Proceedings of IOE Graduate Conference, 2019-Summer
Peer Reviewed
Year: 2019 Month: May Volume: 6
ISSN: 2350-8914 (Online), 2350-8906 (Print)

Abnormal Motion Pattern Detection in Surveillance Video Sequences by Clustering Approach

Himal Acharya [a], Basanta Joshi [b]

[a, b] *Department of Electronics and Computer Engineering, Pulchowk Campus, IOE, Tribhuvan University, Nepal*
[a] *Nepal College of Information Technology, Lalitpur, Nepal*
Corresponding Email: [a] himalacharya@gmail.com

Abstract
Surveillance cameras are widely being used in public places for security and monitoring purposes. Detecting abnormal motion pattern from surveillance video sequences is challenging task. Most of the existing methods are based on supervised technique. Supervised method groups feature points into normal and abnormal motion pattern using classifier. But anomalous event are contextual so this paper focuses on unsupervised learning method of finding abnormal motion pattern. Contextual abnormality can not be detected by supervised method in every surveillance video sequences. This proposed approach works without the need of training phase. By extracting trajectory features by dense optical flow, speed of moving objects are taken into consideration for unsupervised motion pattern in video sequences. K-means clustering approach is simple to implement and computationally efficient. By applying such clustering method, dominant motion group and anomalous group are well separated. Experimental results demonstrate this proposed approach outperforms the state-of-art approaches on standard dataset.

Keywords
Clustering, Motion Pattern, Surveillance Video

1. Introduction

Video surveillance system is being increased in recent years for public security concerns and traffic management. Humans are able to learn, interpret and predict complex interactions and distinguish scenes. The huge amount of data are being generated from existing surveillance systems which makes difficult for human to classify complex events or manually label a video [1].

When there is large amount of data available, vision algorithms need to be implemented for anomaly detection. There are many definitions on term 'anomaly'. Anomaly can be defined as the low probability event in the video sequences or images. So, it is difficult to collect abnormal events in real world. Normal / Abnormal events can be defined in abstract way. The definition of anomaly is contextual. Any event that is normal in one context may be abnormal in another context. For example, man walking in left direction is abnormal if all man walks in right direction.

For increasing security concerns, surveillance cameras are being used popularly. Detection of abnormal events in video sequences is very challenging. The visual attention of human brain is limited and declines after some period. So, there is a problem in manual annotation and analysis of large video sequences. The frequency occurrence of abnormal events is very rare. Due to scarcity and variations in abnormal events, it is hard to find anomalies at training time. Normal and abnormal events definition is vague and cannot be precisely defined as it highly depends on the spatio-temporal context. Due to challenging nature and its application, abnormal detection is seeking considerable attention in both academia and industry. Many supervised methods are employed to find abnormal detection in sequences. Supervised setting requires manually specifying videos containing abnormal frames. Although these supervised methods find a success in detection ; still it's time consuming for human beings to specify normality in real life. It is unrealistic to priorly know every possible normal event. So, unsupervised settings are employed to find abnormal events in frames.

In this paper, real-time method for automatic anomaly detection in surveillance video sequences is proposed which works without training sequences. Pedestrian motion is considered as dominant motion. Events showing motion pattern different from dominant motion are considered as abnormal events. The two main contributions of the proposed approach are:

1. The use of dense trajectories to detect fast and irregular motion in online abnormal situations.
2. An unsupervised and computationally efficient algorithm for abnormality detection with motion speed of dense trajectories.

2. Related Work

Most of anomaly detection works adopt supervised method. Early works in paper [2, 3] detect abnormality by extracting high-level features such as trajectory from object tracking. Such method fails when there is object shading or blurring occurs frequently. The author in paper [4] constructs a dictionary using only normal event training samples in the training phase. In testing phase, this method uses new observation's reconstruction error as a metric for anomaly detection.

The authors in paper [5] propose online real time crowd behavior detection using background modeling and KLT feature extraction. It uses activity map to analyze the trends of image entropy (based on two consecutive frames) and temporal occupancy variation(TOV). This can identify only a few pixels due to changing density of objects in crowded scene. The quality as well as quantities of trajectories using KLT tracker is not sufficient [6]. KLT trajectories based on sparse optical flow are not robust to fast irregular motions. This creates problem when there is dense abnormality. If dense trajectories are adopted, then it can cover motion information (changes in flow field) in videos well. The author in paper [6] also mentions that the efficient solution implemented can remove camera motion by computing motion boundaries descriptors.

The author in paper [7] uses logistic regression to differentiate two consecutive set of frames and abnormality is detected by differentiation from previous frame. The authors in paper [8] use an unmasking scheme which classifies two consecutive set of frames as in paper [7] and removes the most weighted feature. These two papers are based on detecting drastic changes between normal and abnormal events in local context and assume spare abnormality. When the abnormality is relatively high dense, performance of methods used in paper [8, 7] degrades.

The author in paper [9] uses online weigthed clustering in Region of Interest to obtain Adaptive Multi-scale Histogram Optical Flow (AMHOF) features. This paper adopted weighted clustering and simplified Multi-Target Tracker (MTT) algorithm. This algorithm can track and detect only slow changes. Clusters with normal AMHOF are considered as normal frames and frames with above threshold are considered as abnormal frames. After online weighted clustering results, Kalman filter tracking is adopted to detect missing anomaly detection in frame. But also, when there is fast irregular motions, this optical flow features cannot detect those irregular motions. This method cannot be optimized to real time because it uses tracking algorithm with clustering result.

3. Methodology

In this paper, k-means clustering is applied on dense trajectories of surveillance video sequences to identify normal and abnormal event in pedestrian surveillance video sequences. Before abnormal event detection, preprocessing of video sequence is needed.

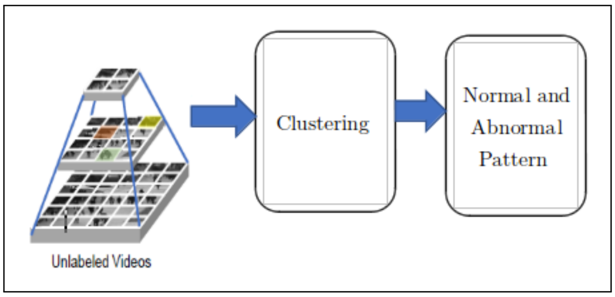

Figure 1: Block Diagram of unsupervised abnormal motion pattern detection approach

At first video, video features are converted to gray scale images and resized to multiple scales to cover image sequence with different sizes. Motion patterns can be grouped by motion speed. For example, carts, skaters, cyclists move at higher speed or opposite direction. Such activities are not the dominant motion in the whole video sequences.

Dense optical flow algorithms are used in place of sparse optical flow method to avoid feature point detection. For dense optical flow features, Farneback dense optical flow is used[10]. The neighborhood around each pixel is estimated as a quadratic function using polynomial expansion.

$$f(x) = x^T A x + b^T x + c \tag{1}$$

where x is position ; A,b,c are calculated from weighted least square fit. This weighting has two matrix component applicability matrix and certainty matrix. Applicability matrix determines the relative weight of points in the neighborhood based on their position in neighborhood. Certainty matrix ensures the certainty of 1 around all center values and reduced certainty near the outside to improve results near the borders. For second frame displaced by d, exact quadratic polynomial is given by:

$$f_2(x) = x^T A_2 x + b_2^T x + c_2 \tag{2}$$

while comparing with first frame coefficients.

$$A_2 = A_1 \tag{3}$$

$$b_2 = b + 1 - 2A_1 d \tag{4}$$

$$c_2 = d^t A_1 d - b_1^T d + c_1 \tag{5}$$

Finally, displacement is calculated as:

$$d = -\frac{1}{2} A_1^{-1} (b_2 - b_1) \tag{6}$$

For every two consecutive frame, optical flow field is calculated for every pixel. Motion speed is used to obtain a feature vector in the video event representation. Distinctive features points are extracted by using dense trajectory features as proposed in paper [6]. Feature points are tracked in dense optical flow for a temporal window of length L frames. Each point at frame t is tracked to next frame t+1 by median filtering in dense optical flow field. flow field.

$$P_{t+1} = (x_{t+1}, y_{t+1}) = (x_t, y_t) + (M * \omega)|_{(\bar{x}_t, \bar{y}_t)} \tag{7}$$

where M is the median filtering kernel and $(\bar{x}_t, \bar{y}_t)$ is the rounded position of (x_t, y_t). After the dense optical flow field is calculated, feature points of subsequent frames are concatenated to form a trajectory of length L: $(P_t, P_{t+1}, P_{t+2}, ..., P_{t+L})$. To avoid drifting, in every trajectory length of L frames samples new feature points to replace previous ones. Here trajectories are represented at multiple temporal scales, in order to recognize activities with different speeds. When a feature point is sampled, then feature point having eigenvalue of autocorrelation matrix below threshold is not included in tracking process.

With classical physics, speed is calculated as the derivative of the position. All feature points of trajectory length L frames have x and y gradients. Taking the magnitude of the x and y gradient, speed is calculated [11]. But, here trajectory of L frames is considered so, speed of feature point is calculated of dense trajectories pixel for temporal window of L frames.

The speed s_i of a feature point p_i is calculated as

$$s_i = \frac{||x_i^{t+L} - x_i^t||}{L} \tag{8}$$

For finding similar motion pattern in video sequences, clustering approach is adopted. Proposed approach for abnormal activity detection is based on assumption that normal motion is smooth and gradual whereas abnormal activities will produce significant and rapid changes in motion[12].

Anomalous motion pattern grouping begins by calculating dense optical flow and grouping similar motion pattern. All extracted feature points' speed features are divided into number of equal intervals. For all interval, interval is characterized by normalized mean speed of that particular interval group and normalized number of features of that particular interval group. Then , k-means clustering is applied to identify normal and abnormal motion patterns. After clustering, anomalous events are localized in video sequences.

4. Experiments and Discussion

This paper work focuses on pedestrian activities analyzed from surveillance video sequences. Here, commonly used dataset UCSD Peds2 is used. This dataset was created by stationary camera and publicly available on internet. UCSD Peds2 Dataset contains 16 training videos and 12 testing videos. Anomalies like skaters, cars , bike are present in testing videos.

Figure 3: Normal Event of UCSD Ped2 at frame 160

Figure 2: Normal Event of UCSD Ped2 at frame 27

Figure 4: Abnormal Event of UCSD Ped2: truck

With the help of dense optical flow, dense trajectories for temporal window of 15 frames of a feature point are extracted. Using classical physics, speed is calculated on Test Video 4 of UCSD Ped2 dataset.

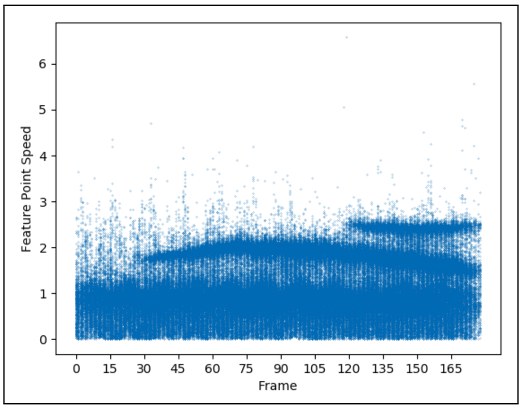

Figure 5: Feature point speed calculation for Test Video 4

In Figure 5, all features points speed are plotted. But, the research work is concerned with temporal window of length 15,so speed is calculated with the help of 15^{th} trajectory point and 1^{st} trajectory point.

Using speed formula, a new figure is plotted which gives clear overview of feature point speed.

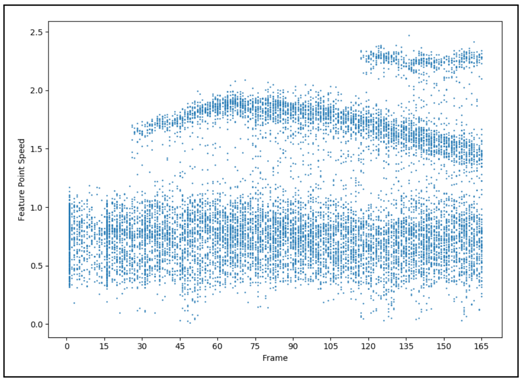

Figure 6: Feature point speed calculation for temporal window of 15 frames: Test Video 4

In Figure 6, there is large group of similar patterns in bottom part of figure. In centre part and top part, there are other two similar patterns of motion. Pedestrian motion is considered as dominant motion with largest size.The motion pattern of dominant group is considered as normal group and other pattern are considered as abnormal group. At first, speed of feature points are divided into 10 intervals of equal width. All the dense trajectories feature speed are kept in 10 interval ranging from minimum speed to maximum speed. For each interval group, normalization group is obtained by dividing each interval group by largest mean speed of interval group and largest number of frames.

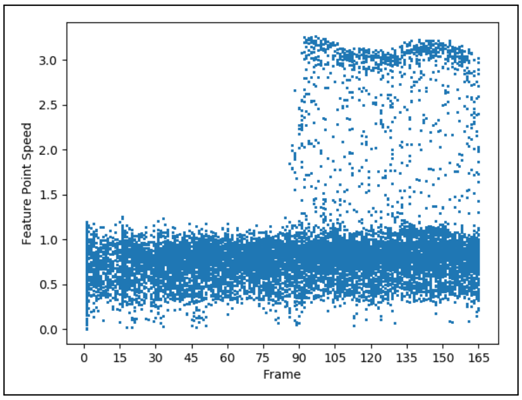

Figure 7: Feature point speed calculation for temporal window of 15 frames: Test Video 2

In Figure 7, there is large group of similar patterns in bottom part of figure. In top part, there is other pattern of motion.

After dividing the similar speed into 10 interval group, k-means clustering is done to separate into normal and abnormal pattern. K-means clustering is adopted to obtain 2 clusters: normal cluster (dominant cluster), abnormal cluster. Different motion patterns are exhibited by pedestrian, bike , truck.

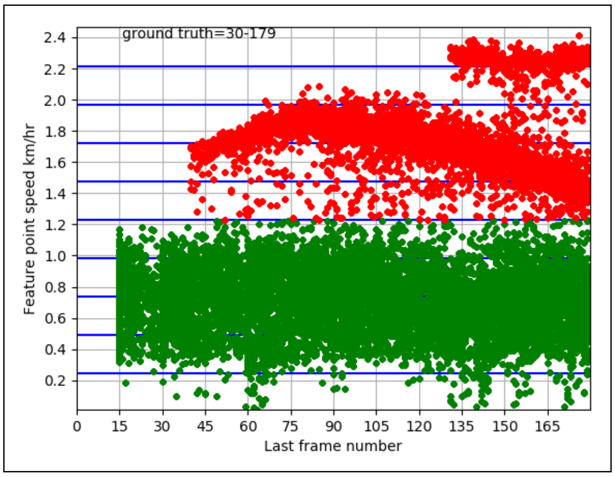

Figure 8: Different Motion pattern of Test Video 4

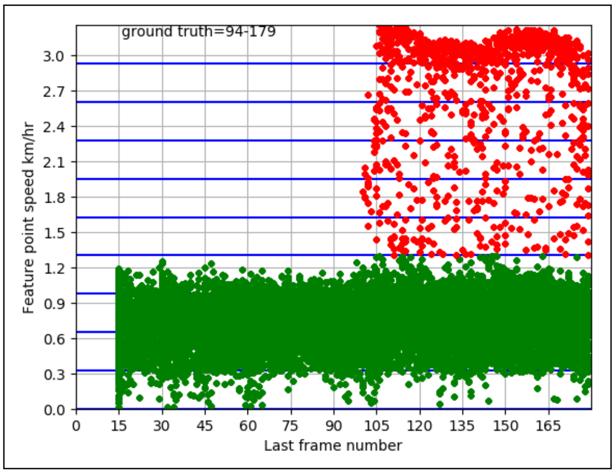

Figure 9: Different Motion pattern of Test Video 2

In Figure 8 and 9, there are two cluster groups. The dominant group (normal group) are represented by green color. Abnormal group is represented by red color. In Test Video 4, two abnormal events truck and bike moves in pedestrian walkway. So, two red color patterns (bike and truck) are classified as abnormal group. Similarly, in Test Video 2, bike is considered as abnormal event as provided in ground truth annotation of UCSD Ped2 dataset. Bike's dense trajectories speed are classified as abnormal events and marked by red color. To every test video, this algorithm gives two clusters: normal group and abnormal event group.

5. Evaluation

With the help of ground truth annotation, feature point's motion pattern are compared. This paper quantitatively evaluates the performance of proposed approach in UCSD Ped2 dataset. Frame-level evaluation is performed whether the frame contains abnormal event or not. In this test video, the ground truth annotation mentions that there are abnormalities starting from frame 30 to 179. But the result in figure 7 shows dissimilar motion pattern starts from frame 27 to last frame. Accuracy in this test video in comparison with ground truth annotation is 87.2%. Similarly for test video 2, ground truth annotation shows dissimilar pattern begins from 94^{th} frame to 179^{th} frame. The result in figure 8 shows dissimilar motion pattern starts from frame 87 to frame 169. Accuracy in this test video in comparison with ground truth annotation is 85.5%.

6. Conclusion

In this paper, a real-time anomalous motion pattern detection algorithm is applied for video sequences. Rather than considering whole scenario of image sequence, only moving object trajectories are considered. So, dense optical flow is utilized to detect dense trajectories in video sequences. By clustering approach, motion patterns with speed attribute are observed to detect normal and abnormal events. This framework is employed in UCSD Ped2 dataset and achieved better results.

Acknowledgments

The authors are grateful to Nepal College of Information Technology (NCIT) for financial support for conference registration.

References

[1] C. Brax, L. Niklasson, and M. Smedberg. Finding behavioural anomalies in public areas using video surveillance data. In *2008 11th International Conference on Information Fusion*, pages 1–8, June 2008.

[2] Zhouyu Fu, Weiming Hu, and Tieniu Tan. Similarity based vehicle trajectory clustering and anomaly detection. In *IEEE International Conference on Image Processing 2005*. IEEE, 2005.

[3] Arslan Basharat, Alexei Gritai, and Mubarak Shah. Learning object motion patterns for anomaly detection and improved object detection. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, jun 2008.

[4] Yang Cong, Junsong Yuan, and Ji Liu. Sparse reconstruction cost for abnormal event detection. In *CVPR 2011*. IEEE, jun 2011.

[5] Andrea Pennisi, Domenico D. Bloisi, and Luca Iocchi. Online real-time crowd behavior detection in video sequences. *Computer Vision and Image Understanding*, 144:166–176, mar 2016.

[6] Heng Wang, Alexander Klaser, Cordelia Schmid, and Cheng-Lin Liu. Action recognition by dense trajectories. In *CVPR 2011*. IEEE, jun 2011.

[7] Allison Del Giorno, J. Andrew Bagnell, and Martial Hebert. A discriminative framework for anomaly detection in large videos. 2016.

[8] Radu Tudor Ionescu, Sorina Smeureanu, Bogdan Alexe, and Marius Popescu. Unmasking the abnormal events in video. 2017.

[9] Hanhe Lin, Jeremiah D. Deng, Brendon J. Woodford, and Ahmad Shahi. Online weighted clustering for real-time abnormal event detection in video surveillance. In *Proceedings of the 2016 ACM on Multimedia Conference - MM*. ACM Press, 2016.

[10] Gunnar Farnebäck. Two-frame motion estimation based on polynomial expansion. In *Proceedings of the 13th Scandinavian Conference on Image Analysis*, SCIA'03, pages 363–370, Berlin, Heidelberg, 2003. Springer-Verlag.

[11] Todd Rowland. Velocity vector. From Mathworld–A Wolfram Web Resource.

[12] Jason M. Grant. *ANALYSIS OF CROWD BEHAVIOR BASED ON OPTICAL FLOW: DETECTION, CLASSIFICATION, AND CLUSTERING*. PhD thesis, University of Notre Dame, 2018.