International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:5, No:12, 2011

Human Detection using Projected Edge Feature

Jaedo Kim, Youngjoon Han, and Hernsoo Hahn

***Abstract*—**The purpose of this paper is to detect human in images. This paper proposes a method for extracting human body feature descriptors consisting of projected edge component series. The feature descriptor can express appearances and shapes of human with local and global distribution of edges. Our method evaluated with a linear SVM classifier on Daimler-Chrysler pedestrian dataset, and test with various sub-region size. The result shows that the accuracy level of proposed method similar to Histogram of Oriented Gradients(HOG) feature descriptor and feature extraction process is simple and faster than existing methods.

***Keywords*—**Human detection, Projected edge descriptor, Linear SVM, Local appearance feature

I. Introduction

HUMAN detection in images is an important part of driver assistance, surveillance system. In the indoor condition, face detection and human body detection successfully used in indoor conditions because illumination is uniform. But human detection has many problems in outdoor conditions because of various clothing and illumination conditions.

Recently, Dalal & Triggs[1] presented a human detection algorithm. Their method uses a dense grid of Histograms of Oriented Gradients(HoG), Tomoki et al.[2] presented a multiple gradient orientation based feature descriptor named Co-occurrence Histogram of Oriented Gradients(CoHOG). Shashua et al.[3] used body parts detectors using SIFT and Mikolajczyk et al.[4] also used jointed SIFT with an SVM classifier. These approachs have slow feature extraction time, so it is hard to real implementation.

Under the motion and moving information, Viola et al.[5] extended their Haar-like wavelets to handle space-time information for moving-human detection. And Dalal & Triggs[6] extended their HOG method to motion feature descriptors. But, it is hard to separate relative motion and movement of object and background in fast camera moving condition.

We propose a simple human detection method. Our approach is based on appearance and shape of human. Projected edge component of local region used as a principle direction.

The rest of the paper is organized as follows: Section 2 explains the overview of our method in brief; Section 3 describes projected edge based our feature descriptor; Section 4 shows experimental results on Daimler-Chrysler dataset; The final section is the conclusion.

Jaedo Kim is with the School of Information Technology, Soongsil University, 1-1, Sangdo-5Dong, Dongjak-Gu, Seoul, 156-743, KOREA (corresponding author to provide phone: 02-821-2050; fax: 02-826-8937; e-mail: duckjd@ssu.ac.kr).
Youngjoon Han and Hernsoo Hahn are with the School of Information Technology, Soongsil University (e-mail: young, hahn@ssu.ac.kr).

II. Overview Of The Method

This section gives an overview of our method per human detection, which is summarized in Fig. 1. The method is based on edge component distribution of local region in a human region. The basic idea is that local appearance and shape of objects can be characterized well.

Firstly, compute edge component using Sobel operator which compute an approximation of the gradient of the image intensity function. The detection window is divided into small spatial regions("blocks"). Edge components in a spatial block are projected to x and y axis. The combined entries of x and y axis form the block vector. For better invariance to illumination, shadowing, etc., block normalization is useful before using them. Feature descriptor is formed by combining block vectors. A conventional SVM based window classifier gives human detection result.

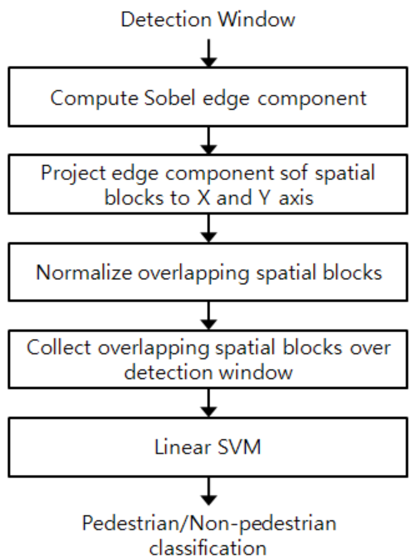
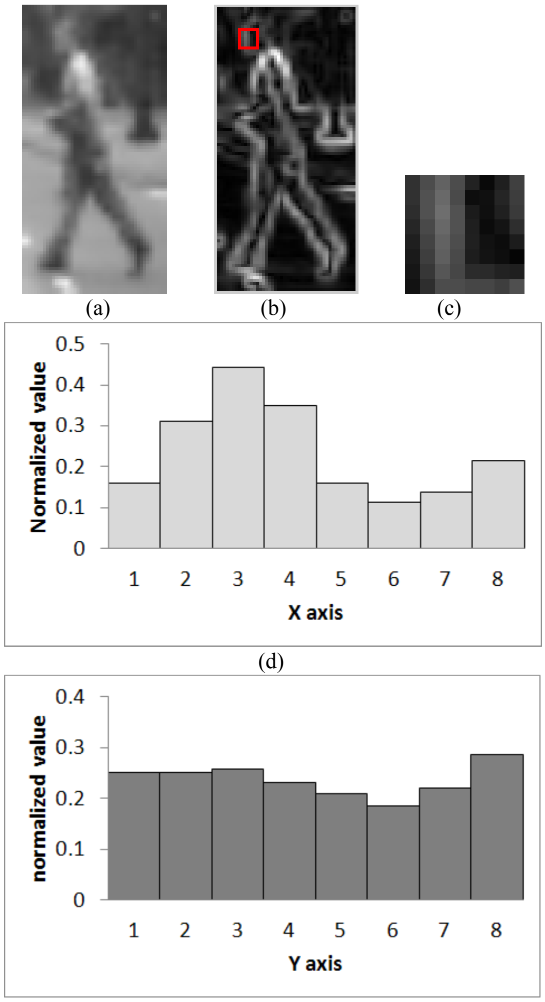


Fig. 1 Block diagram of feature extraction and detection.

III. Projected Edge Based Feature Descriptor

A. Edge Projection of Block

The appearance and shape of human are good feature for recognition. To present the appearance and shape of human, we extract edge components using Sobel operator. And denote sub-region of detection window as a block, project accumulated edge components of a block to X axis and Y axis. Block vector is generated to serialize two projections. Block vector is able to separately represent certain edge distribution and principle edge direction of block. Fig 2 shows an example of edge projection.



(e)

Fig. 2 Edge projection. (a) input image. (b) Sobel edge image of detection window. (c) edge image of block. (d) normalized projection to X axis of block. (e) normalized projection to X axis of block.

B. Overlapping Spatial Blocks

Blocks are overlapping over detection window according block stride to reduce geometrical error. If blocks are located like chessboard, geometrical error is increased because of various pedestrian pose. We use two kinds of block size and block stride: (a) 16x16 block size and 8x8 block stride, (b) 8x8 block size and 4x4 block stride. Fig 3 shows the concept of overlapping block over detection window.

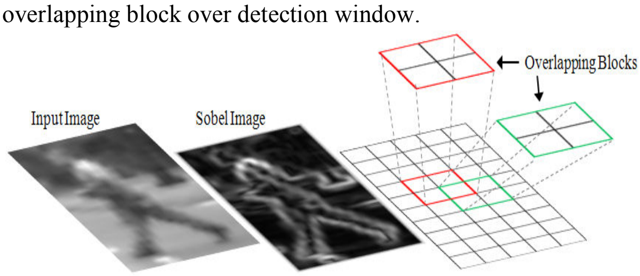


Fig. 3 Overlapping spatial blocks

C. Block Normalization

Block normalization is needed to uniformalize dada before using them. There are many schemes to normalize the data. Let $\boldsymbol{v}$ be the unnormalized descriptor vector, $\|v\|_k$ be its k-norm for k=1, 2, and $\boldsymbol{\varepsilon}$ be a small constant. The schemes are:

(a) L1-norm, $v \to v/(\|v\|_1)+\varepsilon$

(b) L1-sqrt, $v \to \sqrt{v/(\|v\|_1+\varepsilon)}$

(c) L2-norm, $v \to v/\sqrt{\|v\|_2^2+\varepsilon^2}$

We used L2-Norm and $\varepsilon = 0.01$.

IV. Experiment Results

A. Dataset

We conduct experiments on Daimler-Chrysler Dataset[7] (http://www.science.uva.nl/research/isla/downloads/pedestrians/). Daimler-Chrysler dataset has been widely used as a benchmark dataset in performance comparison, containing 14400 training and 9600 test images of human. It contains various pedestrian pose(front, back, and side views) with different illuminations and many common large image(640x480 size). Fig 4 shows some samples of training set.

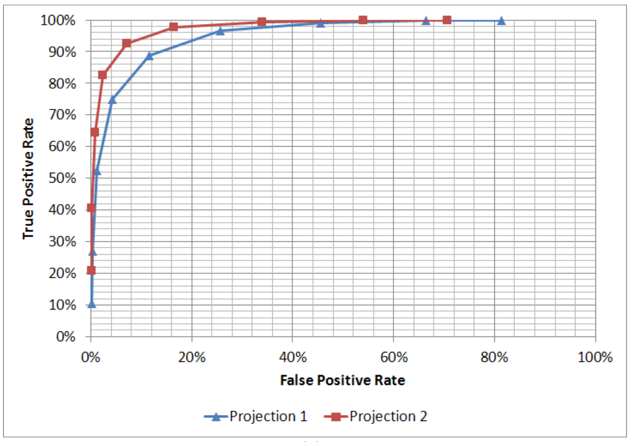
Fig. 4 Sample images of Daimler-Chrysler dataset.

B. Performance Comparison

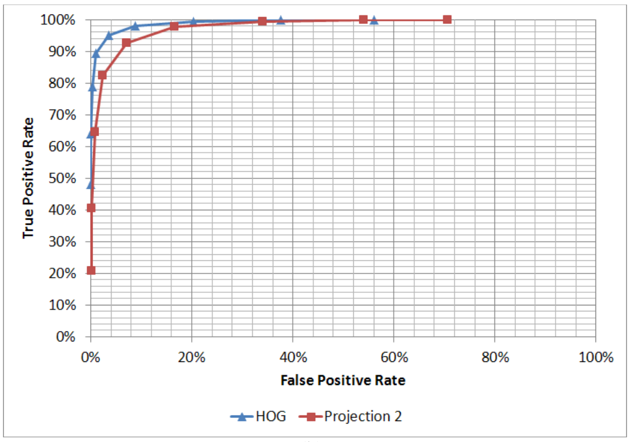
By default we use a soft (C=0.01) linear SVM trained with SVMLight[8]. We evaluate performance and find optimal condition of our method.

The Performances of following conditions are compared by ROC curve on Daimler-Chrysler Dataset.

(a) HOG[1], 64x128 window size, 8x8 cell size, 16x16 block size, 8x8 block stride
(b) Projection 1, 64x128 window size, 16x16 block size, 8x8 block stride, L2-norm
(c) Projection 2, 64x128 window size, 8x8 block size, 4x4 block stride, L2-norm



(a)



(b)

Fig. 5 Performance comparison. (a) performance comparison with Projection 1 and Projection 2 condition of proposed descriptor. (b) performance comparison with Projection 2 and HOG feature descriptor

In Fig 5, the FPR of Projection 1 condition is 14% at 90% TPR, and the FPR of Projection 2 condition is 6% at 90% TPR. Projection 2 condition shows more effective performance than projection 1 condition for human detection.

The FPR of HOG feature is 2% at 90% TPR. The performance of Projection 2 condition is similar to HOG descriptor. And our method reduces the feature extraction time by half compared with HOG feature.

Fig 6 shows some human detection examples of large image.

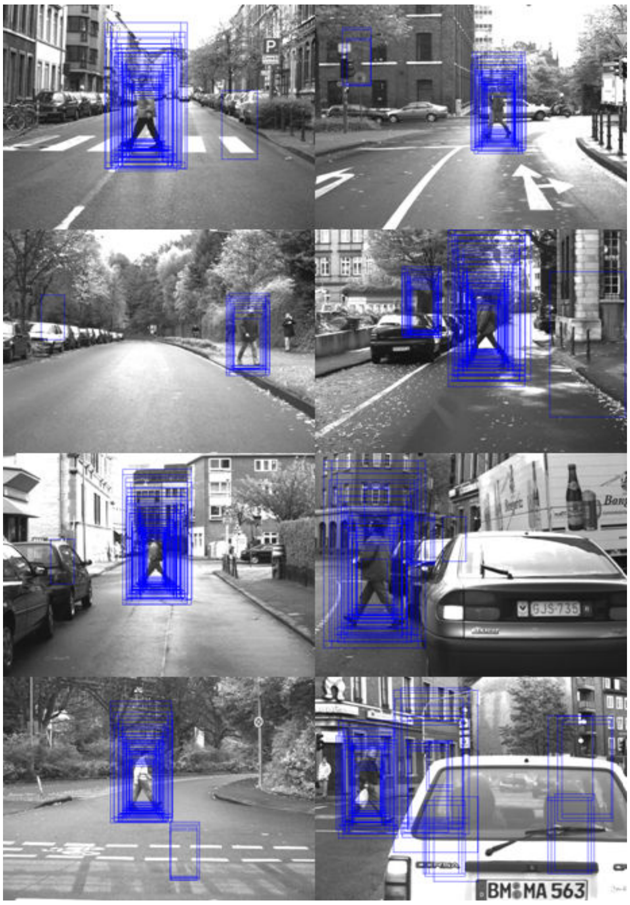


Fig. 6 Detection result images.

V. Conclusion

In this paper, we proposed a feature descriptor, projected edge component series, for human detection. Our feature descriptor consists of X axis and Y axis projection array. The accuracy level of proposed method similar to Histogram of Oriented Gradients(HOG) feature descriptor and feature extraction process is simple and faster than existing methods. Future work involves applying the proposed feature descriptor to other applications.

Acknowledgment

This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency) (NIPA-2011-C1090-1121-0010), and This work was supported by the Brain Korea 21 Project in 2011.

References

[1] N. Dalal, B. Triggs, and C. Schmid, “Histograms of oriented gradients for human detection,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, 2005, pp. 886-893.

[2] T. Watanabe, S. Ito, and K. Yokoi, “Co-occurrence histograms of oriented gradients for pedestrian detection,” in *3rd Pacific Rim Symposium on Advances in Image and Video Technology*, 2009, pp. 37-47.

[3] A. Shashua, Y. Gdalyahu, and G. Hayun, "Pedestrian detection for driving assistance systems: single-frame classification and system level performance," in *IEEE Intelligent Vehicles Symposium*, 2004, pp. 1–6.
[4] K. Mikolajczyk, C. Schmid, A. Zisserman, "Human detection based on a probabilistic assembly of robust part detectors," in *Pajdla, T., Matas, J(G.) (eds.) ECCV 2004*. *LNCS*, vol. 3021, pp. 69–82.
[5] P. Viola, M. Jones, and D. Snow, "Detecting pedestrians using patterns of motion and appearance," in *Int. J. Computer Vision*, vol. 63, no. 2, pp. 153-161.
[6] N. Dalal, B. Triggs, and C. Schmid, "Human detection using oriented histograms of flow and appearance," in *Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS*, vol. 3952, pp. 428–441.
[7] S. Munder, D.M. Gavrila, "An experimental study on pedestrian classification," in *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 28, no. 11, 2006, pp. 1863-1868.
[8] T. Joachims, Making large-scale svm learning practical. In B. Schlkopf, C. Burges, and A. Smola, editors, Advances in Kernel Methods - Support Vector Learning. The MIT Press, Cambridge, MA, USA, 1999.