HK.NCCP International Journal of Applied Mathematics and Soft Computing
Volume 2, Issue 1, June 2016

Research on the tracking of public places based on pedestrian detection

Hean LIU*, Zhike KUANG
School of Information Science & Engineering, Hunan City University, Yiyang Hunan 413000, China

Abstract: Aiming at the defects of the traditional target tracking, the target tracking of pedestrian detection in public places is proposed. In public places, people, the environment is complex, from many aspects and angles to pedestrian detection of tracking and detection in experimental research. The results show that the method can greatly improve the detection accuracy, and has the practical value.

Keywords: Public places; Pedestrian detection; Tracking

1. Introduction

With the development of computer, the popularization and development of surveillance cameras also urgently need newer and better technology to introduce to the video sensor network to use the computer instead of the human senses and the ability to think to achieve intelligent scene monitoring [1-5]. The detection of the target object in the image and video is an important new area of computer science. The typical applications include automotive autopilot, the statistics of the flow of people, content-based image retrieval, advanced man-machine interface, robotics, intelligent control, video surveillance intelligent scene recognition and so on.

Video surveillance plays an important role in the management and security of the public safety, especially in subways, airports and other public transport places at the entrances handle monitoring video, detect passenger information within the monitored area and provide reliable management department data and auxiliary traffic management, and it also can be used as an important reference basis for planning transportation facilities [6-8]. However, in these traffic scenes the intensive passengers posed a great challenge to the application of video image processing technology.

To solve the above problem, the paper proposes a multi-camera machine-based pedestrian target detection and counting tracking technology, which through the synergy of multiple camera points target search, pedestrian feature extraction and classification, the goal of the multi-view fusion and occlusion handling effectively deal with the study. Its search space has been reduced and the detection accuracy has been improved significantly [9-10]. Multiple cameras distributed around the surveillance area. Each camera height from the ground is 5 m and the vertical angle of about 50 degrees and toward the same ground. In order to make the surface area be covered as much as possible, the horizontal direction needs to keep a certain angle between the cameras and preferably uniformly distributed around the monitored region. The multi-camera coordination pedestrian detection network structure is shown in Figure 1.

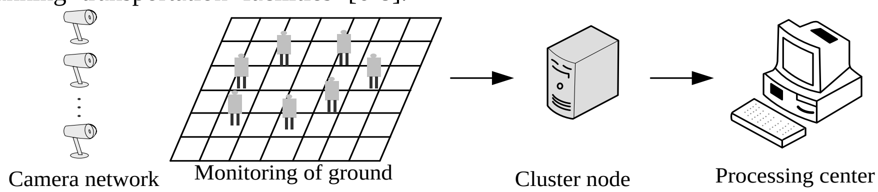


Figure 1. The basic network structure

The structure shown in Figure 1 is the basic network structure. Compared with the structure of the single-camera, multi-camera surveillance can cover a larger area and the obscured issue can be effectively addressed. In order to facilitate the synchronization process, the resolution and frame rate of each camera should be consistent. The camera captured video data and transmitted to the sink node. After finished the data synchronization of each camera, it passed to the processing center. The processing center responsible for the task of pedestrian detection and provide users with a friendly interface. Thus the target of the search in a multi-camera pedestrian detection, pedestrian different appearance and posture change and occlusion of the background changes and pedestrians are the problem need to be solved.

For the existence of pedestrian classification, we proposed extraction method based on multi-level edge dimension reduction and multi-level texture pedestrian characteristics. The process is shown in Figure 2.

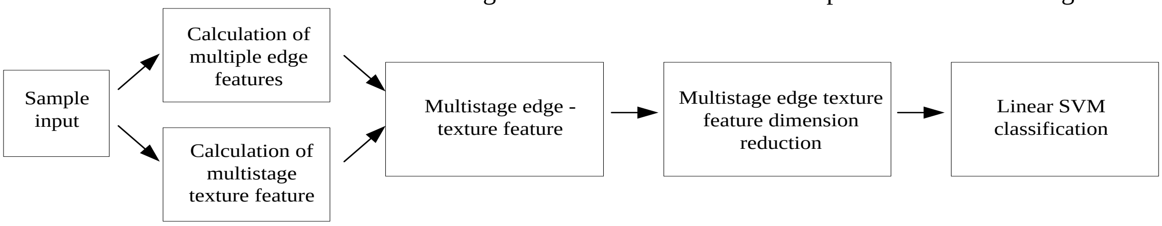


Figure 2. The extraction process based on multi-level dimension reduction edge of a texture feature

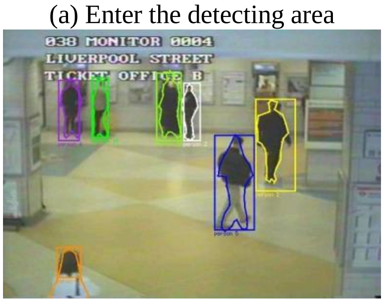
We use a multi-level edge and texture features to describe pedestrians, and reduce this feature dimensionality and eliminate the redundant information. The facts have proved that this method can better solve the background change issues and pedestrians attitude change.

2. Collaborative Multi-camera Machine Pedestrian Detection Count

2.1. Pedestrians counting

We set two detecting line for detecting the number of pedestrians from the direction of coming and going. The region between these two lines is referred to as effective area. When the pedestrians enter the effective area, the pedestrian track count process starts to work, as shown in Figure 3.

(a) Enter the detecting area



(b) Leave the detecting area

Figure 3. Figure of pedestrians counting

If the pedestrian passing the detecting line is a single one, the counter will simply increase the count value. When a width exceeds a single threshold, we think it is a crowd. The trial finds that the projection between single pedestrians of a population has a relatively clearer peak and trough. So as to the count of the population, we use the method of the number of waves in the statistical projection view.

2.2. Moving target detection

The detection of moving target is the premise of realizing the expression of moving behavior. Its detection effect is shown in Figure 4.

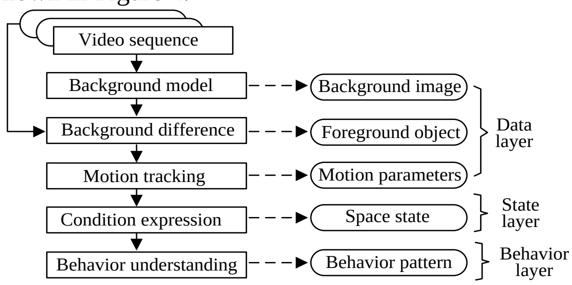


Figure 4. Flowchart of pedestrian crossing state expression and analysis technique

This paper uses the Gaussian mixture model (GMM) in literature [5] to extract the background, and thus to detect the foreground moving target. In the GMM model, the video sequence of the t-th frame of a specific background pixels point I(x,y) can be regarded as a time sequence $\{x_1, x_2, \ldots x_i\} = \{I(x, y, i), 1 \le i \le t\}$,and the time series can be expressed as a superposition of K Gaussian distributions, i.e. a mixed Gaussian distribution. The probability density function of its current pixel point Xt can be expressed as:

$$P(x_i) = \sum_{k=1}^{k} wk\Pi(x_i, u_k, \sum k, i) \tag{1}$$

In this formula, K is the number of Gaussian functions; $w_{k,i}$ is the weight coefficient of the corresponding Gaussian function; $u_{k,i}$ is the mathematical expectation of the k-th Gaussian model. $\sum k, i$ is the covariance matrix of k Gaussian models; η is the Gaussian model, which is expressed as:

$$\eta x_i \mu \sum i = \frac{1}{2(\pi)^{y} \sum 1\Upsilon} x^{-1}(x_i - \mu_i) \tag{2}$$

Due to the influence of the changes of the ambient environment and light, each pixel point of the current image matches with K known distribution models in order to update the model parameters in real time. If the distance between the pixel value and some distribution even value is 2.5 times smaller than the standard deviation, the pixel point will match the distribution successfully, and it will update parameters according to certain rules; The unmatched distributions maintain the original parameters. The updated functions are as follows:

$$\mu_t = (1-p)\mu - 1 + px_t$$

$$O_t^2 = (1-p)\sigma^2 - 1 + p(x_t - \mu_t)^T (x_i - \mu_t) \tag{3}$$

$$p = \mu(x_t, \mu_k, a_k) \tag{4}$$

The K-th distribution weight w_k i will be adjusted according to the following:

$$w_k, t = (1-t)w_k.t - 1 + T(M\ k, t) \tag{5}$$

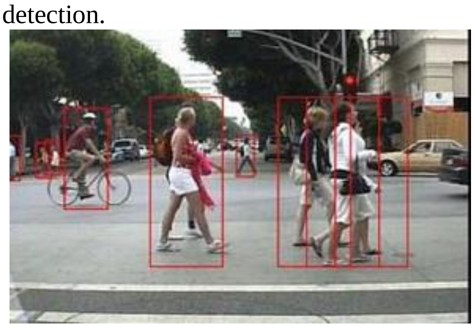
If match, $M_k, t = 1$; otherwise, $M_k, t = 0$.

In the formula, T is the learning rate. If the Gaussian mixed model does not exist the matching Gaussian distribution, re-generate a new Gaussian distribution to replace the distribution of the very bottom in the mixed model according to the descending order of w_k / Q_k. Wherein, the Gaussian distributions which are most likely to describe the process of the stable background will locate at the top of the sequence. The distributions generated by the background transient disturbance will slide to the bottom of the sequence, and eventually will be replaced by the Gaussian distributions which have been given a new assignment value. Thus, we will choose the first η Gaussian distributions in the above sequences as a background pixel model:

$$\eta = \arg\min(\sum_{j=1}^{b} w_j > T \tag{6}$$

Wherein, T is the threshold parameter determining the background model, here is 0.8. The former N Gaussian distributions of the first K Gaussian distributions after ordering is the background.

Through the above GMM model, we can obtain the background image effectively and self-adaptively, thus to detect the prospect moving target by using the background difference. Figure 5 is an effect chart of moving target detection.



(a) Original ima

(b) Background image obtained by GMM model

(c) Foreground image obtained by background difference

Figure 5. Intersection multi-pedestrian moving detection

2.3. Moving target tracking

On the basis of the target location and speed obtained by motion tracking, this paper uses a method which is based on finite state machine (FSM) to effectively express the movement state of the pedestrians crossing street process. Finite state machine is a mathematical model which describes complex systems through the simplistic assumptions. It can effectively support the modeling of a variety of complex behaviors and is widely used in communication protocols, digital circuit design, lexical analysis, text editor, and many other research areas.

According to the empirical model, the pedestrian movement process can be divided to three microscopic states: walk, run and stop. When in a certain state, if the pedestrians are influenced or constrainted by those objective factors like the surrounding environment, road infrastructure or other traffic participants, they will accelerate or decelerate with these factors, thereby changing its current state of motion and enter into the next state; otherwise they maintains the original quo, and so forth until the end of the process of acrossing the street. Here pedestrians' acceleration, deceleration and non-shifting are the input operations. The state conversion process of pedestrians crossing street is shown in Figure 6.

Figure 6. State transition diagram of pedestrian crossing street finite state machine

In Figure 6, s0 is the stopped state; s1 is the walking state; s2 is the running state; x0 is the accelerating operation; x1is the decelerating motion; x2 is the shifting action.

3.Three-dimensional Ground Space Search

Through RANSAC algorithm we obtain a fundamental matrix B, the rotation matrix R and translation matrix few. Thereby, we can obtain the coordinates of the eight pairs of matching points in the world coordinate system, which build a three-dimensional ground.

For any of the pairs of calibration point $p_1(x_1, y_1)$, $p_z(x_z, y_2)$,and its the coordinates $w(x, y, z)$ in the world coordinate system may use linear triangulation method [MA04] to achieve it. The detailed steps are as follows:

Tectonic reconstructions moment $Q = [q_1, q_2, q_3, q_4] \in R^{3\times3}$ in it.

$$q_1 = [0 -1\ y_i\ 403]$$
$$q_2 = [-10\ x_i\ 0]$$
$$q_3 = (R(2, y) + y_2 \times R(3, y) - T(2) + y_2 \times T(3)$$
$$q_4 = (-R(1, y) + x_2 \times R(3, y) - T(1) + x_2 \times T(3)$$
$$(7)$$

Here, R (1, :) indicates all elements of the rotation matrix corpse in the first row, and the meaning of the side 2, :) and R (3, :) can be done in the same manner

4.The Pedestrian Target Fusion and Occlusion Handling of Multi-camera Machine

Single camera can handle some of the local shelter. However, due to the inherent limitations of a single camera, it can only handle relatively small partial occlusion,and is powerless for more serious occlusion . Therefore, the use of multi-camera to handle occlusion is a necessary means [Khanog]. As shown in Figure 7.

The pedestrian who was severely blocked in a perspective, may be in another perspective is completely visible. Using PMHL detection sub, the right woman can be detected, and can not be detected in the left. we can take two center point Pl and p2 in black rectangle. If they meet homograph constraint relations, we can determine pedestrians of them is the same goal. In this way, we can achieve multi-camera collaborative pedestrian counts, that is, if the pedestrians under detection of the multi-camera satisfy the homograph constraint relations, the counting system is to an increase of only 1.

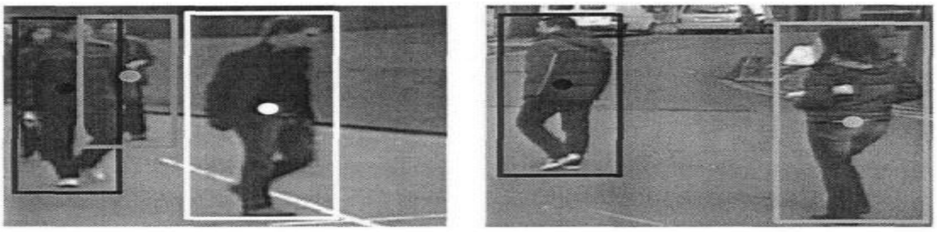


Figure 7. Processing and integration of multi-angle shielding

5. Results of Testing and Analysis

5.1. Multi-view target fusion performance testing

We first tested the proposed multi-view target fusion method. we can calculate the value of the MODA and MODP in each perspective by setting the value of the overlap range between detection window and the marked window is 0.5. Table 1 shows that by using the multi-view goal fusion method, MODA and MODP in each perspective have a better improvement, which shows that multi-view target fusion method can improve the detection accuracy and reduce false alarm in PETS datasets illustrate.

Table 1. Multi-view target fusion method to test in the the PETS datasets

	MODA		MODP	
View 5	Not fusion	Fusion	Not fusion	fusion
View 6	0.5632	0.6863	0.5764	0.6538
View 7	0.6532	0.9235	0.9542	0.6539
View 8	0.5536	0.5634	0.3256	0.4675
	0.6438	0.6897	0.7534	0.5362

5.2. Performance comparison of different types of multi-camera pedestrian detection method

We compared Our methods with the latest two multi-camera pedestrian detection methods . These two methods are the multi-camera sampling method (multi view sampler method) [oelo]and PoM method. By varying closing value: the value of, we get the value the corresponding perspective MoDA and MODP , and then gain the average value of all perspective MODA and MODP , causing the curve shown in Figure 9. Figure 9 shows that at the different closed value, our multi-camera pedestrian detection method is better than the POM method, and is basically the same with performance of the multi-camera sampling method. However, our method does not assume that the pedestrian is moving. Therefore, our methods are effective on stationary and moving pedestrian.

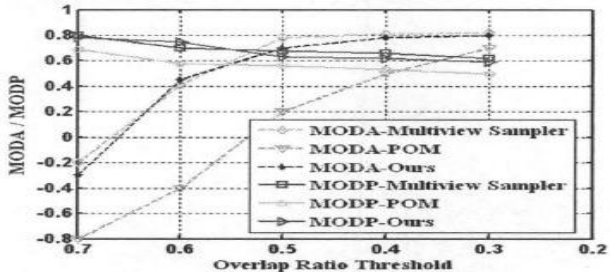


Figure 8. Performance comparison of the MODA and MODP of multi-camera pedestrian detection method

6. Conclusion

In this paper, it uses ground three-dimensional space search and build, tracks the number of pedestrian target under multi-camera collaborative and monitors movement mode to do multi-views of analysis and detection. It uses particle filter method to track the boundary and occlusion problem at abortion advocacy scene has been effectively addressed, which is more accurate, and more efficient than traditional detection methods.

References

[1] Guisheng Yin, Yanan Zhang, Weiwei Yuan,Xiaohui Cui,Xiang Cui,Weijie Cheng, "A Personalized Spam Page Detection Method Using Hidden Markova Model", JCIT: Journal of Convergence Information Technology, Vol. 7, No. 12, pp. 111 ~ 119, 2012

[2] Fei Xiong, Yun Liu, Jiang Zhu,Jie Lian, Ying Zhang, "Hot Post Prediction in BBS Forums Based on Multifactor Fusion", JCIT: Journal of Convergence Information Technology, Vol. 7, No. 12, pp. 129 ~ 137, 2012

[3] Lu Li, Guoyin Zhang, Aihong Yao, Ling Li, "Research on the MP2P-Based Content Distribution Mechanism", JCIT: Journal of Convergence Information Technology, Vol. 7, No. 12, pp. 195 ~ 206, 2012

[4] Zhe Jia,Zhe Gong,Zheng Wei, Jianyi Zhang,Shoushan Luo,Yang Xin, "A Distributed Method on Web Log Sequential Pattern Mining", IJACT: International Journal of Advancements in Computing Technology, Vol. 4, No. 6, pp. 24 ~ 33, 2012

[5] Xu Jin, Luo Jie, "A Study Of Multi-Agent Based Model For Urban Intelligent Transport Systems", IJACT: International Journal of Advancements in Computing Technology, Vol. 4, No. 6, pp. 126 ~ 134, 2012

[6] Huawei Zhang, Wen Zhang, "Large Query Optimization Based on Lessening Complexity", IJACT: International Journal of Advancements in Computing Technology, Vol. 4, No. 6, pp. 157 ~ 164, 2012

[7] Liu Hean, Wu Shuguang, "Research on the Memetic Algorithm Research on Multimodal Function Optimization", JCIT: Journal of Convergence Information Technology, Vol. 7, No. 18, pp. 464 ~ 472, 2012

[8] Jianyuan Li and Jihong Guan, "Learning with Joint Sparse Representation", IJACT: International Journal of Advancements in Computing Technology, Vol. 4, No. 6, pp. 184 ~ 191, 2012

[9] Chih-Chieh Yang, Pei-Ju Lin, Chih-Cheng Sun , "Product Form Design Using Virtual Hand and Deformable Models", JDCTA: International Journal of Digital Content Technology and its Applications, Vol. 6, No. 11, pp. 8 ~ 17, 2012

[10] Qing-long Wu, Yuan-biao Zhang, Yi-qun Yao,Zhong-jie Shen,Han Qi, "Evaluation of Heavy Metal Pollution in Urban Soil and Positioning of Pollution Sources", AISS: Advances in Information Sciences and Service Sciences, Vol. 4, No. 16, pp. 82 ~ 93, 2012

Subscriptions and Individual Articles:

User	**Hard copy:**
Institutional:	800 (HKD/year)
Individual:	500 (HKD/year)
Individual Article:	20 (HKD)