Contrast HOG and Feature Spatial Relocation based Two Wheeler Detection Research using Adaboost

Yeunghak Lee[1,*], Jaechang Shim[2]

Abstract

This article suggests a new algorithm for detecting two-wheelers on the road that have various shapes according to viewpoints. Because of complicated shapes, it is more difficult than detecting a human. In general, the Histograms of Oriented Gradients(HOG) feature is well known as a useful method of detecting a standing human. We propose a method of detecting a human on a two-wheelers using the spatial relocation of HOG (Histogram of Oriented Gradients) features. And this paper adapted the contrast method which is generally using in the image process to improve the detection rate. Our experimental results show that a two-wheelers detection system based on proposed approach leads to higher detection accuracy, less computation, and similar detection time than traditional features.

Key Words: Contrast, Histogram of oriented gradients, Adaboost, Two wheelers.

I. INTRODUCTION

In the past, a vehicle was regarded as a simple tool that transports people and commodities. However, the shapes and functions of vehicles continue to develop rapidly and in diverse ways. The development of vehicles was concentrated on not only improving performance, but also to protect the drivers and passengers in a vehicle in the occurrence of a traffic accident. The degree of injury to passengers in a vehicle caused by a traffic accident is reduced gradually in advantage of many safety device techniques mounted in and outside of the vehicle. But the accidents outside of a vehicle that is caused by a driver's carelessness or by road environments still have many problems, so it is required to detect the risk of a traffic accident in order to save lives [1].

Most of the studies relating to human safety over the past several years have mainly concentrated on increasing the detection rate of pedestrians and automobiles on a road from a still image and framework. Now, the study scope has expanded to protect vulnerable road users(VRUs) such as small automobiles [1, 2]. Because pedestrians and bicycle riders are very vulnerable road participants among VRUs, they are a hot subject for the study field in intelligent transportation system. Therefore, various types of sensors are utilized for an accurate and real time detection and tracking; NIR, FIR, LIDAR, RADAR, Laser Scanner, fusion system, and etc.

Among VRUs on the road, pedestrians are the slowest while the others have the characteristics of moving speedily. Two-wheelers, different from ordinary pedestrians, are expressed with complicated shapes in combination of the rider's dressing style, hair style, whole body attitude, forms and patterns of loading, and various types of two-wheelers, which makes the shapes even further complicated according to the viewing angle, so that it requires a stronger algorithm against changes in their shapes.

The automobile vision-based systems have mainly concentrated on recognizing pedestrians and automobiles [1]. As stated in the foregoing, detection of two-wheelers on the road is similar with that of pedestrians on

road. Extracting characteristics for detecting pedestrians are classified with single characteristics, multiple characteristics, whole area characteristics, and district characteristics pursuant to the usage of extracted characteristics [3]. Especially, HOG characteristics or Improved HOG characteristics are widely utilized in the methods for recognizing pedestrians by using the automobile vision. Zhu et al. [4] applied the HOG characteristics based on variable block size to improve detection speed. Further, Watanabe et al. [5] utilized co-occurrence HOG characteristics, and Wang et al. [2] utilized HOG-LBP human detection to improve detection accuracy.

Two-wheelers are composed of persons, bicycles, or motorcycles together with various shapes of baggage. Therefore, the upper part of a waist is similar to a pedestrian but the lower part is diversified and complicated in shape, so that it becomes another field of challenge besides detection of pedestrians and automobiles for an intelligent automobile. The two-wheeler in the following means a combination of a person and a machine.

This paper consists of the following: Section II introduces basic extracting characteristics methods with HOG and a dramatic algorithm that can increase detection rates significantly. Section III states the framework and training of suggested detecting two-wheelers system. The evaluation and detailed analysis for the experimental results are summarized in section IV. Section V states the conclusion.

II. Feature Extraction

2.1 Histogram of Oriented Gradients (HOGs)

HOG [6, 10] converts the distribution directions of brightness for a local region into a histogram to express them in feature vectors, which is utilized to express the shape characteristics of an object. And it is influenced a little from an effect of illumination by converting the distribution of near pixels for a local region into a histogram, and has a strong feature for a geometric change of local regions. The method for calculating HOG feature is stated in the following.

Magnitude $m(x, y)$ and orientation $\theta(x, y)$ of a given image are computed using Eq. (1) and (2). Since unsigned orientations are desired for this implementation, the orientation range of degree which is less $0°$ is summed up with $180°$.

$$m(x, y) = 2\sqrt{f_x(x, y)^2 + f_y(x, y)} \qquad (1)$$

$$\theta(x, y) = \arctan\left(\frac{f_y(x, y)}{f_x(x, y)}\right) \qquad (2)$$

$$\begin{cases} f_x(x, y) = I(x+1, y) - I(x-1, y) \\ f_y(x, y) = I(x, y+1) - I(x, y-1) \end{cases} \qquad (3)$$
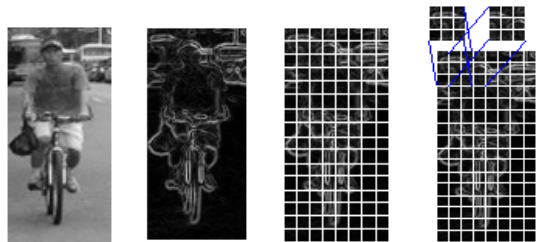

Fig. 1. The example of two wheelers HOG normalization. (a) Original image (b) Calculated magnitude vector (c) cells (d) blocks

In this paper, each cell, as shown Figure 1 (c), is represented by 8x8 pixel size and has 9 bins covering the orientation for [0°, 180°] interval, as shown in figure 1 (c). Directional histograms for brightness prepared in each of the cells were normalized as a block of 3x3 cell. Normalization processes are summarized in figure 1 (d), where a movement of block is based on that it is moved to right side and to lower side by one cell each. A contrast-normalization is used on the local responses to get better invariance regarding illumination, shading, etc. To normalize the cell's orientation histograms, it should be grouped into blocks (3x3 cells). This is done by accumulating a measure of local histogram value over the blocks and the result is then used to normalize the cells in the block. Although there are four different methods for block normalization suggested by Dalal and Triggs [21], L2-norm normalization Π is implemented using equation (4)

$$\Pi = \frac{f}{\sqrt{\|f\|_2^2 + \varepsilon^2}} \qquad (4)$$

For example, the dimension number for the width and height of an input image are 128x64 pixels, the dimension number of the histogram is 9, the size of cell is 8, and the size of block 3, is then calculated, and HOG feature vectors with 6804 dimension is obtained.

2.2 Spatial Relocation of Features

In general, cell characteristic values obtained from each

of the cell are expressed as multi dimension cell space, which are equal to Eq. (5):
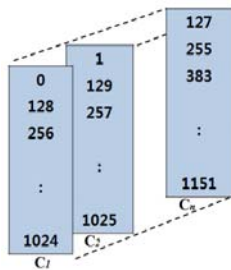
$$\Gamma = \begin{bmatrix} C_1 & C_2 & \dots & C_n \end{bmatrix} \tag{5}$$

(a)

(b)

(c)

(d)

Fig. 2. Compared with proposed spatial relocation method and general feature saving method. (a) General feature saving method for the cells, (b) General method for block, (c) Cell spatial relocation method to save features, (d) Block spatial relocation method.

Numbers of feature vectors in each of the cells are

dependent on numbers of directions. The method suggested in this study to save feature vectors is to save feature vectors using spatial relocation method for each cell, as shown Figure 2.

Feature vectors are saved in spatial relocation method within each cell. The saved feature values are collected by 3x3 cells, processed for blocking, and then normalized. As explained in 2.2, for blocking 3x3 cells, feature vectors of each cell are saved in a block position same as the cells using spatial relocation method. Each block is normalized by using Eq. (4) and final features are saved by the spatial relocation method as shown in Figure 2 (d).

III. Classifier

Adaboost is a simple learning algorithm that selects a small set of weak classifiers from a large number of potential features according to the weighted majority of classifiers. The training procedure of Adaboost is a greedy algorithm, which constructs an additive combination of weak classifier. Our boosting algorithm is basically the same as P. Viola's algorithm [7]. The algorithm takes as input a training set where each belong to some domain and each label is in some label set. The algorithm takes as input a training set $(x_1, y_1), \cdots, (x_n, y_n)$ where each x_i belong to some domain X and each label y_i is in some label set Y.

The final hypothesis H is a weighted majority vote of the T weak hypotheses where α_t is the weight assigned to h_t as below sequence.

Given training set: $(x_1, y_1), \cdots, (x_n, y_n)$

where $x_i \in X$, $y_i \in Y = \{+1, -1\}$

1. Initialize weights $w_{1,i} = \dfrac{1}{2m}, \dfrac{1}{2l}$ for $y_i = +1, -1$

m : the number of positive image (two-wheeler, +1)
n : the number of negative image (non two-wheeler, -1)

2. For $t=1 \cdots T$:
   (a) Normalize the weights,

$$w_{t,i} = \frac{w_{t,i}}{\sum_{j=1}^{n} w_{t,j}}$$

so that $w_{t,i}$ is a probability distribution of ith training image for tth weak classification

(b) For each feature, j, train a classifier h_j which is restricted to using a single feature. The error is evaluated with respect to w_i

$$\varepsilon_j = \sum_i w_j \left| h_j(x_i) - y_i \right|.$$

(c) Choose the classifier, h_t, with the lowest error ε_t

(d) Update the weights:

$$w_{t+1,i} = w_{t,i} \beta_t^{1-\varepsilon_i}$$

where $\varepsilon_i = -1$ if example x_i is classified correctly, $\varepsilon_i = +1$ otherwise, and $\beta_t^i = \dfrac{\varepsilon_t}{1-\varepsilon_t}$

3. Output the final hypothesis:

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

where $\alpha_t = \log(1/\beta_t)$

IV. Experimental Results

The proposed evolutionary algorithm was applied to improve the detection rate using spatial relocation and contrast method. An experiment was carried out with an ordinary user computer environment consisting of a Pentium 3.1 GHz and Visual C++ 6.0 Program. Two-wheelers data used in the experiment includes photos taken directly on the road and others obtained from the Internet. An image of two-wheelers can be expressed with various angles in an automobile. For our purposes, it is hypothesized in the experiment for the following 2 cases: a two-wheeler is running in front of an automobile (rear appearance) and a two-wheeler is coming toward the automobile (front appearance). And the experiment was done for the attitude of 90 degrees and the attitude of within 60 degrees in basis of horizontal line. 2,353 pictures of normalized two-wheelers were used with a size of 128x64 from the taken photos with a size of 640x480. They were utilized by dividing training image and experimental image. Pictures of non two-wheelers were obtained by utilizing randomly extracted pictures from the photos of streets in ordinary cities. The number of non two-wheelers used in the training was equal to the number of two-wheelers, and 3,000 pictures of non two-wheelers were used in the experiment.

The experiment was carried out using HOG Method (the most widely utilized) and spatial relocation (CSR_HOG), which is suggested in the study. A range of thresholds of -20 to 20 was utilized in classification, and confusion matrix, true positive rate (TPR) and false

positive rate (FPR) were used for analyzing experimental results per angles for the methods, and ROC curves are shown in figure. 3, by applying Eq. (6) below:

$$TPR = \frac{TP}{TP + FN} \quad , \quad FPR = \frac{FP}{FP + TN}$$

(6)

where "TP" is True Positive", "FP" is False Positive", "TN" is True Negative and "FN" is False Negative.

In figure 3, "Moto" means motorcycle, "Bike" means bicycle, and "MB" is a mixture of motorcycles and bicycles, respectively. Also, the numerals behind each of the abbreviations "60" signifies within 60 degrees, "90" within 90 degrees, and "90-60" a mixture of 90 and 60 degrees, respectively, as well.
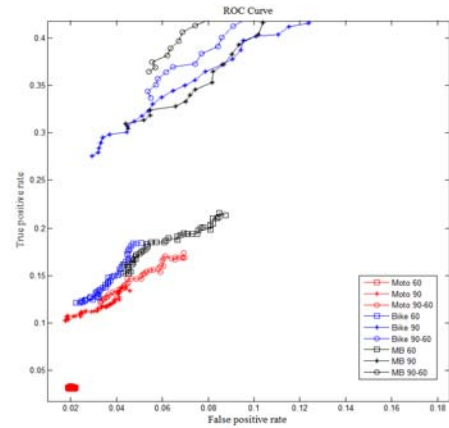
(a)
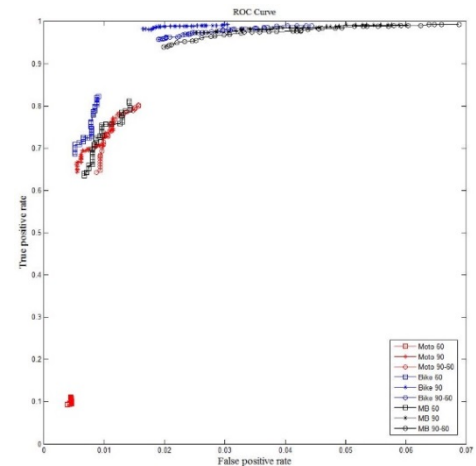
(b)

Fig. 3. Experiment Results (a) Results of an ordinary HOG Method (series data saving method), (b) Results of the contrast HOG feature experiment by applying the suggested algorithm.

In Figure 3 (a), it has shown that the experiment according to "MB" has the best results among these experiments, but

the recognition rate is significantly low. However, in Figure 3 (b), shows that the results of MB experiments according to the suggested method have recognition rate of nearly 1. As shown in Figure 3 (b), because the area by the curve of the suggested method is larger than that of the curve of the ordinary method, it has known that this system has a better performance. When CSR_HOG algorithm was applied for other characteristic vector (serial saving ordinary HOG) method, a higher recognition rate could be obtained, and the results are listed in Table 1.

The highest accuracies for each of the methods were calculated with Eq. (7) [8] and the results are listed in Table 1.

$$Accuracy = \frac{TP}{TP + FP} \qquad (7)$$

Table 1. Accuracies for each of the methods (%)

Angle \ Method		HOG	CSR_HOG
60	M	61.1	81.66
	B	71.2	97.37
	MB	76.7	98.40
90	M	74.9	97.24
	B	78.3	98.32
	MB	76.1	97.43
90-60	M	77.8	96.70
	B	75.5	97.53
	MB	73.1	97.00

As shown in Table 1, Bike (bicycle) has higher accuracies than Moto (motorcycle), signifying that bicycle has a trend of better classifying characteristics than motorcycle, in our opinion because a motorcycle becomes complicated by loading baggage in the rear or by loading high baggage. In the experiment of mixture of the two kinds of two-wheelers, the results of accuracies have partially confused patterns for angles and kinds, but for the whole mixture (90-60) the suggested method has a higher accuracy than the existing algorithm.

In table 1, the CSR_HOG presents not only higher accuracy rate than the traditional HOG for training rate and all degree but also less calculation time than the HOG due to using half image. Contrast is the difference in luminance or colour that makes an object (or its representation in an image or display) distinguishable[9]. The illuminance or brightness can be defined as the amount of energy output by a source of light relative to the source. In some cases, we can easily say that the image is bright, and in some cases, it

is not easy to perceive. Adapted equation is as follows:

$$F = \begin{cases} m \times val, & if \ m > th \\ m \ / \ val, & else \end{cases} \qquad (8)$$

where m is gradient value, val is contrast constant and th is threshold value. The *val* is 2.0, *th* is 50.

V. CONCLUSION

A new system for detecting two-wheelers was suggested by applying the saving method of new characteristic vectors. Characteristic vectors were obtained, saved by collecting local cells once, of which were normalized and collected once again, and then finally the characteristic vectors were extracted. Proposed new feature saving method, called spatial relocated contrast HOG feature, was used to extract cell feature and process block normalization. The Adaboost Method was applied for a speedy classification of two-wheelers; the results compared with existing methods have been shown that it has highly improved system performance.

REFERENCES

[1] H. Jung, Y. Ehara, J. K. Tan, H. Kim, and S. Ishikawa, "Applying MSC-HOG Feature to the Detection of Human on a Bicycle," International Conference on Control, pp. 514-517, Oct. 2012.

[2] H. Cho, P. E. Rybski, and W. Zhang, "Vision-based Bicyclist Detection and Tracking for Intelligent Vehicles," Intelligent Vehicle Symposium, pp. 454-461, June 2010.

[3] L. Yu, F. Zhao, and Z. An, "Locally Assembled Binary Feature with Feed-forward Cascade for Pedestrian Detection in Intelligent Vehicle," Int. Conf. on Cognitive Informatics, pp. 458-463, Jly 2010.

[4] S. Paisitkriangkrai, C. Shen, and J. Zhang, "Fast Pedestrian detection using a cas cade of boosted covariance features," IEEE Transaction on Circuit and System for Video Technology, Vol. 18, pp. 1140-1151, Aug. 2008.

[5] X. B. Cao, Y. W. Xu, D. Chen, and H. Qiao, "Associated evolution of a support vector machine-based classifier for pedestrian detection," Information Sciences, Vol. 179, pp. 1070-1077, Mar. 2009.

[6] N. Dalal and B. Triggs, "Histogram of Oriented Gradients for Human detection," CVPR 2005, pp. 53-60, Jun. 2004.

[7] P. Viola, M. Jones, and M. Snow, "Detection pedestrians using patterns of motion and appearance," ICCV 2003, pp. 153-161, Oct. 2003.

[8] http://en.wikipedia.org/wiki/Receiver_operating_characteristic, 2017, Mrch.

[9] https://en.wikipedia.org/wiki/Contrast_(vision), 2017, March.

[10] Y. Lee, "Curvature and Histogram of Oriented Gradients based 3D Face Recognition using Linear Discriminant Analysis," Journal of Multimedia Information, Vol.2, No.1, pp. 171-178, March 2015.

Authors

Yeunghak Lee received his Ph.D degree from Yeungnam University, Korea, in 2003. He had a year experience at University of Cardiff as postdoc research fellow and worked at Kyungwoon University as professor. Currently, he is a research professor of department of computer engineering at Andong National University. And he received many best paper awards in several domestic and international conference. His research interests include pattern recognition, embedded system and computer vision

Jaechang Shim received the B. S. degree in 1987 and Ph.D. degree from Kyungpook National University, Korea, in 1997. From 1997 to 1999, he worked as a researcher in IBM T. J. Watson Research Center. From 2005 to 2006, he was a visiting Fellow Professor at Princeton University. Since 1994, he has been a Professor of the Andong National University, Korea. He received the best Industrial-Related Paper Award form International Association for pattern Recognition in 1998. His research interests include computer vision, image processing, and pattern recognition