NIGHTTIME PEDESTRIAN DETECTION WITH NEAR INFRARED USING CASCADED CLASSIFIERS

Jianfei Dong, Junfeng Ge, Yupin Luo

Tsinghua Univiersity, Department of Automation, Beijing, 100084, China

email: dongjf@mails.tsinghua.edu.cn

ABSTRACT

This paper presents a novel nighttime pedestrian detection approach only using a near infrared camera, which can be used in a practical driver assistance systems. This method can be divided into three steps: selection step, preprocess step and recognition step. Firstly, objects in the video are separated with an adaptive dual thresholds segmentation method in the selection step; Secondly, most of non-pedestrians are discarded with some constraints in the preprocess step; Finally, in the recognition step a cascaded classifiers with Histograms of Oriented Gradients and Adaptive Boosting Algorithm are introduced. Experiments on video sequences show that the proposed pedestrian detection approach has a high detection rate as well as a very low false alarm rate and run in real-time.

Index Terms— Pedestrian detection, cascaded classifiers

1. INTRODUCTION

In the last few years, there has been more emphasis on detection pedestrians at night [1] [2] [3]. Detecting pedestrians using a camera in a clutter background is a very challenging problem due to their appearance, and pose variability.

In this paper, we introduce a complete nighttime pedestrian detection system with a normal camera. At first we select the candidate regions in the video with an adaptive dual thresholds segmentation method. Then we use four constraints to reject most of non-pedestrians for saving computation and improving the performance of the system. At last the candidates are sent into the classifier to recognize. We design a cascaded classifier which is made up of two classifiers. Both classifiers are based on Adaptive Boosting Algorithm (AdaBoost) [9] and their weak learners are Classification and Regression Trees (CART). The first classifier uses gray-scale image as feature while the second classifier uses Histograms of Oriented Gradients (HOG) descriptor [5] as feature. After tuning a set of parameters, we achieve high detection rate with very low false alarm rate. The experiment results on videos prove the approach is real-time and can be used in a practical driver assistance systems.

2. SYSTEM DESCRIPTION

Fig. 1 shows the main structure of the pedestrian detection system presented in this paper.

2.1. Dual Adaptive Thresholds Image Segmentation

In the previous systems, an adaptive threshold algorithm is applied in image segmentation [4]. This algorithm performs well in the clear scene. However, when the background is highly complicated or the pedestrian is connected with other objects, this algorithm cannot successfully separate the whole pedestrian. It will dramatically decrease the detection rate of the system. In order to solve the problem, a dual thresholds algorithm for image segmentation is applied instead of single threshold algorithm. This algorithm holds two adaptive thresholds. They can be separately calculated by the following formulas:

$$T_{Low}(i,j) = \frac{\sum_{x=i-N}^{i+N} I(x,j)}{L} \tag{1}$$

$$T_{High}(i,j = T_L(i,j) + \theta \tag{2}$$

where $L = 2N + 1$, L is width of neighborhood, whose value often is semibreadth of pedestrian. θ is a little positive number, and choosing proper value leads to best results.

The dual thresholds image segmentation algorithm is like this:

$$\begin{cases} p(i,j) \in F, \text{ if } I(i,j) > T_{High}(i,j) \text{ or} \\ \text{if } T_{Low}(i,j) \leq I(i,j) \leq T_{High}(i,j) \text{ and } p(i-1,j) \in F, \\ p(i,j) \in B, \text{ if } I(i,j) < T_{Low}(i,j) \text{ or} \\ \text{if } T_{Low}(i,j) \leq I(i,j) \leq T_{High}(i,j) \text{ and } p(i-1,j) \in B. \end{cases} \tag{3}$$

where F represents foreground and B represents background.

The dual thresholds segmentation algorithm performs well on the test images. Figure 2 is some segmentation results.

2.2. Candidates Rejecter

The dual adaptive thresholds segmentation algorithm gives a satisfying result. However, it still produce a large amount of

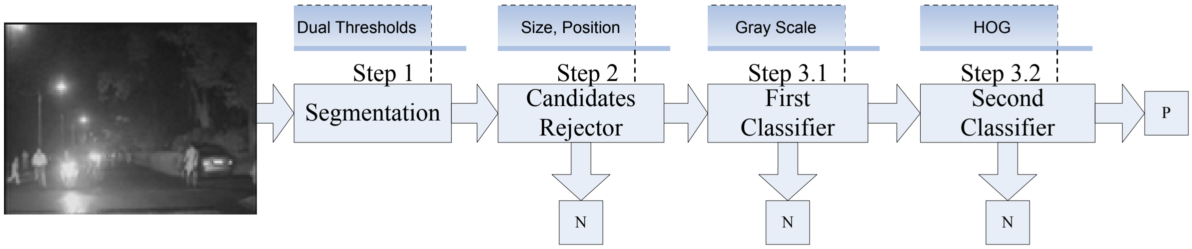

Fig. 1. Structure of System. The input is a frame of video, and the output is a pedestrian's position.

Fig. 2. Dual Adaptive Thresholds segmentation results. The above are original gray-scale images. The below show the segmentation results.

candidate regions, most of which are not pedestrians. Tian [4] gives four constraints: object size and object position in the image coordinates, object width/height ratio and a simple shape constraint. These constraints can be used to design simple filters to reject most of non-pedestrian candidates for saving computation and reducing the false recognition rate of the system.

2.3. Histogram of Oriented Gradients Features

Dalal and Triggs [5] have shown experimentally that HOG as feature significantly outperform other features for human detection. The basic idea is that local object appearance and shape can often be characterized rather well by the distribution of local intensity gradients or edge directions. Their method is mainly based on evaluating well-normalized local histograms of image gradients orientations in a dense grid. In this paper, HOG descriptor is computed as the following three steps:

- Compute gradients of the detection window
- Build histogram of gradient directions for each cell
- Normalize histograms for each block locally

Suard and Broggi [6] have discussed the optimal set of parameters for the HOG descriptors. In this paper, we use the following set of parameters:

- cell: 4×4 pixels
- block: 2×2 cells
- overlap of blocks: 1
- number of bins: 9
- normalization scheme: $\frac{V}{\sqrt{\|V\|_2^2+\epsilon^2}}$

2.4. AdaBoost Learning

Support Vector Machine (SVM) is very popular in human detection [4] [7] [8]. However, AdaBoost Learning method [9] is contracting more and more attention. AdaBoost is a general method for improving the accuracy of any given learning algorithm. Consider the following weak learner:

$$r(x) = \begin{cases} +1 \\ -1 \end{cases} \tag{4}$$

where x is feature of sample.

AdaBoost finally defines a strong classifier:

$$R(x) = \text{sign}\left(\sum_{t=1}^{T} \alpha_t r_t(x)\right) \tag{5}$$

In fact, the output of AdaBoost is weighted vote of the weak learners. The key problem of AdaBoost is how to decide the weight α_t for each weak learner $r_t(x)$.

In this paper we use AdaBoost as pedestrian classifier. The weak learner is classification and regression tree(CART).

2.5. Design of Cascaded Classifiers

The AdaBoost algorithm with HOG feature can perform well in lots of database, such as MIT pedestrian database [5]. However, in our condition, the method has a high false recognition rate for nighttime videos captured by a normal camera, which is not satisfied.

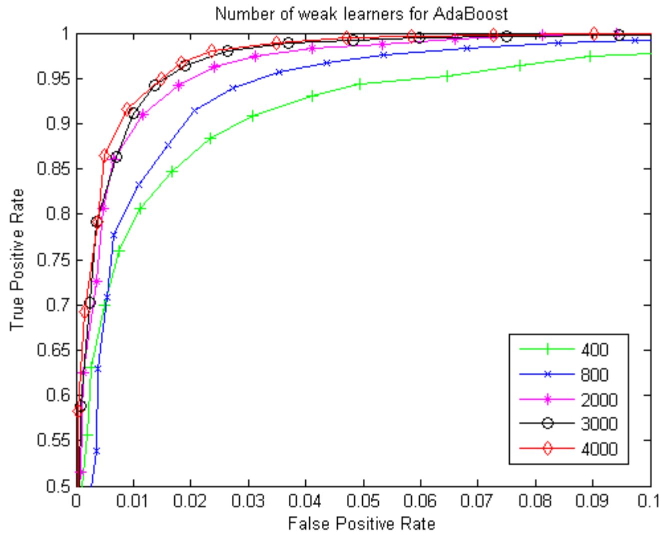

Fig. 3. This figure shows the ROC Curve of the AdaBoost when the number of weak learners varies.

To decrease the false recognition rate, we design the cascaded classifiers each of which is based on AdaBoost. There are three problems we should consider carefully:

- How many weak learners each classifier should have.
- What features each classifier should use.
- How large the threshold of the classifier should be.

To solve these problems, we collect a training set which are 2944 normal images with a size of 24×60: 1598 pedestrians and 1346 non-pedestrians. And we collect a testing set which are 4709 normal images with a size of 24×60: 2138 pedestrians and 2571 non-pedestrians.

(1) To choose the proper number of weak learners, we test five sizes: 400, 800, 2000, 3000 and 4000 weak learners for AdaBoost with HOG feature on the data set. Fig. 3 presents the results of ROC Curve. It clearly shows that when the number of weak learners increase, the performance of AdaBoost is better. However, when the number is 3000 or 4000, the ROC Curves nearly make no difference. Therefore, in our approach we use 4000 weaker learners for AdaBoost.

(2) The feature is important for classification. To choose the proper feature for the cascaded classifiers, three cascaded classifiers are designed. They are all based on AdaBoost with 4000 weak learners and the only difference is the feature: grayscale with HOG, binary image with HOG, or gradient image with HOG. The experiment condition is the same as the above. Fig. 4 presents the result of ROC Curve. The result shows that generally the performance of cascaded classifiers is better than the single classifier. In addition, the ROC Curve of grayscale with HOG is nearly the same as that of gradient with HOG. Because grayscale feature is easier to extract from image than gradient feature, in our approach we designed the first classifier with grayscale feature and the second classifier with HOG feature.

Table 1. The results of the cascaded classifiers

Threshold	Pedestrian detection	False detection	True rate	False rate
0.004	2079	24	97.240%	0.933%
0.011	2032	12	95.042%	0.467%
0.018	1925	6	90.037%	0.233%

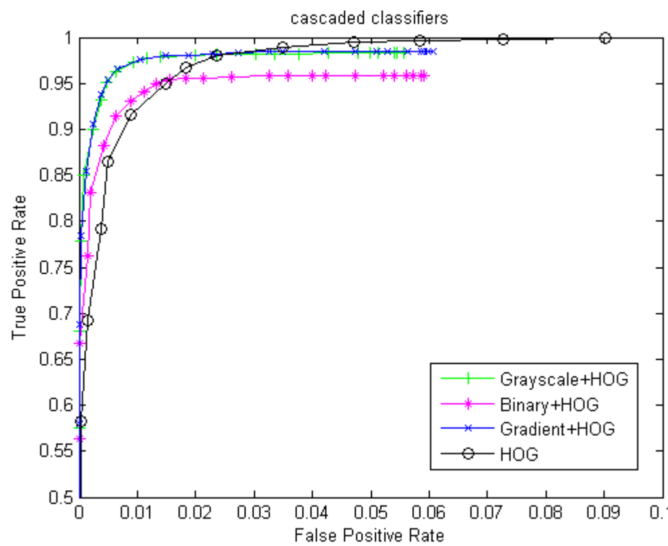

Fig. 4. This figure shows the Roc Curve of various cascaded classifiers.

(3) To design a practical system, we also have to choose the cascaded classifiers' proper threshold θ. The ROC Curve shows that when θ is higher the pedestrian detection rate as well as the false detection rate is higher. Therefore, we should decide the proper threshold in a practical system. In our system, experiment results show that $\theta = 0.01$ is the best threshold. when $\theta = 0.01$, the pedestrian detection rate is 95% and the false detection rate is 0.5%. Table 1 presents the results on samples. The training set is 2944 samples, and the testing set is 4709 samples. Among testing samples, there are 2138 pedestrians and 2571 non-pedestrians.The results also show that the proposed method in this paper performs better than the approach in Suard and Broggi' s paper [6] on samples.

3. EXPERIMENTAL RESULTS

To evaluate the overall performance of the presented system, we collect a training set, which are 7114 normal images with a size of 24×60. Among these train samples, 3707 samples are pedestrian while 3407 samples are non-pedestrian. Fig .5 shows some examples of images used for training.

The HOG feature is computed following section II-C and

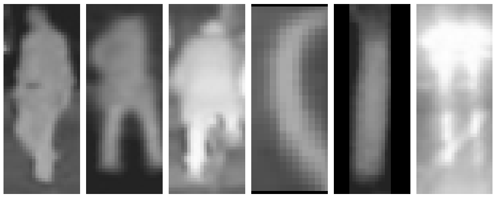

Fig. 5. This figure shows some training samples. (a), (b) and (c) are pedestrians while (d), (e), (f) are non-pedestrians.

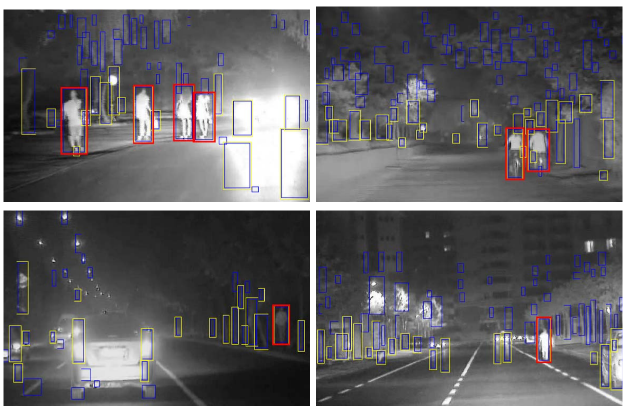

Fig. 6. results of pedestrian detection. The blue rectangle represents candidate rejected. The yellow rectangle represents the candidate passed into classifiers. The red rectangle represents pedestrian detection.

4000 weak learners (CART) are used in AdaBoost as well as the threshold 0.01. We use three videos to test the system. The three videos are from different scenes. The videos are 13m36s (24465 frames). There are 64 pedestrians and 58 of them are detected successfully. The number of all false alarms is 181. So the detection ratio:

$$R_1 = \frac{pedestriandetected}{all \;\; pedestrians} = \frac{58}{64} = 90.63\% \tag{6}$$

$$R_2 = \frac{falsealarms}{allframes} = \frac{181}{24465} = 0.73\% \tag{7}$$

The 0.73% false alarms ratio is very low, which means that 1 false alarm in near 200 frames. If the tracking module is intorduced, the false alarms can be decreased more. Figure .6 shows some examples of results.

Our nighttime pedestrian detection system with a normal camera performs near perfect on the test videos and it runs 15fps on a PC with CPU P4 3.0G and RAM 1G. Therefore, the system in this paper is a true real-time practical pedestrian detection system.

4. CONCLUSIONS

In this paper we have presented a new nighttime pedestrian detection system. It can detect various pedestrians in videos captured by a near infrared camera. The system has three parts. In the image segmentation part, a adaptive dual thresholds segmentation is used to improve the results of segmentation. In the candidate rejection part, four constraints are used to reject most of non-pedestrians to improve the performance of the system. In the classification part, a cascaded classifiers based on AdaBoost are designed. Experiment results have shown that the method has a high detection ratio and a very low false alarm ratio. In addition it runs in real-time on a normal PC. These features make it be used in a practical driver assistance systems.

5. REFERENCES

[1] F. Xu and K. Fujimura, ”Pedestrian Detection and Tracking with Night Vision,” *In Proceedings of IEEE Intelligent Vehicles Symposium*, 2002.

[2] H. Nanda and L. Davis, ”Probabilistic Template Based Pedestrian Detection in Infrared Videos,” *In Proceedings of IEEE Intelligent Vehicles Symposium*, 2002.

[3] M. Szarvas, A. Yoshizawa, M. Yamamoto and J. Ogata, ”Pedestrian Detection with Convolutional Neural Networks,” *In Proceedings of the IEEE Symposium on Intelligent Vehicles*, 2005.

[4] Q. M. Tian, H. Sun, Y. P. Luo D. C. Hu, ”Nighttime Pedestrian Detection with a Normal Camera Using a SVM Classifier,” *In Proceedings of Second International Symposium on Neural Networks*, Vol. 2, 2005, pp189-94.

[5] N. Dalal and B. Triggs ”Histograms of Oriented Gradients for Human Detection,” *International Conference on Computer Vision and Pattern Recognition*, Vol. 2, 2005, pp886-893.

[6] F. Suard, A. Rakotomamonjy, A. Bensrhair, A. Broggi, ”Pedestrian Detection using Infrared images and Histograms of Oriented Gradients,” *In Proceedings of IEEE Intelligent Vehicles Symposium*, 2006.

[7] C. Bellotti, F. Bellotti, A. De Gloria, L. Andreone and M. Mariani, ”Developing a Near infrared Based Night Vision System,” *In Proceedings of IEEE Intelligent Vehicles Symposium*, 2004.

[8] G. Grubb, A. Zelinsky, L. Nilsson and M. Rilbe, ”3D vision sensing for improved pedestrian safety,” *In Proceedings of IEEE intelligent Vehicles Symposium*, 2004.

[9] Y. Freund and R. E. Schapiere, ”Experiments with a new boosting algorithm,” *In Machine Learning: Proceedings of The Thirteenth International Conference*, 1996, pp148-156.