2015 5th International Conference on Information Science and Technology (ICIST)
April 24–26, 2015, Changsha, Hunan, China

A Robust Pedestrian Detection Based on Corner Tracking

Dongmei Liu,Xuan Wang and Junfang Song

Abstract—**How to detect pedestrian quickly and accurately in complex traffic scenes is the key to pedestrian detection. In contrast to most standard approaches for pedestrian detection and tracking, the approach in this paper has better robust and accuracy. The core part of the approach is to extract local corner features of the objects using Moravec algorithm in video image and achieve tracking these corner features in different image sequence by block matching. Experiment results show the capacity of the approach to detection and tracking is effective in different complicated traffic scenes.**

I. INTRODUCTION

COMPUTER vision is a recent domain with a large set of applications where a lot of research has been elaborated so far. Reliable and robust object detection and tracking is a fundamental component for traffic surveillance. However, pedestrians are one of the most challenging categories for object detection. During the past few decades, a large amount of studies[1-3]have been conducted on pedestrian detection and tracking from video surveillance. According to what is tracked for the change detection in a video sequence, this methods for pedestrian detection can be classified into three categories: region-based, contour-based and feature point-based.

In this paper, we aim to recognize and track pedestrians based on feature points, such as corners, keeps tracking a set of points, rather than an individual region or contour, so that it might provide rather stable detection despite partial occlusions in complicated traffic scenes in real-world. The corner of object refers to the point of which gray scale variation steeply in two-dimensional image or which is the maximum of curvature in image edge curve. Previous studies have lots of approaches to extract corner features, including using edge [4], morphology [5] and grayscale [6], etc.

As the experiment object is mainly for the gray-scale video images, thus we extract the corners of object using the method which based on the gray-scale. According to the realization principle, the methods can be classified as following: Harris algorithm [7] advanced in Gaussian filtering has good anti- interference and exactitude, but the algorithm complexity is too high; Sobel filter based on discrete difference is a rapid edge detection method. It is based on convolve the image with a small, separable, and integer valued filter in horizontal and vertical direction, thus it is relatively inexpensive in terms of computations; Moravec [8] based on gradient defined interest point where corner points could be computed though maximum inhibition. Moravec algorithm can satisfied the requirement of real-time and have good anti-interference performance. In this paper, we aim to recognize and track pedestrians in complicated traffic scenes, thus we adopt Moravec algorithm to extract corner points of pedestrians.

The rest of the paper is structured as follows. The next section describes the method of Moravec. In section 2, we describe the tracking approach and its realization of matching. Then we separate analysis the trajectory of pedestrian and the movement features in section 3 and section 4. Finally, experimental results for the detection and tracking of pedestrians are given in section 5. A final section contains the paper conclusions of our work.

II. MORAVEC ALGORITHM

In this paper, the classical Moravec algorithm which gets the feature points by gray-scale images variance is adopted to extract the pedestrian corner features. This algorithm detected the corner points by closing operation and non-maximal suppression (NMS) based on the definition of "interest value". The concrete realization of this method is as follows:

Select one pixel as a center from the image, then construct a window in the size of $n \times n$ (see Fig.1).

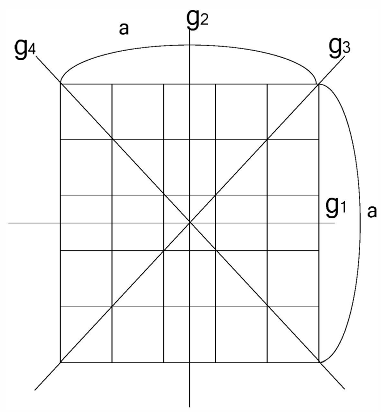


Fig.1.Moravec algorithm's detection window

When we move the window along the horizontal, vertical and two diagonal directions, the interest values (the sum of squared differences (SSD) between a detected pixel and its adjacent pixels) can be calculated in each direction, the expression can be defined as:

Dongmei Liu is with the School of Information Engineering , Chang' an University , Xi'an ,Shaanxi, China (corresponding author to provide phone: 18729532308; e-mail: 1404452367.qq.com).
Xuan Wang is with the School of Information Engineering , Chang' an University , Xi'an ,Shaanxi, China (corresponding author to provide phone: 18729532308; e-mail: 1404452367.qq.com).
Junfang Song is with the School of Information Engineering , Chang' an University , Xi'an ,Shaanxi, China (corresponding author to provide phone: 18729532308; e-mail: 1404452367.qq.com).

978-1-4799-7489-4/15/$31.00 ©2015 IEEE

$$\begin{cases} g_1 = \sum_{i=-k}^{k-1}(f(x+i,y)-f(x+i+1,y))^2 \\ g_2 = \sum_{i=-k}^{k-1}(f(x,y+i)-f(x,y+i+1))^2 \\ g_3 = \sum_{i=-k}^{k-1}(f(x+i,y+i)-f(x+i+1,y+i+1))^2 \\ g_4 = \sum_{i=-k}^{k-1}(f(x+i,y-i)-f(x+i+1,y-i-1))^2 \end{cases} \quad (1)$$

Where k is the image width and height, then we can get four values, g_1 is the interest value of horizontal direction, g_2 is the interest value of vertical direction, g_3 and g_4 are the values of the two diagonal directions. Finally, we obtain the minimum as the interest value of this center pixel as follows:

$$Moravec(x,y) = \min\{g_1, g_2, g_3, g_4\} \quad (2)$$

The greater the value is, the greater the difference of the pixels gradient direction can be and the better the performance of the corner points show. So we put forward a hypothesis T_1, which is 120 obtained from many experiments, if $Moravec(x,y)$ is greater than T_1, we extract these pixels as the corner points.

III. BLOCK MATCHING

The block matching method is very simple and easy to realize. It is the method that obtains the target track trajectories.

A. *The Basic Theory of Block Matching.*

The basic principle of block matching is to divide each frame of the traffic surveillance video into many non-overlapping macro blocks which have the same displacement and find the most similar block in the next video frame in specific matching criteria. The relative displacement between the present block and the matching block is called the motion vector. The process to get the motion vector is motion estimation.

Block matching method is described as follows: suppose we have detected a corner point $P(x,y)$ in the frame t, then regard this point as a center to select a size of $a \times b$ block as the template, finally do sliding search for image block which is meeting the matching criterion in the given $A \times B$ search range in the frame $t+1$ (see Fig.2)

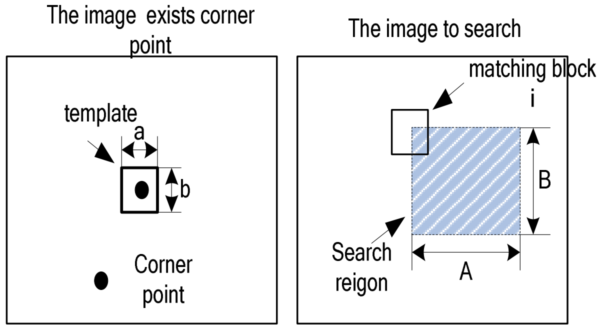


Fig.2. The schematic diagram of block matching. The right is the image of corner points. The left is the image for matching

B. *Block Matching Criterion.*

Block matching criterion is used to measure the level of similarity of two homologous image blocks. It determines the accuracy of the motion estimation and the algorithm validity. In this paper, we adopt the sum of absolute pixel block difference (SAD) as the matching criterion.When the SAD is the smallest, the corresponding block is the best matching block.

$$SAD(i,j) = \sum_{x=0}^{a-1}\sum_{y=0}^{b-1} | F_n(x,y) - F_{n+1}(x+i,y+j) | \quad (3)$$

IV. TRACK IMPLEMENTATION

Firstly, we obtain the pedestrian object's corner points by Moravec algorithm and use the block matching to realize the object tracking in video sequence. Thus, the track trajectories can be obtained. The concrete implementation steps as follows:

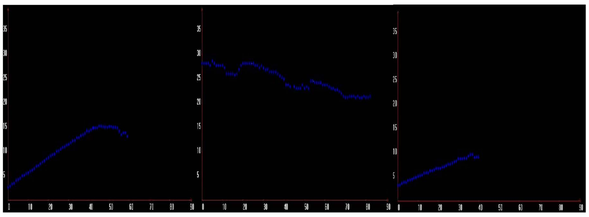
Step 1: in the frame i of the video sequence, we use the method in [9] to get the region of interest (ROI) as the detection region of pedestrian object, and extract object pedestrian corner point $P_i(x,y)$ by Moravec algorithm, and then center around $P_i(x,y)$ to select a $a \times b$ size image block as the template;

Step 2: in the frame $i+1$, the same as Step 1, Select a block of size $A \times B$ and center on $P_i(x,y)$ as the search region, in which do sliding search under the matching rule SAD to find the new matching block. Replace the center point with the new matching point $P_{i+1}(x,y)$, and the new block is substituted as the matching template next time;

Step 3: in the frame $i+2$, ... , frame $i+N$ of the traffic monitoring video, repeat step 2, when the SAD is greater than 480, we can concern the track is end. We can get the trajectory $Track = \{P_i, P_{i+1}, ...P_{i+n}\}$.

(a) (b) (c)

Fig.3.The pedestrian's tracking trajectories based on corner point of different scenarios in video frame. (a)The highway trajectory (b)The urban road trajectory (c)The tunnel trajectory



(a) (b) (c)

Fig.4. The pedestrian's tracking trajectories based on corner point of different scenarios in the real world.(a)The highway tracking trajectory in fact(b)The urban road tracking trajectory in fact(c)The tunnel tracking trajectory in fact

We adopt traffic surveillance videos in many different scenarios to verify this tracking method. Fig.3 and Fig.4 show the experimental results. The method of combining corner points extracting and block matching theory can realize the tracking of pedestrian objects in different scenarios (see Fig.3). The Fig.4 shows that the method we proposed can track the pedestrian objects effectively and obtain smooth movement curve which conforms to pedestrian actual motion.

V. TRAJECTORY ANALYSIS AND PEDESTRIAN DETECTION.

The human motion is randomly, so trajectories are not smooth and have some special points (which are inflection points) dividing the curve into different sections. In order to obtain more accuracy movement information, we separate the trajectories into different sections according to their different motion characteristics.

A. *Subsection Process.*

To analysis the trajectory, there must be a mapping sheet from pixels in the image to coordinates in the real world, thus we get a mapping sheet[10]. By inquiring the mapping sheet, we can get the actual motion trajectory (4) from the image trajectory (5).

$$Track' = \{(s_i,0),(s_{i+1},1),\cdots,(s_{i+n},n)\} \tag{4}$$

$$Track = \{P_i, P_{i+1},\cdots,P_{i+n}\} \tag{5}$$

Where s_n is the distance of P_{i+n} in real world and n is the index of points on the trajectory. The end point and the start point can determine a straight line $y = kx + b$. The distance d_i of any point in the trajectory to the line can be defined as (6):

$$d_i = \frac{|kS_i - i + b|}{\sqrt{k^2+1}} \tag{6}$$

$d_{\max}$ which is the maximum value of all d_i compared with the threshold T_2 which is set in value 100cm in advance based on the experience with different scenes of traffic monitoring videos. If $d_{\max}$ is more than T_2 ,we concern this point as the inflection point $(s_{i,r}, i_r)$,and define $(s_0, 0)$, $(s_{i,r}, i_r)$ as the start points respectively, $(s_{i,r}, i_r)$, $(s_{i+n}, i+n)$ as the end points. Continuing computing the inflection points of each section until all the d_i of this section smaller than T_2, we can get a set of inflection points $\{(s_{i,0}, i_0),(s_{i,1}, i_1),\ldots,(s_{i,m}, i_m)\}$.

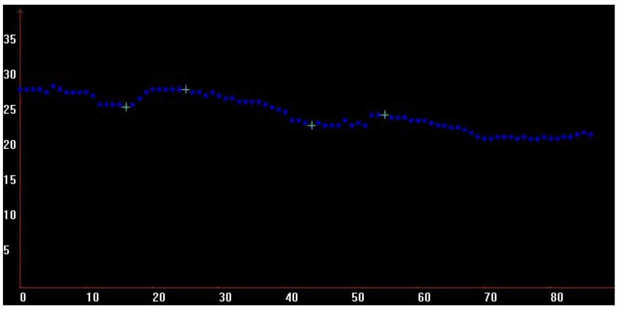


Fig.5 subsection process of urban road trajectory in the real world coordinate.

Fig.5 shows the trajectory of a street cleaner in the urban road. From the holistic monotonicity of the trajectory, we can know that the target was moving from far to near with pause and shift. So there are many not smooth points in the front part of the track, the later part is almost smooth. Thus this approach can realize the segmentation of the track based on different moving stations, and make preparation for the future trajectory analysis.

B. *Linear Analysis.*

Fig.6 shows one section of a trajectory after subsection process, which is not smooth. If calculating the trajectory directly, we can't get the accurate motion information. To each motion state such as walking and running, its trajectory usually meets the linear equation $y = kx + b$. With one time linear fitting, we can obtain the parameters of the equation. But the quality of linear fitting directly influences the accuracy of the detection. The standard to measure the quality of linear fitting is the correlation coefficient, which can be obtained by

$$r = \frac{\sigma_{xy}^2}{\sigma_x * \sigma_y} = \frac{\frac{\sum(x-\bar{x})(y-\bar{y})}{n}}{\sqrt{\frac{\sum(x-\bar{x})^2}{n}} * \sqrt{\frac{\sum(y-\bar{y})^2}{n}}} = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2} * \sqrt{\sum(y-\bar{y})^2}} \tag{7}$$

Where x is distance s_i , y is time $i * \Delta t$, r is correlation coefficient.

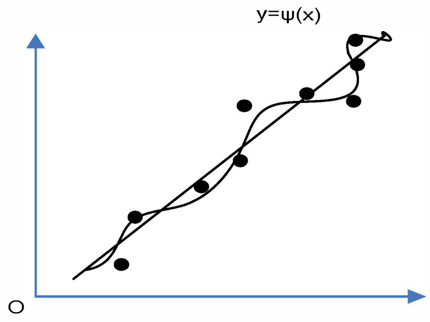


Fig.6 The diagram of Trajectory linear fitting.

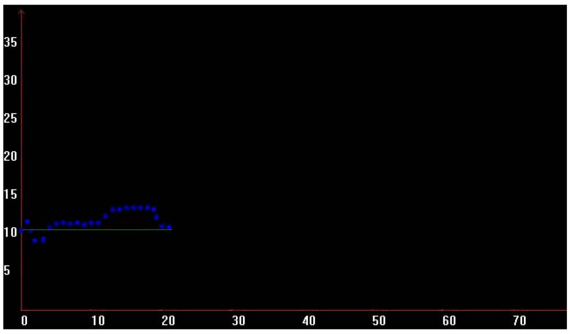


Fig.7 The trajectory of interference target without linear features

The closer r to 1, the higher of the correlation is and the better of the linear fitting quality is. Set a threshold T_3 as 0.5 which is get based on experiments, and calculate r of each section of trajectory. If the r is smaller than T_3, remove this section and reserve the else part to further analysis. Fig.7 shows a interference trajectory caused by the watermark which was left by the passing vehicle in the tunnel pavement. Because r of this track is smaller than T_3, so we remove it .

After the above processing, we obtain the trajectories which can accurately reflect the movement of the target. Assuming that we have gotten a $Track'$ which is eligible. Because $Track'$ has well linear correlation, we can use the start and end point of the trajectory to calculate the speed of the target. We set the thresholds of the speed are T_4, T_5. T_4 is 0.3m/s. T_5 is 2.0m/s.

$$v = \frac{|s_n - s_0|}{(n-1) * \Delta t} \tag{8}$$

Where Δt is the moving time interval. If $T_4 < v < T_5$,we consider this target as pedestrian.

VI. EXPERIMENTS RESULT

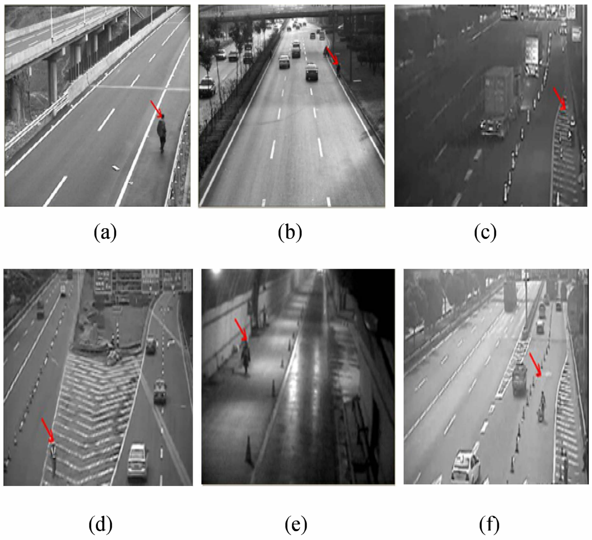


Fig.8 Test results in various complex traffic scenes. (a)The highway of Chongqing (b)The Southern 2nd Ring Road of Xi'an (c)The highway of Shanghai in night (d)The highway of Shanghai(e)The tunnel of Huaying Mountain (f)The highway of Shanghai

We did mounts of experiments in different traffic scenes to verify our method(see Fig.8). The approach proposed in this paper is fast and real time. The detection time is smaller than 20ms, which is faster than the method we used before. It can detect pedestrian accurately in complex traffic scenes such as far from the camera, being covered, in night, in the tunnel, etc. This approach can reduce the latency smaller than 1ms.

In the scenery of Fig.8(a), the pedestrian is walking along the road lane, which is a very well camera version for our tracking detection method. In the Fig.8(b), the pedestrian object is far from the camera, our method can detect the pedestrian accurately and quickly. In the Fig.8(c), it is in the night, which has little light except the car light and street light, the experiment result is well too. In the Fig.8(d) and (f), it is the result of common road scenes. In the Fig.8(e), it is the tunnel in the night, the light is very nonuniform, but our method can recognize the pedestrian correctly.

TABLE 1
DETECTION RESULTS

Scene	False Positive Rate	Miss Rate	Positive Detection Rate
(a)	0.53%	2.4%	97.6%
(b)	0.42%	3.4%	96.6%
(c)	0.67%	2.1%	97.9%
(d)	0.40%	3.0%	97.0%
(e)	0.54%	4.1%	95.9%
(f)	0.62%	1.9%	98.1%

The method was tested on multiple gray-scale video sequences captured by a 25-Hz camera and digitized at a 720*576 resolution. The final tested performance results are shown in Table 1. We can conclude that the average detection accuracy of this method is 97%. The positive detection rate of (e) is 95.9%, because it is in the tunnel.

VII. CONCLUSION

In this paper, we proposed a tracking detection method of pedestrian target based on corner points. To improve the reliability of target tracking, we adopted Moravec algorithm to extract the local corner point which can represent the uniqueness and distinctiveness of pedestrian target. Set the position of corner point in last frame as the center to select a image block in appropriate size as the template. Do sliding search under the criterion of SAD in the region around the corner point as the center to realize pedestrian tracking in video sequence. Through the subsection process, we can extract motion information of every section of the trajectory by linear analysis, include speed,position, etc. Finally we can realize the pedestrian detection. The experiment confirmed my theory tracking well, it can overcome the influence of light rapid change and overlap and have well stability and applicability.

ACKNOWLEDGMENT

The authors would like to thank the support of the National High-Tech Research and Development Program (863program) of China (No.2012AA112312) and the Fundamental Research Funds for the Central Universities of China (No.2014G3242005).

REFERENCES

[1] P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, IEEE Conference on CVPR,2001, pp.511–518.

[2] B. Leibe, E. Seemann, B. Schiele, Pedestrian detection in crowded scenes, IEEE Conference on CVPR, 2005, pp.878-885.

[3] Z. Lin, Hierarchical part-template matching for human detection and segmentation, ICCV, 2007, pp.1-8.

[4] B.F. Buxton, H. Buxton, Computation of optic flow from the motion of edge features in image sequences, Image and Vision Computing.2 (1984) 59-75.

[5] Xiu-yang Zhao, Bo Yang, Shu-ming Gao, Yue-hui Chen,Multi-contour registration based on feature points correspondence and two-stage gene expression programming, Neuro computing. 145(2014)512-29.

[6] Fang Yang, Feature extraction and analysis on X-ray image of Xinjiang Kazak Esophageal cancer by using gray-level histograms, Medical Imaging Physics and Engineering (ICMIPE), 2013 IEEE International Conference on ,2013, pp.61-65.

[7] Chen-guang Guo, Xiang-long Li,Lin-feng Zhong, A fast and accurate corner detector based on Harris Algorithm,2009 Third International Symposium on Intelligent Information Technology Application,2009,pp.49-52.

[8] Yu-xi Zhu, Cui Ru-chun, The algorithm design of wavelet image fusion based on interest operator on the determined condition, Mechanic Automation and Control Engineering (MACE), 2011 Second International Conference on, 2011, pp.749-752.

[9] Huan-sheng Song,Sheng-nan Lu,Xiang Ma,Vehicle behavior analysis using target motion trajectories, IEEE Transactions on Vehicular Technology .63(2014)3580-3591.

[10] Xiao-ping Zhu,Yang Fu, A camera calibration method based on traffic video, Electronic Design Engineering.20(2012)48-52.