Detecting a Taxi from a Video for Visually Handicapped People

著者	Nishimura Akihiro, Tan Joo Kooi, Kim Hyoungseop, Ishikawa Seiji
journal or publication title	SICE Annual Conference 2015
volume	2015
page range	89-92
year	2015-07-28
URL	http://hdl.handle.net/10228/5682

SICE Annual Conference 2015
July 28-30, 2015, Hangzhou, China

Detecting a Taxi from a Video for Visually Handicapped People

Akihiro Nishimura[1], Joo Kooi Tan[1], Hyoungseop Kim[1] and Seiji Ishikawa[1]

[1]Department of Mechanical and Control Engineering, Kyushu Institute of Technology, Kitakyushu, Japan
(Tel : +81-93-884-3191; E-mail: {nishimura, etheltan, ishikawa}@ss10.cntl.kyutech.ac.jp, kim@cntl.kyutech.ac.jp)

Abstract: This paper proposes a method of detecting a specific moving object, a taxi in particular, on a road from a video provided from a camera attached to a user. In order to raise the quality of life of visually handicapped people, a computer vision system which works in place of their eyes and a brain may be useful. As one of such systems, this study focuses its attention on finding a taxi on a road which is a convenient vehicle to such people as a means of transfer outdoors. The novel idea of this study is that a camera and a PC system for finding a taxi is carried by a user, a visually handicapped person, for example. The proposed method employs the HOG features to represent a vehicle, and finds a taxi by Real AdaBoost and color information with the detected vehicle. The performance of the proposed method is shown experimentally.

Keywords: Object detection, taxis, motion vectors, HOG, Real AdaBoost, color.

1. INTRODUCTION

According to the 2013 annual report issued by Cabinet Office, Government of Japan [1], there were approximately 310,000 visually handicapped people in 2006 in Japan. Although various kinds of support have been provided to them, they still have difficulties in enjoying daily activities. It is therefore strongly requested even to computer vision community to raise their quality of life (QoL) more by inventing a tool or a system which can substitute for their eyes and a brain.

As one of such a system, this paper proposes a method of detecting a specific moving object, a taxi in particular, from a video automatically by image processing. It intends to find a cruising taxi outdoors to raise the mobility of a visually handicapped person. Han et al. [2] proposes a method of detecting people and vehicle. But they do not concentrate on a taxi. Detection of vehicles is a popular research subject in car vision field, but the detection of a specific vehicle such as a bus or a taxi is not very popular and is more difficult than detecting general vehicles. Finding public transportation vehicles such as a bus by a computer vision system may also be worth studying. It is left for another study [3], however.

The present study concentrates on developing a system that finds a taxi on a road by the employment of a computer vision system. The novel idea of the proposed system is that a user carries the computer vision system, i.e., a camera and a PC, by himself/herself. Since a blind user can find a road by sound, the camera takes images of the road to detect a taxi employing the PC. The method is addressed and the experimental results are shown.

2. OVERVIEW OF THE METHOD

The flow of the proposed method is shown in Fig. 1. The procedure is explained in the following.

Step 1: Initially two image frames are fed into the program: From the second step, a single image frame is fed.
Step 2: Motion vectors are detected from successive image frames by use of Harris corner detector and pyramidal Lucas-Kanade tracker.
Step 3: Regions of interest (ROIs) which contain locally distributed motion vectors having a similar orientation are set on the motion vector image.
Step 4: The ROI is analyzed if it contains a vehicle employing the HOG features and Real AdaBoost.
Step 5: If the ROI is judged as containing a vehicle, it is further analyzed to judge a taxi employing color information.
Step 6: The result is displayed on the LCD.
Step 7: The above procedure is repeated until all the image frames are processed.

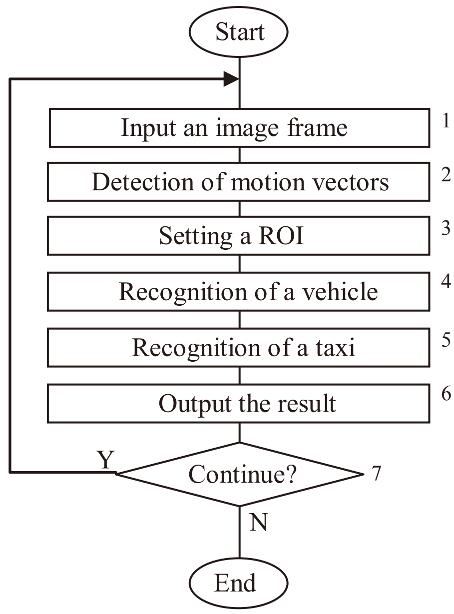


Fig. 1 Flow of the proposed method.

3. DETECTING A SPECIFIC MOVING OBJECT

Obviously the familiarity in utilizing public transportation raises the QoL of the handicapped, particularly of visually handicapped people. Public bus services deserve for it. But taxi services could be contributing more except for its fare because of its anytime and anywhere nature.

The proposed method first detects motion vectors different from camera motion on successive image frames by the employment of Harris corner detector [4] and Lucas-Kanade tracker [5]. If several motion vectors gather in a local area, they are bounded by a box. Then the image area in the bounding box is transformed into the HOG features [6] representation and examined by a strong classifier produced by Real AdaBoost [7] if it contains a vehicle, actually a sedan-type car. If it is detected, the existence of a green number plate and a specific colored mark on the roof are searched, which certifies a taxi.

3.1 Detecting motion vectors

In detecting motion vectors from an image sequence, a stationary camera is assumed in this paper for simplicity. Since the proposed system is to be carried by a user, the camera attached to the user's head or body may move according to the user's motion. In this case, camera motion is compensated from a fed image sequence and the motion vectors are solely detected. However this won't be discussed in this particular paper.

An example of motion vector detection is given in Fig. 2, in which (a) is the result of feature points detection and (b) is that of motion vectors detection.

3.2 Setting a region of interest

Using the detected motion vectors, region of interests (ROIs) that contain a vehicle are set on the image. The motion vectors are grouped referring to their magnitude and orientation, and the centroid of the motion vectors is calculated with each group. Then a predefined-size window whose center coincides with the centroid is set on the image. This gives a ROI. The size of the window is experimentally given. This works when successive vehicles are mutually distant. If they are close with each other, the ROI of the main vehicle might contain part of another vehicle. But this difficulty may be overcome at the recognition stage.

(a) (b)

Fig. 2 Example of motion vectors detection on a vehicle: (a) Detected feature points, (b) detected motion vectors.

3.3 Recognition of a vehicle

The detected ROI is analyzed if it contains a sedan-type car employing the HOG features and AdaBoost technique. As is well known, this paired recognition technique is strong with human detection [6]. It may also give satisfactory performance in vehicle recognition even if the setting of a ROI is not very strict, only if it contains main part of the vehicle on the background.

4. IDENTIFYING A TAXI BY COLOR INFORMATION

Once a sedan is found in an image, specific color information is examined in the ROI. Figure 3 depicts a typical taxi found everywhere in Japan. Its number plate circled below must be dark green by law, whereas the shape and color of the mark circled on the roof depends on a company. Instead of evaluating the shape of the mark, the present method examine its color. Thus the judgment on if the vehicle in the extracted ROI is a taxi is done using local color information. If the number of the pixels in the ROI having a specified color exceed a certain threshold, the vehicle in the ROI is judged as a taxi.

Alternatively, detection of a part protruded from the roof of a sedan-type vehicle may be an enough proof for a taxi. But it is not employed in this study, since it might be embedded in the messy background.

In a practical environment, a color changes its brightness and saturation, and the relation is given by $f(R,G,B)=0$. Namely a single color corresponds to this relation. Then the number of pixels S having a particular color in a ROI W is given by

$$S=\sum_{|f(R,G,B)(x,y)|<\varepsilon \text{ for } p(x,y)\in W} 1 . \tag{1}$$

Here $p(x,y)$ is a pixel at (x,y) and R, G and B are the RGB values of $p(x,y)$. The threshold ε is a small positive number.

Let us denote S at frame k by S_k. Then the vehicle in the ROI W at frame k is judged as a taxi, if S_k exceeds a certain threshold θ, i.e., if

$$S_k \geq \theta \tag{2}$$

holds. This judgment is repeated with every successive frame.

It is noted that the function $f(R,G,B)=0$ can be multiple, if multiple colors represent a taxi.

Fig. 3 A typical taxi with characteristic color and shape information.

5. EXPERIMENTAL RESULTS

The data set employed for training the Real Adaboost is 1,956 sedan-type vehicle images and 10,000 negative images. Some of them are shown in Fig. 4.

The particular taxi chosen for recognition has a black body and the main color of its roof-top mark is close to blue. This color was collected from a number of the specified taxis and the RGB relations were analyzed. Figure 5 shows the relation between *R* and *G*, and *R* and *B* with the color. They are approximately formulated by

$$G = -0.0003R^2 + 1.0142R - 0.9182$$
$$B = 0.0003R^2 + 0.8793R + 41.157 \tag{3}$$

This gives the function $f(R,G,B)=0$ in Eq. (1). The thresholds are $\varepsilon_G = 3$ and $\varepsilon_B = 5$.

It is noted that the green of the number plate is not used as color information in this experiment, since the frontal part of the taxi sometimes become darker and the green color cannot be extracted well.

Two evaluation functions given by Eq.(4) and Eq.(5) are introduced to judge if a ROI contain a vehicle appropriately.

$$cover = \frac{n(GT \cap OA)}{n(OA)} \tag{4}$$

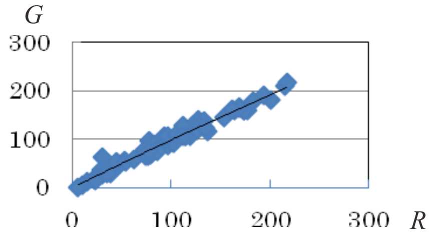
$$overlap = \frac{n(GT \cap OA)}{n(GT)} \tag{5}$$

Here *GT* is the Grand Truth region containing a vehicle, and *OA* is a ROI (Object Area). Figure 6 shows them visually. Function $n(R)$ gives the number of the pixels contained in the image area *R*.

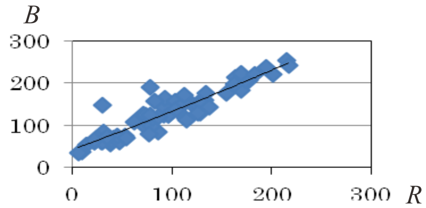
Above two evaluation functions are calculated with each ROI and the following judgment is done.

$$\begin{cases} A\ Moving\ Object & \text{......}\ cover\ \&\ overlap > 0.4 \\ Not\ a\ Moving\ Object & \text{......}\ otherwise \end{cases} \tag{6}$$

Fig. 4 Examples of the sedan-type vehicle images employed for training.



(a)



(b)

Fig. 5 The RGB relations with respect to the blue of the roof-top mark: (a) *R*-*G* relation, (b) *R*-*B* relation.

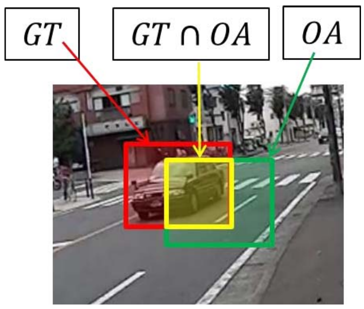
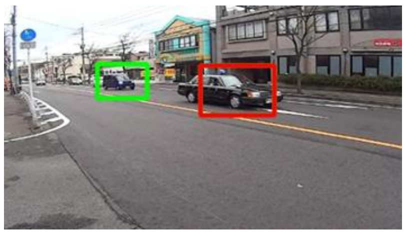


Fig. 6 Definition of GT and OA regions on an image.



(a)

(b)

Fig. 7 Performance of the proposed method: (a) Video_1, (b) video_2. A green box indicates a normal car, whereas a red box indicates a taxi.

If both values of cover and overlap are larger than 0.4, a moving object is judged as detected correctly, which is *TP* (True Positive). If the background area was detected, it is *FP* (False Positive).

Two videos, video_1 and video_2, are used to examine the performance of the proposed method. Both videos contain a taxi followed by a sedan (not a taxi). The vehicles in video_1 run faster than those in video_2. Table 1 shows the results. The detection rate of a moving object was 90% in average. But the detection rates of a taxi cruising a distant place were 24.5% and 40.6%, respectively. If the distance is restricted to within 30m, the detection rate was 84.5%, which is given in Table 2. The performance of vehicles extraction is shown in Fig. 7, in which a car in a green box was judged as a normal car (not a taxi), whereas the one in a red box was judged as a taxi.

6. DISCUSSION

When we look for a taxi on the street, we normally search for a sedan and check if it has a small protruded part on the roof. It is the mark of a taxi company having a particular shape, color and often letters. It is not difficult for us to find such a mark on a vehicle. But it is not very simple to find it by a camera and a computer. Although the proposed method obtained some results experimentally, further improvements are necessary.

It is obvious that a taxi cruising a distant place is not easy to be detected only by color information, since the roof-top mark is small and the number of the pixels having a particular color is small. It may be advantageous to add shape information by examining the roof of a detected vehicle if it has a protruded part on it. This needs investigation in the next stage of this study.

Even in a crowded traffic, respective vehicles need to be discriminated. It is difficult particularly on closer side of the road than the farther side of the road. The method has to take occlusion between successive vehicles into account. This also remains for further study.

7. CONCLUSION

This paper proposed a method of detecting a specific moving object, a taxi in particular, from a video taken in the street. It was shown by the experiment that a taxi distant from the present position is not very easy to be detected, whereas, when it comes closer, the detection becomes easier.

As the purpose of the present study is to realize a system which can help those visually handicapped people with their outdoor mobility, the proposed method will be improved further to achieve higher detection rate of a cruising taxi.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant No.25350477.

REFERENCES

[1] Cabinet Office of Japanese Government: Annual Report on Government Measures for Persons with Disabilities, pp.5-6, 2013.

[2] F. Han, Y. Shan, R. Cekander: "A Two-stage approach to people and vehicle detection with HOG-based SVM", PerMIS, pp.133-140, 2006.

[3] S. Hatano: Detecting a Specific Moving Object from a Self-wearable Camera Images Using Co-occurrence Feature, Master Thesis, Kyushu Inst. Tech. (2015)

[4] C. Harris, M. Stephens: "A combined corner and edge detector", Proc. 4th Alvey Vision Conf., pp.147-151 (1988).

[5] B. D. Lucas, T. Kanade: "An iterative image registration technique with an application to stereo vision", Proc. 7th Int.. J. Conf. on Artificial Intelligence, pp. 674-679 (1981).

[6] N. Dalal, B. Triggs:"Histograms of oriented gradients for human detection", Proc. IEEE Conf. on Comput. Vision and Pattern Recogn., pp.886-893 (2005).

[7] R. E. Schapire, Y. Singer:"Improved boosting algorithms using confidence-rated predictions", Machine Learning, Vol. 37, No. 3, pp.297-336 (1999).

Table 1 Evaluation on the precision of the detection.

	Number of frames	Precision on Moving objects (%)	FPR on Moving objects (%)	Precision on Vehicles (%)	Precision on Taxis (%)	Process. Time (msec/frame)
Video 1	140	85.5	30.4	55.1	40.6	43.1
Video 2	80	95	13.2	91.4	24.5	49.2

Table 2 Evaluation on the precision of the detection. The case where a vehicle is within 30m from the camera.

	Number of frames	Precision on Moving objects (%)	FPR on Moving objects (%)	Precision on Vehicles (%)	Precision on Taxis (%)	Process. Time (msec/frame)
Video 1	71	87.3	11.9	85.9	84.5	78.5