International Journal of Engineering & Technology, 7 (3.34) (2018) 156-159

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Optical Flow Approach Followed by SVM Classification Model to Recognize Abnormal Behavior of a Crowd

Basavaraj G M[1], Dr. Ashok Kusagur[2]

[1]*Associate professor, Dept. of ECE, Nagarjuna College of Engineering and Technology Bengaluru, India*
[2]*Associate Professor, Dept Electrical and Electronics Engg. UBDT College of Engineering, Davangere, Karnataka India,*
**Corresponding author E-mail: baswaraj832@gmail.com, ashok.kusagur@gmail.com*

Abstract

A many of researches have been carried out in the field of the crowd behavior recognition system. Recognizing crowd behavior in videos is most challenging and occlusions because of irregular human movement. This paper gives an overview of optical flow model along with the SVM (Support Vector Machine) classification model. This proposed approach evaluates sudden changes in motion of an event and classifies that event to a category: Normal and Abnormal. Geometric means of location, direction, and displacement of the feature points of each frame are estimated. Harris corner Detector is used in each frame for tracking a set of feature points. Proposed approach is very effective in real time scenario like public places where security is most important. After analyzing result ROC curve (receiver operating characteristics) is plotted which gives classification accuracy. We also presented frame level comparison with Ground truth and social force model (SFM) techniques. Our proposed approach is giving a promising result compare to all state of art methods.

Keywords: *Optical flow descriptor, Motion map, Harris corner detector, SVM classifier, ROC curve*

1. Introduction

The crowd is often seen at airport, railway station, shopping mall, multiplex, hotel, museums hospitals and other public places. It is growing day by day because of increasing population rapidly. Security of people in these places is very crucial. This topic is achieving more and more attention by researchers. Many challenges come when we talk about the understanding of a crowded event. Group level relationship, complex interaction and various semantics are challenges in crowd analysis.Recognition of crowd event involves multiple agents, which is described in most of previously paper. Motion pattern of individual or group defines the crowded event. Some well-known existing methods like particle flow [1] [2], optical flow [3] and local gradients methods simulate the crowd flow instead of tracking the individual object. Implementing this type of approaches are always related to the macroscopic model. Macroscopic models are not close to the behavior semantic. By implementing group structure [4], particle trajectory [5] [6], energy potentials [7], dense complex movement can also be captured. Group level architecture also offers a macroscopic interpretation of the crowd which is good for understand the crowd behavior.

In this paper, we are proposing an approach based on optical flow followed by Support Vector Machine classification model. The combined approach gives a good result in terms of classification accuracy. It is powerful enough to match or exceed the performance of state of art method.UMN dataset scenario is taken for experimental evaluation. Optical flow gives characteristics of a crowd and based on this characteristics SVM classifies the behavior of the crowd. In initial step motion the heat map is computed, which states the region of motion activity. Based on heat map points of interest is extracted. Once the point of interest is calculated, we start evaluating the optical flow pattern. Optical flow pattern gives a feature vector in the form of direction and distance parameter.

In [8] author describes the measurement of optical flow through the help of histogram representation, then based on threshold value we predict the abnormalities in the video. In [9] texture dynamic model is applied to model the appearance of the crowded scene. This method can explicitly detect the spatial and temporal anomalies. In [10] Spatio-temporal gradients are modeled with HMM model to detect irregularities in a densely crowded scene. Particle advection schemes [11] [12] makes a rectangular grid of particles for each frame and detect abnormalities using underlying motion. In this approach, each particle is supposed to be individual, so it overcomes the restriction of the tracking people in high crowd density. [11] describes a Social Force Model (SFM) which is carried out for analysis of crowd behavior.

In section 2 and 3, we introduced motion map and feature extraction process. We also included an overall block diagram of our proposed algorithm. An SVM classifier is described in section 4 and Section 5 describes the steps how our proposed algorithm works. In section 6 results are performed and finally concluding remarks.in section 7.

2. Motion Map Generation

Motion map is a 2D histogram representation of data, which shows the main region of motion activity. This is built from the accumulation of binary blobs of the objects which are in motion in the frame.Background subtraction method is implemented in detecting the region of motion in a frame [13].

The obtained motion map is responsible to define the region of interest (RoI), which is required for further processing such as feature detection. Detection of heat map improves result accuracy and reduces the processing time which is very important in real time application. We extract heat map for all the frames of the taken scenario.

3. Feature Extraction

For each input frame of the taken scenario, a set of the interested point is extracted. Usually, we take a mask to define these points of interest. This mask is got from motion map of the frame. Our approach implements Harris corner detector [14] for extracting interest points. After getting points of interest in one frame, we start finding these points on next frame using optical flow pattern. For tracking purpose, we applied Kanade-Lucas-Tomasi feature tracker [10, 14].

After matching features points between individual frames we get a set of vectors.

$$O = \{O_1, \dots O_N\} | O_i = (x_i, y_i, d_i, \theta_i)\} \quad (1)$$

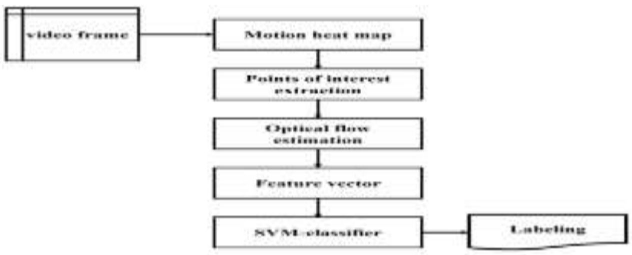
Where x_i *and* y_i are coordinates of feature$i$,$d_i$is the distance between matched feature points of two consecutive frame.$\theta_i$is the motion direction between two matched points of consecutive frames.

$$d_i = \sqrt{(q_{x_i} - p_{x_i})^2 + (q_{y_i} - p_{y_i})^2} \quad (2)$$

$$\theta_i = \operatorname{atan}\left(\frac{q_{y_i} - p_{y_i}}{q_{x_i} - p_{x_i}}\right) \quad (3)$$

Here p(x_i,y_i) and q(x_i,y_i) are location of interest points moving from two consecutive frame.

.In normal frames of the video, individuals are moving in all direction where as in the abnormal frame of the video individuals are running or walking in one direction. After getting feature vector we convert all feature vectors in the form of Histogram representation. Histogram representation of each frame is trained by SVM method in order to get support vectors.



These support vectors are analyzed by SVM classifier to detect abnormal event of the video. The complete flow diagram of the proposed approach is shown in figure 1

4. SVM Classifier

Support Vector Machine is classification method, which was initially based on the theory of statistical learning. Later by supporting of kernel method it can deal with the nonlinear problem also. The problem of SVM can be represented as:

$$\min_{\omega, \varepsilon, \rho} \frac{1}{2}\|\omega\|^2 + \frac{1}{vn}\sum_{i=1}^{n} \varepsilon_i - \rho \quad (4)$$

Subject to $(\omega, \Phi, (x_i) \geq \rho - \varepsilon_i, \varepsilon_i \geq 0)$

Where ωwidth of margin and$x_i \in X$, $i \in [1 \dots n]$ are n training samples of*X*. $\varepsilon_i$is the stack variable for penalizing the outliers. The hyper parameter $v \in (0,1] is$ the weight for the controlled slack variable. Φ is a map of non-empty set of data X .

5. Abnormal Detection by Optical Flow with SVM Classifier

By adopting the optical flow feature descriptor and the SVM classification method, the abnormal detection in video stream is summarized in following steps:

I. Image frames conversion
II. Extraction of Motion Heat Map which represents the active area of motion
III. Extracting the point of interest
IV. Tracking the feature points in each frame
V. compute the optical flow of each frame $[I_1, \dots, I_m]$ $[I_1, I_2 \dots, I_m] \rightarrow [O_1, O_2 \dots, O_m]$
VI. Compute the histogram of each frame $[O_1, O_2 \dots, O_m] \rightarrow [H_1, H_2 \dots, H_m]$
VII. Training data are learned by SVM classifiers to get support vector $[H_1, H_2 \dots, H_m] \rightarrow [S_1, S_2 \dots, S_m]$
VIII. Each frame is classified by SVM classifier with the help of analyzing support vectors.

6. Result and Analysis

The experiment performance of the proposed approach is evaluated using the dataset is taken from University of Minnesota (UMN). First, the dataset is converted into frames. Then these frames are exposed to proposed approach. We compared our result to ground truth and different state of arts method. Our proposed approach is clearly outperforming SFM [16] and other state of art technique like chaotic invariants [15], sparse recons.[17], Local statistics [18], MDT [19] and GLCM [20]. We have presented this comparison result in form of table1. The UMN dataset scenario is shown in figure 2 - figure 5. Figure 2 and 4 represents normal activity wh ere as figure 3 and figure 5 represents abnormal activity.

Fig. 2: Court-yard Scenario

Fig. 3: Abnormal court-yard scenario

Fig. 4: Hit and Run Scenario

Fig. 5: Abnormal hit run scenario

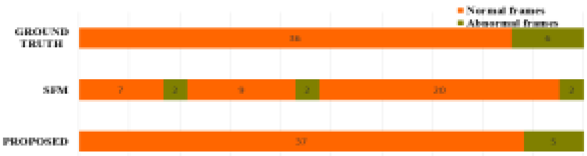


Fig. 6: Frame level comparison of proposed algorithm with Ground truth and existing algorithm (SFM) for dataset courtyard scenario.

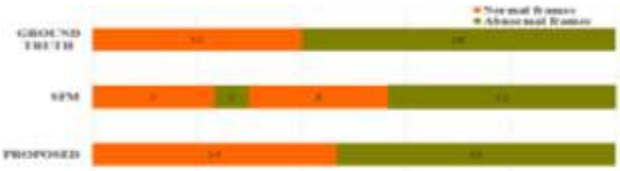


Fig. 7: Frame level comparison of proposed algorithm with Ground truth and existing algorithm (SFM) for dataset hit run scenario.

The UMN dataset that is used for experiment, contains two type of scenarios are courtyard and hit run. The number of frames extracted from these scenarios is, 42 and 30 respectively. An individual frame level comparison is shown in figure 6 and 7 for each scenario.

The courtyard dataset contains 42 frames . In ground truth, 36 frames are normal whereas rest of 6 frames are abnormal. , The hit and run scenario contains 30 frames. In ground truth, 12 frames are normal and rest of 18 frames are abnormal. An individual frame level comparison is shown in figure 6 and 7 for each scenario with Social force model. SFM doesn't provide accurate information which should be close to the ground truth. It is clearly shown that, when we compare frame level comparison, our proposed model gives better than SFM and other state of art methods discussed previously. Receiver operating characteristics (ROC) graph gives Area under curve (AUC). For the taken scenario the corresponding ROC graph is shown in figure 8 and figure 9. AUC represents classification accuracy of the model.AUC value of different algorithm is shown in table 1.

Table 1: AUC Comparison of our proposed technique with different state of arts techniques

SFM [16]	Chaotic Invariants [15]	Sparse recons. [17]	Local statistics[18]	MDT [19]	GLCM[20]	OUR PROPOSED
94.9	99.4	99.5	99.5	99.5	99.5	99.5

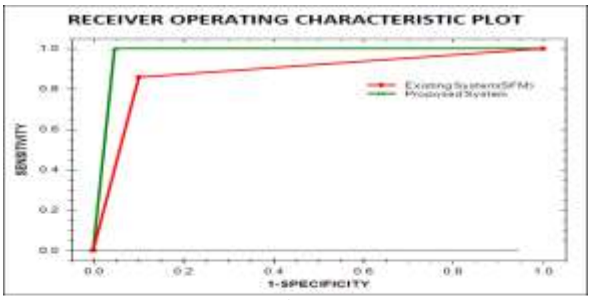


Fig. 8: ROC curve for scenario courtyard

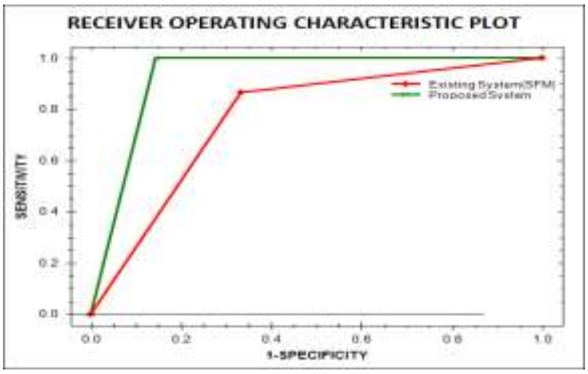


Fig. 9: ROC curve for scenario Hit-run

7. Conclusion

We have presented a technique which is based on optical flow approach and SVM classification system to detect abnormal video events. The dataset used in our experiment is UMN dataset. Optical flow techniques tracks low-level information like points of interest. The displacement and angle between points of interest are estimated. These parameters give the feature vector. For each individual feature, we have a feature vector. These feature vectors are converted into histogram representation. For each frames in a video, histogram representation is estimated. These histogram representations are converted to support vectors. Support vectors are analyzed by SVM classifiers to give results. ROC results are shown in the form of graphs which is comparable result against state of art methods in terms of area under curve and performance.

References

[1] S. Ali and M. Shah, "A lagrangian particle dynamics approach for crowd flow segmentation and stability analysis," in CVPR 2007. IEEE, 2007, pp. 1–6.

[2] S. Ali and M. Shah, "Floor fields for tracking in high density crowd scenes," ECCV 2008, pp. 1–14, 2008.

[3] N. Ihaddadene and C. Djeraba, "Real-time crowd motion analysis," in ICPR 2008. IEEE, 2008, pp. 1–4.

[4] W. Ge, R.T. Collins, and B. Ruback, "Automatically detecting the small group structure of a crowd," in WACV 2009. IEEE, 2009, pp. 1–8.

[5] R. Mehran, B. Moore, and M. Shah, "A streakline representation of flow in crowded scenes," ECCV 2010, pp. 439–452, 2010.

[6] S. Wu, O. Oreifej, and M. Shah, "Action recognition in videos acquired by a moving camera using motion decomposition of lagrangian particle trajectories," in ICCV 2011. IEEE, 2011, pp. 1419–1426.

[7] Cui, Q. Liu, M. Gao, and D.N. Metaxas, "Abnormal detection using interaction energy potentials," in CVPR 2011. IEEE, 2011, pp. 3161–3167.

[8] Amit Adam, Ehud Rivlin, IlanShimshoni, and DavivReinitz. Robust real-time unusual event detection using multiple fixed-location monitors. IEEE Transactionson Pattern Analysis and Machine Intelligence, 30(3):555 560, 2008.

[9] V. Mahadevan, Weixin Li, V. Bhalodia, and N. Vasconcelos. Anomaly detection in crowded scenes. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pages 1975 – 1981, 2010.

[10] L. Kratz and K. Nishino. Tracking with local spatio-temporal motion patterns in extremely crowded scenes. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 693 – 700, 2010.

[11] R. Mehran, A. Oyama, and M. Shah. Abnormal crowd behavior detection using social force model. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pages 935 – 942, 2009.

[12] S.Ali and M. Shah. A lagrangian particle dynamics approach for crowd flow segmentation and stability analysis. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pages 1 – 6, 2007.

[13] L. Li, W. Huang, I. Y. H. Gu, and Q. Tian. Foreground object detection from videos containing complex background. In MULTIMEDIA '03: Proceedings of the eleventh ACM international conference on Multimedia, pages 2–10, New York, NY, USA, 2003. ACM.
[14] C. Harris and M. Stephens. A combined corner and edge detector. In Alvey Vision Conference, page 147-152, 1988.
[15] S. Wu, B. E. Moore, and M. Shah, "Chaotic invariants of lagrangian particle trajectories for anomaly detection in crowded scenes," in IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2010, pp. 2054–2060.
[16] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2009, pp. 935–942.
[17] Y. Cong, J. Yuan, and J. Liu, "Sparse reconstruction cost for abnormal event detection." in Computer Vision and Pattern Recognition, CVPR, 2011, pp. 3449–3456.
[18] V. Saligrama and Z. Chen, "Video anomaly detection based on local statistical aggregates," in IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2012, pp. 2112–2119.
[19] W. Li, V. Mahadevan, and N. Vasconcelos, "Anomaly detection and localization in crowded scenes," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 36, no. 1, pp. 18–32, 2014.
[20] Lloyd, Kaelon, et al. "Detecting violent and abnormal crowd activity using temporal analysis of gray level co-occurrence matrix (GLCM)-based texture measures." Machine Vision and Applications 28.3-4 (2017): 361-371.