AN INDEXING SYSTEM FOR PHOTOS BASED ON SHOOTING POSITION AND ORIENTATION WITH GEOGRAPHIC DATABASE

Kiyoko Iwasaki, Kazumasa Yamazawa and Naokazu Yokoya

Graduate School of Information Science
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara, 630-0192, Japan
E-mail: { kiyoko-i, yamazawa, yokoya } @is.naist.jp

ABSTRACT

With the spread of digital cameras, shooting photos has been becoming an everyday affair. However, there are few methods or systems to manage photos simply, and a huge amount of photo data remains unorganized. Although it is possible to add appropriate words explaining the contents of the photo as one of the methods to manage photos, it requires much time and effort to input such indexes manually. It is also difficult to add indexes intended by a user automatically. This paper proposes a semi-automatic photo indexing system that enables users to generate indexes simply and browse a photo library efficiently. Index candidates are acquired from map database retrieval and relevant word extraction using web retrieval based on shooting position and orientation information. We have implemented an indexing system prototype based on the proposed method, and have carried out some experiments.

1. INTRODUCTION

In recent years, a number of methods for managing photos based on shooting position information have been investigated [1–3]. The Exif [4] specifies the formats of metadata for photos including caption, camera parameters such as focal length, shooting position such as latitude/longitude and so on. Since cellular phones with both a camera and a GPS device are already in practical use, photos with shooting position information will be generalized further. Shooting position information can be associated with the place names and the facility names through the matching with a map database, and users can browse and retrieve photos by keywords based on the shooting position [3]. Such conventional works use only prepared map databases. They often output indexes which the user does not intended. The user is required to input the intended indexes manually.

In this paper, we propose a semi-automatic photo indexing system that enables users to make location-based indexes for digital photographs easily. Place and facility names as index candidates corresponding to the subject position estimated from the shooting position and orientation are acquired from a map database prepared in advance. If there are no appropriate candidates in the database, new candidates are obtained by relevant word extraction using web retrieval. The map database is updated by adding the indexes selected by users as feedback, and hereafter it can present candidates that are more appropriate.

In Section 2, we describe the proposed indexing system. Section 3 describes a prototype system and some experiments. Finally, Section 4 summarizes the present work.

2. INDEXING PHOTOS BASED ON SHOOTING POSITION AND ORIENTATION

2.1. Overview of photo indexing system

Figure 1 shows the flowchart of the proposed method. First, a user acquires a photo in JPEG format with shooting position and orientation information by a camera to which sensors such as GPS device, gyro sensor and digital compass are attached. As for the shooting position, the latitude/longitude and the altitude are acquired by the GPS device. As for the shooting orientation, the angle of elevation and the direction are acquired by the gyro sensor and the compass. The camera parameters such as focal length and F-number are also acquired by the Exif header of the JPEG files. To acquire index candidates based on the subject position of photo, the subject position is estimated from the shooting position and orientation and the camera parameters.

Next, the place and facility names corresponding to the subject position are acquired from a map database prepared in advance. The acquired names are shown to the user as a list of index candidates, and then the user selects an index that is appropriate for the context of photo from the list of candidates. When there can be found no appropriate candidates such as more detailed names than those stored in the map database, they are obtained by relevant word extraction

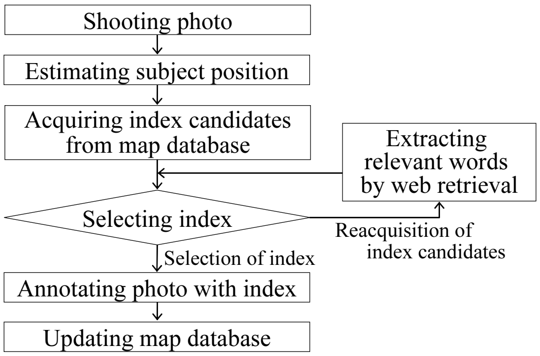


Fig. 1. Flow diagram of photo indexing.

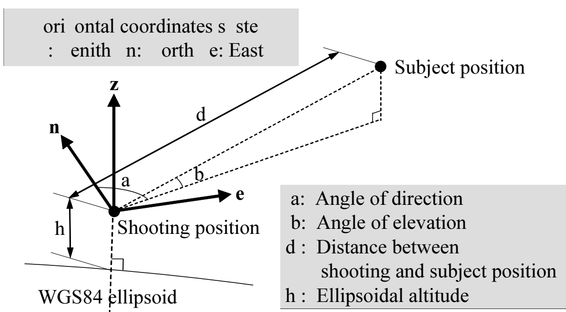


Fig. 2. Subject position estimation.

via Internet. The user selects a word that is the most relevant to the desired word. The system extracts relevant words using web retrieval with the selected word and shows them as new index candidates to the user. The index selected by the user is stored in a personal photo database with the photo and the metadata such as the shooting date, time and position. Moreover, the index is used as feedback to the map database. After that, the map database is expected to present candidates that are more appropriate.

2.2. Subject position estimation

Figure 2 shows the relation between the shooting position and the subject position. First, the shooting position whose latitude/longitude and altitude are based on WGS84 (GPS coordinate system) is defined as the origin of the horizontal coordinate system. On the coordinate system, the subject position is estimated from the angle of direction acquired by the compass, the angle of elevation acquired by the gyro sensor and the distance to the subject acquired from the Exif. The following processes use the latitude/longitude of subject position projected onto WGS84 coordinate.

2.3. Index candidate acquisition from map database

Index candidates are first automatically acquired by inquiring with estimated subject position to a map database. The initial map database consists of place data and facility data included in map software on the market, and it is updated by using the indexes selected by a user as feedback. Each data in the database is composed of its name, position (latitude/longitude) and frequency of user selection. The index candidates are presented to the user in ascending order of the distance between the estimated subject position and the position of the data.

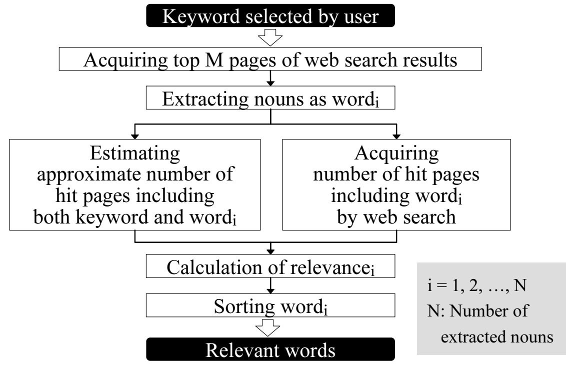


Fig. 3. Relevant word extraction diagram.

2.4. Relevant word extraction using web retrieval

When the index candidates acquired from the map database are not appropriate, new candidates are obtained by relevant word extraction. Sato et al. [5] proposed a relevant word extraction method using web retrieval. Since the method aims to apply the extracted words to a dictionary, the main purpose of the method is to extract words accurately and the processing time is not considered. On the other hand, our system should consider the processing time to realize interactive indexing.

Figure 3 shows the flowchart of relevant word extraction. The user selects a word related to the desired index from shown candidates. The selected word is used for the input of extraction. Hereafter the selected word is referred to as "keyword". First, our system requires the URL list of relevant web pages by web retrieval with the keyword, and acquires the top M web pages of the list. Next, the system extracts words without HTML tags from the web pages, and obtains the parts of speech of the words. Our system selects the nouns as the appropriate words from the extracted words.

The indicator of relevance between a keyword and an extracted noun is given by Eq.(1) in the present study. When the value of Eq.(1) is large, the relevance is large. The extracted nouns are sorted in descending order of the relevance, and are presented to the user as new index candidates. The value of $hit_{key \cap word_i}$ is estimated approximately to compute $relevance_i$ in practical processing time.

$$
\begin{aligned}
relevance_i &= hit_{key \cap word_i} / hit_{key \cup word_i} \\
&= hit_{key \cap word_i} / (hit_{key} + hit_{word_i} - hit_{key \cap word_i}) \\
&\quad (i = 1, 2, ..., N), \qquad (1)
\end{aligned}
$$

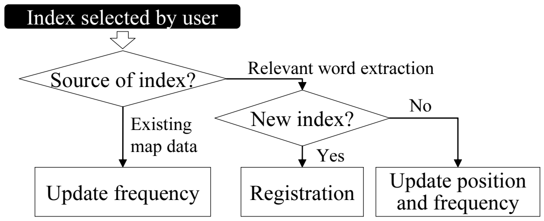


Fig. 4. Update of map database using selected index.

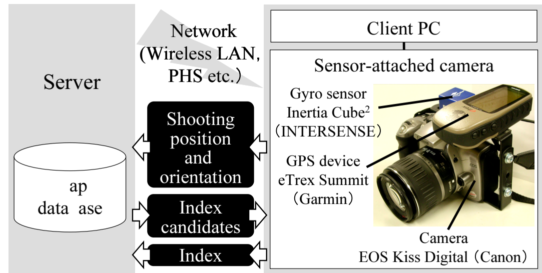


Fig. 5. Prototype system.

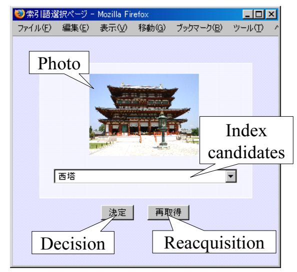


Fig. 6. Indexing user interface.

where

$relevance_i$:Relevance of word$_i$,

hit_{key} :Number of hit pages including keyword,

hit_{word_i} :Number of hit pages including word$_i$,

$hit_{key \cap word_i}$:Number of hit pages including both keyword and word$_i$,

$hit_{key \cup word_i}$:Number of hit pages including either keyword or word$_i$,

N :Number of extracted nouns.

2.5. Feedback on map database based on user selection

Figure 4 shows the flowchart of updating the map database by using the selected indexes as feedback. The positions of the indexes acquired from existing map data such as map software on the market are reliable. On the other hand, the positions of the indexes obtained from relevant word extraction may have errors which depend on the sensors at shooting. A new index is registered on the map database with the estimated subject position, and the position of the existing index is updated by Eq.(2)-(4). Since the feedbacks are repeated, the corresponding positions of indexes are averaged.

$$lat_{new} = (lat_{prev} \times freq_{prev} + lat_{sbj})/freq_{new}, \quad (2)$$

$$lon_{new} = (lon_{prev} \times freq_{prev} + lon_{sbj})/freq_{new}, \quad (3)$$

$$freq_{new} = freq_{prev} + 1, \quad (4)$$

where

lat_{new}, lon_{new} :New lat/lon corresponding to index,

lat_{prev}, lon_{prev} :Previous lat/lon corresponding to index,

lat_{sbj}, lon_{sbj} :lat/lon of subject position,

$freq_{prev}, freq_{new}$:Frequency of user selection.

3. EXPERIMENTS

3.1. Prototype system

We have implemented a prototype system based on the proposed method and have made indexes of photos obtained in real environments.

Figure 5 shows the structure of the prototype system. It is composed of a camera, GPS device and gyro sensor with a compass. Photos and their shooting positions and orientations are acquired by the sensors. The acquired data is stored in a client PC. When indexing, a user selects photos with the shooting position and orientation on a user interface. The user interface has input forms built on a web browser. The selected photos as well as their shooting positions and orientations are sent from the client PC to the server which has the map database via network. The server generates index candidates for the photos received and returns them to the client PC.

Figure 6 shows a screen shot of the user interface for indexing. The user selects an index from the drop-down list of index candidates. When there is an index which the user intended, the user selects the index and pushes the decision button for sending the data to the server. When there are no appropriate candidates, the user selects a word related to the index which the user intended and pushes the reacquisition button to request new candidates acquisition to the server.

The map database in the server consists of the facility data provided by Japanese government and the data in map software on the market. In relevant word extraction, the system uses Google API [6] as a search engine and acquires top 10 pages of search results.

3.2. Indexing for photos

Figure 7 shows examples of photos which were shot at Yakushiji, a temple in Nara, Japan. The initial map database does not have any building names in "Yakushiji", but it has data of "Yakushiji". We have carried out the indexing from

(a) Kondo 1 (b) Kondo 2 (c) Genjosanzoin

Fig. 7. Photos to be indexed.

Figures 7(a) to (c).

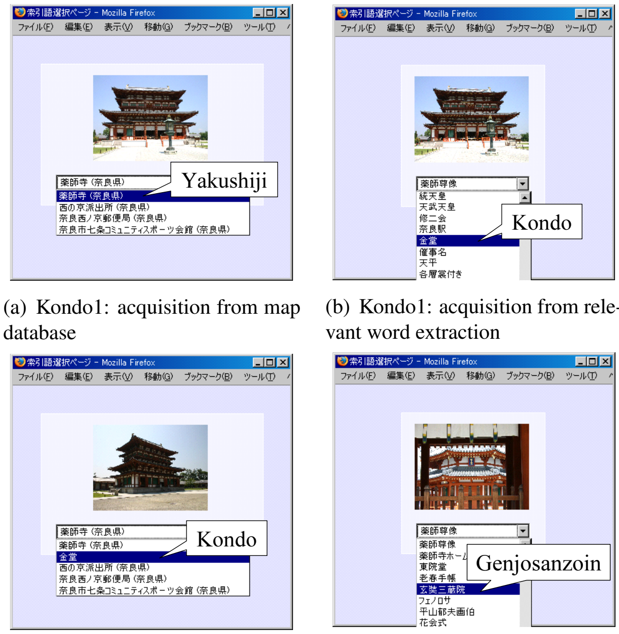
Figure 7(a) shows a photo whose subject is the building "Kondo". Suppose that the user aims to append index "Kondo" to the photo. Since the server has the initial map database, "Yakushiji" is presented as one of the index candidates to the user as shown in Figure 8(a). The user selects "Yakushiji" relevant to "Kondo" from the list of index candidates and pushes the reacquisition button. The server extracts the words relevant to "Yakushiji". The extracted words are shown to the user as new index candidates (See Figure8(b)). Since the list includes "Kondo" (ranked 19th), the user selects it as the index to the photo. After that, the map database is updated with the index as feedback. The number of index candidates obtained by the relevant word extraction is 402 in this case, and almost all of the building names in Yakushiji are included in the top 100 words of them.

Figure 7(b) shows a photo of the same "Kondo", but the shooting position is different from the previous one. As shown in Figure 8(c), "Kondo" is included in the list of index candidates, because "Kondo" has already been stored in the map database at the previous photo indexing. Hence indexing has became easier. As for Figure 7(c), "Genjosanzoin" (ranked 5th) is selected from the list of index candidates shown in Figure 8(d) and is appended to the photo.

4. CONCLUSIONS

In this paper, we have proposed a semi-automatic photo indexing system that enables users to make indexes easily and browse a photo library efficiently. In experiments, we have confirmed the feasibility of adding appropriate index to photos based on shooting position and orientation. We have also confirmed that the system can present more appropriate index candidates from the database after its update during the indexing process.

Future work includes the improvement of the matching between the position data in map database and the estimated subject positions. The current system acquires index candidates in order of the distance between a subject position and a position data in map database. We will consider the sensor accuracy, the depth of field and the area of subject. Other points of future work are concerned with experiments in a wide area by various users. To achieve the experiments, more simple devices like hand-helds will be needed.



(a) Kondo1: acquisition from map database (b) Kondo1: acquisition from relevant word extraction (c) Kondo2: acquisition from map database (d) Genjosanzoin: acquisition from map database

Fig. 8. Selection of index.

5. REFERENCES

[1] K. Toyama, R. Logan, A. Roseway, and P. Anandan, "Geographic location tags on digital images," in *Proc. 11th ACM Int. Conf. on Multimedia*, pp. 156–166, 2003.

[2] D. D. Spinellis, "Position-annotated photographs: A geotemporal web," *IEEE Pervasive Computing*, vol. 2, no. 2, pp. 72–79, 2003.

[3] M. Naaman, Y. J. Song, A. Paepcke, and H. Garcia-Molina, "Automatic organization for digital photographs with geographic coordinates," in *Proc. 2004 Joint ACM/IEEE Conf. on Digital Libraries*, pp. 53–62, 2004.

[4] J. Electronics and I. T. I. Association(JEITA), "Exchangeable image file format for digital still cameras: Exif version 2.2," 2002.

[5] S. Sato and Y. Sasaki, "Automatic collection of related terms from the web," in *Proc. 41st Annual Meeting of the Association for Computational Linguistics*, pp. 121–124, 2003.

[6] "Google Web API." `http://api.google.com/`.