International Journal of Innovative Computing, Information and Control
Volume **10**, Number **2**, April **2014**

 ISSN 1349-4198
pp. 521–533

DETECTION OF A BICYCLE IN VIDEO IMAGES USING MSC-HOG FEATURE

Heewook Jung, Yusuke Ehara, Joo Kooi Tan
Hyeongseop Kim and Seiji Ishikawa

Department of Mechanical and Control Engineering
Kyushu Institute of Technology
1-1 Sensuicho, Tobata, Kitakyushu, Fukuoka 804-0015, Japan
{jung; ehara; etheltan; ishikawa}@ss10.cntl.kyutech.ac.jp; kimhs@cntl.kyutech.ac.jp

Received February 2013; revised August 2013

Abstract. *Traffic accidents are decreasing due to the influence of technology advancement; however accidents still occur due to carelessness of drivers. Therefore, many researchers have been studying how to realize an advanced safety system. The Histograms of Oriented Gradients (HOG) feature is well known as a useful method of detecting a standing human in various kinds of backgrounds. Unlike a person, a bicycle can appear differently from various angles. In this paper, we propose a method of detecting a bicycle on the road using improved HOG feature named MSC-HOG feature and the Real-AdaBoost algorithm. Experimental results and evaluation show satisfactory performance of the proposed method.*
Keywords: MSC-HOG feature, Bicycle detection, Object recognition, Advanced safety system

1. **Introduction.** Many researchers have been currently studying object detection and recognition. Object detection and recognition have been applied to various fields to date. In the security field, the object detection and recognition technology is used in surveillance cameras [4,22,23]. In the entertainment field, the technologies are applied to games using face recognition and camera controllers [21]. Robots helping us in many ways have made tremendous strides for these few decades. However, environmental recognition and recognition of objects that can be employed in a robot's eye-brain system have not achieved great advancement. Human eyes are very important for recognizing objects. Similarly, robots can be applied to many more fields, if they have eyes as good as human eyes. Therefore, generic object recognition is one of the core technologies for future intelligent robots.

On the other hand, object detection and recognition technologies are recently applied to advanced safety vehicle (ASV) systems [16], so we continue to monitor developments in ASV systems among applied fields. The system can detect pedestrians and cars, etc. with an in-vehicle camera. To name a few, Eyesight of SUBARU and Pedestrian Detection with Full Auto Break of VOLVO are deployed on a commercial scale. Eyesight of SUBARU uses stereo cameras and the speed limit to break up a car is 30 km/h. Other function of Eyesight is recognizing distance from the car ahead. Pedestrian Detection with Full Auto Break of VOLVO uses a wide angle camera and a radar, and the speed limit to break up a car is 35 km/h. Another function of VOLVO system is recognizing distance with pedestrians.

It is practically difficult to recognize objects in real traffic scenes because of cluttered backgrounds, image degradation, shape distortion, illumination change, various occlusions, etc. The systems of SUBARU and VOLVO have some problems that often results in miss detection and false detection when pedestrians should be detected. This is because of the factors which disrupt objects recognition as mentioned above. It is necessary to detect objects more quickly and precisely in order to use these systems in a practical scene. According to the survey of National Police Agency of Japan, there were 144, 018 bicycle accidents among the total of 664, 907 traffic accidents in 2011 in Japan. Although the number of traffic accidents is slowly decreasing in Japan, the rate of bicycle accidents accounts for more than 20 percent of all the traffic accidents in recent five straight years. Furthermore, the accident rate of bicycles vs. cars accounts for more than 80 percent. Here, we feel the necessity to detect bicycles automatically by in-vehicle cameras to reduce bicycle accidents. Therefore, in this paper, we propose an effective method of detecting bicycles on the road. We also show experimental results and evaluation to prove effectiveness of the proposed method.

1.1. **Related work.** Unlike the detection of pedestrians, bicycles detection is more difficult because its appearance changes dramatically according to viewpoints. Conventional researches on vision-based bicycles detection employ a part-based model and pedaling movement [9,10], but its availability in real transportation environment is not examined. Fast and precise detection is required to detect a bicycle in real outdoor scenes. For this purpose, we need to have better object features and classifiers. Many researchers have been studying on extracting better object features: SIFT (Scale Invariant Feature Transform) and HOG (Histograms of Oriented Gradients) are well-known features [1,14]. SITF is robust to variable scales, illumination and rotation, and is mainly applied to image matching and face recognition. HOG feature is, on the other hand, robust to detection and recognition of objects that have appearance change like a human and is often used to detect a standing human in various backgrounds. However, these features have some weaknesses: it takes a long time for SIFT to calculate, and HOG feature has disadvantages against variable scales and variable rotation. Therefore, some researchers improved SIFT features and HOG features to overcome these disadvantages [12,18,19].

To quickly and precisely detect an object, another important factor is better classifiers performance. Now, SVM (Support Vector Machine) and AdaBoost have been proposed as excellent classifiers for classification of two class objects [1,8,12,13]. SVM is a method that separates two categories by maximizing the margin that is the distance between a hyperplane and support vectors [8], whereas AdaBoost is a method of uniting weak classifiers of simple hypotheses and generating a strong classifier [7]. SVM has been used successfully in a human detection and AdaBoost has been used effectively in face detection and recognition [1,15]. To obtain better performance, SVM and AdaBoost algorithm have been employed along with several object features [15,17], and existing researches use a cascade structure and a coarse-to-fine approach to improve detection time [12,24].

1.2. **Approach of the proposed method.** Many papers deal with pedestrians detection using feature descriptors and classifiers [1,3,4,25]. Bicycle detection is more difficult than pedestrians detection because of the variety of its appearance according to viewpoints. Although there are some papers on automatic bicycles or bicyclists detection, they are still insufficient for putting them into practical use. The method proposed in [9], e.g., cannot detect a bicycle without pedaling movement. The method stated in [10] is difficult to track multiple bicyclists and no results based on real traffic scenes are given. In this paper, we focus our attention on bicycles detection in real traffic scenes. It is important to detect bicycles quickly and precisely using an in-vehicle camera from a drivers

seat in order to prevent the accidents against bicycles. To realize the idea, we propose the use of improved HOG feature named as MSC-HOG feature and the RealAdaBoost algorithm. The MSC-HOG feature gains more useful information than the original HOG feature proposed by Dalal and Triggs [1] and Watanabe et al. [25] since it employs variable cell sizes, whereas [1,25] adopt a fixed cell size.

The features used in the proposed method are combination of the MSC-HOG feature and hue in the HSV color space. Since the time for processing data increases according to the increase of feature dimension, we use an integral histogram to reduce calculation cost. We also employ RealAdaBoost and a cascade structure for fast classification of a selected region. The cascade structure can improve classification speed to a large extent by discarding negative cases quickly.

This paper is organized as follows. Section 2 introduces a proposed learning algorithm along with the definition of the MSC-HOG feature. Section 3 explains the employed detection algorithm. In Section 4, experimental results are shown. We discuss about experimental results in Section 5. Finally, the conclusion is given in Section 6.

2. Method.

2.1. **Learning part.** We describe a new method of extracting features and constructing classifiers in this subsection. First, we introduce the HOG feature proposed by Dalal and Triggs [1]. It is a method of extracting outlines of an object in an image. Because gradients of adjacent pixels receive histogramization and normalization at each area called a cell, it is robust to illumination change and to geometric change. The HOG feature is extracted in the following way.

Step 1: Compute the magnitude $m(x, y)$ and orientation $\theta(x, y)$ of a given image using the following formulae:

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

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

Here, given an image $f(x, y)$,

$$\begin{aligned} f_x(x, y) &= f(x + 1, y) - f(x - 1, y) \\ f_y(x, y) &= f(x, y + 1) - f(x, y - 1) \end{aligned} \tag{3}$$

Step 2: Derive the orientation histogram from the orientations and magnitudes at each cell to decrease the amount of information that the oriented gradients image possesses.

Step 3: Perform normalization of the histogram using an overlapping block. Here, if each cell size is 5×5 pixels and the orientation histogram has 9 bins with each cell and each block size is 3×3 cells, a 30×60 sub-image contains a 3240 feature set.

2.2. **MSC-HOG feature.** The MSC-HOG (Multiple-size Cell HOG) feature is based on the original HOG feature, yet it can obtain more comprehensive and useful information than the original HOG feature. The main difference between the two methods is that the cell size is variable in the MSC-HOG feature, whereas it is fixed in the original HOG feature. So the edge length is fixed in the latter as shown in Figure 1(a). On the other hand, one is able to gain long as well as short edges in the MSC-HOG feature as shown in Figure 1(b) because the cell size is not fixed.

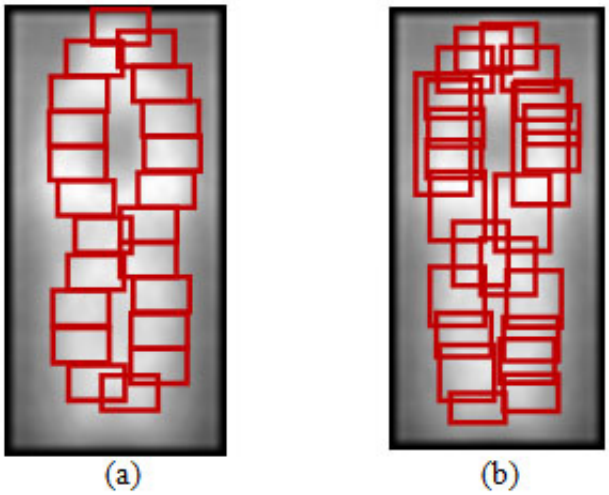

FIGURE 1. Difference in the distributed cells: (a) original HOG feature, (b) MSC-HOG feature

Another difference of the original HOG and the MSC-HOG feature is how to normalize the obtained feature. Original HOG normalizes the feature with each block but MSC-HOG performs the normalization after merging all the cells in order to save computational time.

The MSC-HOG feature is calculated in the following way.

1) Integral histogram computation:

First of all, compute the magnitude $m(x, y)$ and orientation $\theta(x, y)$ of a given image at each cell using Equations (1)-(3). Then calculate a gradient integral image with respect to each orientations bin, say b ($b = 1, 2, \ldots, B$) using Equation (4).

$$ii_b(x, y) = \sum_{x' \leq x} \sum_{y' \leq y} i_b(x', y') \tag{4}$$

Here $ii_b(x, y)$ is called an integral histogram with bin b and $i_b(x, y)$ is the gradient magnitude at bin b.

Finally, the accumulative gradient magnitude of integral image with each bin is computed by Equations (5) and (6).

$$l_b(x, y) = l_b(x, y - 1) + i_b(x, y) \tag{5}$$

$$ii_b(x, y) = ii_b(x - 1, y) + l_b(x, y) \tag{6}$$

Here $l_b(x, y)$ is a total of gradient magnitude with each bin in the same row.

2) Cell size and its location:

The cell size is computed by Equation (7).

$$S_{i,x} = MIN_x + STEP \times i \quad S_{i,y} = MIN_y + STEP \times i \quad (i = 0, 1, 2, \ldots, I) \tag{7}$$

Here $S_{i,x}$ is the width of the ith cell and $S_{i,y}$ is the height of the ith cell. $STEP$ is a variable width of a cell size and MIN_x, MIN_y are the minimum sizes of a cell to the x and the y direction, respectively. Various sizes of cells are placed randomly on an image and they are allowed mutual overlap.

3) Feature calculation:

Let us take the kth cell whose upper left coordinate is (x, y) and having the width S_{kx} and the height k_{ky}. The frequency at the bth bin of the histogram of the cell, denoted by

TABLE 1. RealAdaBoost algorithm

1. Suppose $S = (x_1, y_1), (x_1, y_1), \ldots, (x_1, y_1)$ is a set of samples, where $\mathbf{x} \in X$ $(i = 1, 2, \ldots, N)$ are feature vectors and $y_i \in -1, +1$ are their labels. 2. Initialize weighted value D_t of samples. $D_1(i) = \frac{1}{N}$ $(i = 1, 2, \ldots, N)$ 3. For $t = 1$ to T do for $m = 1$ to M do Compute probability distribution W_t of a weak classifier $h_i(t)$ $W_+^j = \sum_{i,j \in J \bigwedge y_i = +1}^{n} D_t(i)$ $\quad$ $W_-^j = \sum_{i,j \in J \bigwedge y_i = -1}^{n} D_t(i)$ Here i is the learning sample's number, and j is the bin's number of the histogram. Compute estimation Z_m $Z_m = 2 \sum_j \sqrt{W_+^j W_-^j}$ Select the smallest Z_m of the weak classifier $\quad$ $h_t = \arg\min Z_{t,m}$ Renew sample weight $D_t(i)$ $h_t(\mathbf{x}_i) = \frac{1}{2} \ln \frac{W_+^j + \varepsilon}{W_-^j + \varepsilon}$ $\quad$ $D_{t+1}(i) = D_t(i) \exp[-y_i h_i(\mathbf{x}_i)]$ 4. The final strong classifier is given by $H(x) = sign(\sum_{t=1}^{T} h_t(x))$

$v'_{k,b}$, is calculated by Equation (8).

$$v'_{k,b} = ii_b(x, y) + ii_b(x + S_{k,x}, y + S_{k,y}) - ii_b(x + S_{k,x}, y) - ii_b(x, y + S_{k,y}) \tag{8}$$

By denoting the number of the cells on an image by K, the feature vector representing the image is given by

$$\mathbf{v}' = (v'_{1,1}, v'_{1,2}, \ldots, v'_{1,B}, \ldots, v'_{k,1}, v'_{k,2}, \ldots, v'_{k,B}, \ldots, v'_{K,1}, v'_{K,2}, \ldots, v'_{K,B}) \tag{9}$$

Here we assume that the number of bins B is common to all the K cells.

4) Normalization:

Normalization of the feature vector $\mathbf{v}'$ is executed by Equation (10).

$$\mathbf{v} = \frac{\mathbf{v}'}{\|\mathbf{v}'\| + \varepsilon} \tag{10}$$

Here ε is a positive real number to prevent that the denominator becomes zero.

Finally, the feature vector representing the image is given by

$$\mathbf{v} = (v_{1,1}, v_{1,2}, \ldots, v_{1,B}, \ldots, v_{k,1}, v_{k,2}, \ldots, v_{k,B}, \ldots, v_{K,1}, v_{K,2}, \ldots, v_{K,B}) \tag{11}$$

2.3. **RealAdaBoost.** The RealAdaBoost algorithm [2] is a method of uniting weak classifiers of a simple hypothesis and generating a strong classifier. The MSC-HOG model of bicycles are defined and trained via a RealAdaBoost to detect bicycles under various circumstances. The algorithm of RealAdaBoost is given in Table 1.

3. **Detection Algorithm.** A detection algorithm is employed for a bicycle using the MSC-HOG feature. In the detection algorithm, we propose a method using adjustable number of weak classifiers to classify fed images quickly. We then merge detected windows containing a same bicycle into a single window using the mean shift clustering and the nearest neighbor algorithm, because, when detection of a bicycle is performed, many windows containing a single common bicycle are detected in an image.

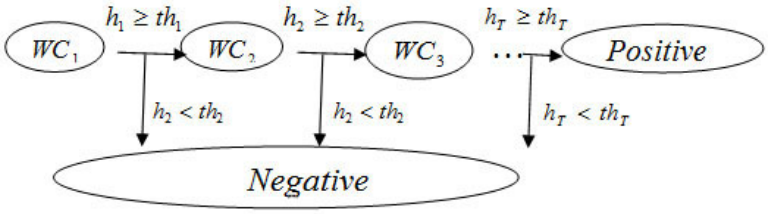

FIGURE 2. Adjustable number of weak classifiers

3.1. **Adjustable number of weak classifiers.** In general, classifying objects using the RealAdaBoost algorithm repeats the above mentioned procedure. However, our method ceases to repeat if the output of the weak classifiers is smaller than the threshold at each round. Therefore, our method is fast in classifying negative images as shown in Figure 2.

In Figure 2, WC_i is the ith weak classifier. h_i is the total of weak classifiers' output until round i, th_i is threshold at round i. The threshold is computed by Equation (12).

$$th_i = \min\{h_{i,1}, h_{i,2}, \ldots, h_{i,ID}\} \quad (ID \in N) \tag{12}$$

3.2. **Mean shift clustering.** The first step of the mean shift clustering is the assumption of a density using a kernel function. We use a kernel function represented by Equation (13).

$$k(\mathbf{x}) = \begin{cases} c(1 - \|\mathbf{x}\|) & \|\mathbf{x}\| < 1 \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

Here $\boldsymbol{x}$ is a vector representing the center of a detected window.

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

$$m(\mathbf{x}) = \frac{\sum_{i=1}^{n} \mathbf{x}_i k\left(\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\|^2\right)}{\sum_{i=1}^{n} k\left(\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\|^2\right)} - \mathbf{x} \tag{14}$$

Here the first term of the mean shift vector assumes density and the second term a shifted mean. And h is a radius of a detecting window. If the mean shift vector is 0, the density is the highest.

3.3. **Nearest neighbor.** Euclid distance d is computed from the point of the highest density to each sample point. Merging is then executed, if d is smaller than a predefined threshold. The result of the merging process is shown in Figure 3.

FIGURE 3. Merge process: (a) before merging, and (b) after merging

4. **Experimental Results.** Two kinds of experiments were performed, i.e., an experiment on the detection of a bicycle (Exp_1), and an experiment on the detection of its driving direction (Exp_2). In the performed experiment, we used not only the collected still bicycle images but also real outdoor scene images captured by a video camera installed in a moving car. Its installation is shown in Figure 4. The camera and PC specifications are given in Table 2.

FIGURE 4. Video cameras

TABLE 2. Specification of the PC and the camera used in the experiment

OS	Microsoft Window 7 Home Premium 64 bit
CPU	$Intel^{R}Core^{TM}$ i5 2.53GHz
Memory	4.0GB
Camera	SONY HDR-HC7
Software Tool	Microsoft Visual 2008 C++

4.1. **Detection of a bicycle.** In the first experiment using still images of a bicycle (Exp_1s), 2,700 positive samples containing 900 bicycle images directing front, left and right, respectively, and 6,000 negative samples (images not containing a bicycle) were used for training of weak classifiers and a strong classifier as shown in Figure 5. The training data was collected on the internet by us, since there is no substantial bicycles database in the world yet. On the other hand, we used 900 positive samples and 1,000 negative samples for test. The performance of the proposed method on the recognition rate and the processing time was compared with Dalal and Triggs [1] and Watanabe et al. [25]. Table 4 compares the dimension of the used feature vector among the three methods. The results of the performance are given in Table 5.

TABLE 3. Parameters of the proposed method used in the experiments

Minimum cell size of width and height	5
Maximum cell size of width and height	23
Size change of a cell	3
The number of bins	9
Total dimension of MSC-HOG	160290

TABLE 4. Comparison of the dimension of a feature vector in Exp_1s

Original HOG [1]	CoHOG [25]	Proposed method (MSC-HOG)
3240	34560	160290

TABLE 5. Recognition rate and processing time in Exp_1s

	Reacognition Rate [%]	Processing Time [ms]
Original HOG [1] + RealAdaBoost	96.5	412
CoHOG [25] + RealAdaBoost	96.8	238
Proposed method	97.1	136

FIGURE 5. Samples of the employed training data: (a) right direction, (b) left direction, (c) front direction, (d) background

TABLE 6. Performance of the detection in real outdoor scenes

	precision [%]		recall [%]		FPR [%]	
	Exp_1v1	Exp_1v2	Exp_1v1	Exp_1v2	Exp_1v1	Exp_1v1
Original HOG	96.0	84.3	91.0	79.0	4.0	15.7
CoHOG	96.0	86.5	91.0	83.0	4.0	13.5
MSC-HOG	96.0	89.0	91.0	85.0	4.0	11.0

The proposed method was also applied to real outdoor videos (Exp_1v). In this experiment and in the following Exp_2, we used totally 3,600 positive samples (1,200 sheets of samples directing front, left and right, respectively) and 6,000 negative samples for training of weak classifiers and a strong classifier. The results of two cases in a town (Exp_1v1 and Exp_1v2) are shown in Figure 6 and Figure 7, respectively, in which a bicycle moves from right to left on a zebra-crossing. It is noted that the occlusion of the bicycle does not occur in Exp_1v1 (Figure 6), whereas it occurs in Exp_1v2 (Figure 7). The employed image frames are 100 and each image has pixels. A detected bicycle is indicated by a yellow rectangle in the image. The performance of the proposed method is shown in Table 6.

FIGURE 6. Result of Exp_1v1: Detection of a bicycle – The case where there is no occlusion with the bicycle

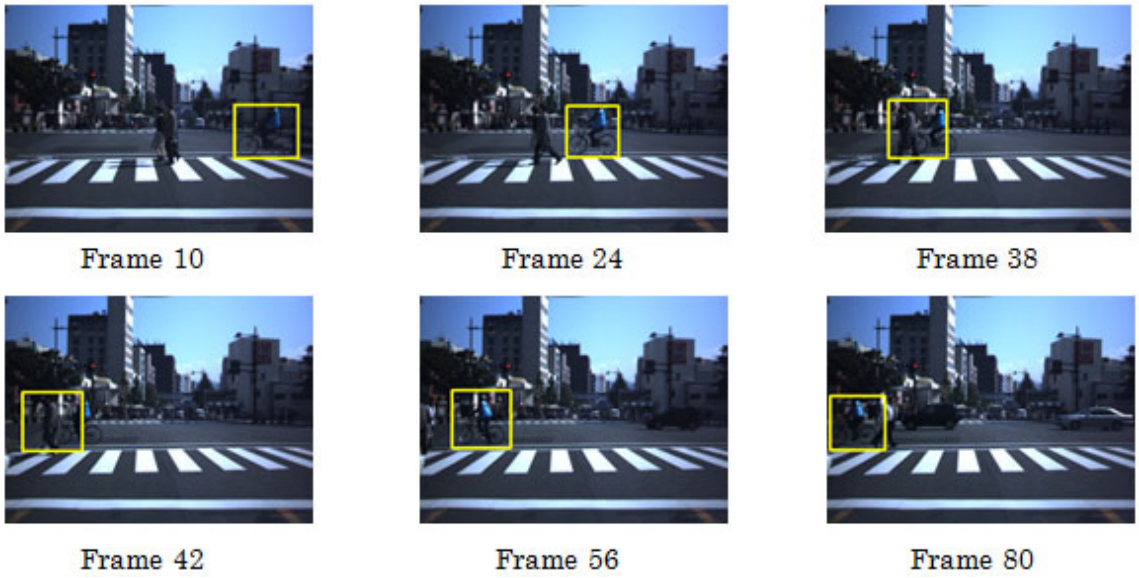

FIGURE 7. Result of Exp_1v2: Detection of a bicycle – The case where there is occlusion of the bicycle

The performance of the detection of a bicycle in real outdoor scenes is evaluated in Table 6. Precision, recall and FPR in Table 6 are computed by Equation (15)

$$\begin{aligned} precision &= TP \times 100/(TP + FP)[\%] \\ recall &= TP \times 100/(TP + TN)[\%] \\ FPR &= 100 - precision[\%] \end{aligned} \tag{15}$$

where TP True Positive, FP False Positive, TN True Negative, and FPR False Positive Rate. As shown in Figure 8, we judge success or failure using the following index δ.

$$\delta = \frac{R_{gt} \bigcap R_d}{R_{gt} \bigcup R_d} \tag{16}$$

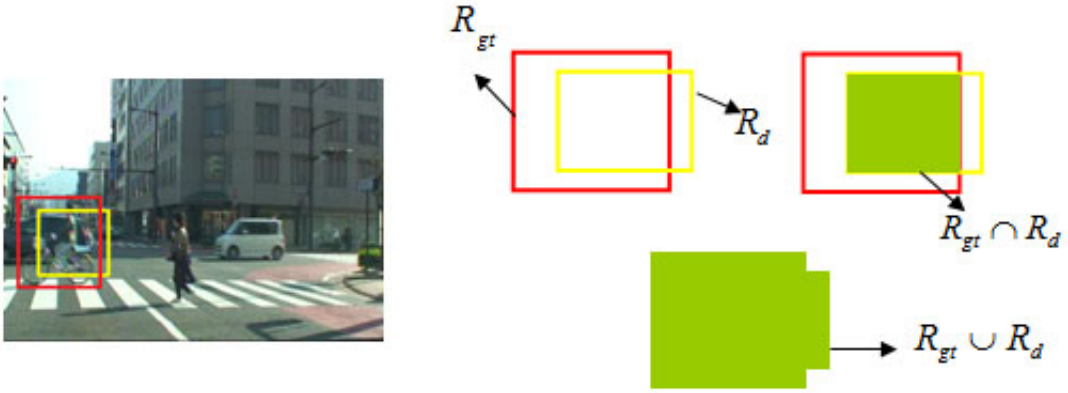

FIGURE 8. Criterion on success in the detection

Here, R_{gt} is a ground truth rectangle area containing a bicycle, and R_d is a detected bicycle area using the proposed method. If δ is 0.5 or more, we judge that the detection of a bicycle is successful.

4.2. **Detection of the driving direction.** In the detection experiment of a bicycle's driving direction in real outdoor scenes (Exp_2), we used totally 3,600 positive samples (1,200 sheets of training samples directing front, left and right, respectively) and 6,000 negative samples for training. Three detectors are used for detecting the driving directions. Experimental results on the detection of driving direction of a bicycle are given in Figure 9. The detected bicycles moving left, right and front are indicated by a blue, red and green rectangle, respectively. The evaluation on this experiment is given in Table 7. The precision, recall and FPR in Table 7 are computed by Equation (15).

TABLE 7. Performance of the detection of a bicycle driving direction

	Reacognition Rate [%]	recall [%]	FPR [%]
Left	92.0	83.0	8.0
Right	95.0	85.0	5.0
Front	95.0	81.0	5.0

5. **Discussion.** We performed two main experiments to test the performance of the proposed method. The first experiment is on the detection of a bicycle. First of all, we conducted an experiment on whether or not positive (bicycle) and negative (non-bicycle) images are classified correctly to examine the performance of the proposed method. In this experiment, the proposed method has achieved more outstanding results than the methods employing the original HOG feature [1] and CoHOG feature [25] with respect to classification accuracy and speed as shown in Table 5. This is because we can get more effective information by use of the MSC-HOG feature and color information for accuracy, and the employment of adjustable number of weak classifiers is highly effective in speed.

We also conducted an experiment on bicycle detection in real outdoor scenes. In outdoor scenes, bicycle detection is not very easy because occlusion often occurs with bicycles and there are many objects having similar shape with a bicycle. The occlusion occurred in the scene of Exp_1v2 (see Figure 7) and not in the scene of Exp_1v1 (see Figure 6). Therefore, the detection rate in Exp_1v2 was lower than that of Exp_1v1 as shown in Table 6. However, we still achieved a better result than the method employing the original HOG feature and CoHOG feature whose results are also given in Table 6. This may be because

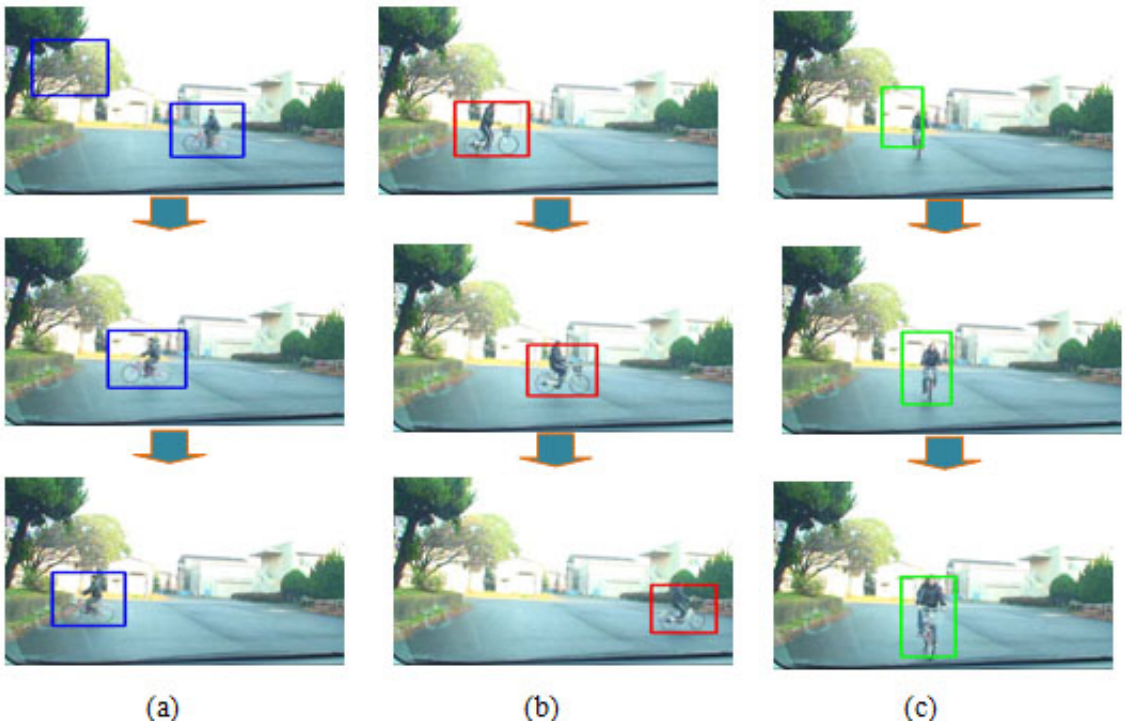

FIGURE 9. Result of Exp_2: Detection of a bicycle driving direction: (a) left, (b) right, (c) front

the MSC-HOG feature can enhance the feature of non-occluded part better than the other methods.

The second experiment is on the detection of driving direction of a bicycle. We used three classifiers to detect the front direction, the left direction and the right direction. In detecting the front direction, bicycle scale changes as a bicycle moves to the front. If the size of the bicycle is too small, its detection is difficult. In detecting the left and the right direction, they were sometimes detected in the reverse direction. This may come from the fact that a bicycle is potentially symmetric with left and right.

The proposed method is applicable to the detection of various objects and it can detect objects more precisely by the use of the MSC-HOG feature than using the original HOG feature [1] or CoHOG feature [25]. Many papers dealing with human detection show classification results using merely trimmed positive and negative images as experimental results, but when they are actually applied to real outdoor scenes, they cannot achieve good results. It is because the real outdoor scenes include occlusion of the object in question and many other objects of similar shapes. The method proposed in this paper shows useful results in real outdoor scenes because we can get more information using the MSC-HOG feature and color information.

In this paper, we proposed a method of detecting a bicycle ridden by a human in an outdoor environment. To the best of our knowledge, there have been no proposals on the method of object detection focusing on bicycles. It is not difficult to think of increased use of bicycles in the near future because of its convenience, yet a possible consequence may be an increase of traffic accidents involving bicycles. It is therefore vital to establish a method of automatic detection of a bicycle in a traffic environment to reduce such accidents.

6. **Conclusion.** In this paper, we proposed a method of detecting a bicycle using the MSC-HOG feature and the RealAdaBoost algorithm.

The proposed method has achieved, as shown in Table 5, outstanding results with respect to the precision of detection by the employment of MSC-HOG feature compared with the method using the original HOG feature or CoHOG feature. The speed of the detection was also improved by the employment of a cascade structure. This may make it possible to detect a human on a bicycle in real-time. In future, we plan to enhance further the performance of the detection of a bicycle and to develop a method which detects various objects that might become risky factors for traffics.

Acknowledgment. This study was supported by JSPS KAKENHI Grant Number 25350 477, which is greatly appreciated.

REFERENCES

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

[2] R. E. Schapire and Y. Singer, Improved boosting algorithms using confidence-rated predictions, *Machine Learning*, vol.37, no.3, pp.297-336, 1999.

[3] R. Benenson, M. Mathias, R. Timofte and L. V. Gool, Pedestrian detection at 100 frames per second, *Proc. of Conf. on Computer Vision and Pattern Recognition*, pp.2903-2910, 2012.

[4] M. Wang and X. Wang, Automatic adaptation of a generic pedestrian detector to a specific traffic scene, *Proc. of Conf. on Computer Vision and Pattern Recognition*, pp.3401-3408, 2011.

[5] M. Munaro and A. Cenedese, Scene specific people detection by simple human interaction, *IEEE International Conference on Computer Vision Workshops*, pp.1250-1255, 2011.

[6] D. Comaniciu and P. Meer, Mean shift: A robust approach toward feature space analysis, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.24, no.5, pp.603-619, 2002.

[7] A. Bertoni, P. Campadelli and M. Parodi, A boosting algorithm for regression, *Proc. of the 7th International Conference on Artificial Neural Networks*, pp.343-348, 1997.

[8] B. E. Boser, I. M. Guyon and V. N. Vapnik, A training algorithm for optimal margin classifiers, *Proc. of the 1st Annual Workshop on Computational Learning Theory*, pp.144-152, 1992.

[9] K. Takahashi, Y. Kuriya and T. Morie, Bicycle detection using pedaling movement by spatiotemporal Gabor filtering, *International Journal of Innovative Computing, Information and Control*, vol.8, no.6, pp.4059-4070, 2012.

[10] H. Cho, P. E. Rybski and W. Zhang, Vision-based bicyclist detection and tracking for intelligent vehicles, *Intelligent Vehicles Symposium*, pp.454-461, 2010.

[11] F. Porikli, Integral histogram: A fast way to extract histograms in Cartesian spaces, *Proc. of Conf. on Computer Vision and Pattern Recognition*, pp.829-836, 2005.

[12] Q. Zhu, S. Avidan, M. C. Yeh and K. T. Cheng, Fast human detection using a cascade of histograms of oriented gradients, *Proc. of Conf. on Computer Vision and Pattern Recognition*, pp.1491-1498, 2006.

[13] C. R. Wang, J. Wu and J. J. Lien, Pedestrian detection system using cascaded boosting with invariance of oriented gradients, *International Journal of Pattern Recognition and Artificial Intelligence*, vol.23, no.4, pp.801-823, 2009.

[14] C. Wojek, S. Walk and B. Schiele, Multi-cue onboard pedestrian detection, *Proc. of Conf. on Computer Vision and Pattern Recognition*, pp.794-801, 2009.

[15] P. Viola and M. J. Jones, Robust real-time face detection, *International Journal of Computer Vision*, vol.57, no.2, pp.137-154, 2004.

[16] A. Geiger, P. Lenz and R. Urtasun, Are we ready for autonomous driving? The KITTI vision benchmark suite, *Proc. of Conf. on Computer Vision and Pattern Recognition*, pp.3354-3361, 2012.

[17] S. Hussain and B. Triggs, Feature sets and dimensionality reduction for visual object detection, *British Machine Vision Conference*, pp.112.1-112.10, 2010.

[18] H. Bay, T. Tuytelaars and L. V. Gool, SURF: Speeded up robust features, *Proc. of European Conf. on Computer Vision*, pp.404-417, 2006.

[19] X. Wang, T. X. Han and S. Yan, An HOG-LBP human detector with partial occlusion handling, *Proc. of Int. Conf. on Computer Vision*, pp.32-39, 2009.

[20] A. Vedaldi, V. Gulshan, M. Varma and A. Zisserman, Multiple kernels for object detection, *Proc. of Int. Conf. on Computer Vision*, pp.606-613, 2009.

[21] S. Scher, R. Crabb and J. Davis, Making real games virtual: Tracking board game pieces, *Proc. of Int. conf. on Pattern Recognition*, pp.1-4, 2008.

[22] V. Atienze-Vanacloig, J. Rosell-Ortega, G. Andreu-Garcia and J. M. Valiente-Gonzalez, People and luggage recognition in airport surveillance under real-time constraints, *Proc. of Int. Conf. on Pattern Recognition*, pp.1-4, 2008.

[23] C. Chen and J. Odobez, We are not contortionists: Coupled adaptive learning for head and body orientation estimation in surveillance video, *Proc. of Conf. on Computer Vision and Pattern Recognition*, pp.1544-1551, 2012.

[24] M. Pedersoli, A. Vedaldi and J. Gonzalez, A coarse-to-fine approach for fast deformable object detection, *Proc. of Conf. on Computer Vision and Pattern Recognition*, pp.1353-1360, 2011.

[25] T. Watanabe, S. Ito and K. Yokoi, Co-occurrence of histograms of oriented gradients for human detection, *IPSJ Transactions on Computer Vision and Applications*, pp.39-47, 2010.