

## **Fast collective photographic subject detection without pixels by an assumption about a shoot and its elevation angle**

**Sora Tanioka**

*Department of Computer Science, National Defense Academy of Japan, hasirimizu 1-10-20  
yokosuka, kanagawa 239-8686, Japan*

**Masao Kubo**

*Department of Computer Science, National Defense Academy of Japan, hasirimizu 1-10-20  
yokosuka, kanagawa 239-8686, Japan*

**Hiroshi Sato**

*Department of Computer Science, National Defense Academy of Japan, hasirimizu 1-10-20  
yokosuka, kanagawa 239-8686, Japan  
E-mail: {masaok,hsato}@nda.ac.jp  
www.nda.ac.jp/cs/*

### **Abstract**

In this paper a method that discovers socially attracted photographic subjects in real time is proposed. This needs only non-pixel information of a digital photograph for this discovery which include a time of shoot, GPS location where a photograph is taken, a bearing of the shoot and so on. This information is sufficiently lightweight so that we expect that a congestion-free photo sharing wireless network service can be achieved. In this paper, a photography behavior model which considers an elevation angle of a shoot makes the accuracy better.

*Keywords:* Reality mining; Photograph, Event detection, Geo tagged images

### **1. Introduction**

Interests at the present moment of others can be observed by current Information Communication Technology. Photograph based communication is more vivid and secure than a text based information sharing like Twitter. However, this communication system usually becomes much slow because the heavy congestion caused by transmission of huge pixel data. Therefore this framework has not been used for a real-time information sharing system. So far researches which attempt to improve this lower real time capability have been proposed. It has been proposed a method of transmitting only the best photos that match the viewing

request to the server while they are categorized only by not-pixel information of digital photograph. The information includes time to shoot, GPS and bearing of shoot so on. According to this method, photo pixel data congestion is unlikely because they are transmitted sequentially. However, in the conventional method, there is a practical problem that estimation accuracy of the object taken from the biased direction is worse. In order to solve this issue in this paper, we propose a method for estimating sterically an object by using the elevation angle (Fig.1). It is intended to improve the estimation accuracy by noting the characteristic of

tilting the camera in order to capture the entire subject into photographic frame.

This paper is organized as follows. First, related studies are introduced. Next the algorithm by using the elevation angle information is proposed. Finally, computer simulations perform, we verify the validity of the algorithm.

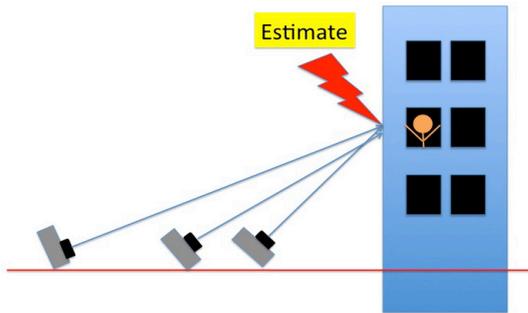


Fig.1 The concept of the proposed method: estimation of subject by a collection of elevation angle data in non pixel information in digital photographs.

## 2. Related works

Kubo et al. [1] introduces the photographer behavior model using the angle of view by composition. Thereby it realizes a method of automatically detecting its main subject in each image. As shown in Fig.2(a) the method that adopts a planar estimation in estimating the subject can be very accurately estimated when it is taken with a wide angle to the object. However, such a examples are rare in the real environment. Also when it is taken from a narrow angle, it can be pointed out that the accuracy is hardly improved even if information is increased as shown in Fig.2 (b).

## 3. The proposed system

### 3.1. Input data

Here, small mobile devices such as a smart phone is our main subject. The photographic information includes latitude, longitude, altitude, azimuth of GPS, elevation angle by accelerometer, and time to shoot in TIFF and Exif data .

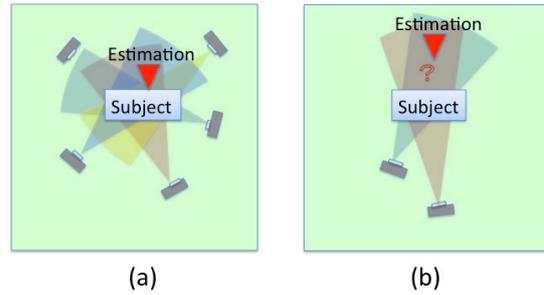


Fig.2 (a)A case conventional methods solve. (b) A difficult case for the conventional methods.

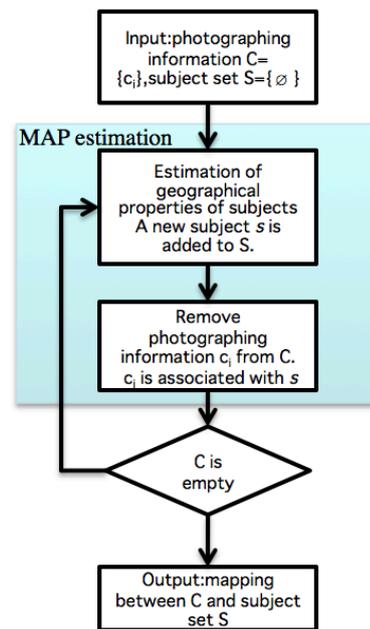


Fig.3 Subject location discovery algorithm by MAP estimation

### 3.2. Flow of the proposed system

Fig.3 describes the proposed algorithm. In this figure a set of photographic information is described by  $C = \{c_1, \dots, c_N\}$ . The each subject is estimated by MAP (Maximum a posteriori) estimation iteratively. Once a subject is found,  $c_i$  which seems to be a consequence of shooting the subject is removed from  $C$ . By this procedure, the proposed method can discover multiple subjects.



Fig.4 Search result of ‘building+photograph’ by Google Image Search

### 3.3. Photographer behavior model

The photographer behavior model is a probability distribution  $P_{shoot}(c_{Ti}|s_j)$  (that a photographer at  $c_{Ti}$  takes a photo for a given subject  $s_j$ ).

$$P_{shoot}(c_{Ti}|s_j) = K_{motivation} \cdot V_{composition}(s_j). \quad (1)$$

where  $K_{motivation}$  represents the motivation to take a photo and  $V_{composition}(s_j)$  means the fitness of the composition. This consists of the elevation angle  $E$  and the bearing  $B$ .

$$V_{composition}(s_j) = f(E)f(B) \quad (2)$$

However, in this paper, we assume that the photographs that place a subject in the center of the frame are useful to viewers and  $f(B)=1$ .

### 3.4. Algorithm for estimation of subjects

First, the overview of the MAP estimation is described.

$$\log \varpi_{j,k} = \sum_{x=1}^{|c_j|} \max \{ \log P_{shoot}(c_{Tx}|o_{j,k}, \Delta_x), l_0 \} + l_0 \quad (3)$$

For each grid  $s_j$  in a given area, a weight as  $P_{shoot}(c_{Ti}|s_j)$  is added about  $C$ . After this weighting step, a grid of maximum weight is selected as a subject  $s$  in this iteration.  $c_i$  which seems to take a photo of this subject is removed and  $s$  with  $c_i$  is added the subject set  $S$ . The next iteration starts with the new  $C$ . If  $C$  is empty the algorithm stops..

### 3.5. Photographer behavior and elevation angle

Fig.4 illustrates the top 20 of the search result of “building” by Google image search. As these images tell, the elevation angle is nearly horizontally when a photo is taken from far. On the other hand, when it is taken from the vicinity of a building the elevation angle becomes large in order to include the top of the building in the photo. Thus it was found to be estimated by the elevation angle at the time of photographing the distance from the size and shooting point of the object.

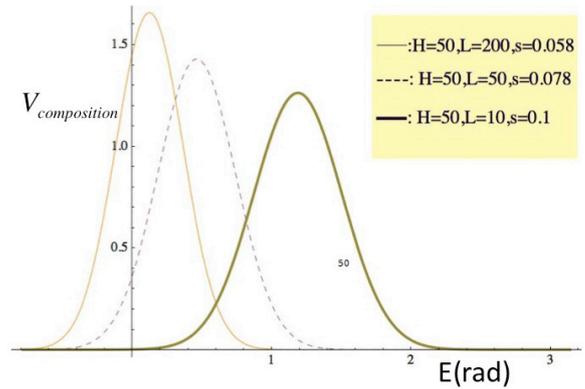


Fig.5 Examples of the proposed photographer behavior model of elevation angle

### 3.6. The proposed photographer behavior model of elevation

The photographer behavior model of elevation  $f(E)$  is proposed in this section. This  $f(E)$  is used for the Map estimation procedure in the algorithm in 3.4.

$$V_{composition}(s_j) = f(E) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(E-D)^2}{2\sigma}\right) \quad (4)$$

$$D = \arctan\left(\frac{H}{2L}\right) \quad (5)$$

The composition of a photograph is evaluated by an elevation angle  $E$ . We suppose that distribution of  $E$  obeys a normal distribution.  $D$  is the best elevation angle and  $H$  is the height of the subject and  $L$  is the distance to the subject.  $\sigma$  is a parameter.

Fig.5 shows examples of this  $f(E)$ . The motivation  $K_{motivation}$  is constant among all photographer. The y axis indicates  $E(\text{rad})$ . The x axis shows the evaluation  $V_{composition}$ . There are 3 lines, namely the thin line, the dotted line, bold line which represent the evaluation in eq.(4) for a tall building of the height of 50 m. The thin line shows a case when a photographer takes it from far and it can takes a good photograph but it seems to be

difficult because the suitable elevation angle is less than the 2 remaining cases. This trend seems to be consistent with our experience.

#### 4. Experiment

By performing an experiment by computer simulations to evaluate the performance of the proposed method. The experiment is conducted in the configuration of estimating the subject that can be taken only from the direction of the difficult had been part in the conventional technique. Only one subject with the height of 40 meters sets at  $(x,y) = (0,0)$ . The dataset of photographic information is generated by Metropolis-Hastings algorithm. For reduction of memory resources, each elevation angle is discretized into 18 pattern from 0 degree to 85 degree.

Fig.6 shows the result of estimation of location of the subject. The bullet at  $(x,y)=(0,0)$  indicates the subject. The result by the traditional method is  $(x,y)=(4,-46)$  represented by '+' mark. On the other hand, the result of the proposed method is  $(x,y)=(-6,8)$  shown by a 'x' mark. The proposed method can improve the accuracy. The Fig.7 shows the result when there are 2 subjects in a given field and the permitted area to take a photo is restricted also. In this case, our method can estimate the subjects' location simultaneously.

#### 5. Conclusion

Collect only the shooting information of digital photos, we have proposed a method for estimating a subject that has attracted the social attention. The proposed a photographer model using elevation and improved conventional been difficult case. Consequently, it was possible to improve the practicability.

This work is partially supported by JSPS 15K00433.

#### References

1. Y. Kubo, M. Kubo, H. Sato, M. Hirano, and A. Namatame, Understanding Geographic Attentions of Crowd from Photographing Information in *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol. 17, No. 6(2013) 890-903.

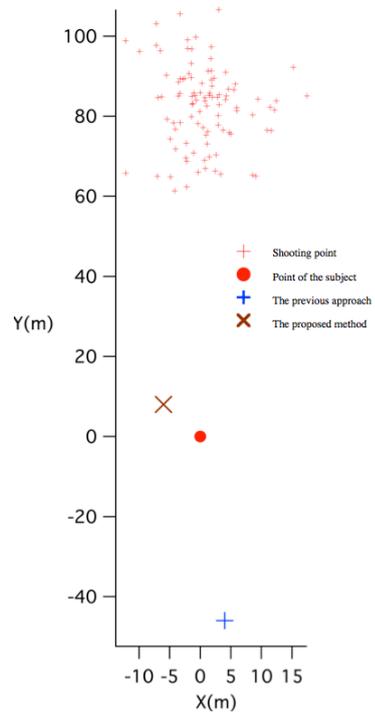


Fig.6 the result of the experiment

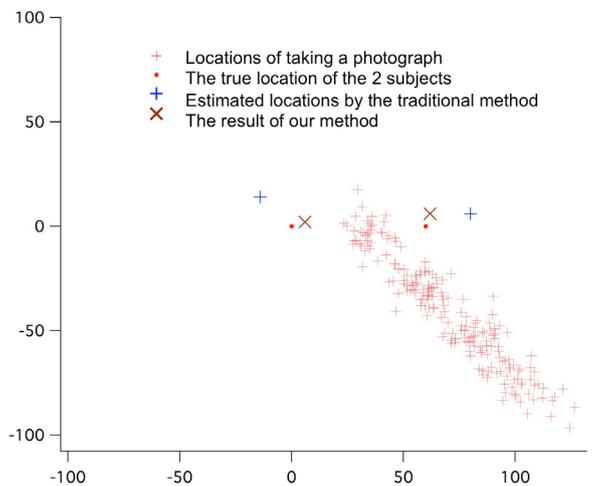


Fig.7 the result of the experiment