Implementation of Fast Pedestrian Detection Algorithm Based on CENTRIST Feature

First A. Author, *Fellow, IEEE*, Second B. Author, and Third C. Author, Jr., *Member, IEEE*

# INTRODUCTION

I

n the research, pedestrian detection is based on traditional feature detection method and SVM classifier. Compared with deep learning, traditional feature detection method has higher universality. The traditional feature detection method is mainly divided into two parts: one is the description algorithm of pedestrian features, and the other is the classifier. Among them, the commonly used feature detection algorithms mainly include HOG feature [1], ChnFtrs feature [2], CENTRIST feature [3] and LBP feature, etc. Commonly used classifiers include SVM [4], AdaBoost [5], K nearest neighbor classifier, Neural network.

HOG+SVM is the most classic detection method in pedestrian recognition, but HOG feature detection needs to calculate the gradient of each pixel and count the gradient direction histogram to establish the pedestrian contour feature. The calculation is complicated, and the speed is slow. For CENTRIST feature, the descriptor encodes the contour feature by Comparing the size of adjacent pixels can also retain the contour feature of the target well. It does not need to calculate the specific gradient or know the specific gray value, so it has a higher calculation rate.

# METHOD

## Sobel image edge detection

Whether it is pedestrian recognition based on HOG feature or CENTRIST feature, it is aimed at the contour information of the target, so before extracting the CENTRIST feature value of the image, we must first perform edge detection on the image. Edge detection eliminates local texture features and irrelevant information while retains the most important structural features of the image. It is an indispensable primary link in pedestrian detection.

The Sobel operator has good anti-noise performance and is one of the most used operators in edge detection. Its main principle is to find the brightness change in each direction to confirm the set of points with the most obvious brightness change, that is, the image edge. Finally set the threshold, if the gray gradient is higher than the threshold, the point is considered as an edge point. Fig. 1 shows the Sobel edge detection results of different kernel sizes. We can know that too small a Sobel kernel will cause insufficient detection, and too large a Sobel kernel will cause detection failure.

|  |  |  |
| --- | --- | --- |
| 图片包含 户外, 照片, 站, 田地  描述已自动生成 | 人在树林里  中度可信度描述已自动生成 | 图片包含 户外, 小孩, 草, 小  描述已自动生成 |
| (a) | (b) | (c) |

Fig. 1. Sobel detection results for different kernel sizes. **(a)** Sobel kernel size is 1. **(b)** Sobel kernel size is 3. **(c)** Sobel kernel size is 5

## CENTRIST feature extraction

In practical applications, a detection window with a size of 108×36 is used to traverse the entire image and extract its CENTRIST feature. In the detection window, it is divided into 9×4 image blocks, and the 2×2 cell is regarded as a super-block, And there will be 24 super blocks according to go through each cell. The CENTRIST feature histogram is formed by taking the super block as the unit image, and the CENTRIST feature dimension of the final detection window is 6144.

Before the image is smoothed by Sobel processing, due to the limitation of the size of the CENTRIST feature detection window, the proportion of pedestrians in the image is like 108×36 but the size may have large deviations. Unequal sizes may cause missed detection and false detection, so we need to zoom the image to ensure the accuracy of the detection. This function is implemented in the program by setting the zoom factor and the function. After the image extracted by Sobel edge, the detection window is used to scan and extract the feature vector, and the CENTRIST feature of each detection window is obtained. The test results are shown in Fig. 2.

|  |  |  |
| --- | --- | --- |
| 狗站在地上的男人和女人  中度可信度描述已自动生成 | 图片包含 户外, 照片, 建筑, 男人  描述已自动生成 | 砖墙边  中度可信度描述已自动生成 |
| (a) | (b) | (c) |

Fig. 2. CT feature detection results. **(a)** original image. **(b)**Sobel image. **(c)**CT image

## Cascade classifier (SVM)

In the last section, the CT image of the target image is obtained after CENTRIST feature extraction, and then the CT image obtained needs to be classified. In this paper, linear SVM classifier and HIK SVM classifier are used for two detections, and finally real-time accurate pedestrian detection is implemented.

The overall steps are as follows:

(1) Train linear SVM and HIK SVM classifier to obtain linear SVM classifier and HIK SVM.

(2) Use linear SVM to perform preliminary classification on the CT images of the detection window in each image. If it is determined that the detection window contains pedestrians, the detection window will be retained and recorded; if no pedestrians are included, the next detection window will continue until all detection windows are traversed.

(3) Use HIK SVM classification to classify and distinguish the detection window in detail. The windows detected by the upper classifier that are considered to contain pedestrians are traversed and reclassified, the windows without pedestrians are discarded, and the successful detection windows are retained, the final output confirms the window containing the pedestrian and includes its location

# RESULT

This paper collected pedestrian videos of different sections in the campus environment to evaluate the classification effect of classifier and compared this method with the traditional classification method HOG+SVM. In addition to the difference in feature detection and classification methods, edge detection and other processing methods remain the same. Its detection indexes are detection rate DR, false detection rate FPR, total time T and average processing time . The detection statistical results are shown in Table I. Fig. 3 shows the pedestrian detection results of different roads.

TABLE I

Pedestrian detection data statistics

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Identify  Way |  |  | /ms | /ms |
| CENTRIST | 83.38 | 7.21 | 22120 | 110.60 |
| HOG | 71.27 | 13.06 | 275619 | 1378.10 |

|  |  |
| --- | --- |
| 小孩们在路上走  描述已自动生成 | 一些人在街道上走  描述已自动生成 |
| (a) | (b) |
| 路上的行人  描述已自动生成 | 图片包含 站, 男人, 女人, 穿着  描述已自动生成 |
| (c) | (d) |

Fig. 3. Pedestrian detection results of different sections. **(a)**Nanmen Road section during the day. **(b)**teaching building during the day. **(c)** Evening canteen section. **(d)** Evening Flyover Section

In terms of detection, the CENTRIST+ cascade classifier can be used to achieve more accurate detection, and its detection rate is higher than that of the traditional HOG+SVM detection method. TableⅡ is a separate comparison result of detection efficiency. The traditional HOG detection method cannot meet the real-time requirements of detection in time. This paper uses the fast pedestrian detection algorithm to detect the CENTRIST feature, which also speeds up the detection rate. In addition, the combination of SVM+HIK SVM also ensures the correct rate of detection, which has a better detection effect than a single linear detector.

TABLE Ⅱ

Comparison of processing time between the two algorithms

|  |  |  |
| --- | --- | --- |
| Size | HOG/ms | CENTRIST/ms |
| 240\*320 | 576 | 45 |
| 480\*640 | 1187 | 102 |
| 1440\*1080 | 7821 | 533 |

# CONCLUSION

This paper mainly studies fast pedestrian detection methods and compares HOG+SVM detection methods. And uses CENTRIST feature detection and fast SVM classification methods to increase the detection rate; by using SVM+HIK SVM— two-level classifier detection, the detection accuracy is guaranteed; In the target selection, the non-maximum value suppression method is used to select the optimal target detection frame. Finally, experiments verify the advantages of this method in detection rate and efficiency.

References

1. Dalal N . Histograms of oriented gradients for human detection[J]. Proc of Cvpr, 2005.
2. P Dollár, Tu Z , Perona P, et al. Integral Channel Features[C]. British Machine Vision Conference, BMVC 2009, London, UK, September 7-10, 2009. Proceedings. DBLP, 2009.
3. Wu J, Rehg J M. CENTRIST: A Visual Descriptor for Scene Categorization[J]. IEEE transactions on pattern analysis and machine intelligence. 2011, 33(8): 1489-1501.
4. Vempati S, Vedaldi A, Zisserman A, et al. Generalized RBF Feature Maps for Efficient Detection[C]. British Machine Vision Conference. DBLP, 2010.
5. Freund Y, Schapire R E. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting[J]. Journal of computer and system sciences. 1997, 55(1): 119-139.