An Efficient Method of HOG Feature Extraction Using Selective Histogram Bin and PCA Feature Reduction

Chi Qin LAI, Soo Siang TEOH
Intelligent Biometric Group
School of Electrical and Electronic Engineering, USM Engineering Campus
Universiti Sains Malaysia, 14300 Nibong Tebal, Penang, Malaysia
chi.qin.c.lai@intel.com, eeteoh@usm.my

***Abstract*—Histogram of Oriented Gradient (HOG) is a popular image feature for human detection. It presents high detection accuracy and therefore has been widely used in vision-based surveillance and pedestrian detection systems. However, the main drawback of this feature is that it has a large feature size. The extraction algorithm is also computationally intensive and requires long processing time. In this paper, a time-efficient HOG-based feature extraction method is proposed. The method uses selective number of histogram bins to perform feature extraction on different regions in the image. Higher number of histogram bin which can capture more detailed information is performed on the regions of the image which may belong to part of a human figure, while lower number of histogram bin is used on the rest of the image. To further reduce the feature size, Principal Component Analysis (PCA) is used to rank the features and remove some unimportant features. The performance of the proposed method was evaluated using INRIA human dataset on a linear Support Vector Machine (SVM) classifier. The results showed the processing speed of the proposed method is 2.6 times faster than the original HOG and 7 times faster than the LBP method while providing comparable detection performance.**

***Index Terms*—feature extraction, image analysis, object detection, pattern recognition, computer vision.**

I. INTRODUCTION

In line with the increasing demand for intelligent systems with visual perception capability, research on the application of computer vision has become more intensive and important. In general, computer vision interprets the visual data captured from the surrounding to detect and recognize objects of interest. Human detection is one of the most researched areas in computer vision. This is because it is useful in many applications such as pedestrian detection in driver assistance system [1], surveillance system [2], crowds monitoring system [3], human activity classification [4] and fall detection system for the care of elderly people [5]. However, human detection is a challenging task as human may appear in different types of clothing and postures. In recent years, many approaches for vision-based human detection have been proposed in the literature [6-9]. Some representative works include Local Binary Pattern (LBP) [10], HAAR descriptor [11], SIFT feature [12] and Histogram of Oriented Gradient (HOG) [13].

The most common approach of human detection makes use of a feature-classifier based technique. A set of representative features is extracted from images and used to train a classifier. The trained classifier will then be used to distinguish between human and non-human. One of the most common features for human detection is the Histogram of Oriented Gradient (HOG) which was originally proposed by Dalal and Triggs [13]. HOG is a grid descriptor which extracts the feature by using well normalized local histogram of image gradient orientation in a dense grid. However, HOG feature extraction is computationally intensive and may pose a challenge for implementation in a real time system. In this paper, an efficient method which can speed up the extraction of gradient orientation features for human detection is proposed.

The paper can be summarized as below. Section II reviews the HOG feature and its related works. This is followed by Section III which describes the proposed method in detail. Section IV explains the experiments conducted to evaluate of the proposed method and presents the test results. Finally, a conclusion is given in Section V.

II. RELATED WORKS ON HOG FEATURES

The original HOG feature extraction method makes use of image gradient orientation and normalized histogram. The steps of the feature extraction are summarized as follows.

First, the input image is resized to 64×128 pixels. Gamma normalization is then performed on the image, followed by the computation of gradient's magnitude and angle for each pixel. The resulting image is then divided into grid of cells with size 8×8 pixels. Next, a 16×16 pixels sliding window is placed on top of the grid and slide through the cells. At each step, the sliding window overlaps four cells forming a block. Subsequently for each block, trilinear interpolation is used to vote the gradient magnitude into a histogram based on their gradient's orientation. The histograms are then normalized and assembled to form a 1-D feature vector.

The original HOG feature extraction is computationally intensive and generates a large feature pool. Some improvements to speed up the feature extraction have been proposed by several researchers. One of the improvement is to use 'integral histogram' technique [14] similar to the 'integral image' calculation introduced by Viola and Jones

This work was financially supported by The Ministry of Higher Education Malaysia through the FRGS Grant: 203/PELECT/6071292 (USM)

Digital Object Identifier 10.4316/AECE.2016.04016

[15]. This technique allows efficient calculation of histogram in any rectangular region in the image. Yahia et al. [16] applied this technique to speed up the calculate of histogram in each cell. Their experiments on INRIA and MIT databases showed that this method can speed up the feature extraction process while maintaining the accuracy of the original HOG.

There are some research works that propose to combine HOG features with other feature extraction methods, such as LBP [17], Gabor [18] and Discrete Wavelet Transform (DWT) [19]. The advantage of combining with other methods is that the combined features inherit the beneficial properties from both methods. However at the same time, the extraction process may become more complex and the feature size may also be larger and thus requires longer processing time.

There is another approach proposed by Hemmati et al. [20] which speed up the HOG feature extraction by using hardware acceleration on a field-programmable gate array (FPGA). In their method, parallel hardware architecture with special defined memory access pattern is implemented to improve the processing speed. Their proposed hardware acceleration is able to extract HOG features at 60 fps on a HDTV (1080 × 1920 pixels) frame. In another research, Chen et al. [21] implemented a hardware accelerator integrated circuit for HOG feature extraction using the 0.13-μm technology. Their experiment showed that the circuit is able to give comparable working speed as the FPGA implementation but at a lower hardware cost.

The hardware acceleration method to speed up the HOG feature extraction requires specific hardware and specialized programming technique to implement the algorithm. It may require more expensive hardware for the implementation.

In this paper, a method to efficiently extract the orientation histogram feature based-on selective histogram bin and Principal Component Analysis (PCA) is proposed. Opposed to the original HOG method which uses the same number of orientation bins throughout the image, the proposed technique selectively extracts the features using different number of histogram bins depending on the regions in the image. Regions which may belong to part of a human figure are given higher priority and thus extracted using higher number of histogram bins, while the rest of the regions are using lower number of bins. This will generate less number of features and speed up the extraction and classification time.

To further reduce the feature size, PCA is used to rank the feature and remove some of the unimportant features. The details of the proposed feature extraction process are discussed in the next section.

III. THE PROPOSED METHOD

The general idea of the proposed method is to extract the HOG feature using selective number of histogram bins. Instead of using the same number of histogram bins throughout the image, the proposed method uses a higher number of histogram bins on the regions of the image that may belong to part of a human figure while lower number of bins on the rest of the regions.

To make the features to be more invariant to illumination changes in the image, multiple blocks normalization is used. In addition, to further reduce the feature size, PCA [22, 23] is used to remove some unimportant features and obtain a subset of reduced features which is sufficient to represent the full feature set. SVM classifier [24, 25] is used in the proposed method to classify between human and non-human.

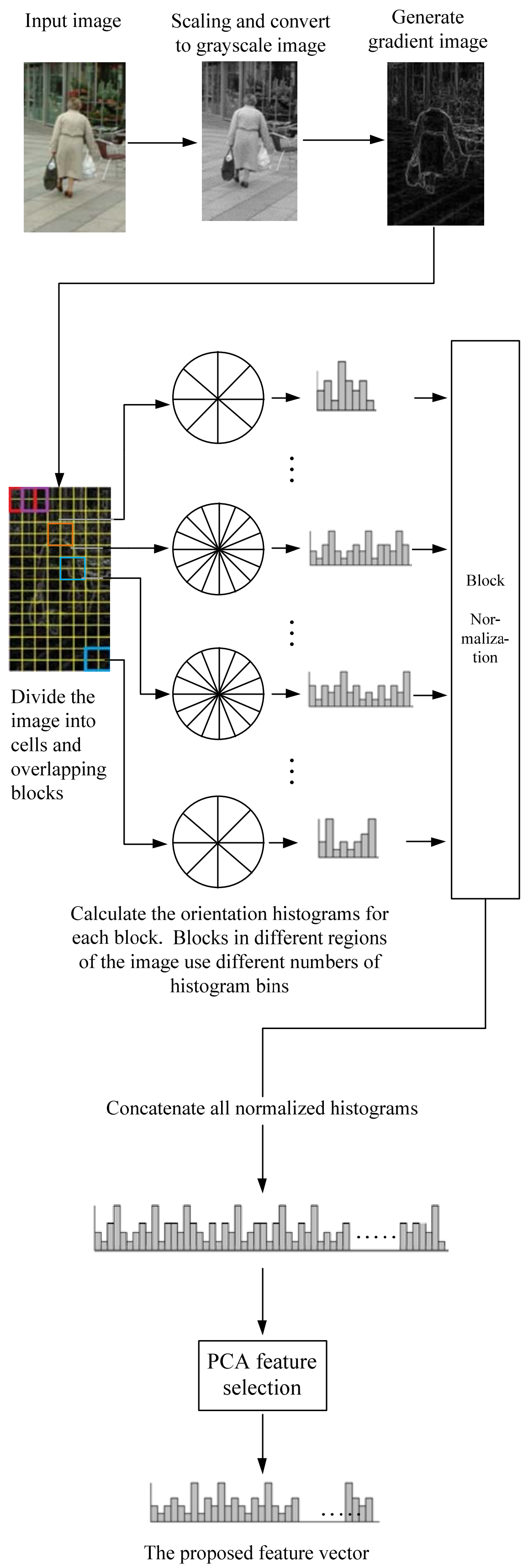


Figure. 1 Summary of the proposed method

The details of each step of the proposed method are explained in the following subsections.

A. Image pre-processing

First, the input image is scaled to 64 × 128 pixels to ensure that consistent size of image is used in the feature extraction. The image is then converted into grayscale.

B. Gradient computation

Next, the gradients for every pixel in the image are calculated using the following equations:

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

$$dy = I(x, y+1) - I(x, y) \tag{2}$$

where dx and dy are the horizontal and vertical gradient respectively, *I(x, y)* is the pixel value at *(x,y)* position.
The gradient orientation, θ is then calculated using:

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

C. Dividing the input image into cells and blocks

The resulting gradient image is divided into cells of 8 × 8 pixels. A sliding window with the size of 16 × 16 pixels is then slides through the cells, with each step covering four neighboring cells. Each group of four neighboring cells forms a block. The neighboring blocks are overlapping with each other as shown in Fig. 1. Through this process, a total of 105 blocks are formed on the 64 × 128 pixels' image. The division of image into blocks is similar to the original HOG. It is needed to facilitate the feature extraction in the subsequent steps.

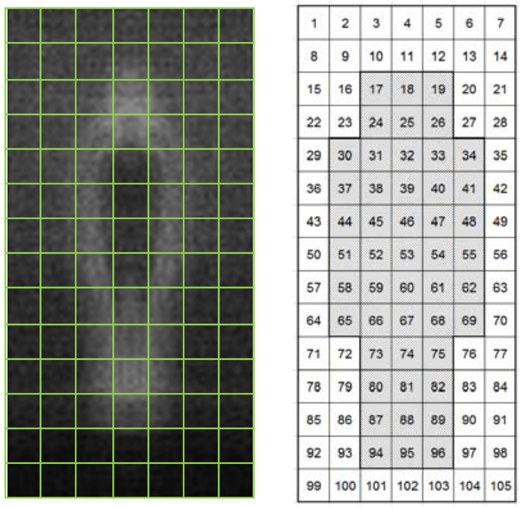
D. Construct the histogram of oriented gradient using selective number of histogram bins

For each block, a histogram for the gradient's orientation is constructed. This is done by voting of the orientation angles of each pixel into a predefined number of histogram bins. Using higher number of bins will extract more detailed orientation information from the image but it will generate higher number of features.

In order to reduce the feature size and yet retain the important details in the feature, different number of histogram bins is used for different regions in the image. Higher number of histogram bins is used to extract features for the regions which may belong to part of a human figure while lower number of bins is used for the rest of the regions.

To identify the regions which may belong to part of a human figure, an average image is constructed from 739 positive training samples. A gird that shows the position of the blocks is then placed on the average image as shown in Fig. 2(a). From this image, the blocks that may contain human figure are identified. These blocks are shown in shaded in Fig. 2(b).

Higher number of histogram bins is used to extract features for the shaded blocks while lower number of histogram bins is used for the rest of the blocks. The optimum values for the high and low number of bins to be used are determined empirically and the experiment results are presented in Section IV.



(a) (b)

Figure 2. (a) A grid to show the position of blocks overlaid on the generated average image (b) The identified blocks that cover the human figure (the shaded blocks)

E. Block normalization

L2 normalization scheme is used to normalize the histogram values to make the features more invariant to changes in illumination.

Based on the experiments conducted by [13], they showed that *L2-norm* is the best normalization schemes for HOG feature extraction and therefore it is used in the proposed method. The *L2* normalization scheme is defined as:

$$v_n = v / \sqrt{\|v\|_2^2 + \varepsilon^2} \tag{4}$$

where v is the unnormalized feature vector, v_n is the normalized feature vector, $\|v\|_2$ is the *L2-norm* value and ε is the small normalization constant to avoid division by zero. The value of *L2-norm* is calculated using the following equation:

$$\|v\|_2 = \sqrt{\sum_{i=1}^{n} v_i^2} \tag{5}$$

Multiple-block normalization is performed by grouping several neighboring blocks according to their number of orientation bins used in the histogram generation. A total of 30 groups were formed as shown in Fig. 3.

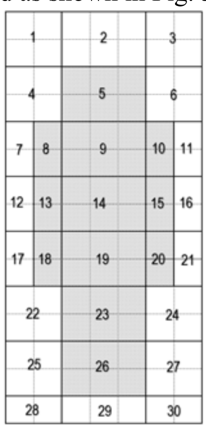


Figure 3. Grouping of blocks to perform normalization

L-2 normalization scheme is then used to normalize all the histogram values in each group. The normalized histograms from all the groups are finally concatenated to form the proposed feature vector.

F. PCA feature selection

Conventionally PCA is used for feature dimension reduction where the *n*-dimensional original features are transformed linearly onto *n* orthogonal axes. Projecting observation onto each of these axes will generate a new set of *n* uncorrelated variables. A subset of these variables with high eigenvalue is then used as the new feature vector. However, the calculation for each derived feature requires $n \times n$ multiplications and the use of all original features. This will increase the computation time for feature extraction. In this study, PCA is used for feature selection to remove some unimportant features. We do this by utilising the information inferred in the coefficients of the Principal Components (PC) for features ranking [23].

First, the Covariance Matrix, *C* of the *n*-dimensional features vector extracted from the positive training samples is generated. Next, the eigenvalues and eigenvectors of *C* are calculated to obtain the Principal Components (*PCs*). A total of *n* possible *PCs* can be obtained. Each of the *PCs* consists of *n* number of coefficients and each coefficient is linked with a correlated feature from the original feature pool.

By starting off with the first PC, the feature which is paired up with the largest coefficient of the PC is placed on the highest rank. The same procedure is applied on the subsequent PCs to obtain a list of features with descending ranking. Based on the ranking, a different number of low ranking features are removed to form a subset of reduced features. Experiments have been conducted to find the best reduced number of features which can represent the full feature set. The results are presented in Section IV. By reducing the size of the proposed feature, the time required for the classification process can also be reduced.

IV. Experiment, Result and discussion

An image classifier is used to evaluate the performance of the proposed feature. In this study, an SVM classifier with Radial Basis Function (RBF) kernel is used as it is shown in [26] that RBF kernel presents high performance. The parameters that need to be optimized in this classifier are *C*, a noise parameter and γ from the RBF kernel. The optimum values of these parameters are obtained by exhaustive search during the experiment.

The training dataset consists of 738 positive images and 4065 negative images. For the evaluation, an independent set of 925 positive and 34184 negative samples were used. The positive samples were taken from INRIA dataset [13] while the negative samples were generated from outdoor images that do not contain any human. Some examples of positive and negative images are shown in Fig. 4.

The evaluation results are presented in the form of Detection Error Trade-off (DET) curves which show 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 performance.

The experiments were conducted on a standard Intel core-*i5* computer without any hardware acceleration. All the algorithms were implemented using the Microsoft Visual Studio C++ programming language and OpenCV [27] image processing library.

The following subsections present the evaluation results. First, the results for the experiments to find the best parameters settings for the proposed method are presented. Then the results of performance comparison between the proposed method and three other methods which are the original HOG, LBP and integral HOG are given.

(b)

Figure 4. Examples of the training images. (a) Positive training samples. (b) Negative training samples.

A. Evaluation on different methods of histogram normalization for features extracted using different number of histogram bins

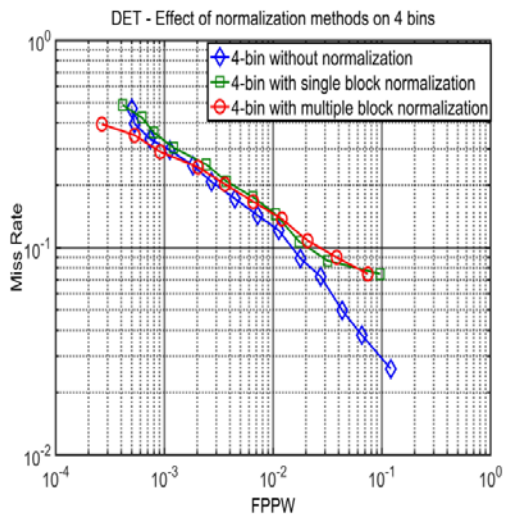
Histogram normalization is done to reduce the effect of large illumination variations on the detection performance. To select the best normalization method to be used in the proposed method, several experiments have been conducted using different ways of histogram normalization based on the *L2-norm*.

The first method performed *L2*-normalization on the histogram values of each block individually while the second method performed *L2*-normalization on a group of neighboring blocks. The grouping of blocks for the second method is as given Fig. 3. The evaluation on the normalization methods was conducted on features extracted using different number of histogram bins, which are 4-bin, 8-bin, 16-bin and 32-bin.

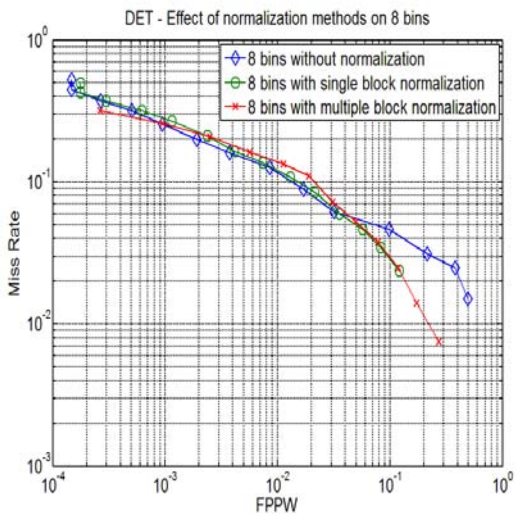
The results for each method of histogram normalization are presented in the DET curves in Fig. 5. For better comparison, DET curves of similar number of histogram bins were plotted on the same graph. The results of feature extraction without any normalization were also plotted as a baseline for comparison. Other than that, Miss Rate at FPPW = 10^{-3} for each of the curves was recorded and given in Table I to compare their performance.

From the results, it was found that multiple blocks normalization is able to improve the performance of the feature compared to single block normalization or without normalization. This is true for features extracted using 4, 8, 16 and 32 bins. By normalizing the histogram values of a group of blocks extracted using similar number of bins, the performance of the feature is improved since it becomes more invariant to illumination changes in the image.

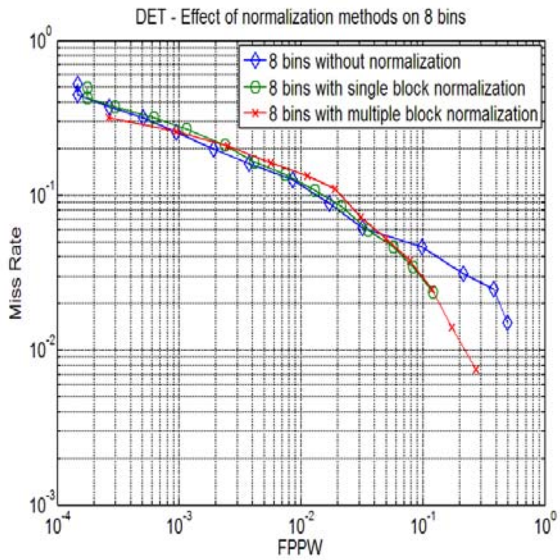
Table II shows the processing time for features extracted using different number of histogram bins and different normalization methods. Extractions using 4, 8, 16 and 32 histogram bins generated 420, 840, 1680 and 3360 features respectively. The result shows that as the feature size increases, the total processing time for feature extraction and classification will also increase. This is because using higher number of histogram bin will require longer processing time to construct the histogram for the feature extraction. Besides, for SVM classification, the computational complexity of prediction using the RBF kernel is $O(d \times N)$ where N is the number of support vectors and d is the feature dimension [28]. This resulted in the increase in classification time when the feature size increases.



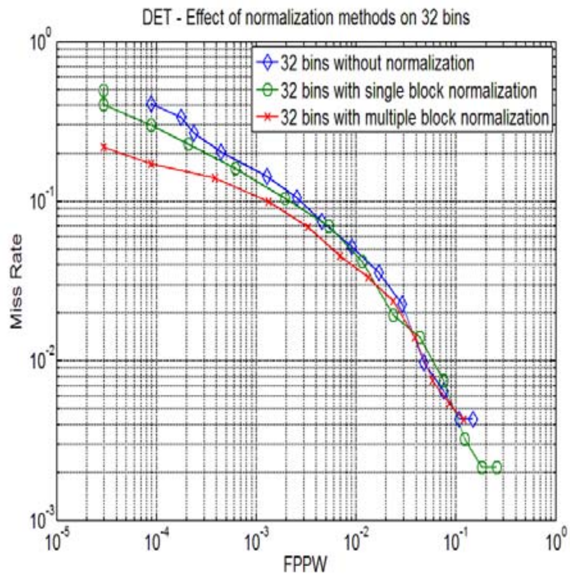
(a)



(b)



(c)



(d)

Figure 5. DET curves to show the classification performance of different normalization methods. Tested on features extracted using different number of histogram bins: (a) 4-bin (b) 8-bin (c) 16-bin and (d) 32-bin

TABLE I. MISS RATE AT FPPW = 10^{-3} FOR DIFFERENT NORMALIZATION METHODS

Number of histogram bins	Miss Rate at FPPW = 10^{-3} (%)		
	Without normalization	**Single block normalization**	**Multiple blocks normalization**
4-bin	30.0	33.2	28.7
8-bin	25.3	28.2	25.2
16-bin	16.3	21.9	14.8
32-bin	16.2	14.5	11.3

TABLE II. TOTAL PROCESSING TIME FOR DIFFERENT NORMALIZATION METHODS

Number of histogram bins	Total Processing Time (ms)		
	Without normalization	**Single block normalization**	**Multiple blocks normalization**
4-bin	5.08	6.63	7.09
8-bin	6.36	7.39	8.15
16-bin	9.35	11.20	14.02
32-bin	14.93	16.1	18.01

B. Evaluation on the selective number of histogram bins for different regions in the image

As mentioned in Section III (D), a higher number of histogram bins will be used on blocks located in the regions which may belong to part of a human figure, while a lower number of histogram bins will be used for the rest of the regions. The purpose of this experiment is to determine the best combination of the high and low numbers of histogram bins to be used for the feature extraction. The combinations of (high/low) numbers of histogram bins that have been tested are 32-bin/16-bin, 32-bin/8-bin, 32-bin/4-bin and 16-bin/4-bin.

The DET curves for each of the combinations are presented in Fig. 6 and their Miss Rates at FPPW = 10^{-3} are given in Table III. From the results, it was found that the combination of 32-bin/16-bin gave the lowest Miss Rate (14.1%) followed by the combination of 32-bin/8-bin (14.7%). For the combinations of 32-bin/4-bin and 16-bin/4-bin, both of them showed higher Miss Rate (16.8%). These results indicate that using 32-bin can effectively extract the important information from the regions which may belong to part of a human figure. For the background region, 4-bin and 8-bin are not sufficient to extract the remaining information. Therefore, the combination of 32-bin/16-bin gave the best detection result and it will be used in the proposed method.

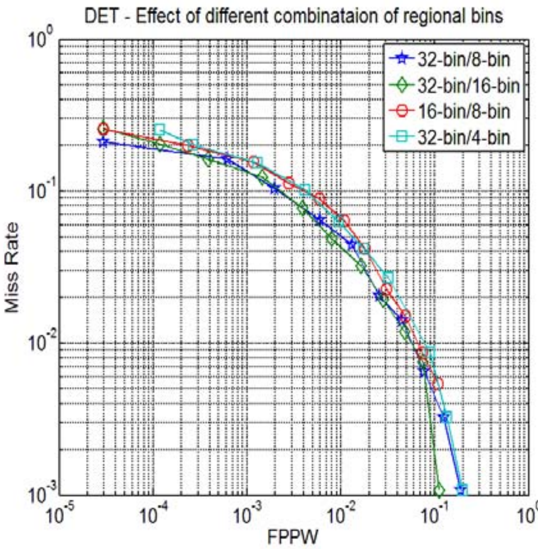


Figure 6. Performance of features extracted using different combinations of high/low number of histogram bins

TABLE III. MISS RATE AT FPPW = 10^{-3} FOR DIFFERENT COMBINATIONS OF HIGH/LOW NUMBER OF HISTOGRAM BINS

Combinations of high/low number of histogram bins	**Miss Rate at FPPW = 10^{-3} (%)**
32-bin/16-bin	14.1
32-bin/8-bin	14.7
32-bin/4-bin	16.8
16-bin/4-bin	16.8

C. Experiment on feature reduction using PCA

The size of the features extracted using 32-bin/16-bin combination is 2041. This size will be further reduced using PCA as discussed in Section III (F). PCA will be used to rank the features according to their importance and some unimportant features will be removed. The purpose of this experiment is to determine the suitable numbers of top ranking features that are sufficient to represent the original feature set.

First, the Principal Components (PCs) of the positive training images are calculated. Next, the coefficients of each PC were analyzed to identify the dominant features. In Fig. 7, a plot to show the cumulative percentage of variance explained by the first 800 PCs is given.

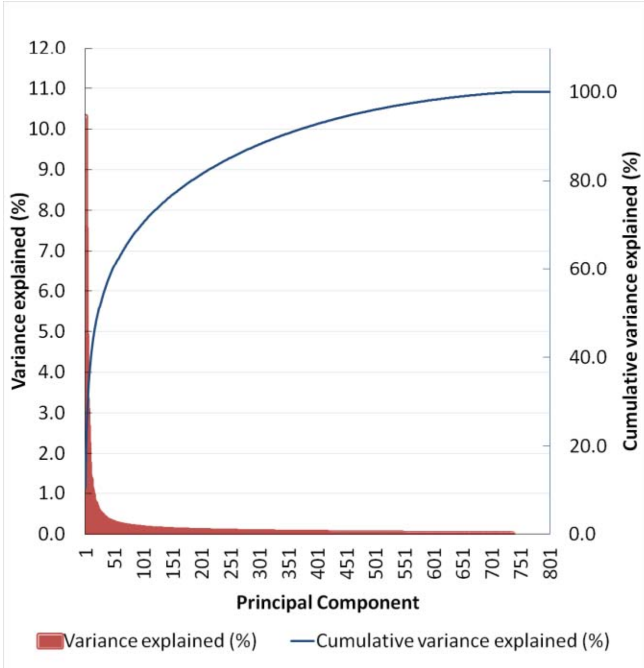


Figure 7. Percentage of variance explained by the first 800 PCs

It can be seen that almost 100% of the total variance comes from the first 729 PCs. This indicates that a subset of the total features should be able to represent the full features set. In this experiment, the performance of features subsets formed by picking the top ranked 500 features (Top500), 1000 features (Top1000) and 1500 features (Top1500) were evaluated.

Fig. 8 shows the DET curves for the three features subsets. The Miss Rate at FPPW = 10^{-3} for each of the curves is recorded in Table IV and their processing time is given in Table V.

From Fig. 8, it can be observed that the DET curve for the top 1500 feature is very close to the DET curve of the full feature set. Table IV shows that the top 1500 feature subset has a Miss Rate of 14.1% at FPPW =10^{-3}, which is similar to the full feature set. On the other hand, both the top 500 and top 1000 feature subsets show higher Miss Rates, which indicates that their detection performances are inferior. This could be due to too much reduction in the feature size which has removed some of the representative features that are important for the classification.

From Table V, it can be seen that in general, the total processing time is reduced with the reduction in feature size. The processing time for the top 1500 features subset (10.02 ms) is 18.36% faster compared to the full feature set (11.86 ms). Using the top 1500 feature subset, a total of 948 (38.7%) unimportant features can be removed.

As a summary, the top 1500 feature subset is sufficient to represent the full feature set. This compact set of features is the final proposed feature in this paper.

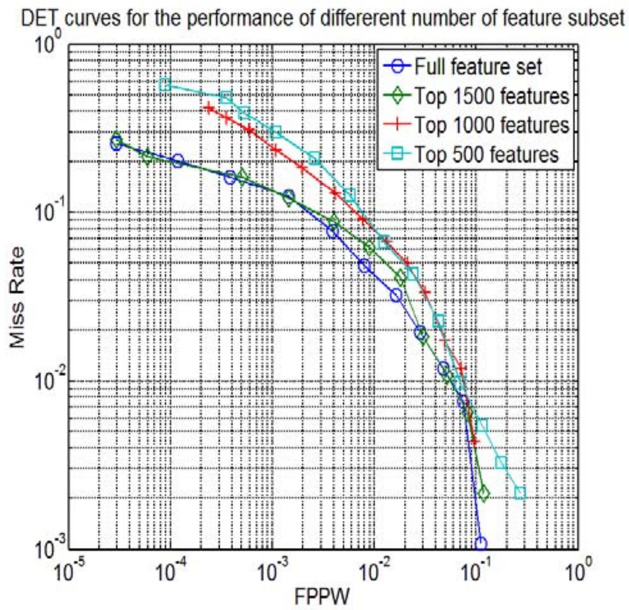


Figure 8. DET curves to compare the performance of different feature subsets

TABLE IV. MISS RATE AT FPPW = 10^{-3} FOR DIFFERENT FEATURE SUBSETS

Feature subset	Miss Rate at FPPW = 10^{-3} (%)
Top 500	31.4
Top 1000	23.4
Top 1500	14.1
Full feature set (2448 features)	14.1

TABLE V. PROCESSING TIME FOR DIFFERENT FEATURE SUBSETS

Feature subset	Total processing time (ms)
Top 500	6.36
Top 1000	7.71
Top 1500	10.02
Full feature set (2448 features)	11.86

D. Performance comparison of the proposed method with other methods

The proposed method was compared with the original HOG feature [13], the integral HOG [16] and the LBP feature , which are commonly used for pedestrian detection.

In the experiments, the same performance measures were used for each of the methods to get a fair comparison. These measures include using the same training and evaluation data sets, running all experiments on the same computer and using SVM classifier for the classification. The results were plotted on the DET curves in Fig. 9. The Miss Rate at FPPW = 10^{-3} are given in Table VI. The total processing time for each of the methods was recorded and presented in Table VII to compare their speed performance.

The results showed that the performance of the proposed method is slightly inferior compared to the other three methods. The Miss Rate at FPPW = 10^{-3} for the proposed method is 4.4% higher than the original HOG, 7% higher than the LBP and 4.7% higher than the integral HOG. However, comparing their speed performances, the proposed method is 2.6 times faster than the original HOG, 2.5 times faster than the integral HOG and 7 times faster than the LBP method.

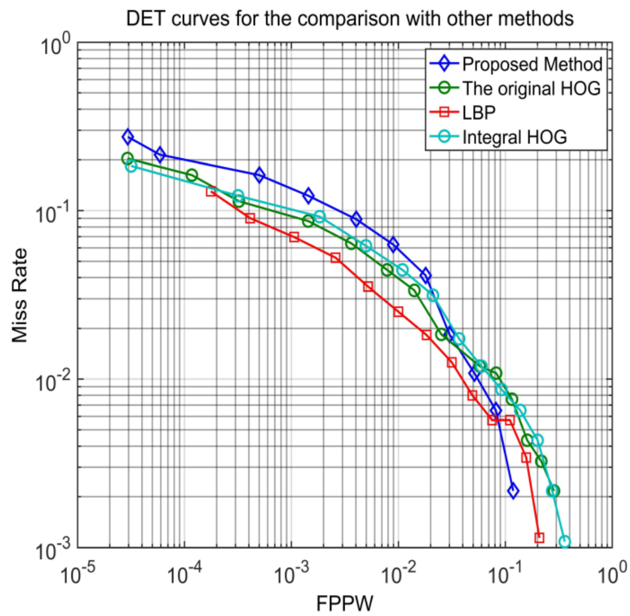


Figure 9. Comparison of the proposed method with the original HOG, the integral HOG and the LBP method

TABLE VI. MISS RATE AT FPPW = 10^{-3} FOR THE PROPOSED METHOD, THE ORIGINAL HOG, THE INTEGRAL HOG AND THE LBP METHOD

Type of Features	Miss Rate at FPPW = 10^{-3} (%)
Original HOG	9.7
Integral HOG	10.0
LBP	7.1
The proposed method	14.1

TABLE VII. PROCESSING TIME OF THE PROPOSED METHOD, THE ORIGINAL HOG, THE INTEGRAL HOG AND THE LBP METHOD

Types of Features	Feature extraction time (ms)	Prediction time (ms)	Total processing time (ms)
Original HOG	12.63	13.09	25.72
Integral HOG	8.00	15.75	23.75
LBP	55.34	14.93	70.27
The proposed method	4.46	5.56	10.02

The proposed method generates a feature set containing 1500 features, while the original HOG and LBP contain 3780 and 2880 features respectively. The integral HOG generates the same number of features as the original HOG. By comparing the size of the feature sets, it can be seen that the proposed method has the smallest size and it requires the shortest processing time. Although LBP contains fewer features than the original HOG, it uses a more complicated method of feature extraction and hence requires the longest feature extraction time among the three methods.

The proposed method is able to reduce the processing time by reducing the number of features using selective number of histogram bins for different regions in the image. The feature size is then further reduced by using PCA. This has effectively shortened the processing time while maintaining comparable detection accuracy.

V. CONCLUSION

In this paper, an efficient method for the extraction of histogram of orientated gradient feature for human detection is proposed. The method performs feature extraction using selective number of histogram bins to reduce the feature size. A higher number of histogram bins is used to extract features from regions which may belong to part of a human figure while lower histogram bins is used for the rest of the regions. This method has successfully reduced the extracted feature size, while providing comparable results. PCA feature selection is used to further reduce the dimension of the feature size by ranking the features according to their importance and removing some irrelevant features. The performance of the proposed feature was compared with the original HOG, the integral HOG and the LBP features. From the evaluation results, it was shown that the proposed method is able to speed up the feature extraction time by 2.6 times compared to the original HOG, 2.5 times compared to the integral HOG and 7 times compared to the LBP method, while providing comparable detection performance.

REFERENCES

[1] P. Dollar, C. Wojek, B. Schiele, and P. Perona, “Pedestrian Detection: An Evaluation of the State of the Art,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 4, pp. 743–761, Apr. 2012. doi: 10.1109/TPAMI.2011.155

[2] J. L. Raheja, S. Deora, and A. Chaudhary, “Cross border intruder detection in hilly terrain in dark environment,” Optik - International Journal for Light and Electron Optics, vol. 127, no. 2, pp. 535–538, Jan. 2016. doi:10.1016/j.ijleo.2015.08.234

[3] J. E. Mallah, F. Carrino, O. A. Khaled, and E. Mugellini, “Crowd Monitoring,” in Distributed, Ambient, and Pervasive Interactions, vol. 9189, N. Streitz and P. Markopoulos, Eds. Cham: Springer International Publishing, 2015, pp. 496–505. doi:10.1007/978-3-319-20804-6_45

[4] R. M. Mueid, C. Ahmed, and M. A. R. Ahad, “Pedestrian activity classification using patterns of motion and histogram of oriented gradient,” Journal on Multimodal User Interfaces, May 2015. doi: 10.1007/s12193-015-0178-3

[5] J.-L. Chua, Y. C. Chang, and W. K. Lim, “A simple vision-based fall detection technique for indoor video surveillance,” Signal, Image and Video Processing, vol. 9, no. 3, pp. 623–633, Mar. 2015. doi: 10.1007/s11760-013-0493-7

[6] C. Q. Lai and S. S. Teoh, "A review on pedestrian detection techniques based on Histogram of Oriented gradient feature," in Proc. IEEE Student Conference on Research and Development (SCOReD), 2014, pp. 1-6. doi: 10.1109/SCORED.2014.7072948

[7] B. Li, Q. Yao, and K. Wang, "A review on vision-based pedestrian detection in intelligent transportation systems," in Proc. 9th IEEE International Conference on Networking, Sensing and Control (ICNSC), 2012, pp. 393-398. doi: 10.1109/ICNSC.2012.6204951

[8] G. Zheng and Y. Chen, "A review on vision-based pedestrian detection," in Proc. IEEE Global High Tech Congress on Electronics (GHTCE), 2012, pp. 49-54. doi: 10.1109/GHTCE.2012.6490122

[9] D. T. Nguyen, W. Li, and P. O. Ogunbona, "Human detection from images and videos: A survey," Pattern Recognition, vol. 51, pp. 148-175, March 2016. doi: 10.1016/j.patcog.2015.08.027

[10] Y. Mu, S. Yan, Y. Liu, T. Huang, and B. Zhou, "Discriminative local binary patterns for human detection in personal album," in Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2008, pp. 1-8. doi: 10.1109/CVPR.2008.4587800

[11] A. Mohan, C. Papageorgiou, and T. Poggio, "Example-Based Object Detection in Images by Components," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 4, pp. 349-361, Apr. 2001. doi: 10.1109/34.917571

[12] O. Biglari, R. Ahsan, and M. Rahi, "Human Detection Using SURF and SIFT Feature Extraction Methods in Different Color Spaces," Journal of Mathematics and Computer Science, vol. 11, p. 111, 2014.

[13] N. Dalal and B. Triggs, “Histograms of Oriented Gradients for Human Detection,” in Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005, pp. 886–893 vol 1. doi: 10.1109/CVPR.2005.177N

[14] F. Porikli, "Integral histogram: a fast way to extract histograms in Cartesian spaces," in Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005, pp. 829-836 vol. 1. doi: 10.1109/CVPR.2005.188

[15] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001, pp. I-511-I-518 vol.1. doi: 10.1109/CVPR.2001.990517

[16] Y. Said, M. Atri, and R. Tourki, "Human detection based on integral Histograms of Oriented Gradients and SVM," in Proc. International Conference on Communications, Computing and Control Applications (CCCA), 2011, pp. 1-5. doi: 10.1109/CCCA.2011.6031422

[17] X. Wang, T. X. Han, and S. Yan, "An HOG-LBP human detector with partial occlusion handling," in Proc. IEEE 12th International Conference on Computer Vision, 2009, pp. 32-39. doi: 10.1109/ICCV.2009.5459207

[18] C. Conde, D. Moctezuma, I. Martín De Diego, and E. Cabello, "HoGG: Gabor and HoG-based human detection for surveillance in non-controlled environments," Neurocomputing, vol. 100, pp. 19-30, 1/16/ 2013. doi:10.1016/j.neucom.2011.12.037

[19] G.-S. Hong, B.-G. Kim, Y.-S. Hwang, and K.-K. Kwon, "Fast multi-feature pedestrian detection algorithm based on histogram of oriented gradient using discrete wavelet transform," Multimedia Tools and Applications, pp. 1-17, 2015. doi: 10.1007/s11042-015-2455-2

[20] M. Hemmati, M. Biglari-Abhari, S. Berber, and S. Niar, "HOG Feature Extractor Hardware Accelerator for Real-Time Pedestrian Detection," in Proc. 17th Euromicro Conference on Digital System Design (DSD), 2014, pp. 543-550. doi:10.1109/DSD.2014.60

[21] P. Y. Chen, C. C. Huang, C. Y. Lien, and Y. H. Tsai, "An Efficient Hardware Implementation of HOG Feature Extraction for Human Detection," IEEE Transactions on Intelligent Transportation Systems, vol. 15, no. 2, pp. 656-662, Apr. 2014. doi: 10.1109/TITS.2013.2284666

[22] I. Fodor, "A Survey of Dimension Reduction Techniques," technical report, Lawrence Livermore National Lab., CA (US), 2002.

[23] I. Jolliffe, "Principal component analysis," Wiley StatsRef: Statistics Reference Online, 2002. doi: 10.1007/b98835

[24] V. N. Vapnik, "The nature of statistical learning theory", Springer-Verlag New York, Inc., 1995. doi: 10.1007/978-1-4757-3264-1

[25] S. Abe, "Support Vector Machines for Pattern Classification" Advances in Pattern Recognition, Springer-Verlag New York, Inc., 2005. doi: 10.1007/978-1-84996-098-4

[26] C.-C. Hsu, and C.-W. Lin, "A Practical Guide to Support Vector Classification," Department of Computer Science, National Taiwan University, Taipei 106, Taiwan 2003.

[27] G. Bradski, "The OpenCV Library," Doctor Dobb's Journal of Software Tool, vol. 25 (11), pp. 120-126

[28] A. V. S. Vempati, A. Zisserman and C. V. Jawahar, "Generalized RBF feature maps for efficient detection," in Proc. British Machine Vision Conference, pp. 2.1-2.11, 2010. doi:10.5244/C.24.2