

PAPER • OPEN ACCESS

Pedestrians Detection Based on the Integration of Human Features and Kernel Density Estimation

To cite this article: Guo-hua Ji *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **490** 042027

View the [article online](#) for updates and enhancements.



IOP | ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the **collection** - download the first chapter of every title for free.

Pedestrians Detection Based on the Integration of Human Features and Kernel Density Estimation

Guo-hua Ji^{1,*}, Xiao-hua Zhang² and Xian-yi Cheng³

¹Silicon lake vocational and technical college, Kunshan, JiangSU, China

²Silicon lake vocational and technical college, Kunshan, JiangSU, China

³Silicon lake vocational and technical college, Kunshan, JiangSU, China

*Corresponding author e-mail: 549962712@qq.com,
383423268@qq.com, xycheng@ntu.edu.cn

Abstract. Kernel density estimation doesn't need the distribution hypothesis of the background characteristics, and also it does not need to estimate the parameters. Therefore, it can deal with moving object detection under complicated background, but the algorithm application is limited by the selection of kernel function bandwidth. In view of this problem, the method of integrating the human features and kernel density estimation is presented in this paper, aiming at the pedestrian detection. First of all, selecting the kernel function bandwidth through the priori information of moving object, then, based on kernel density estimation, extracting the foreground (i.e. moving object), and finally, once again using the human characteristic to detect the video pedestrians. The experiments show, compared with the traditional methods, the method introducing the priori information can greatly reduce the computing burden of kernel density estimation. Even with the changes of the light and the outside noise interference, the method presented in the paper can accurately detect pedestrians and no-pedestrians. This method can be applied to the vehicle, animal detecting, but a priori information plays an important role.

1. Introduction

Moving object detection is an important issue in video processing. The purpose of moving object detection is to find out the moving object, and then to extract the size, the position, the velocity and some other information of the moving object, so as to provide technical support for deeper image analysis. Concerning the research of the moving object detection, many scholars at home and abroad have put forward many different systems and algorithms.

The system of Video Surveillance And Monitoring (VSAM) which is developed by the project of DARRA in the United States^[1], can enable users to monitor some complex scenes in the future world. The w4 (what, where, when, who) systems^[2] developed by the company IBM are mainly used for the feature extraction and movement-trend track of the pedestrians. In view of the existing shortcomings



of parameter density estimation methods, Yi et al^[3] proposed a non-parametric density estimation method. In domestic, Chinese Academy of Sciences has made a research on the analysis of the behavior and events in traffic scenes and the pattern recognition algorithm for human motion^[4]. Tsinghua University developed a long-distance ID system^[5].

Aiming at the shortages of the above-mentioned methods, this paper presents a kernel density estimation method based on the human features. Through the selection of prior information of human body features and the function bandwidth and compared with the traditional method, the presented method can greatly reduce the computational burden of kernel density estimation.

2. Kernel Density Estimation of Moving Object Detection Combined with Prior Knowledge

The system block diagram of moving object detection by using kernel density estimation method is shown as Fig.1:

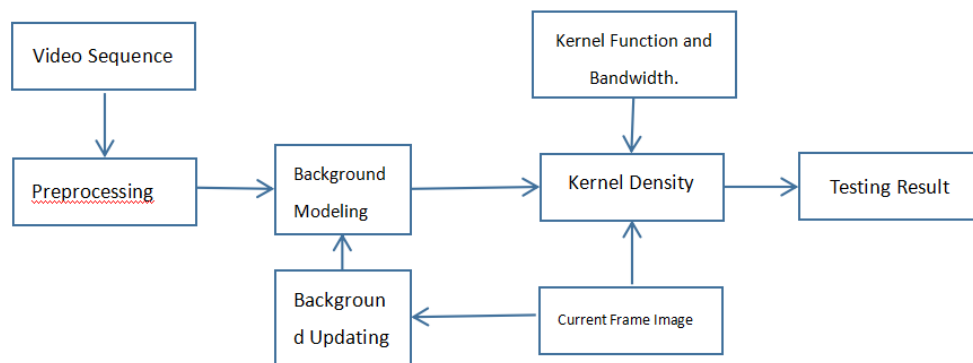


Figure 1. System Block Diagram of Moving Object Detection by using Kernel Density Estimation

2.1 Background Modelling

Kernel density estimation can simulate all kind of unknown density distribution function, and this advantage makes kernel density estimation method well suitable for computer vision analysis.

Assuming that X_1, X_2, \dots, X_n are n sample values of pixel value features in the feature space of a certain pixel point (i, j) in the video image, then the probability density of the pixel value X_t using the kernel density estimation at time t is:

$$\hat{f}(x_t) = \frac{1}{n} \sum_{i=1}^n K(x_t - x_i) \quad (1)$$

Wherein $K(x)$ is the kernel function, and as well Gaussian function is selected as the kernel function :

$$\hat{f}(x_t) = \frac{1}{n} \sum_{i=1}^n \frac{1}{(2\pi)^{d/2} |\Sigma|^{0.5}} e^{\varphi(x_i - x_t)} \quad (2)$$

$$\text{Here, } \varphi(x_i - x_t) = -\frac{1}{2} (x_t - x_i)^T \Sigma^{-1} (x_t - x_i)$$

If the video is of the visible type, then we select the gray value of the image as the feature space, so as to estimate the kernel density by using the gray difference between the background and the moving object. In regard of applying kernel density estimation methods to other practices, the most suitable features can be selected by analyzing all the features one by one. In fact, the value of n has a direct impact on the ability of anti-noise interference, therefore, we should consider that the selection of n will not only ensure the anti-interference ability but also not affect the calculation speed.

3. Pedestrian Detection by Kernel Density Estimation integrating Human Features

The algorithm has a low computational cost and a good robustness against the deformation of the moving object. Therefore, it is a flexible and effective moving object modeling method. However, when the background is complex or the moving object is occluded, this method is not so ideal for tracking accurately the moving objects. Combined with the characteristics of the human body, this section makes full use of these salient features in pedestrian detection to improve the detection effect.

3.1 Human Body Structure

Human movement is a non-rigid movement, and there is not exaggerated distortion of human body during the human movement. Normally, the body motion is created by the body's limbs transformation of position in space. However, no matter how the human body moves, the position change of torso and head is always very small. Therefore, it is ideal to select the part with the smallest change for extracting the features of the moving object.

In the binary image extracted by the frame difference method based on the histogram image enhancement technology, the head of the human body is a figure approaching the circular shape.

When the distance between the human body and the camera is becoming bigger or is becoming smaller, the approximate circular shape of the binary image of the human head will correspondingly become smaller or larger. As the human body moves away from the camera or near to the camera, the feature-changes in other parts of the human body cannot be easily calculated, except that the image changes of the head can be calculated by Hough transform. Therefore, when using the algorithm of the kernel density estimation to detect human movement, the problem of fixing the bandwidth can be resolved by the calculation result of Hough transform.

3.2 Radius Extraction of Human Head

Hough transform can not only detect the circle of which the radius is known, but also detect the circle of which the radius is unknown. Concerning the extract of the feature of the human body, since the human body is changing all the time, the radius of the circular area of the head in the extracted binary image of the human body is regarded as unknown. Therefore, how the Hough transform detects the circles of unknown radii will be a focus here. In an x - y plane image, to determine a circle of unknown radius, it is obligatory to identify three elements, namely the respective coordinates of the circle radius and the circle dot on the y -axis and the x -axis. Through the Hough transform, the coordinates of the circle radius and the circle dot can be detected.

Suppose that in the x - y plane the equation of the circle is the formula (3):

$$(x_i - a)^2 + (y_i - b)^2 = r^2 \quad (3)$$

Wherein, x_i, y_i ($i=1,2,\dots,n$) is the set of points on the circle to be predicted in the binary image, a, b represent the coordinates of the circle dot, and r represents the circle radius. From the formula (3), it can be found that any definite point in the original image corresponds to a three-dimensional cone in the parameter space. The equation of this cone is as follows:

$$(a - x_i)^2 + (b - y_i)^2 = r^2 \quad (4)$$

For any given set of points in the circle $\{(x_i, y_i)\}$, the conical surface clusters formed by these three-dimensional conic surfaces are shown in Figure 2:

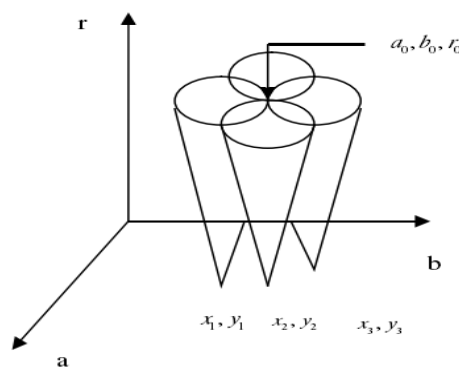


Figure 2. Hough Transform of Circle

As shown in figure 2, If a_0, b_0, c_0 is a set of points located on the circle, then these cone clusters will intersect at one point a_0, b_0, c_0 in the parameter space, and then in the original image plane, this point corresponds exactly to the circle radius and the coordinate of the circle dot. The coordinates of the circle dot and the circle radius is the intersection point of a circle-cones mapped from the head binary image space to the parameter space. The circle dot and the circle radius can be represented by a three-dimensional cumulative array $P(a, b, r) = P(a, b, r) + 1$, then the local maximum of the co-ordinates for the corresponding $P(a, b, r)$ can be found in the image plane, through this maximum the parameters of the circle dot and circle radius can be gotten.

4. Experiment and Result Analysis

Figure 3 shows the comparative experimental results, a1, a2, a3 and a4 are the 31th, the 38th, the 44th, the 62nd and the 82nd frame on the image of the original video Test1. b1, b2, b3, b4, b5 are the detection results using the conventional method (the diameter of the connected area). c1, c2, c3, c4 and c5 are the experimental results of pedestrian detection by the kernel density estimation integrating human head features.

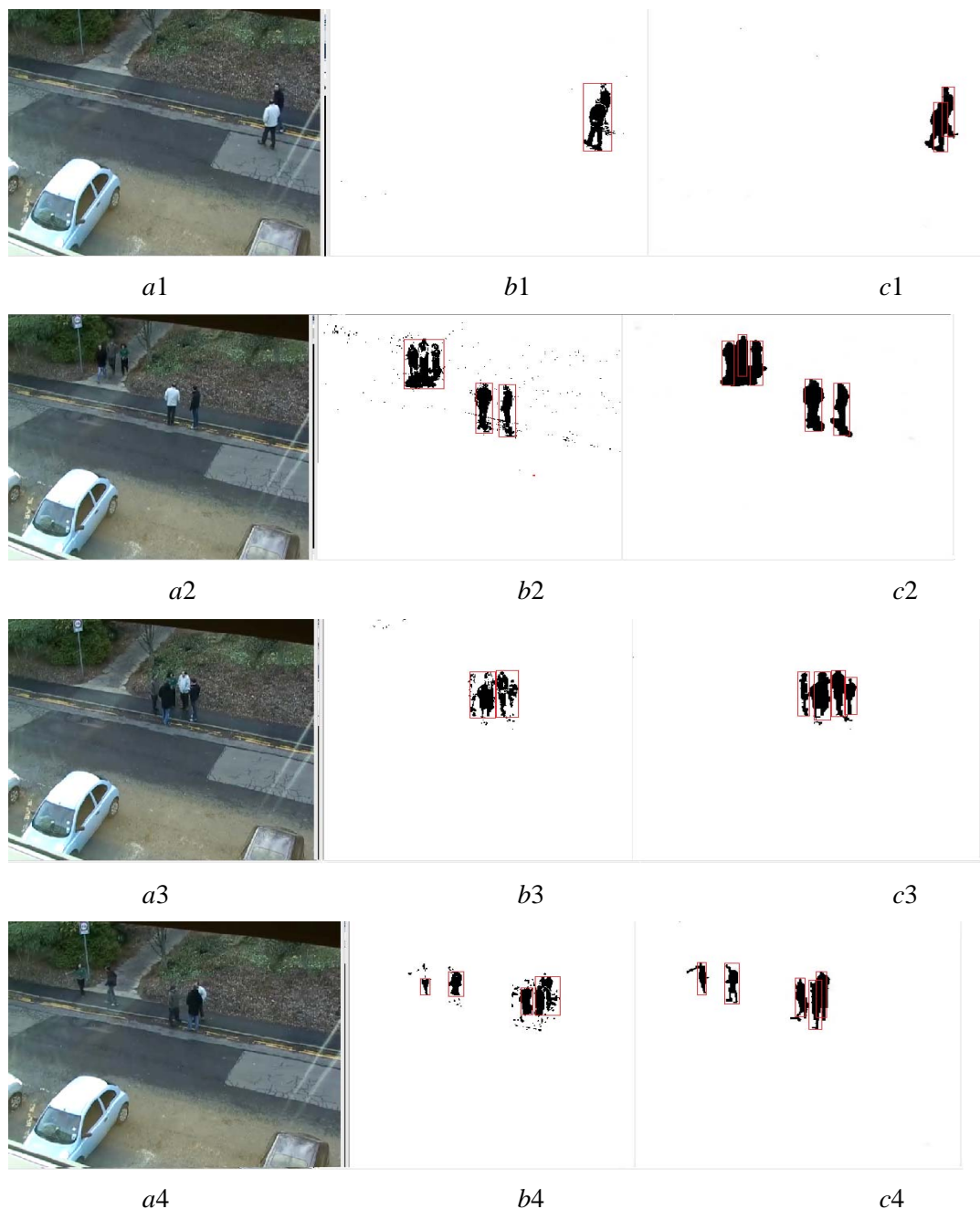


Figure 3. Comparative Experimental Results of Pedestrian Detection

Normally, traditional methods cannot distinguish the people when they are walking side by side (just like b1, b2, b3 shown in Figure 3), but the methods described herein can do it (just like c1 shown in Figure 3). If the human head is blocked, this method described herein cannot well detect all the persons (just like c3 in the Figure 3, the second rectangle on the left contains only two persons, and one person has not been detected.). The traditional method takes the human body as the detect object (just like the two rectangles on the left in picture of b4 shown in Figure 3), since the traditional method does not take into account the features of the human body, the detection accuracy will be affected, and this will mistake other moving objects as the human body. The detection results in the picture C4 is comparatively more ideal. In addition, the method in this paper can avoid the influence of the detection from the abnormal limb movement of the human body. Just like the first object on the

left in the picture c4 of Figure 3, although the arms are outstretched, but the arms have not been detected as part of the human body or as an independent human body. Therefore, concerning the distinction between human body and non-human body, the method described in this paper has obvious advantages(just like the picture c4 of Figure 3).

5. Conclusions

Pedestrian detection involves a wide range of applications and research value, and for moving object detection, the kernel density estimation method has become a relatively perfect detection method. However, despite its obvious advantages, in the kernel density estimation method, the selection of bandwidth of the kernel function is still a problem which is not resolved so far. Therefore, the paper proposes a pedestrian detection method combining human body feature and kernel density estimation.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (61171132,61340037), the Natural Science Foundation of Nantong University (BK2014064),Modern education technology research project of Jiangsu (2017-R-54131) .

References

- [1] Ling Shao,Yan Liu,Jianguo Zhang.Human action segmentation and recognition via motion and shape analysis[J].Pattern Recognition Letters 2012,33(4):438-445.
- [2] Takamitsu Matsubara,Sang-Ho Hyon,Jun Morimoto.Real-time stylistic prediction for whole-body human motions[J].Neural Networks,2012,25:191-199.
- [3] Yi D,Lei Z,Liao S C,et al. Deep metric learning for person re-identification[C]//Proceedings of the 22nd International Conference on Pattern Recognition. Stockholm, Sweden: IEEE, 2014: 34-39.[DOI: 10. 1109/ICPR. 2014. 16].
- [4] Schroff F,Kalenichenko D,Philbin J.FaceNet: a unified embedding for face recognition and clustering[C]//Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston,MA,USA: IEEE,2015: 815-823.[DOI:10. 1109/CVPR. 2015. 7298682].
- [5] Hu L F,Jiang J G,et al. Person re-identification based on multi-features fusion and independent metric learning[J]. Journal of Image and Graphics,2016,21(11):1464-1472.