The Sixteenth International Symposium on Artificial Life and Robotics 2011 (AROB 16th '11),
B-Con Plaza, Beppu,Oita, Japan, January 27-29, 2011

Detection of a Bicycle and Its Driving Directions Using HOG Feature

Heewook Jung, Joo Kooi Tan, Hyoungseop Kim, Seiji Ishikawa

Department of Mechanical & Control Engineering, Kyushu Institute of Technology, Kitakyushu, Japan
(Tel: 81-93-884-3183; Fax: 81-93-884-3183)
{*jung, etheltan, ishikawa*}*@ss10.cntl.kyutech.ac.jp*, *kim@cntl.kyutech.ac.jp*

Abstract: Studies on car vision have currently been practiced around recognizing a human enthusiastically. The Histograms of Oriented Gradients (HOG) feature has been proposed as useful feature for recognizing a human standing in various kinds of background. On the other hand, although a bicycle is important transportation vehicle in urban environment, its automatic recognition or detection is not an easy task for a computer vision system, because bicycle's appearance can change dramatically according to viewpoints and a person riding on the bicycle is a non-rigid object. Thus, automatic bicycle detection is an important research subject in an intelligent perception system using car vision. In this paper, we propose a method of detecting a bicycle and its driving direction using the HOG feature and RealAdaboost algorithm. Experimental results show satisfactory performance of the proposed method.

***Keywords*:** Bicycle detection, car vision, HOG feature, RealAdaboost.

I. INTRODUCTION

An automated perception system with a car is currently very helpful in preventing traffic accidents and contributes to decreasing traffic accidents. But it is difficult to prevent traffic accidents perfectly. Factors that endanger traffic safety still remain around a moving car. If the detection of the risk of traffic accidents (a human, a bicycle, a car, a bicyclist, etc.) is employed in automated perception system, the number of traffic accidents will further decrease.

The Histograms of Oriented Gradients (HOG) feature has been proposed by Dalal and Triggs [1] as useful feature for recognizing a human standing in various kinds of background. The HOG feature proposed by them seems an effective feature for representing and recognizing a human image. But, in their paper, it includes an unnecessary part of the image such as the background. It is not conceivable that the background feature contributes to high recognition rate of a human image. Zhu et al. [2] employ the HOG feature based on variable block size, but they also include the background in the feature. In this paper, we propose an object detection method employing an improved HOG feature.

The research on car vision is currently practiced around recognizing a human and not a bicycle which is important transportation in urban environment. Although the number of whole traffic accidents is decreasing, the rate of bicycles' accidents to the whole traffic accidents is increasing. Detection of a bicycle by a computer vision system is, however, not an easy task, because bicycle's appearance can change dramatically among viewpoints and a person riding on a bicycle is a non-rigid object.

In this paper, we propose a method of detecting a bicycle and its driving direction using improved HOG feature and RealAdaboost algorithm.

II. LEARNING ALGORITHM

A learning algorithm is used to HOG feature and RealAdaboost. The proposed method can detect bicycle's driving direction, because we make a bicycle model using HOG feature. The algorithm of learning is given in **Fig 1**.

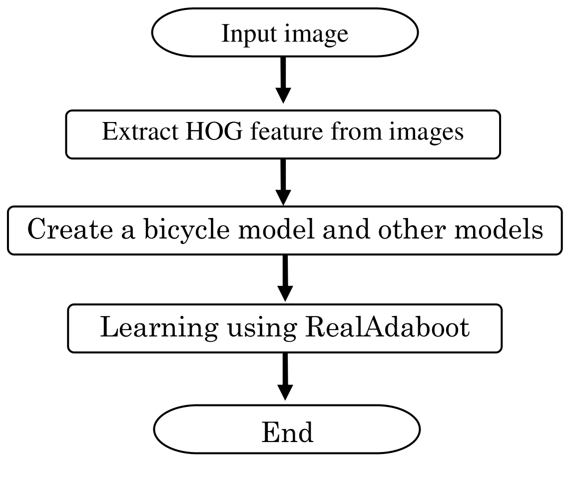


Fig. 1 Overview of the learning algorithm.

1. HOG feature

The HOG feature [1] is a method that extracts feature of outlines in images. Because gradients of adjacent pixels receive histogramization and normalization at each area, it is robust to illumination change and to geometric change.

The HOG feature is extracted in the following way.

Step1: Compute the magnitude $m(x, y)$ and orientation $\theta(x, y)$ of images using equation (1). Given coordinates of images $I(x, y)$,

$$m(x, y) = \sqrt{f_x(x, y)^2 + f_y(x, y)^2}$$
$$\theta(x, y) = \tan^{-1} \frac{f_x(x, y)}{f_y(x, y)} \tag{1}$$

Here,

$$f_x(x, y) = I(x+1, y) - I(x-1, y)$$
$$f_y(x, y) = I(x, y+1) - I(x, y-1)$$

The result of the oriented gradients image is shown in **Fig. 2**.

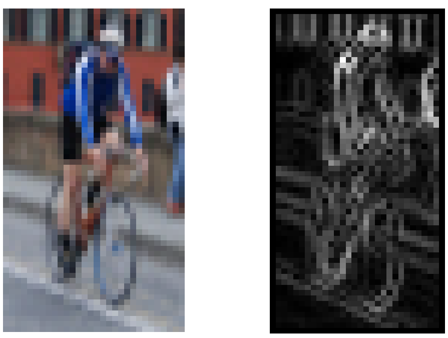
Step2: Derive the orientation histogram from the orientations and magnitudes.

Oriented gradients image has a large amount of information. But if there is such an amount of information, it needs more time to classify and to learn in the procedure of pattern recognition. Deriving the orientation histogram from the orientations and magnitudes has an effect of making the amount of information decrease.

Here each cell size is 5×5 pixels and the orientation histogram has 9 bins with each cell.

Step3: Perform histogram's normalization using an overlapping block.

Here, each block size is 3×3 cells. Each block contains 9×9=81 features and each 30×60 sub-image contains 4×10×81=3240 features.



(a) (b)

Fig. 2 A processed bicycle image: (a) An input image, and (b) its oriented gradients image.

2. A bicycle model

Many researchers calculate the HOG feature from an entire image irrespective of the foreground or the background. We have proposed an improved HOG feature calculation by the employment of a bicyclist mask and a bicycle mask. This employment shows better performance in the bicycle detection/recognition. This improved HOG feature is employed in the proposed method of this paper.

The main idea of the proposed method is to calculate the HOG feature on a bicyclist and a bicycle mask instead of calculating it on the whole image. A bicyclist and a bicycle mask are created in the following way.

Step1: The intensity of gradient is calculated of all the images in an image database.

Step2: The average values of the intensity of the gradient are calculated at every pixel on the images and a normalized average gradient intensity image is produced.

Step3: An edge image is produced from the average gradient intensity image.

Step4: A silhouette image is created from the average gradient intensity image and a skeleton image is made from the silhouette image.

Step5: Finally a bicyclist and a bicycle mask are yielded by performing logical OR operation between the edge image obtained at step 3 and the skeleton image obtained at step 4.

3.RealAdaboost

The adaboost algorithm is a method of uniting weak classifier of a simple hypothesis and generating a strong classfier. A bicyclist and a bicycle model are defined and trained via a RealAdaboost to detect bicycles under various circumstances. In training using RealAdaboost, we use 1000 positive training samples and 2000 negative samples and we train a two-component (a bicyclist and a bicycle) model. Effectiveness of the proposed method is shown by experiment. The algorithm of RealAdaboost is given in **Table 1**.

Table.1 RealAdaboost algorithm.

1. Suppose $S = \{(x_1, y_1), (x_2, y_2), \cdots, (x_N, y_N)\}$ as a sample space, where $x \in X$ are feature vectors and $y \in \{-1, +1\}$ are labels.

2. D_t is the initial distribution

$$D_1(n) = \frac{1}{N}$$

3.
for t=1 to T do
for m=1 to M do
Compute probability distribution W_t of weak classfier $h_i(x)$

$$W_+^j = \sum_{i,j \in J \wedge y_i = +1}^{n} D_t(i)$$

$$W_-^j = \sum_{i,j \in J \wedge y_i = -1}^{n} D_t(i)$$

Compute estimation Z_m

$$Z_m = 2\sum_j \sqrt{W_+^j W_-^j}$$

Select the smallest Z_m of the weak classifier

$$h_t = \arg\min Z_{t,m}$$

Renew sample weight $D_t(i)$

$$h_t(x_i) = \frac{1}{2}\ln\frac{W_+^j + \varepsilon}{W_-^j + \varepsilon}$$

$$D_{t+1}(i) = D_t(i)\exp[-y_i h_i(x_i)]$$

4. The final strong classifier is given by

$$H(x) = sign\left(\sum_{t=1}^{T} h_t(x)\right).$$

III. DETECTION ALGORITHM

A detection algorithm is used to a bicycle model using HOG feature. And we use five mask models for detecting driving directions. The algorithm of detection is given in **Fig 3**.

1. Merge processing

When detection of a bicycle is performed using HOG feature, many windows containing a single common bicycle are obtained in an image. They are then merged into a single window using mean shift clustering and the nearest neighbor algorithm.

A. Mean shift clustering

The first step of mean shift clustering is assumption of density using kernel function. Here we use a kernel function represented by

$$k(x) = \begin{cases} c(1-\|x\|) & \|x\| < 1 \\ 0 & otherwise \end{cases} \qquad (2)$$

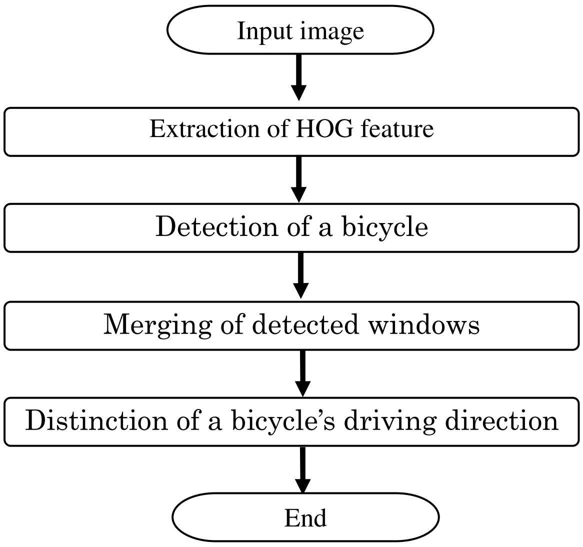


Fig. 3 Overview of the detection algorithm.

The second step is shifting the mean to high density employing a mean shift vector. A mean shift vector is given as follows;

$$m(x) = \frac{\sum_{i=1}^{n} x_i k\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} k\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} - x \qquad (3)$$

Here, the first term of the mean shift vector assumes density and the second term shifted mean. When the mean shift vector is 0, the density is the highest.

B. Nearest neighbor

Euclid distance d is computed from the point of the highest density to each sample point.

$$d = \sqrt{(x-x_i)^2 + (y-y_i)^2} \qquad (4)$$

Merging is executed, if d is smaller than a predefined threshold. The result of the merging processing is shown in **Fig. 4**.

(a) (b)

Fig. 4 Merge processing: (a) Before merging, and (b) after merging.

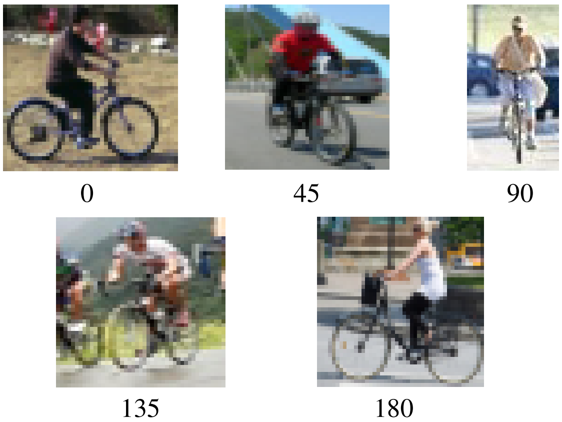


Fig. 5 Bicycle's five driving direction.

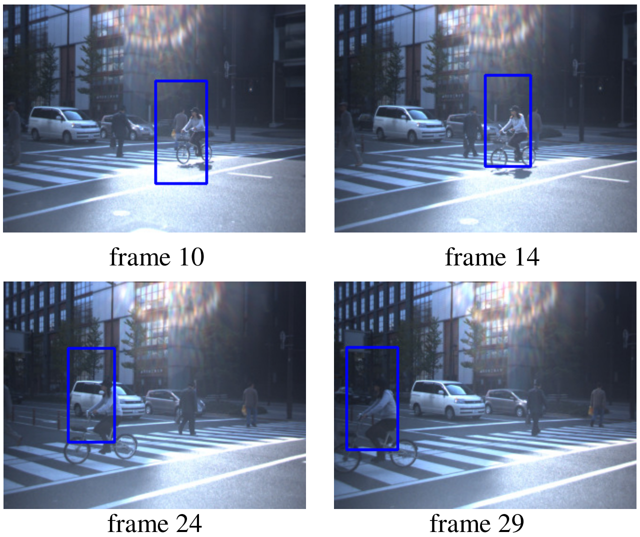


Fig. 6 Experimental results (135).

2. Detection of driving direction

We use five mask models (See **Fig. 5**) for detecting driving directions. They are 0, 45, 90, 135 and 180 degrees.

IV. EXPERIMENTAL RESULTS

Experiments have been done on a video containing a bicyclist and a human passing in front of a camera. **Fig. 6** and **Fig. 7** show successful cases, but a human itself is not detected. Because learning data of the positive class is a bicycle and a bicyclist.

V. CONCLUSIONS

In this paper, we propose a method of detecting a bicycle and its driving direction using improved HOG feature and RealAdaboost algorithm. This method has several advantages over existing methods. This method calculates the HOG feature on a bicyclist mask and on a bicycle mask instead of calculating it on the whole image. Therefore, the computation time of this method is fast. Moreover, this method is applicable to the detection of various objects. In future, we plan to enhance the performance of the detection and to detect various objects.

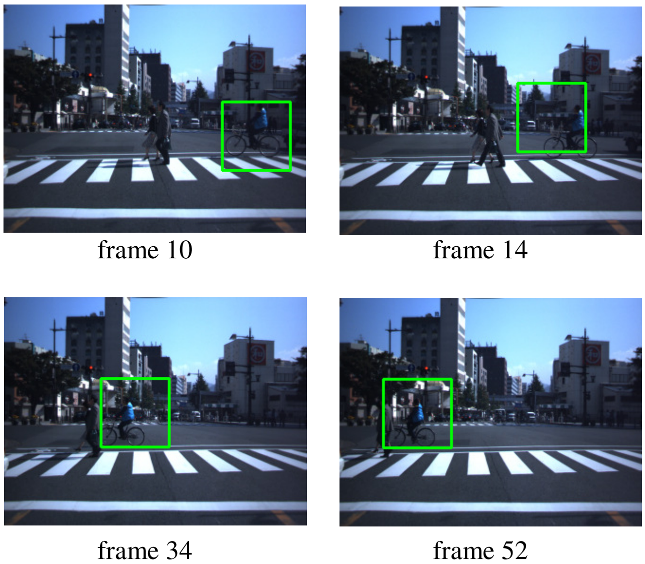


Fig. 7 Experimental results (180).

ACKNOWLEDGEMENT

This work was supported by JSPS Grant-in-Aid for Scientific Research (C) 22510177

REFERENCES

[1] N. Dalal, B. Triggs, "Histograms of oriented gradients for human detection", Proc. IEEE Conf. on Computer Vision & Pattern Recognition, 886-893, 2005.

[2] Q. Zhu, et al, "Fast human detection using a cascade of histograms of oriented gradients", Proc. IEEE Conf. on Computer Vision & Pattern Recognition, 2006.

[3] K. Khattab, J. Dubois, and J. Miteran, "Cascade boosting-based object detection from high-level description to hardware implementation", EURASIP Journal on Embedded Systems, 2009.

[4] Y. Nakashima, J. K. Tan, S. Ishikawa, T. Morie, "On the detection of a human and its body direction from a video", Proc. of 15^{th} Int. Symposium on Artificial Life and Robotics, 294-297, 2010.

[5] R. E. Schapire and Y. Singer, "Improved boosting algorithms using confidence-rated predictions", Machine Learning, **37**, 3, 97-336, 1999.