

# A Scenes-Detection and Recommendation Method Using GPS Information Embedded Photographs from WWW

Hsien-Tsung Chang, Hao-Chen Wang

Computer Science and Information Engineering, Chang Gung University  
 No.259, Wunhua 1st Rd., Gueishan Township, Taoyuan County 333, Taiwan (R.O.C.)  
 {smallpig, ppschachen}@widelab.org

**Abstract:** The sharing of travel photos on web albums, online communities and personal blogs has become increasingly popular (e.g. through Flickr or Facebook web albums) in this age of fast-growing markets for tourism, ever shortening IT development life cycles and ubiquity of digital cameras. The most critical challenge is how to analyze and filter out the useful clues of travel photographs to mine relevant travel information from the vast troves of photographs on the Internet. We design and implemented a system to detect the scenes and recommend their rank with a experiment which is restricted the range of photographs to those taken in Taiwan.

**Keywords:** GPS, Scenes-Detection, Tourist Rank

## I. Introduction

### A. Research Background

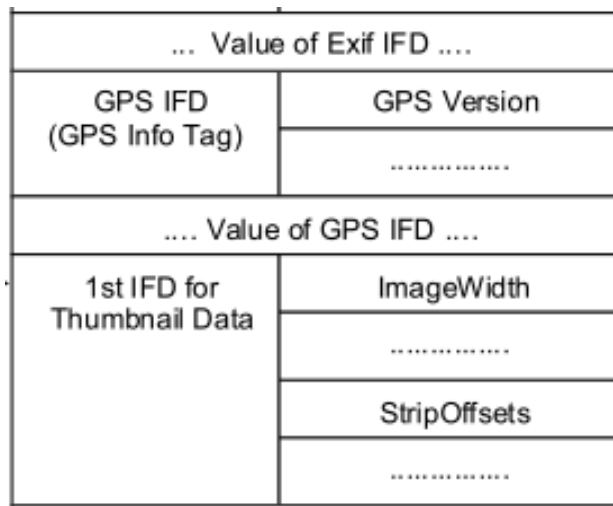


Figure 1. EXIF Data Structure [4]

According to the World Travel & Tourism Council (WTTC) [1], the global tourism industry currently accounts for 9.4% of the global gross domestic product (GDP), and is projected to increase to 9.5%, or about 10.5 trillion U.S. dollars, in 2019. In addition, statistical data in 2009 showed that international tourists visiting the Asia-Pacific region in 2007 grew 10.5% in

comparison to 2006, a significant increase. The number of tourists in 2008 was approximately 180,000,000, up just 1.6 % from the previous year, but the global tourism industry is nevertheless an important sector of growth for the global economy.

ItemName	Information
Saturation	Normal
Sharpness	Normal
SubjectDistanceRange	Unknown
<b>GPS Information</b>	
GPSVersionID	2.2.0.0
GPSLatitudeRef	N
GPSLatitude	24 3152.95 [DMS]
GPSLongitudeRef	E
GPSLongitude	121 541.4888 [DMS]
GPSAltitudeRef	Sea level
GPSAltitude	1061/1 meters
GPSMapDatum	WGS-84

Figure 2. Examples of EXIF Application

In the past, digital cameras were capable only of generating an electronic image file. Nowadays, most digital cameras are further equipped with the Exchangeable Image File Format (EXIF) photograph information capability. EXIF is an exchangeable image format, specifically designed for incorporating digital camera photograph information. The Japan Electronic Industry Development Association established the standards for version 1.0 in 1996, and the latest official standards currently stand at 2.2. It can also be tagged onto JPEG, TIFF, RIFF and other document formats. Figure 1 shows the basic EXIF structure of the uncompressed files. The GPS IFD values can be read on digital cameras supporting GPS, and when one travels with GPS satellite navigation, one can embed and record GPS information in a photograph by artificial means [3]. In Figure 2, with the use of the relevant applications, one can see a list of information that can be recorded by photograph EXIF, such as GPS information. There are currently several online albums communities that enable display of the EXIF information. For example, Picasa provides information ranging from the model of camera used,

ISO, latitude, longitude, etc. From version 3 of the wildly popular software ACDSEE, the Metadata field information will also be displayed if the photograph contains EXIF information.

### B. Motivation

In accordance with our research motivation, we started from travel photograph collections of the general public and the tourism information embedded in the photographs. From the photographs and embedded EXIF information obtained from the web, we located scenic spots by means of filtering noise detection. Since the popularity of scenic spots will vary with time, we isolated scenic spots of interest by automatic detection, and identified trends in the rankings thus obtained.

Our starting point was the EXIF information embedded in the photographs. When crawling for photographs on the web, we also captured other relevant information and employed analysis filtration accordingly, as outlined in the following chapters. We conducted exploration of scenic spots and provided tourism-related recommendations, using the Tourist PageRank calculation method to compile relevant rankings.

PageRank [5], also known as page rank, was an algorithm invented by Sergey Brin and Larry Page, the founders of Google. Its main application is search engine technology, for indicating the relevance and importance of web pages. Based on the hyperlinks between web pages, one can calculate a page's rank using the core concept of the vote for the hyperlink. The complete equation is as following.

$$\text{PageRank}(p_i) = \frac{q}{N} + (1-q) \sum_{p_j \in M(p_i)} \frac{\text{PageRank}(p_j)}{L(p_j)}$$

In the above equation,  $p_1, p_2, \dots, p_N$  are the pages in our collected set,  $M(p_i)$  is the set of pages that link to  $p_i$ ,  $L(p_j)$  is the number of page links out from  $p_j$ , and  $N$  is the number of all pages. The variable  $q$  is known as the damping factor, and it represents the probability that a random visitor will continue on to other pages of a website after he or she has browsed a particular page.

The PageRank of individual pages is determined by the voting scores, and the number of external links to an individual page is representative of its scores. For example, if someone links to your page, it represents his or her vote for you. However, the vote scores will also depend on the global total of external links. If the PageRank value is 5, and 10 persons have voted (10 external links), then score is only 0.5 points. Therefore, the greater the number of linked pages the higher the value of the PageRank.

### C. Paper Structure

This paper is divided into five sections. Section 1 consists of the foreword; Section II details our approach, Section III describes the experimental design, analysis, results and discussion; Section IV presents our survey of the relevant literature; Section V outlines the advantages, limitations and possible applications of our study.

## II. System Design and Implementation

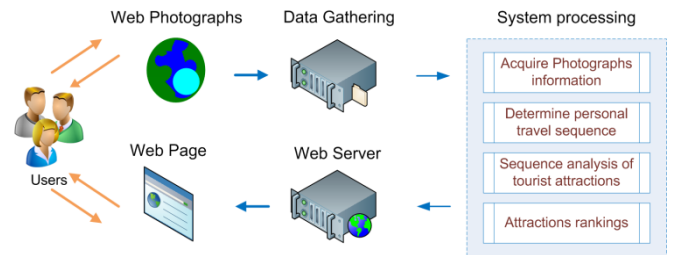


Figure 3. System Architecture

### A. System Architecture

Figure 3 shows our system structure. First, we obtained photographs published by users on the Internet, with the sources encompassing the entire network (such as Internet photograph albums, groups and blogs, etc.) as well as more concentrated sources such as Flickr [6] and other online communities for sharing photograph albums. We next wrote a program to gather data according to techniques fine-tuned over the years in our laboratory. We analyzed the EXIF information of each photograph and filtered “noises” in the photograph information using a set rules we formulated to determine the user’s travel behavior sequence. Next, we used the processed information to identify the scenic spots frequented by the users. Finally, we ranked and presented the results, and provided relevant recommendation.

### B. Information Processing

We used the Data Gathering function to capture photographs on the Internet to our servers, and recorded the source URLs of the photographs on the Internet, to enable follow-ups for users. This way, a user interested in certain scenic spots will not need to go through the search engine a second time to unearth more information. The data filtering and analysis consisted of four stages. First, we obtained the relevant photograph related information. Second, we determined the chronological order of the travel. Third, we conducted sequence analysis of the scenic spots. Fourth, we ranked the scenic spots.

#### 1) To obtain information about photograph

We captured the location of a photograph on the server, using the ReadExif function to obtain the EXIF information, and recorded the EXIF fields: “photography time, file name, camera model, latitude, longitude, version”, e.g. <Jun 28, 2008, P6272809.jpg, OLYMPUS IMAGING CORP., latitude 50.444469 °, longitude 18.855607 °, 2.2.0.0>, using “EXIF\_field name” to represent each field. In the program flow diagram shown in Figure 4, we first determined the EXIF information of the photograph. If the information did not exist, we then set the version of field to 0, to resolve the different versions of the format. If the EXIF\_photography time was not present, we would then select the earlier of the community site file upload time recorded or the time the photography was said to be taken from the website.

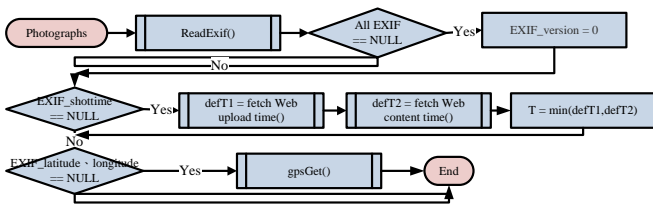


Figure 4. To obtain information about the photograph

Figure 5 illustrates how we used the longitude and latitude fields obtained with the GPSGet function to determine the GPS location where the photograph was taken. In the tempName array, we stored the three more frequently appearing geographic terms, before sending out inquiries to the search engines. Take array [0] as an example: for "tempName [0] GPS", we used the geographic term stored in array [0] + space + GPS to send out the inquiry.

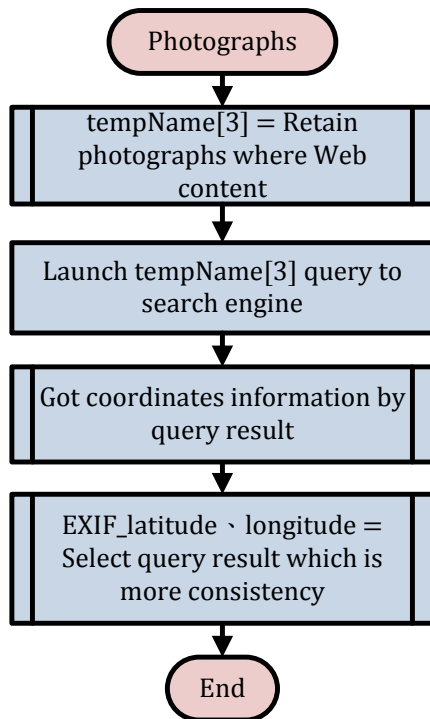


Figure 5. Photograph GPS information

Following the above procedures and functions, we saved the information for each photograph to the photograph data tables. The data table fields are: < ID, photograph name, longitude, latitude, photography time, camera model, EXIF version, where the web addresses >. When adding a record, the system automatically generates a serial number to the ID.

2) Determine personal travel sequence

During this stage, we analyzed the results generated from the previous stage. We defined the size of the tourism sequence to be  $1 \sim n$ , with time in days as the unit of  $\alpha$  and distance in km as the unit of  $\beta$ .

Figure 6 illustrates the flow of how to determine personal travel information. We performed the following three steps for all the information in the table. The first was cluster, and we filtered the noise with three conditions. The first condition

was to ensure that there was one and only one unique user ID for each user, with the assumption that there existed no external factors interfering with the experiment (e.g. the use of more than one camera for a particular trip). With the user ID in place, we obtained the user's name automatically with the program we wrote. Its capture was based on the following three criteria: source URLs from the same web directory, description on the originating webpage (photograph album of the user ID), hyperlinks on the originating webpage. The second condition: The same group of photograph taken over more than  $\alpha$  filters days (set  $\alpha = 3$ ). Condition 3: Filter the distance of the photograph by more than  $\beta$  km ( $\beta = 200$ ), with the distance commonly assumed for travel in the western half of Taiwan as an example. The distance from Taipei to Pingtung is about 400 km, a distance that can be used to determine the scope of tourism in northern Taiwan or southern Taiwan. The second step was the cluster of the previous step, according to the chronological order, with each cluster of information used to save a sequence. The third step was to analyze the travel sequence, a sequence that did not exceed more than two hours for each photograph, failing which we undid the process and reverted back to step one.

This way, we can infer the individual behavior travel sequence, a sequence comprising of several data points, each containing a photograph and the geographic location information, a sequence that was effectively a record of the user's travel through time. Although the use of the time at which the photographs were taken to reenact the user's travel itinerary leaves some margins of error with respect to time, we do not believe they have significant impact on our experimental results. Should we be unable to read the page by EXIF and the photograph's photography time, we would consider photographs sharing the same directory to taken at the same time. For the above mentioned filter and the method for determining travel, we do not foresee significant effects on our experimental results.

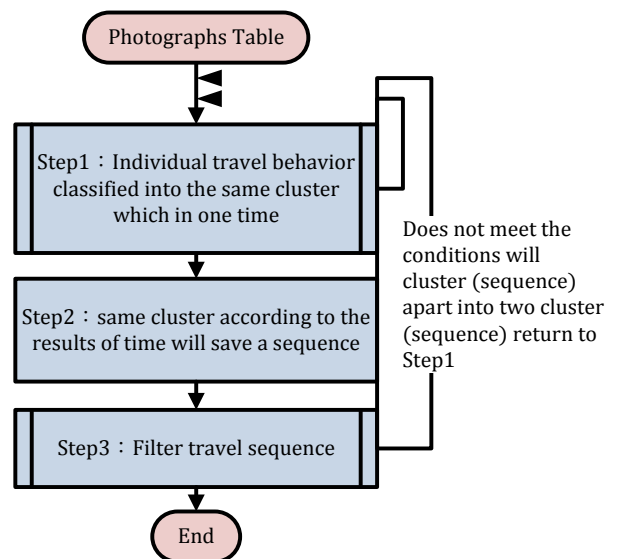


Figure 6. Determine personal travel information

3) Sequence analysis of tourist attractions

In this section, we based our analysis on the travel sequence determined in the previous section and determined the distance from the GPS information of each photograph. The distance of  $\gamma$  km ( $\gamma = 1$ ) within the photographs was chosen

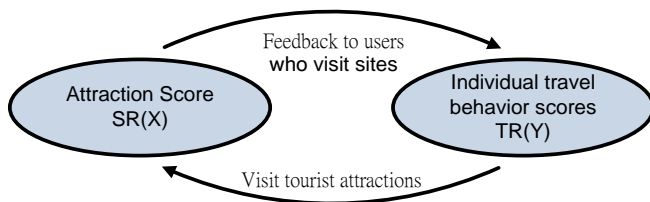
based on the assumption that few visitors will traverse a distance exceeding one km from the scenic spot. We then analyzed the name of scenic spots and obtained the names of the scenic spots from the web as well as the user's description.

We retained five candidate names for each scenic spot (e.g.: Truku, Taroko, Taroko National Park, Hualien County, Taiwan) and compiled statistical rankings for each candidate name of a scenic spot. When all the scenic spots statistical ranking was complete, we then filtered the name of the scenic spots.

According to the statistical rankings of the scenic spots' names, we selected the three most frequently appearing ones from the five candidates as the name of the scenic spots (e.g.: Truku, Taroko, Taroko National Park). This way, we obtained a more popular but not overly broad range of names for the scenic spots.

For names in other languages, we selected visitors speaking the respective languages, obtained descriptions of the photographs and titles on the Internet, and then chose the three spot names which appeared the most frequently. We then filtered in the same manner as described above.

#### 4) Ranking tourist attractions



**Figure 7.** The relationship between tourist attractions and individual behavior

We next applied the commonly used method for computing search engine rankings, as described in the introduction, to determine our ranking of the scenic spots — Tourist Rank. First of all, we assumed that the users took pictures only at locations considered scenic spots, and that some of these photographs also depict the users in their own homes or personal settings or some non-tourist spots. We filtered the noise by means of a ranking score threshold.

As formula (1) and formula (2) in the following,  $SR(X)$  represents the scores of the scenic spots, and  $TR(Y)$  represents the scores of personal travel behavior, with the relationship as shown in Figure 7.  $V_Y$  represents the set of tourists that ever visited  $Y$  and  $S_X$  represents the set of scenic spots that were visited by tourist  $X$ . We treat a scenic spot as a web page and the personal travel behaviors as links those are similar to HITS algorithm for ranking the web pages.

$$TR(Y) = \sum_{X \in V_Y} SR(X) \quad (1)$$

$$SR(X) = \sum_{Y \in S_X} TR(Y) \quad (2)$$

In order to match real situation of scenic spots, we introduce the number of photos as  $P(X)$  to indicate the total number of photos that tourist took in the scenic spot. We think that if a scenic spot is amazing for a tourist and he would take more photos than one scenic spot which is bored. So the value

$P(X)/\text{Numberof}(Y)$  indicates the attraction score for a scenic spot according to the average number of photos when a tourist visit that spot. We also use the PageRank concept to divide the  $TR(Y)$  by  $\text{Numberof}(Y)$ . The variable  $t$  is the total person-time of visiting  $Y$ . We modified the formula (2) to formula (3) as following to make it more match the real situation of a scenic spots.

$$SR(X) = \log(t) * \log\left(\sum \frac{TR(Y)}{|Y|}\right) + \frac{P(X)}{|Y|} \quad (3)$$

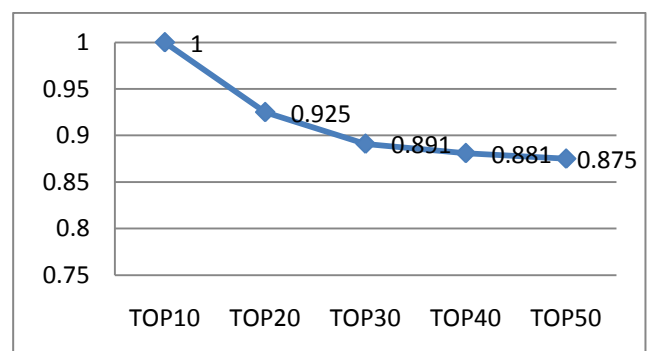
We applied the above formulas and calculated the spots scores  $SR(X)$  and personal travel scores  $TR(Y)$  iteratively. With the four stages of analysis and filtering as defined in Chapter II.B, we finally generated a ranking of the spots. We filtered the scenic spot scores which were lower than predefined threshold and presented our list of recommended scenic spots based on the scores in the ranking.

### III. Experiments

Our study made use of four servers to gather web photographs, and then a personal computer for performing the analysis to obtain and filter for the required information, with the final results stored on a server.

In our experimental set-up, the range of tourist attractions was restricted to Taiwan. However, we did not gather photographs only from websites based only in Taiwan, for the simple reason that many foreign tourists took photographs in Taiwan on their travels, in order to enhance our multi-language support. To enable users to obtain more information on the attractions, we can increase the volume of information covered by the query results. This will also facilitate future experiments to further expand the range of exploration sites worldwide.

In the experimental methods section, we grouped the results of our analysis into two categories. The first category caters to the assessment of attractions for which we offered relevant information. We attempted to verify from the users the accuracy of the information provided. The second category was mining attraction assessment. We attempted to better understand this kind of analysis and evaluation, how we could validate the accuracy of our system by comparison with real attractions. We will describe in details the experimental method as follows.



**Figure 8.** The average value for the top 50 detected spots.

We check the top 50 ranked scenic spots for evaluate manually. Figure 8 illustrates the results.

The second experiment is design as a questionnaire. We ask 64 master students what top 10 scenic spots in Taiwan is. And then check if the scenic spots from the questionnaire are detected by our system. We found that 93.43% of the scenic spots from questionnaire were detected by our system.

But still 6.57% of the scenic spots were not detected by our system. We check the un-detected scenic spots and discover an interesting fact. One of the undetected scenic spot named “士林夜市” (ShiLin night market). Tourist will go to ShiLin night market for delicious food. They wont take photograph for the food. Even they take the photograph, but we cannot retrieve the GPS information because we only can eat the food indoor. There exactly are a lot of photographs taken in the night market, but we have no GPS information to extract from almost all of them.

#### IV. Related Works

Y. Yang [8] applied the OWL approach in the collection and establishment of tourist information, where OWL is the acronym for Web Ontology Language. OWL is intended for use in cases where the data is process by computer applications rather than by human. In other words, it must be understood by the machine language. OWL can be used to clarify the meaning of terminologies, as well as the relationship between terminologies. When it comes to expressing meanings and semantics, OWL enjoys more means of expression compared to XML, RDF, and RDF-S type syntaxes. Therefore, to express a language readily understandable to computers in the web which is stronger than the above, which was designed primarily to process the contents of information, we took advantage of a feature for network and employed this method on a large number of travel sites to parse and integrate information. They used an OWL editor for graphical representations of concepts, where a line represents the relationship between concepts, allowing users to easily establish the relationship between self-concept and semantics. In the future, we hope to apply these methods to achieve automated deployment of travel information, but the syntax of such sites is still rather small, meaning the information thus obtained is very limited.

In 2009, A. Popescu [9] et. al attempted to match photographs supplied by users and the names of attractions in the photographs. They analyzed photographs whose file names were changed by the user, or web pages where the photo was described by the title and context. They made use of Wikipedia to obtain the names of attractions and construct a database, and then leveraged the vast resources of Wikipedia to obtain their counterparts in over 10 foreign languages, thereby creating geographic content with vocabularies from multiple languages to increase their matching range.

In Figure 9, the attractions on the left had 100 or less travel photographs whereas those on the right had 1000 or more photographs. It can be seen that there existed more famous attractions from before, and they attracted more visitors. We can identify several less popular attractions from this sample collection. In contrary to the methods of this paper, Wikipedia may not have included all the attractions, and those included are likely to be the most widely known but not necessarily the

most popular attractions. In addition, the validity of a sample collected by manually altering nicknames and information on webpages is debatable.

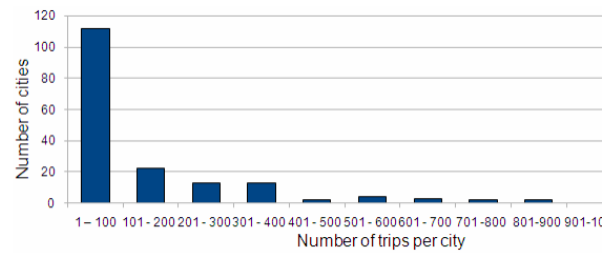


Figure 9. Cities and tourist attractions Chart[9]

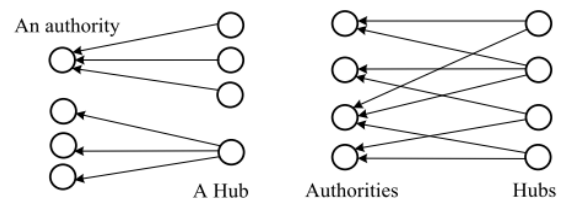


Figure 10. TBHG Structure[10]

Y. Zheng [10] also attempted to mine interesting sights by recording user travel behavior with GPS tracking. First, they defined the destinations, which contained two parameters: time thresholds and distance threshold, to determine the tourist attractions. They subsequently re-used the density of cluster algorithms to eliminate the noises in the information, with the information collected clustered at different levels. Here, the noise referred to areas where the cluster density was relatively sparse areas, and where TBHG was used to modulate the travel sequence.

After the completion of the above, one would obtain the clusters at different levels, as shown in Figure 10 on the right. One can then apply the HITS inference module to directly link the users and attractions. Here, the authors used two parameters. Authority represents the scores for the attractions, which is the total of the scores for hubs pointing toward this attraction. The higher the scores, the larger the number of high-frequency users. Hub represents the scores of the user, which is the sum of the scores of attractions it points toward. A higher score indicates that the user has visited many high-frequency tourist attractions. One then iteratively computes the scores to rank the attractions and users. This paper made extensive use of GPS log, and since not everyone owns such a device, it would be more difficult to gather travel information effectively. On the other hand, the use of the HITS method for score is more easily manipulated, since a user may travel frequently and thus has a higher score, but he or she may not necessarily visit the more interesting attractions.

#### V. Conclusions

We have effectively exploited photographs that recorded the sources of information and mined relevant travel information, using the analysis and filtering method as outlined above to remove the noise in the tourist attractions information. Lastly, we used the Tourist-PageRank to remove homes, offices and

other spaces for personal activities, and presented a system that ranks scenic spots in order of popularity as well as generates recommendations for users.

The source of our data is the entire body of photographs on the Internet around the world, but we limited our range of tourist attractions to Taiwan by restricting exploration out of range, in order to facilitate a proper assessment of the feasibility of our system. We gathered data at the same time, to provide information in different languages in terms of attraction names, descriptions from foreigner, etc. and we will obtain more alias names for popular attractions to increase the scope of the user query. A user with an interest in a particular tourist attraction can access photographs from a direct link on our system to the web page where the photographs originated without the use of additional search engines, websites, tools to obtain more information on the tourist attractions. Finally, we designed experiments using previously formulated methodologies to effectively get the system to automatically obtain the name of tourist attractions and verify feasibility and accuracy of information.

In the future, we hope to further expand the scope of our study to cover tourist attractions around the world, enabling more and more users to plan their budget travel accordingly. We intend also to study how best to help recommend planning of travel itinerary to cater to the needs of different individuals, a research topic we believe to be of utmost importance to the travel industry.

## Acknowledgment

The financial support by the National Science Council, Republic of China, through Grant NSC 98-2221-E-182-050- of the Chang Gung University is gratefully acknowledged.

## References

- [1] World Travel & Tourism Council (WTTC), <http://www.wttc.org/>
- [2] Taiwan Experimental Research Foundation National Science and Technology Policy Research and Information Center - Market reports reference IDC 2007.08.
- [3] Geotagging Photo Process, Taiwan academia sinica computing center.
- [4] EXIF 2.2 Official standard, <http://www.exif.org/Exif2-2.PDF>
- [5] Google Technology, <http://www.google.com/technology/>
- [6] Flickr: Taiwan , <http://www.flickr.com/groups/tw/>
- [7] Google map , <http://maps.google.com.tw/>
- [8] Yang, Y. et al , 2006. A Budget Travel Planning System using Ontologies for Web Information Aggregation. Information and Communication Technologies, 2006. ICTTA'06. 2nd, 1.
- [9] Popescu, A., Grefenstette, G. & Moëllic, P.A., 2009. Mining Tourist Information from User-Supplied Collections. ACM CIVR'09.
- [10] Zheng, Y. et al, 2009. Mining interesting locations and travel sequences from GPS trajectories. In Proceedings of the 18th international conference on World Wide Web. pp. 791-800.

## Author Biographies



**Hsien-Tsung Chang** received his M.S. and Ph.D. degrees in Computer Science and Information Engineering (CSIE) from National Chung Cheng University in July 2000 and July 2007, respectively. He joined the faculty of Computer Science and Information Engineering Department at Chang Gung University and served as the Assistant Professor in August. Dr. Chang's research areas focus on search engine, data engineering, information retrieval, web services and web 2.0.  
E-mail: [smallpig@mail.cgu.edu.tw](mailto:smallpig@mail.cgu.edu.tw)



**Hao-Chen Wang** was born in Hualien Taiwan on October 1 1985. He received his B.S. and M.S. degrees in Information Management and Computer Science and Information Engineering (CSIE) from Chinese Culture University and Chang Gung University in 2008 and 2011, respectively.  
E-mail: [ppshachen@widelab.org](mailto:ppshachen@widelab.org)