splits each geotag into six degrees. The termRelationQualifierType in MPEG-7 is also expanded to defining relations between multiple geotags. The explanations of the six degrees shown in Fig.2 are described below:

***Figure 2**. Concept of Proposed Geotag Related Ontology*

1. Broader Term (BT), defines a geotag with a spatially larger term compares to the current geotag. For example, the BT of "National Chengchi University" might be "Taipei City" as National Chenchi University is located in Taipei City.
2. Narrower Term (NT), defines a geotag with a spatially narrower term compares to the current geotag. For example, the NT of "National Chengchi University" might be "DaRen Building" or "International Building" as both DaRen Buiding and International Building are located within National Chengchi University.
3. Past Term (PT), defines a geotag with a temporally earlier term than the current geotag. For example, the PT of "National Chengchi University" might be its previous name "Central Political School" as National Chenchi University was previously named as Central Political School.
4. Future Term (FT), defines a geotag with a temporally later term compares to the current geotag.
5. Synonymous, has two attributes: Use Instead (US) describes a linked geotag that is a preferred term to the current geotag. For example, "Zhengda" is a preferred term for "National Chengchi University". Use For (UF) assigns a geotag terms that is preferred to the linked geotag. For example, "NCCU" is a preferred term used for "National Chengchi University".
6. Related Term (RT) / Nearby Term (NB). The former Related Term (RT) denotes a linked geotag term related to the current geotag. For example, "MOS Zhengda Store" is a related geotag to "National Chengchi University" as MOS Zhengda Store is frequently visited when people are located in National Chengchi University. The latter Nearby Term (NT) is a linked geotag term, which has geographic relationship with the related geotag. For example, "Taipei Zoo" is a NT of "National Chengchi University" as Taipei Zoo is geographically close to National Chengchi University.

We proposed to implement a Geotag Related Ontology with six degrees of geotag relationships, which can compensate the deficiency of the current geotag structure provided by recent social network services. Moreover, there is no limitation to the number of linked geotags for each degree. Therefore, we are able to structure a hierarchical but also flexible Geotag Related Ontology. For example, we can link "Taipei City" with Narrower level elements like "Wenshan District" and below "Wenshan District" we have "National Chengchi University". Under "National Chengchi University", there are, "International Building" or other geographical related tags. Another example is the "Taichung Commercial College", which was reinstituted as "National Taichung Institute of Technology" in 1999, and later "National Taichung University of Science and Technology" in 2011. This relation can be expressed as the PT/FT. We can also describe the competitive relationship between "National Taichung Institute of Technology" and "Taipei College of Business" (RT), and geographical link with "Taichung First Senior High School" and "Chungyo Department Store" (NB) that are located nearby. The Geotag Related Ontology can be used on single photograph or photograph collections to enhance recognition and understanding of geographic information related to the photographs, which improves the originally rigid geographic annotations with flexibility and a hierarchical retrieving possibility.

Geotag Related Ontology provides a methodology for annotating personal digital photograph with geographic related information. The purpose of our proposal is to improve the applicability of the iDDDC scheme and the expanded Spatial Ontology provides a basis for realizing more advanced schemes such as Spatial-Temporal Ontology, Personal History Ontology, Human relation Ontology, and Hobby, Taste and Preference Ontology [12].

Semi-Automatic iDDDC Generation

While modern mobile devices can easily retrieve GPS data, it is still difficult to automatically map the data with a specific location name. Moreover, most current digital cameras have yet equipped with GPS receiving sensors. This posed potential difficulties in realizing the geotagging process. For annotating personalized photograph collections the idea of implementing an automatic location tagging is even more challenging. We thus propose a semi-automatic geotagging process for personal digital photograph collections.

When annotating photograph collections, a recommendation that reduces the time of complete manual input is required in order to improve annotation speed and efficiency. Firstly, we try to construct a server side POI database using distributed crawling process to retrieve check-in data with the aid of different service provider APIs. The data in the database mainly comprises of Gowalla's spot data of four different locations – Taiwan, Tokyo, Paris, and Greece –containing over 40,000 spot entries. Each entry includes data concerning location name, latitude and longitude, check-in counts, district, category, and description. Besides this data, we also crawled user's personal data from Facebook including personal and friend check-in data in order to accumulate more personalized user data in our database. In this system, we use the POI database to allow greater efficiency and ease in annotating photograph collections with a suggested geotag lists. Meanwhile,

we take into consideration of user's preference, friend check-in behavior relatedness, and spot popularity to calculate our recommended geotags when user does not manually type any annotating texts. We envision to realize a collaborative filtering algorithm based on above description in order to provide users with a semi-automatic methodology in constructing iDDDC descriptions towards personal photograph collections.

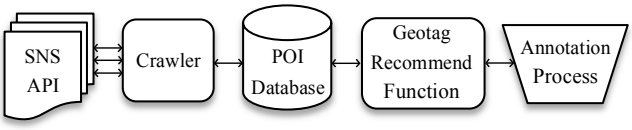

***Figure 3**. Semi-automatic iDDDC metadata generation system*

Conclusion & Future Works

PARIS system was designed to improve user experience in photograph collection management. However, as technology advances, people spend more and more time on social network services, making it very easy for users to manage their photograph collections online. While various geographical information as well as geographic related web resources cannot be easily connected with personal photograph annotations, we try to integrate resources provided by existing social network services with an improved MPEG-7 multimedia description scheme based annotation scheme. We modify the previously proposed when and where descriptions within the DDDC description scheme with geotags provided by current social network services and rename the modified scheme as iDDDC (Integrated Dozen Dimensional Digital Content). Meanwhile, we also propose the Geotag Related Ontology to strengthen geographical relationship information for the related geotags in order to enhance retrieving precisions. Manually annotating personal photograph collections is extremely difficult, and it requires a lot of time and work. Therefore, we implemented a semi-automatic annotating system, more precisely, a geotag recommendation system, that allows users to conveniently add social geotags and add six degrees of relations between tags.

This paper explains our approach towards a better way of personal photographs management with the aid of social network services. We try to reuse the already available geotags to achieve potential location based photograph retrievals. We provide a hierarchical and flexible geotagging methodology with proposed Geotag Related Ontology. In the future, based on iDDDC mechanism, we plan to provide more specific ontology assistance in personal photograph collection management, and design easier and more precise photograph retrieval methodologies. We also hope to develop a collaborative annotation mechanism [15] that allows geotag sharing and combining to further enhance personal photograph collection management experiences.

References

[1] Marios C. Angelides & Harry Agius, The Handbook of MPEG Applications: Standards in Practice (Wiley, 2011) pg. 221.
[2] Gao Wei et al., Face Recognition Based on PCA and SVM Optimized by MEB, Proc. ICINIS, pg.68. (2010).
[3] B. S. Manjunath et al., Introduction to MPEG-7 Multimedia Content Description Interface (Wiley, 2002) pg. 83.
[4] Ana B. Benitez et al, Semantics of Multimedia in MPEG-7, Proc. ICIP, pg. 137. (2002).
[5] Von-Wun Soo et al., Automated semantic annotation and retrieval based on sharable ontology and case-based learning techniques, Proc. JCDL, pg. 61. (2003).
[6] Eero Hyvönen et al., Ontology-Based Image Retrieval, Proc. XML Finland 2002 Conference, pg. 15. (2002).
[7] Alejandro Jaimes & John R. Smith, Semi-automatic, data-driven construction of multimedia ontologies, Proc. ICME, pg. 781. (2003).
[8] A. Th. (Guus) Schreiber et al., "Ontology-Based Photo Annotation," Jour. IEEE Intelligent Systems, 16, 66 (2001).
[9] Pei-Jeng Kuo et al., Continuous Archiving of Personal Digital Photograph Collections with a MPEG-7 Based Dozen Dimensional Digital Content Architecture, Proc. Archiving, pg. 188. (2005).
[10] Pei-Jeng Kuo et al., Building Personal Digital Photograph Libraries: An Approach with Ontology-Based MPEG-7 Dozen Dimensional Digital Content Architecture, Proc. CGI, pg. 482. (2004).
[11] Pei-Jeng Kuo et al., MPEG-7 Based Dozen Dimensional Digital Content Architecture for Semantic Image Retrieval Services, Proc. EEE, pg. 517. (2004).
[12] Peg-Jeng Kuo, Continuous Personal Digital Photograph Archival and Retrieval with Spatial and Temporal Coordinated MPEG-7 Description Scheme, Ph.D. Thesis, University of Tokyo. (2007).
[13] Marios C. Angelides & Harry Agius, The Handbook of MPEG Applications: Standards in Practice (Wiley, 2011) pg. 263.

Author Biography

Po-Yen Chen received his B.S in Geomatics from National Chengchi University, Taiwan. He is currently a Master program student in the Department of Computer Science at National Chengchi University. His research topics include geographic information system, digital content archiving, social computing and digital publishing.

Pei-Jeng Kuo received her B.S. in Physics from National Taiwan University, Taiwan; M.E. from Massachusetts Institute of Technology, USA; and PhD from The University of Tokyo, Japan. She is currently an Assistant Professor at National Chengchi University. Her research topics include indexing, archiving, delivering, and retrieving of multimedia contents with MPEG-7 technology. She also works on digital publishing in recent years.