Intelligent Video Event Detection and Behavior Analysis Research

Xiong Fang

Experimental Center, Hunan International Economics University, Changsha, 410205, China
Tel.: 86-0731-88127113
E-mail: matlab_wjf@126.com

Abstract: Intelligent video surveillance has been developed rapidly and broadly in daily public security and various civil applications. To establish an effective and standard performance evaluation, an appropriate metric is very important. This survey summarizes the state-of-the-art of performance evaluation for intelligent video algorithms, especially focuses on the performance evaluation of motion detection and object tracking. It also highlights the future direction for performance evaluation of intelligent video technology. *Copyright © 2014 IFSA Publishing, S. L.*

Keywords: Intelligent video, Motion detection, Tracking, Performance evaluation, Video surveillance.

1. Introduction

With the rapid development of video surveillance in recent years, there is rapid increase in the number of cameras, a lot of video data bring challenges to real-time monitoring alarm and the effective use of video data. In standard video surveillance systems, there are often dozens or even thousands of road-way video and corresponding digital video data. Operator discovers the where there is alarm event by observing each video, and these become very difficult. Further, when the alarm event is analyzed later, the operator needs to find the relevant record in time, intelligent video analysis is lacked in traditional video surveillance, and video data can not be efficiently retrieved, only manual searches are done according to the approximate time periods, the video data analysis is caused to consume a lot of work [1, 2]. An effective way to solve the above problem is automatic intelligent video analysis. The event is extracted and recorded in real-time, which appeared in the video and the user is interested, a timely warning is achieved, and video data is efficiently retrieved by using the information of the stored event [3, 4].

Event detection and behavior analysis system is an advanced intelligent video analysis system. Perimeter video can be monitored and abnormal behavior is analyzed, it can be applied to a variety of public places including airports, railway stations, ports, buildings, streets, residential, and other places, for the detection, classification, tracking and recording passers, vehicles and other suspicious objects, that can determine whether there is a long wandering, stay, retrograde behavior of pedestrians and vehicles, such as occurred in the penalty area, in addition to testing personnel running, fighting and other unusual behavior [5, 6].

Event detection and behavior analysis system can simultaneously monitor the different behavior of the same scene in multiple targets, and these can be set flexibly according to the characteristics of the prevention target, including the type and size of the moving object [7, 8]. The system can adapt to

different environmental changes, including light, seasons, day and night, rain or shine, etc., and it can be a good on anti-camera shake. Intelligent video analysis surveillance system solution is shown in Fig. 1.
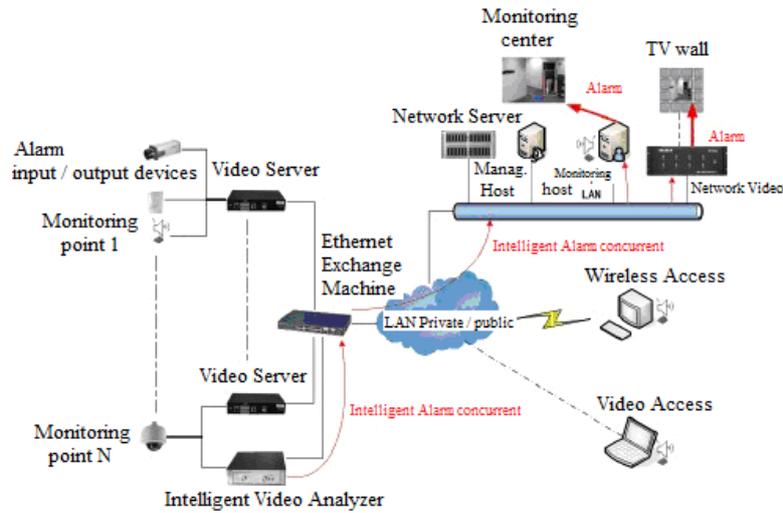

Fig. 1. Intelligent video analysis surveillance system.

2. Intelligent Video System Function

By event detection and behavior analysis system, we can achieve the following main functions [9, 10]:

1) To enter illegally the penalty area: there is the ability to detect whether there is a pedestrian or vehicle into a predetermined area, the user can flexibly adjust the sensitivity. More advanced systems can effectively detect the target of not less than 10 pixels; (Fig. 2: ①,②)

2) To retrograde illegally: When there are the reverse movement of pedestrians and vehicles in the designated area, recording and alarm; (Fig. 2: ③)

3) Exception running: to test whether there is rapid movement in the designated area; (Fig. 2: ④)

4) Fights: to detect whether counterfeiting brawl; (Fig. 2: ⑤)

5) The dynamic image is enlarged: the PTZ camera can be controlled to capture and zoom on a moving target, and high-resolution images are generated. (Fig. 2: ⑥)

6) The identification and classification: the pedestrians, vehicles and other targets in video are classified; (Fig. 2: ⑦)

7) To strand sensitive areas: to detect whether suspicious people or vehicles wander or stay for a long time in the designated area, when the residence time exceeds a preset value, the system will issue a warning; (Fig. 2: ⑧)

8) Access trailing monitoring: detecting whether someone followed through access control, to alarm instantly if it is found; (Fig. 2: ⑨)

9) Mob: to detect groups gather, retention (Fig. 2: ⑩)

10) Items left behind: When legacy article phenomenon occurs in the environment, to alert immediately (Fig. 2: (11))

11) Sudden change: when the video image has a dramatic change (such as a camera is blocked and substantial move), the system alarms (Fig. 2: (12)).
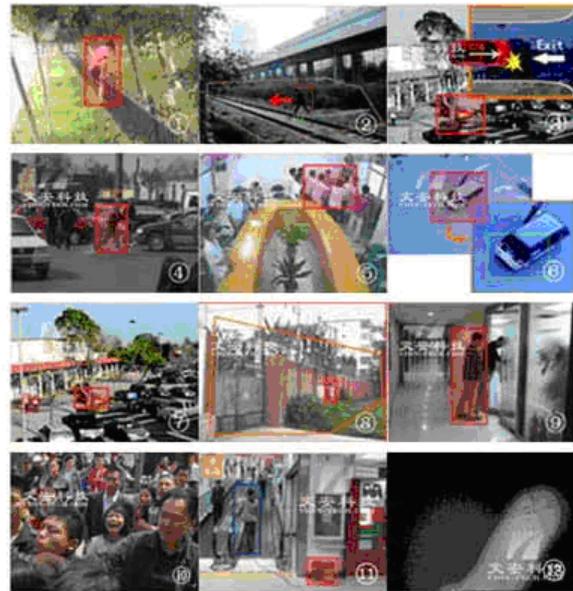

Fig. 2. Event detection and behavior analysis system.

3. Event Detection and Behavior Analysis System Morphology and Architecture

In event detection and behavior analysis, there are advanced technologies with system combines video processing, image processing, pattern recognition, artificial intelligence, and multiple areas. There are front-end and rear-end center treatment approach in the form of main products.

224

By using backend central processing, video analysis is done in high-performance intelligent video server. In intelligent analysis server, digital video stream usually is decompressed by soft decoding, and then event detection and behavioral analysis are done. Central processor can be integrated into third-party video surveillance platform, it is linked with mainstream video surveillance platform seamlessly [11, 12]. Event analysis results and related parameter settings also communicate through intelligent video transmission protocol and with the video surveillance platform. Center handling system per server can simultaneously handle eight generally to 16 video. In this mode, all of the front camera only has basic video capture capabilities, and all the video analysis must be brought together to back or key node is unified and is processed by a computer [13, 14]. The structure of the back-end intelligent video analysis system is shown in Fig. 3. Another is intelligent video processor based on the front-end solution. In this work mode, all target tracking, behavior judgment, alarm trigger is done by the front-end intelligent analysis equipment, only the alarm information is transmitted to the monitoring center over the network. Video analysis equipment is placed after the IP camera, so the bandwidth of video streaming can be effectively saved. The structure of the front-end intelligent video analysis system is shown in Fig. 4. Intelligent Video Architecture is shown in Fig. 5.
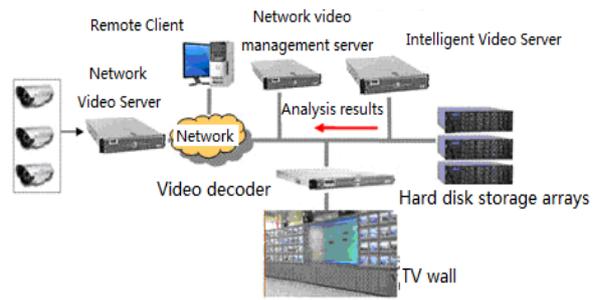

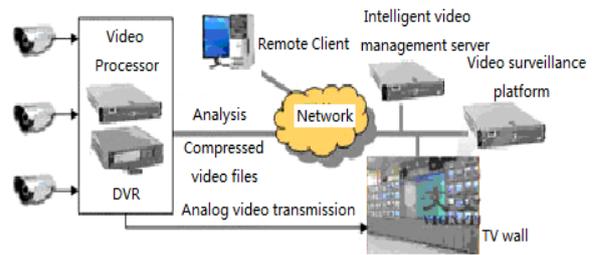**Fig. 3.** Event Detection System – backend center system topology.

Fig. 4. Event Detection System – front center system topology.

Fig. 5. Intelligent Video Architecture.

4. System Technology Theory and Evaluation of Event Detection and Behavior Analysis

4.1. Perimeter Detection

Technical principle is the foreground object detection technology. This technology is that the user interest region is monitored by the fixed camera surveillance, background model is built, foreground objects are filtered out and the movement routes of objects are determined. The behavior discrimination of moving objects is achieved based on movement routes of foreground objects. Typical applications of the technology include the following four categories.

1) Regional and out: user defines a polygon area in the scene in the monitoring, when there are objects entering or leaving that trigger an alarm event.

2) Regional residence: user defines a polygon area within monitoring scenario, when there are

objects in the area, if its residence time exceeds a predetermined threshold, an alarm event is triggered.

3) Tripwire: Users define a segment in the monitored area. When the object moves across the line segment, an alarm event is triggered.

4) Directed tripwire: user defines line segment in the monitored area. When the object across the line in a specified direction, that trigger an alarm event.

Application premise: the camera needs to be fixed, camera is moved only when moving the monitoring scene changes, otherwise it is impossible to establish a stable background model, it is fit to open a scene, such as border lines, the environment surrounding of the plant, it is available to a larger vision on the one hand, another aspects of this technique is not applicable to the foreground congested object, then different scenarios foreground objects can not be distinguished due to mutual adhesion, movement routes are obtained based on prospects, it also is fail.

Reference Test Environment: In addition to routine testing recognition rate, false positive rate and false negative rate, we should examine the results of the following targeted testing.

1) Jitter: outdoor camera chance or periodic jitter occurs due to various reasons, because the technology generally requires a fixed camera, and therefore the specifications need to be tested in a slight shake environment, stability of system performance indicators are researched.

2) Unrelated objects: there is a higher probability test environment, such as leaves jitter, flowers shaking, clouds and other objects interference, the false happens.

3) Weather: general positioning of the system is real-time monitoring systems in working around the clock, these technical indicators are tested in the rain, fog, snow weather.

4) Light: outdoor environments or indoor lighting environment changes significantly, the test should be carried out in 24 hours, the three technical indicators are obtained, the system's ability is determined to adapt to light changes.

Reference index: the recognition rate is > 80 %, the false alarm rate is < 20 %.

4.2. Dynamic Tracing

Technical principle: The purpose of this technique is to achieve close to a specified target object to keep track. The target object track is completed within the specify panoramic camera scene, the target is generally designated by the personnel operator in a non-empty scene, which is selected for open scene of border lines by a computer of their choice. The target track can be divided into two stages, which is lock target and continuous tracking, tracking camera and the panoramic camera may be physically the same, which depend on the

specific application needs and level of the product itself.

There are two technology in the targeted stages, one is tracking based on feature pattern analysis, for example, an RGB image template matching, because the camera is gradually narrowed, while the position and size of the object change, then the higher the lock speed, then the greater the risk of losing the object from view, if the size and variance the target speed are no a priori case, once the target is lost, relock is cost. By contrast, the slow lock consideration will have much smaller, especially for the application of sensitive military authorities, you should first make sure that the target is not lost, consideration is followed by the lock-consuming. Another three-dimensional information is achieved by calibrating the scene, the locking speed can be significantly improved, but the calibration process is complicated, and the panoramic camera angle is once changed, calibration is generally required, there is greater consideration in maintenance.

Because the dynamic tracking applications are with obvious enemy confrontation, in continuous tracking phase, technically must have some ability to adapt for the direction change of the target object, movement velocity, temporary obscured or even deliberately taken refuge in obstructions. On the other hand, the tracking stability is also directly affect by the dynamic performance of tracking camera itself, and the dynamic performance here include the time constant of camera mechanical motion control part, the signaling transmission delay, the reasonableness of the control protocol parsing mechanism.

Application premise: the technology application is avoided in crowded scenes, if the crowding degree of the scene is the higher, the technique practicality also decreases. In particular, the multiple cameras are used in the tracking relay systems, in the domestic applications scene, the whole system needs to be verified in tracking stability, especially through the actual test. When cameras are installed, the target object and moving objects around it are avoided to be obscured and sustained seriously.

Evaluation Methods: In addition to routine tests, the following specific test should be selected according to the application requirements.

1) Lock Test: the success rate of the individual test at the lock target stage.

2) Static test: when long stationary target, to investigate the stability of the system itself.

3) Speed test: in continuous tracking stage, the target is stopped suddenly, to change direction or to change speed.

4) Shielded test: Target is periodically obscured by trees or telephone poles, or target deliberately takes in refuge obstacles.

5) The suspected target test: when the objectives meet with other moving objects, or when the target is temporarily obscured, there are other moving objects around.

4.3. Retrograde Detection

Technical principles: The technology is used to detect people or objects, which reverse out of the entrances. The main technical feature is to detect the retrograding objects based on the object motion model, which is still applicable in a crowded scene.

Application premise: camera mounting angle should meet to minimize the occlusion between people or objects. Overlooking angle is recommended to be between 45 degrees to 90 degrees, the camera sight should be consistent with the flow direction, the horizontal angle is within plus or minus 30 %.

Evaluation Methods: In addition to the described test mode (basic part), the following targeted testing can be added.

Targeted testing in the crowded scene: for example, in the pedestrian entrance, if the horizontal entrances can accommodate N individuals, 2N ~ 4N people can be made to pass convoy in a more natural way, while a person is arranged to enter from the opposite direction, the leak reported rates of system is detected mainly.

Targeted testing of short / local behavior: in the detection area, one is arranged to do the body motion (such as the arms, upper body, etc.) for the opposite direction, or the backward movement for short time (according to application needs), the system false alarm rate is mainly detected.

Reference index: the recognition rate is > 80 %, the false alarm rate is < 20 %.

4.4. Pedestrian Abnormal Behavior

Technical principles: The technology is used to detect the high-risk behavior of the suspected robbery or fleeing after criminal within the monitored area. There are two main technologies. A class implementation is still based on the foreground object detection, according to the movement route of the foreground objects, its speed is estimated, if speed exceeds a certain value, the police is called. The accuracy of this speed valuation is low, but it could not overcome that the larger the near, the smaller the far off the camera. Meanwhile, people can not distinguish between pedestrians and bicycles, the practicality is not high. The second class implements use characteristic movement pattern analysis techniques, by using pattern recognition method, movement patterns and velocity may appear and are to learned in each part of the scene, abnormal acceleration can be detect not only, the criminals fled when crossing railings and parapets, the behavior is usually low probability, it can be identified.

Application premise: if walking people are the dominated sport in scene, the overlooking angle of the camera is not less than 30 degrees, when the viewed perspective angle is small, and erection height should be improved to avoid serious occlusion between people. Intelligent Video Analysis encoding

service application examples are shown in Fig. 6, Fig. 7, Fig. 8 and Fig. 9.

Fig. 6. Intelligent Video Analysis encoding service application examples.
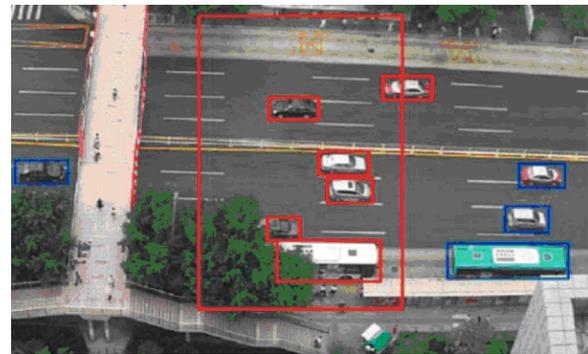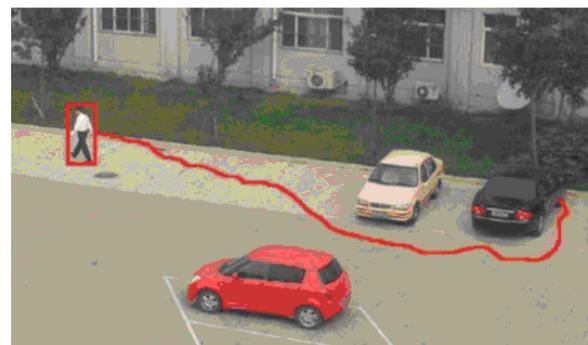

Fig. 7. Traffic statistics.

Fig. 8. Trajectory tracking.

4. Conclusions and Outlook

Different video surveillance industry demand generally has a very significant difference, especially in applications for intelligent video analysis technology, which also determines the specificity of behavior detection and anomaly events in different industry. With the deepening of applications in various industries as well as to further enhance the level of security control requirements, various

security areas will face more and more different challenges, the demand for video surveillance is also increasingly diverse and complex. How to recognize and analyze the behavior has become more of a problem which intelligent video analysis technology had to face in the industry deepen application process. Only combined with practical application in industry, in-depth understanding of the specific requirements of the different sectors, the needs of users can be better grasped, the function of intelligent video analysis technology will be implemented in the application reality. This is that the intelligent video analysis technology is the ultimate expression of the future industrialization value [15-18].
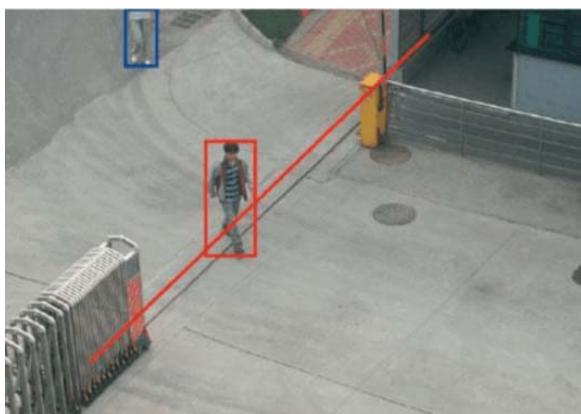
Fig. 9. Cross line tracking.

Intelligent video analysis application related products develop with the development of video surveillance market size, which extends from the relative concentration of applications to all walks of life significantly deeper. Some manufacturers of professional video analysis is also working to improve existing core technology, the structure of existing algorithms are further enhanced and optimized, the actual needs of users are looking for more on the function and application, the best system is achieved in the form of products. As high-end video surveillance applications, such as perimeter detection, object tracking and other functions, it has been successfully applied in various key industries, and gradually emerging power [19-22].

Video surveillance has a large proportion in the future, but the road to the future is relatively flat, this is reflected in the form of a solid, it gradually meets the needs of people, before the curtain of the future video analysis is opened, to do and to accumulate.

References

[1]. J. Aguilera, H. Wildenauer, M. Kampel, Evaluation of motion segmentation quality for aircraft activity surveillances, in *Proceedings of the 2nd Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS'05)*, Beijing, China, October 2005, pp. 293-300.

[2]. D. Thirde, M. Borg, V. Valentin, Visual surveillance for aircraft activity monitoring, in *Proceedings of the 2nd Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance*, Beijing, China, 15-16 October 2005, pp. 255-262.

[3]. D. P. Young, J. M. Ferryman, PETS metrics: On-line performance evaluation service, in *Proceedings of the 2nd Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS'05)*, Beijing, China, October 2005, pp. 317-324.

[4]. N. Lazarevic-McManus, J. Renno, D. Makris, Designing evaluation methodologies: the case of motion detection, in *Proceedings of the 9th IEEE International Workshop on PETS*, New York, 2006, pp. 23-30.

[5]. H. Grabner, P. M. Roth, H. Bischof, Is pedestrian detection really a hard task, in *Proceedings of the 10th IEEE International Workshop on Performance Evaluation of Tracking and Surveillance*, Rio de Janeiro, 2007, pp. 1-8.

[6]. T. Ellis, Performance metrics and methods for tracking in surveillance, in *Proceedings of the 3rd IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS'02)*, Copenhagen, 2002, pp. 26-31.

[7]. A. T. Nghiem, F. Bremond, M. Thonnat, ETISEO, performance evaluation for video surveillance systems, in *Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS'07)*, London, 2007, pp. 476-481.

[8]. J. Black, T. Ellis, P. Rosin, A novel method for video tracking performance evaluation, in *Proceedings of Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS'03)*, Nice, France, 2003, pp. 125-132.

[9]. J. C. Nascimento, J. S. Marque, Performance evaluation of object detection algorithms for video surveillance, *IEEE Transactions on Multimedia*, Vol. 8, Issue 4, 2006, pp. 761-774.

[10]. D. Hall, J. Nascimento, P. Ribeiro, Comparison of target detection algorithms using adaptive background models, in *Proceedings of the 2nd Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS'05)*, Beijing, China, October 2005, pp. 113-120.

[11]. V. Manohar, M. Boonstra, V. Korzhova, PETS vs. VACE evaluation programs: A comparative study, in *Proceedings of the 9th IEEE International Workshop on PETS*, New York, 2006, pp. 1-6.

[12]. A. T. Nghiem, F. Bremond, M. Thonnat, A new evaluation approach for video processing algorithms, Austin, in *Proceedings of the IEEE Workshop on Motion and Video Computing (WMVC'07)*, A Texas, USA, 2007, pp. 15-22.

[13]. R. Kasturi, D. Goldgof, P. Soundararajan, Framework for performance evaluation of face, text, and vehicle detection and tracking in video: Data, metrics, and protocol, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 31, Issue 2, 2009, pp. 319-336.

[14]. R. Stiefelhagen, K. Bernardin, R. Bowers, The CLEAR 2006 evaluation, *Lecture Notes in Computer Science*, Vol. 4122, 2007, pp. 1-44.

[15]. R. Stiefelhagen, K. Bernardin, R. Bowers, The CLEAR 2007 evaluation, *Lecture Notes in Computer Science*, Vol. 4625, 2008, pp. 3-34.

[16]. E. D. Gelasca, T. Ebrahimi, Application dependent video segmentation evaluation: a case study for video surveillance, 2009. (http://infoscience.epfl.ch/record/91012).

[17]. Chu Yiping, Zhang Yin, Ye Xiuzi, Based on hidden conditional random adaptive video segmentation algorithm, *Acta Automatica Sinica*, Vol. 33, Issue 12, 2007, pp. 1252-1258.

[18]. C. J. Needham, R. D. Boyle, Performance evaluation metrics and statistics for positional tracker evaluation, *Lecture Notes in Computer Science*, Vol. 2626, 2003, pp. 278-289.

[19]. S. Muller-Schneiders, T. Jager, H. S. Loos, Performance evaluation of a real time video surveillance system, in *Proceedings of the 2nd Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS'05)*, Beijing, China, October 2005, pp. 137-144.

[20]. F. Bashir, F. Porikli, Performance evaluation of object detection and tracking systems, in *Proceedings of the 9th IEEE International Workshop on PETS*, New York, 2006, pp. 7-14.

[21]. F. Yin, D. Makris, S. Velastin, Performance evaluation of object tracking algorithms, in *Proceedings of the 10th IEEE International Workshop on Performance Evaluation of Tracking and Surveillance*, Rio de Janeiro, 2007, pp. 17-24.

[22]. Z. Zivkovic, B. Krose, A probabilistic model for an EM-like object tracking algorithm using color-histograms, in *Proceedings of the 6th IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (in connection with ECCV'04)*, Prague, 2004, pp. 73-82.

[23]. P. Guha, A. Mukerjee, K. S. Venkatesh, Efficient occlusion handling for multiple agent tracking by reasoning with surveillance event primitives, in *Proceedings of the 2nd Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS'05)*, Beijing, China, October 2005, pp. 49-56.

————————————————