Personalized Multimedia Web Summarizer for Tourist

Xiao Wu, Jintao Li,
Yongdong Zhang, Sheng Tang
Key Laboratory of Intelligent Information Processing
Institute of Computing Technology, CAS, Beijing, China
{wuxiao, jtli, zhyd, ts}@ict.ac.cn

Shi-Yong Neo
Department of Computer Science
National University of Singapore
Singapore
neoshiyo@comp.nus.edu.sg

ABSTRACT

In this paper, we highlight the use of multimedia technology in generating intrinsic summaries of tourism related information. The system utilizes an automated process to gather, filter and classify information on various tourist spots on the Web. The end result present to the user is a personalized multimedia summary generated with respect to users queries filled with text, image, video and real-time news made retrievable for mobile devices. Preliminary experiments demonstrate the superiority of our presentation scheme to traditional methods.

Categories and Subject Descriptors

H.3.3 [**Information Search and Retrieval**]: Search Process

General Terms

Design, Experimentation

Keywords

Classification, Summarization, Video and Image Analysis

1. INTRODUCTION

The growth of internet in recent years results in an explosion of free multimedia content. Being a free platform for information sourcing, it is not surprising to see more and more users surfing or "googling" places of interest online to obtain extra information before even visiting them. Besides the usual official websites depicting information on these places of interest, there is also a huge amount of traveler's blogs, sightseeing photos and videos posted from end-users. Often, users waste much of the time trying to search for the information he needs rather than the reading through the information. Most information presented to the user is at website basis, which means users have to scrutinize through the chunks of details to obtain what they really wants as different users might be looking for different content. The challenge is to provide a personalized multimedia tourism information retrieval by analyzing the user query and providing most relevant materials which are most interesting.

Most previous researches and official travel agencies uses GPS based mobile tourism application and generalized tourist spot introduction. For example, personalized multimedia presentation for sightseeing is studied in [1] to facilitate the individual decision for budget traveling. Location based mobile tourism application and system proposed in [2] aims to provide guidance everywhere. However, less attention is paid to the problem of how to exploit previous tourists' experiences on Web into personalized summary, i.e. mining the huge amount of multimedia content generated during user's sightseeing and personalize summary for different queries. The problem is of great importance because the summary is valuable for future visitors to know about a place.

In this research, our motivation is to present summaries using online tourism information such from official websites, online encyclopedia, blogs, photos and videos, which exhibit the views of tourist spots in a personalized manner to user. We first obtain tourism related data from various websites. This is followed by analyzing them to distinguish the more important ones, denote as Keys. State-of-art algorithms such as high level concept and near duplicate detection are adopted to understand visual content. Finally Keys is returned to the user depending on the user's query.

2. OBTAINING AND CLASSIFYING MULTIMEDIA TOURISM INFORMATION

Previous tourist generated multimedia information available online are often used by users as tour guidance and suggestions. For example, millions of travelers are attracted to China every year by its brilliant ancient culture. The output of these travelers is a huge amount of tourism articles, photos and videos on various tourist spots publish on the Web.

2.1 Information Extraction

We first perform Web information extraction to obtain tourism metadata. Taking tagging quality and user popularity into consideration, we extract organized texts, geo-tagged images and categorized videos separately from Wikipedia, Flickr, Youtube and official tourism website. Using the tourist spot name to query the websites, as shown in Fig.1, tourism multimedia data in result lists is crawled and regrouped into location based data pools. In simple terms, each set of data relevant to a specific tourist spot contains three groups: extracted texts, images and videos.
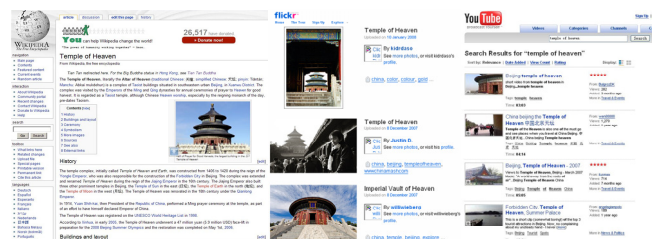

Figure 1. Query result lists from wikipedia, flickr and youtube using "Tian Tan" (temple of heaven) as the query

2.2 Classification of Keys

In this work, we attempt to classify elements or Keys (can be a key-point like a paragraph of text, an image or a segment of video) in the set of extracted multimedia data so as to handle user's query effectively. The main idea is to give each Key an intuitive description class. For example, a picture showing the building exterior is "outdoor scenery"; while a paragraph of text depicting past events is "history". The classes we predefined are: history, landscape, indoor scenery, outdoor scenery and general. By knowing the genre of Keys, we can provide users with more relevant information through query analysis.

Textual Keys are directly extracted automatically from Wikipedia and official tourism websites. They are classified using a text classifier which indicates whether the text provide descriptions to the landscape, history or general. The text classifier is built on [3] which is used for identifying news genre from news articles. Visual Keys (images and videos) adopt high level concept detection schemes. We choose 9 visual concept detectors from [4], i.e. Person, Crowd, Waterscape, Sky, Mountain, Vegetation, Building, Indoor and Outdoor to give semantic to visual Keys. Table.1 shows the correlation between Keys and respective genre.

Table 1. The correlation between Keys and genre

attributes / Keys genre	Visual Concepts of Keys for ranking	Textual and visual Keys combination
general	-	Text
history	-	Text
landscape	Sky, Water, Mountain, vegetation, outdoor	Text + Photos + videos
Indoor scenery	Person, indoor, building	Photos + videos
outdoor scenery	Building, outdoor, sky, person, crowd, outdoor	photos + videos

2.3 Giving Importance to Keys

We use the reliability of text classifier as the Importance for textual Keys re-ranking. For each visual Key, 9 high level concept [4] scores are calculated and the linear weighting scheme is adopted as follow, to compute the relevance to each visual genre.

$$GE(K_i) = \sum_j Concept\_Score_j(K_i) \qquad (1)$$

where K_i denotes the Key and GE is the relevance to a visual genre. Hence Keys could be mapped as sample points in 3D genre space as illustrated in Fig.2 (a). By projecting each Key on 3 axes, Keys are reranked into 3 orders to evaluate Importance for genres.
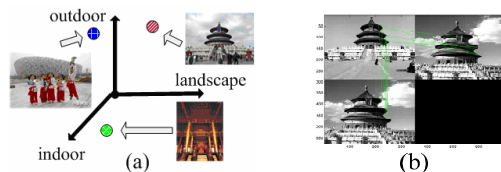

Figure 2. Ranking importance of visual Keys of "Tian Tan"
(a) Photos and keyframes are mapped into 3D genre space
(b) Near duplicate detection using local feature points matching

2.4 Filtering Noise and Duplicate Removal

Visual data clawed from sharing websites contain noises which is not relevant to tourism. We use genre space model and near duplicate detection [5] to filter out noises. First irrelevant Keys with very low GE score are discarded. Second duplicated photos and shots are detected (as in Fig.2 (b)) and clustered to find out visual data of tourist spot and then eliminate the noises.

3. QUERY ANALYSIS AND SUMMARY

By analyzing query keywords (as in Table.2), Keys of certain class of a tourist spot are summarized and returned to users. First, we determine the required tourist spot by capturing key query words. Subsequently, we return Keys particular in respect to the user's requirement. For example: the query "What is the historical background of Tian Tan?" is directed towards the history aspect and history Keys should be returned. Table 2 shows how the various types of queries interact with the genre of Keys.

Table 2. Relating query to genre of tourism query

Query Genre	Relevant keywords in Query Text (query is given by users in English)
general	Queries only contain tourist site name; queries with general terms (e.g. how, what, weather ("气候"))
history	Queries with culture terms (e.g. famous person names, history ("历史"), custom ("风俗"))
Land-scape	Queries with sightseeing terms (e.g. landscape, name of nature scene, spectacle ("奇观"))
indoor	Queries with entertainment terms (e.g. buy, games)
outdoor	Queries with view terms (e.g. name of building and city sites, park ("公园"), sites ("景点"))

To facilitate tourists who surf from mobile devices, we aim to provide a concise summary by: (a) returning Keys which are ranked highest; (b) porting to formats such as 3GPP or FLV which is playable directly on mobile phones. In addition, latest news reports (crawled automatically from online news sites) if available, about the required tourist spot is also returned. This information can prepare the tourist for unforeseen circumstances like temporary closure for maintenance or bad weather.

4. EXPERIMENTS

We carry out a prelim assessment on this work. A group of 20 users are chosen to compare the results of this summarizer to searching for information online. From the feedback, 19 users found our system to provide sufficient details on what they need, and 12 users found more information using our system than surfing online themselves. This work is currently in the prelim stages and we are looking towards more comprehensive testing.

5. CONCLUSION

In this work, we introduce an automated process to gather, filter and classify information known as Keys on various tourist spots on the Web. A personalized multimedia summary filled with text, image and video is generated with respect to user's queries Prelim experiments with users demonstrate the superiority of our presentation scheme to traditional methods.

6. ACKNOWLEDGMENT

7. REFERENCES

[1] A. Scherp, S. Boll. Generic Support for Personalized Mobile Multimedia Tourist Applications. ACM Multimedia, 2004.

[2] Schmidt-Belz, B., Poslad. Personalized and Location-based Mobile Tourism Services. Workshop on "Mobile Tourism Support Systems" in conjunction with Mobile HCI, 2002.

[3] S.-Y. Neo, Y. Ran, H.-K. Goh, Y. Zheng, T.-S. Chua, J. Li. The Use of Topic Evolution to help Users Browse and Find Answers in News Video Corpus. ACM Multimedia, 2007.

[4] S. Tang, Y. Zhang, etc. TRECVID 2007 High-Level Feature Extraction by MCG-ICT-CAS. TRECVID workshop, 2007.

[5] Y. T. Zheng, S. Y. Neo, etc. Fast Near-duplicate Keyframe Detection in Large-scale Corpus for Video Search. IWAIT07.