International Journal of Hybrid Information Technology
Vol. 9, No.7 (2016), pp. 217-226
http://dx.doi.org/10.14257/ijhit.2016.9.7.20

Adaboost based Two-wheeler Detection System using Normalized Cross Correlation

Yeunghak Lee

Department of Avionics Electronic Engineering, College of Aviation, 39196, Gyeongbuk, Kumi, Republic of Korea
annaturu@ikw.ac.kr

Abstract

There are several types of transport on the road; vehicles, bikes, motorcycles (two and four wheels), kids' bikes, construction vehicles, etc. In this paper, we present an enhanced algorithm for the detection rate to classify the two-wheelers with riding people using the normalized cross correlation (NCC) and its application of histogram of oriented gradients (HOG). HOG descriptors which are one of the well known methods for object detection are the feature descriptor using edge intensity in a local region. Using the NCC algorithm which is used to make a match from template images, local weighting values are calculated from template cell features. And the combined cell features are classified by using a boosting algorithm (Adaboost). The improvement detection rates are confirmed through experiments using bicycles and motorcycles data set for the front direction.

Keywords*: Adaboost, Normalized Cross Correlation, Two-wheeler, HOG, vulnerable road users*

1. Introduction

In addition to advanced technology, the means of transportation has developed in several types or shapes and including variety function. Now the development of transportation systems has been concentrated on not only improving performance, but also to protect the drivers and passengers in a vehicle in the occurrence of a traffic accident, as an environment-friendly vehicle. And the road environment is changing to protect not only the vehicles and passengers, but also pedestrians [1].

Although many researchers have used a variety of algorithms to detect the pedestrian and vehicles, the research of vulnerable road users (VRUs), such as a small device for riding (bicycle-adults and kids, motorcycle-two and four wheels, *etc*.) is still a hot subject for the study field in the intelligent transportation system. Recently, the research scope has gradually expanded to protect VRUs, such as those consisting of humans, two-wheelers, small moving devices and other small vehicles [1][2]. VRUs which is combined human body and complicated devices also move more slowly than vehicles on the road, except motorcycles. Even though there are many kinds of safety devices to protect VRUs, people on the two-wheelers are always exposed to a harmful road environment. So this is similar to pedestrian detection research.

As this paper previously alluded to the similarity of the shape of pedestrians and two-wheelers, the two-wheeler detection algorithm also closely resembles the pedestrian. The feature extraction method from vision-based images has been primarily studied in the Haar wavelet-based method, HOG, which has direction of gradient and local receptive field (LRF). The SVM, cascade method, neural network, and Adaboost algorithm are applied to a lot of categorizing methods [3][4][5].

Two wheelers consist of a human and machine; usually the human is the upper part and the machine is the lower part in the shape. HOG [6] based detector system has slow

ISSN: 1738-9968 IJHIT

performance because of its dense encoding scheme and multi-level scale images. Porikli [7] solved this problem using the concept of "Intelligent Histogram" to speed up the feature extraction process. Because of the above reasons, we tried to use a modified HOG algorithm to select the best features and Adaboost to improve detection rate.

In this study, we invented a new algorithm based on an adapted HOG value which is normalized correlation coefficient between a global area and each cell. The motivation of this paper is as follows. Firstly, the two wheeler detection system is still not a considerable time investment needed to find a good algorithm. Secondly, even though it is familiar with pedestrian detection, it is one of the most difficult studies due to the range of various image poses, as well as environmental conditions, cluttered backgrounds, and composite objects. Thirdly, the normalized form of cross correlation can be preferred for a global cell histogram featuring matching applications that do not have a simple cell histogram domain expression. This paper is organized as follows. In section 2 and 3, this paper explains the original feature set using general HOG and NCC algorithm and the associated evolutionary algorithm which has improved the detection rate. Section 4 describes Adaboost algorithm used in training and classification. The results of their evaluation and a detailed performance analysis are presented in section 5. Section 6 concludes this paper.

2. Histogram of Oriented Gradients

Histograms of Oriented Gradients (HOGs) are feature descriptors used in computer vision and image processing for the purpose of object detection. It is little influenced from the 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.

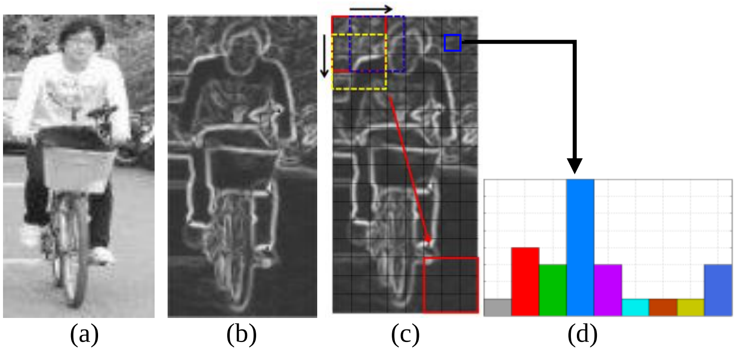


Figure 1. The Example of Two Wheelers HOG Normalization. (A) Original Image (B) Calculated Magnitude Vector (C) Cells and Blocks (A Block Is 3x3 Cells) Sliding (D) A Cell Histogram

Proposed main idea [6] is that local object appearance and shape can often be characterized rather well by the distribution of local intensity gradients or edge directions. This is achieved by dividing the image into cells and for each cell a one-dimension histogram of gradient directions over the pixels of the cell is calculated. Then each block in the image consists of a number of cells, as shown in Figure 1. After calculating *x*, *y* derivatives (*dx* and *dy*), the magnitude $|m(x,y)|$ and orientation $\theta(x,y)$ of the gradient for each pixel *I(x, y)* is computed from

$$\begin{aligned} dx &= I(x+1,y) - I(x-1,y) \\ dy &= I(x,y+1) - I(xy-1) \end{aligned} \tag{1}$$

$$|m(x,y)| = (dx^2 + dy^2)^{0.5} \tag{2}$$

$$\theta(x,y) = \tan^{-1}(dy/dx) \tag{3}$$

One thing to note is that, at orientation the computation radian to degree method is used, which returns values between -180° and 180°. Each histogram divides the gradient angle range into a predefined number of bins. In this paper, each cell, as shown Figure 1 (c) and (d), is represented by 8x8 pixel size and has 9 bins covering the orientation for [0°, 180°] interval. For each pixel's orientation, the corresponding orientation bin is found and the orientation's magnitude $|m(x, y)|$ is voted to this bin. To normalize the cell's orientation histograms, it should be grouped into blocks (3x3 cells). A cell feature is expressed as $F_i = [f_1, f_2, \cdots f_9]$. The characteristic quantities of *k'th* block may be expressed as:

$$B_k = [F_1, F_2, \cdots F_9] \tag{4}$$

This is done by accumulating a measure of the local histogram value over the blocks. Normalization processing is expressed in Figure 1 (c), where the movement of the block is based on the fact it is moved to the right side and to the lower side by a cell each way (back arrow). Although there are four different methods for block normalization suggested by Dalal and Triggs [6], *L2*-norm normalization Π is implemented using equation (5),

$$\Pi = \frac{B_k}{\left(\|B_k\|^2 + \varepsilon^2\right)^{0.5}} \tag{5}$$

where ε is the small constant instead of zero for the denominator. Here, the overlapping processing is done to ensure the important feature information of each cell is saved as a concatenated method.

3. Normalized Cross Correlation (NCC)

In signal processing, cross correlation is an important tool to estimate the degree of similarity in which two series as a signal of the lag of one relative to another are correlated, in image processing, pattern recognition, and other fields. To overcome several disadvantages for the cross correlation, NCC (Normalized Cross Correlation) algorithm [4][8] which is an improved method to measure the resemblance between two signals is used to decide the similarity between two images. Equation (6) is shown as a basic definition for the normalized cross correlation coefficient γ.

$$\gamma = \frac{1}{n} \sum_{x,y} \frac{[f(x,y) - \bar{f}_{u,v}][t(x-u, y-u) - \bar{t}]}{\{[f(x,y) - \bar{f}_{u,v}]^2 [t(x-u, y-u) - \bar{t}]^2\}^{0.5}} \tag{6}$$

where $f(x, y)$ is the intensity value of image f of the size $M_x \times M_y$ at the point (x, y), $x \in \{0, \ldots, M_x - 1\}$, $y \in \{0, \ldots, M_y - 1\}$. Template t of size is $N_x \times N_y$. And $\bar{t}$ is the mean of the feature and $\bar{f}_{u,v}$ is the mean of $f(x, y)$ in the region under the feature which is calculated by

$$\bar{f}_{u,v} = \frac{1}{N_x N_y} \sum_{x=u}^{u+N_x-1} \sum_{y=v}^{v+N_y-1} f(x, y) \tag{7}$$

A common way to calculate the pixel (u, v) of the pattern in the image *f* is to evaluate the normalized cross correlation value γ at each point *(u, v)* for *f* and the template *t*, which has been shifted by *u* steps in the *x* direction and by *v* steps in the *y* direction. The denominator in (6) is the variance of the zero mean image function $f(x,y)$ - $\bar{f}_{u,v}$ and shifted zero mean template function $t(x-u, y-v)$ - $\bar{t}_{u,v}$.

4. Adaboost Sequence

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 the weak classifier. Our boosting algorithm is basically the same as P. Viola's algorithm [9].

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 $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$ 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, $w_{t,i} = w_{t,i} \Big/ \sum_{j=1}^{n} w_{t,j}$, 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 $\beta_t^i = \varepsilon_t / (1 - \varepsilon_t)$

3) Output the final hypothesis:

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

The final hypothesis H is a weighted majority vote of the T weak hypotheses where α_t is the weight assigned to h_t . Using the two strong classifications, this paper suggests the use of 2nd stage cascade method. It improves the recognition rate due to the complementary role for two feature vectors of quite different types.

5. Experimental Results

Two-wheeler data used in the experiment includes photos taken on the street and others obtained from the internet randomly. For our purposes, it is hypothesized in the experiment for the following two cases: rear and front appearance. The experiment was done for the attitude of 90 degrees and within 60 degrees on the basis of a horizontal line. 2,353 pictures were used with a size of 128x64. The number of non two-wheelers used in the training was equal to the number of two-wheelers, and 3,000 pictures were used as non two-wheelers. The experiment was carried out with an ordinary user computer environment consisting of a Pentium 3.1 GHz and Visual C++ 6.0 program and Matlab. The negative (non-two wheelers) samples used in our experiments were extracted randomly from general street images, as shown in Figure 2. In Figure 2, ***P*** means positive and ***N*** means negative. Each training set ratio of positive and negative is 1:1. For more details, the bicyclist training images are 340 examples for the 60 degree view and 845 examples for 90 degrees. And motorcycle driver training images are 96 examples for the 60 degree view and 234 examples for 90 degrees. We also mixed the two degrees and types (bicyclist and motorcycle driver).

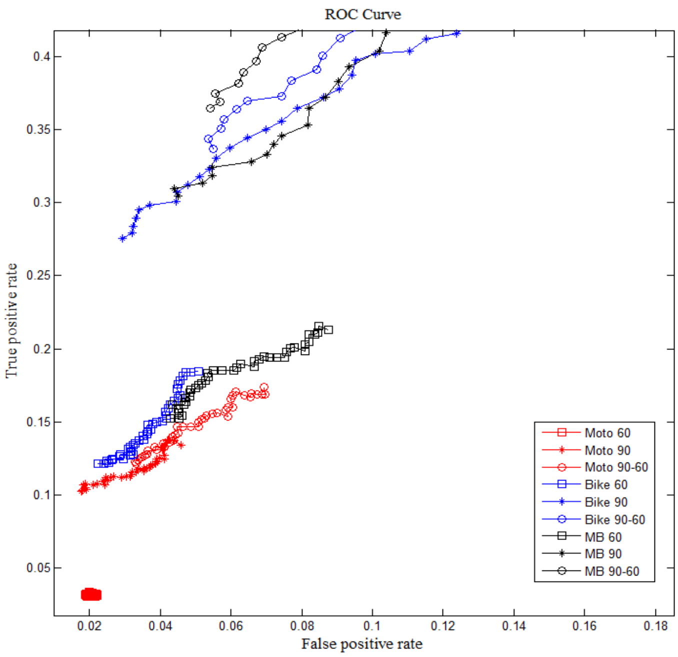
Figure 2. Examples of Training Samples for Positive and Negative

In the first experiment, the performance of the traditional HOG feature is shown in Figure 3 (a), according to the data type and degree.

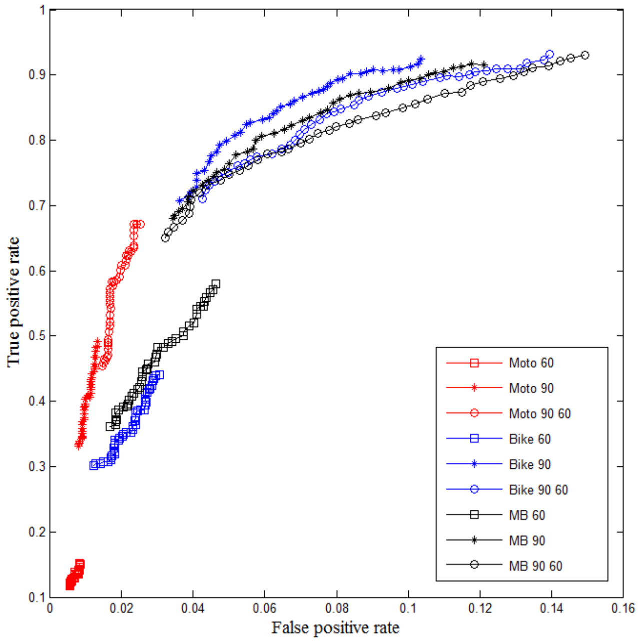
Secondly, the experiment was carried out by the proposed NCC1 and NCC2 method used weighting value which is suggested in the study. A range of thresholds of -20 to 20 was utilized in classification, and confusion matrix, true positive rate (TPR) and false positive rate (FPR) were used for analyzing experimental results per angles for the methods, and ROC curves are shown in Figure 3, by applying Eq. (8) below:

$$TPR = \frac{TP}{TP + FN}\ ,\ FPR = \frac{FP}{FP + TN} \tag{8}$$

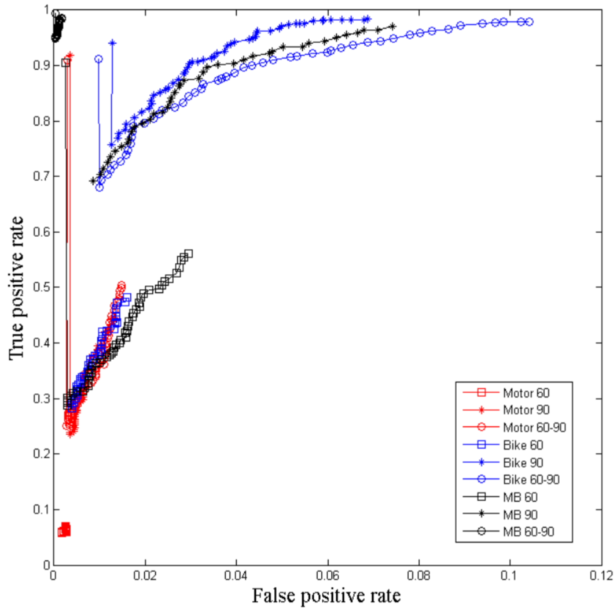
where “*TP*” is True Positive”, “*FP*” is False Positive”, “*TN*” is True Negative and “*FN*” is False Negative. In Figure 3, “MB” is a mixture of motorcycles and bicycles. Also, the numerals behind each of the abbreviations “60” signifies within 60◦, “90” within 90◦, and “90-60” a mixture of 90◦and 60◦, respectively, as well.


ROC Curve
True positive rate
False positive rate
Moto 60
Moto 90
Moto 90-60
Bike 60
Bike 90
Bike 90-60
MB 60
MB 90
MB 90-60

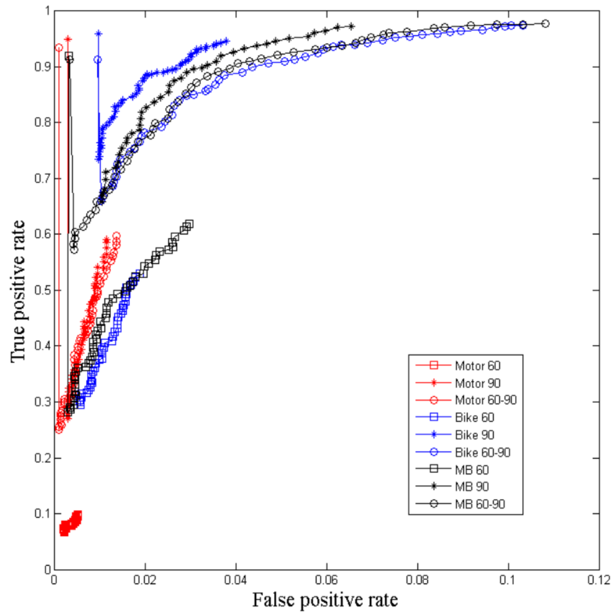
(a)


True positive rate
False positive rate
Moto 60
Moto 90
Moto 90 60
Bike 60
Bike 90
Bike 90 60
MB 60
MB 90
MB 90 60

(b)



(c)



(d)

Figure 3. Experiment Results (A) Results of an Ordinary HOG Method, (B) The Result of Local Binary Pattern (LBP). (C) Results of the NCC Experiment by Applying the Suggested Method-No Crossing Local Area (D) Results of the NCC Experiment by Applying the Suggested Method-Crossing Local Area

In Figure 3, we confirmed that the mixed two degrees (90◦ and 60◦) showed a bigger area for the receiver operating characteristic (ROC) curves than non-mixed degree. For the HOG features, we have used 8x8 pixels for a cell and 3x3 cells for a block. To use the NCC, consider two series $f(x)$ and $t(x)$ which are histogram features where $x \in \{0, \ldots, N-1\}$. By the equation (9), the proposed NCC(W_γ) is defined as

$$W_\gamma = \frac{1}{n}\sum_{x,y}\frac{[f(x)-\bar{f}][t(x-u)-\bar{t}]}{\{[f(x-\bar{f}]^2[t(x-u)-\bar{t}]^2\}^{0.5}} \tag{9}$$

where $f(x)$ is local cell histogram features, $t(x)$ is target histogram features, $\bar{f}$ and $\bar{t}$ are the means of corresponding histogram features.

In general, a positive image which contains a bicycle or motorcycle consists of a bicycle or motorcycle, human, and background, as shown in Figure 3 (c) and (d). In this paper, the special area was established as the blue region which is used as the target histogram features in the Figure 4. The target histogram features are the average value of each bin of cells for the target area.

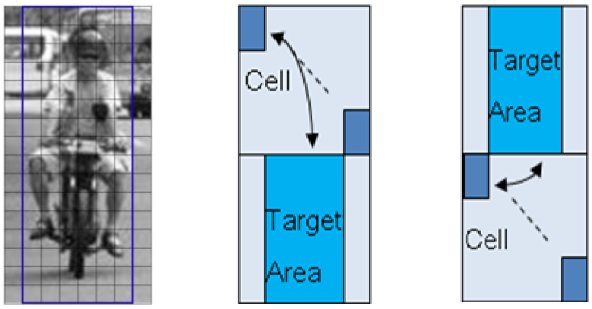


Figure 4. Proposed NCC Method to Calculate the NCC Weighting Value (A) NCC1 Calculate Weighting Value Using Global Area and (B) NCC2 Method to Calculate Weighting Value Using Crossing Local Area

Using this target histogram, the proposed new two wheeler detection method using the HOG is calculated as follows:

$$H_j = C_j * |W_\gamma| \tag{10}$$

where C_j means jth cell histogram and H_j is new histogram features which is weighted NCC value between local cell histogram features and target histogram features.

In Figure 3 (a), the ordinary HOG method, it is shown that the experiment according to "MB 90-60" has the best results among these experiments, but the recognition rate is significantly low. However, Figure 3 (c) and (d) show that the results of M and B experiments according to the proposed method have a higher recognition rate than ordinary HOG. When the proposed algorithm was applied to the other characteristic vector method, a higher recognition rate could be obtained, and the results are listed in Figure 3 (c) and (d). The highest accuracies for each of the methods were calculated with equation (11) [10] and the results are listed in Table 1.

As shown in Table 1, MB (motorcycle and bike) has higher accuracies than Bike and Motor for the proposed method, signifying that MB has a trend of better classifying characteristics than Bike and Motor. And among proposed methods, NCC2 indicated higher recognition rate than NCC1. In our opinion even if an MB becomes more complicated by loading baggage at the rear or by the higher loading of baggage than a bicycle, our proposed algorithm improves the recognition rate better than others. Because of similarity for the object, crossing NCC method (Figure 4(b), NCC2) is higher than NCC1 method. For the same database, a comparison of the performance results for other different feature types (LBP) is shown in Figure 3 (b). The result of LBP features shows lower performance than HOG and the proposed algorithm, because of a concentrated lower area in the Figure 3 (b).

Table 1. Accuracies for each of the Method

Angle(◦) \ Method		HOG	LBP	NCC1_HOG	NCC2_HOG
60	***M***	61.1	53.7	67.5	79.4
	B	71.2	72.2	90.2	91.6
	MB	76.7	71.1	90.6	92.0
90	***M***	74.9	48.0	91.8	95.0
	B	78.3	77.8	95.7	95.4
	MB	76.1	74.5	94.8	95.2
90-60	***M***	77.8	71.8	91.1	93.5
	B	75.5	73.4	94.1	93.8
	MB	73.1	72.0	99.4	94.2

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{11}$$

6. Conclusion

For accuracy and efficiency, two wheeler detection with riders in a still image is one of the most difficult works due to a variety of poses, as well as environmental conditions and cluttered backgrounds. This representation yields not only a suggested solution for two wheelers on the road but a computationally efficient algorithm using normalized cross correlation. In this study, we have introduced a novel practical implementation of the solution for vulnerable road users on the road using NCC weighting value for HOG features. It has been experimentally demonstrated that the proposal using the NCC1 and NCC2 method as weighting value leads to better classification results than other traditional methods from ROC.

References

[1] H. Jung, U. Ehara, J. K. Tan and H. Kim, “Applying MSC-HOG Feature to the Detection of Human on a Bicycle”, Intl. Conference on Control, **(2012)**, October 17-21.
[2] H. Cho, P. E. Rybsti and W. Zhang, “Vision-based Bicyclist Detection and Tracking for Intelligent Vehicles”, Intelligent Vehicle Symposium, **(2010)**, July.
[3] G. Anant Bhaskar, D. Parag and S. Mukesh, “Artificial Neural Networks for Internal Combustion Engine Performance and Emission Analysis”, International Journal of Computer Applications, vol. 87, no.5, **(2014)**, pp. 23-27.
[4] Y. H. Lee, “Histogram of Oriented Gradients and Normalized Correlation Coefficient based Two-wheeler Detection System using Adaboost”, Advanced Science and Technology Letters, vol. 129, **(2016)**, pp. 59-64.
[5] Y. H. Lee, J. Y. Ko, J. H. Suk, T. M. Roh and J. C. Shim, “Pedestrian Recognition using Adaboost Algorithm based on Cascade Method by Curvature and HOG”, The Journal of KIISE, vol. 16, no. 6, **(2010)**, pp. 652-662.
[6] N. Dalal and B. Triggs, “Histogram of Oriented Gradients for Human Detection”, IEEE CVPR, **(2005)**, June 25.
[7] F. Porikli, “Integral histogram: a fast way to extract histograms in Cartesian spaces”, IEEE CVPR, **(2005)**, June 25.
[8] D. M. Tsai and C. T. Lin, “Fast normalized cross correlation for defect detection”, Pattern Recognition Letters, vol. 24, no. 15, **(2003)**, pp. 2625-2631.
[9] P. Viloa, M. Jones and M. Snow, “Detecting pedestrians using patterns of motion and appearance”, The 9th ICCV, **(2003)**, October 13-16.
[10] http://en.wikipedia.org/wiki/Receiver_operating_characteristic, **(2016)**, June.

Author

Yeunghak Lee, he received his Ph.D degree from Yeungnam University, Korea, in 2003. He had a year experience at University of Cardiff as postdoc research fellow. Currently, he is a professor of department of avionic electronic engineering at Kyungwoon University. And he is contributing himself as a management editor for the Journal of Multimedia and Information Systems. His research interests include pattern recognition, embedded system and computer vision.