



Detecting anomalous crowd behavior using correlation analysis of optical flow

Navneet Nayan¹ · Sitanshu Sekhar Sahu¹ · Sanjeet Kumar¹

Received: 24 June 2018 / Revised: 20 March 2019 / Accepted: 2 April 2019 / Published online: 11 April 2019
© Springer-Verlag London Ltd., part of Springer Nature 2019

Abstract

Defining anomaly criteria is inherently a challenging and critical task. In current scenario, accurate and fast detection of any anomalous events by the visual surveillance system is a major target. To tackle this challenge, a novel method based on correlation analysis of the optical flow has been proposed in this paper for accurate and fast detection of anomalous behavior of a crowd. Exhaustive experimentation has been carried out on the available standard UMN and PETS 2009 datasets for performance comparison. The experimental results demonstrate that the proposed method provides fast detection compared to the existing methods with an accuracy of 97.32%. Further, several factors such as frame gap (the gap of frames for processing during event detection) and illumination condition have been studied. It has been found that the range of correlation value is large under proper illumination condition, whereas its range is small under improper illumination condition. Based on the empirical study, the optimum threshold of 0.75 justifies all types of illumination conditions for the event detection. Again, the optimal value of a gap of four frames has been found suitable to accurately detect the anomalous crowd behavior with less computational burden.

Keywords Anomalous crowd behavior · Optical flow · Correlation analysis · Correlation coefficient · Frame gap · Optimal threshold value

1 Introduction

At present, intelligent video surveillance is one of the major and challenging areas of research which is gaining importance due to its widespread applications. Detection of an abnormal activity, a suspicious person, suspicious objects, and most important unwanted circumstances (like road accident, bomb explosion etc) is really a big challenge for any surveillance system. Accurate detection of any anomalous events by the visual system is a critical task firstly due to the inherent difficulty in defining anomaly criteria and secondly due to the high volumes of data generated [1]. Currently, operators in control rooms use multiple cameras in around public buildings and along streets to detect an abnormal activity through monitoring round the clock. In most of the cases, the recorded videos are explored after the occurrence of a suspicious event. Even the number of operators will be also huge to handle increasing

demand of camera nodes for larger coverage area. Accurate and timely detection of an event in real-time scenario is one of the key concerns in this area. Timely detection of these events provides sufficient time margin for taking action to reduce the impacts of those events. There is a need for real-time monitoring with smart cameras which are able to detect, track and analyze suspicious activity and minimize the impact of any unwanted circumstances.

An event is defined as ‘any action or incidence satisfying the threshold condition of some predefined criteria.’ Hence, depending on various criteria, the events may be categorized in different ways like accidental event, events related to crowd density, event related to crowd escape behavior [1, 2], and many more. Recently, detection of anomaly in the crowd behavior has become a very interesting and challenging area of research. In [3, 4], crowd is defined as ‘a collection of individuals’ and their mind and their mood may shape the behavior of crowd. One of the major factors influencing the mind and mood of the crowd is incidences occurring around them [5]. And this is how the crowd behavior originates [6]. Definition of anomaly depends on the user and varies from scene to scene and situation to situation. Here, anomaly is defined as ‘all of a

✉ Sitanshu Sekhar Sahu
sitanshusekhar@gmail.com; sssahu@bitmesra.ac.in

¹ Department of Electronics and Communication Engineering,
Birla Institute of Technology, Ranchi 835215, India

sudden dispersion in crowd or all of a sudden abrupt change in the speed of crowd.' Regarding this type of anomaly in crowd, several tools and methods are popularly used. Optical flow is one of these, and it has been used in several applications by several researchers. In [7], the authors have proposed an algorithm for the weighted velocity of optical flow vectors and introduced the concept of divergent center for crowd escape behavior detection and localization. Optical flow has been used for crowd motion modeling. An improvement was seen in [7] compared to social force model [8] in terms of both performance metrics accuracy and timely detection. A new concept of object-based abstraction and acceleration feature was proposed in [9]. Its performance was superior to [8]. But the major limitation was an unsatisfactory performance for high-density crowds and poor real-time performance. Optical flow was again used in [10] but in a different way. Using the concept of Riemannian manifolds, a novel idea of optical flow manifolds and optical flow bundle was proposed for detecting various crowd behaviors like walking, running, dispersing, evacuation, merging and splitting. In [11], authors have developed a new method for learning motion pattern of crowded scenarios using the hierarchical agglomerative clustering algorithm. A less accuracy of 78% was obtained in this method. Graph modeling and matching [12] based on Delaunay triangulation have been proposed for human crowd behavior analysis. This method outperforms [8, 11] in both accuracy and timely detection. But in terms of accuracy, it was inferior to [7]. Concept of Delaunay triangulation has been used in [13] where crowd has been represented as an evolving Delaunay graph, and a novel set of local mid-level visual descriptors has been used for crowd analysis. In [8], social force modeling was introduced to detect abnormal crowd behavior but it was less accurate and as an improvement of this work, social attribute-aware force model [14] came into light. In [14] scene scale estimation along with congestion and disorder attributes was added as an improvement of [8]. This improvement resulted in a supreme performance in terms of accuracy over other existing works. However, it may work unsatisfactorily in complex situations like grouping and formation of crowds. Pedestrian behavior modeling [15], spatiotemporal motion concept [16], unsupervised k -means clustering and semi-supervised HMM model [17], unusual event detection using multiple local monitors [18] are other methods for crowd modeling and anomalous detection of the crowd [19]. Velastin et al. [20, 21] use the image processing concept for the analysis of crowd movement and automated crowd density estimation. In the era of artificial intelligence, several machine learning concepts have played important role in crowd behavior analysis. Some of the works based on machine learning and deep learning are discussed in subsequent paragraphs.

One of the important fast and low-power field-programmable gate array (FPGA) implementations of crowd anomaly detection has been discussed in [22]. Here, authors have

proposed a new hardware-friendly classifier KUGDA (k -means with univariate Gaussian discriminant analysis) to achieve the desired targets. The proposed method is based on outlier rejections. According to authors, this work is the first dedicated implementation of crowd anomaly detection. As future work, authors have suggested to use proposed features with deep neural networks. In another work [23], the authors have used deep learning architectures with slow feature analysis learning methods for video anomaly detection. For measuring the degree of abnormality, the authors have proposed a novel summed squared derivative value concept. This approach can detect both local and global anomaly. However, there is some scope of improvement for local anomaly detection in [23]. Continuing with the deep learning concept, the authors of [24] have proposed a deep learning-based anomaly detection system. They incorporated four modules, namely background estimation, object segmentation, feature extraction and activity recognition module, to achieve their goals. As future work, they focus on reducing the feature dimension while maintaining its robustness. In this continuation concept of hybrid features, using level set method has been introduced in [25]. Five different image descriptors such as color moments, edge histogram descriptor, color layout descriptors, color and edge directivity descriptors, and scalable color descriptors were extracted from an input video for the robust detection of an abnormal event. A novel method based on sparse reconstruction using two dictionaries, namely global dictionary and an online local dictionary, has been discussed in [26]. One-class support vector machine (SVM) model along with histogram of oriented gradients-local binary patterns (HOG-LBP) descriptor and histogram of oriented optical flow (HOF) descriptor has been used for abnormal event detection in crowded scenes in [27].

In this work, optical flow [28, 29] is used as the basic tool followed by correlation analysis of magnitude matrices [30] of optical flow vectors. An optical flow vector has four parameters. Here the magnitude parameter of optical flow is utilized [31]. Then a threshold value of correlation coefficients has been empirically estimated, and a new approach has been proposed to detect an anomaly in the crowded scenes.

2 Methodology

The basic steps of the entire methodology for anomalous crowd behavior detection are (a) capturing actions using CCTV cameras, (b) checking all threshold conditions, and (c) making decision whether event occurred or not (Fig. 1).

2.1 Anomalous crowd behavior detection

Whenever there is a crowded place, movement of people is very less or there is only slight movement in crowd. Hence, frame by frame there is little change or almost negligible

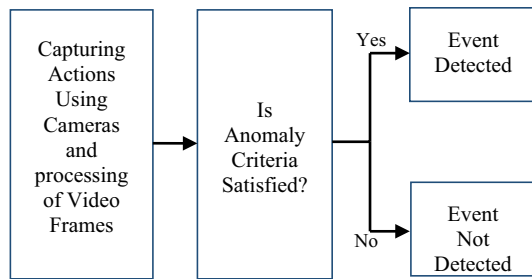


Fig. 1 A framework for anomalous crowd behavior detection

change in the scene. Hence, there will be high correlation between consecutive frames, and consequently, gradient values of consecutive correlation values will be very less. But once crowd will start dispersing all of a sudden, the consecutive frames will be differ from one another by a considerable amount and hence the correlation between consecutive frames will start decreasing and hence gradient values of the consecutive correlation values will increase. The flowchart of the entire process is shown in Fig. 2.

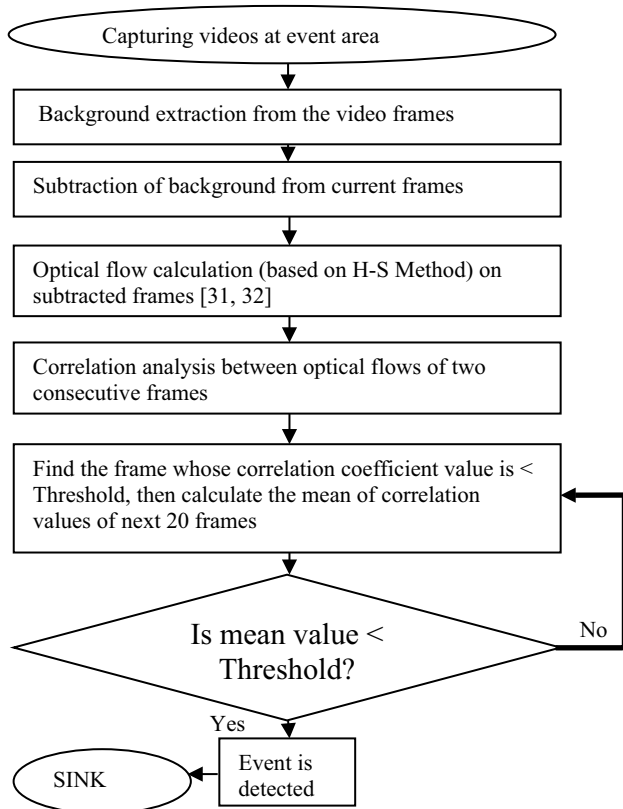


Fig. 2 Flowchart of entire process of anomalous crowd behavior detection

3 Proposed algorithm for anomalous event detection

Algorithm 3.1: *Anomalous crowd behavior detection.*

Input: Captured frames of video signal.

Output: Decision whether event is detected or not based on anomaly criteria.

Step Involved:

- i. **Read the video and Obtain total number of frames as ‘n’**

```

For iFrame = 1: n
  Extract all the frames and store them in a
  folder named as ‘folder’.
End
    
```

- ii. **Obtaining magnitude parameter of optical flow of subtracted frames and storing them in a matrix**

```

Make the first frame as background ‘bg’.
For iOpt = 1:4: length (folder)
  Opticflow = opticalFlowHS;
  (making an optical flow object);
  Read all the extracted frames
  stored in the folder as ‘I’
  S = bg-I; (subtract every frame
  from background)
  Obtaining optical flow on
  subtracted frames as ‘Flow’
  Obtain magnitude parameter of the
  optical flow and store them in
  matrix ‘M’
End
    
```

- iii. **Calculation of correlation between consecutive frames by calculating correlation between consecutive magnitude matrices**

```

For j = 1:iOpt-1
  R1 = M (:,j);
  R2 = M (:,j+1);
  Correlation(C) between
  consecutive frames (i.e. consecutive
  matrices)
    
```

$$C = \frac{\sum_m \sum_n (R_{1mn} - \overline{R_1})(R_{2mn} - \overline{R_2})}{\sqrt{\left(\sum_m \sum_n (R_{1mn} - \overline{R_1})^2\right) \left(\sum_m \sum_n (R_{2mn} - \overline{R_2})^2\right)}}$$

where R_1 and R_2 are m-by-n matrices.

$\overline{R_1}$ and $\overline{R_2}$ are the mean of matrices R_1 and R_2 .

End

- iv. **Checking anomaly criteria to make a decision whether an event occurred or not**

```

For k = 1:j-20
    
```

```

  if C < 0.75 (Threshold value)
    
```

```

    Store all such frames in a variable
    'index'
    Cal. mean of 20 consecutive frames
    from the 1st value of 'index'
    
$$M = \frac{I_1 + I_2 + I_3 + \dots + I_{18} + I_{19} + I_{20}}{20}$$

    End
    if  $M < 0.75$ 
        Display "Anomalous event is
        detected"
        else increment the index by 1 and
        calculate the mean as in 13 from
        incremented index value
    End

```

4 Results and discussion

The proposed methodology has been assessed with the videos obtained from University of Minnesota (UMN) and PETS 2009 datasets. The UMN dataset contains a combination of 11 videos of three different scenarios. Two of these scenarios are in the broad daylight with perfect illumination condition, whereas the third scenario is inside a hall where the illumination condition is little poor. In PETS 2009 dataset, the videos have been taken in perfect illumination conditions. Several video streams (112 combinations) have been generated from the video datasets by random placement of event locations. In this way, total 84 event locations and 28 non-event locations have been generated for the study. All the simulations have been carried out on Intel(R) Core(TM) i7-6700 CPU @ 3.40 GHz processor with 8.00 GB RAM.

Test video-1 The length of the video (*Test video-1*) is of duration 21 s and contains 625 captured frames. The size of the each frame is 320-by-240. Sub-sampling has been done on test video-1 frames. In total, 156 frames have been sub-sampled with a frame gap of 4 frames out of 625 frames for analysis. The video frame of normal movement (people are walking at normal speed) and abnormal crowd movement (people are running in fear) is shown in Fig. 3.

Figure 4 shows the plot of correlation coefficient versus frame number in which a pattern is observed which tells that during normal movement of crowd the correlation was high, but during abrupt movement correlation values started falling down, and once very few people were left and entire seen was almost unchanged again, high correlation values obtained. The event started at 120th frame and end at 146th frame.

Test video-2 This video is extracted from UMN dataset. This scenario contains a total of six videos. The length of the video is of duration 32 s and has 766 frames. The frame size is 320-by-240. The frames were sub-sampled at a gap of four frames and extracted.

Figure 5 shows three frames where the first frame is of normal movement of crowd and the other two frames show anomalous behavior of crowd. The correlation analysis of the frames is shown in Fig. 6. During normal movement, correlation values are high and as the random movement started corresponding to the event occurrence, the correlation values decreased. It again increased as the event ends. The event region is identified between 127th and 148th frame.

Test video-3 This video is also extracted from UMN dataset. The length of video is 34 s and has 807 frames. The frame size is 320-by-240. The frames were extracted at a gap of four frames; hence, there are 201 frames.

In Fig. 7, the first frame shows normal crowd movement and the other two frames show anomalous behavior of crowd, and Fig. 8 contains the simulated results of this video. In Fig. 8, initially correlation values are high representing the normal movement of crowd, a dip followed by a rise again indicates that the crowd started dispersing and after some time there were no one in the scene. The frame numbers between which the event was detected is 179th frame and 198th frame.

The proposed method has been assessed on several videos generated from UMN dataset [32] and PETS 2009 dataset [33]. In total, 112 videos have been generated, out of which 84 video are categorized as event videos and 28 are categorized as non-event videos. The value of correlation coefficient is considered here as an anomaly criteria, and its threshold for the event detection is chosen as 0.75 under proper as well as improper illumination condition. The performance of the proposed method is assessed with receiver operating characteristic curve (ROC), and the results are shown in Fig. 9. Various parameters like area under the curve (AUC), precision, accuracy and detection rate are presented in Table 1. Precision also known as positive predictive value is defined as $TP/(TP + FP)$, and accuracy is defined as $(TP + TN)/(TP + FN + FP + TN)$. The detection rate also known as recall or sensitivity is defined as $TP/(TP + FN)$. It is also defined as the number of detected events divided by number of ground truth events. A brief explanation of these terms can be found in [12]. The metric timely detection is measured as experimentally determining the frame at which event started and the frame at which event finished. The result shows that the average detection rate of

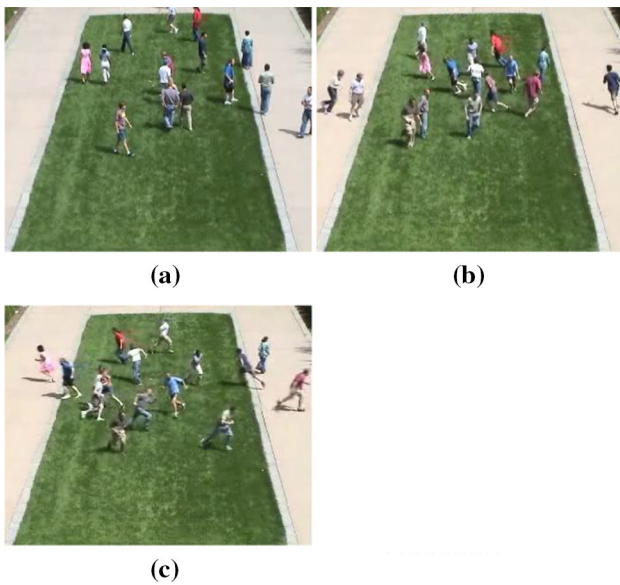


Fig. 3 a Normal movement of crowd (normal walk) whereas b, c crowd’s abnormal movement (running)

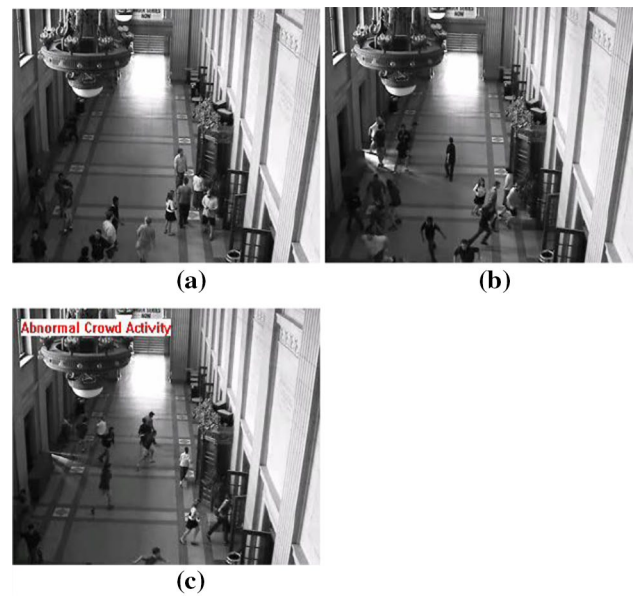


Fig. 5 a Normal movement of crowd and b, c crowd’s abnormal movement

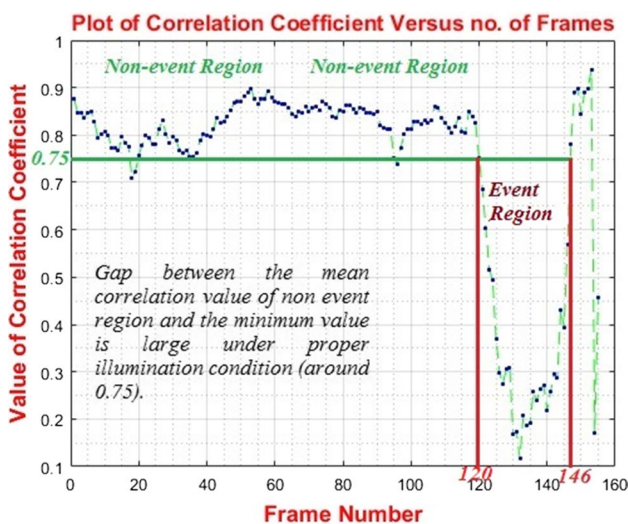


Fig. 4 Plot of correlation coefficient versus frame number when correlation analysis was done on the magnitude parameter of optical flow of consecutive frames. Here the non-event region and the region in which event detected is shown

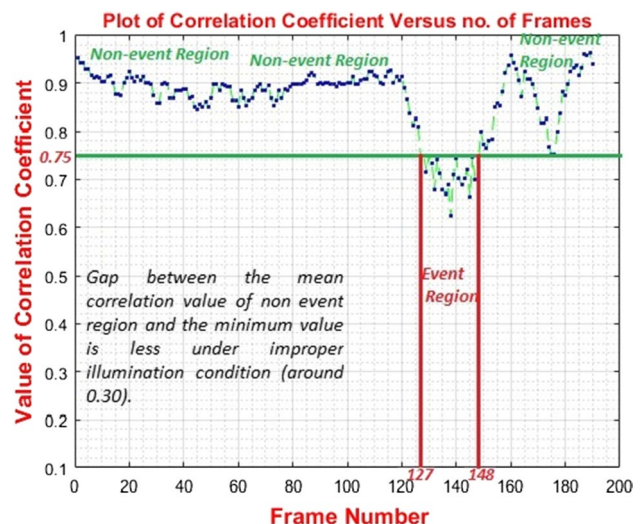


Fig. 6 Plot of correlation coefficient versus frame number. The correlation analysis is done on the optical flow of consecutive frames

the proposed method is 97.61%, i.e., almost all the events are detected properly.

Table 2 presents the performance comparison of the proposed method with other existing methods. Different authors have used different metrics to show their performances and superiority over other state of the art. It is elucidated that performance of the proposed method is either comparable or better than the existing methods. Another metric timely detection is also used for comparison. So, the simulation

time and frame around which the event occurred have been assessed and the results are shown in Table 3. The results show an improvement regarding these factors compared to existing works. We have used the same datasets for the comparison.

From Table 3, it is concluded that the proposed method is superior to other existing state-of-the-art methods regarding timely detection. Compared to other works, the proposed method detected the events earlier and results were more nearer to the actual time of event occurred.

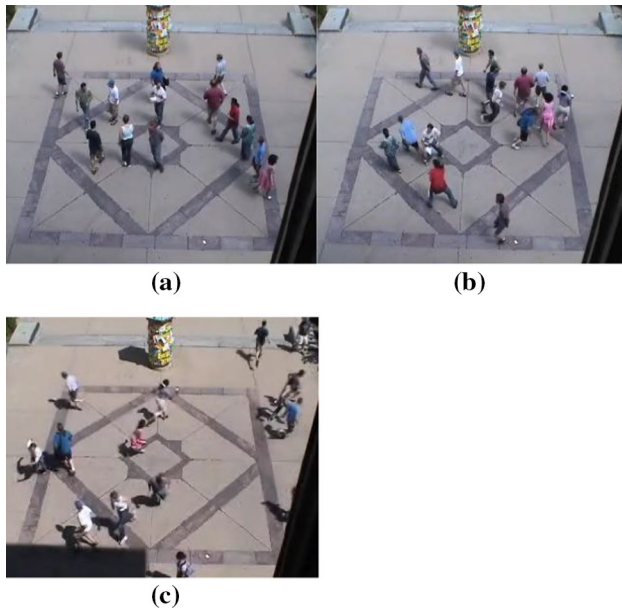


Fig. 7 a Normal movement of crowd and b, c crowd’s abnormal movement

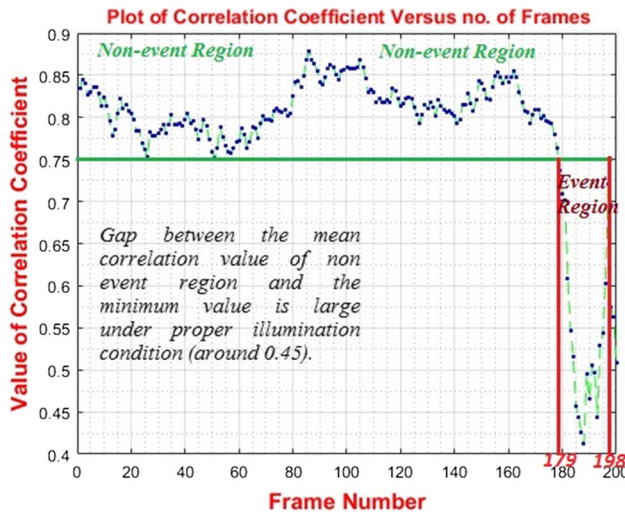


Fig. 8 Plot of correlation coefficient values of optical flow of consecutive frames against frame number

Further, the study has been carried out to observe the accuracy of the event detection with various threshold and the results are shown in Table 4.

The results clearly show that the optimal threshold value is 0.75 giving an accuracy of 97.32%. Also this threshold value handles the illumination condition, the proper one and the improper one.

Table 4 and Fig. 10 show the impact of variation of threshold values on the accuracy of detection of events. The effect is more pronounced in case of improper illumination

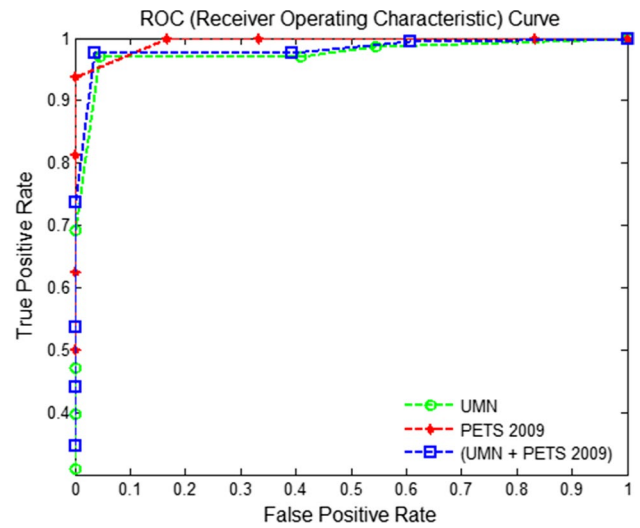


Fig. 9 ROC curve for the proposed method on the UMN and PET’09 dataset

Table 1 Performance analysis of the proposed method for several videos generated from UMN and PETS 2009 datasets

| Dataset | AUC | Precision | Accuracy (in %) | Detection rate (in %) |
|-----------------|------|-----------|-----------------|-----------------------|
| UMN | 0.97 | 1 | 97.77 | 97.05 |
| PETS 2009 | 0.99 | 0.96 | 96.45 | 100 |
| (UMN+PETS 2009) | 0.98 | 0.98 | 97.32 | 97.61 |

condition. As we can see in Fig. 9, the accuracy varies in a range of 0.76 and 0.98 by varying the threshold values in case of proper illumination condition, whereas its variation range is between 0.19 and 0.95 under improper illumination condition. Thus, the optimal threshold value is judiciously chosen as 0.75. At this value, the accuracy percentage is 98 and 95 for proper and improper illumination condition, respectively. Thus, it is clearly observed that the threshold value of 0.75 is suitable under any illumination condition. The overall accuracy of the proposed method in the optimum threshold is 97.32%.

5 Conclusion

In this paper, a method based on correlation analysis of the optical flow is proposed for detection of anomalous behavior of crowd. Exhaustive experimental results demonstrate that the proposed method provides more accurate and fast results compared to the existing methods. The correlation coefficient of the optical flow of consecutive frames provides a pattern of non-event region and event region. It has been

Table 2 Performance comparison of the proposed method with existing state-of-the-art methods

| Approach | AUC (UMN dataset) | AUC (PETS 2009 dataset) | Detection rate (UMN dataset) (%) |
|--|-------------------|-------------------------|----------------------------------|
| Pure optical flow [11] | – | – | 78 |
| Graph modeling and matching [12] | – | – | 91 |
| Bag of trajectory graphs [34] | 0.944 | 0.912 | – |
| Weighted average of non-extensive entropies [35] | 0.95 | – | – |
| Social force model [8] | 0.96 | – | 85 |
| Deep incremental slow feature analysis network [23] | 0.9692 | 0.9906 | – |
| Histogram of oriented swarm [36] | 0.9702 | – | – |
| Optical flow manifolds | – | – | – |
| Weighted velocity of optical flow | – | – | – |
| Proposed method (correlation analysis of optical flow) | 0.9755 | 0.9948 | 97.05 |
| FPGA implementation using KUGDA [22] | 0.9818 | – | – |
| Visual descriptors [13] | 0.9861 | – | – |
| Interacting energy potentials [37] | 0.985 | – | – |
| Social attribute-aware force model [14] | 0.986 | – | – |
| GLCM based texture measures [38] | 0.9956 | – | – |
| OADC-SA [39] | 0.9967 | – | – |

Basis of this comparison has been done in terms of AUC and detection rate

Table 3 Performance comparison of the proposed method with the method of [7, 9, 10] based on the simulation time and frame numbers around which event occurred

| Method used | Dataset | Videos description | No. of frames | Average simulation time (in s) | Occurrence of event (frame no.) experimental | Occurrence of events (frame no.) real |
|--|-----------|--------------------|---------------|--------------------------------|--|---------------------------------------|
| Weighted velocity of optical flow [7] | UMN | Video 1(a) | 625 | 0.7895/frame | 485–580 | 477–590 |
| | | Video 2(c) | 770 | 0.7895/frame | 3200–3260 | 3180–3297 |
| | | Video 3(c) | 807 | 0.7895/frame | 7660–7720 | 7649–7730 |
| Concept of acceleration feature [9] | UMN | Video 1(a) | 625 | Poor real-time performance | 490–580 | 477–590 |
| | | Video 2(a) | 548 | Poor real-time performance | 1778–1855 | 1770–1875 |
| | | Video 3(a) | 658 | Poor real-time performance | 6160–6239 | 6144–6254 |
| | PETS 2009 | View 1 (14–33) | 378 | Poor real-time performance | 340–377 | 336–377 |
| | | View 2 (14–33) | 378 | Poor real-time performance | 350–377 | 336–377 |
| Concept of optical flow manifolds [10] | PETS 2009 | View 1 (14–33) | 378 | Not given | 343–377 | 336–377 |
| Proposed method (correlation analysis of optical flow) | UMN | Video 1(a) | 625 | 16.974 | 480–584 | 477–590 |
| | | Video 2(a) | 548 | 15.307 | 1771–1869 | 1770–1903 |
| | | Video 2(c) | 770 | 19.731 | 3195–3279 | 3180–3297 |
| | | Video 3(a) | 658 | 17.341 | 6146–6238 | 6144–6254 |
| | | Video 3(c) | 807 | 21.137 | 7649–7725 | 7649–7730 |
| | PETS 2009 | View 1 (14–33) | 378 | 12.402 | 339–377 | 336–377 |
| | | View 2 (14–33) | 378 | 12.402 | 340–377 | 336–377 |

Table 4 Performance comparison of accuracy of detection of events for proper and improper illumination condition by varying the threshold values

| | TV | IC | TNE | TNNE | TP | TN | FP | FN | Accuracy (in %) | Total accuracy (in %) |
|------|----|----|-----|------|----|----|----|----|-----------------|-----------------------|
| 0.50 | | P | 33 | 16 | 29 | 16 | 00 | 04 | 92 | 50.89 |
| | | I | 51 | 12 | 00 | 12 | 00 | 51 | 19 | |
| 0.55 | | P | 33 | 16 | 29 | 16 | 00 | 04 | 92 | 50.89 |
| | | I | 51 | 12 | 00 | 12 | 00 | 51 | 19 | |
| 0.60 | | P | 33 | 16 | 30 | 16 | 00 | 03 | 94 | 58.03 |
| | | I | 51 | 12 | 07 | 12 | 00 | 44 | 30 | |
| 0.65 | | P | 33 | 16 | 31 | 16 | 00 | 02 | 96 | 65.17 |
| | | I | 51 | 12 | 14 | 12 | 00 | 37 | 41 | |
| 0.70 | | P | 33 | 16 | 32 | 15 | 01 | 01 | 96 | 79.46 |
| | | I | 51 | 12 | 30 | 12 | 00 | 21 | 66 | |
| 0.75 | | P | 33 | 16 | 33 | 15 | 01 | 00 | 98 | 97.32 |
| | | I | 51 | 12 | 49 | 12 | 00 | 02 | 95 | |
| 0.80 | | P | 33 | 16 | 33 | 10 | 06 | 00 | 88 | 88.39 |
| | | I | 51 | 12 | 49 | 07 | 05 | 02 | 88 | |
| 0.85 | | P | 33 | 16 | 33 | 05 | 11 | 00 | 76 | 83.03 |
| | | I | 51 | 12 | 49 | 06 | 06 | 02 | 87 | |

TNE, TNNE, TV, IC, P and I are acronyms for total number of events, total number of non-events, threshold value, illumination condition, proper and improper, respectively

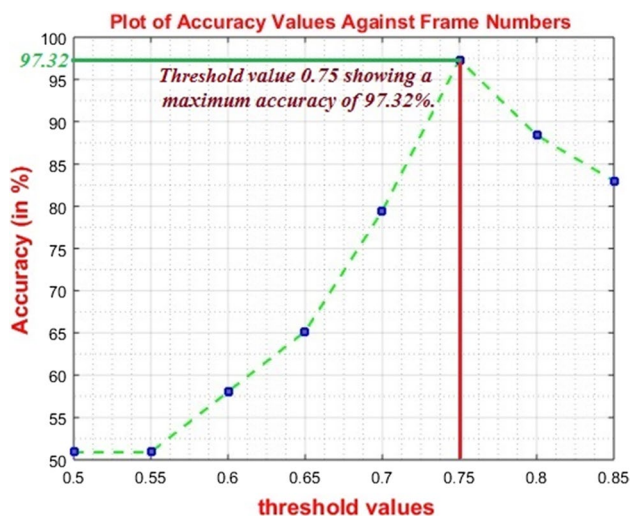


Fig. 10 Variation of accuracy values with the threshold value of the correlation coefficient

found that the range of correlation value is large to decide a threshold under proper illumination condition, whereas its range is small under improper illumination condition. Based on the empirical study, the optimum threshold value for all the illumination condition was found to be 0.75 for accurate detection. In this threshold, the proposed method provides an accuracy of 97% on standard datasets.

References

- Li, T., Chang, H., Wang, M., Ni, B., Hong, R., Yan, S.: Crowded scene analysis: a survey. *IEEE Trans. Circuits Syst. Video Technol.* **25**(3), 367–386 (2015)
- Junior, J.C.S.J., Musse, S.R., Jung, C.R.: Crowd analysis using computer vision techniques: a survey. *IEEE Signal Process. Mag.* **27**, 66–77 (2010)
- Bon, G.L.: *The Crowd: A Study of the Popular Mind*. Dover Publications, New York (2012)
- Still, G.K.: *Introduction to Crowd Science*. CRC Press, Boca Raton (2013)
- Dogbe, C.: Modeling crowd dynamics from a kinetic theory view point. *J. Math. Sci. Adv. Appl.* **30**, 71–88 (2014)
- Bellomo, N., Berrone, S., Gibelli, L., Pieri, A.: First order models of human crowds with behavioral-social dynamics. *Phys. Soc. (Physics.soc-ph)*, 1–13 (2015)
- Chen, C.Y., Shao, Y.: Crowd escape behavior detection and localization based on divergent centers. *IEEE Sens. J.* **15**(4), 2431–2439 (2015)
- Mehran, R., Oyama, A., Shah, M.: Anomalous crowd behavior detection using social force model. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 935–942 (2009)
- Chen, C., Shao, Y., Bi, X.: Detection of anomalous crowd behavior based on the acceleration feature. *IEEE Sens. J.* **15**(12), 7252–7261 (2015)
- Rao, A.S., Gubbi, J., Marusic, S., Palaniswami, M.: Crowd event detection on optical flow manifolds. *IEEE Trans. Cybern.* **46**(7), 1524–1537 (2016)
- Hu, M., Ali, S., Shah, M.: Learning motion patterns in crowded scenes using motion flow field. In: *IEEE International Conference on Pattern Recognition (ICPR)*, pp. 1–5 (2008)
- Chen, D.Y., Huang, P.C.: Visual based human crowds behavior analysis based on graph modeling and matching. *IEEE Sens. J.* **13**(6), 2129–2138 (2013)

13. Fradi, H., Luvison, B., Pham, Q.C.: Crowd behavior analysis using local mid-level visual descriptors. *IEEE Trans. Circuits Syst. Video Technol.* **27**(3), 589–602 (2017)
14. Zhang, Y., Qin, L., Ji, R., Yao, H., Huang, Q.: Social attribute-aware force model: exploiting richness of interaction for abnormal crowd detection. *IEEE Trans. Circuit Syst. Video Technol.* **25**(7), 1231–1245 (2015)
15. Yi, S., Li, H., Wang, X.: Pedestrian behavior modelling from stationary crowds with application to intelligent surveillance. *IEEE Trans. Image Proc.* **25**(9), 4354–4368 (2016)
16. Cong, Y., Yuan, J., Tang, Y.: Video anomaly search in crowded scenes via spatio-temporal motion context. *IEEE Trans. Inf. Forensic Secur.* **8**(10), 1590–1599 (2013)
17. Andersson, M., Gustafsson, F., Laurent, L.S., Prevost, D.: Recognition of anomalous motion patterns in urban surveillance. *IEEE J. Sel. Top. Signal Process.* **7**(1), 102–110 (2013)
18. Amit, A., Rivlin, E., Shimshoni, I., Reinitz, D.: Robust real-time unusual event detection using multiple fixed-location monitors. *IEEE Trans. Pattern Anal. Mach. Intell.* **30**(3), 555–560 (2008)
19. Gunduz, A.E., Ongun, C., Temizel, T.T., Temizel, A.: Density aware anomaly detection in crowded scenes. *IET Comput. Vis.* **10**(5), 374–381 (2015)
20. Velastin, S.A., Yin, J.H., Davies, A.C., Vicencio-Silva, M.A., Allsop, R.E., Penn, A.: Analysis of crowd movements and densities in built-up environments using image processing. In: *IEE Colloquium on Image Processing for Transport Applications*, pp. 8/1–8/6 (1993)
21. Velastin, S.A., Yin, J.H., Davies, A.C., Vicencio-Silva, M.A., Allsop, R.E., Penn, A.: Automated measurement of crowd density and motion using image processing. In: *IEEE International Conference on Road Traffic Monitoring and Control (ICRTMC)*, pp. 127–132 (1994)
22. Khan, M.U.K., Park, H., Kyung, C.: Rejecting motion outliers for efficient crowd anomaly detection. *IEEE Trans. Inf. Forensic Secur.* **14**(2), 541–556 (2019)
23. Hu, X., Hu, S., Huang, Y., Zhang, H., Wu, H.: Video anomaly detection using deep incremental slow feature analysis network. *IET Comput. Vis.* **10**(4), 258–265 (2016)
24. Revathi, A.R., Kumar, D.: An efficient system for anomaly detection using deep learning classifier. *Signal Image Video Process.* **11**(2), 291–299 (2017)
25. Kangwei, L., Jianhua, W., Zhongzhi, H.: Abnormal event detection and localization using level set based on hybrid features. *Signal Image Video Process.* **12**(2), 255–261 (2018)
26. Li, S., Liu, C., Yang, Y.: Anomaly detection based on sparse coding with two kinds of dictionaries. *Signal Image Video Process.* **12**(5), 983–989 (2018)
27. Amraee, S., Vafaei, A., Jamshidi, K., Adibi, P.: Abnormal event detection in crowded scenes using one-class SVM. *Signal Image Video Process.* **12**(6), 1115–1123 (2018)
28. Horn, B.K.P., Schunck, B.G.: determining optical flow. *Artif. Intell.* **17**(1–3), 185–203 (1981)
29. Hassner, T., Itcher, Y., Kliper-Gross, O.: Violent flows: real-time detection of violent crowd behavior. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR)*, pp. 1–6 (2012)
30. Bai, L., Velichko, A., Drinkwater, B.W.: Ultrasonic characterization of crack-like defects using scattering matrix similarity metrics. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **62**(3), 545–559 (2015)
31. Barron, J.L., Fleet, D.J., Beauchemin, S.S., Burkitt, T.A.: Performance of optical flow techniques. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 236–242 (1992)
32. UMN dataset: Detection of Unusual Crowd Activity. http://mha.cs.umn.edu/proj_events.shtml. Accessed 1 Dec 2016
33. PETS 2009 Benchmark Data. <http://www.cvg.reading.ac.uk/PETS2009/a.html>. Accessed 15 Mar 2017
34. Zhang, Y., Huang, Q., Qin, L., Zhao, S., Yao, H., Xu, P.: Representing dense crowd patterns using bag of trajectory graphs. *Signal Image Video Process.* **8**(1), 173–181 (2014)
35. Susan, S., Hanmandlu, M.: Unsupervised detection of nonlinearity in motion using weighted average of non-extensive entropies. *Signal Image Video Process.* **9**(3), 511–525 (2015)
36. Kaltsa, V., Briassouli, A., Kompatsiaris, I., Hadjileontiadis, L.J., Strintzis, M.J.: Swarm intelligence for detecting interesting events in crowded environments. *IEEE Trans. Image Process.* **24**(7), 2153–2166 (2015)
37. Cui, X., Liu, Q., Gao, M., Metaxas, D.N.: Abnormal detection using interaction energy potentials. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3161–3167 (2011)
38. Lloyd, K., Rosin, P.L., Marshall, D., Moore, S.C.: Detecting violent and abnormal crowd activity using temporal analysis of grey level co-occurrence matrix (GLCM) based texture measures. *Signal Image Video Process.* **28**, 361–371 (2017)
39. Yuan, Y., Fang, J., Wang, Q.: Online anomaly detection in crowd scenes via structure analysis. *IEEE Trans. Cybern.* **45**(3), 548–561 (2015)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.