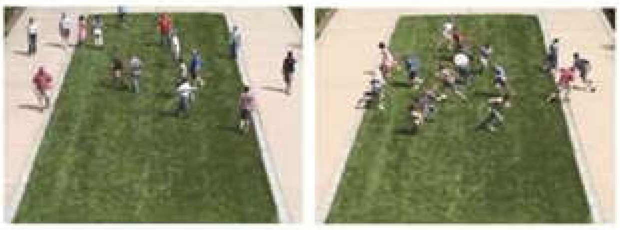
DETECTION AND LOCALIZATION OF UNUSUAL HUMAN ACTIVITY IN CROWDED SCENES BY USING A MOTION INFLUENCE MAP

[1]Thaviti. Venkateswarlu, [2]B. Vasanthi, [3]Y.Swamy, [4]Alikepalli Sumithra
[1,2,3]Assistant Professor, [4]PGStudent, Dept. of Master of Computer Application, Newton's Institute of Engineering, Macherla, Andhra Pradesh, India.

ABSTRACT
In this study, we offer a unique approach for detecting atypical human behavior in crowded environments or at home. In particular, we developed an effective approach, dubbed a motion influence map, for portraying human activities instead of identifying or segmenting persons. The suggested motion influence map's primary advantage is that it accurately captures the motion characteristics of the size, speed, and direction of the objects or subjects, as well as their interactions within a frame sequence.
Index Terms:-Unusual activity detection, vision-based surveillance, motion influence map, crowded scenes/home.

INTRODUCTION
A significant number of surveillance cameras have been deployed in both public and private spaces due to the growing relevance of security. The number of video surveillance systems, however, is overwhelming the staff members in charge of keeping them under control. A smart surveillance system that can automatically identify strange or abnormal actions has garnered a lot of interest in this regard. The detection of human activity and human-human interaction in video sequences has drawn the attention of several researchers in computer vision and pattern recognition over the past few decades [1], [2], and [3]. Lately, academics have been more interested in the identification of anomalous or unexpected behavior in crowded environments. Unlike human action or interaction recognition, conventional methods are not applicable to the detection and/or tracking of human subjects in a crowded scene owing to the presence of occlusions, small objects sizes, and other factors. For unusual activity detection in a crowded scene, texture information such as a spatio-temporal gradient [4], mixture of dynamic textures [5], and spatio-temporal frequency [6], [7] has been considered an efficient means of detection. In the meantime, other groups have used optical flows that directly characterize motion features in a sequence, e.g., a motion heat map [8], clustered motion patterns [9], [10], spatial saliency of the motion feature [11], crowd predictionusing a force field model [12], optical flow fields [13], particle trajectory [14], a social force model [15], and a local motion histogram [16].

(a) Local usual activity (b) Local unusualactivity: bicycle in the middle of the frame



(c) Global usual activity (d) Global unusualactivity: running people across the frame Fig. 1. Examples of two unusual activities.(a) Persons walk in either side. (b) Riding a bicycle passing through the walking persons. (c) Persons walking around in the scene. (d) Every person starts to run suddenly. Although motion flow based approaches have shown their efficacy in previous works, we believe it is still important to consider the information on the size of the objects and their interactions. For example, in Fig. 1b, where riding a bicycle is considered an unusual activity, the size of the object and its effect to the nearby pedestrians' moving directions are important information along with themovement speed. To the best of our knowledge, none of the previous methods has explicitly considered this information,the use of which can be helpful in enhancing the performance. However, as stated above, owing to the inapplicability of human segmentation and tracking in a crowded scene, an alternative approach is needed. In this paper, we propose a novel method to represent the motion characteristics of moving objects by considering their motion flows, sizes, and interactions, simultaneously. Specifically, we define a "motion influence map" that efficiently depicts the underlying motion patterns in a crowded scene/home.

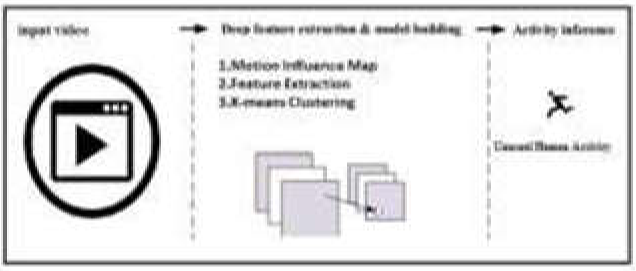
RELATED WORK

Abnormal event or activity detection has recently gained great interest from researchers in vision-based surveillance. Xiang et al. addressed the problem of behavior modeling for surveillance videos [17]. Anomalies were detected by means of the likelihood ratio test with normal behavior classes of an individual person, which were modeled in an unsupervised learning. Jiang et al. proposed a new framework for anomaly detection using aspatio-temporal context [18]. Theypresented instant behaviors of a singleobject using an atomic event, which contained the location, movement direction, and velocity of an object. Normal eventswere described using a combination of atomic events under three categories.Anomalous activities in a spatio-temporalcontext were detected based on a sequence of atomic events. Owing to the huge variations in appearance, scale, illumination, and pose, it is difficult to detect or track individual persons within crowded scenes, and the above-mentioned methods are therefore not applicable to such a scenario. To this end, recent researches have focused on the direct use of motion patterns in an image. Wang et al. used Kanade- LucasTomasi (KLT) corners [19] to represent moving objects and clustered similar motion patterns in an unsupervised manner [9]. They detected anomalies in a frame sequence using two types of historical motion descriptors: the self history and the neighboring history [10]. Xiong et al. proposed a cameraparameter independent method by counting people [20]. They used both an optical flow and a foreground distribution. The kinetic energy was measured using an optical flow to distinguish running activities from walking activities, and a crowd index distribution,which was defined by the foreground pixel distribution values, was also measured todetect the gathering and scattering activities.Some other researchers have focused on crowd behavior modeling, which has beenan interesting research issue in various fields [21], [22], [23], [24]. A number of techniques have been adopted for global unusual activity detection by modeling the behavior of the crowd itself. Mehran et al. described crowd behaviors by means of the social force model [22], with no human detection or tracking processes involved [15]. They measured the interaction force bycomputing the difference between the desired and actual velocities obtained from the particle advection on the optical flowfield [14], [25]. Latent Dirichlet allocation was also used to discover the distribution of normal behaviors based on the social force. Cui et al. considered social behavior and its action using the interaction energy potential [26]. They detected the space-time interest points [27] and tracked them using a KLT feature tracker [19] to obtain human motion within a video sequence. The interaction energy potential was estimated from the velocity of the space-time interest points to explain whether they will meet in the near future [28]. Meanwhile, other research groups have concentrated more on local unusual activity detection. Mancas et al. quantifiably represented the global rarity to select irrelevant motions from the spatial context using bottom-up saliency [11]. Theymeasured the saliency index in multiple channels, which consisted of different speeds and directions. Local unusualactivity was finally detected using the corresponding saliency maps. Ihaddadene etal. observed motion variations of a set of interest points [8]. They built a motion heat map based on the motion intensities, andmade a comparison with the variations in local motion. Mahadevan et al. modeled the appearance information and dynamics of normal behavior in crowded scenes with a mixture of dynamic textures [5]. Theyconsidered both spatial saliency and temporal saliency to detect and localize unusual events in a crowded scene. Finally, there have been attempts at conducting acrowd behavior analysis by extracting local spatio-temporal cuboids from an optical flow or the gradient pattern features. Kratz et al. analyzed a volume of extremely crowded video sequences by constructing a motion-pattern distribution that captured thelocal spatio-temporal motion patterns [29]. They encoded the motion patterns into a distribution-based Hidden Markov Model (HMM) [30]. Wang et al. measured the change in intensity frequencies over time in a spatio-temporal cuboid using a wavelet transform [6]. They proved that an abnormalregion shows a high-frequency within a certain period of time. Moreover, the optimal spatio-temporal cuboid selectionhas also been studied [7]. Since a spatio- temporal cuboid is extracted from a small part of the frame, the size and location ofthe cuboid are important factors affecting the quality of the features. The quality of a spatio-temporal cuboid was improved by choosing local maximum points in a Gaussian distribution.

PROPOSED METHOD

In this section, we describe a method for representing motion characteristics for thedetection and localization of unusualactivities within a crowded scene. Here, we should note that, we considered two types ofunusual activities: local and global. Localunusual activities occur within a relatively small area. Different motion patterns may appear in a portion of the frame, such as the unique appearance of non-human objects or the fast movement of a person when most ofthe other pedestrians are walking slowly. Global unusual activities occur across theframe, for example, when every pedestrian within a scene starts to run suddenly to escape from the scene.

SYSTEM ARCHITECTURE



An overview of the proposed method

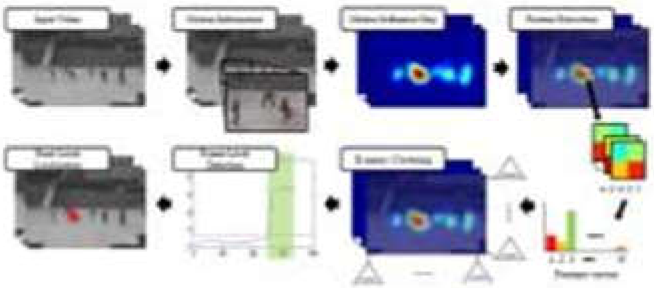


Fig. 2 illustrates the overall framework ofthe proposed method.

Given a sequence of frames, the motion information at both the pixel-level and block-level is computed sequentially. Based on the block-level motion information, the motion influence energy is computed and a motion influence map is then constructedfrom the energies in each frame. Theproposed motion influence map represents both the spatial and temporal characteristics within a single feature matrix. For the classification, we divide the motion influence map into an uniform grid, and perform the k-means clustering for each region. The distances between the center of the clusters and each extracted spatio-temporal motion influence feature are used as the feature values for unusual activity detection at the frame-level. Once a frame isclassified as unusual, we further localize theexact position of the unusual activity at the pixellevel.

Overview of the Proposed Method

Fig. 2.An overview of the proposed methodfor unusual activity detection and localization in crowded scenes.

Motion Descriptor

In our work, we estimate the motion information indirectly from the optical flows [9], [12]. Specifically, after computing the optical flows for every pixel within a frame, we partition the frame into M by N uniform blocks without a loss of generality, where the blocks can be indexed by $\{B_1,B_2,\cdots,B_{MN}\}$, and then compute a representative optical flow for each block bytaking the average of the optical flows of thepixels within the block:

$$\mathrm{b}_i = \frac{1}{J}\sum_j \mathrm{f}_i^j$$

where b_i denotes an optical flow of the i-th block, J is the number of pixels in a block, and f^j denotes an optical flow of the j-th pixel in the i-th block. We define two operators, ∠a and ||a||, which compute the orientation and magnitude of optical flows ∠a, respectively. Regarding the orientation of the optical flow of the i-th block, forcomputational efficiency, we perform hard assignment using the following rule:

$$q(\angle \mathrm{b}_i) \equiv k \quad \text{s.t. } (2k-3)\times\frac{\pi}{8} < \angle \mathrm{b}_i \leq (2k-1)\times\frac{\pi}{8}$$

Where k ∈ {1, 2, 3, 4, 5, 6, 7, 8}. Here, we should note that we consider a block in a frame to be a virtual object, irrespective of the reality, and use two interchangeably. That is, instead of detecting and trackingreal objects such as a pedestrian or cart, which is infeasible for a video clip of a crowded scene, we estimate the motion characteristics of the blocks and utilize themas motion descriptors for unusual activitydetection.

Motion Influence Map

Note that the movement direction of apedestrian within a crowd can be influencedby various factors such as obstacles alongthe path, nearby pedestrians, and movingcarts. This interaction characteristic, whichwe call the "motion influence," has beensuccessfully used in previous crowd motionanalysis studies [22], [23], [24], [28]. In thispaper, we also exploit the interactioncharacteristic for unusual activity detection. We assume that the blocks under influenceto which a moving object can affect are determined by two factors: the motiondirection and motion speed. The faster anobject moves, the more neighboring blocksthat are under the influence of the object.Neighboring blocks have a higher influence than distant blocks. With regard to theimpact of moving object i to the block j, wefirst define two indicator variables, δ_{ij}^d and δ_{ij}^ϕ, which denote whether block j is underthe influence of object i by considering thedistance between them and by taking intoaccount the visibility of block j to object i, respectively, as follows:

$$\delta_{ij}^d = \begin{cases} 1 & D(i,j) < T_d \\ 0 & \text{otherwise} \end{cases}$$

$$\delta_{ij}^\phi = \begin{cases} 1 & -\frac{F_i}{2} < \phi_{ij} < \frac{F_i}{2} \\ 0 & \text{otherwise} \end{cases}$$

where D(i,j) is the Euclidean distance between object i and block j, T_d is a threshold, φ_{ij} denotes the angle between a vector from object i to object j and the motion direction of object i, and F_i is the field of view[1]of object i. Fig. 3 describes these variables graphically. We then define the influence weight w_{ij} of object i to block jas follows:

$$w_{ij} = \delta_{ij}^d \delta_{ij}^\phi \exp\left(-\frac{D(i,j)}{\|\mathrm{b}_i\|}\right).$$

After computing the influence weights of allblocks, $\{w_{ij}\}_{i,j\in\{1,2,\cdots,MN\}}$, we finallyconstruct a motion influence map representing the motion patterns occurring within a frame. Every block in the motion influence map consists of an 8dimensional vector. Each component of the motioninfluence vector represents the quantized motion vector orientation of block i. Note that in our computation of the influence weight, we consider only a pair of blocks.

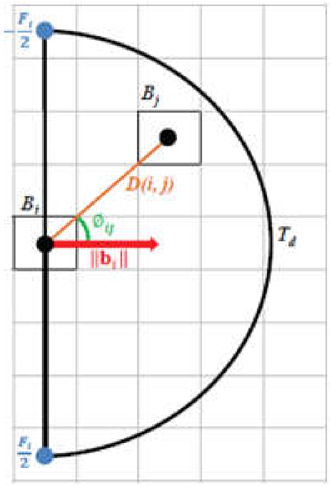


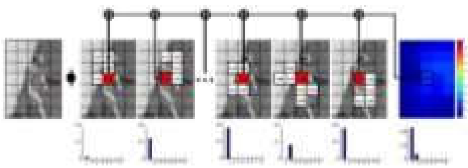
Fig. 3. The schematic description of thevariables used to compute an influence weight.

That is, w_{ij} reflects only the influence of block i on block j. Therefore, to compute themotion influence vector of block j, i.e., $H^j(k)$, within a frame, we need to considerall other blocks that potentially affect themotion of block j as follows:

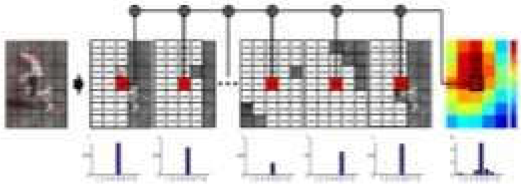
$$H^j(k_i) = \sum_{i \neq j} w_{ij}$$

where j ∈ {1,2,··· ,MN}, k_i denotes the quantized orientation index of block i,which is used as a component index of block j. In Fig. 4, we present a graphical explanation to build a motion influence mapand compare the motion influence maps for three different scenarios. In Figs. 4b-4d, we denote the target block, for which we compute a motion influence value, in red, and the numbers in blocks denote theinfluence weights for the target block. The histograms below the maps depict the motion influence value of the target block component. The bin index, k, is the orientation of the motion vector of block i. To show the value of the motion influence map through simple graphics, the right-mostcolored matrices (in Figs 4b-4d) illustratethe scalar-value representation of a motion influence vector, which is the aggregate of eight component values. Because there are more than five blocks affecting the target block, the ellipses in the middle of Figs. 4b- 4d denote the implied motion influence values affecting the target block. Note that owing to the high movement speed of the subject in Fig. 4c, a larger number of blocks are considered when computing the influence weight than in the other cases shown in Figs. 4b and 4d. It should also be noted that the proposed motion influence map considers the motion speed, direction, object size, and interactions of nearby objects, simultaneously. Concretely, for the case of fast movement among slowly moving subjects and/or objects, owing tothe large magnitude of motion flows for a subject, a larger number of nearby blocks are affected when computing the influence weights, which further results in high valuesfor a motion influence map (Fig. 4c).

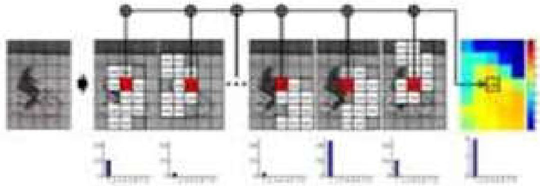
(a) A global view of constructing a motion influence map: (first) optical flow in a pixel-level, (second) motion vector in a block- level, (third and fourth) computing a motion influence weight and the corresponding amotion influence vector for red-coloredblock.



(b) A walking person



(c) A running person



(d) A person riding a bicycle

Fig. 4. An illustration of constructing a motion influence map and exemplar maps for different scenarios. Regarding a rigid object, e.g., a cart or bicycle, the motionflows of the object are relatively stable, biased, and consistent over time when compared to those of a human subject, for which there are large motion variations withcomplicated motion directions from non- rigid body parts, e.g., arms and legs. For thisreason, rigid objects tend to have consistent motion patterns over time in terms of the direction and magnitude of the motion, thereby resulting in high influence weights and thus high and biased vectors in the respective motion influence map. In the meantime, since a motion influence map is constructed by summing the influence weights related to the target block, it can represent reciprocal interactions among objects. For example, if two bicyclists are coming toward each other, the two opposite directions appear in a block, and the sum of motion influence weights for this case will be much higher than for the case of a cyclist approaching a walking pedestrian. Utilizing these characteristics, we can predict the occurrence of unusual activities in the current frame. Moreover, we can also pinpoint the location of an unusual activity. That is, the proposed motion influence map can be utilized to detect the occurrence of an unusual activity and find its location. Furthermore, unlike previous methods that focus mostly on either local or global activity detection, it is possible for our method to detect both local and global activities using a unified framework based on the proposed motion influence map Herein, we provide a pseudo algorithm for the construction of a motion influence map:

INPUT: $B \leftarrow$ motion vector set, $S \leftarrow$ block size, $K \leftarrow$ a set of blocks in a frame
OUTPUT: $H \leftarrow$ motion influence map
$H^j (j \in K)$ is set to zero at the beginning of each frame
for all $i \in K$ **do**
  $T_d = \|\mathbf{b}_i\| \times S$;
  $\frac{F_d}{2} = \angle \mathbf{b}_i + \frac{\pi}{2}$;
  $-\frac{F_d}{2} = \angle \mathbf{b}_i - \frac{\pi}{2}$;
  for all $j \in K$ **do**
    if $i \neq j$ **then**
      Calculate the Euclidean distance $D(i, j)$ between $\mathbf{b}_i$ and $\mathbf{b}_j$
      if $D(i, j) < T_d$ **then**
        Calculate the angle ϕ_{ij} between $\mathbf{b}_i$ and $\mathbf{b}_j$
        if $-\frac{F_d}{2} < \phi_{ij} < \frac{F_d}{2}$ **then**
          $H^j(\angle \mathbf{b}_i) = H^j(\angle \mathbf{b}_i) + \exp\left(-\frac{D(i,j)}{\|\mathbf{b}_i\|}\right)$
        end if
      end if
    end if
  end for
end for

Feature Extraction, Detection, andLocalization

In the proposed motion influence map, a block in which an unusual activity occurs, along with its neighboring blocks, have unique motion influence vectors. Furthermore, since an activity is captured by multiple consecutive frames, in this work, we extract a feature vector from a cuboid defined by n×n blocks over the most-recentt number of frames. Specifically, we partition the frames into non-overlapping "mega" blocks, each of which is a combination of multiple motion influenceblocks. We then extract spatio-temporal features for each mega block by adding all motion vectors in the mega blocks at each frame, and finally concatenate the motion influence vectors of the recent t number of frames. As a result, we extract an 8×t dimensional concatenated feature vector from a mega block within the frame (Fig. 5).

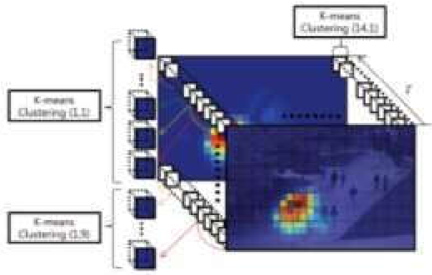


Fig. 6. Illustration of k-means clustering with frame division.
(1,1), (1,9),7 and (14,1)denote, respectively, the coordinate of the respective mega blocks.

For each mega block, we then perform K-means clustering using the spatio-temporal features, and set the centers as codewords.That is, for the (i,j)-th mega block, we have K codewords, $\{\mathbf{w}_k^{(i,j)}\}_{k=1}^{K}$. Here, we should note that in our training stage, we use only video clips of normal activities. Therefore, the codewords of a mega block model thepatterns of usual activities that can occur inthe respective area. In the testing state, after extracting the spatio-temporal feature vectors for all mega blocks, we construct a minimum distance matrix E over the mega blocks, in which the value of an element is defined by the minimum Euclidean distancebetween a feature vector of the current testframe and the codewords in the corresponding mega block as follows:

$$\mathcal{E}(i,j) = \min_{k} \ \|\mathbf{f}^{(i,j)} - \mathbf{w}_k^{(i,j)}\|^2$$

where E(i,j) denotes the (i,j)-th element in E, and $f^{(i,j)}$ is the feature vector of the (i,j)-th mega block in the test frame. In a minimum- distance matrix, the smaller the value of an element, the less likely an unusual activity isto occur in the respective block. On the other hand, we can say that there are unusual activities in t consecutive frames ifa higher value exists in the minimum- distance matrix. Therefore, we find the highest value in the minimum-distance matrix as the frame representative feature value. If the highest value of the minimum distance matrix is larger than the threshold, we classify the current frame as "unusual". The localization is also performed using the same strategy with the same threshold for each mega block to localize the unusual activity or activities.

EXPERIMENTAL RESULTS

The application is implemented in python using OpenCV library in Ubuntu windows/Linux environment. The architecture of the application is madeflexible in order to load different types of video clips. In Figure 5(a), a person is running with dog that result is shown in Figure 5(b) as a abnormal activate.

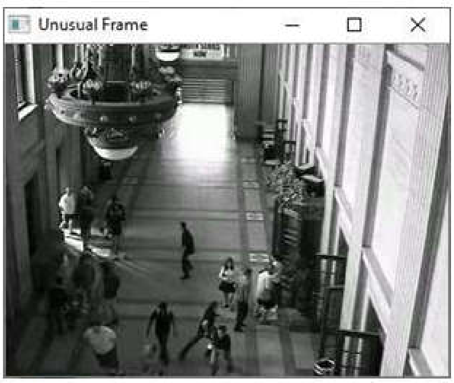


Figure 5(a)

Figure 5(b).

CONCLUSION

With the growing number of surveillance cameras installed in private and public areas/ home, there has been a demand for the automatic and intelligent analysis of video sequences using computers. Unusual event or activity detection in a crowded/home scene has recently been of great interest in the area of vision based surveillance. in this project, we proposed a method for representing the motion characteristics within a frame to detect and localize unusual human activities in a crowded scene/home. Owing to the Representational power of the proposed motion influence map for both space and time, we can classify a frame as usual or unusual, and localize the areas of unusual activities within a frame. For a real application, a smart surveillance system needs to efficiently detect both local and global unusual activities within a unified framework.

FUTURE WORK

The proposed method has a limitation when there is a strong perspective distortion in the input video as the motion influence map is built based on the motion direction and magnitude of the moving objects. However, the main focus of this work is to detect unusual activities within a crowded scene, for which the cameras usually cover a wide area, resulting in small objects being presentin the scene without significant perspective changes. Also our experiments were limited to a fixed viewpoint, and there is a limitation in the applicability of the approach for surveillance cameras with pan, zoom, or tilt functionality. At this moment the proposed method deals only with static cameras. However, it can be easily extended to PTZ cameras using localization results.

REFERENCES

1. M. Mancas, N. Riche, J. Leroy, and B. Gosselin, “Abnormal Motion Selection in Crowds Using Bottom-Up Saliency,” Proc. 18th IEEE International Conference on Image Processing, Bruxelles, Belgium, Sep. 11-14, 2011,pp. 229-232.
2. D. Y. Chen and P. C. Huang, “Motion- Based Unusual Event Detection in Human Crowds,” Journal of Vision and Communication and Image Representation, Vol. 22, No. 2, 2011, pp. 178-186.
3. S. Ali and M. Shah, “Floor Fields for Tracking in High Density Crowd Scenes,” Proc. 10th EuropeanConference on Computer Vision,Marseille, France, Oct. 12-18, 2008, pp. 1-14.
4. S. Wu, B. Moore, and M. Shah, “Chaotic Invariants of Lagrangian Particle Trajectories for Anomaly Detection in Crowded Scenes,” Proc. IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, USA, June 13-18, 2010, pp. 2054-2060.
5. R. Mehran, A. Oyama, and M. Shah, “Abnormal Crowd Behavior Detection Using Social Force Model,” Proc. IEEE Conference on Computer Vision and Pattern Recognition, Miami, USA, June

20-25, 2009, pp. 935942.

6. Adam, E. Rivlin, I. Shimshoni, and D. Reinitz, “Robust Real-Time UnusualEvent Detection Using Multiple Fixed- Location Monitors,” IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 30, No. 3, 2008, pp. 555-560.
7. T. Xiang and S. Gong, “Video Behavior Profiling for Anomaly Detection,” IEEETrans. on Pattern Analysis and Machine Intelligence, Vol. 30, No. 5, 2008, pp. 893-908.
8. F. Jiang, J. Yuan, S. A. Tsaftaris, and A.K. Katsaggelos, “Anomalous Video Event Detection using Spatiotemporal Context,” Computer Vision and Image Understanding, Vol. 115, No. 3, 2011, pp. 323-333.
9. D. Lucas and T. Kanade, “An Iterative Image Registration Technique with anApplication to Stereo Vision,” Proc. 7th International Joint Conference on Artificial Intelligence, San Francisco, USA, Aug. 3-9, 1981, pp. 674-679.
10. G. Xiong, J. Cheng, X. Wu, Y. Chen, Y.Ou, and Y. Xu, “An Energy Model Approach to People Counting forAbnormal Crowd Behavior Detection,” Neurocomputing, Vol. 83, 2012, pp. 121-135.
11. Zhan, D. N. Monekosso, P. Remagnino, S. A. Velastin, and L. Q. Xu, “Crowd analysis: a survey,” International Journalof Machine Vision and Applications, Vol. 19, No. 5-6, 2008, pp. 345-357.
12. Helbing and P. Molnar, “Social ForceModel for Pedestrian Dynamics,” Physical Review E, Vol. 51, No. 5, 1995, pp. 4282-4286.
13. Lerner, Y. Chrysanthou, and D. Lischinski, “Crowds by Example,” Computer Graphics Forum, Vol. 26, No.3, 2007, pp. 655-664.
14. S. Ali and M. Shah, “A Lagrangian Particle Dynamics Approach for Crowd Flow Segmentation and Stability Analysis,” Proc. IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, USA, June 17-22, 2007, pp. 1-6.
15. X. Cui, Q. Liu, M. Gao, and D. N. Metaxas, “Abnormal Detection Using Interaction Energy Potentials,” Proc.IEEE Conference on Computer Vision and Pattern Recognition, Colorado,USA, June 20-25, 2011, pp. 3161-3167.
16. Laptev, M. Marszalek, C. Schmid, and Rozenfeld, “Learning Realistic Human Actions from Movies,” Proc. IEEE Conference on Computer Vision and Pattern Recognition, Anchorage,USA, June 2328, 2008, pp. 1-8. S. Pellegrini, A. Ess, K. Schindler, and
17. L. V. Gool, “You’ll Never Walk Alone: Modeling Social Behavior for Multi- Target Tracking,” Proc. 12th IEEE International Conference on Computer Vision, Kyoto, Japan, Sep. 29 - Oct. 2, 2009, pp. 261-268.
18. Y. Cong, J. Yuan, and J. Liu, “Abnormal Event Detection in Crowded Scenes Using Sparse Representation,” Pattern Recognition, Vol. 46, No. 7, 2013, pp. 1851-1864 .
19. W. Li, V. Mahadevan, N. Vasconcelos, ”Anomaly Detection and Localization inCrowded Scenes,” IEEE Transactions on Pattern Analysis and MachineIntelligence, Vol. 36, No. 1, 2014, pp. 18-32.
20. Kim and K. Grauman. “Observe Locally, Infer Globally: A SpaceTime MRF for Detecting Abnormal Activities with Incremental updates,” Proc. IEEE Conference on Computer Vision and Pattern Recognition, Miami, USA, June 20-25, 2009, pp. 2921-2928.