Optimal Feature Selection for Pedestrian Detection based on Logistic Regression Analysis

Jonghee Kim, Jonghwan Lee, Chungsu Lee, Eunsoo Park, Junmin Kim, Hakil Kim
School of Information and Communication Engineering
Inha University
Incheon, Republic of Korea
{jhkim, jhlee, cslee, espark, jmkim}@vision.inha.ac.kr, hikim@inha.ac.kr

Jaeeun Lee, Hoeri Jeong
School of Statistics
Inha University
Incheon, Republic of Korea
{jennalee831, hrjeong0328}@naver.com

***Abstract*—This paper describes a pedestrian detection method using feature selection based on logistic regression analysis. As the parent features, Haar-like and Histograms of Oriented Gradients (HOG) features are used manually. For the statistical analysis, stepwise forward selection, backward elimination, and Least Absolute Shrinkage and Selection Operator (LASSO) methods are applied to our Logistic Regression Model for Pedestrian Detection (LRMPD). The experimental results shows that the average of 48.5% of a full model are selected for LRMPD and this classifier shows performance of up to 95% for detection rate with an approximately 10% false positive rate. Processing time for one test image is about 1.22ms.**

Keywords-pedestrian detection; logistic regression; multi-feature; feature selection;

I. INTRODUCTION

For many modern people, it is necessary to drive a private car. As roads get busier, traffic accidents have been steadily increasing. Thus, Advanced Driver Assistance Systems (ADAS) have come into existence. Pedestrian detection is the most important issue in this field. Pedestrian detection depends on computer vision for ADAS, for which various approaches have been proposed. It is easier to detect pedestrians using infrared cameras or laser scanners, but such devices are more expensive than simple normal cameras and getting people to install such expensive sensors on their cars is difficult. Thus, research with webcams or cameras on smart phones is proceeding.

There are several methods of feature extraction and classification for pedestrian detection in images. In 2003, P. Viola, M. J. Jones, and D. Snow proposed Haar-like wavelet features and an Adaboost classification cascade [1]; Dalal and Triggs in 2005 proposed Histograms of Oriented Gradients (HOG) with a linear SVM for training and classifying certain samples [2]. Haar-like and Adaboost methods can recognize objects rapidly but have high numbers of false positives. On the other hand, a combination of HOG and SVM gives much better power but is time-consuming. These two are the most interesting methods of pedestrian detection, and many studies are being done with them, such as slightly changing their parameters, like [3], or integrating several of the above features. The part-based method is one other way of performing human detection. Using this method, template matching [4] and the use of HOG features [5] have been presented.

Unlike the use of single features, as detailed above, multi-feature or feature selection methods have also been recently studied for pedestrian detection. In 2008, P. Geismann used Haar and HOG features in a two-staged approach [6]. He generated hypotheses using a Haar detector in the first stage and verified whether the hypotheses were right using HOG feature extraction. Nevertheless, these separated stage methods are complex, and in the processes, some insignificant features are included.

Regression analysis is a statistical technique for estimating the relationships among variables. From such an analysis, we can learn how a dependent variable changes when any one of several independent variables is varied, while the other independent variables are held fixed. On the contrary, logistic regression is used for estimating expected values of a categorical dependent variable in a qualitative response model, based on the explanatory variables. After setting a model, we can pick out the meaningful variables by some variable selection methods. In this model, there left only significant variables or features. In addition, it is considered a more efficient method than step-by-step detection as it has reduced dimensions.

The outline of this paper is as follows. In section 2, the feature extraction process is described; in section 3, we propose the LRMPD (Logistic Regression Model for Pedestrian Detection) based on logistic regression analysis using feature selection methods. Section 4 explains the utilized data set; section 5 shows some experiments with and results obtained using our LRMPD. Finally, summarized contents and conclusions of this paper are shown in section 6.

II. FEATURE EXTRACTION

A. Haar-like

1) Mask selection: There are various types of masks for extracting Haar-like features. So, it is important to select appropriate masks for pedestrian detection; we did this manually for the statistical analysis, because, we might say, vertical shapes may be approprate for pedestrian detection.

978-1-4799-0652-9/13 $31.00 © 2013 IEEE
DOI 10.1109/SMC.2013.47

2) Feature value calculation: For each selected mask, the feature value $f_{Haar}(x)$ is calculated as the difference between the sum of the pixel values within the black and white mask regions. Equation (1) explains this. Normalization progressed after calculation.

$$f_{\mathrm{Haar(x)}} = s_{black(x_1)} - s_{white(x_2)} \tag{1}$$

$$where,$$

$$\mathrm{s(x)} = \sum_{region} pixel(region)$$

Figure 1. Some results of Haar-like features

Figure 1 shows some selected Haar-like feature masks and their value, size and position in the image as a result.

B. HOG

1) Gamma normalization: We used grayscale images and performed square root gamma compression to normalize them.

2) Gradients computation: To acquire the gradient information of each image, i.e., the magnitude m(x,y) and the orientation θ(x,y) of the pixel I(x,y), we used following equations (2) and (3).

$$m(x,y) = \sqrt{d_x(x,y)^2 + d_y(x,y)^2} \tag{2}$$

$$\theta(x,y) = tan^{-1}\frac{d_y(x,y)}{d_x(x,y)} \tag{3}$$

$$where,$$

$$d_x(x,y) = I(x+1,y) - i(x-1,y)$$

$$d_y(x,y) = I(x,y+1) - I(x,y-1)$$

After calculation, we categorized the orientation of the pixels between 0 and 8; In other words, a pixel with an orientation from 0 to 19 degrees became 0 and one with an orientation from 140 to 159 degree became 7.

3) Cell histogram generation: First, a cell has 8×8 pixels, and we generated a cell histogram that has a bin of categorized orientation from 0 to 8 and values for the accumulated magnitude of all pixels included in each bin.

4) Block histogram generation: A block has 2×2 cells so that each block has 16×16 pixels. The values of the cells were recalculated according to a weighted vote and were collected into bins over local spatial regions. Each block histogram was generated with a block spacing stride of 8 pixels.

5) Normalization and accomplishment descriptor: After generating, we normalized the sum of the histogram values between 0 and 1 for each block histogram and concatenated all block histograms. This is a HOG descriptor.

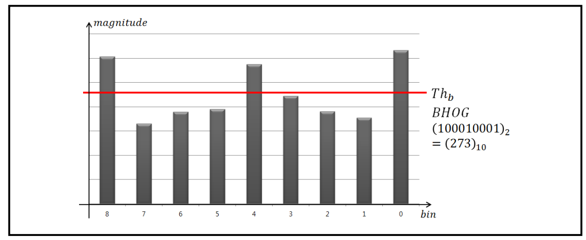


Figure 2. Part of the process of BHOG for one block of an image

C. BHOG

All processes are the same as those of HOG before generating the block histogram but we calculate the average value of the block magnitudes and apply this value Th_B as the threshold to binarize each histogram value as 0 or 1. After binarization, we convert the value to a decimal number. This is a BHOG descriptor. The following equations (4) and (5) were used in the BHOG process; Figure 2 shows this.

$$\mathrm{Th_B} = \frac{1}{9}\sum_{bin=0}^{8} mag(bin) \tag{4}$$

$$BHOG = \sum_{bin=0}^{8} f(mag(bin) - Th_b) \times 2^n \tag{5}$$

$$function\ f\ defined\ as$$

$$f(x) = \begin{cases} 0, & if\ x < 0, \\ 1, & otherwise. \end{cases}$$

III. Logistic Regression

A. Logistic regression analysis

Logistic regression analysis is a type of regression analysis used to predict the outcome of a categorical dependent variable based on one or more predictor (independent) variables. That is, it is used to estimate the expectation values of the parameters in a qualitative response model. Generally, logistic regression is used for problems in which the dependent variable is binary. We used some regression analysis to see if each feature was meaningful or not.

1) Stepwise forward selection and backward elimination: Stepwise forward selection and backward elimination methods are most widely used in regression analysis. Stepwise method starts with no variables in the model. At each step, indepentent variables are included one by one if they are significant and gotten rid of if not. This process goes on until all variables in the model are significant. Backward elimination, on the other hand, starts with all variables. Then, we exclude one at a step from the most insignificant one. What makes this method different from the stepwise method is that variables gotten rid of never comes in again. This process is repeated until all variables in the model are signficant.

2) Least Absolute Shrinkage and Selection Operator (LASSO): Because of dissatisfaction with the ordinary least squares (OLS) method, that is, with its poor prediction accuracy and interpretation, R. Tibshirani proposed LASSO [7]. LASSO, unlike normal ridge regression, shrinks some coefficients and sets others to 0; hence, we can obtain meaningful features from the full model.

After regression analysis, we were able to obtain the linear regression model. Equation (6) is the normal linear regression model.

$$y = \beta_0 + \sum_{i=1}^{p} x_i \beta_i + \varepsilon \qquad (6)$$

The expectation of this linear regression model is continuous. However, in a study to pedestrian detection, the response is binary. Therefore, it is necessary that we transform this into a logistic function by using the following logit transformation.

$$ln \frac{P(y = 1 \mid x)}{1 - P(y = 1 \mid x)} = \alpha + \beta x, \qquad P = \frac{e^{\alpha + \beta x}}{1 + e^{\alpha + \beta x}} \qquad (7)$$

B. Logistic regression model

According to the results of the above logit transformation, we are able to propose the following logistic regression models, which we call LRMPD, the Logistic Regression Models for Pedestrian Detection. We also need a classifier to divide images into classes of 'positive' or 'negative'. Equation (8) mathematically represents the logistic regression model.

$$P = (y = 1 \mid x) = \frac{1}{1 + e^{-z}}, \qquad z = \beta_0 + \sum_{i=1}^{p} x_i \beta_i \quad (8)$$

$P(y = 1|x)$ is the probability of 'y = 1' given the values of independent variables. After transformed, it has continuous values lying between 0 and 1. Hence, in order to distinguish an observation, a classification criterion is needed. The specifications of our model are shown in Table I.

TABLE I. SPECIFICATIONS OF LRMPD

LRM Type	Number of variables (features)			
	Stepwise Selection	***Backward Elimination***	***LASSO***	***Full Model***
HOG + Haar	51 (37.2)	56 (40.9)	79 (57.7)	137
BHOG + Haar	58 (42.3)	65 (47.4)	90 (65.7)	137

( ) : percentage for the full model

LRM type 'HOG + Haar' is selected feature from the combination Haar and HOG; 'BHOG + Haar' is from the Haar and BHOG combined model. As a result, the full model can be reduced by about 37 to 65% by using LRMPD. We can also reduce the processing time required to calculate the expectations of the model.

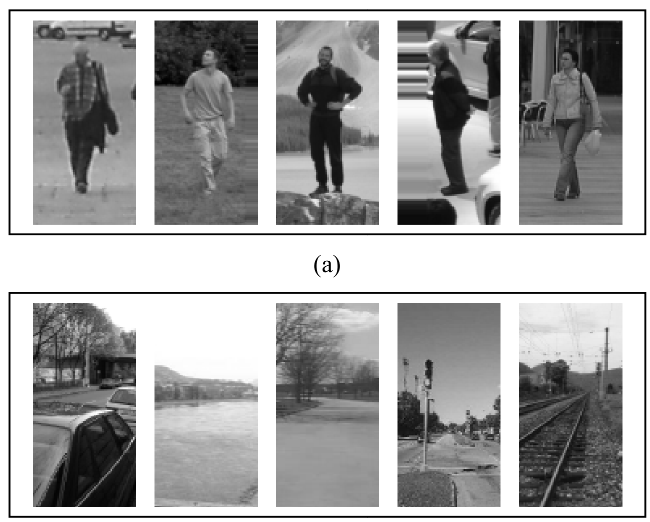
IV. DATA SET

A. Training Set

We trained our LRM with the 'INRIA' person data set, which contains thousands of positive and negative images. Figure 3 shows some examples. We used a set of 2,000 positive images and 4,000 negative images to make the classifier. Each positive training sample has a size of 64×128 and is and image of people. Negative images have irregular sizes and aspect ratios, so we randomly cropped negative images into images of regular size and aspect ratio.

B. Testing set

We also used the 'INRIA' person data set to test our LRM as a classifier, with 500 positive and 500 negative images not used in the training process.



(a)

(b)

Figure 3. Examples of data set: (a) positive (b) negative

V. EXPERIMENTS AND RESULTS

A. Experimental proceduree

First, we extracted the feature values and obtained an appropriate model; then, we calculated the model value by multiplying each feature value by the coefficient of its LRM and summing up. Each observation (image) is classified by its own expectation. When $P(y = 1|x)$ is the expectation of each observation $\hat{y}$, the criterion of classification is shown as in equation (9).

$$\hat{y} = \begin{cases} 1, & P(y=1|x) \geq 0.3, \\ 0, & otherwise. \end{cases} \quad (9)$$

Because a negative trained image has twice as many trained as a positive image, the decision value for classifying is '0.3'.

B. Results

The best LRMPD is the BHOG + Haar-like feature with backward elimination or LASSO method. The experimental results are shown in the Table II, below. Table III shows the processing time of each LRM.

TABLE II. PERFORMANCE OF EACH LRMPD

LRM Type	Stepwise Selection		Backward Elimination		LASSO	
	DR	*FPR*	*DR*	*FPR*	*DR*	*FPR*
HOG + Haar	0.936	0.135	0.930	0.131	0.940	0.138
BHOG + Haar	0.942	0.100	0.948	**0.097**	**0.950**	0.105

Unit : percentage(%)

TABLE III. PROCESSING TIME OF EACH LRM

LRM Name	Stepwise Selection	Backward Elimination	LASSO	Full Model
HOG + Haar	7.044	7.001	7.013	7.399
BHOG + Haar	1.049	1.061	1.068	1.078

Unit : milliseconds(ms)

C. Experimental environments

System environments of experiment are shown as Table IV.

TABLE IV. SYSTEM ENVIRONMENTS

CPU	Intel® Core™ i7 CPU 920 @ 2.67GHz
RAM	DDR3 6GB RAM
OS	Microsoft Windows 7 Enterprise (64bit)
Tool	Visual Studio 2010 (C++, OpenCV 2.4.4)

VI. CONCLUSIONS AND FUTURE WORK

This paper introduces a logistic regression model for pedestrian detection. As parent features, Haar-like, HOG, and BHOG, HOG's derivative feature, are used. Stepwise forward selection, backward elimination, and LASSO methods are used for logistic regression analysis. As for the results, the BHOG + Haar LRM had the best performance as its detection rate was about 95% and its false positive rate was about 10%. And, the processing time of this LRM is about 1.22ms, including feature extraction. The advantages of LRMPD are efficiency, which arose because of the exclusion of meaningless features and the ease of interpretation. Moreover, using LRMPD, there is no necessity for a separate classifier. However, high FPR is the drawback of this system. So, improvement of FPR and reduction of processing time are needed soon. For these improvements, we can apply other parent features or can vary the used features.

ACKNOWLEDGMENTS

This research was funded by the MSIP(Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2013, and the Technology Innovation Program (10040723) funded by the Ministry Of Trade, Industry & Energy(MOTIE, Korea).

REFERENCES

[1] P. Viola, M. J. Jones and D. Snow, "Detecting pedestrians using patterns of motion and appearance," in *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference On,* 2003, pp. 734-741 vol.2.

[2] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in 2005, pp. 886-893.

[3] B. J. Jeon and D. J. Kim, "Real-time Pedestrian Detection based on Binary HOG," in *Human Computer Interation KOREA, 2012 HCI Conference,* vol. 2012, pp. 500-502, 2012.

[4] D. T. Nguyen, W. Li and P. Ogunbona, "A part-based template matching method for multi-view human detection," in *Image and Vision Computing New Zealand, 2009. IVCNZ '09. 24th International Conference,* 2009, pp. 357-362.

[5] A. C. Cosma, R. Brehar and S. Nedevschi, "Part-based pedestrian detection using HoG features and vertical symmetry," in *Intelligent Computer Communication and Processing (ICCP), 2012 IEEE International Conference On,* 2012, pp. 229-236.

[6] P. Geismann and G. Schneider, "A two-staged approach to vision-based pedestrian recognition using haar and HOG features," in *Intelligent Vehicles Symposium, 2008 IEEE,* 2008, pp. 554-559.

[7] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society.Series B (Methodological),* vol. 58, pp. 267-288, 1996.