

Efficiency Improvement in the Extraction of Histogram Oriented Gradient Feature for Human Detection Using Selective Histogram Bins and PCA

Chi Qin Lai and Soo Siang Teoh

Abstract Histogram of Oriented Gradient (HOG) feature which was originally proposed by Dalal and Triggs is widely used in vision-based human detection. However, HOG feature extraction method produced a large feature pool which is computationally intensive and very time consuming, causing it not so suitable for real time application. This paper proposed a method to reduce the HOG feature extraction time without affecting too much on its detection performance. The proposed method performs feature extraction using selective number of histogram bins. Higher number of histogram bins which can extract more detailed orientation information is applied on the regions of image that may contain human figure. The rest of the regions in the image are extracted using lower number of histogram bins. This will reduce the feature size without compromising too much on the performance. To further reduce the feature size, Principal Component Analysis (PCA) is used to rank the features and select only the representative features. A linear SVM classifier is used to evaluate the performance of the proposed method. Experiment was conducted using the INRIA human dataset. The test results showed that the proposed method is able to reduce the feature extraction time by 2.6 times compared to the original HOG and 7 times compared to the LBP method while providing comparable detection performance.

Keywords Human detection · Feature extraction · HOG feature · Object classification

C.Q. Lai (✉) · S.S. Teoh (✉)

Intelligent Biometric Group, School of Electrical and Electronic Engineering,
USM Engineering Campus, Universiti Sains Malaysia, 14300 Nibong Tebal, Malaysia
e-mail: lailai_comecome@msn.com

S.S. Teoh

e-mail: eeteoh@usm.my

© Springer Science+Business Media Singapore 2017

H. Ibrahim et al. (eds.), *9th International Conference on Robotic, Vision, Signal Processing and Power Applications*, Lecture Notes in Electrical Engineering 398,
DOI 10.1007/978-981-10-1721-6_29

1 Introduction

In conjunction with technology growth, the development of computer vision has become an important part of artificial intelligent, which makes use of the interpretation of data for object recognition. Human detection is one of the important research areas in computer vision. This is because the technology can be applied to various fields, which include pedestrian detection in smart vehicles system, intruder detection in security system and crowd monitoring in public surveillance system. However, pedestrians always appear in various types of clothing and in different postures and manners of body movement, causing pedestrian detection to become a difficult task. In recent years, different approaches of vision-based human detection have been proposed in the literature [1–3]. Existing representative works includes the Local Binary Pattern (LBP) [4], the HAAR features descriptor [5], the SIFT feature extraction [6] and Histogram of Oriented Gradient (HOG) [7].

The most common approach of human detection uses features-classifier based technique. A set of representative image features is extracted from the image and a classifier is trained using the extracted features to distinguish between human and non-human. Dalal and Triggs [7] had proposed a grid descriptor based on Histogram of Oriented Gradient, which extracts the features by using well-normalized local histograms of image gradient orientation in a dense grid. This method is computationally intensive and requires long processing time. In this paper, a time efficient HOG extraction method is introduced, targeting to speed up the feature extraction and classification time of the original HOG.

This paper can be summarized as below: Sect. 2 explains about the HOG feature. Section 3 describes the methodology of the proposed method, while Sect. 4 presents the experimental results. Finally, a conclusion is given in Sect. 5.

2 HOG Feature Extraction

HOG feature which was initially proposed by Dalal and Triggs makes use of the image gradient orientation and their normalized histogram. In their proposed method, each color channel of the input image will first be gamma normalize, followed by gradient calculation for each pixel. The gradient image is then divided into grids of cells with 8×8 pixels. A sliding window with size 16×16 pixels will slide through the cells, forming overlapping blocks. Each pixel in a block will then vote based on their gradient orientation into a histogram using trilinear interpolation. Next, the histograms of each block are locally normalized. Finally, the resulting histograms from all the blocks are concatenated to form the HOG feature vector.

However, the original HOG method creates a large pool of feature vector, causing the method to be time-consuming and not so suitable for real-time application. There are a few approaches done up to date by researchers to improve the

execution speed on the original HOG. These include using hardware acceleration on Graphical Processing Unit (GPU) [8], application of integral image [9], as well as combining HAAR with the original HOG [10].

In this paper, a method for efficient HOG feature extraction is introduced. The main idea of the proposed method is to use selective number of histogram bins in the feature extraction and PCA for feature reduction.

3 The Proposed Method

3.1 Dividing the Input Image into Cells and Blocks

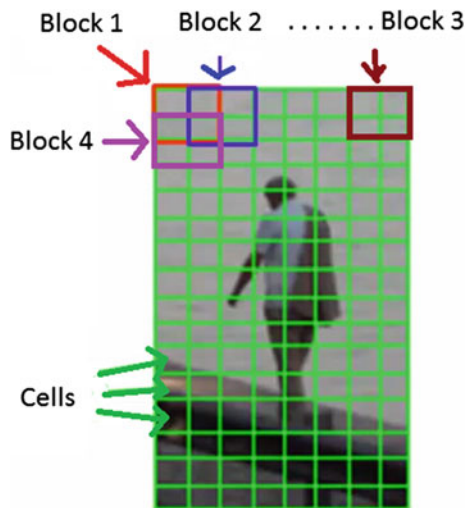
In the proposed method, the input image is first scaled to 64×128 pixels. It is then converted into greyscale image. The resulting image is divided into cells of 8×8 pixels. Every four neighboring cells form a block as shown in Fig. 1. Consequently, a total of 105 overlapping blocks can be formed to facilitate the feature extraction in the subsequent steps.

3.2 Computation of Gradient Orientation

The gradient of each pixel is calculated using the following equations:

$$dx = I(x + 1, y) - I(x, y) \tag{3.1}$$

Fig. 1 Dividing the input image into cells and overlapping blocks (Note not all blocks are shown in the figure)



$$dy = I(x, y + 1) - I(x, y) \quad (3.2)$$

where dx and dy are the horizontal and vertical gradients respectively. $I(x, y)$ is the pixel value.

Next, the gradient orientation, θ is calculated by using:

$$\theta(x, y) = \tan^{-1} \left(\frac{dx}{dy} \right) \quad (3.3)$$

3.3 Construct the Histogram of the Gradients' Orientation

After the pixels' gradient orientations are calculated, a histogram of the orientations is constructed for each block. Different numbers of orientation bins can be used to construct the histogram. Using higher number of bins will capture more orientation details from the image. However, this will generate large feature size and thus slow down the processing time. From our test, it was found that the classification performance improved with the number of bins used in the extraction, with 32 bins giving the best result.

In order to reduce the feature size and yet retaining the important details in the feature, this paper proposed a method to use selective number of histogram bins for different regions in the image. Blocks that are located in the regions that may contain human figure use a higher number of histogram bins, while the rest of the blocks use a lower number of histogram bins. In order to determine the regions in the image that may contain human figure, an average image was constructed from hundreds of positive human images (Fig. 2a). A grid of the blocks is then placed on the average image to identify the blocks that may contain human figure and thus extracted using a higher number of histogram bin (Fig. 2b). Different combinations of high and low numbers of histogram bins have been tested and the results are presented in Sect. 4.1.

After the histograms for each block are constructed, the histograms' values from neighboring blocks are normalized. This is to reduce the effect of illumination variations in different blocks in the image. Finally, all the normalized histograms are concatenated to form the proposed feature vector.

3.4 Feature Selection by Principal Component Analysis

After the feature extraction, we attempted to further reduce the feature size by removing some of the unimportant features. We used Principal Component Analysis (PCA) to rank the extracted features from the most important to the least. The least important features are then removed from the feature vector. Several

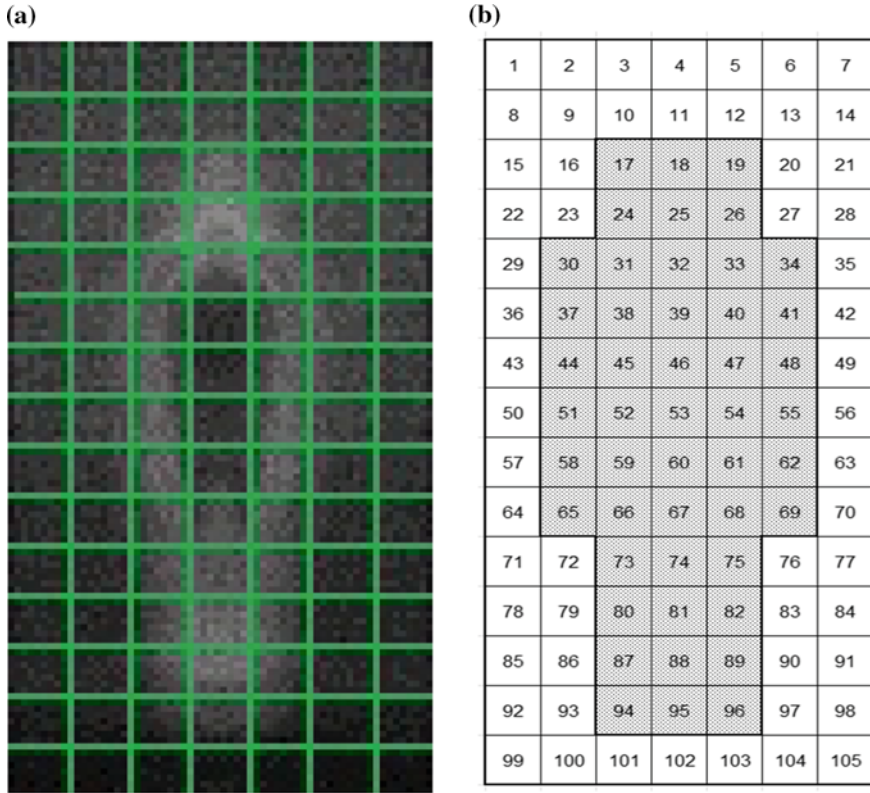


Fig. 2 a From the average image, the important regions in the image that contain human features are determined. Blocks located in these regions (as shown in the shaded blocks in (b)) are extracted with higher number of histogram bins while the rest are extracted with lower number of histogram bins

experiments have been conducted to determine the best number of features to be removed which does not deteriorate too much on the classification performance. The reduced features set forms the optimum features vector proposed in this paper.

4 Results and Discussion

In this section, the performance of the proposed feature will be evaluated and compared with other methods. The algorithms were implemented using the Visual Studio C++ programming language and OpenCV [11] image processing library. The experiments were conducted on a standard Intel core i5 PC without any hardware acceleration. The training dataset used in the experiment consists of 738 positive images and 4065 negative images. Besides, an independent set of 925

positive and 34,184 negative samples were used as the evaluation dataset. The positive samples were taken from INRIA dataset [12] while the negative samples were generated from outdoor images that do not contain any human. For all the experiments, the performances of the features were tested using a linear SVM classifier. The results are presented in the form of Detection Error Trade-off (DET) curves which shows the miss rate against the False Positive per Window (FPPW). On top of that, the feature extraction and classification times were also recorded to evaluate their speed performances.

4.1 Evaluation Using Different Number of Histogram Bins

As discussed in Sect. 3.3, the proposed algorithm uses selective number of bins to construct the gradient histograms for different regions in the image. Higher number of histogram bins is used on blocks that may contain human figure while lower number of bins is used on the rest of the blocks. In this experiment, we evaluated the performance of the features extracted using different combinations of high and low numbers of histogram bins. The combinations of high/low numbers of bins tested are 32bin/16bin, 32bin/8bin, 32bin/4bin and 16bin/8bin. The results are presented in the DET plot in Fig. 3a. As shown in the figure, the curve for the combination of 32bin/16bin is lower among the four curves and thus it gives the best performance. The number of features extracted using this combination is 2448.

This feature size is still large and therefore we further reduced it by using PCA as explained in Sect. 3.4. A total of 948 least important features were removed and thus reduced the feature size to 1500. The performance of the reduced feature set is

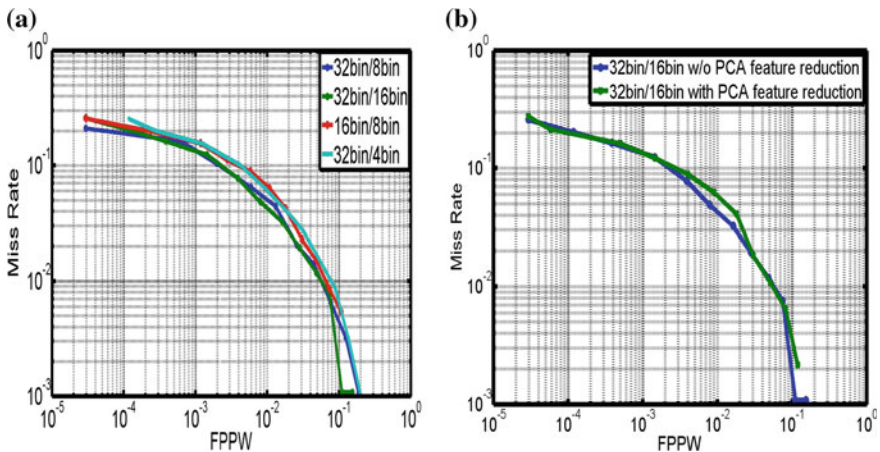
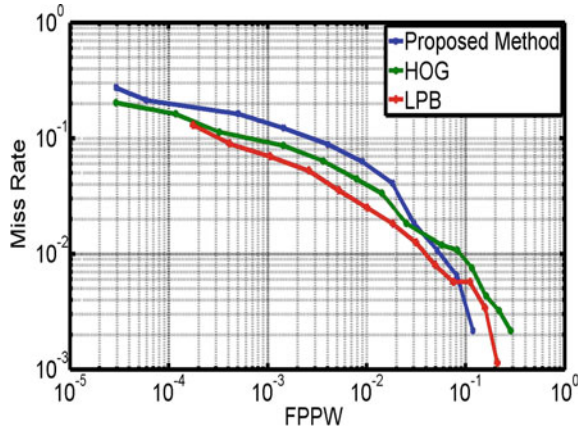


Fig. 3 DET plots to evaluate the performance of the proposed feature. **a** Features extracted using different configurations of high and low numbers of histogram bins. **b** Performance of reduced feature using PCA

Fig. 4 Performance comparison of the proposed method and other methods



given in Fig. 3b. The result showed that its performance does not deteriorate too much compared with the full feature set. By reducing the feature size, the speed efficiency of the feature is improved since the feature extraction and classification time is reduced.

4.2 Performance Comparison with Other Methods

The performance of the proposed feature is compared with the original HOG and the LBP features. The DET plots are shown in Fig. 4. It can be observed that the proposed feature is slightly inferior compared to the other two features. Table 1 shows the missed rate at $FPPW = 10^{-3}$. The miss rate for the proposed system is 10.5 %, which are 1.5 % higher than the original HOG and 3.5 % higher than the LBP feature. Nevertheless, the proposed method is more time efficient in the feature extraction. Table 2 shows the processing time for different methods. It can be seen that both the extraction and classification times for the proposed method are significantly lower compared to other methods. The time required by the proposed method with PCA feature reduction is 2.6 times shorter than the original HOG and 7 times shorter than the LBP feature.

Table 1 Performance of different methods

Types of features	Miss rate at False Positive Per Window (FPPW) = 10 ⁻³ (%)
The original HOG	9.0
LBP	7.0
The proposed method (with PCA feature reduction)	10.5

Table 2 Feature extraction and prediction time of different methods

Types of features	Feature extraction time (ms)	Prediction time (ms)	Total time (ms)
The original HOG	12.63	13.09	25.72
LBP	55.34	14.93	70.27
The proposed method (without PCA feature reduction)	5.47	10.69	16.16
The proposed method (with PCA feature reduction)	4.46	5.56	10.02

5 Conclusion

This paper has proposed a method that improves the efficiency in the HOG feature extraction by using selective histogram bins and PCA. Higher numbers of histogram bins which can extract more detailed orientation information are used only in the regions of the image that may contain human figure. This will reduce the number of features without compromising too much on the performance. To further reduce the feature size, PCA is used to rank the features and discard the unimportant features. Experiment results showed that the processing time of the proposed method is 2.6 times shorter compared to the original HOG, while only tolerate for a small increment in the miss rate. Its processing time is also 7 times shorter compared to the LBP feature. In many applications that required human detection, improvement in the processing time is very critical in order to achieve real-time performance.

Acknowledgments The authors would like to thank the financial support provided by The Ministry of Education Malaysia through the FRGS Grant: 203/PELECT/6071292 (USM).

References

1. Bo L, Qingming Y, Kunfeng W (2012) A review on vision-based pedestrian detection in intelligent transportation systems. In: 9th IEEE international conference on networking, sensing and control (ICNSC)
2. Lai CQ, Teoh SS (2014) A review on pedestrian detection techniques based on Histogram of Oriented gradient feature. In: IEEE student conference on research and development (SCOReD)
3. Gang Z, Youbin C (2012) A review on vision-based pedestrian detection. In: IEEE global high tech congress on electronics (GHTCE)
4. Yadong M et al (2008) Discriminative local binary patterns for human detection in personal album. In: IEEE conference on computer vision and pattern recognition
5. Viola P, Jones M (2001) Rapid object detection using a boosted cascade of simple features. In: IEEE computer society conference on computer vision and pattern recognition
6. Biglari O, Ahsan R, Rahi M (2014) Human detection using SURF and SIFT feature extraction methods in different color spaces. *J Math Comput Sci* 11:111

7. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: IEEE computer society conference on computer vision and pattern recognition
8. Lillywhite K, Dah-Jye L, Dong Z (2009) Real-time human detection using histograms of oriented gradients on a GPU. In: Workshop on applications of computer vision (WACV)
9. Said Y, Atri M, Tourki R (2011) Human detection based on integral histograms of oriented gradients and SVM. In: International conference on communications, computing and control applications (CCCA)
10. Hui-Xing J, Yu-Jin Z (2007) Fast human detection by boosting histograms of oriented gradients. In: Fourth international conference on image and graphics
11. Intel (2013) OpenCV (Open Source Computer Vision) Library Wiki. <http://opencv.willowgarage.com/wiki/>. Accessed 1 Oct 2013
12. Dalal N (2013) INRIA person dataset. <http://pascal.inrialpes.fr/data/human/>. Accessed 15 Oct 2013