Received October 12, 2018, accepted October 21, 2018. Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2018.2878733

Physics Inspired Methods for Crowd Video Surveillance and Analysis: A Survey

XUGUANG ZHANG[1], QINAN YU[1], AND HUI YU[2], (Senior Member, IEEE)
[1]School of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, China
[2]School of Creative Technologies, University of Portsmouth, Portsmouth PO1 2DJ, U.K.

Corresponding author: Xuguang Zhang (zhangxg@hdu.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61771418 and Grant 61271409.

ABSTRACT Crowd analysis is very important for human behavior analysis, safety science, computational simulation, and computer vision applications. One of the most popular applications is video surveillance that plays an important role in crowd behavior analysis including real-time crowd behavior detection and information retrieval. In the field of video surveillance, many kinds of methods have been proposed for analyzing crowds, such as machine learning, signal processing, and physical model-based methods. As a kind of collective movements, crowd behavior contains many physical attributes, such as velocity, direction of motion, interaction force, and energy. Therefore, a lot of methods and models derived from physical ideas have been applied in many frameworks of crowd behavior analysis. This survey reviews the development of physical methods of crowd analysis in detail. The physics-inspired methods in crowd video analysis are summarized into three categories including fluid dynamics, interaction force, and complex crowd motion systems. Furthermore, the existing public databases for crowd video analysis are collated in this paper. Finally, the future research directions of the open issues of crowd video surveillance are also discussed.

INDEX TERMS Crowd behavior analysis, video surveillance, physics model, crowd abnormal behavior detection, crowd motion segmentation.

I. INTRODUCTION

Crowd behavior analysis has become a hot research topic in many disciplines, such as statistical physics [1], [2], computer science [3]–[5] and psychology and behavior [6], [7]. To analyze crowd behavior effectively, research methods of different disciplines have often been integrated together. Crowd motion is essentially a kind of collective motion. Physical modeling is very useful for resolving the issue around crowd analysis, since many physics based methods have been used to understand collective motion successfully. Collective motion has been introduced in [8]. The effective approach to demonstrate collective motion is statistical physics. In that article, Vicsek also points out several physical concepts which can be used to describe collective motion, such as velocity, correlation function and fluid dynamics. These types of collective motion are summarized as non-living systems, bacteria colonies, macro-molecules, amoeba, cells, insects, fish, birds, mammals and crowd.

Many review and survey papers have been published to introduce the research methods of crowd surveillance and behavior analysis. We collect the review papers published during the periods of 2004-2017. First, we collect some review papers based on [9]. Then, we collect some other related review papers from ACM, IEEE, Springer and Elsevier databases including both journal and conference papers. It is worth mentioning that we pay more attention to the review about crowd motion analysis. Therefore, some survey papers about pedestrian detection, tracking, and action analysis are not be included here, although in a broad sense they can also be counted as part of crowd analysis. We have calculated the number of these review and survey papers published from 2004 to 2017 [9]–[34], as can be seen in Fig. 1. We can see the number of review papers increases significantly in recently years (2012-2017). This indicates that the crowd surveillance behavior analysis has attracted increasing interests from researchers. In these summary articles, some articles focus on application topics, such as crowd counting and density estimation [22], [27], [29], [30], abnormal crowd behavior detection [14], [16], [32], human behavior recognition [13], [15], [17], [21], [23], and crowd dynamics simulation and understanding [10] [34]. Some articles give a comprehensive review of methods of crowd monitoring

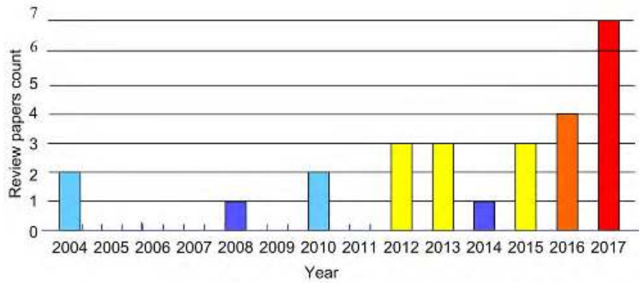

FIGURE 1. **The graph shows the number of review and survey papers for crowd video surveillance and analysis from 2004-2017.**

and behavioral analysis, focusing on the background of crowd safety management [9], [11], [12]. Research methods of crowd video surveillance and analysis can be divided into two categories, one being pattern recognition and machine learning, and the other being physics inspired method. Several review papers introduce the physics-based models and methods used in crowd behavior analysis. In [11], they point out that there are three kinds of physics inspired models (microscopic, mesoscopic and macroscopic) that have been used to describe a crowd. Thida *et al.* [18] divide the crowd video analysis model into two categories. i.e., the macroscopic and microscopic model. In [19], they introduce the physics based methods to deal with crowd and group behavior analysis, from the crowd video analysis and crowd simulation. In [26], they give a physics and biologically inspired perspective to help us understand crowd behavior. Although the existing reviews have covered most of the research fields and methods of crowd behavior analysis and a small number of papers have also involved the physics inspired in crowd behavior analysis, there is still a lack of an in-depth and complete review of crowd video surveillance and behavior analysis based on physical inspired models and methods, however. Here, we introduce the physical models and methods used for crowd video surveillance and analysis, following the framework below (Fig. 2). Firstly, three kinds of research strategies are presented. Then, the importance and significance of video surveillance among the three strategies are discussed. Secondly, the research methods of crowd video surveillance are summarized into two kinds of systems, based on physical model and machine learning. Finally, some physical models and methods are described, such as fluid dynamics, energy, entropy, force model. In fact, many approaches for crowd behavior analysis are overlapped with each other. The purpose of classification is only to facilitate understanding these methods.

The main research about crowd behavior analysis can be divided into three methods: (1) Controlled experiment: The participants are asked to perform designated movements according to the predetermined behavior patterns in preset experimental scenarios and environments to reveal the regularity and behavior characteristics of crowd movements. In [35] the experiments are conducted in an eleven story building, the participants who took part in the test were

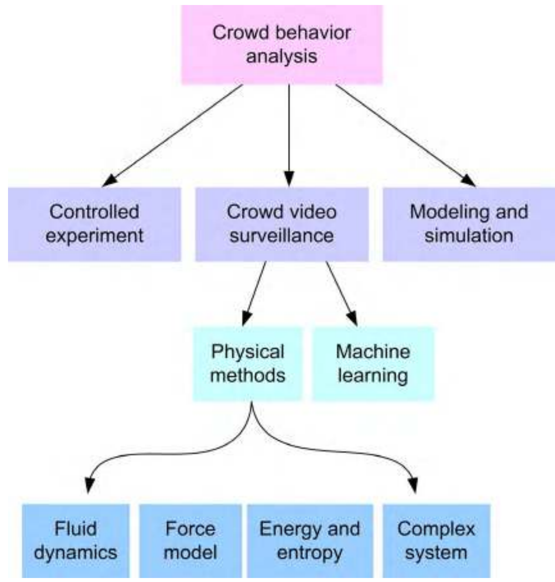

FIGURE 2. **The arrangement of idea in this paper.**

comprised of different small groups. The experiment revealed the effects of groups of different social relationships on crowd movements. In [36], several experiments were conducted to show how the merging architectural features influence crowd motion; (2) Crowd modeling and simulation: The controlled experiment may be impractical in certain scenarios or in a high density crowd. The movement and characteristics of the crowd should be simulated by computer software [37]. Karamouzas and Overmars [38] simulated the pedestrian walking behavior of small groups by considering the interaction between a pedestrian with its group members, other groups and individuals. In [39], a mutual information of these interacting agents was proposed to determine the level of order within a crowd, this was integrated into the social force model to simulate the crowd behavior during crowd evacuation. (3) Crowd video surveillance: The crowd movement is captured by a camera, and the crowd behavior can be analyzed from the image sequence online or offline. In [40], the crowd motion was segmented based on a local-translation domain segmentation model, by treating the crowd as a scattered motion field. In [41], a method based on sparse representation was proposed for detecting abnormal events in a crowd scene. In their research, a sparse reconstruction cost is used to measure crowd behavior. The example of different kinds of research strategies can be seen from Figure 3. Crowd video surveillance plays a very important role for crowd understanding and analysis. Analyzing crowd behavior from video sequences can detect abnormal behavior automatically and provide an effective video retrieval function. The research field about crowd analysis in video surveillance includes pedestrian detection and re-identification [42]–[47], target tracking [48]–[51], crowd counting and density estimation [52]–[55], crowd motion segmentation [56]–[58], abnormal behavior detection [59]–[62] and crowd behavior classification and understanding [63], [64].

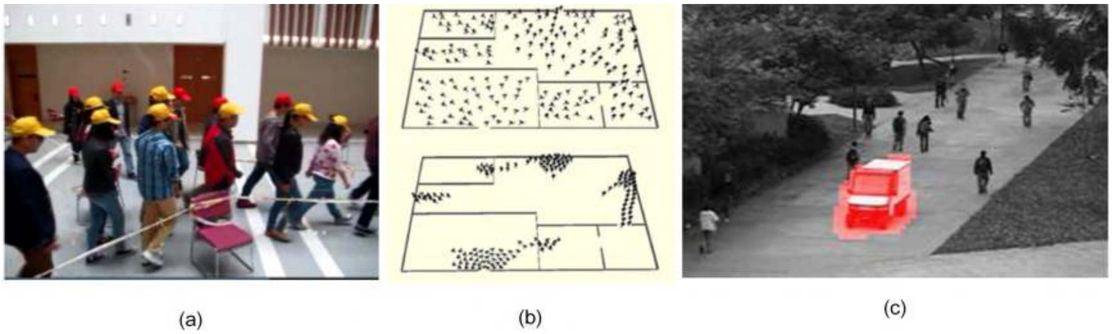

FIGURE 3. Three kinds of research fields of crowd analysis. (a) shows a snapshot of the merging process in a merging corridor from the controlled experiment [36]; (b) the crowd motion simulation in complex environments with many rooms and exits in [39]; (c) a crowd abnormal behavior detection result of [41].

The methods used for crowd video surveillance can be summarized into two categories: physical model or machine learning based methods. (a) Physical modeling based method: paying attention to the physical movements and principles contained in the crowd behavior, such as describing pedestrian movement of individual at the micro level using a social force model [65], detecting anomalies in the behavior on the macro scale using the energy and entropy of the population [66], and integrating the macro and micro information into a space-time cube based structure to build the local and global social network model for crowd behavior description [67]. (b) Machine learning based methods: focusing on the crowd motion data processing, analysis and learning, such as using the signal processing method of sparse expression for crowd attributes recognition [68] and abnormal behavior detection [41], and extracting the texture features in the video to identify abnormal crowd behavior [69].

This survey focuses on the crowd video analysis using physics-based methods, by giving a comprehensive review about the physics-based approach. The fluid dynamics, energy and entropy, force model and complex system science that are used in crowd analysis are summarized in this paper. Using these methods, the application of crowd motion segmentation, crowd counting and density estimation, and crowd abnormal behavior detection can be deal with effectively.

The remainder of this paper is organized as follows. Section II gives a physical viewpoint for crowd video surveillance. Section III introduces the flow field analysis based method in crowd surveillance field. In Section IV, the force model used for crowd analysis is summarized. Section V describes the crowd motion as a system, where the methods using energy, entropy and complex system analysis are collected. Section VI introduces the public crowd analysis datasets. Conclusions and future developments are made in Section VII.

II. CROWD VIDEO SURVEILLANCE FROM PHYSICAL VIEWPOINT

Many approaches have been proposed for crowd video surveillance. The ideas of these approaches to analyze crowd behavior are usually from individual to crowd. These methods can be divided into the following categories, such as, gestures and action of pedestrians [70], pedestrian trajectory analysis, reactions between pedestrians, small group activities, the state and behavior of the whole crowd. The purpose of this paper is to help readers understand the crowd behavior from the perspective of physics. A crowd can be regarded as a complex system. In physics, dealing with complex systems usually has three levels of detail: microscopic, mesoscopic and macroscopic. Similarly, in the field of crowd video analysis, there are three types of perspectives, i.e., microscopic, mesoscopic and macroscopic.

Macro based approach: a crowd is treated as a whole in this kind of method. The behavior of a crowd is analyzed from its overall external performance. Many contributions have been proposed for crowd surveillance. Ernesto et al. used optical flow and unsupervised feature extraction method to recognize the emergency event in a large-scale crowd [71], A Bernoulli statistical shape model was employed to count the number of pedestrians in a crowd [72], Kratz and Nishino [73] used a spatio-temporal motion model for large-scale crowd abnormal behavior detection, Xiong *et al.* [66] detected the abnormal crowd behavior according to the energy based model. Analyzing crowds from the macro point of view is suitable for processing a large-scale crowd or pedestrians that have the same movement pattern (Fig. 4. a). However, the individuals' position and movement features are neglected in this kind of method, thus, macro-based approaches are not suitable for dealing with a small-scale and loose movement crowd with individual characteristic.

Micro based approach: this kind of methods focuses on the behavior of each individual. The trajectories or gestures of the individuals can be used to recognize the crowd behavior. There are many methods based on a micro point of view, such as locating the position of pedestrians by detecting the head and shoulder of individuals [74], using Bayesian clustering to detect moving individuals, and analyzing their behavior [75], detecting the fight scenes in a video by AdaBoost classifier [76], and using shape and motion template to detect the individuals' movements in a crowd [77]. These methods are suitable for processing a small-scale crowd (Figure 4 (c)). It is very hard to recognize the gestures and trajectories of

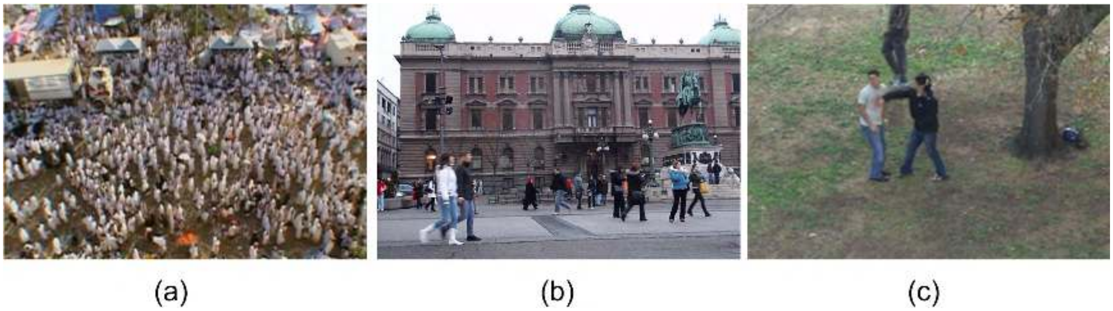

FIGURE 4. Different types of crowd (a) large-scale crowd (b) middle-scale and loose crowd (c) small-scale crowd

an individual when there are many pedestrians, and occlusion occurring among them.

The large-scale and small-scale crowd can be analyzed from a macro and micro point of view respectively. However, in real life, many crowd scenes are constructed by a mid-scale crowd. In this kind of crowd, the pedestrians do not have the same motion pattern. We cannot calculate its statistical information from a macro scale observation. In these scenes, pedestrians usually appear in small groups (Fig. 4. (b)), with relationships between these small groups still possible. Some contributions focus on the behavior of small groups in a crowd or deal with crowds from both macro and micro characteristics; arguably, we can call it a mesoscopic point of view. In [78], the individuals are detected and tracked in a crowd, and their common fate and features are considered using Gestalt psychology, which determines whether several individuals belong to the same group. Using this strategy, the small groups can be identified using agglomerative clustering methods. Chaker *et al.* [67] combined both the macro and micro information of a crowd. They represented each windows of a crowd video as a set of spatio-temporal cuboids, and a local and global social network model was constructed for detecting and localizing the crowd's abnormal behaviors. Zhang *et al.* [79] tracked the individuals by a covariance tracking method. A complex network model was constructed using the relationship between individuals. And then, five crowd behaviors (gather, meet, together, separation and dispersion) are classified by the k-nearest neighbor method, based on the characteristic parameters of the crowd complex network.

III. FLOW FIELD ANALYSIS FOR CROWD SURVEILLANCE

Many researchers consider crowd motion as fluid flow. To analyze the behavior of a crowd, the first important task is to represent a crowd as a flow field. Optical flow estimation can be used to change a crowd video to a vector field. After getting the crowd motion vector field, the methods of flow field analysis can be used for crowd analysis.

A. CROWD FLOW FIELD REPRESENTATION

The popular theories for describing fluid motion depend on the Lagrange and Euler systems. The former focuses on the movement of each particle of a fluid, and the latter (Euler's law) focuses on the whole flow field, i.e., describing the physical quantity of fluid flow characteristics as vector field and scalar field. Based on the Lagrange and Euler systems, path-lines, stream-lines and streak-lines (Figure 5) can be used for particle flow moving curve representations. Path-lines indicate the trajectory that a particular particle passes during the flow process. The path-line corresponds to the Euler system. Stream-line refers to, at any instant, the velocity vector at any point on the line being tangent to it. The stream-line also corresponds to the Euler system. Streak-lines refer to the curve of a particle of fluid passing through a fixed point, and the positions of the particles at that present moment.

The streak-line has been calculated based on a Lagrangian system to represent a crowd as fluid flow in [80]. The performance for crowd motion representation of streak-lines has been compared with that of path-lines and stream-lines. By combining this with potential functions, streak-line based methods show good performance in crowd segmentation and abnormal behavior detection. Wang et al. improved the traditional optical flow method using a high accurate variational model. The streak-lines and streak flow can be calculated by the improved optical flow method. Furthermore, the corresponding formulation was modified to represent the similarity of the streak flow to achieve a high accuracy crowd flow segmentation [81]. In [82], they computed the streak-lines of the crowd image sequence to classify the abnormal activity for crowd surveillance. Beyond the computing of streak-lines, Zhang et al. used the similarity of streak-lines to segment high density crowd behaviors [83]. As for flow field representation, texture based methods can reveal more detail of a flow field than line based methods [84]. By using a line integral convolution technique after calculating the stream-lines of a crowd flow field, Zhang et al. represented crowd motion and background regions as different texture images. A crowd counting method was described according to the relationship between the area of foreground region and the number of individuals in a crowd [85].

B. DYNAMIC ANALYZATION FOR CROWD FLOW FIELD

Another kind of flow-field based crowd analysis method focuses on analyzing the dynamic system of a flow field.

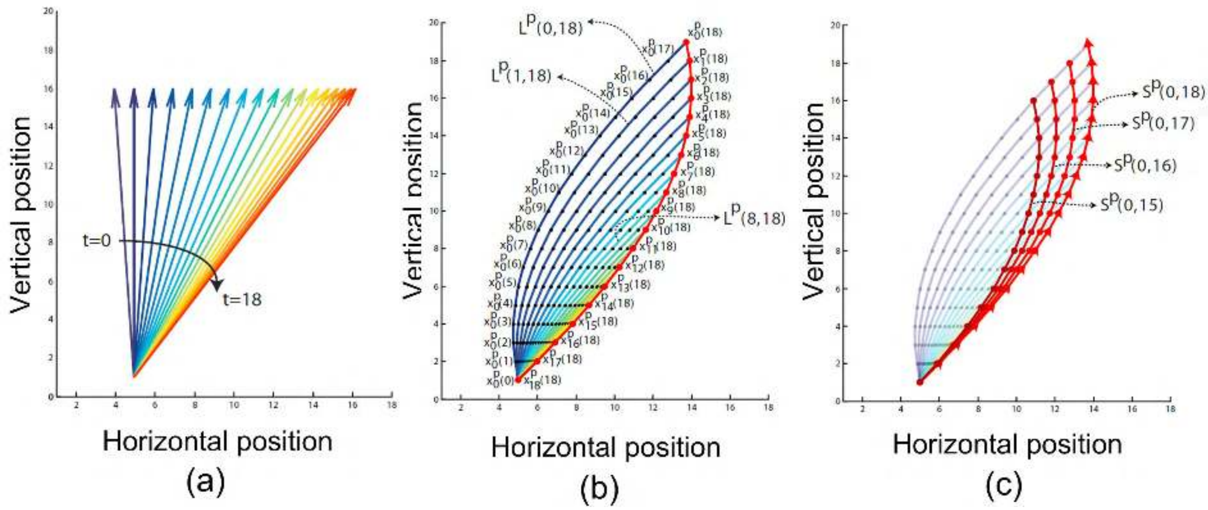

FIGURE 5. An illustration of stream-lines (a), path-lines (b) and streak-lines (c) in [80].

FIGURE 6. The detection result of 15 crowd scenes in [86].

Based on the crowd motion flow field, Ali et al. constructed a Finite Time Lyapunov Exponent field. Based on the Lagrangian particle dynamics framework, an effective crowd segmentation and stability analysis method can be achieved [87]. Solmaz *et al.* dealt with a crowd by initializing a dynamic system based on optical flow estimation. The eigenvalues of the Jacobian matrix can be used to determine the stability of the crowd flow field. Five crowd behaviors (Bottlenecks, Fountainheads, Lane Formation, Ring/Arch Formation and Blocking) can be identified in their research without the step of object detection and tracking [86]. In Figure 6, the crowd behavior identification results in 15 scenes are shown. Wu et al. extracted the curl and divergence features of the crowd motion, then, a vector pooling was constructed based on curl and divergence feature for crowd video classification and retrieval [88]. They also used the curl and divergence of motion trajectories to describe the motion structures of a crowd. The curl and divergence of motion trajectories descriptors are useful for identifying five kinds of crowd behavior (lane, clockwise arch, counterclockwise arch, bottleneck and fountainhead) [89].

Some contributions focus on the detection of the salient regions in the crowd; Lim *et al.* [90] proposed a framework to analyze the temporal variations in a crowd flow field. In this framework, they regarded the crowd motion flow field as a dynamic system, and used the stability theory of dynamic system to detect the salient regions in the crowd. Hu *et al.* [91] detected the sinks modes in a crowd, and

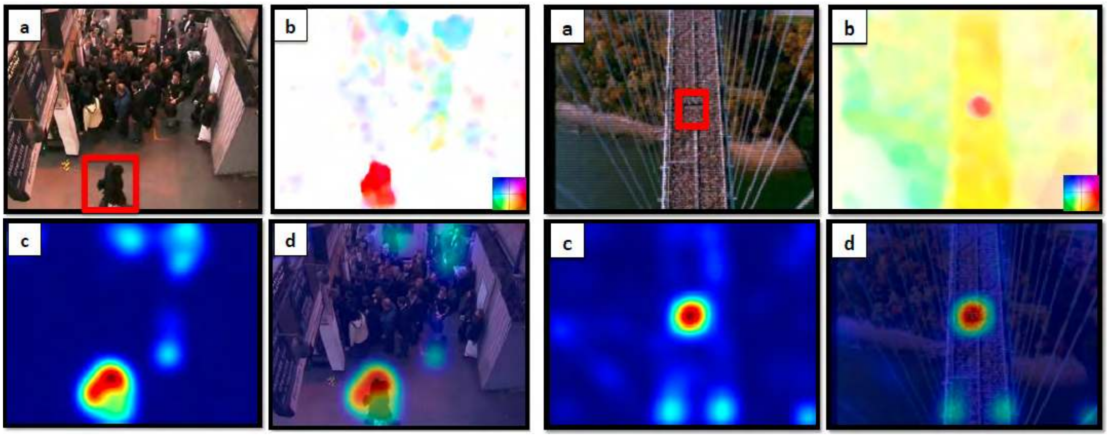

FIGURE 7. The experimental results of salient motion detection in [58]. Left is the result of counter flow detection. Right is the result of crowd instability detection. (a), (b), (c), (d) represent the input frames, optical flow field, motion saliency map and saliency map overlaid on the input frame respectively.

defines a group of flow vectors which have the same physical motion pattern in an instantaneous motion field. The spectral analysis has been used for saliency detection of crowd motion in [58]. The spectral analysis method can detect the interested regions (salient action, counter flow and unstable region) of a crowd scene and does not depend on prior knowledge and training. The results can be seen in Fig. 7, in which the counter flow and crowd instability can be detected. Khan et al. initialized a crowd scene as a dynamic system based on optical flow. Then, an unsupervised clustering algorithm was employed to extend the short particle trajectories to longer tracks. The sinks (pedestrians disappear) and sources (pedestrians appear) structures can be identified using their method [92], [93].

IV. FORCE MODEL FOR CROWD ANALYSIS

A force-based model is an effective tool to analyze crowd behavior in the microcosmic level. This kind of method has been used in crowd motion simulation and video surveillance extensively. One of the most famous models for simulating the crowd motion is social force model which has been proposed in [94]. In a social force model, the psychological and physical forces are integrated to describe a pedestrian's behavior in a crowd. Helbing *et al.* [95] also used this model to investigate the crowd motion mechanisms in panic and escape situations. In a social force model, three kinds of forces have been considered: the driving force toward to a destination, influence force from other pedestrians, and repulsive force from other objects such as a wall. Many contributions have been proposed to modify the social force model. Parisi *et al.* [96] integrated a self-stopping mechanism into the traditional social force model to avoid the pedestrians pushing each other over. By using this self-stopping mechanism, the simulation data of pedestrian flows in normal conditions was reduced significantly. Gao *et al.* modified the social force model by considering the relative velocity of pedestrians. In their model, different weights were assigned to pedestrians according to their moving speeds. When predicting possible collisions between pedestrians, they considered not only the distance factor but also the velocity factor [97]. Han and Liu [98] added an information transmission mechanism that gathered information from the neighbor of the pedestrians to calculate the driving force.

A. SOCIAL FORCE MODEL IN CROWD VIDEO SURVEILLANCE

The application of social force model has been extended to crowd video analysis fields. Ramin *et al.* used social force models to detect and localize the abnormal behaviors of a crowd [65], [99]. In their research, the represented the crowd video as space-time flow field. Each particle is treated as an individual, and then, the interaction force of particles were gained by the social force model. The performance of social force-based method is better than pure optical flow based method for abnormal crowd behavior detection. The Particle Swarm Optimization algorithm has also been applied to optimize the social force to model the normal and abnormal behavior of crowd [100]. The local density feature was estimated by Local Binary Pattern in [101]. The social force model and local density estimation were integrated to form a local pressure model. Therefore, a histogram of oriented pressure was used to describe the behavior of a crowd. A graph representation based method was proposed in [102] to reveal the interaction groups in a crowd.

The social force model and visual focus of attention model are used to form a graph to reveal the socially interacting groups. A social attributes-aware force model has been proposed in [103] to represent the social disorder and congestion attribute in a crowd scene. In a social attributes-aware force model, social force, disorder and congestion attribute, and scale estimation are integrated together. Fig. 8 shows the framework of the proposed method based on a social attributes-aware force model in [103]. Similar to social force, the attractive force and repulsive force are often used for

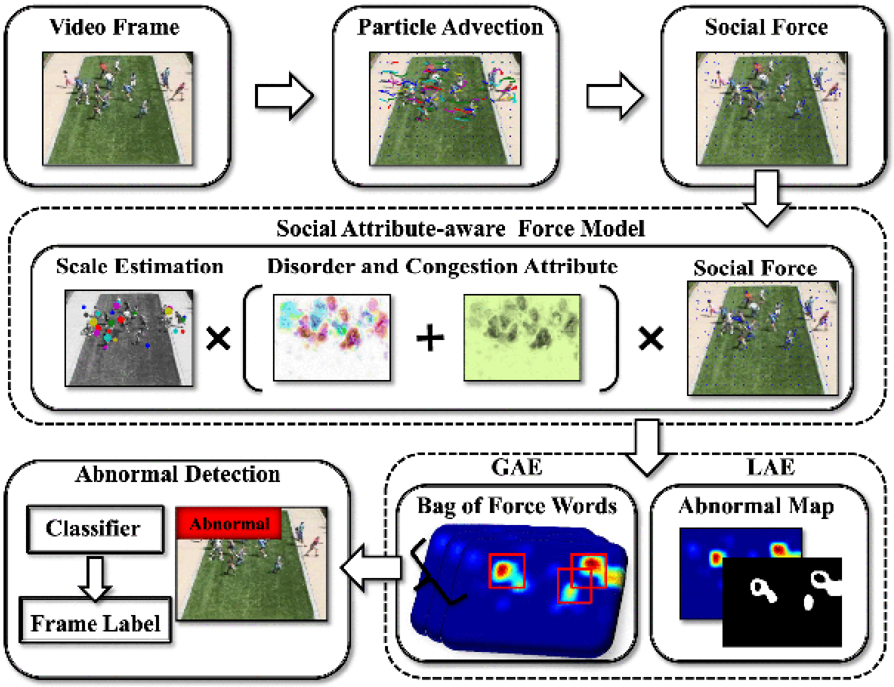

FIGURE 8. **The framework of the proposed method in [103]. The social force is first calculated after the particle advection processing. Furthermore, the social attribute-aware force model is constructed. Finally, abnormal crowd behavior can be detected based on bag of force word.**

crowd analysis [104]. Linear cyclic pursuit is used to capture the attractive and repulsive acting on the pedestrians. Their method can predict crowd motion in real-time [105].

B. FORCE MODEL IN FLUID DYNAMICS

Another kind of force model used in crowd video analysis derives from fluid dynamics. To track pedestrians in high density crowd scene, Ali *et al.* proposed a scene structure based force model. The force model was made from static floor field, dynamic floor field and boundary floor field [106]. Shear force was calculated to represent the interaction force of different particles in a crowd, and a crowd was treated as a viscous force field. The shear force will be different if the motion pattern of pedestrians is different. For example, if the pedestrians have the same motion pattern, the shear force will be small. The shear force based method can be used to analyze large-scale crowd behaviors [107]. The framework of the crowd behavior perception algorithm in [107] can be seen from Fig. 9. There are four major modules in their algorithm, i.e., spatio-temporal variation fluid field construction, spatio-temporal viscous force field construction, sodebook generation based on spatio-temporal viscous fluid features and crowd event perception based on LDA. Another force model that comes from smoothed particle hydrodynamics is used to detect crowd coherency motion. Three forces (pressure force, viscous force, and external force) are calculated to construct the density independent hydrodynamics model [108].

Similar to using a force model, an agent based model is also useful for understand crowd motion. Zhou proposed a mixture model of dynamic pedestrian-agents for learning crowd behavior [109]. This agent based model can be used for crowd behavior simulation, classification and prediction.

V. ANALYZING CROWD AS A SYSTEM

The crowd motion issue can be treated as a physical system. For a closed system, no energy and information can be exchanged with the outside. For an open system, the energy and information can be exchanged with outside [110]. Some physical ideas with regard to the crowd as a system have been used for crowd analysis, such as energy, entropy, chaos and complex networks.

A. ENERGY CALCULATION IN CROWD SYSTEM

The position and velocity of each pedestrian and the interaction between different individuals have been used to construct the energy function in many contributions. To detect the abnormal behavior in a group of pedestrians, the interaction energy potential function was calculated after the processing of interesting point detection and tracking in [111]. The interaction energy potential function can be used to represent the interaction between a pedestrian with other pedestrians around him. The energy change can be used to reflect the normal and abnormal behavior of a group of people. Based on several behavior factors (damping, speed, direction, attraction, grouping and collision) that influence the choice of a

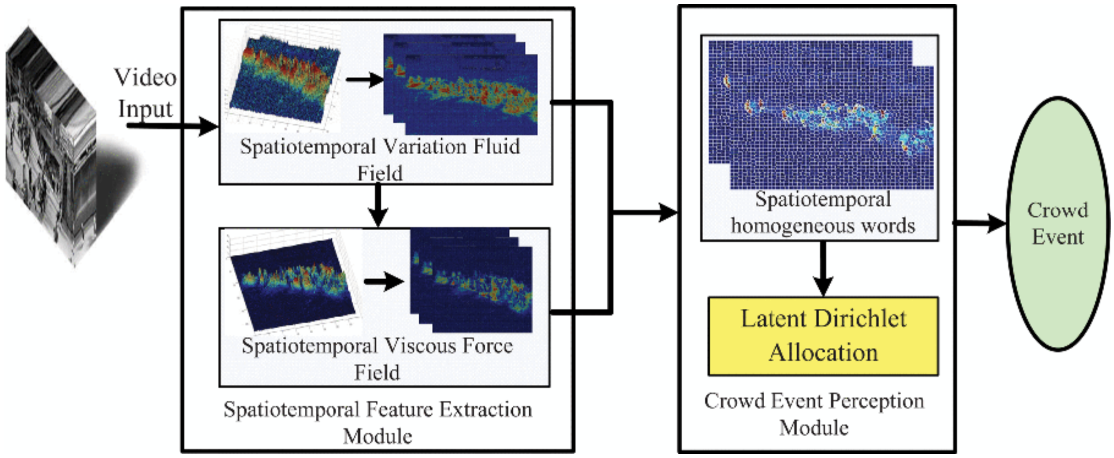

FIGURE 9. The four modules of the crowd behavior perception algorithm in [107].

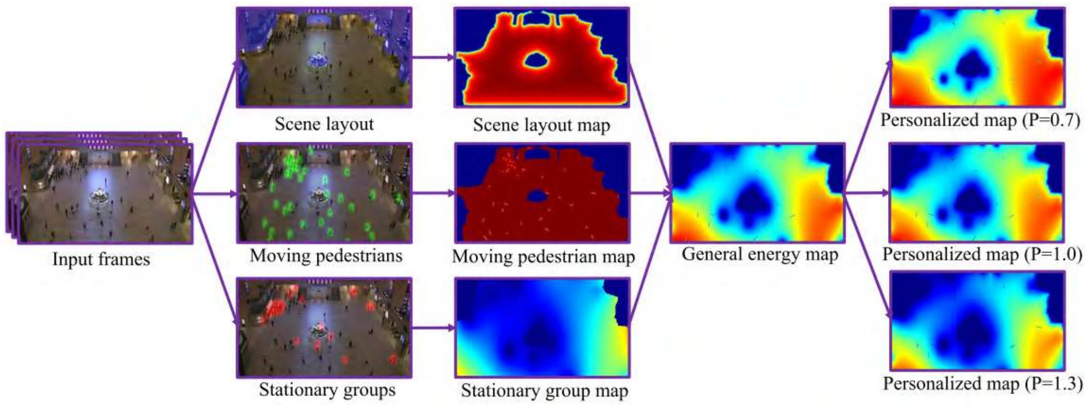

FIGURE 10. Flowchart of the energy map model proposed in [115]. Firstly, scene layout, moving pedestrians, and stationary group factors are extracted from the input frames. Furthermore, three energy maps are gained from the three corresponding influence factors respectively. And the general energy map can be calculated by integrating the three energy maps. Finally, personalized maps can be calculated based on the general energy map.

pedestrian, Yamaguchi et al. proposed an energy function for pedestrian navigation. Using the energy function, a good performance about behavior prediction and pedestrian tracking can be achieved [112]. Taking knowledge from solid-state physics, a potential energy function of a particle's inter force was originally used for online anomaly crowd behavior detection in Yuan's research. For crowd anomaly behavior detection, crowd structure representation is an important step. The potential energy function of a particle's inter force is useful for describing the relationship of different individuals [113]. The image potential energy is related to the position of the pixel in the image. In [114], the distance between the pedestrians and the camera was considered. Using the pinhole perspective projection model, the depth information of an image can be estimated. Integrating the depth information into the image potential energy model, an effective method can be achieved for crowd counting and density estimation [66].

Another kind of energy-based method describes a crowd as an energy map. Some useful information can be found from the energy map. Three influence factors (Scene Layout, Moving Pedestrians and Stationary Groups) are separated or located from the image sequences to construct an energy map model; the flowchart of the energy map model can be seen from Fig. 10. Based on the energy map model, the behavior of stationary crowds can be analyzed such as pedestrian walking path prediction, travel time estimation, and abnormal event detection [115]. Lin *et al.* [116] also proposed a heat map based method for group activity recognition. The trajectories of pedestrians were treated as a set of heat sources and the energies of different heat sources were extracted to describe the group activities.

B. THE ENTROPY IN A CROWD SYSTEM

Entropy is an effective method which can be used to measure the disorder of a system [117]. In [110], a crowd was considered as an open system, which can exchange energy with outside. Shannon information entropy was employed in their paper to detect crowd behavior from a macro scale. They pointed out that if the crowd motion was disorderly, the value

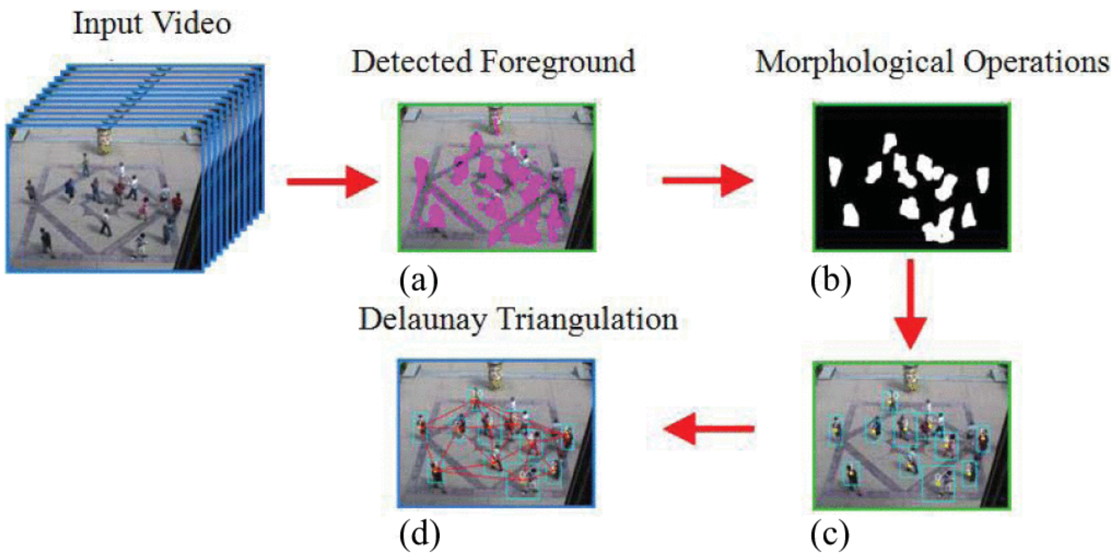

FIGURE 11. The framework of the graph based crowd behavior analysis model proposed in [125]. (a) Position of each pedestrian or a group of individuals occluded each other. (b) Foreground detection by morphological operations (c) The connected components are gained. (d) A graph is constructed by Delaunay triangulation.

of the entropy would be higher than the threshold. Otherwise, if the crowd motion state was ordered, the entropy value would be smaller than the threshold. In [118], the motion speed and direction of moving pedestrians were calculated. Based on the motion information of individuals, the probability distribution could be worked out based on a histogram statistic. Furthermore, entropy is used to describe the direction probability distribution to detect the abnormal event of a crowd. For an input crowd video, the spatio-temporal information of some interest points in an image region can be extracted based on the degree of randomness of the directions and displacements. After that, an entropy model was used to judge the disorganization of a crowd [119]. A particle entropy based method is proposed in [120]. Moving particles was first extracted by optical flow estimation, then, the particle entropy was computed in the horizontal and vertical direction. Using the particle entropy, the abnormal crowd behavior can be detected in time.

C. TREATED CROWD AS A COMPLEX SYSTEM

A complex system is composed of many components which may interact with each other. Every pedestrian in a crowd will be interacted by other individuals. We can regard a crowd as a complex system. To describe the situation of a system, many physics methods have been researched, such as phase space, complex network and chaotic invariant. In [121] a chaotic invariants method was proposed to recognize the abnormal crowd behavior. The chaotic invariants (largest Lyapunov exponent and correlation dimension) were calculated after the clustering of particle trajectory in a moving crowd. Finally, a Gaussian mixture model was used to calculate the threshold to judge the normal and abnormal crowd behavior. Bellomo et al. discussed the mathematics model for describing crowd dynamics from a complex system viewpoint. That paper contains the idea of physics, the mathematics models that are proposed focus on the multiscale analysis, i.e., microscopic, macroscopic, and mesoscopic scales [122]. Coherent motion is an important description for crowd system; a coherent neighbor invariance method was proposed in [123]. The invariance of spatio-temporal relationships and the invariance of velocity correlations were calculated to form a coherent filtering to detect the coherent motion in a crowd. To detect motion activity in a video, Sethi and Roy-Chowdhury constructed a physics-based model in phase space. The Multi-Resolution Phase Space descriptor was formed by Sethi Metric, the Hamiltonian Energy Signature, and the Multiple Objects, Pairwise Analysis descriptors, to represent complex activities in crowd video [124].

Complex networks and graph analysis is very useful for dealing with the problem of complex systems. In the field of crowd surveillance and analysis, complex network and graph analysis is also an effective method. Chen et al. proposed a graph modeling and matching methods for crowd behavior analysis [125]. First, the pedestrians were detected as the foreground region by background subtraction and morphological operations. Then, every foreground region was considered as a node, and the Delaunay triangulation was used to connect nodes to form a graph to represent the motion of a crowd. Finally, the topology variation of the graphs was calculated to detect the anomaly behavior of a crowd. The framework of the proposed method in [125] are shown in Figure 11. A complex network model was used in [67] to detect and locate anomaly behavior of crowd. In their research, the global social network was constructed in spatio-temporal cuboids in a video scene. Following the time evolve, the global social network was updated based on local social networks.

TABLE 1. Datasets for crowd video surveillance and analysis.

Datasets name	Year	References	Application issues
Violent-flows	2012	[129]	Violent and non-violent behavior classification and violence outbreak detection
Mall	2012	[137]	Crowd counting
Grand Central	2012	[109]	Crowd motion analysis
Fudan Pedestrian Dataset	2011	[134]	Crowd counting
UCSD	2008	[127]	Crowd anomaly behavior detection
Pets2009	2009	[126]	Crowd count and density estimation, tracking in a crowd, detection of flow and crowd events
Shanghai Tech Part A and Part B	2016	[140]	Crowd counting and density estimation
UMN	2009	[128]	Abnormal crowd behavior detection
Crowd Saliency dataset	2014	[130]	Crowd saliency motion detection
CUHK	2014	[131] [132]	Group State Analysis, crowd Video Classification
UCF-Crowd Segmentation Dataset	2007	[87]	Crowd Segmentation
UCF Crowd counting Dataset	2013	[133]	Crowd counting
UCF-Tracking in High Density Crowds Data Set:	2008	[106]	Tracking in High Density Crowds
AGORASET	2012	[135]	Crowd motion simulation and analysis
SHOCK	2015	[136]	Analyzing the behavior of spectator crowd
WorldExpo '10	2015	[138]	Crowd counting in cross-scene
Large scale pedestrian walking route dataset	2015	[139]	Crowd behavior analysis

VI. DATASET FOR CROWD VIDEO SURVEILLANCE

For crowd video surveillance, the collection of crowd motion video is not an easy job. In order to facilitate researchers to verify the effectiveness of their methods, many datasets have been open to use. The dataset of crowd video analysis focuses on these research field, such as: tracking targets in crowd scene, crowd counting and density estimation, crowd motion segmentation, crowd behavior analysis and crowd abnormal event detection. Table 1 shows some of these datasets and their application issues. The sample images for each dataset are presented in Figure 12.

PETS2009 [126]: Different crowd activities are contained by multisensor sequences in this dataset. The research issue focuses on crowd count and density estimation, tracking of individuals in a crowd and detection of flow and crowd events.

Grand Central [109]: The surveillance scene of this dataset is the New Yotk Grand Central station. The length of the video is 33:20 minutes. Totally the frame number is 50010. The frame rate is 25fps. The resolution is 720×480.

UCSD [127]: The video of crowd anomaly detection in this dataset was captured by a stationary camera. There are two anomaly events in this dataset. One is that non-pedestrians appear on the pedestrian road such as a cart, wheelchair, skater and biker, the other is anomalous pedestrian motion patterns.

UMN [128]: The dataset consists of 11 different videos with the escape events in 3 different indoor or outdoor scenes. This dataset can be used to detect abnormal (panic) crowd behavior.

Violent-flows [129]: This dataset is used for classifying violent and non-violent behavior of crowds and violence outbreak detection. All of the 246 videos in this dataset can be downloaded from YouTube.

Crowd Saliency dataset [130]: The crowd saliency dataset is collected from other crowd datasets such as the UCF and Data-driven crowd dataset. The typical population saliency movement is covered in this database, such as counter flow, source, sink and instability motion.

CUHK Crowd Dataset [131], [132]: This dataset is collected for analyzing the group behavior in crowd scene. There are 474 crowd video sequences captured from 215 scenes in this dataset. The trajectories in each video sequence are extracted by gKLT trackers after deleting short trajectories, stationary points, and some errors.

UCF-Crowd Segmentation Dataset [87]: This dataset collects 38 videos about crowd and other high density moving targets from the BBC Motion Gallery and the Getty Images website.

UCF-Crowd counting Dataset [133]: This dataset contains 50 images of extremely dense crowds for crowd counting. The images are collected mainly from the FLICKR.

UCF-Tracking in High Density Crowds Data Set [106]: Three video sequences of marathon motions and their corresponding static floor fields are contained in this dataset.

Fudan Pedestrian Dataset [134]: The image sequences are captured in Guanghua Tower, Fudan University,

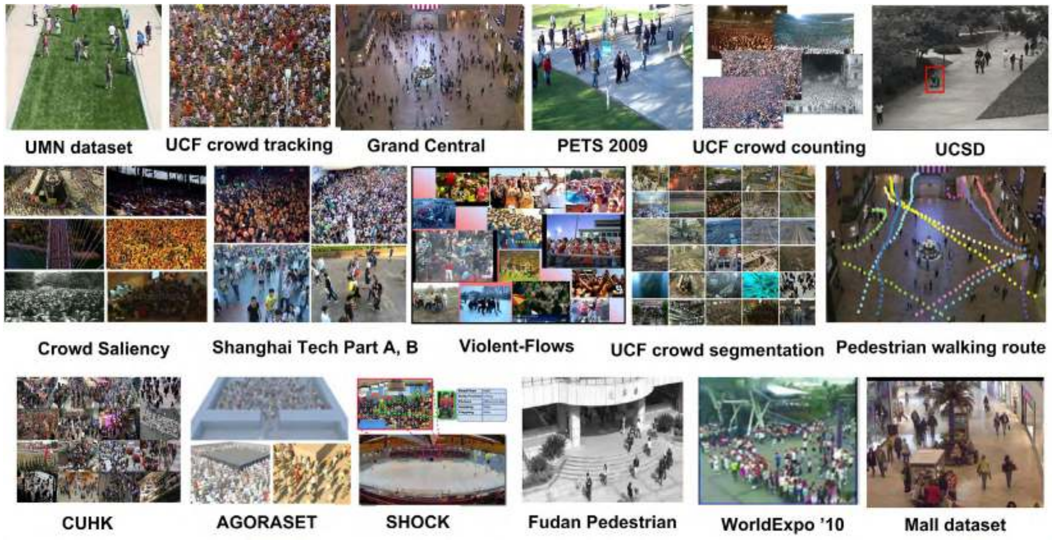

FIGURE 12. Sample images from various datasets for crowd video surveillance

for crowd counting. The number of pedestrians in this dataset is 0-15. In the Fudan pedestrian dataset, the pedestrians often have shadows under their feet.

AGORASET Dataset [135]: Agoraset is a dataset which provides the crowd motion simulation of eight scenes. Various situations are corresponded in this dataset such as viewing angle, illumination, stress of the crowd.

SHOCK [136]: This dataset is collected for analyzing the behavior of spectator crowds at stadiums/theaters/events. The individuals in this kind of crowd are semi-static. The aims of this dataset are to deal with the new questions of spectator crowd analysis, such as spectator detection and segmentation, global head orientation, automatic highlight generation, gesture segmentation and social signal processing.

Mall dataset [137]: The Mall dataset is collected using a public surveillance camera in a shopping mall. The head positions of each pedestrian are provided in this dataset. This dataset contains 2000 video frames which can be used for crowd counting.

WorldExpo '10 [138]: The purpose of this dataset is to focus on crowd counting in a cross-scene. All of the video sequences are captured using one of108 cameras in Shanghai 2010 WorldExpo. Most of the cameras are set as disjoint bird views to cover a large variety of scenes.

Large scale pedestrian walking route dataset [139]: This dataset is constructed for providing accurate pedestrian walking route in a long and crowd video. The total frame number is 100000, and the max pedestrian number in one frame is 332.

Shanghai Tech Part A and Part B [140]: This dataset is constructed for crowd counting and density estimation. Two parts are contained in this dataset. In Part A the images are chosen from the internet randomly, in Part B the images are captured in the metropolitan areas of Shanghai.

VII. CONCLUSIONS AND FUTURE DEVELOPMENTS

This paper focus on the physics method for the application on crowd video surveillance. Firstly, we have introduced three ways including controlled experiment, crowd modeling and simulation, crowd video surveillance for crowd behavior analysis. Furthermore, we have summarized three levels of crowd video surveillance from the physics point of view, i.e. macroscopic, microscopic and mesoscopic. The method based on physics has also been described in detail from three aspects, i.e. flow field analysis, force model and crowd motion system. Using physics based methods to represent and analyze crowd behavior can be used in many applications, such as abnormal crowd detection, crowd counting and density estimation, crowd segmentation, and motion detects. The physics based method almost covers the application of crowd video surveillance.

There are some open issues for crowd video analysis for future work. The authors believe the research topics below are becoming important. (1) Crowd safety status prediction: The existing analysis methods for crowd behavior analysis have mainly concentrated in the detection of abnormal crowd behavior. It means when the alarms are triggered in the video surveillance system, there must be crowd abnormal behavior. In order to improve the crowd safety, it is very important to give early warning before abnormal behavior occurs. To predict crowd safety state is very difficult. However, there are still possible solution to this problem through multi-disciplinary cross research. Here are two thoughts for this issue: (a) The research methods of video surveillance may not be limited to only machine learning and artificial intelligence techniques. More attention can be paid to the contributions of crowd disasters management based on emergency management, sociology and psychology, so as to extract the changes of crowd before disaster event occurred, such as rumor spreading, crowd gathering, etc. We suggest building a more effective model of crowd emotion based on psychological research. (b) Multiple data sources such as GPS and WIFI should be combined with video data to obtain more accurate and extensive pedestrians' positions. Crowd scene is not closed, so we suggest that the combination of traffic system data to predict the crowd status of the current scene to avoid too many pedestrians squeezing in the already crowded space. (2) Recognizing the social relationship in a crowd. For middle scale crowds, the behavior of a crowd will be different if the crowd is formed by individuals with different social relationships. Faria *et al.* [141] designed experiments to analyze leadership in the crowd. They found that if the pedestrians were provided social cues, they would move to their destination more accurately. Pedestrians are expected to identify more social relationships and cues in a crowd. To solve this problem, we recommend the following methods. (a) It will be useful to add identity recognition and emotion recognition elements to video surveillance, such as gait recognition, face recognition and expression analysis. (b) Complex networks can be used to depict the relationships among individuals. By constructing a complex network from the crowd scene, the importance of each person in the scene can be gained, which is conducive to reveal their social relationships. (3) Surveillance crowd in a large space region. One camera can only survey a fixed region, in real life, the crowd motion region is usually very large such as a big square, scenic spot and airport. How to integrate the information gained from multiply cameras is also an important direction for future work.

REFERENCES

[1] A. M. Ibrahim, I. Venkat, and P. De Wilde, "Uncertainty in a spatial evacuation model," *Phys. A, Statist. Mech. Appl.*, vol. 479, pp. 485–497, Aug. 2017.

[2] V. Dossetti, S. Bouzat, and M. N. Kuperman, "Behavioral effects in room evacuation models," *Phys. A, Statist. Mech. Appl.*, vol. 479, pp. 193–202, Aug. 2017.

[3] V. Kountouriotis, S. C. A. Thomopoulos, and Y. Papelis, "An agent-based crowd behaviour model for real time crowd behaviour simulation," *Pattern Recognit. Lett.*, vol. 44, pp. 30–38, Jul. 2014.

[4] J. Zhong, W. Cai, M. Lees, and L. Luo, "Automatic model construction for the behavior of human crowds," *Appl. Soft Comput.*, vol. 56, pp. 368–378, Jul. 2017.

[5] Y. Li, M. L. Mekhalfi, M. M. Al Rahhal, E. Othman, and H. Dhahri, "Encoding motion cues for pedestrian path prediction in dense crowd scenarios," *IEEE Access*, vol. 5, pp. 24368–24375, 2017.

[6] M. Haghani and M. Sarvi, "Following the crowd or avoiding it? Empirical investigation of imitative behaviour in emergency escape of human crowds," *Animal Behav.*, vol. 124, pp. 47–56, Feb. 2017.

[7] N. W. F. Bode and E. A. Codling, "Human exit route choice in virtual crowd evacuations," *Animal Behav.*, vol. 86, no. 2, pp. 347–358, Aug. 2013.

[8] T. Vicsek and A. Zafeiris, "Collective motion," *Phys. Rep.*, vol. 517, no. 3, pp. 71–140, 2012.

[9] B. Yogameena and C. Nagananthini, "Computer vision based crowd disaster avoidance system: A survey," *Int. J. Disaster Risk Reduction*, vol. 22, pp. 95–129, Jun. 2017.

[10] R. Leggett, "Real-time crowd simulation: A review," 2004.

[11] B. Zhan, D. N. Monekosso, P. Remagnino, S. A. Velastin, and L.-Q. Xu, "Crowd analysis: A survey," *J. March. Vis Appl.*, vol. 19, nos. 5–6, pp. 345–357, 2008.

[12] J. C. S. J. Junior, S. R. Musse, and C. R. Jung, "Crowd analysis using computer vision techniques," *IEEE Signal Process. Mag.*, vol. 27, no. 5, pp. 66–77, Sep. 2010.

[13] J. Candamo, M. Shreve, D. B. Goldgof, D. B. Sapper, and R. Kasturi, "Understanding transit scenes: A survey on human behavior-recognition algorithms," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 1, pp. 206–224, Mar. 2010.

[14] N. N. A. Sjarif, S. M. Shamsuddin, and S. Z. Hashim, "Detection of abnormal behaviors in crowd scene: A review," *Int. J. Adv. Soft Comput. Appl.*, vol. 4, no. 1, pp. 1–33, 2012.

[15] W. Hu, T. Tan, L. Wang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 34, no. 3, pp. 334–352, Aug. 2004.

[16] A. A. Sodemann, M. P. Ross, and B. J. Borghetti, "A review of anomaly detection in automated surveillance," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 42, no. 6, pp. 1257–1272, Nov. 2012.

[17] A. A. Chaaraoui, P. Climent-Pérez, and F. Flórez-Revuelta, "A review on vision techniques applied to human behaviour analysis for ambient-assisted living," *Expert Syst. Appl.*, vol. 39, no. 2, pp. 10873–10888, 2012.

[18] M. Thida, Y. L. Yong, P. Climent-Pérez, H.-L. Eng, and P. Remagnino, "A literature review on video analytics of crowded scenes," *Intelligent Multimedia Surveillance*. Berlin, Germany: Springer, 2013, pp. 17–36.

[19] H. Jo, K. Chug, and R. J. Sethi, "A review of physics-based methods for group and crowd analysis in computer vision," *J. Postdoctoral Res.*, vol. 1, no. 1, pp. 4–7, 2013.

[20] S. Vishwakarma and A. Agrawal, "A survey on activity recognition and behavior understanding in video surveillance," *Vis. Comput.*, vol. 29, no. 10, pp. 983–1009, Oct. 2013.

[21] D. Gowsikhaa, S. Abirami, and R. Baskaran, "Automated human behavior analysis from surveillance videos: A survey," *Artif. Intell. Rev.*, vol. 42, no. 4, pp. 747–765, 2014.

[22] S. A. M. Saleh, S. A. Suandi, and H. Ibrahim, "Recent survey on crowd density estimation and counting for visual surveillance," *Eng. Appl. Artif. Intell.*, vol. 41, pp. 103–114, May 2015.

[23] P. Afsar, P. Cortez, and H. Santos, "Automatic visual detection of human behavior: A review from 2000 to 2014," *Expert Syst. Appl.*, vol. 42, no. 20, pp. 6935–6956, 2015.

[24] T. Li, H. Chang, M. Wang, B. Ni, R. Hong, and S. Yan, "Crowded scene analysis: A survey," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 3, pp. 367–386, Mar. 2015.

[25] M. S. Zitouni, H. Bhaskar, J. Dias, and M. E. Al-Mualla, "Advances and trends in visual crowd analysis: A systematic survey and evaluation of crowd modelling techniques," *Neurocomputing*, vol. 186, pp. 139–159, Apr. 2016.

[26] V. J. Kok, M. K. Lim, and C. S. Chan, "Crowd behavior analysis: A review where physics meets biology," *Neurocomputing*, vol. 177, pp. 342–362, Feb. 2016.

[27] N. A. Nemade and V. V. Gohokar, "A survey of video datasets for crowd density estimation," in *Proc. Int. Conf. Global Trends Signal Process., Inf. Comput. Commun.*, 2016, pp. 389–395.

[28] S. Maheshwari and S. Heda, "A review on crowd behavior analysis methods for video surveillance," in *Proc. ICTCS*, 2016, Art. no. 52.

[29] J. M. Grant and P. J. Flynn, "Crowd scene understanding from video: A survey," *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 13, no. 2, 2017, Art. no. 19.

[30] V. A. Sindagi and V. M. Patel, "A survey of recent advances in CNN-based single image crowd counting and density estimation," *Pattern Recognit. Lett.*, vol. 107, pp. 1–14, May 2017.

[31] S. Lamba and N. Nain, "Crowd monitoring and classification: A survey," *Adv. Comput. Comput. Sci.*, vol. 553, pp. 21–31, May 2017.

[32] K. Rohit, K. Mistree, and J. Lavji, "A review on abnormal crowd behavior detection," in *Proc. Int. Conf. Innov. Inf., Embedded Commun. Syst.*, 2017, pp. 1–3.

[33] H. Y. Swathi, G. Shivakumar, and H. S. Mohana, "Crowd behavior analysis: A survey," in *Proc. Int. Conf. Recent Adv. Electron. Commun. Technol.*, 2017, pp. 169–178.

[34] H. Gayathri, P. M. Aparna, and A. Verma, "A review of studies on understanding crowd dynamics in the context of crowd safety in mass religious gatherings," *Int. J. Disaster Risk Reduction*, vol. 25, pp. 82–91, Oct. 2017.

[35] Y. Ma, L. Li, H. Zhang, and T. Chen, "Experimental study on small group behavior and crowd dynamics in a tall office building evacuation," *Phys. A, Statist. Mech. Appl.*, vol. 473, pp. 488–500, May 2017.
[36] N. Shiwakoti, Y. Gong, X. Shi, and Z. Ye, "Examining influence of merging architectural features on pedestrian crowd movement," *Safety Sci.*, vol. 75, pp. 15–22, Jun. 2015.
[37] O. J. Akinwande, H. Bi, and E. Gelenbe, "Managing crowds in hazards with dynamic grouping," *IEEE Access*, vol. 3, pp. 1060–1070, 2015.
[38] I. Karamouzas and M. Overmars, "Simulating and evaluating the local behavior of small pedestrian groups," *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 3, pp. 394–406, Mar. 2012.
[39] M. Xu, Y. Wu, P. Lv, H. Jiang, M. Luo, and Y. Ye, "miSFM: On combination of mutual information and social force model towards simulating crowd evacuation," *Neurocomputing*, vol. 168, pp. 529–537, Nov. 2015.
[40] S. Wu and H. Wong, "Crowd motion partitioning in a scattered motion field," *IEEE Trans. Syst., Man, and*, vol. 42, no. 5, pp. 1443–1454, Oct. 2012.
[41] Y. Cong, J. Yuan, and J. Liu, "Abnormal event detection in crowded scenes using sparse representation," *Pattern Recognit.*, vol. 46, no. 7, pp. 1851–1864, 2013.
[42] Y. Wang, S. Piérard, S.-Z. Su, and P.-M. Jodoin, "Improving pedestrian detection using motion-guided filtering," *Pattern Recognit. Lett.*, vol. 96, pp. 106–112, Sep. 2017.
[43] S.-I. Jung and K.-S. Hong, "Deep network aided by guiding network for pedestrian detection," *Pattern Recognit. Lett.*, vol. 90, pp. 43–49, Apr. 2017.
[44] Z. Zhang, W. Tao, K. Sun, W. Hu, and L. Yao, "Pedestrian detection aided by fusion of binocular information," *Pattern Recognit.*, vol. 60, pp. 227–238, Dec. 2016.
[45] C. Zhu and Y. Peng, "Discriminative latent semantic feature learning for pedestrian detection," *Neurocomputing*, vol. 238, pp. 126–138, May 2017.
[46] N. Aparajita, P. K. Sa, S. K. Choudhury, S. Bakshi, and B. Majhi, "A neuromorphic person re-identification framework for video surveillance," *IEEE Access*, vol. 5, pp. 6471–6482, 2017.
[47] Z. Boukhers, Y. Wang, K. Shirahama, K. Uehara, and M. Grzegorzek, "Convoy detection in crowded surveillance videos," in *Proc. 7th Int. Workshop Hum. Behav. Understand. (HBU)*, in Lecture Notes in Computer Science, vol. 9997, 2016, pp. 137–147.
[48] H. Idrees, N. Warner, and M. Shah, "Tracking in dense crowds using prominence and neighborhood motion concurrence," *Image Vis. Comput.*, vol. 32, no. 1, pp. 14–26, Jan. 2014.
[49] Z. Jin and B. Bhanu, "Analysis-by-synthesis: Pedestrian tracking with crowd simulation models in a multi-camera video network," *Comput. Vis. Image Understand.*, vol. 134, pp. 48–63, May 2015.
[50] H. Fradi, V. Eiselein, J.-L. Dugelay, I. Keller, and T. Sikora, "Spatio-temporal crowd density model in a human detection and tracking framework," *Signal Process., Image Commun.*, vol. 31, pp. 100–111, Feb. 2015.
[51] I. Ali and M. N. Dailey, "Multiple human tracking in high-density crowds," *Image Vis. Comput.*, vol. 30, no. 12, pp. 966–977, Dec. 2012.
[52] M. Fu, P. Xu, X. Li, Q. Liu, M. Ye, and C. Zhu, "Fast crowd density estimation with convolutional neural networks," *Eng. Appl. Artif. Intell.*, vol. 43, pp. 81–88, Aug. 2015.
[53] Y. Zhang, F. Chang, M. Wang, F. Zhang, and C. Han, "Auxiliary learning for crowd counting via count-net," *Neurocomputing*, vol. 273, pp. 190–198, Jan. 2018.
[54] H. Fradi and J.-L. Dugelay, "Towards crowd density-aware video surveillance applications," *Inf. Fusion*, vol. 24, pp. 3–15, Jul. 2015.
[55] D. Ryan, S. Denman, C. Fookes, and S. Sridharan, "Scene invariant multi camera crowd counting," *Pattern Recognit. Lett.*, vol. 44, pp. 98–112, Jul. 2014.
[56] H. Ullah and N. Conci, "Structured learning for crowd motion segmentation," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2013, pp. 824–828.
[57] V. J. Kok and C. S. Chan, "GrCS: Granular computing-based crowd segmentation," *IEEE Trans. Cybern.*, vol. 47, no. 5, pp. 1157–1168, May 2017.
[58] C. C. Loy, T. Xiang, and S. Gong, "Salient motion detection in crowded scenes," in *Proc. 5th Int. Symp. Commun., Control Signal Process.*, 2012, pp. 1–4.
[59] Y. Yuan, Y. Feng, and X. Lu, "Statistical hypothesis detector for abnormal event detection in crowded scenes," *IEEE Trans. Cybern.*, vol. 47, no. 11, pp. 3597–3608, Nov. 2016.
[60] P. Liu, Y. Tao, W. Zhao, and X. Tang, "Abnormal crowd motion detection using double sparse representation," *Neurocomputing*, vol. 269, pp. 3–12, Dec. 2017.
[61] C. Chen, Y. Shao, and X. Bi, "Detection of anomalous crowd behavior based on the acceleration feature," *IEEE Sensors J.*, vol. 15, no. 12, pp. 7252–7261, Dec. 2015.
[62] X. Chong, W. Liu, P. Huang, and N. I. Badler, "Hierarchical crowd analysis and anomaly detection," *J. Vis. Lang. Comput.*, vol. 25, no. 4, pp. 376–393, 2014.
[63] S. Yi and X. Wang, "Profiling stationary crowd groups," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2014, pp. 1–6.
[64] J. Shao, K. Kang, C. C. Loy, and X. Wang, "Deeply learned attributes for crowded scene understanding," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 4657–4666.
[65] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 935–942.
[66] G. Xiong, J. Cheng, X. Wu, Y.-L. Chen, Y. Ou, and Y. Xu, "An energy model approach to people counting for abnormal crowd behavior detection," *Neurocomputing*, vol. 83, pp. 121–135, Apr. 2012.
[67] R. Chaker, Z. Al Aghbari, and I. N. Junejo, "Social network model for crowd anomaly detection and localization," *Pattern Recognit.*, vol. 61, pp. 266–281, Jan. 2017.
[68] A. N. Shuaibu, I. Faye, Y. S. Ali, N. Kamel, M. N. Saad, and A. S. Malik, "Sparse representation for crowd attributes recognition," *IEEE Access*, vol. 5, pp. 10422–10433, 2017.
[69] J. Wang and Z. Xu, "Spatio-temporal texture modelling for real-time crowd anomaly detection," *Comput. Vis. Image Understand.*, vol. 144, pp. 177–187, Mar. 2016.
[70] G. Cheng, Y. Wan, A. N. Saudagar, K. Namuduri, and B. P. Buckles, "Advances in human action recognition: A survey," *Comput. Sci.*, 2015.
[71] S. Blunsden, E. L. Andrade, and R. B. Fisher, "Modelling crowd scenes for event detection," in *Proc. Int. Conf. Pattern Recognit.*, vol. 1, 2006, pp. 175–178.
[72] W. Ge and R. T. Collins, "Marked point processes for crowd counting," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 2913–2920.
[73] L. Kratz and K. Nishino, "Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 1446–1453.
[74] P. Tu, T. Sebastian, G. Doretto, N. Krahnstoever, J. Rittscher, and T. Yu, "Unified crowd segmentation," in *Proc. Eur. Conf. Comput. Vis.*, 2008, pp. 691–704.
[75] G. J. Brostow and R. Cipolla, "Unsupervised Bayesian detection of independent motion in crowds," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 1, Jun. 2006, pp. 594–601.
[76] S. J. Blunsden and R. B. Fisher, "Pre-fight detection-classification of fighting situations using hierarchical AdaBoost," in *Proc. Int. Conf. Comput. Vis. Theory Appl.*, Lisbon, Portugal, vol. 12, 2009, pp. 303–308.
[77] P. Guo and Z. Miao, "Action detection in crowded videos using masks," in *Proc. Int. Conf. Pattern Recognit.*, 2010, pp. 1767–1770.
[78] W. Ge, R. T. Collins, and R. B. Ruback, "Vision-based analysis of small groups in pedestrian crowds," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 5, pp. 1003–1016, May 2012.
[79] X. Zhang, M. Ouyang, and X. Zhang, "Small scale crowd behavior classification by Euclidean distance variation-weighted network," *Multimedia Tools Appl.*, vol. 75, no. 19, pp. 11945–11960, 2016.
[80] R. Mehran, B. E. Moore, and M. Shah, "A streakline representation of flow in crowded scenes," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Crete, Greece, 2010, pp. 439–452.
[81] X. Wang, X. Yang, X. He, Q. Teng, and M. Gao, "A high accuracy flow segmentation method in crowded scenes based on streakline," *Optik-Int. J. Light Electron Opt.*, vol. 125, no. 3, pp. 924–929, Feb. 2014.
[82] G. M. Basavaraj and A. Kusagur, "Optical and streakline flow based crowd estimation for surveillance system," in *Proc. IEEE Int. Conf. Recent Trends Electron. Inf. Commun. Technol.*, May 2016, pp. 414–416.
[83] D. Zhang, J. Xu, M. Sun, and Z. Xiang, "High-density crowd behaviors segmentation based on dynamical systems," *Multimedia Syst.*, vol. 23, no. 5, pp. 599–606, Oct. 2017.
[84] D. Weiskopf and G. Erlebacher, "Overview of flow visualization," in *The Visualization Handbook*. New York, NY, USA: Elsevier, 2005, pp. 261–278.

[85] X. Zhang, H. He, and S. Cao, "Flow field texture representation-based motion segmentation for crowd counting," *Mach. Vis. Appl.*, vol. 26, nos. 7–8, pp. 871–883, 2015.
[86] B. Solmaz, B. E. Moore, and M. Shah, "Identifying behaviors in crowd scenes using stability analysis for dynamical systems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 10, pp. 2064–2070, Oct. 2012.
[87] S. Ali and M. Shah, "A Lagrangian particle dynamics approach for crowd flow segmentation and stability analysis," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–6.
[88] S. Wu, H. Su, H. Yang, S. Zheng, Y. Fan, and Q. Zhou, "Bilinear dynamics for crowd video analysis," *J. Vis. Commun. Image Represent.*, vol. 48, pp. 461–470, Oct. 2017.
[89] S. Wu, H. Yang, S. Zheng, H. Su, Y. Fan, and M.-H. Yang, "Crowd behavior analysis via curl and divergence of motion trajectories," *Int. J. Comput. Vis.*, vol. 123, no. 3, pp. 499–519, 2017.
[90] M. K. Lim, C. S. Chan, D. Monekosso, and P. Remagnino, "Detection of salient regions in crowded scenes," *Electron. Lett.*, vol. 50, no. 5, pp. 363–365, Feb. 2014.
[91] M. Hu, S. Ali, and M. Shah, "Detecting global motion patterns in complex videos," in *Proc. Int. Conf. Pattern Recognit.*, 2008, pp. 1–5.
[92] S. D. Khan, G. Vizzari, and S. Bandini, "Identifying sources and sinks and detecting dominant motion patterns in crowds," in *Proc. Transp. Res. Procedia, Conf. Pedestrian Evacuation Dyn.*, Amsterdam, The Netherlands, vol. 2, Oct. 2014, pp. 195–200.
[93] S. D. Khan, S. Bandini, S. Basalamah, and G. Vizzari, "Analyzing crowd behavior in naturalistic conditions: Identifying sources and sinks and characterizing main flows," *Neurocomputing*, vol. 177, pp. 543–563, Feb. 2016.
[94] D. Helbing and P. Molnar, "Social force model for pedestrian dynamics," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 51, no. 5, pp. 4282–4286, 1995.
[95] D. Helbing, I. Farkas, and T. Vicsek, "Simulating dynamical features of escape panic," *Nature*, vol. 407, pp. 487–490, Sep. 2000.
[96] D. R. Parisi, M. Gilman, and H. Moldovan, "A modification of the social force model can reproduce experimental data of pedestrian flows in normal conditions," *Phys. A, Statist. Mech. Appl.*, vol. 388, no. 17, pp. 3600–3608, 2009.
[97] Y. Gao, P. B. Luh, H. Zhang, and T. Chen, "A modified social force model considering relative velocity of pedestrians," in *Proc. IEEE Int. Conf. Autom. Sci. Eng. (CASE)*, Aug. 2013, pp. 747–751.
[98] Y. Han and H. Liu, "Modified social force model based on information transmission toward crowd evacuation simulation," *Phys. A, Statist. Mech. Appl.*, vol. 469, pp. 499–509, Mar. 2017.
[99] B. E. Moore, S. Ali, R. Mehran, and M. Shah, "Visual crowd surveillance through a hydrodynamics lens," *Commun. ACM*, vol. 54, no. 12, pp. 64–73, Dec. 2011.
[100] R. Raghavendra, A. Del Bue, M. Cristani, and V. Murino, "Abnormal crowd behavior detection by social force optimization," in *Proc. Int. Conf. Hum. Behav. Unterstand.*, 2011, pp. 134–145.
[101] H. Yang, Y. Cao, S. Wu, W. Lin, S. Zheng, and Z. Yu, "Abnormal crowd behavior detection based on local pressure model," in *Proc. Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, 2012, pp. 1–4.
[102] K. N. Tran, A. Gala, A. Kakadiaris, and S. K. Shah, "Activity analysis in crowded environments using social cues for group discovery and human interaction modeling," *Pattern Recognit. Lett.*, vol. 44, pp. 49–57, Jul. 2014.
[103] Y. Zhang, L. Qin, R. Ji, H. Yao, and Q. Huang, "Social attribute-aware force model: Exploiting richness of interaction for abnormal crowd detection," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 7, pp. 1231–1245, Jul. 2015.
[104] E. L. Andrade and R. B. Fisher, "Simulation of crowd problems for computer vision," in *Proc. 1st Int. Workshop Crowd Simulation*, 2005, pp. 71–80.
[105] V. Srikrishnan and S. Chaudhuri, "Crowd motion analysis using linear cyclic pursuit," in *Proc. Int. Conf. Pattern Recognit.*, 2010, pp. 3340–3343.
[106] S. Ali and M. Shah, "Floor fields for tracking in high density crowd scenes," in *Proc. 10th Eur. Conf. Comput. Vis.*, Marseille, France. Berlin, Germany: Springer, Oct. 2008, pp. 1–14.
[107] H. Su, H. Yang, S. Zheng, Y. Fan, and S. Wei, "The large-scale crowd behavior perception based on spatio-temporal viscous fluid field," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 10, pp. 1575–1589, Oct. 2013.
[108] H. Ullah, M. Uzair, M. Ullah, A. Khan, A. Ahmad, and W. Khan, "Density independent hydrodynamics model for crowd coherency detection," *Neurocomputing*, vol. 242, pp. 28–39, Jun. 2017.
[109] B. Zhou, X. Wang, and X. Tang, "Understanding collective crowd behaviors: Learning a mixture model of dynamic pedestrian-agents," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2871–2878.
[110] Y. Zhao, M. Yuan, G. Su, and T. Chen, "Crowd macro state detection using entropy model," *Phys. A, Statist. Mech. Appl.*, vol. 431, pp. 84–93, Aug. 2015.
[111] X. Cui, Q. Liu, M. Gao, and D. N. Metaxas, "Abnormal detection using interaction energy potentials," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 3161–3167.
[112] K. Yamaguchi, A. C. Berg, L. E. Ortiz, and T. L. Berg, "Who are you with and Where are you going?" in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 1345–1352.
[113] Y. Yuan, J. Fang, and Q. Wang, "Online anomaly detection in crowd scenes via structure analysis," *IEEE Trans. Cybern.*, vol. 45, no. 3, pp. 548–561, Mar. 2015.
[114] G. Xiong, X. Wu, J. Cheng, Y.-L. Chen, Y. Ou, and Y. Liu, "Crowd density estimation based on image potential energy model," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, Dec. 2011, pp. 538–543.
[115] S. Yi, H. Li, and X. Wang, "Pedestrian behavior modeling from stationary crowds with applications to intelligent surveillance," *IEEE Trans. Image Process.*, vol. 25, no. 9, pp. 4354–4368, Sep. 2016.
[116] W. Lin, H. Chu, J. Wu, B. Sheng, and Z. Chen, "A heat-map-based algorithm for recognizing group activities in videos," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 11, pp. 1980–1992, Nov. 2013.
[117] G. Bressan, C. Barnaba, S. Gentili, and G. Rossi, "Information entropy of earthquake populations in northeastern Italy and western Slovenia," *Phys. Earth Planetary Interiors*, vol. 271, pp. 29–46, Oct. 2017.
[118] Y. Nam, "Crowd flux analysis and abnormal event detection in unstructured and structured scenes," *Multimedia Tools Appl.*, vol. 72, no. 3, pp. 3001–3029, 2014.
[119] M. H. Sharif and C. Djeraba, "An entropy approach for abnormal activities detection in video streams," *Pattern Recognit.*, vol. 45, no. 7, pp. 2543–2561, Jul. 2012.
[120] X. Gu, J. Cui, and Q. Zhu, "Abnormal crowd behavior detection by using the particle entropy," *Optik-Int. J. Light Electron Opt.*, vol. 12, no. 14, pp. 3428–3433, 2014.
[121] S. Wu, B. E. Moore, and M. Shah, "Chaotic invariants of Lagrangian particle trajectories for anomaly detection in crowded scenes," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 2054–2060.
[122] N. Bellomo, B. Piccoli, and A. Tosin, "Modeling crowd dynamics from a complex system viewpoint," *Math. Models Methods Appl. Sci.*, vol. 22, no. 2, pp. 296–351, 2012.
[123] B. Zhou, X. Tang, and X. Wang, "Coherent filtering: Detecting coherent motions from crowd clutters," in *Computer Vision—ECCV*. Berlin, Germany: Springer-Verlag, 2012, pp. 857–871.
[124] R. J. Sethi and A. K. Roy-Chowdhury, "Physics-based activity modelling in phase Space," in *Proc. ICVGIP*, 2010, pp. 170–177.
[125] D.-Y. Chen and P.-C. Huang, "Visual-based human crowds behavior analysis based on graph modeling and matching," *IEEE Sensors J.*, vol. 13, no. 6, pp. 2129–2138, Jun. 2013.
[126] J. Ferryman and A. Shahrokni, "PETS2009: Dataset and challenge," in *Proc. Int. Workshop Perform. Eval. Tracking Surveill. (PETS)*, 2009, pp. 1–6.
[127] A. B. Chan, Z.-S. J. Liang, and N. Vasconcelos, "Privacy preserving crowd monitoring: Counting people without people models or tracking," in *Proc. Comput. Vis. Pattern Recognit. (CVPR)*, 2008, pp. 1–7.
[128] "Unusual crowd activity dataset," Dept. Comput. Sci. Eng., Univ. Minnesota, Minneapolis, MN, USA, Tech. Rep. [Online]. Available: http://mha.cs.umn.edu/Movies/Crowd-Activity-All.avi
[129] T. Hassner, Y. Itcher, and O. Kliper-Gross, "Violent flows: Real-time detection of violent crowd behavior," in *Proc. 3rd IEEE Int. Workshop Socially Intell. Surveill. Monitoring (SISM), IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Providence, RI, USA, Jun. 2012, pp. 1–6.
[130] M. K. Lim, V. J. Kok, C. C. Loy, and C. S. Chan, "Crowd saliency detection via global similarity structure," in *Proc. Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2014, pp. 3957–3962.
[131] J. Shao, C. C. Loy, and X. Wang, "Scene-independent group profiling in crowd," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2014, pp. 2227–2234.

[132] J. Shao, C. C. Loy, and X. Wang, "Learning scene-independent group descriptors for crowd understanding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 6, pp. 1290–1303, Jun. 2016.
[133] H. Idrees, I. Saleemi, C. Seibert, and M. Shah, "Multi-source multi-scale counting in extremely dense crowd images," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2013, pp. 2547–2554.
[134] B. Tan, J. Zhang, and L. Wang, "Semi-supervised elastic net for pedestrian counting," *Pattern Recognit.*, vol. 44, nos. 10–11, pp. 2297–2304, Oct./Nov. 2011.
[135] P. Allain, N. Courty, and T. Corpetti, "AGORASET: A dataset for crowd video analysis," in *Proc. Int. Workshop Pattern Recognit. Crowd Anal.*, 2012, pp. 1–6.
[136] D. Conigliaro *et al.*, "The S-Hock dataset: Analyzing crowds at the stadium," in *Proc. CVPR*, 2015, pp. 2039–2047.
[137] K. Chen, C. C. Loy, S. Gong, and T. Xiang, "Feature mining for localised crowd counting," in *Proc. Eur. Conf. Comput. Vis.*, 2012, pp. 1–11.
[138] C. Zhang, H. Li, X. Wang, and X. Yang, "Cross-scene crowd counting via deep convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 833–841.
[139] S. Yi, H. Li, and X. Wang, "Understanding pedestrian behaviors from stationary crowd groups," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Patter Recognit.*, Jun. 2015, pp. 3488–3496.
[140] Y. Zhang, D. Zhou, S. Chen, S. Gao, and Y. Ma, "Single-image crowd counting via multi-column convolutional neural network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 589–597.
[141] J. J. Faria, J. R. G. Dyer, C. R. Tosh, and J. Krause, "Leadership and social information use in human crowds," *Animal Behav.*, vol. 79, no. 4, pp. 895–901, Apr. 2010.

XUGUANG ZHANG received the B.S. degree in electrical technology from Northeast Normal University in 2000 and the Ph.D. degree in machinery and electronics engineering from the Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, China, in 2008. He was a Professor at the School of Electrical Engineering, Yanshan University, China. He is currently a Professor at the School of Communication Engineering, Hangzhou Dianzi University, China. His research interests include video and image processing, crowd behavior analysis, and human behavior understanding.

QINAN YU received the B.S. degree in communication engineering from Anqing Normal University, China, in 2016. She is currently pursuing the master's degree with Hangzhou Dianzi University, China. Her research interests are crowd video analysis.

HUI YU is currently a Professor with the University of Portsmouth, U.K. His research interests include vision, computer graphics, and application of machine learning to above areas, particularly in human machine interaction, image processing and recognition, virtual/augmented reality, 3-D reconstruction, robotics, and geometric processing of human/facial performances. He is an Associate Editor of the IEEE Transactions on Human-Machine Systems and *Neurocomputing* journal.