Pedestrian Detection using HOG-based Block Selection

Minsung Kang and Young Chul Lim
IT Convergence Research Division, DGIST, Daegu, Hyeonpung Myeon, Korea

Keywords: Pedestrian Detection, Intelligent Vehicle, HOG, Camera, Computer Vision.

Abstract: Recently, pedestrian detection methods have been popularly used in the field of intelligent vehicles. In most previous works, the Histogram of Oriented Gradients (HOG) is used to extract features for pedestrian detection. However HOG is difficult to use in the real-time operating system of an intelligent vehicle. In this paper, we proposed a pedestrian detection method using a HOG-based block selection. First, we analyse the HOG block and select the parts of the block with a high hit rate. We then use only 20% of the total HOG blocks for the pedestrian feature. The proposed method is 5 times faster than methods using the entire feature, while performance remains almost the same.

1 INTRODUCTION

Pedestrian detection methods have been used recently for intelligent vehicle, intelligent robot and video security applications. Pedestrian detection is the technical methodology for finding the position of pedestrians from a camera image. In the detection process, first, a feature is extracted for pedestrian classification. Then a pedestrian is detected using a feature from a searched image. The performance and computation speed are typically different when using features that are extracted with various shapes.

The Histogram of Oriented Gradients (HOG) is one of the well-known features used for pedestrian detection. The HOG feature is robust to variations of illumination. However, The HOG feature needs a high amount of image processing because the dimensions of the feature are high. Hence, pedestrian detection based on HOG is impractical for the real time operation of vehicles. In an intelligent vehicle, real time operation is important because reaction time is directly connected to the safety of the driver and pedestrian.

As a result, many pedestrian detection methods based on the HOG feature are being researched with the goal of reducing computation time. Many of these existing methods change the process of searching the image to reduce computation time. Other methods use a GPU to improve computation speed but these need an NVIDIA graphic card. Such methods use the high dimensions of HOG and improve computation speed in post-processing. However these methods do not solve the fundamental problem.

Accordingly, this paper proposes a pedestrian detection method using HOG-based block selection. The structure of this paper is as follows: Chapter 2 introduces related works about pedestrian detection. Chapter 3 describes the proposed algorithm. Chapter 4 deals with the verification of the proposed algorithm through experiments. And finally Chapter 5 presents conclusions.

2 RELATED WORKS

Pedestrian detection methods involve the extraction of features from a pedestrian dataset and a training feature using a classifier such as SVM. Then the feature is used to detect the pedestrian in a whole camera image. Existing methods typically change the process of searching the image to reduce computation time. As shown in Figure 1, the sliding window method makes an image pyramid from the original image in order to search the image.

The computation speed of the sliding window method is very slow because the area being searched is big. The classifier of the sliding window method is fixed. Hence, to address this issue, as shown in Figure 2, an alternative method makes various sizes of classifier to improve the search speed. But this method is hard to use because HOG is an invariant feature. For this reason, a hybrid method has been proposed, as shown in Figure 3. The computational

Kang M. and Lim Y..
Pedestrian Detection using HOG-based Block Selection.
DOI: 10.5220/0005147607830787
In *Proceedings of the 11th International Conference on Informatics in Control, Automation and Robotics* (IVC&ITS-2014), pages 783-787
ISBN: 978-989-758-040-6

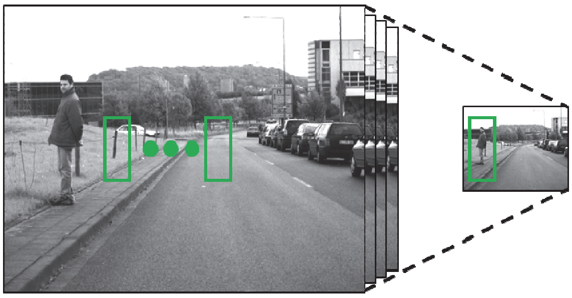

Figure 1: Dense image pyramid.

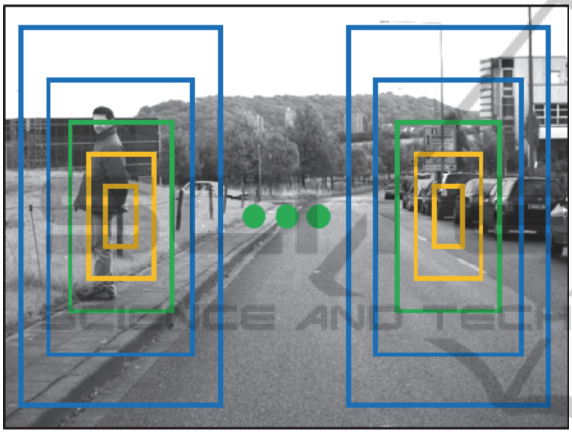

Figure 2: Classifier pyramid.

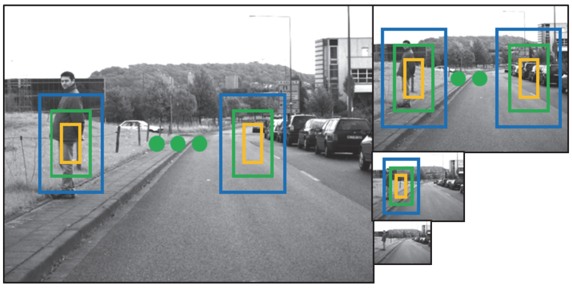

Figure 3: Hybrid approach.

time is improved but this method results in the loss of data and its structure is highly complex.

Another method uses a GPU. That method uses a parallel calculated block histogram for improving computation time. However, it depends on having equipment available because it requires an NVIDIA graphic card. The method uses all of the blocks of the HOG for pedestrian detection.

The method proposed in this paper, in contrast, does not use all the blocks of the HOG. As shown in Figure 4, this new method selects only a part of the blocks. Figure 4-(a) shows the extracted feature based on HOG from a general pedestrian database. The figure of the pedestrian does not fill the entire

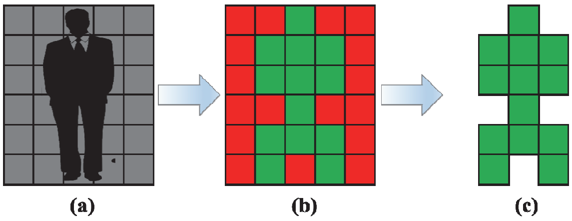

Figure 4: Proposed algorithm; (a) HOG feature, (b) Analysis feature (green: pedestrian, red: another), (c) Select feature.

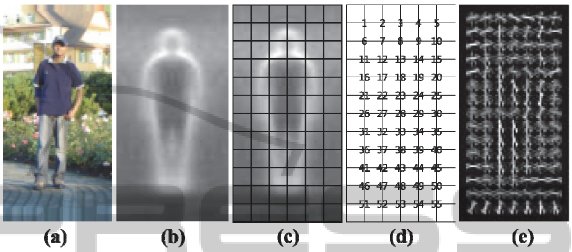

Figure 5: HOG (a) original image, (b) oriented gradients, (c) divided cells, (d) divided blocks, (e) visualization of HOG.

feature. Accordingly, the proposed method performs an analysis to determine which part of the block contains a high probability of the pedestrian existing, as in Figure 4-(b). Then the proposed method selects that part of the blocks, such as the green blocks of Figure 4-(c).

As a result, the proposed method uses only 20% of all HOG blocks. Hence, the overall computational time is 5 times faster than methods using all the blocks. And the performance is almost the same.

3 PROPOSED ALGORITHM

The HOG method of pedestrian detection divides an image into units of cells after calculating orientation gradients, such as in Figure 5-(b) and then calculates a histogram, as shown in Figure 5-(e). The oriented gradients is divided cell such as Figure 5-(c) and mixed 4 cell to block. Then the block is normalized by the L1-norm or L2-norm method. In a 48 x 96 image, the feature is 1980 because all block number is 55 and each block is calculated 9 histogram.

However in Figure 5-(c), the non-pedestrian blocks which correspond to the background or other objects (for example, blocks 1, 5, 51 and 55) can be used for extracting the feature. Blocks which do not contain a pedestrian block, such as those for a building, tree and road, are not required for

55.31	70.31	74.17	70.84	55.72
59.83	73.68	75.61	73.47	60.29
67.73	78.89	81.41	78.46	67.77
70.18	76.52	73.37	76.23	69.91
67.35	72.44	66.90	72.03	67.35
65.49	69.03	64.57	68.14	65.34
63.30	72.08	72.69	71.72	63.22
58.81	71.37	74.01	71.90	59.16
56.91	71.30	73.75	71.65	57.17
61.41	73.33	72.33	72.29	60.40
63.49	75.98	78.12	74.89	61.72

Figure 6: Hit rate per block (green: high rank 40%, red: low rank 60%).

55.31	70.31	74.17	70.84	55.72
59.83	73.68	75.61	73.47	60.29
67.73	78.89	81.41	78.46	67.77
70.18	76.52	73.37	76.23	69.91
67.35	72.44	66.90	72.03	67.35
65.49	69.03	64.57	68.14	65.34
63.30	72.08	72.69	71.72	63.22
58.81	71.37	74.01	71.90	59.16
56.91	71.30	73.75	71.65	57.17
61.41	73.33	72.33	72.29	60.40
63.49	75.98	78.12	74.89	61.72

Figure 7: Hit rate per block (green: high rank 30%, red: low rank 70%).

55.31 59.83	70.31 73.68	74.17 75.61	70.84 73.47	55.72 60.29
67.73	78.89	81.41	78.46	67.77
70.18	76.52	73.37	76.23	69.91
67.35	72.44	66.90	72.03	67.35
65.49 63.30	69.03 72.08	64.57 72.69	68.14 71.72	65.34 63.22
58.81	71.37	74.01	71.90	59.16
56.91	71.30	73.75	71.65	57.17
61.41	73.33	72.33	72.29	60.40
63.49	75.98	78.12	74.89	61.72

Figure 8: Hit rate per block (green: high rank 20%, red: low rank 80%).

information processing because the environment exists in various other shapes, unrelated to pedestrians.

Accordingly, in this paper, rather than using all 1980 features, only the relevant features are used. As shown in Figure 6, all of the blocks are each trained and tested using SVM, and a hit rate is calculated for each block. Then blocks with a high hit rate are selected for the feature after the unnecessary data, having a low hit rate, are removed. In Figure 6, the green areas are good blocks with a high hit rate. And these blocks are similar in shape to a human head, shoulder, waist and leg.

In Figure 6, the green block is high rank 40% of hit rate such as human. And the red block is low rank 60% of hit rate such as environment. In Figure 8, the high rank 20% of hit rate similar to pedestrian head, shoulder, waist and leg. As shown in Figure 8, in this paper we propose a block selection method. The proposed method is to analyse each block to determine their relationship to a pedestrian figure, then to remove unnecessary blocks such as those corresponding to the environment, and select good blocks which correspond to a human head, shoulder, waist and leg.

4 EXPERIMENTS

This study used the Daimler pedestrian dataset for pedestrian detection experiments. The Daimler pedestrian dataset includes 48 x 96 images that were acquired while driving on roads. The training dataset is composed of 52,112 positive images and 32,465 negative images. The test dataset is composed of 25,608 positive images and 16,235 negative images. In Figure 9, the test result is the detection rate per block number. All 55 of block numbers 1 to 55 were trained and tested.

In Figure 9, the performance is almost the same as compared with methods using all the blocks. But the detection rate rapidly decreases when the method uses less than 11 blocks. The detection rate using all the blocks is 94.48% and the method using 11 blocks is 93.19%. This means the detection rate decreases 1.3% when going from using 55 blocks for detection to using 11 blocks. This is slight change in performance considering the block number has been decreased by 80%.

As shown in Figures 10, 11 and 12, we provide an analysis of the classifier comparing the method using all blocks and the proposed method. In Figure 10, the classifier result for the method using all blocks is a false negative in a complex background, but the proposed method is a true positive. This is because the proposed method used just a part of the

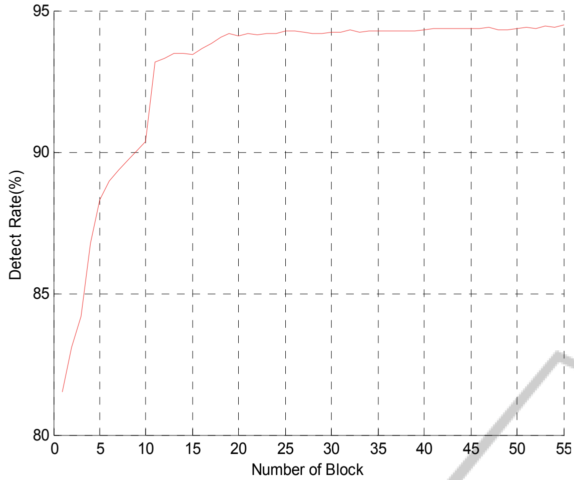

Figure 9: Detection rate per block number.

Figure 10: Proposed method: true positive; method using all blocks: false negative.

Figure 11: Proposed method: false negative; method using all blocks: true positive.

Figure 12: Proposed method: false negative; method using all blocks: false negative.

HOG blocks. However in Figure 11, the classifier result is the opposite of the image in Figure 10, with a simple background. In Figure 12, the method using all blocks and the proposed method are false negatives due to overlapping of the pedestrians.

Figure 13: Proposed method: true negative; method using all blocks: false positive.

Figure 14: Propose method: false positive; method using all blocks: true negative.

Figure 15: Proposed method: false positive; method using all blocks: false positive.

Also we analysed the classifier with negative images. As shown in Figures 13, 14 and 15, the classifier results are similar to the positive images. The performance of the proposed method is better than the method using all blocks in a complex background. And the method using all blocks and the proposed method are false positive for large objects such as a vehicle.

We have experimentally tested the proposed method in pedestrian detection and found the performance of pedestrian detection to be good with persons in a complex background, such as Figure 16. However vertical or large objects produce a false positive, as in Figure 17. And overlapping persons were missed in detection, such as in Figure 18.

5 CONCLUSIONS

In this paper, the proposed algorithm effectively reduced unnecessary features by feature analysis and improved computational speed. Existing methods to

Figure 16: True Positive of pedestrian detection using the proposed method (black box: detection result).

Figure 17: False Positive of pedestrian detection using the proposed method (black box: detection result).

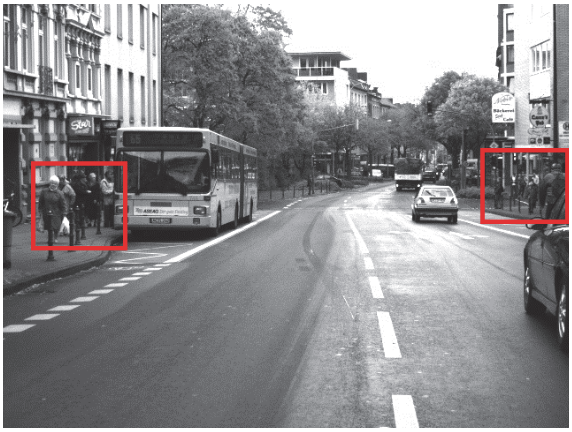

Figure 18: False negative of detection using the proposed method (red box: missed detection).

improve computational speed require either high complexity or special equipment. But the proposed algorithm greatly improved computation speed by using a simple method. The experimental results show that the computation speed of the proposed method is 5 times faster than the method using all HOG blocks. However the performance is almost the same. The classifier result of the proposed method is better than the method using all blocks for complex backgrounds.

Also, pedestrian detection using the proposed method can provide a real time operating system for intelligent vehicles. In future works, we will use a more advanced probability method for partial block analysis. These studies will be helpful for developing a system for pedestrian detection for intelligent vehicles.

ACKNOWLEDGEMENTS

This work was supported by the DGIST R&D Program of the Ministry of Education, Science and Technology of Korea.

REFERENCES

Geronimo, D., Lopez, A., M., Sappa, A., D., 2010. Survey of pedestrian detection for advanced driver assistance systems. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Dalal, N., Triggs, B., 2005. Histogram of Oriented Gradients for Human Detection. In *Proceeding of IEEE Conference Computer Vision and Pattern Recognition*.

Dollar, P., Wojek, C., Schiele, B., Perona, P., 2012. Pedestrian detection: an evaluation of the state of the art. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Dollar, P., Belongie, S., Perona, P., 2010. The Fastest Pedestrian Detector in the West. In *Proceeding of Conference British Machine Vision*.

Viola, P., Jones, M., 2001. Fast multi-view face detection. In *Proceeding of IEEE Conference Computer Vision and Pattern Recognition*.

Munder, S., Schnorr, C., Gavrila, D., 2008. Pedestrian detection and tracking using a mixture of view-based shape-texture models. In *IEEE Transactions on Intelligent Transportation Systems*.

Prisacariu, V., A., Reid, I., 2009. fastHOG – a real-time GPU implementation of HOG. *Technical Report*.