CrossMark

Abnormal events' detection in crowded scenes

Mariem Gnouma[1] (ORCID) **· Ridha Ejbali[1] · Mourad Zaied[1]**

Abstract In this paper, two new methods are developed in order to detect and track unexpected events in scenes. The process of detecting people may face some difficulties due to poor contrast, noise and the small size of the defects. For this purpose,the perfect knowledge of the geometry of these defects is an essential step in assessing the quality of detection. First, we collected statistical models of the element for each individual for time tracking of different people using the technique of Gaussian mixture model (GMM). Then we improved this method to detect and track the crowd(IGMM). Thereafter, we adopted two methods: the differential method of Lucas and Kanade(LK) and the method of optical flow estimation of Horn Schunck(HS) for optical flow representation. Then, we proposed a novel descriptor, named the Distribution of Magnitude of Optical Flow (DMOF) for anomalous events' detection in the surveillance video. This descriptor represents an algorithm whose aim is to accelerate the action of abnormal events' detection based on a local adjustment of the velocity field by manipulating the light intensity.

1 Introduction

Surveillance cameras have invaded all public areas. These cameras needed a regular human presence to monitor the captured scenes and interference if necessary to alert those responsible in case of problems. These intelligent surveillance systems have not yet reached the desired level of suitability and robustness resulting from the huge amount of data analysis

✉ Mariem Gnouma
mariem.gnouma.tn@ieee.org

[1] Research Team on Intelligent Machines, National School of Engineers of Gabes,
University of Gabes, Gabes, Tunisia

and wide algorithmic assumptions. Yet, the ability to monitor large geographical areas with minimal human resources is a very hard manual process and renders the detection of events difficult since it requires a constant analysis of all paths; which is typically a big challenge for a human analyst due to information overload, fatigue and inattention.

Now monitoring would be done by systems that allow the recognition of different inconsistent clauses to make network managers be more attentive. Unfortunately, nowadays, the security problem is a big challenge. Indeed,there are limitations to the efforts made from the research community, because it was thought to build a vision system that can handle the different scenes congested by providing the information needed to ensure public safety.

In the literature, there are several definitions for abnormal events. However, the most widely accepted definition defines abnormal events as low probability events compared to a probabilistic model of normal behavior.

The probability of the detection of an abnormal event means the discovery and response to abnormal changes or behaviors of humans or objects in the videos. Now, there are several techniques for detecting abnormal actions for the vision system, such as [19, 57, 58], which allows the detection of emergency conditions. Furthermore, to detect those events, many novelties in video surveillance such as people counting [29] people detection [21],the detection of a host [40], traffic monitoring [56], abnormal crowd behaviors [12] and temporal object tracking [7] have been made in recent years. Abnormal events in the video surveillance can be divided into two categories: Global Abnormal Events (GAE) and Local Anomalous Events (LAE) [34].

Local events can be defined as an energy model that is used in many rare event detection tasks.

Local EIA, as suggested by Cao et al. [6], uses motion directions and kinetic energy to detect aberrations in overpopulated scenes. Xu et al. [55] have proved the energy model in spatio-temporal contexts.

The GAEs are the events happening over a stage such as "fight the group", for example, where the pedestrians suddenly dispersed due to an explosion.

Some other works combine the two such as in [24] and the sparse representation by projecting the histogram of maximal optical flow projection (HMOFP) of features from a set of normal data.

In our work, we exceptionally focus on global abnormal crowd behavior detection.

Almost, there exist two main approaches for understanding Global abnormal crowd behaviors[18]: holistic approaches and object-based approaches.

In object-based methods, a crowd is considered a collection of objects [46]. It is necessary to detect and track each individual to understand the behavior of the crowd.

On the other hand, For holistic approaches [8, 11, 30] the crowd is presented as a global entity and extracts the different characteristics such as the use of the optical flow to identify the movement in the data-set for the analysis of a higher level.To detect this type of event, Wang et al. in [52] proposed a method based on an image descriptor combined with a nonlinear classification method; by introducing an optical flow orientation histogram as a descriptor encoding the displacement information of each video frame. The results obtained by Cui et al [39] suggest that the potential energy of interaction explores the relationship between the state of a person and these reactions. Spatio-temporal grid-based framework was introduced by Nam [36] to deal with the complexity of structures of the motion flows that can group optical flows into crowds. Meanwhile, several other approaches are highly effective in detecting global abnormal events [45].

In our research, we focus primarily on the global detection of aberrant events of a crowd in surveillance video sequences. we can define an anomaly as the detection of patterns in a

given set of data that do not conform to an established standard. This means that aberrant events can be identified as irregular activities with respect to usual normal ones, such as "pedestrians suddenly start running in different directions".

We can define an anomaly as the detection of patterns in a given set of data that do not conform to an established standard. This means that aberrant events can be identified as irregular activities with respect to usual normal ones, such as "pedestrians suddenly start running in different directions".

This paper aims at presenting the following contribution:

–   A new real-time technique for pedestrian tracking in a crowd by improving the method of Gaussian Mixture Model to track all the crowded scene in the video which is adaptively modeled so that it follows the crowd throughout the surveillance with high performance.
–   We propose a novel DMOF descriptor to distinguish abnormal events in a crowded scene. Since our method was established in specific cells of frame for anomaly motion detection, this not only avoids unnecessary computation and speeds up the system but also contributes to a high detection rate. Because the descriptor is scale and rotation-invariant, the new DMOF is very effective

We work on two stages that require segmentation prior background, no earlier normal movement training with motion recognition and tracking of aberrant movement in the object is performed. We analyze the trajectory to follow the crowd by improving the Gaussian mixture model method (IGMM).Also, more details of our proposed approach DMOF are provided for clarity and integrity.

The rest of this paper is organized as follows: Section 2 introduces the related work in the literature. The proposed methods for both tracking and recognition of abnormal events are described in Section 3. Section 4 provides a comparison between our method and other online approaches in the literature to demonstrate the advantages of the proposed methods. The final section gives a conclusion of this study.

2 Related work

In this part, we present the various aspects of the problem of crowd abnormality detection. Anomaly detection has been identified as a technique to uncover rare events. Among the several topics related to detection of anomalies are detection of a sudden movement, noise detection and detection of violence.

Many of the state of the art approaches for unexpected behavior detection in crowded scenes are based on complex models, which require high processing time and several parameters to be adjusted.

Usually, in public places, people do not always follow the laws of physics; they have no preservation of momentum and can stop, run and start at will. To detect certain types of exceptional events; Mehran et al. [34] argue that the model of social work is the best solution to detect and locate aberrant behavior in certain videos of the crowd. To solve the measurement of this problem of anomalies, most conventional algorithms [1, 3] detect the test sample based on the low probability as an anomaly by fitting a model of training data. In addition, several other algorithms such as the Hidden Markov Model (HMM) [1], the Markov Random Field (MRF) [2] and Temporal MRF [3] have been proposed.

We focus on the detection of abnormal events in overcrowded scenes using the optical flow characteristics. The optical flow is to compute pixel wise instantaneous motion

between two consecutive frames. This influence of optical flow is frequently a point of hypothetical argument when dealing with crowds. The most interesting approach to this issue has been proposed in [9, 13, 16, 20, 47]. In early attempts, some aberrant methods focused on the use of the optical flow algorithm. Ryan et al. [42] used measures of optical flow with great results. The authors encode optical flow vectors using a three dimensional gray level co-occurrence matrix structure, created the dynamics of a local region by the uniformity of the motion. They created a new model of normality using a GMM. Their method is both powerful at astute between normal and abnormal scenes while keeping up an arguably real-time processing speed of approximately 9 frames/s.

In [31, 33, 50], the algorithm of LK is used to estimate optical flow for each pixel. Each action is labeled by movement direction and the position of an empty square cell to produce the representation of words in videos. In this context, Marques et al. [23] argue that dense crowd tracking algorithms are impossible and that optical flow estimation can become unreliable. Reference [43] analyzes and compares various aspects of abnormal detection in dense scenes by measuring the uniformity of a flow field by the use of three dimensional gray level co-occurrence matrix structure. This method is very effective to help differentiate between normal and abnormal movements while respecting a speed of time slightly questionable in real time. We choose to adapt this technique because of the robustness of the multiple and simultaneous camera and object motions and it is widely used in dense crowd segmentation and motion detection.

3 Proposed methods overview

Unlike to the various approaches mentioned above, our proposed method is created for the detection of abnormal motion in a crowd in a large sense. We cite the example of unexpected human behaviors. We try to evolve techniques that are generally valid to be applied to different video surveillance.

3.1 Proposed method for crowd detection

Detection of the crowd is a difficult task due to occlusions in individuals and because of scene clutter.Detection of moving objects is the first stage of all basic since reduced vision systems process on analytical measurements. Nevertheless, with the sudden changes in different natural scenes such as repetitive movements of people and climate change, the detection of a crowd is a difficult problem to deal with in an effective way. Major research efforts are emerging to fix this problem. Assorted solutions have been suggested as in [22, 26, 35, 44] where the authors relied on the spatio-temporal modeling of the scene of the crowd to localize and discover abnormal events by modeling movement variations of various space time volumes as in including their spatio-temporal statistical behaviors. Other authors using aerial images offer a method based on finding the nearest neighbor relationships for each person in the image to detect crowd regions [25]. In general, the duration of attention of the human being deteriorate after 20 minutes, so, the task of manual surveillance re- quires prolonged and demanding attention [32]. Now, anomaly detection and monitoring of the dynamics of the crowd become increasingly popular.

In the current surveillance systems, there is considerable interest in the analysis of the crowd, especially the study of crowd behavior. Then, we must focus on safety in the crowd in order to detect dangerous situations. But, analysis of crowd behavior hides many problems such as the discovery of the event from the crowd and the change of the crowd or anomaly

detection. In contrast to these two methods [32, 50], which are based on low-level functions to symbolize crowd motion, we suggest a solution based on a GMM as it is tested on each Image location to better model the local distribution of the image.

Detecting people in these scenes relies on detecting movements of each single, then adapting the method of IGMM for the tracking. In the final stage of the detection process of the GMM method[14], a number of characteristics of the object such as the delimitation area, perimeter, and the color histogram of the regions corresponding to objects are extracted from the image using the foreground pixel map.

The foreground extraction step is developed by identifying the pixels of each frame I by comparing each pixel $X_1...X_t$ with each Gaussian. The pixel is defined by its intensity in the RGB color space. Next, the probability is determined from the noticed value of every pixel. As a first step, we compare each pixel to the associated components 'μ' [43]. If this value is close enough to 'μ', then this pixel is considered the appropriate element. This difference must be less than that of the standard deviation of the element measured by the component in the algorithm. As well, we update the mean variance as well as Gaussian weight to find the new pixel value. Thereafter, we determine the components that belong to the background model by applying a threshold to the weights of the 'w' elements. As for the last step, it consists of identifying the pixels of the first plane which do not correspond to the element identified as the background.

$$P(X_t) = \sum_{i=1}^{k} \omega_{i,t} \eta \left(X_t; \mu i, t, \sum_{i,t} \right) \tag{1}$$

Where $w_{i,t}$ is an estimate of the weight (the portion of the data is recorded with this Gaussian) Gaussian th($G_{i,t}$) in the mixture at time t. $\mu_{i,t}$ is the average value of ($G_{i,t}$), $\sum_t^i$ is the covariance matrix and η is a Gaussian probability density.

$$\sum_{i=1}^{k} \omega_{i,t} = 1 \tag{2}$$

The average of the gaussian mixture is:

$$\mu_t = \sum_{i=1}^{k} \omega_{i,t} \mu_{i,t} \tag{3}$$

With

K: the number of distributions
$w_{i,t}$:   the weight of the ith Gaussian
T: the time
$\sum_{i=1}^{k}$:   the standard deviation.

Foreground region detection outputs of algorithms typically contain noise and thus they are not suitable for further processing without a special post-treatment. Morphological operations such as erosion and dilation are operations applicable to the pixel map foreground to eliminate noise that is caused by the elements listed above. Our goal of using these operations consists in:

–   Removing the foreground pixels that do not match the real leading regions;
–   Removing background pixels that are noisy near and within areas of sensing objects.

After that, we use the Blob, which is a mathematical method, on the levels of regions of a digital image that allows the detection and analysis of connected regions [43]. Its a set of connected pixels. Blob's analysis is the study of connected regions in an image. It can detect the position, quantity, orientation and shape in each picture.

Blob's Analysis removes blobs that are of no importance and take only the relevant blobs for a deep analysis. After, the detection of the leading regions and application of post-processing operations to eliminate noise areas and shade, the foreground pixels are filtered by grouping them in related areas (blobs) and labeled using connected component labeling algorithms. After finding the individual blobs that correspond to objects, the bounding boxes of these regions are calculated. Once we have increased the regions, we extract the features of the corresponding objects from the current image. These characteristics are the size (S), the center of mass (Cm) and the histogram color (Hc). The calculation of the size of the object is common and we just count the number of foreground pixels that are contained within the bounding box of the object. To calculate the point, the center of mass = cm $(X_{cm}; Y_{cm})$ of an object O, we use the following equation:

$$\text{Xcm} = \sum_{i^n}^{k} \frac{\sum_i^n}{n}, Y_{cm} = \frac{\sum_{i^{ny1}}}{n} \tag{4}$$

n is the number of pixels in O. The histogram color, HC, is calculated based on the RGB intensity values of object pixels in the current image. To reduce the computational complexity of operations that use the color histogram, color values are quantized. Assuming that N is the number of bins in the histogram, then all lids become 255 / N color values. Therefore, for an object O, the histogram color is updated as follows:

$$\text{H}_c\left(\frac{C_i}{N}, \frac{C_j}{N}, \frac{C_k}{N}\right) = Hc\left(\frac{C_i}{N}, \frac{C_j}{N}, \frac{C_k}{N}\right) + 1\forall_c = ((C_i, C_j, C_k), \in 0 \tag{5}$$

C represents the color value of (i, j, k) th pixel. In the above equation, i, j and k are index variables in three color channels. In the next step; the color his- togram is standardized to

Fig. 1 Result of the bounding box

permit a proper comparison with other histograms in later stages. The normalized histogram Hc is calculated as follows:

$$\hat{H}_c[i][j][k] = \frac{H_{cj}[i][j][k]}{\sum\limits_{ijk}^{n} H_c[i][j][k]} \tag{6}$$

Finally, the resulting histogram is a 3D matrix (N * N * N). After that, we discover the bounding boxes for all connected components equivalent to the motion of people with the Blob analysis of objects (Figs. 1 and 2).

We construct the boundaries of each rectangle that follows each person during his movement as shown in the figure above (Fig. 1). After monitoring the stages of our system for the detection of items, the bounding boxes surround most people with perfect location.

To find the location of each individual in the scene, we based on the result of the selection box (BBOX) to determine the rectangle boundaries for crowd tracking. In case of the detection of two persons whose coordinates of the rectangle encompass this person this coincides with the coordinates with others then we consider him committing a single object as presented in the figure (Fig. 2).

The position of the rectangle that encompasses the crowd is defined as follows:

– ABS, ORD: the upper left corners of the bounding box

$$W = Round\left(\sum bbox(:,3)/size(bbox,1)\right) \tag{7}$$

$$H = Round\left(\sum bbox(:,4)/size(bbox,1)\right) \tag{8}$$

$$ABS = min(bbox(:,1)) \tag{9}$$

When bbox(:,1) represents a vector of values of all abscissa for pedestrian detection

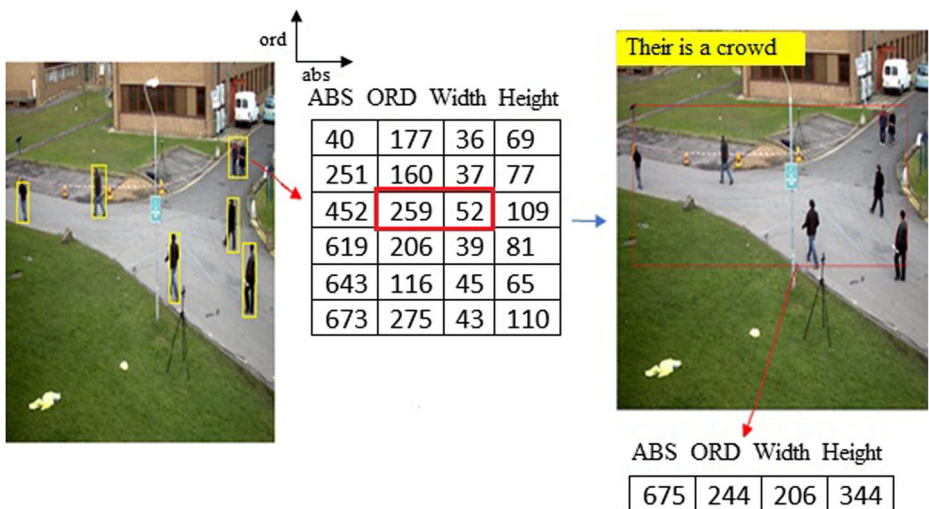$$ABS2 = max(bbox(:,1)) \tag{10}$$

$$ORD = min(bbox(:,2)) \tag{11}$$

Fig. 2 Crowd detection process

$$ORD2 = max(bbox(:, 2)) \tag{12}$$

$$Width = ||[ABS, ORD2 + H; ABS2 + W, ORD2 + H]|| \tag{13}$$

With Width represent the distance between these two terms: ABS,ORD2+H and ABS2+W,ORD2+H.

$$Height = ||[ABS, ORD; ABS, ORD2 + H]|| \tag{14}$$

The final rectangle position encompassing all the people in the scene is:

$$position = [ABS \quad ORD \quad Width \quad Height] \tag{15}$$

3.2 Proposed method for abnormal event's detection

In this paper, we propose a new method based on optical flow characteristics for the abnormal detection of crowd behavior based on two algorithms to detect abnormal behavior in congested scenes.

1. **Method of Lk**

   The method of Lucas Kanade [33] takes into account additional assumptions on a smaller area to particularize the optical flow. It minimizes a criterion on a small area, thus obtaining the optical flow of this small area. This procedure maintains the luminance between the two successive images whose pixels calculate the displacement between the two frames with fixed optical flow in the vicinity. The function to be minimized is:

$$E_{lk} = \sum_{\Omega} \left[\frac{\partial E(x, y, t)}{\partial x} u(t) + \frac{\partial E(x, y, t)}{\partial y} v(t) + \frac{\partial E(x, y, t)}{\partial t} \right]^2 \tag{16}$$

with

$$\begin{cases} u(x, y, z) = u(t) \\ v(x, y, z) = v(t) \end{cases} (x, y) \in \Omega \tag{17}$$

2. **Method of HS**

   To achieve this optical flow estimation, we obtain a formula which combines the luminous variation of the frame with a point displacement pattern containing the same luminance distribution [15]. At a point of brightness distribution (x, y) of the image at the time t, E (x, y, t), when this structure moves, the luminosity of an object caught is constant, thus:

$$\frac{dE}{dT} = 0 \tag{18}$$

The use of the rule of differentiation chain gives us

$$\frac{\partial E}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial E}{\partial y} \frac{\partial y}{\partial t} + \frac{\partial E}{\partial t} = 0 \tag{19}$$

$$E_x u + E_y v + E_t = 0 \tag{20}$$

We have also filed in with additional abbreviations such as: Ex and Ey, for the incomplete derivatives of the brightness of the frame concerning x, y and t. The constraint in the equation in another way is,

$$(E_x, E_y)(u, v) = -E_t \tag{21}$$

3.2.1 Proposed work

In this section we present our proposed descriptor as mentioned earlier for anomaly detection based on both results of optical flow for the method of Lk and HS, illustrated in Fig. 3. We divide the frame into 8.8 blocks for each line.

As previously mentioned, optical flow information (magnitude) is used to construct the feature vector for each global. To do this, we calculate the optical flow field between the frame t − 1 and t. We obtain a matrix A; then, we compute it (1 using the method of LK, 2 using the method of HS) between the frame t and t + 1. We obtain the matrix B.

We construct a feature matrix based on the magnitude provided by the vector field resulting from the optical flow (note that the magnitude of the optical flow indicates the speed of movement of the pixel).

As we mentioned earlier, assuming that the brightness of each moving pixel in the frame is constant,

$$v = \begin{bmatrix} d_x \\ d_y \end{bmatrix} \tag{22}$$

With d_x and d_y are its displacements along the axis x and y.

To a given motion vector v, we compute the magnitude Mgn as follows:

$$Mgn = \sqrt{dx^2 + dy^2} \tag{23}$$

The motion field is a rich dynamic descriptor of the flow which is related to flow velocity, density and pressure in the flow. However, the optical flow is to compute pixel wise instantaneous motion between two consecutive frames. The speed vectors are depicted by the colors shown in this circle. All velocity vectors are coded by the color it means by placing its origin at the center of the circle. Depending on this angle, the point takes a different color, its intensity varies from black to full color according to the norm of the vector. The maximum speed is represented by white dots. This mapping finally returns to build an application that combines the angle and standard velocity vector at a point in space (RGB). This type of representation can give a color for each pixel, and can quickly view the consistency of results.
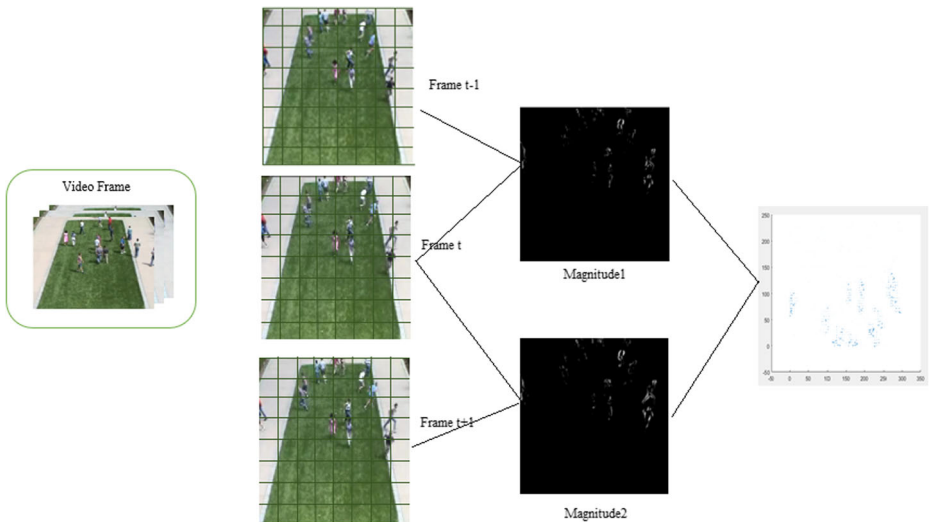

Fig. 3 Extraction process of the feature

Each block contains a set of interest points. We are only interested in a sufficient number of these points. So, we apply a filtering process in order to discards irrelevant points that does not contain movements. We note that an interfering movement is marked in white.

We cannot, however, determine the component of motion at right angles with the component in the direction of the gradients.

Let us denote the motion features in term of velocity magnitude mgn at any time t. For an the optical flow field (u,v) of each pixel following (17), the velocity magnitude mgn for a 8*8 cell can be calculated as:

$$mgn = 1/64 \ (23) \tag{24}$$

With $c(mgn(c, t))$ is the probability of observing the current cell c at the time t.

We begin with the initialization of the the Sum=0 at the t=1, then we calculate the sum of white intensity of each cell after the processing of filtration of the second obtained image:

$$Sum = Sum + \sum S \tag{25}$$

With S representing the sum of the white intensity of each cell in the image.

Consequently, the speed can not be calculated without offering an additional imposition. To view the route of optical flow field, we use the image of velocity vectors, which can depict a color to symbolize the flow direction and its intensity as shown in figure (Fig. 4).
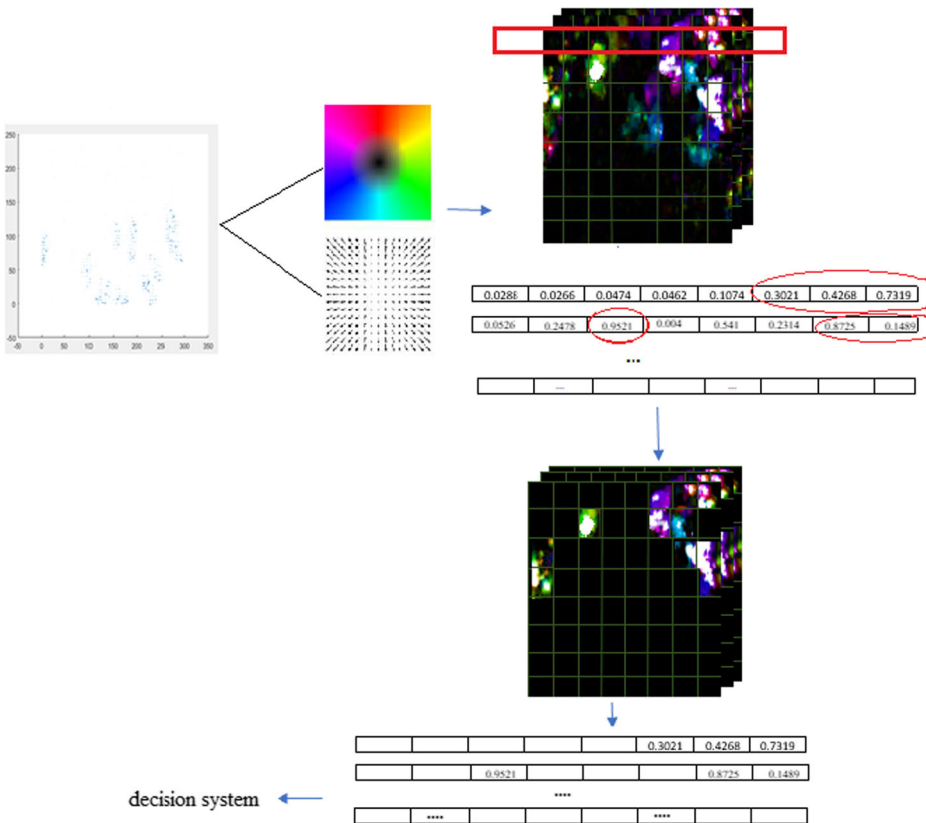

Fig. 4 Abnormal crowd detection process

A suitable threshold is used to generate an abnormality frame; Following [30] if the light intensity (see Fig. 5) is very important, that is, greater than the maximum of threshold; then, we can notice that there is an abnormal move- ment in the scene. Else, if the value of the intensity is less than the minimum of the threshold, we label it as a normal event (experimental results). Correct detection should satisfy the fact that light intensity should surpass the threshold: For a visual representing, the route of the optical flow field are used in the vectors of the original image with superimposed speeds, can further be used to map a color to symbolize the direction of flow as well as its intensity.

4 Experimental results and discussion

In this section, we present, first, the benchmark datasets all throughout the surveillance to evaluate the performance of our proposed methods. Then, the experimental protocol is detailed. Finally, the obtained results are analyzed and compared to other methods.

4.1 Datasets

Our surveillance system is of the "smart camera" capable of detecting an abnormal situation automatically. For the detection of the crowd, we improved the GMM [14] method to extract various useful features for time tracking of a crowd. Then, we tested our approach on three databases that contain various human activities. These activities included walking and running: Pets' [38], Avenue [27] and UMN database [49].

We used 2 classes from Pets' datasets: S1 and S2 for walking and S3 for event recognition.

The resolution of video frame is 768*576.

This dataset contains different crowd activities and the task is to provide a probabilistic estimation of each of these activities: walking, running, evacuation (rapid dispersion), local dispersion, crowd formation and splitting at different time instances.

The UMN database is publicly accessible to videos of normal and abnormal crowd at the University of Minnesto. It includes videos of 11 different scenarios in three different indoor and outdoor stages. Each video consists of an initial part of normal behavior and ends with sequences of abnormal behavior. The video size is of 320×240 pixels. The movements of persons are essentially multi-directional.

The Avenue dataset includes 15 sequences, each one is about 2 minutes long. It contains in total 14 abnormal events including throwing objects, loitering and running.

4.2 Experimental protocol and parameters' setting

We adopted as evaluation measures for the Area Under ROC curve (AUC). The AUC is derived from the ROC curve which is very often used to compare the performances of

Fig. 5 Light intensity range

Table 1 Efficiency tests on different datasets (unit: ms/frame)	UMN dataset	Method			
		Ours	STT	STC	IBC
	Scenario 1	15.4	15	17	1818
	Scenario 2	17.2	18	19	2200
	Scenario 3	15.6	16	15	2712

different methods. For a good recognition algorithm, larger the AUC accuracy is, he better the algorithm is.

In our experiments, we have made few attempts to optimize the parameters used in our method, and so improved performance may be obtained by exploring other values. The DMOF is performed on each 8*8 cell and updated every 3 frames.

In our proposed method, two parameters are needed to be adjusted. The first one corresponds to pyramid parameters of LK. We define a height Lm of pyramid (in application, $Lm = 2, 3$ or 4). At each level of the pyramid, the image is sub-sampled by a factor of 2 for the two successive images considered.

Second, we empirically fixed the threshold that differs the abnormal event from a normal event. We calculate all the numbers of pedestrians in each frame by following this equation:

$$Size = rg(bbox) \tag{26}$$

Then we multiply the result obtained by calculating the sum of luminous intensity for a single pedestrian who is walking three times to obtain the sum of intensity in a running rain (max of threshold).

4.3 Results and analysis for crowd detection

During the evaluation, for the processes of detection of the crowd; we compared our work with other benchmarking approaches such as composition inference (IBC) [5], spatio-temporal compositions (STC) [41] and spatio-temporal texture (STT) [51]. The STC highlights its real-time performance like the IBC which has been advised as one of the most accurate methods for anomaly detection. The STT algorithm developed its robustness in performance compared to other conventional image/frame-based approaches. All of these methods have been tested under the same hardware configurations and software settings.

The temporal consumption of our algorithm has been compared with popular approaches as illustrated in the following table.

The time used by our algorithm for crowd detection has been compared with well-know approaches as presented in Table 1.

In the table, IGMM for crowd tracking performs faster than other benchmarking approaches.

We also provide in Table 2 a comparison of the execution time measured in processed s/frames for several anomaly detection methods. Noting that these numbers correspond to

Table 2 Frames per second processing of various methods	Method	BM [55]	SS	LOF [37]	Our
	FPS	1.037	5	1.302	1.054
	CPU	3.16	3	2.5	2.3
	Platform	Matlab	n.c	c++	Matlab

Fig. 6 Our proposed approach to the detection of the crowd

different implementations, so this comparison is only indicative. Indeed, the size of the images may vary as well as the execution time, which may not encompass the entire flow, which causes a comparison of the calculation time. For example, the run-time of the SS [17] method does not include the calculation of optical flow fields. Nevertheless, our method is prone to treat abnormality detection more efficiently than other methods.

Our method of detection of the crowd has been successfully tested on video sequences. This method gives the result in a display of a yellow bar at the top left if there is a crowd (This is a crowd), otherwise it displays this message (No crowd).

The detected and tracked subjects are marked by a red rectangle on the output frame. It is built using the bounding boxes that define all related components equivalent to a movement of a pedestrian with Blob Analysis object. As shown in Fig. 6, our system detects the
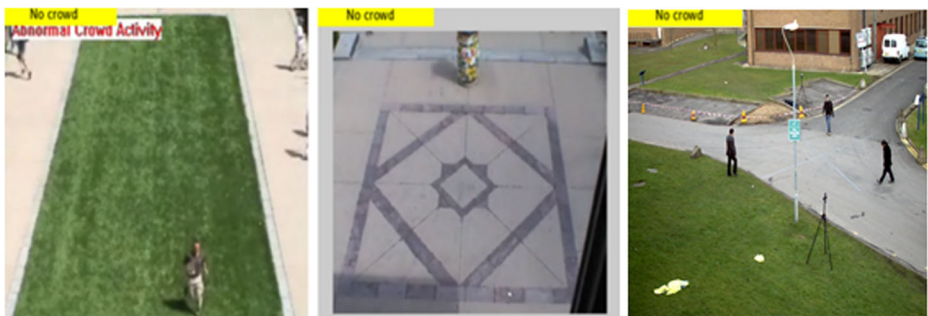

Fig. 7 Our proposed approach to the detection of the crowd (in case of the no existence of crowd
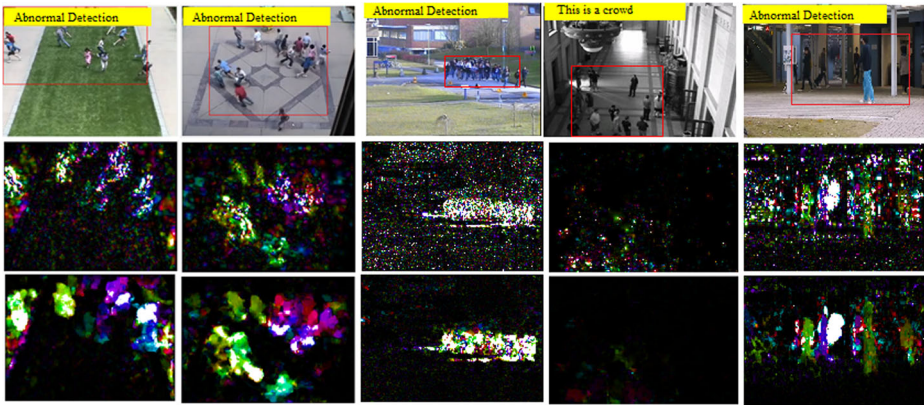
Fig. 8 Sample snapshots of abnormal events' detection: first line:using HS second line: using LK

existence of a crowd and surrounded it with a rectangle. The dimensions of this rectangle depend on the movement of pedestrian in scene.

In Fig. 7, the system displays an alert at the top of each window to describe the absence of crowds. If the system detects the presence of the crowd, we go to the next step which is the step of detecting abnormal movements in the crowd if the detection process is completed.

4.4 Results and analysis for abnormal detection

Our approach is tested on movies directed by a vision of the day, but we might consider an application running in night vision with the use of cameras with infrared illumination and does not affect the quality of work since our application is based on the estimate of the intensity of movements and lighting has no effect on the application. In order to detect
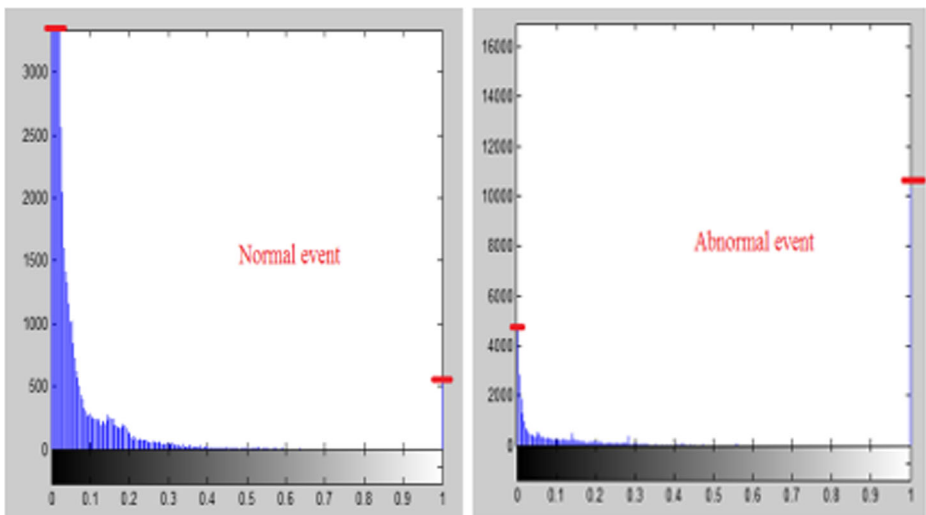

Fig. 9 The histogram showing the variation of normal and abnormal events
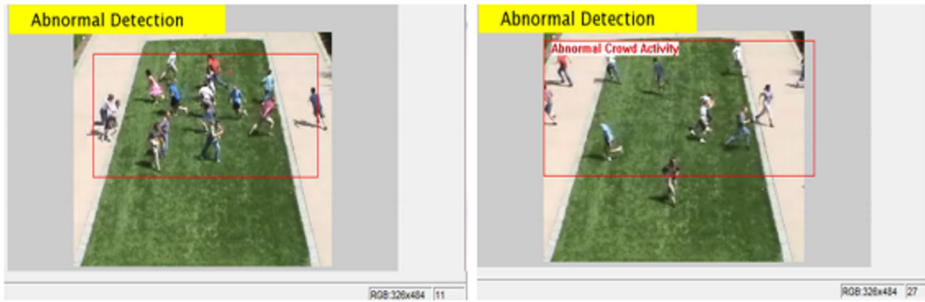
Fig. 10 Comparisons of our method with other approaches in terms of speed

abnormal events, we rely on the calculation of the intensity movements of each person in motion, using the calculation of the optical flow between two successive images (Fig. 8).

These figures show the influence of the light intensity in three datasets.

The results provided by the implementation of the method of HS is very sensitive to noise. Moreover, the overall nature of this method gives it a very smooth character, resulting in an inaccuracy of the result, notably an absence of identification of local movements. This explains the success of the LK method in making rapid coarse estimates of the optical flow. The best result is therefore given by the refined pyramidal implementation of LK (5 levels of pyramid, 10×10 patch size and 10 iterations) to detect abnormal movements in public places.

In addition, the use of the LK technique is better performed because it divides the video frame into small portions and solves it with respect to HS which assumes that the optical flow is smooth over the whole image.

The results obtained with the Lk method are very satisfactory, the contours of the objects are very well respected with good segmentation. Finally, the flow is dense and as accurate as possible.

The histogram drawn in Fig. 9 shows the variation in luminous intensity, in order to explain the distribution of the intensities (or colors) of each frame.

Generally, surveillance video needs security agents in order to control its different streams. In public places, more control and speed are needed when detecting abnormal events.In this field, our work has been improved in a way that it outperforms other similar

	Methods	Area under ROC
Table 3 Anomaly detection results on the whole UMN dataset	Pur optical flow [34]	0.840
	BM [54]	0.964
	FF [48]	0.810
	SF [53]	0.850
	OFF [11]	0.964
	CI [53]	0.879
	SRC [24]	0.847
	Biswas [4]	0.983
	Ours(HS)	0.930
	Ours(LK)	0.984

Table 4 Accuracy (%) comparison for abnormal crowd behavior detection in the UMN dataset

	Ours LK	OFF	BM	FF	CI	SF	SRC
View 1	98.84	99.10	99.03	88.69	90.62	84.41	90.52
View 2	97.72	94.85	95.36	80.00	85.06	82.35	78.48
View 3	98.7	97.76	6396.	77.92	91.58	90.83	92.70

Table 5 Accuracy (%) comparison of the proposed methods with OFF, BM, FF, CI and SF for crowd escape behavior detection in PETS2009 dataset scenario 1

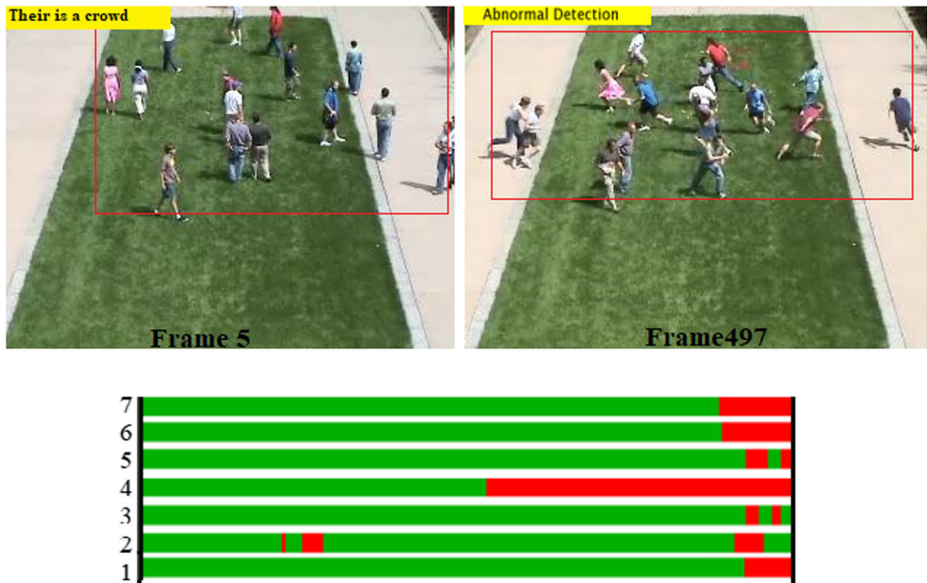|	Proposed LK	Proposed HS	OFF	BM	FF	CI	SF
view 1	99.34	99.01	99.07	92.45	37.74	56.60	63.21
view 2	98.64	97.53	98.13	83.02	37.74	83.02	70.76
view 3	84.21	82.32	62.62	89.62	37.74	81.13	52.83
view 4	89.53	84.019	97.20	90.57	37.74	52.83	48.11

Fig. 11 Crowd escape behavior detection for scene 1 in the UMN dataset. (7) Ground truth, (6) OFF method, (5) result of BM, (4) result of FF, (3) result of CI, (2) result of SF, (1) Our method

Table 6 Accuracy (%) comparison for crowd escape detection of the PETS database for scenario 2

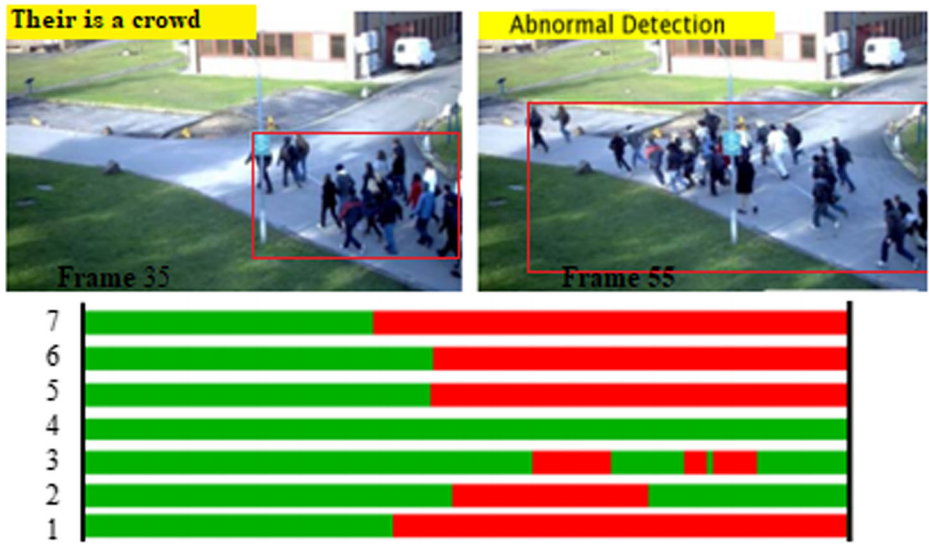|	Proposed LK	Proposed HS	OFF	BM	FF	CI	SF
view 1	99.00	98.20	98.66	96.01	94.50	94.95	91.22
view 2	99.70	99.42	99.20	94.15	63.83	92.02	89.36
view 3	99.34	99.47	99.47	95.21	95.48	94.15	94.68
view4	87.72	85.39	89.57	91.49	96.81	89.36	64.63

Fig. 12 Crowd escape detection of the PETS database for scenario 1 view 1. (7) Terrestrial Truth, (6) OFF, (5) BM, (4) FF, (3) CI, (2) SF and (1) proposed method DMOF

works. This is foreseen by the following figure which shows the speed of our approach compared to other approaches. And to further support our approach, we can see that the abnormal event is detected from the frame number 11, but in another algorithm [34] the abnormal movement is detected in frame number 27 (Fig. 10).

This experiment aims to assess the performance of anomaly detection along the video clip.

Thereafter, we calculate the real positive rate and the false positive rate in the movie. We compare our method with state of the art in detection performance on UMN dataset (see Table 3); our measure achieved higher accuracy in the anomaly detection, better than all the state of the art methods.

The recognition results of our method for the detection of abnormal movements appear to be very reliable and reach 95.5%.

Table 4 shows the accuracy comparison of six methods for the three views in the first and second scenario of UMN datatset.
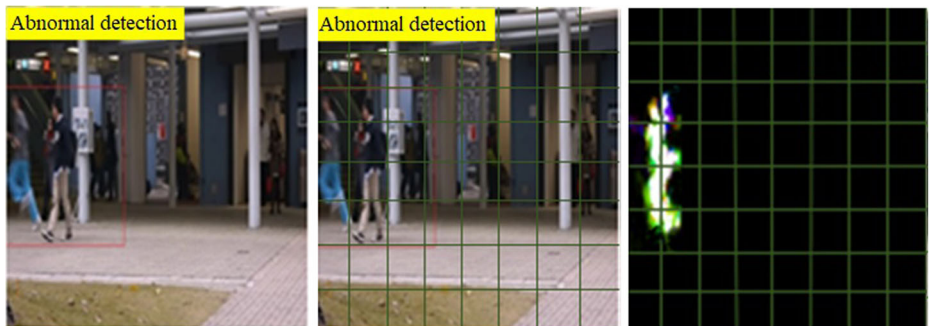

Fig. 13 An abnormal event and its corresponding abnormal patches in the Avenue dataset

Table 7 Comparison with other sparsity-based method on the Avenue dataset		WD	LT	MISK	Total	FA
	Ground Truth	9	3	7	19	0
	Avenue [28]	9	3	7	19	2
	Our LK	8	3	7	18	0
WD wrong direction, LT loitering, FA false alarm	Our HS	8	3	7	18	1

The proposed feature detection DMOF for abnormal event detection performs and is better than other techniques. It's achieves the best accuracy with an average of 98.4% using our DMOF + LK representation, which is higher than the accuracy of OFF (96.46%), CI (87.91%), FF (81.04%), SRC (84.70%), SF (85.09%) and BM (96.40%) ...

The accuracy of this methods below the same evaluation settings was taken from [10].

For the Pets dataset, we compare our methods with others methods. We follow the same evaluation settings as outlined in [11] (Table 5).

In the second scenario, we again follow the same evaluation settings as described in [11]. In the second scenario, we do that for each view.

In this dataset, especially for views three and four, the pedestrians are occluded by a tree.

Additional, as soon as the crowd appear, from these camera angles, the crowd get away and start quickly, that's why the rate was decreased.

Our algorithm DMOF for the second scenario achieved the second best accuracy with an average of 96.44 %, compared to the other results of OFF 96.72%, BM (94.22%), SF (84.97%), CI (92.62%) and FF (87.66%).

The abnormal event detection results for a frame of Avenue dataset is shown in Fig. 11. The experimental results in Table 6 lists the comparison with other methods and shows that DMOF descriptor performs better than other methods, especially using the LK method for optical flow representation.

For different databases, the feature of DMOF + LK is better than DMOF + HS.

This method greatly reduces the cumbersome intermediate operation compared with other methods (Figs. 12, 13 and Table 7).

5 Conclusion

In this paper, a real time tracking people and detection of abnormal events in crowded scenes is described. Our algorithm is based on the criterion of light intensity which corresponds to the rate of movement of individuals. A method based on the Gaussian mixture model was tested and improved to try to deduce better detection and tracking of the crowd. Then, a novel optical flow magnitude of features for abnormal crowd behavior is presented. This approach can be used to exactly predict abnormal events. It's used successfully in the 11 different scenarios of the UMN database. Scenario 1,2 and 3 Pets' dataset and Avenue database. The advantages of our proposed methods are that no longer contains a learning phase and that they supply a method for the temporal tracking over time. To test the detection accuracy of the proposed method prove that the method of LK is rapid in estimating optical flow and in detecting abnormal events and it is better than many previous methods.

References

1. Adam A, Rivlin E, Shimshoni I, Reinitz D (2008) Robust real-time unusual event detection using multiple fixed-location monitors. IEEE Trans Pattern Anal Mach Intell 30:555–560
2. Benezeth Y, Jodoin PM, Saligrama V, Rosenberger C (2009) Abnormal events detection based on spatio-temporal co-occurences. In: IEEE conference on computer vision and pattern recognition, CVPR, pp 2458–2465
3. Benezeth Y, Jodoin P-M, Saligrama V (2011) Abnormality detection using low-level co-occurring events. Pattern Recognit Lett 32(3):423431
4. Biswas S, Gupta V (2017) Abnormality detection in crowd videos by tracking sparse components. Mach Vis Appl 28(1-2):35–48
5. Boiman O, Irani M (2007) Detecting irregularities in images and in video. Int J Comput Vis 74(1):17–31
6. Cao T, Wu X, Guo J, Yu S, Xu Y (2009) Abnormal crowd motion analysis. In: IEEE international conference on robotics and biomimetics (ROBIO) 2009, IEEE, Guilin, pp 1709–1714
7. Chaturvedi PP, Rajput AS, Jain A (2013) Video object tracking based on automatic background segmentation and updating using RBF neural network. International Journal of Advanced Computer Research 3:866
8. Colque RVHM, Caetano C, De Andrade MTL et al (2017) Histograms of optical flow orientation and magnitude and entropy to detect anomalous events in videos. IEEE Trans Circuits Syst Video Technol 27(3):673–682
9. Cong Y, Yuan J, Liu J (2011) Sparse reconstruction cost for abnormal event detection. In: Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, 2011. pp 3449–3456
10. Direkoglu C, Sah M, O'Connor NE (2017) Abnormal crowd behavior detection using novel optical flow-based features. In: 14th IEEE international conference on advanced video and signal based surveillance (AVSS), 2017. IEEE
11. Direkoglu C, Sah M, O'Connor NE (2017) Abnormal crowd behavior detection using novel optical flow-based features. In: Advanced Video and Signal Based Surveillance (AVSS), 2017 14th IEEE International Conference on. IEEE, 2017. pp 1–6
12. Drews P, Quintas J, Dias J et al (2010) Crowd behavior analysis under cameras network fusion using probabilistic methods. In: 13th conference on information fusion (FUSION), p 18
13. Fang Z, Fei F, Fang Y et al (2016) Abnormal event detection in crowded scenes based on deep learning. Multimedia Tools and Applications 75(22):14617–14639
14. Hauhan AK, Krishan P (2013) Moving object tracking using gaussian mixture model and optical flow. International Journal of Advanced Research in Computer Science and Software Engineering 3(4)
15. Horn BKP, Schunck BG (1981) Determining optical flow. Artif Intell 17(1–3):185–203
16. Hu M, Ali S, Shah M (2008) Detecting global motion patterns in complex videos. In: 19th international conference on pattern recognition. ICPR 2008. IEEE, pp 1–5
17. Hu Y, Zhang Y, Davis L (2013) Unsupervised abnormal crowd activity detection using semiparametric scan statistic. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 767–774
18. Junior J, Mussef S, Jung C (2010) Crowd analysis using computer vision techniques. IEEE Signal Process 27:66–77
19. Kaviani R, Ahmadi P, Gholampour I (2014) Incorporating fully sparse topic models for abnormality detection in traffic videos. In: Proceeding of the international econference on computer and knowledge engineering, Mashhad, Iran, pp 586–591
20. Khatrouch M, Gnouma M, Ejbali R, Zaied M (2017) Deep learning architecture for recognition of abnormal activities. In: The 10th international conference on machine vision, Vienna, Austria
21. Kim J, Grauman K (2009) Observe locally, infer globally: a space-time MRF for detecting abnormal activities with incremental updates. In: Proceeding of the IEEE conference on computer vision and pattern recognition, pp 2921–2928
22. Kratz L, Nishino K (2010) Tracking with local spatio-temporal motion patterns in extremely crowded scenes. In: IEEE Conference on computer vision and pattern recognition (CVPR), 2010. IEEE, pp 693–700
23. Kratz L, Nishino K (2012) Tracking pedestrians using local spatio-temporal motion patterns in extremely crowded scenes. IEEE Trans Pattern Anal Mach Intell 34(5):9871002
24. Li A, Miao Z, Cen Y et al (2017) Anomaly detection using sparse reconstruction in crowded scenes. Multimedia Tools and Applications 76(24):26249–26271
25. Li J, Hospedales TM, Gong S, Xiang T (2010) Learning rare behaviours. In: Proceeding of the Asian conference on computer vision, Queenstown, New Zealand, pp 293–307

26. Li W, Mahadevan V, Vasconcelos N (2014) Anomaly detection and localization in crowded scenes. IEEE Trans Pattern Anal Mach Intell 36(1):18–32
27. Lu C, Shi J, Jia J (2013) Abnormal event detection at 150 fps in matlab. In: IEEE international conference on computer vision (ICCV), 2013, IEEE 27202727
28. Lu C, Shi J, Jia J (2013) Abnormal event detection at 150 fps in matlab. In: Proceedings of the IEEE international conference on computer vision, pp 2720–2727
29. Lui X, Rittscher J, Perera A, Krahnstoever N (2005) Detecting and counting people in surveillance applications. In: Advanced video and signal based surveillance, pp 306–311
30. Mahadevan V, Li W, Bhalodia V, Masconcelos V (2010) Anomaly detection in crowded scenes. In: Proceeding of the IEEE conference on computer vision and pattern recognition
31. Mariem G, Ridha E, Mourad Z (2016) Detection of abnormal movements of a crowd in a video scene. International Journal of Computer Theory and Engineering 8(5):398
32. Marques JS, Jorge PM, Abrantes AJ, Lemos JM (2003) Tracking groups of pedestrians in video sequences. In: IEEE computer society conference on computer vision and pattern recognition workshops, vol 9, p 101101
33. Marzat J (2008) INRIA - Estimation temps réel du flot optique, ISA
34. Mehran R, Oyama A, Shah M (2009) Abnormal crowd behavior detection using social force model. In: IEEE conference on computer vision and pattern recognition, 2009. CVPR 2009. IEEE, pp 935–942
35. Mousse AM (2016) Reconnaissance dactivits humaines partir de squences multi-camras: application la detection de chute de personne. Universit du Littoral-Cted'Opale
36. Nam Y (2014) Crowd flux analysis and abnormal event detection in unstructured and structured scenes. Multimedia Tools and Applications 72(3):3001–3029
37. Pérez-Rúa J-M, Basset A, Bouthemy P (2017) Detection and localization of anomalous motion in video sequences from local histograms of labeled affine flows. Frontiers in ICT 4:10
38. PETS Dataset. http://www.cvg.reading.ac.uk/PETS2009/a.html
39. Rao AS, Gubbi J, Rajasegarar S, Marusic S, Palaniswami M (2014) Detection Of anomalous crowd behaviour using hyperspherical clustering. In: International conference on digital lmage computing: techniques and applications (DlCTA). IEEE, pp 1–8
40. Rodrigues de Almeida I, Jung CR (2013) Change detection in human crowds. In: Proceeding of the conference on graphics, patterns and images, pp 63–69
41. Roshtkhari MJ, Levine MD (2013) An on-line, real-time learning method for detecting anomalies in videos using spatio-temporal compositions. Comput Vis Image Underst 117(10):1436–1452
42. Ryan D, Denman S, Fookes C, Clinton B, Sridharan S (2011) Textures of optical flow for real-time anomaly detection in crowds. In: 2011 8th IEEE international conference on advanced video and signal based surveillance (AVSS), pp 230–235. https://doi.org/10.1109/AVSS.2011.6027327
43. Santosh DHH, Venkatesh P, Poornesh P, Rao LN, Kumar NA (2013) Tracking multiple moving objects using gaussian mixture model. International Journal of Soft Computing and Engineering (IJSCE) 3(2):114–119
44. Shah AJ (2016) Abnormal behavior detection using tensor factorization. (Doctoral Dissertation, Cole de Technologie Suprieure)
45. Shi Y, Liu Y, Zhang Q, Yi Y, Li W (2016) Saliency-based abnormal event detection in crowded scenes. J Electron Imaging 25(6)
46. Tu P, Sebastian T, Doretto G, Krahnstoever N, Rittscher J, Yu T (2008) Unified crowd segmentation. In: European conference on computer vision, vol 5305, pp 691–704
47. Tziakos I, Cavallaro A, Xu LQ (2010) Local abnormality detection in video using subspace learning. In: Seventh IEEE international conference on advanced video and signal based surveillance (AVSS), 2010, pp 519–525
48. University of Reading, PETS 2009 Dataset S3 Rapid Dispersion, available from http://www.cvg.reading.ac.uk/PETS2009/a.html
49. Unusual crowd activity dataset of University of Minnesota, from http://mha.cs.umn.edu/movies/crowdactivity-all.avi
50. Varadarajan J, Odobez J-M (2009) Topic models for scene analysis and abnormality detection. In: Proceeding of the international conference on computer vision workshops, Kyoto, Japan, pp 1338–1345
51. Wang J, Xu Z (2016) Spatio-temporal texture modelling for real-time crowd anomaly detection. Comput Vis Image Underst 144:177–187
52. Wang T, Snoussi H (2014) Detection of abnormal visual events via global optical flow orientation histogram. IEEE Trans Inf Forensics Secur 9(6):988–998
53. Wu S, Moore BE, Shah M (2010) Chaotic invariants of lagrangian particle trajectories for anomaly detection in crowded scenes. In: IEEE conference on computer vision and pattern recognition, pp 2054–2060

54. Wu S, Wong HS, Yu Z (2014) A bayesian model for crowd escape behavior detection. IEEE Trans Circuits Syst Video Technol 24(1):85–98
55. Xu D, Song R, Wu X, Li N, Feng W, Qian H (2014) Video anomaly detection based on a hierarchical activity discovery within spatio-temporal contexts. Neurocomputing 143:144–152
56. Zhan B, Monekosso DN, Remagnino P, Velastin SA, Il-Qun X (2008) Crowd analysis: a survey. Mach Vis Appl 19(5–6):345–357
57. Zhang T, Yang Z, Jia W et al (2016) A new method for violence detection in surveillance scenes. Multimedia Tools and Applications 75(12):7327–7349
58. Zhang Y, Qin L, Yao H et al (2012) Abnormal crowd behavior detection based on social attribute-aware force model. In: 19th IEEE international conference on image processing (ICIP), 2012. IEEE, pp 2689–2692

Mariem Gnouma received the license degree in Computer Science from Faculty of Sciences of Gabes (FSG), and the master degree of computer science and multimedia of the higher Institute of computer and Multimedia of Gabes (ISIMG) respectively in 2012,2015. Currently, she is pursuing the Ph.D. degree in Computer Engineering in the Research Team in Intelligent Machines (RTIM), University of Gabes.

Ridha Ejbali received the Ph.D. degree in computer engineering, the Master degree and computer engineer degree from the National Engineering School of Sfax Tunisia (ENIS) respectively in 2012, 2006 and 2004. He was an assistant technologist at the Higher Institute of Technological Studies, Kebili Tunisia since 2005. He joined the Faculty of Sciences of Gabes Tunisia (FSG) where he is an assistant in the Department computer sciences since 2012. His research area is now in pattern recognition and machine learning using Wavelets and Wavelet networks theories.
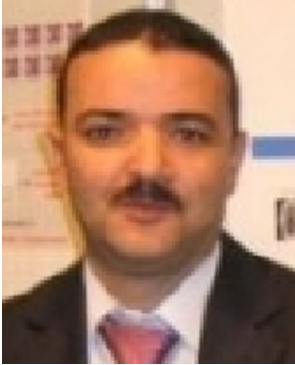
Mourad Zaied received the HDR, the Ph.D. degrees in computer engineering and the master degree of science from the National Engineering School of Sfax respectively in 2013, 2008 and in 2003. He obtained the degree of computer engineer from the National Engineering School of Monastir in 1995. Since 1997 he served in several institutes and faculties in university of Gabes as teaching assistant. He joined in 2007 the National Engineering School of Gabes (ENIG) as where he is currently an associate professor in the Department of Electrical Engineering. His research interests include computer vision and image and video analysis. These research activities are centered around wavelets and wavelet networks and their applications to data classification and approximation, pattern recognition and image, audio and video coding and indexing. He was the chair of the Workshop on Intelligent Machines: Theories & Applications (WIMTA II 2009) and he organized two winter schools on wavelet and its applications (2005) and on Matlab toolkits (2004).