A Patch-based Framework for Detecting Abnormal Activities with a PTZ Camera

Yisi Tao[1], Yuanzhe Chen[1], Weiyao Lin[1*], Xintong Han[1], Hongxiang Li[2], Zheng Lu[1]

[1] Department of Electronic Engineering, Shanghai Jiao Tong University, China (*Corresponding author)

[2] Department of Electrical and Computer Engineering, University of Louisville, USA

ABSTRACT

In this paper, a novel patch-based (PB) framework is proposed for detecting abnormal activities using a Pan-Tilt-Zoom (PTZ) camera. We first propose a new scene-patch-based (SSB) algorithm which can efficiently extract the target object's global trajectory from the PTZ camera. Furthermore, we propose an extended network-based (ENB) algorithm for detecting abnormal activities. The proposed ENB algorithm models the entire scene as a network where each node in the network corresponds to a patch of the scene and each edge between nodes corresponds to the activity correlation between the scene patches. Based on this network, a recursive training strategy is proposed to train the edge weights in the network such that abnormal activities can be effectively detected through these trained edge weights. Experimental results demonstrate the effectiveness of our proposed framework.

***Index Terms*—PTZ Camera Tracking, Event Detection, Patch-based Method**

1. INTRODUCTION

Abnormal activity detection is of increasing importance in many applications including video surveillance, human-computer interaction, and video retrieval [1-13].

Many algorithms have been proposed to detect abnormal activities [1-4, 9-13]. However, most of the existing algorithms only focus on the detection from a fixed camera while in practice many video surveillance systems require detecting activities with a moving camera such as the Pan-Tilt-Zoom (PTZ) camera where the camera keeps tracking the target object and thus wider scenes can be monitored [4-8]. In this paper, we focus on two key issues for abnormal activity detection with a PTZ camera.

1.1. Global Trajectory Extraction from a PTZ Camera

When the camera is moving with the target object, each frame captured from the camera will only include part information of the entire scene. Therefore, it is very important to recover the global trajectory of the target object to facilitate the follow-up activity detection steps [4-8].

Many algorithms have been proposed for extracting global trajectories from a PTZ camera. Some methods [4-6] pre-label or pre-detect some sets of self-calibrated points or curves over the ground plane of the scene. When the carema moves, some of these points or curves will be included in the captured frame. Based on these points or curves, the current postion of the moving object can be calculated. However, since these "pre-labeling based" methods require pre-labeling or pre-detecting steps beforehand, they are often inconvenient and less flexible in many scenarios. Other resarchers [7-8] first create a ground plane image of the entire scene, and then by matching the interest points between the current frame and the entire ground plane image, the global position of the current frame can be decided. Although these "interest point matching-based" methods show good performance, they are not applicable in many practical applications due to the high computational complexity. Furthermore, the requirement for creating the entire ground plane image also limits their applications.

1.2 Trajectory-based Abnormal Activity Detection

After extracting the global trajectory, activity detection algorithms need to be performed for analyzing the activities of the target object. Many trajectory-based algorithms have been proposed to detect abnormal activities [1-3, 9-10]. For example, Zelniker et al. [3] create global trajectories by tracking people across different cameras and detect abnormal activities if the current global trajectory deviates from the normal paths. However, this method is vulnerable to the tracking performance and it cannot work well when the tracking step is less accurate. Loy et al. [9] and Li et al. [10] segment the scene of multiple cameras into different patches and then estimate the probability distributions or graphical structures over these patches for activity detection. Although these algorithms are more robust to tracking errors, their training processes are complicated and thus still have great limitations.

In this paper, a novel patch-based (PB) framework is proposed for detecting abnormal activities using a single pan/tilt/zoom camera. The contributions of our proposed framework can be summarized as follows: (a) we propose a new scene-patch-based (SSB) algorithm for extracting the target object's global trajectory from a Pan-Tilt-Zoom camera. The proposed SSB algorithm can extract global trajectories from very simple on-the-fly information without the inconvenient pre-processing steps. Therefore, it is flexible to be applied in various scenarios. Furthermore, we also propose to segment the entire scene into patches and map the extracted trajectory into these patches accordingly. By this way, the possible trajectory extraction errors can be further reduced. (b) We propose an extended network-based (ENB) algorithm for detecting abnormal activities. The proposed ENB algorithm models the entire scene as a network where each node in the network corresponds to a patch of the scene and each edge between nodes corresponds to the activity correlation between the scene patches. Based on this network, a recursive training strategy is proposed to train the edge weights and the thresholds in the network such that abnormal activities can be effectively detected through these trained edge weights. Experimental results demonstrate the effectiveness of our proposed framework.

The rest of the paper is organized as follows: Section 2 describes the motivations of our work. Section 3 discusses the details of proposed PB framework. The experimental results are shown in Section 4 and Section 5 concludes the paper.

2. THE MOTIVATIONS

As mentioned, many of the existing global trajectory extraction algorithms require some pre-processing steps (such as pre-labeling points or creating the ground plane image), which make them inconvenient in many cases. Therefore, new algorithms are needed to minimize these pre-processing steps. We observe that since the size of the detected target object and the angle change of the camera are often available during PTZ camera tracking. If we can extract the global trajectory based on these on-the-fly information, the pre-processing steps can be effectively excluded. Besides, since the extracted global trajectory may include errors due to tracking or calibration inaccuracy, if we map the trajectory into patches (i.e., the global trajectory can be represented by the patches that the global trajectory passes, as in Fig. 1) and perform activity recognition based on these patch-based trajectories, the errors can be further reduced. Thus, in this paper, we propose a new scene-patch-based (SSB) algorithm which (a) extracts global trajectories from the on-the-fly information such as the size of the object and the angle change of the camera, and (b) maps the extracted global trajectory into patches for further improving its robustness.

Furthermore, as discussed, most of the existing trajectory-based activity recognition algorithms are either less efficient or time-consuming. Thus, it is also desirable to develop more effective activity detection algorithms. In Gao et al's work [2], a Network-Transmission-Based (NTB) algorithm is proposed which models the entire scene as a network where the nodes and the edges represent the patches and the activity-correlations between patches, respectively. With this network, human activities can be modeled as the signal transmission problem over the network where

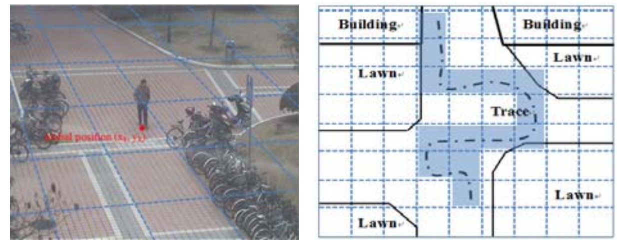


Fig. 1 The process for mapping the global trajectory into patches (Left: The red point is the calculated global position of the current person and the dashed blocks are the defined patches of the scene. Right: The curve is the extracted global trajectory. And the grey blocks are the patch-based trace by connecting the patches that the global trajectory passes).

activities with 'high transmission energies' are detected as abnormal [2]. In this algorithm, the way to calculate the activity-correlations (i.e., the edges weights of the network) is crucial to the recognition accuracy. However, the NTB algorithm has some limitations in calculating these activity-correlations because: (a) it calcuates activity-correlations only based on normal activity samples while features of the abnormal activity samples are not considered. (b) It calculates the activity-correlations in a heuristic way (i.e., taking the inverse of the crossing time between spatches). This often makes the algorithm fail to maximize the distance between normal and abnormal activity samples. Therefore, in this paper, we propose an extended network-based (ENB) algorithm for detecting abnormal activities. The proposed ENB algorithm introduces a new recursive training strategy for training the edge weights and the thresholds in the network such that abnormal activities can be effectively detected through these trained edge weights.

Based on the above discussions, we can propose a patch-based (PB) framework which uses SSB algorithm for global trajectory extraction from a PTZ camera and the ENB algorithm for abnormal activity detection. The proposed SB algorithm is described in detail in the following.

3. THE PATCH-BASED FRAMEWORK

The framework of our proposed PB framework can be described by Fig. 2.

Fig. 2 mainly includes two steps: the global trajectory extraction step (i.e., the upper dashed block named "the SSB algorithm" in Fig. 2) and the activity detection step (i.e., the lower dashed block named "the ENB algorithm" in Fig. 2).

In Fig. 2, the target object is first detected. Based on the detected object, the camera angles will be changed accordingly such that the target object is located at the center of the captured frame. By this way, the PTZ camera can suitably keep track of the target object. Then, based on the size information of the target object and the angle change information of the camera, the object's global position in the current frame can be calculated by our SSB algorithm. This global position will be further mapped into pre-defined

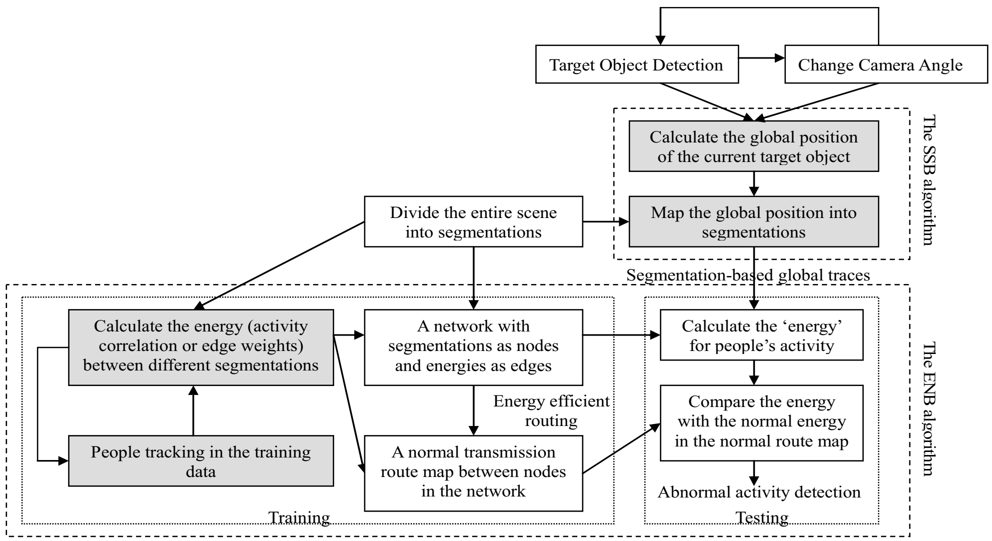


Fig. 2 The framework for the proposed SB framework.

patches for the follow-up activity recognition step. In our experiments, we use tracking-by-detection-like strategies to track and locate the object's position in the current frame. That is, we use particle filter [12] to track the target object while using object detector [11] to correct and re-initialize the object's position when necessary.

The activity detection step mainly includes two parts: the testing part and the training part (i.e., the two dotted blocks in Fig. 2). In this step, the process of target object moving among segments is viewed as the 'energy' consumed to transmit a group of packages in the network, and the abnormal activity can be detected when its 'transmission energy' is obviously larger than the normal case (i.e., for example, if a person move to an irregular patch, the 'energy' used will become large and this will be detected as an abnormal activity), by:

The current activity $R(u, q)$ is abnormal if:

$$E_c(u, q)>T_1 \text{ or } |E_c(u, q)- E_{min}(u, q)|>T_2 \qquad (1)$$

where $R(u, q)$ is the trajectory of the target object with u being the starting patch and q being the current patch of the target object. $E_c(u,q)=\sum_{(i,j)\in R(u,q)} e_{i,j}$ is the total transmission energy for the target object's activity where $e_{i,j}$ is the energy edge (or edge weights) between patches i and j [2]. $E_{min}(u,q)=\sum_{(i,j)\in R_{min}(u,q)} e_{i,j}$ is the minimum possible energy between u and q where $R_{min}(u, q)$ is the trajectory corresponding to this minimum energy. $R_{min}(u, q)$ and $E_{min}(u, q)$ can be pre-calculated in the training process by estimating all the 'energy' edges between nodes in the network and performing an energy efficient routing algorithm [2]. T_1 and T_2 are two thresholds for detecting abnormal activities. In this paper, the edge weights $e_{i,j}$ and the two thresholds are estimated by our proposed recursive training strategy.

It is noted that our PB framework has following features:

(a) After mapping the global trajectories into patches in the SSB algorithm, the same patch definition is used for detecting abnormal activities (i.e., the patch-based trajectories from SSB algorithm will be directly used as the input trajectory in the ENB algorithm). Thus, the global trajectory extraction step and the activity detection step can be perfectly integrated in our SB framework.

(b) The major difference between our ENB algorithm and the NTB algorithm [2] is that our ENB algorithm introduces a new recursive training strategy for calculating the energy weights and the thresholds. With this strategy, the performance of our NTB algorithm is obviously improved.

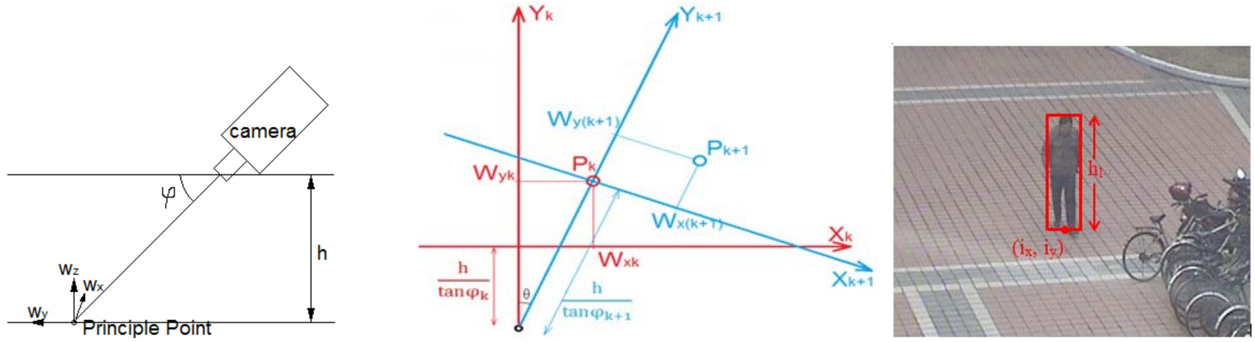
Based on the above discussions, we can see that SSB algorithm (i.e., the two upper grey blocks in Fig. 2) and the recursive training strategy (i.e., the two lower grey blocks in Fig. 2) are the key parts of our PB framework. Therefore, we will describe them in details in the following subsections.

3.1 The scene-patch-based (SSB) global trajectory extraction algorithm

In our SSB algorithm, assume that the object's global position at time k is (x_k, y_k, z_k). If we further assume that the object is standing on the ground such that the object's z coordinate value (i.e., z_k) is always 0, then the global position can be simplified into (x_k, y_k) and they can be calculated by Eqn. (2):

$$\begin{cases} x_k = \sqrt{w_{xk}^2 + \left(w_{yk} + \dfrac{h}{\tan\varphi_k}\right)^2} \cdot \cos\left(\sum_{i=1}^{k} \Delta\theta_i\right) \\ y_k = \sqrt{w_{xk}^2 + \left(w_{yk} + \dfrac{h}{\tan\varphi_k}\right)^2} \cdot \sin\left(\sum_{i=1}^{k} \Delta\theta_i\right) \end{cases} \qquad (2)$$

where (w_{xi}, w_{yi}) is the object's local coordinate in the



(a) (b) (c)

Fig. 3 (a) The relationship of (w_{xi}, w_{yi}), h, and φ_i in a camera's side view [5]. (b) The basic idea for deriving Eqn. (2). (c) The bottom center pixel location of the target object in the image (i_x, i_y) and the height of the detected object in the image h_l.

coordinate system of the camera at time i. Note that the camera's coordinate system will change when the camera moves. h is the height of the camera and it is assumed to be fixed in our algorithm. φ_i is the tilt angle of the camera at time i. $\triangle\theta_i$ is the camera's entire angle change at time i. The difference between φ_i and $\triangle\theta_i$ will be discussed later. The relationship of (w_{xi}, w_{yi}), h, φ_i and the basic idea for deriving Eqn. (2) are described by Fig. 3 (a)-(b), respectively.

In Fig. 3 (b), the coordinate named "X_k-Y_k" is the camera's coordinate system at time k while the coordinate named "X_{k+1}-Y_{k+1}" represents the camera's coordinate system for time $k+1$. P_k is the position of the detected target object at time k. The local coordinate of P_k in "X_k-Y_k" system is (w_{xk}, w_{yk}). And $D(P_k)=\sqrt{w_{xk}^2+\left(w_{yk}+\frac{h}{tan\varphi_k}\right)^2}$ is the distance between P_k to the camera. Based on our camera moving strategy, the camera is moved according to P_k such that P_k's coordinate will be (0, 0) in the new coordinate system (i.e., the blue "X_{k+1}-Y_{k+1}" system in Fig. 3 (b)). By this way, we can guarantee that the detected object is always located at the center area of the camera view. Based on the above strategy, the camera's angle change $\triangle\theta_i$ at time i can be derived by:

$$\Delta\theta_i = arctan\left(\frac{w_{xi}}{w_{yi}+\frac{h}{tan\,\varphi_i}}\right) \tag{3}$$

Based on Eqn. (3), the camera's absolute angle can be calculated by accumulating the angle change $\triangle\theta_i$ over all times. After that, the object's global position (x_k, y_k) can be calculated based on the distance $D(P_k)$ and the camera's absolute angle, as in Eqn. (2).

Furthermore, the tilt angle φ_k and the object's local coordinate (w_{xk}, w_{yk}) in Eqn. (2) can be calculated by Eqns. (4) and (5), respectively.

$$\varphi_k = \varphi_0 + \sum_{i=1}^{k} \Delta\varphi_i \tag{4}$$

where φ_0 is the initial tilt angle of the camera and $\triangle\varphi_i$ is the change of the tilt angle at time i. It should be noted that $\triangle\varphi_i$ differs from $\triangle\theta_k$ in Eqn. (3) since $\triangle\varphi_i$ only corresponds to the tilt angle change while $\triangle\theta_k$ corresponds to the entire angle change including both the tilt and the pan angles.

$$\begin{cases} w_x = \dfrac{sec\varphi}{2\cdot h_l\cdot i_y}\Big(-hh_l i_x + h_0 i_x i_y + h_0 h_l i_x sin^2\varphi \\ \qquad \pm i_x\sqrt{h_0^2 i_y (h_l - i_y) sin^2 2\varphi + (hh_l + h_0 i_y cos2\varphi - h_0 h_l sin^2\varphi)^2}\,\Big) \\ w_y = \dfrac{w_x\cdot i_y}{i_x\cdot sin\varphi} \end{cases} \tag{5}$$

where (i_x, i_y) is the bottom center pixel location of the target object in the image, and (i_x+$\triangle i_x$, i_y+h_l) is the top center pixel location of the target object in the image, as in Fig. 3 (c). h_l is the height of the detected object in the image and $\triangle i_x$ is the possible deviation of the top center pixel. h_0 is the estimated average height of the target person. The results in Eqn. (5) are derived from the calibration models [5], as in Eqn. (6-7).

$$\begin{bmatrix} i_x \\ i_y \\ 1 \end{bmatrix} = \varepsilon \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & fsin\varphi & fcos\varphi & 0 \\ 0 & -cos\varphi & sin\varphi & -h/sin\varphi \end{bmatrix} \cdot \begin{bmatrix} w_x \\ w_y \\ 0 \\ 1 \end{bmatrix} \tag{6}$$

$$\begin{bmatrix} i_x + \Delta i_x \\ i_y + h_l \\ 1 \end{bmatrix} = \varepsilon_2 \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & fsin\varphi & fcos\varphi & 0 \\ 0 & -cos\varphi & sin\varphi & -h/sin\varphi \end{bmatrix} \cdot \begin{bmatrix} w_x \\ w_y \\ h_0 \\ 1 \end{bmatrix} \tag{7}$$

From Eqn. (2) and (4)-(5), we can see that our proposed global trajectory extraction algorithm can extract the object's global location only based on the on-the-fly information (i.e., i_x, $\triangle i_x$, i_y, h_l, and $\triangle\varphi_i$). Thus, it can efficiently avoid the inconvenient pre-processing steps and is flexible in various scenarios.

After calculating the global position of the current object, the calculated global position will be further mapped into patches by Eqn. (8):

$$q_k = M\left(F\left(x_k, y_k\right)\right) \tag{8}$$

where $M(\cdot)$ is the mapping function that maps the object's global position (x_k, y_k) to its corresponding patch q_k. $F(\cdot)$ is the filtering function to filter out the unreliable global positions. In our algorithm, the object detector [11] is introduced such that a global position is viewed as unreliable if this position is far from the one estimated by the object detector. From Eqn. (8), we can see that our SSB algorithm improves the trajectory extraction robustness in two aspects: (a) It introduces a filtering functionality (i.e., the object detector) to filter out the unreliable global positions. (b) It maps the global positions into patches to allow the possible global trajectory extraction inaccuracies (i.e., even if the extracted global trajectory is deviated from the ground truth, their patch-based global trajectories are the same as long as they belong to the same patch.

3.2 The Recursive Training Strategy for the ENB Algorithm

As mentioned, the main purpose of our recursive training strategy is to provide a more efficient way for achieving the 'energy' edges $e_{i,j}$ and the thresholds (T_1 and T_2) in Eqn. (1) for more efficient detection performance. The energy edges $e_{i,j}$ between patches i and j can be calculated by:

$$e_{i,j} = 1 \Big/ \sum_k tw_k(i,j) \tag{9}$$

where $tw_k(i,j)$ is the impact weight of k-th training trajectory. $tw_k(i,j)$ is calculated in an recursive way by:

$$tw_k^{n+1}(i,j) = \begin{cases} tw_k^n & \text{if trace } k \text{ is correctly recognized} \\ tw_k^n\left(1+\dfrac{E_k^n - T_i^n}{E_k^n}\right) & \text{if } k \text{ is a normal miss} \\ tw_k^n\left(1-\dfrac{T_i^n - E_k^n}{2\cdot T_i^n}\right) & \text{if } k \text{ is an abnormal miss} \end{cases} \tag{10}$$

where $E_k^n = \sum_{(i,j)\in R(k)} e_{i,j}^n$ is the energy for trajectory k calculated by the energy edge values in n-th iteration (i.e., $e^n_{i,j}$). T_i^n ($i=1$, 2) is the threshold in n-th iteration. The selection of i depends on which threshold is used for detection in Eqn. (1). Before iteration, each $tw_k(i, j)$ is initialized to be 1 if the k-th training trajectory transposes patches i and j, or initialized to be 0 otherwise. From Eqn. (10), we can see that the impact weight of trajectory k will be increased if the normal activity k is miss detected, thus the corresponding energy $e_{i,j}$ will be decreased to better facilitate the normal activity recognition. On the contrary, if the abnormal activity k is miss detected, the impact weight of k will be decreased (or the energy $e_{i,j}$ is increased) for increasing the ability to detect abnormal activities. Similarly, in each iteration, the thresholds T_1 and T_2 are calculated by:

$$T_i^n = argmin\left({err_{nor}^n}^2 + {err_{ab}^n}^2\right) \tag{11}$$

where err^n_{nor} and err^n_{ab} are the miss rates [1-2] for the normal and abnormal activities in the n-th iteration, respectively.

From Eqn. (10) and (11), we can see that our recursive training strategy includes both the normal and abnormal trajectories in the training data and recursively updates the parameters according to the detection performance. By this way, the parameters can be more efficiently trained for better recognizing the results, as will be shown in the next section.

4. EXPERIMENTAL RESULTS

In this section, we show experimental results for our SB framework. We use a MG-TK 3518 Pan-Tilt-Zoom (PTZ) camera for tracking the target objects. 400 video sequences are created. For each sequence, we manually label the objects' ground truth global trajectories in order for evaluating the algorithms' performance in global trajectory extraction. Furthermore, in order to evaluate the algorithms' performance in detecting abnormal activities, we further label each trajectory as normal activity (i.e., people follow regular trajectories) or abnormal activity (i.e., people follow irregular trajectories such as moving back-and-force, move onto the lawn, etc).

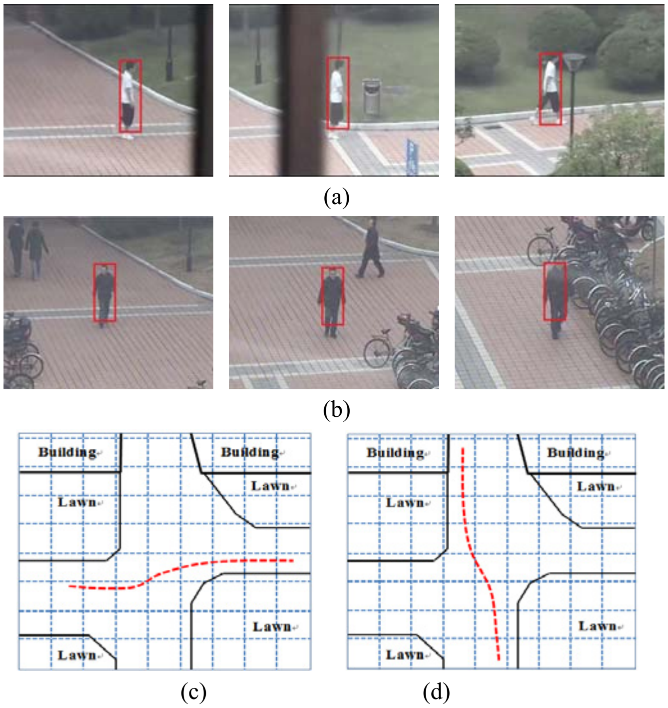
Also, it should be noted that when training the weights and thresholds in Eqn. (10) and (11), the training data is separated into training sets and validation sets [1]. Due to the limited space, only parts of the experimental results are shown in this paper. Fig. 4 shows some example sequences and the extracted global trajectory results by using our PTZ SSB framework. Table 1 compares the global trajectory extraction accuracy and the time complexity for different methods. When calculating global trajectory accuracy, we map all the global trajectories into patches and compare the resulting patch-based trajectories with the ground truth.

Table 1 Comparison for different global trajectory extraction methods

	Pre-labeling method [6]	Interest point matching method [7-8]	**SSB**
Accuracy (%)	91.1	94.8	**94.2**
Time per frame (ms)	45.7	961.9	**39.0**

From Table 1, we can see that although the interest point matching-based algorithm [7-8] can produce good results, its complexity is high, e.g., more than 20 times higher than that of [6] or our SSB. Thus, it restricts its applications. The accuracy of the pre-labeling based method [6] is lower than the other two since the pre-labeled points are often blocked by the target object. It makes this method not applicable in many applications. Compared to these, our SSB algorithm is more flexible by only using the simple on-the-fly information for trajectory extraction. As a summary, our SSB algorithm achieves good performance in both the accuracy and time complexity.

We compare the activity detection results for three



(a) (b) (c) (d)

Fig. 4 (a)-(b): Two results by our moving camera tracking framework. (c)-(d): The global trajectories of the examples in (a)-(b) created by our SSB algorithm.

methods, as shown in Table 2: (a) The Trajectory-Similarity Based (TSB) method [3], (b) The Network-Transmission-Based (NTB) method [2], (c) Our proposed Extended Network-Based (ENB) algorithm. In order to have fair comparisons, the trajectories used in the TSB method are the ground-truth trajectories, while the trajectories used in the NTB and ENB methods are the patch-based trajectories extracted by our SSB algorithm.

Table 2 Miss rates and TER comparisons of abnormal activity detection

	75% training-25% testing			90%training-10%testing		
	TSB	NTB	**ENB**	TSB	NTB	**ENB**
Normal Miss (%)	15.1	10.7	**3.6**	16.8	9.1	**0**
Abnormal Miss (%)	32.8	15.4	**15.4**	27.9	20.0	**16.0**
TER (%)	27.6	13.0	**9.3**	22.4	16.7	**11.1**

In Table 2, three error rates are compared: Normal Miss (i.e., the number of miss detected normal activities divided by the total number of normal activities), Abnormal Miss (i.e., the number of miss detected abnormal activities divided by the total number of abnormal activities), and Total Error Rate (TER, the total number of miss detection in both activities divided by the total number of trajectories in both activities). From Table 2, we can see that the performance of our ENB algorithm is better than the other methods since the recursive training strategy in the ENB method can adjust the energy weights and thresholds properly to maximize the distance between normal and abnormal samples. Thus the final detection performances can be improved. Moreover, we can also see that although TSB utilizes the ground truth trajectory data, its performance is poorer than the patch-based methods (NTB and ENB). This further implies that algorithms based on the patch framework are more effective in detecting trajectory-based activities.

5. CONCLUSION

In this paper, we propose a novel patch-based framework for abnormal activity recognition using a PTZ camera. In this framework, we first propose a new scene-patch-based (SSB) algorithm which can efficiently extract the target object's global trajectories from the PTZ camera. Furthermore, we propose an extended network-based (ENB) algorithm which introduces a recursive training strategy for training the parameters of the network-based model. Thus, the activity detection results can be efficiently improved. Experimental results demonstrate the effectiveness of our proposed framework.

5. ACKNOWLEDGEMENTS

This paper is supported in part by the following grants: National Science Foundation of China grants (61001146, 61025005, 61103124), SMC grant of SJTU, Chinese national 973 grants (2010CB7314016), and the Open Project Program of the National Laboratory of Pattern Recognition (NLPR).

REFERENCES

[1] W. Lin, M.-T. Sun, R. Poovendran and Z. Zhang, "Activity Reco gnition using a Combination of Category Components and Local Models for Video Surveillance," *IEEE Trans. Circuits Syst. Video Technol.*, 2008.

[2] H. Gao, W. Lin, X. Yang, H. Li, N. Xu, J. Xie, Y. Li, "A new network-based algorithm for multi-camera abnormal activity detection," *ISCAS*, 2011.

[3] E. E. Zelniker, S. Gong and T. Xiang, "Global abnormal behaviour detection using a network of CCTV cameras," *Int'l Workshop. Visual Surveillance*, 2008.

[4] F. Lv, T. Zhao, and R. Nevatia, "Self calibration of a camera from a video of a walking human," *Int'l Conf. Pattern Recognition*, 2002.

[5] H. Dinh, H. Tang, E. Kwon, "Camera Calibration for Roundabout Traffic Scenes", *Transportation Research Board Annual Meeting*, 2011.

[6] R. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision," *Cambridge University Press*, 2nd ed. 2003.

[7] L. G. Mirisola and J. Dias, "Tracking from a Moving Camera with Attitude Estimates," *ICR*, 2008.

[8] L. G. Mirisola, J. Dias, and A. T Almeida, "Trajectory recovery and 3d mapping from rotation-compensated imagery for an airship," *IEEE Int'l Conf. Intelligent Robots and Systems,* 2007.

[9] C. C. Loy, T. Xiang and S. Gong. "Modelling activity global temporal dependencies using time delayed probabilistic graphical model," *ICCV*, 2009.

[10] J. Li, S. Gong, and T. Xiang, "Discovering multi-camera behaviour correlations for on-the-fly global prediction and anomaly detection," *Int'l Workshop. Visual Surveillance*, 2009.

[11] J. Wu, C. Geyer and J. Rehg, "Real-Time Human Detection Using Contour Cues" *ICRA*, 2011.

[12] R. Hess and A. Fern, "Discriminatively Trained Particle Filters for Complex Multi-Object Tracking," *CVPR*, 2009.

[13] F. Jiang, J. Yuan, S. Tsaftaris and A. Katsaggelos, "Anomalous Video Event Detection Using Spatiotemporal Context," *Computer Vision and Image Understanding*, vol. 115, no. 3, pp. 323-333, 2011.