H-005

A Research on Automatic Picture Classification and Image Retrieval

ワファー・イード† 齊藤 剛†

Wafaa Eid Tsuyoshi Saitoh

Abstract

This paper presents an approach for classifying photos based on the image visual contents. Photos can be classified into faces and places. A face detector is used to extract faces from the photos, while a face recognizer is used to classify the detected faces. GPS information which is used to detect places is extracted from geo-tagged photos in a database by image matching. The reverse geo-tag technique enables to find keywords related to the place, and the keywords help to retrieve photos.

1. Introduction

Nowadays the development of technology is so fast that the market of digital cameras is getting bigger as the user demands are getting higher. The number of people who have cameras, either digital one or one included on the mobile phone, is increasing. The minimum price is getting lower, while the memory space is getting bigger. People document their life by taking a lot of photos per day. Storing and organizing such a huge amount of photos can be a challenging task. It will be hard to retrieve a desired photo taken in a certain place with a certain person. Most of the current available image classification systems use manual annotation to tag the GPS data though it's a time consuming task. For these reasons, it's necessary to have an automated classification and retrieval system that can classify photos and suggest keywords for easy retrieval process.

2. Automated Classification and Retrieval System

2.1 System Overview

The system will allow users to upload photos and then it will detect whether the photos contain GPS information or not. If they contain GPS information, keywords will be extracted and added to the photos. If not, the system will search for a similar photo in a geo-tagged database, extract the GPS information from the matched geo-tagged photo and add the keywords to the users' photos. The system flowchart is shown in figure 1.

2.2 System Features

The Global Positioning System (GPS) is a space-based satellite navigation system that provides location and time information in all weather conditions, anywhere on or near the Earth where there is an unobstructed line of sight to four or more GPS satellites [1].

† Graduate School of Information System and multimedia Design, Tokyo Denki University

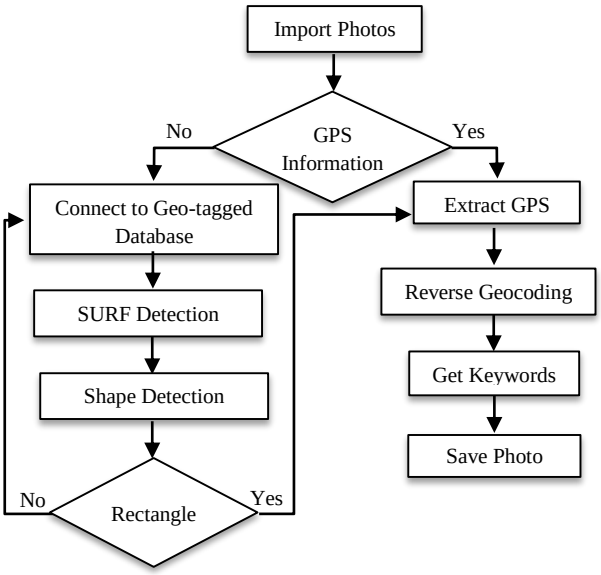


Figure 1. The system flowchart

GPS data can be recorded live in Exchangeable Image File (EXIF) [2] attached to the photo. The system can extract the GPS data by reading the EXIF file attached to the photo. The extracted GPS data can help to get the name of the place using reverse geocoding. Reverse geocoding is the process of translating a point from a map into a human-readable address [3]. Using this feature available from google developer, the system can extract keywords which are related to the place. Such keywords will be helpful in retrieving photos. However, not all the photos can have GPS data in their EXIF. In such cases the system will find a match to the user's photo from a geo-tagged database. The image matching process is explained in details in section 2.3. The system will be able to detect and recognize faces. Hence face detection and recognition have not been developed in the system yet.

2.3 Image Matching

Image matching is to find similar visual contents between 2 photos (a user's photo and a geo-tagged photo in the database). The process of image matching is as follows. First, photos from the user are compared with photos from the database using SURF detection. And then the system will filter the output of SURF detection process by color. Shape detection will be applied to the photo to detect a blue rectangle.

2.3.1 SURF Detection

SURF (Speeded Up Robust Features) is a robust local feature detector which is mainly used in object recognition. It works

by finding a set of distinctive interest points. These points are detected by calculating the Hessian-Laplace detector in different scale spaces. The descriptor will be extracted from the normalized region around the interest points. The matching interest points are calculated by comparing the contrast. If they have the same type of contrast, it's matched [4]. Figure 2 shows the interest points and the matching lines.

Figure 2. Top: Interest points are indicated by blue circles. Bottom: Green lines indicate matched interest points.

2.3.2 Color Filtering

Color filtering is applied to the output of SURF detection process to eliminate all the colors of the image except what's specified. Filtering the image by one color can help in detecting shapes more easily and faster.

2.3.3 Shape Detection

SURF detection is applied to photos imported by users and geo-tagged photos in the database until a rectangle is detected. To make the detection easier, the system will filter the resulted image by color to avoid wrong detection. After that, it will apply an edge detector [5] to detect the rectangle, and the GPS data will be extracted from the geo-tagged photo.

3. Experiment Results

Results show that the system can detect a match between user's photos and geo-tagged photos in the database and extract the GPS data and keywords. Figure 3 shows the result of matching 2 photos taken at the same place.

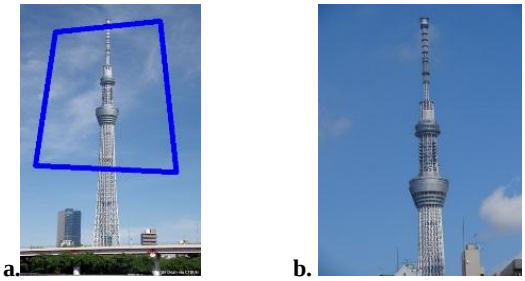
a **b.**

Figure 3. a. User's photo after SURF detection.
b.The matched geo-tagged photo from the database.

Extracted keywords are shown as follows.

"２,３,１丁目,Asakusa,Taito,Tokyo,Japan,JP,浅草雷門,111-0034,Kaminarimon,Tawaramachi Station, １,Nishiasakusa,Asakusa Station,４,Hanakawado,２５,111-0032"

However, objects like tall buildings can be photographed from different locations. Even so, the system will extract the GPS data and keywords from photos in the database no matter where user's photos were taken, which will bring wrong results. An example is shown in figure 4.



a. **b.**

Figure 4. a. User's photo after SURF detection.
b.The matched geo-tagged photo from the database.

4. Conclusion

This paper introduces how photos can be classified into faces and places and how visual contents of photos can be valuable resource for organizing and retrieving purposes. The fast image matching technique enables to add GPS data from geo-tagged photos in a database to user's photos. Based on the results, this method can detect similar geo-tagged photos. It helps to retrieve place information and provide keywords for fast text-based image retrieval. Face detection and recognition should be completed as our future work.

References

[1] Wikipedia, http://en.wikipedia.org/.
[2] EXIF: Exchangeable image file format for digital still cameras: EXIF version 2.2. Technical report, Japan Electronics and Information Technology Industries Association (2002).
[3] Google Developers, https://developers.google.com.
[4] Herbert Bay , Andreas Ess , Tinne Tuytelaars , Luc Van Gool, Speeded-Up Robust Features (SURF), Computer Vision and Image Understanding, v.110 n.3, p.346-359, June, 2008.
[5] J. Canny, A computational approach to edge detection, IEEE Trans. Pattern Anal. Mach. lntell. PAMI-8, 679-698 (November 1986).