Crowd Monitoring: State-of-the-art and Future Directions

Utkarsh Singh, Jean-François Determe, François Horlin, Philippe De Doncker

U. Singh was with the OPERA – Wireless Communications Group at Université libre de Bruxelles. He is now with the Artificial Intelligence Lab, Vrije Universiteit Brussel, Brussels 1050, Belgium. (Corresponding author, E-mail: utkarsh.singh@ulb.ac.be)

J.-F. Determe is with the OPERA – Wireless Communications Group at Université libre de Bruxelles, Brussels 1050, Belgium. (E-mail: jdeterme@ulb.ac.be)

F. Horlin is with the OPERA – Wireless Communications Group at Université libre de Bruxelles, Brussels 1050, Belgium. (E-mail: fhorlin@ulb.ac.be)

Philippe De Doncker is with the OPERA – Wireless Communications Group at Université libre de Bruxelles, Brussels 1050, Belgium. (E-mail: pdedonck@ulb.ac.be)

Abstract— With the growing concerns over public safety, the importance of crowd monitoring is being realized by various security and event management agencies worldwide. Estimation of crowd dynamics can help such agencies in prevention of any unanticipated accidents or issues. Research on crowd monitoring has been underway since the past few decades. Conventional crowd monitoring systems mainly rely on computer vision approach. Due to predominant use of videos/ image sequences, the existing techniques may raise data privacy concerns. This has led to development of new crowd monitoring techniques which are privacy preserving and require minimum public participation. This paper aims to serve as a single and sufficient source of information to the concerned researchers on various aspects of crowd monitoring and also provide future directions which can be helpful in developing advanced crowd monitoring techniques.

Keywords—Artificial intelligence, Crowd monitoring, Computer vision, Detection, Forecasting, Sensor Networks, Public safety

1. INTRODUCTION

Crowding is a common phenomenon observed during major events such as concerts, festivals, sports, games, and entertainment. There is a potential risk to public safety during such large scale gatherings. Recent examples include the mishaps during 2006 Mecca pilgrimage in Saudi Arabia and 2010 Love Parade in Germany [1]. Crowd monitoring helps in estimation of crowd dynamics; which can help in better event management and ensuring public safety. It is therefore, imperative for the event organizers to monitor the crowd, detect/predict possible risks and take possible mitigative measures. Besides, crowd monitoring has various added advantages, such as:

- It can help in better layout of the venue of event. For example, well-designed food points, shopping points, entries, and exits can prevent overcrowding during an event.
- Analysis of crowd behaviour can help in preparing crowd management strategies in overcrowded events such as concerts, Olympics, and demonstrations.
- When used at a micro-level (e.g. facial recognition), it can help in tracking and apprehending suspects.
- It can be useful in developing intelligent technologies for automated crowd assistance in large-scale events.
- Crowd monitoring can also give ideas for virtual simulation and testing of crowd behaviour.

The early research has mainly focused on crowd monitoring systems using image or video sequences (vision based approach) [2]. However, there is limited possibility of using such systems for real-time situational awareness [3]. Efforts have also been made to identify and define parameters associated with crowd dynamics. Parameters like crowd density, velocity, and flow direction can be crucial in risk analysis [4]-[6]. With the technological advancements, the research is being gradually redirected from vision based approach towards sensor based crowd monitoring systems. Modern technologies like Smartphone, GPS and Bluetooth have already been proposed for estimation of crowd location, density, and dynamics during major events [7]-[10]. However, even the recent crowd monitoring

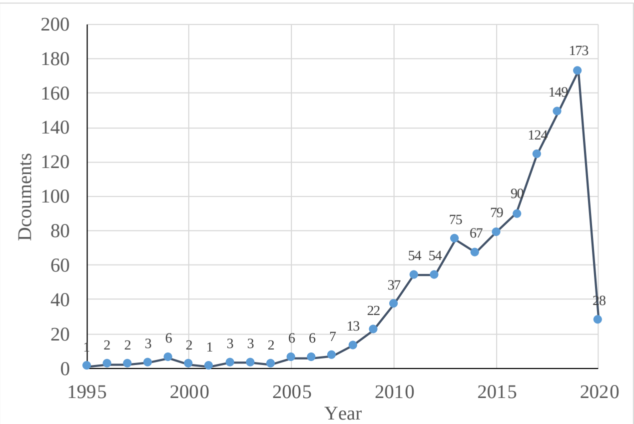

Fig. 1. Documents published per year on 'crowd monitoring'. [Source: Scopus]

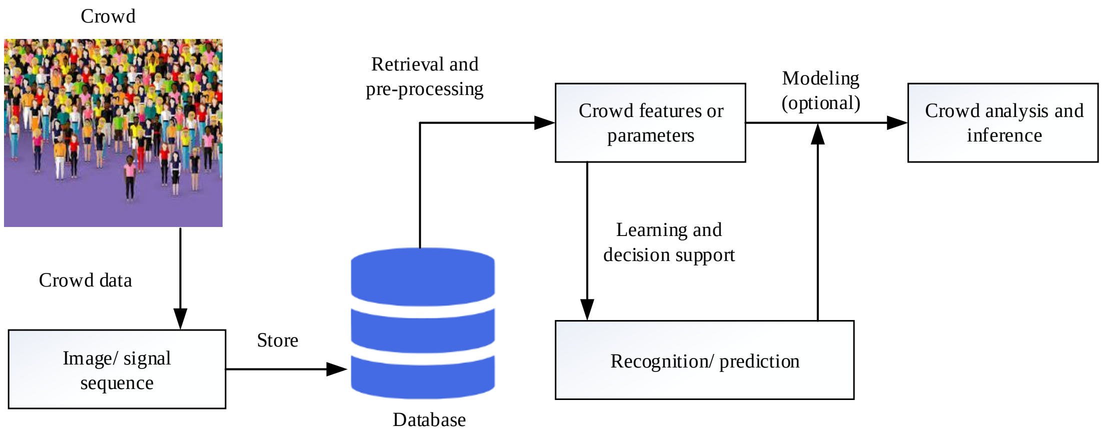

Fig. 2. Generalized representational model of a crowd monitoring system.

technologies might become obsolete with the emergence data privacy laws. Thus, there is a need to develop new methods which are less intrusive yet more efficient. A Scopus database search with the keyword 'crowd monitoring' in all categories, shows 1009 published documents till the time of writing this paper. The surge in the published documents over the past decade (Fig. 1), clearly shows a rising interest in crowd monitoring. Only those publications are used in this review which have made unique theoretical or experimental contributions to the field.

This paper starts with a generalized idea of a crowd monitoring system in Section 2. Section 3 provides a discussion on the features and parameters which are useful for crowd characterization. Sections 4 and 5 provide discussions on crowd recognition and tracking, respectively. Possible use cases are described at the beginning of sections 3, 4, and 5. For the ease of reading, sections 4 and 5 are divided into vision and non-vision categories. Section 6 provides an insight of various crowd modeling and inference techniques divided into vision and non-vision categories. The paper provides additional useful information in Section 7 and ends with future directions in Section 8.

2. CROWD MONITORING SYSTEM

A crowd monitoring system can be realized as a combination of image/signal data, storage database, learning and decision support, modeling and analysis. Fig. 2 shows a generalized representational model of a crowd monitoring system for the purpose of explanation. Initially, the crowd data is directly/ indirectly acquired using devices such as camera, smartphones, sensors, Bluetooth, WiFi, GPS, etc. The crowd data can be an image, sequence of images, or signals. Data pertaining to crowd is usually stored using several identification factors (location, device number, time-stamp, etc.), so that it can be easily retrieved as and when required [11]. A pre-processing may be required based on the data type. For example, sampling and noise filtering may be required in case of images [12]. Additionally, the data might need sorting or segregation based on the identification factors for ease of analysis. The next stage involves image/ signal processing which help in extracting key features or parameters pertaining to the crowd. Lin et al. in [13] estimated the crowd size by employing Haar wavelet for image processing. Similarly, Fourier-Fractal analysis was used for crowd density estimation in [14] and Kalman filter was used for background removal in [15]. Crowd features and parameters provide important information pertaining to crowd. For example, too many people per in a given area might signify high crowd density and need for immediate attention. Such information can be utilized for crowd analysis and inference. Crowd modeling is an optional aspect of crowd analysis [16]. It just helps in providing a better inference of dynamic crowd behaviour using mathematical models. However, the necessity of modeling depends on the objective of crowd analysis. If the objective is as simple as inferring the level of crowd density, it can be done using crowd features and supervised learning. Intelligent systems can additionally be employed for learning and decision support. Such systems can help in crowd recognition, risk prediction, and mitigation. Authors in [13], utilize trained support vector machines to estimate the crowd size in test images. Marana et al. in [17], have classified image pixels into different textures using self-organizing neural networks and thus estimated the crowd density. Factors such as occlusion, huge data handling, cost of implementation and privacy concerns inhibit the possibility of using vision based crowd monitoring systems for real-time situational awareness [3]. An insight into new technologies shows WiFi sensors being the popular choice [18]. Possible reasons for this is that WiFi systems are not only free from occlusion, but are also non-intrusive and more cost effective [19]. Some relevant applications of WiFi technology include counting pedestrians in urban traffic [20], indoor counting system [21], crowd speed estimation [22], and occupancy based lightning control in smart buildings [23].

3. CROWD CHARACTERIZATION

A crowd can be characterized by the features or parameters associated with it. This section provides an overview of the features/ parameters which can be useful in crowd characterization. The motivation for this discussion comes from the various uses of crowd characterization. Some use cases for crowd characterization may include behaviour analysis, ascertaining crowd simulation environment, better event management, personal assistance, and turbulence detection. Crowd parameter is a measurement or estimation which helps in inference of crowd dynamics, whereas a crowd feature can be considered as the distinctive attribute of a crowd which is used to compute crowd parameter. Various parameters found in the literature are as follows: crowd count, crowd density, crowd velocity, crowd turbulence, crowd pressure. Out of these parameters, crowd count and crowd density have been most extensively researched, given their prominence in vision based crowd monitoring. The empirical connections are not intended to be at the forefront of this paper. But to ensure that a correct and context based usage is adopted by the researchers, following articles are recommended as an introduction [24] and [25]. Crowd count and density are empirically similar because both signify the estimation of number of people in a crowded scenario. Crowd density is slightly different in the sense that it is estimated as the ratio of number of people per unit area (e.g. pedestrian) or number of people detected per segment (in an image). Chan et al. [26] presented a privacy preserved crowd monitoring technique which used 28 features (based on segment, edge, and texture) extracted from image segments for crowd counting. Though the given approach does not focus on an individual's identity, but it cannot be considered privacy preserved approach because it ultimately uses crowd images for training and testing. A similar approach to crowd counting is also reported in [27], where segment, structure, and local texture based features have been extracted from crowd images. Xi et al. in [18] use Grey theory to study the effect of moving people on wireless channel state information. Crowd count is estimated through aggregation over all the receivers associated to channel. This approach is truly non-intrusive but has a limited scope due to limited range of WiFi devices. Authors in [12] have counted crowd using a group of image sensors. Each sensor distinguishes the foreground objects from background. The actual count is obtained by aggregating resulting silhouettes for all sensors in the network. An algorithm is also implemented in each sensor to check false counts of the number of people and locations. As already mentioned, crowd density is another way of perceiving the number of attendees in an event. Authors in [28] measure crowd density in terms of number of pedestrians per unit area. Based on this, the crowd flows have been classified as: free, restricted, dense, and jammed. Such a classification can be useful in necessary follow up by the event organizers. A lot of researchers have used pixel analysis or background removal to estimate crowd density in vision based crowd monitoring. In [29], classification of the image pixels into background and people has been used as a basis for crowd density estimation. Authors in [30] have used background removal and proposed a geometric correction to derive a relation between the pixel and persons. It must be noted that these mathematical relations between pixel and persons do not account for occlusion and are thus unreliable. This problem has been addressed in [31], where global shape features obtained via Fourier descriptors are directly mapped on different human configurations. Marana et al. [32] showed that there is a direct relation between the image texture and crowd density, such that coarse texture corresponds to low density and fine texture corresponds to high density. Wavelet transform based image processing can also give useful features for crowd density estimation, such as: image energy [12], edge and colour [33]. A representation of crowd parameters through heat maps using GPS based information can be found in [10]. The authors have provided mathematical expressions for estimation of crowd density, velocity, pressure, and turbulence. Those definitions are case specific and there are further possibilities to define crowd parameters in general context. For further information on crowd characterization and underlying parameters, the readers may refer to references [34] and [35].

4. CROWD RECOGNITION

This section elaborates on the techniques used to recognize individuals in crowd or the crowd as a single entity. For example, recognizing an individual through body features or recognizing a crowd en masse by fingerprinting based density estimation. Some use cases for crowd recognition may include crowd counting, preparing crowd time series, forecasting crowd count/ density, checking abnormal densities, evacuation, crowd control, crowd modeling, identifying crowd attributes, and image labelling. While this section may not necessarily include references directly addressing crowd recognition, it does discuss the idea of crowd recognition possible through various approaches.

4.1. Vision based crowd recognition

In vision based approach, it is easier to simply count the number of people with the help of image processing and learning techniques. Face is the most distinctive feature of people in a crowded scenario and is thus it has been extensively studied in the literature. Most of the work have employed supervised learning techniques to train a learner on different faces obtained from open access datasets. However, the demerit of this approach is that the face is a unique feature and in order to deliver good performance the learners must be trained on an excessively large number of images; which takes considerable amount of time. Even after a significant amount of training, high accuracy cannot be guaranteed. Authors in [36] use genetic algorithm for image segmentation and recognition. A mean image is used for the training whereas testing has been performed on random images. A texture based recognition strategy has been demonstrated in [37]. Jones et al. in [38] present a fast technique for classification of viewpoints using decision tree. Authors in [39], have offered the advantages of both speed and accuracy in detection of pose and perspective of individuals by employing width-first search mechanism in a decision tree. Another way of head recognition is proposed in [40], where scale-adaptive filtering, spurious clue suppression, and mean-shift algorithm have been utilized. In general, most of the crowd recognition techniques in computer vision exploit the texture, color and shape in images. A critical issue in crowd recognition is occlusion, which occurs due to high cluttering. Wu et al. [41] have tackled partial occlusion by employing edgelet features for body part detection. The detection results are then utilized to prepare a likelihood model for inter-occluded individuals. Appearance and spatial distribution have been utilized to prepare a crowd model in [42]. A 2D translation and occlusion reasoning is used for recognition thereafter. Leibe et al. [43] have utilized iterative evidence integration for pedestrian recognition in occluded conditions. Another challenge in recognition is identifying individuals in a moving crowd. A coarse detection may be useful in such cases, attained through image processing under assumed features of a human body [44]. Heisele et al. [45] suggested that a moving pedestrian can be recognized by analyzing their motion of legs parallel to the image plane. Each image is segmented into several pixels having colour/ position information. This serves as a feature space for a classifier, which ultimately helps in pedestrian recognition. A functional and architectural breakdown of pedestrian detection system is presented in [46], which breaks down single classification problem into several individual classification stages. Situational features and stationary objects have also been focused upon in this work. Some researchers have also highlighted the importance of spatio-temporal cues in crowd recognition. Authors in [47] demonstrate a probabilistic grouping (based on space-time proximity and trajectory coherence) of image features into clusters which can recognize individual movements. Spatio-temporal cues can also be exploited for detection of direction of movement, as shown in [48]. Herein, a probability density function is calculated for each direction of motion.

4.2. Non-vision based crowd recognition

Among sensor based monitoring systems, WiFi has been more popular recently. Crowd recognition in these systems is based on fingerprinting. Channel state information (CSI) and received signal strength indicator (RSSI) can help in macroscopic level crowd recognition, whereas probe requests (PRs) help in both recognition of both, individuals and crowd as a whole. It is however important to mention that, crowd recognition in fingerprinting (WiFi and Bluetooth) based approach can be easily affected by the fact that not all of the crowd will have WiFi switched ON on their phones and certainly the majority will have Bluetooth switched OFF. Some notable research on channel state information (CSI) based recognition can be found in [18] and [49]. The CSI approach estimates the presence of crowd by quantifying its impact on the channel state. Though this approach has the merit of being non-intrusive, it does not assert the technical aptness for large scale public events, where the crowd density surges during peak hours. Authors in [50] have used RSSI based approach for occupancy estimation (recognizing crowd as a single entity) in indoor and outdoor environments, but the crowd size was very small. A more robust way of crowd count estimation is to utilize probe requests (PRs). PRs help a smartphone (or any WiFi enabled device) in searching nearby access points (APs). Authors in [19] and [20] have used this method to estimate pedestrian density /count. These works show that a PR contains device address and AP address. In PR based crowd counting, it is important to be aware of variable PR frequency in devices. As mentioned in [51] and [52], it can be tackled by having two or more sensors separated by up to 100 meters for efficient crowd counting. Though PRs can provide more accurate crowd count estimation, it could also lead to privacy concerns. Authors in [52], give a solution for preserving privacy in terms of SHA256 encrypted media access control (MAC) addresses. Weppner et al. [53] used 31 scanners to monitor a motor show for 13 days based on sensing of WiFi /Bluetooth enabled devices. The collected data was utilized only to create heat maps, which does not give a quantitative representation and recognition of the crowd. Besides, there is no discussion on the possibility of duplicate counts arising due to same MAC address being picked up by multiple sensors at the same time [54]. Some works also try to utilize crowd recognition for crowd prediction. Authors in [55] use WiFi signal detectors for crowd counting to estimate shopper volume in a store. Another work on indoor crowd queuing time estimation is presented in [56]. The authors used WiFi positioning data and statistical time series analysis to predict the queuing time. A recent and reliable approach for estimating and forecasting crowd counts using WiFi sensors in large-scale events is shown in [57]. Concerns like overestimation of crowd counts and privacy are duly addressed.

5. CROWD TRACKING

Crowd tracking may be comprehended as recognition of an individual through consecutive image frames or a continuous localization in time and space. Some possible use cases for crowd tracking could be individual tracking, surveillance, security, parameter breach detection, crowd learning pedestrian flows, and estimation of crowd velocity.

5.1. Vision based crowd tracking

Table 1. Comparative assessment of Tracking Systems [75]

Tracking Criteria	Computer vision	Global Positioning System	Radar/ Lidar	Sensor fusion	WiFi sensor Network
Indoor tracking	Yes	No	Yes	Yes	Yes
Large scale tracking	No	Yes	No	Yes	Yes
Trajectory accuracy	Low	High	Low	Low	High
Tracking latency	High	Low	Low	High	Low
System complexity	High	Low	Low	High	Low

Much like recognition, tracking also faces severe issues due to occlusion. Authors in [58] present two approaches to perceive occlusion, merge-split and straight-through. The former focuses on re-detection of an object post-occlusion, whereas the latter considers continuous tracking of subject at all times. Colour, edge, and geometry based features are very popular in crowd tracking, since they can be helpful in tackling occlusion related problem. Authors in [58] and [59] demonstrate that salient features of an image and the interest points obtained using a colour detector can be very helpful in tracking under occlusion. Zhao et al. [60], [61] developed 3D human models and Bayesian model for detection and tracking. Authors in [62] have used edgelet features for body part detection and a probabilistic data association for tracking. Some publications have shown that accounting for interpersonal interactions in a crowd model can enhance the tracking performance [63], [64]. A motion model based joint tracker has been developed in [64] which tracks target identity throughout an interaction to enhance tracking efficiency. Interactions may also take place, such that an individual departs from or enters in a crowd. A two-step solution is proposed in [65] to address this issue. The first step tracks active region using spatio-temporal details, whereas the second step involves recursive labelling of detected segments using Bayesian network based statistical model. Another solution is given in [66], where background subtraction removes shadows and unreliable cues; and colour information accounts for occlusions, depth, and position. Sullivan et al. [67] associated the targets with trajectories. Interaction of targets is signified by the end of a trajectory. This proposition is suitable for offline crowd analysis but requires huge storage for the resulting trajectories. Large crowds may require multi-camera setup with a multi-angle view for precise tracking. Mittal et al. [68] presented a multiple synchronized camera system named M2Tracker, for tracking multiple people in a cluttered scene. A combination of static and pan-tilt-zoom (PTZ) cameras is used in [69]. Special techniques have also been developed for multi-angle tracking, normally a planar homography constraint would be included. For example in [70], a contiguous spatio-temporal region is formed by feet positions belonging to the same person and is clustered as the movement track. In [71], ground points associated with a target are located and a multi-hypothesis framework using particle filter is developed for tracking. For proper understanding and modeling of pedestrian dynamics, reliable empirical data are necessary for analysis and verification. Authors in [72], have presented a software based strategy for time-efficient automatic extraction of accurate pedestrian trajectories, which is capable of tracking people on planar and uneven terrain without markers. Convolutional neural networks (CNN) have recently gained popularity in vision based crowd monitoring. Though there might be resolution issues in density maps generated by CNN, authors in [73] show that, tracking performance can be improved through a fusion of kernel correlation filters (KCF) and crowd density maps. A broader perspective on the degree of difficulty raised by a crowd dataset toward tracking tasks is presented in [74]. The proposed measures highlight in a compact way the motion as well as the appearance variability of the data, in the form of local and global entropy measures. These entropies can help in understanding crowd homogeneity and isotropy, which are essential indicators of behaviour.

5.2. Non-vision based crowd tracking

Recently, efforts have also been made to develop scalable systems for WiFi based detection and tracking [75]. In this study, the researchers have shown that unregistered smartphones with an active screen, send probe requests more frequently as compared to registered smartphones with or without active screen, leading to better detection rates. It is also shown that, detection and tracking is affected negatively by the speed of walking. A good description of WiFi based tracking systems is also presented in [52]. Wireless systems acquire information on a smartphone’s MAC using wireless sniffing and uses an RSSI based localization method for positioning. The purpose of this system is to monitor pedestrian traffic and monitor the density of people based on tracking smartphones in a street and explain how this information can be used to improve the service provided to people. Authors in [53] share three different approaches to tracking. First, where user’s mobile devices scan the environment for signals from stationary beacons such as WiFi access points or Bluetooth iBeacons. Second are systems where users mobile devices are used to detect the presence of other mobile devices. This approach has been widely used for the tracing social interactions. Third case has stationary scanners detecting, counting, and tracking mobile beacons carried by the users. Such mobile beacons can either be dedicated devices or the WiFi or Bluetooth interfaces of standard mobile devices. A comparative summary of the tracking systems is provided in Table 1. Note that the comparison is based on state-of-the-art detection and tracking system proposed in [75], tested in both indoor/ outdoor environments. Crowd trajectories are formulated as a sequence of IDs which correspond to time-synchronized wireless gateway nodes. The number of people moving along a particular trajectory is obtained by using longest common subsequence (LCS) algorithm.

6 . MODELING AND INFERENCE

Modeling is not a pre-requisite for crowd inference. Crowd dynamics can be inferred even without modeling, but it requires some prior knowledge about the crowd and event. Given the dynamic nature of crowd, modeling can be used for better interpretation of the changes in crowd behaviour at different instances. The modeling and inference techniques can be divided into two main categories, which are discussed hereafter.

6.1. Vision based models

In this category, the crowd models can be developed using the information extracted from processed visual data. Crowd models based on motion and foreground information were proposed in [76], [77], and [78]. These models involved two types of probability density functions, occurrence and orientation, which help in inferring the preferred crowd path. Andrade et al. in [79], [80], [81] used optimal flow and unsupervised feature extraction to model crowd motion patterns and recognize irregularities in crowd flow. The feature information can not only help in crowd modeling, but also help in analyzing crowd involvements. A complete review on application of image processing to crowd monitoring can be found in [82]. Boghossian et al. [83] have used vision based techniques for crowd path and direction estimation and thus improve crowd inference. A detection, tracking, and monitoring system for crowd and vehicles has been proposed in [84]. It utilizes optical flow based detection of pedestrians, crowd, and vehicles. Cupillard et al. in [85] and [86] present a model which uses multi-camera setup and is capable of tracking and detection of low level motion and behaviour related crowd scenarios. A method for detection of complex events is described in [87], which relies on constraints based on appearance, kinematics, and a hypothetic event model. Such methods usually rely on certain assumptions, which indicates that apriori knowledge of event conditions is necessary for crowd inference. These techniques can be fast, efficient, and economic, because the inference procedure is relatively simpler.

6.2. Non-vision based models

Crowd models focus on prediction and apprehension of crowd behaviour by identifying feature correlations. These models can be useful in quantitative analysis of crowd dynamics. Ref. [88] classifies crowd models in three different types: (i) microscopic- where pedestrians are treated as discrete individuals, (ii) macroscopic- where crowd is considered as a single entity, (iii) mesoscopic- where characteristics of pedestrians and crowd are collectively considered. This section highlights some unique non-vision based crowd models inspired from various sources [16]. The first set of models are inspired from physics. Helbing [89] gave different formulations analogous to each type of crowd model: stochastic (microscopic), gas kinetic (mesoscopic), and fluid dynamic (macroscopic). A social field theory based model was also given in [90], which explores the factors for crowd behaviour. This model was later used to analyze collective patterns of motion in [91]. Authors in [92] have explored the factors behind uneven information propagation in a crowd. Another macroscopic model studying the goal-oriented behaviour of pedestrians is presented in [93]. Physics inspired crowd models have wide possibilities of analyzing crowd behaviour from different perspectives, e.g. effect of robots on crowd [94] and modeling of a historical event [95]. Forecasting crowd densities is also a macroscopic task which models crowd count as a time series. Authors in [55] use WiFi signal detectors for crowd count estimation and an autoregressive integrated moving average (ARIMA) model for forecasting. Another work on indoor crowd queuing time estimation is presented in [56]. The authors used WiFi positioning data and nonstandard autoregressive (NAR) model to predict the queuing time. The next set of models are agent based, where fuzzy methods are used to describe crowd motion and associated factors. For example, emotional parameters have been explored in [96] to analyze the behaviour of an individual in a goal-oriented crowd. Pan et al. [97] have modelled an agent to analyze non-adaptive or destructive behaviour that a crowd may face in adverse conditions. Authors in [98] develop a model assuming that collocated individuals undergo similar psychological and environmental influences. Commercial version of some agent based models are also available [99], [100] which simulate the crowd behaviour in different environments. Another variant of crowd model is based on cell theory, where a crowded area is perceived as a collection of cells. Each cell represents the minimum occupancy area for an individual. The movement of individuals within the cells is guided by a set of rules. A commercially available product based on this model is EGRESS [101]. A hybrid model using the concept of floor field and agents in presented in [102]. In this model a static floor field contains exit distance information, a dynamic floor field is associated with crowd motion, and the last field keeps track of distance from cell to adjacent wall. The last category of non-vision model is inspired from nature. In these models, the cognitive and transitive crowd behaviours are perceived through auxiliary models which closely resemble a human crowd. For example, the emotional ant model in [103] studies the psychological behaviour using biologically inspired ant agents as crowd. Authors in [104] have studied the interaction between the pedestrians using the concept of chemotaxis. The simulation of the evacuation is also presented to show the ability of the model to represent different types of crowd behaviours. A detailed review about crowd motion simulation models in presented in [105] which might be a useful reference for readers. A general guidance for developing computer based building evacuation models can be found in [106].

7. ADDITIONAL INFORMATION

This section provides some additional information related to crowd monitoring, which can serve as a quick reference for the readers.

- **Learning Algorithms:** Intelligent learning algorithms are being used for counting and density estimation since the inception of crowd monitoring. Such algorithms involve training of a classifier on features extracted from an image. During crowd counting or density estimation, test images are classified using the trained learner. The ratio of correctly classified density labels (e.g. low, medium, high) to the total number of tested images gives the classification accuracy. A neural network has been trained on edge and colour features for crowd density estimation in [33]. The authors used a hybrid least-square and global learning algorithm which does not get trapped in a local minima and offers quick convergence. A better performance is claimed in [12], where support vector machines trained on wavelet based statistical features are used for crowd estimation. The classifier was designed using three different kernels: polynomial, Gaussian radial basis function (RBF), and a multilayer perceptron. Chan et al. [26] used a Gaussian regression to estimate the number of people per segment. To account for the linearity and non-linearity in extracted features, the authors combined a linear and a squared exponential RBF kernel. Most of the techniques either use a single regressor for global crowd counting or use multiple regressors for localized crowd counting, which requires extensive training. A single multi-output regression model utilizing global and local features for crowd counting is presented in [27], as a solution to the above problem. A k-nearest neighbour based crowd classification approach is presented in [107]. In case of crowd counting the learners are trained on number of people instead of density labels. Alternatively, crowd count can also be achieved by formulating a time series forecasting problem. In this approach an artificial intelligence technique such as neural networks can be employed to learn crowd counts at specific time intervals; and then crowd counts for future time values can be predicted for a particular time horizon. A recent work on WiFi and deep learning based crowd forecasting is available in [57], which has been used for crowd count forecasting at a large scale public event. In fact, there is much more focus use of deep learning techniques in crowd monitoring as they provide better learning and prediction accuracy. While preparing a crowd forecasting model, it is important to consider the influence of factors such as traffic and weather on the crowd behaviour. Authors in [108] have developed a Support Vector Regression and Adaptive Particle Swarm Optimization (APSO-SVR) based forecasting algorithm for predicting pedestrian flows in Tiananmen Square. Crowd influencing factors like holidays, rain, wind, temperature, relative humidity, and air quality index have been considered in problem formulation.

- **Devices:** A wide range of devices have been utilized in crowd monitoring, with vision based devices being the most popular ones. The monitoring approach and underlying mathematical models depend on the choiceof device. Therefore, it is necessary to be aware of the device capabilities and type of crowd data acquired with a particular device. Static CCTV camera can provide visual information of a location over a range of time [33]. Multi-view [68] and PTZ cameras [69] can provide multi-angular visual information for better recognition and tracking. Bluetooth [9] and WiFi [18] sensors can help in estimation of crowd count via probe requests. Cell-phones [7] can provide crowd count or even location information based on the granted permissions. GPS [8, 9] can help in inferring position and velocity information of individuals in crowd, which can eventually help in crowd count/ density estimation. A high-level comparison between the capabilities of some monitoring devices is shown in Table 2 [109, 110]. The values specified for accuracy and coverage are given in form of intervals wherein most approaches reside. The coverage is to be regarded as the direct measuring range of an unextended implementation, i.e. the spatial scalability which many system approaches offer has not been taken into account (eg. deployment of additional sensor nodes). Thus, there are many exceptions exceeding these intervals. Similarly, only the main measuring principles are mentioned in the table.

- **Datasets:** Open access datasets can be very useful in training and testing of crowd monitoring techniques. Almost all the available datasets in crowd monitoring are based on vision approach of crowd estimation. It is notable that despite the rise in sensor based crowd monitoring, there is almost no open access datasets available. The researchers working in this domain must come forward and contribute more to sensor based open access datasets on crowd monitoring. It is important to mention the role of synthetic/ simulated datasets, which can not only compensate for lack of publicly available datasets and annotated data, but also provides diverse testing environments. Table 3 gives

Table 2. Comparison between some commonly used monitoring devices [109, 110]

Device/ Technology	Typical accuracy	Typical coverage [m]	Measurement principle	Use cases in monitoring
Cameras	0.1mm – dm	1 – 10	Image metrology	Facial recognition, counting, detection, tracking, trajectory estimation, surveillance, behaviour analysis.
Infrared	cm – m	1 – 5	Thermal imaging, active beacons	Detection, tracking, counting
Tactile/ Polar Systems	µm – mm	3 – 2000	Mechanical, interferometry	Automotive, metrology
Sound	cm	2 – 10	Echo reflection	Speech recognition, tracking, hospitals
WLAN/ WiFi	m	20 – 50	Fingerprinting	Counting, detection, trajectory tracking, location based services
Bluetooth	m	1-10	Fingerprinting	Counting, detection
RFID	dm – m	1 – 50	Proximity, fingerprinting	Detection, location finding, contact discovery
Ultra-Wideband	cm – m	1 – 50	Reflection, time of arrival	Robotics, automation
Global Navigation Satellite System (GNSS)	10 m	Global	Parallel correlation, assistant GPS	Location based services, navigation, tracking
Pseudolites	cm – dm	10 – 1000	Carrier phase ranging	Positioning, navigation
Other radio frequencies	m	10 – 1000	Proximity, fingerprinting	Detection, tracking
Inertial navigation	1%	10 – 100	Dead reckoning	Pedestrian navigation
Magnetic systems	mm – cm	1 – 20	Fingerprinting, ranging	Detection, positioning, healthcare
Infrastructure systems	cm – m	Building	Fingerprinting, capacitance	Detection, assisted living

detailed information on crowd monitoring datasets. The authors have made best efforts to include all the publicly available datasets to serve as a single-point reference for researchers in this domain. A lot more work has been done in sensor based crowd monitoring, but the works could not be included in the table due to inaccessibility of datasets. The publications related to datasets can be found via the referenced links.

- **Privacy:** As acknowledged in the introduction, privacy is a big concern for this field. The focus on privacy in existing research, has already been highlighted to some extent in Table 3. It is evident that not much effort has been dedicated towards privacy-preserved and non-intrusive crowd monitoring. With the implementation of recent and more definitive guidelines like those in General Data Protection Regulation (GDPR) in the European Union and Data Protection Act 2018 in United Kingdom; it is expected that the emerging crowd monitoring techniques will be obliged to be privacy preserving. While this means that the existing monitoring techniques might become obsolete, it will also accelerate the research in crowd monitoring. Different monitoring techniques have different data requirements and thus different privacy concerns. This leads to the issues related to public consent, when being monitored. Vision based monitoring captures videos and images of individuals. Specially with the PTZ cameras [69] and facial recognition technology, it might compromise individual identities. However, this type of intensive surveillance is generally needed by security agencies, which will need prior approval from the Government. An example of this can be found in recent report on trial of live facial recognition technology by London Metropolitan Police [126], where data collection is only approved for safety and security purpose. From perspective of research ethics, someone avoiding the cameras is an indication that they are exercising their entitlement not to be a part of particular trial or are protecting their own right to privacy. From a policing perspective, this same behaviour may acquire a different meaning and serve as an indicator of suspicion. In case of WiFi and Bluetooth based monitoring, the device MAC addresses are the point of privacy concern [9, 52, 54]. From the research perspective, it can be ensured that individual identities are protected through data encryption before data storage. Obtaining a certification for compliance with local privacy laws, is also advisable. In any case, prior notice should be given related to system deployment. In case of a mobile app based tracking/ monitoring, user should be well informed and permissions should be taken for data logging. Other lesser used but relevant techniques like infrared, sound, and RFID based monitoring systems [109], also pose similar concerns, and should adhere to the recommended practices.

8. FUTURE DIRECTIONS

Various aspects of crowd monitoring were discussed in this paper. The most notable point is the exceptional research on vision based crowd monitoring techniques. Important observations and suggestions with respect to future research is as follows:

Table 3. Description of Crowd Monitoring Datasets

Dataset	Ref.	Synthesis	Privacy emphasis	Description	Use case tags
UCSD/ PETS2009	[111]	Camera	Yes	Grayscale videos	Modeling, clustering, segmentation
Mall	[112]	Public webcam	No	Large scale, video frames	Crowd counting, profiling, crowd density, feature mining
BEHAVE	[113]	Camera, simulations	No	Real and simulated video frames	Crowd behaviour classification, anomaly detection
CBE	[114]	Camera, simulations	No	Simulations based on real video segments	Crowd simulation, crowd behaviour, tracking, trajectory
CUHK-SYSU	[115]	Camera	No	Pedestrian images	Person search, crowd counting
CUHK Person Re-identification	[115]	Camera	No	Pedestrian images	Person re-identification, crowd counting
CUHK Occlusion	[115]	Camera	No	Images from 6 different cities/ research groups	Pedestrian detection, mobile scene analysis, feature extraction, classification
Immediacy	[115]	Camera	No	Large scale, interaction images	Immediacy prediction, pose/ interaction/ posture analysis
CUHK Square	[115]	Camera	No	1 hour video sequence	Pedestrian detection, traffic scene analysis, surveillance, classification
Shanghai World Expo'10	[115]	Camera	No	Large scale, video sequences/ images	Crowd counting, crowd behaviour analysis, feature extraction, classification
WWW Crowd Attribute	[115]	Camera, Assorted	No	Large scale, 10000 videos, 8257 scenes, 94 crowd-related attributes	Crowd scene analysis, attribute learning (who, where, why)
MIT Traffic	[115]	Camera	No	1.5 hour video, 20 clips	Crowd detection, activity perception, traffic analysis
Train Station	[115]	Camera	No	Video, trajectories	Crowd behaviour classification, velocity & position prediction, behaviour & trajectory modeling
MIT Trajectory	[115]	Single/ multi-cam view	No	Videos, activity/ motion simulations	Crowd trajectory, activity modeling/ learning, anomaly detection
CelebA	[115]	Web, Assorted	No	Celebrity images	Face attribute recognition, face detection,
CUHK Face Sketch	[115]	Assorted	No	Facial images from existing databases	Face recognition, facial feature analysis, face sketch recognition
CityStreet	[116]	Multi-view camera, assorted	No	Wide area city street scenes, collection of 3 datasets	Crowd counting, crowd density mapping
CityUHK-X	[116]	Camera	No	Images with extrinsic camera parameters	Crowd counting, image deconvolution, crowd density map based learning
Line Counting	[116]	Camera	No	Collection of UCSD, Grand Central and LHI datasets	Line counting, crowd velocity estimation, segmentation, parameter breach detection
MADS	[116]	Multi-view camera	No	Multi-view human pose videos/ images	Activity recognition, discriminative tracking, 3D human modeling
Highway Traffic	[116]	Camera	No	Traffic videos	Time series based video clustering, crowd scene segmentation
YACVID	[117]	Camera, sensor-fusion	Only in few cases	Huge repository of global videos/ images	Crowd detection/ classification, tracking, surveillance, behaviour analysis, segmentation, crowd attributes.
TRANCOS	[118]	Camera	No	Traffic images	Object detection, vehicle counting
Daimler	[119]	Stereo camera	No	Pedestrian videos/ images	Crowd detection, segmentation, path prediction, classification
WiAR	[120]	WiFi (cards)	Yes	RSSI & CSI signatures	Activity attributes and recognition
Winter Wonders 2018-19	[121]	WiFi (sensor network)	Yes	RSSI based counts	Crowd counting, time series based forecasting
ShanghaiTech	[122]	Camera, Assorted	Yes	Images from Internet & Shanghai streets	Foreground segmentation, density map, crowd counting
UCF-QNRF	[123]	Web, Assorted	No	High resolution, dense crowd images	Crowd counting, density map estimation, localization
UMN-MHA	[124]	Camera	No	Videos for activity monitoring	Event detection, real-time tracking, action recognition, pattern learning, behavioral analysis
GCC	[125]	Simulations	Yes	Large scale, simulated crowd images	Synthetic data, pre-trained crowd counting models, diverse environments

- In order to address the privacy laws, focus needs to be shifted to less-intrusive and privacy preserving crowd monitoring techniques. Both, vision and non-vision techniques can compromise the privacy of an individual, either by facial recognition or by tracking device signatures. While security-oriented technologies might have permissions to identify and track individuals, general crowd monitoring techniques need to focus more on intrusion and privacy. Automatic facial blurring in images and SHA encryption for device addresses should be used before uploading data on the server. These techniques have been already used by some researchers.

- Sensor fusion has a great potential, where the synergy of different technologies can be used to develop an advanced and robust crowd monitoring system. For example, the pros and cons of WiFi and cellular networks (cost, coverage, and data quality) are complimentary to each other. Therefore, the use of sensor fusion and multi-fidelity networks must be considered in developing future monitoring systems. This will ensure that any shortcoming of one system is fulfilled by the other.

- Crowd monitoring is an important research area and in order to develop effective technologies, it is necessary to test the newly proposed techniques in different crowd scenarios. However, the possibility of getting real life scenarios for testing is either limited or biased with assumptions. Emulation software like Autodesk 3ds Max, can be really useful to evaluate the performance of developed techniques under such conditions.

- Applications of artificial intelligence (AI) to crowd monitoring has mainly witnessed the usage of supervised learning algorithms. The performance of these algorithms heavily depends on the type of features used and the amount of training. Therefore, the consistency of performance in such techniques is not guaranteed. Fast paced advancements in deep learning, multiagent systems, knowledge representation and explainable AI must be actively utilized in anomaly detection, activity recognition and behaviour analysis.

- There is a need to focus on the forecasting aspect in addition to crowd counting and classification. This can help in predicting critical crowd scenarios and possibly (with the implementation of decision support systems) alert the security agencies to take precautionary actions in due time. Crowd forecasting can also help in making timely arrangements at large-scale public events.

- Despite the huge potential of learning-based intelligence, there are still some major challenges to their application in crowd monitoring, such as: poor data quality, insufficient data and imbalanced training data. For example, data with missing values and outliers is likely to cause incorrect and unreliable decision making or analysis. Therefore, research must also be focussed on data diagnosis and synthesis.

- Lastly, there is a need to encourage the availability of datasets, especially the ones corresponding to non-vision based crowd monitoring techniques so that research can be accelerated further.

ACKNOWLEDGMENTS

The authors thank INNOVIRIS (The Brussels Institute for Research and Innovation) for funding this research through the MUFINS (Multi-Fidelity Intelligent Network System for Crowd Monitoring) project.

REFERENCES

[1] Cardone, Giuseppe, et al. "Crowdsensing in urban areas for city-scale mass gathering management: Geofencing and activity recognition." IEEE Sensors Journal 14.12 (2014): 4185-4195.

[2] J. C. S. Jacques, Jr., 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.

[3] S. Gong, C. C. Loy, and T. Xiang, “Security and surveillance,” in Visual Analysis of Humans, T. B. Moeslund, A. Hilton, V. Krüger, and L. Sigal, Eds. London, U.K.: Springer-Verlag, 2011, pp. 455–472.

[4] Corbetta, A., Menkovski, V., Benzi, R. and Toschi, F., 2019. Deep learning velocity signals allows to quantify turbulence intensity. arXiv preprint arXiv:1911.05718.

[5] D. Helbing, A. Johansson, and H. Z. Al-Abideen, “Dynamics of crowd disasters: An empirical study,” Phys. Rev. E, vol. 75, no. 4, pp. 46109–46116, 2007.

[6] Z. Fang, S. M. Lo, and J. A. Lu, “On the relationship between crowd density and movement velocity,” Fire Safety J., vol. 38, no. 3, pp. 271—283, 2003.

[7] F. Calabrese, M. Colonna, P. Lovisolo, D. Parata, and C. Ratti, “Real-time urban monitoring using cell phones: A case study in Rome,” IEEE Intell. Transp. Syst. Mag., vol. 12, no. 1, pp. 141–151, Mar. 2011.

[8] N. Koshak and A. Fouda, “Analyzing pedestrian movement in mataf using GPS and GIS to support space redesign,” in Proc. DDSS, pp. 1–14, 2008.

[9] M. Versichele, T. Neutens, M. Delafontaine, and N. Van de Weghe, “The use of bluetooth for analysing spatiotemporal dynamics of human movement at mass events: A case study of the Ghent festivities,” Appl. Geograph., vol. 32, no. 2, pp. 208–220, 2012.

[10] M. Wirz, T. Franke, D. Roggen, E. Mitleton-Kelly, P. Lukowicz, and G. Troster, “Inferring crowd conditions from pedestrians’ location traces for real-time crowd monitoring during city-scale mass gatherings,” in Proc. IEEE WETICE, Jun. 2012, pp. 367–372.

[11] De Moraes, Rodrigo Saar, and Edison Pignaton De Freitas. "Multi-UAV Based Crowd Monitoring System." IEEE Transactions on Aerospace and Electronic Systems (2019).

[12] Xiaohua, L., Lansun, S. and Huanqin, L., 2006. Estimation of crowd density based on wavelet and support vector machine. Transactions of the Institute of Measurement and Control, 28(3), pp.299-308.

[13] Lin, S., Chen, J., Chao, H.: Estimation of number of people in crowded scenes using perspective transformation. IEEE Trans. Syst. Man Cybernet. A, 31(6), 645–654 (2001)

[14] Marana, A., da Costa, L., Lotufo, R., Velastin, S.: On the efficacy of texture analysis for crowd monitoring. In: Proceedings of the International Symposium on Computer Graphics, Image Processing, and Vision, vol. 00, p. 354 (1998)

[15] Velastin, S., Yin, J., Davies, A., Vicencio-Silva, M., Allsop, R., Penn, A.: Automated measurement of crowd density and motion using imageprocessing. Road traffic monitoring and control, 1994. In: Seventh International Conference, pp. 127–132 (1994)

[16] Zhan, Beibei, et al. "Crowd analysis: a survey." Machine Vision and Applications 19.5-6 (2008): 345-357.

[17] Marana, A., Velastin, S., Costa, L., Lotufo, R.: Automatic estimation of crowd density using texture. Safety Sci. 28(3), 165–175 (1998)

[18] Xi W, Zhao J, Li XY, Zhao K, Tang S, Liu X, Jiang Z. Electronic frog eye: Counting crowd using wifi. InInfocom, 2014 proceedings IEEE 2014 Apr 27 (pp. 361-369). IEEE.

[19] A. Kurkcu and K. Ozbay, “Estimating pedestrian densities, wait times, and flows with WiFi and bluetooth sensors,” Transportation Research Record, vol. 2644, no. 1, pp. 72–82, 2017.

[20] A. Guillén-Pérez and M. D. C. Baños, “A wifi-based method to count and locate pedestrians in urban traffic scenarios,” in 2018 14th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob). IEEE, 2018, pp. 123–130.

[21] L. Ascorti, S. Savazzi, G. Soatti, M. Nicoli, E. Sisinni, and S. Galimberti, “A wireless cloud network platform for industrial process automation: Critical data publishing and distributed sensing,” IEEE Transactions on Instrumentation and Measurement, vol. 66, no. 4, pp. 592-603, 2017.

[22] P. Giri, K. Ng, and W. Phillips, “Wireless Sensor Network System for Landslide Monitoring and Warning,” IEEE Transactions on Instrumentation and Measurement, vol. 68, no. 4, pp.1210-1220, 2018.

[23] L. Lombardo, S. Corbellini, M. Parvis, A. Elsayed, E. Angelini, and S. Grassini, “Wireless sensor network for distributed environmental monitoring,” IEEE Transactions on Instrumentation and Measurement, vol. 67, no. 5, pp.1214-1222, 2017.

[24] Haghani, M. and Sarvi, M., 2018. Crowd behaviour and motion: Empirical methods. Transportation research part B: methodological, 107, pp.253-294.

[25] Banerjee, A., Maurya, A.K. and Lämmel, G., 2018. Pedestrian flow characteristics and level of service on dissimilar facilities: A critical review. *Collective Dynamics*.

[26] Chan AB, Liang ZS, Vasconcelos N. Privacy preserving crowd monitoring: Counting people without people models or tracking. InComputer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on 2008 Jun 23 (pp. 1-7). IEEE.

[27] Chen, Ke, Chen Change Loy, Shaogang Gong, and Tony Xiang. "Feature mining for localised crowd counting." In BMVC, vol. 1, no. 2, p. 3. 2012.

[28] Polus A, Schofer JL, Ushpiz A. Pedestrian flow and level of service. Journal of transportation engineering. 1983 Jan;109(1):46-56.

[29] Yin, J., Velastin, S., Davies, A.: Image Processing Techniques for Crowd Density Estimation Using a Reference Image. Proc. 2nd Asia-Pacific Conf. Comput. Vis. 3, 6–10 (1995)

[30] Ma, R., Li, L., Huang, W., Tian, Q.: On pixel count based crowd density estimation for visual surveillance. IEEE Conf. Cybernet. Intell. Syst. 1 (2004)

[31] Dong, L., Parameswaran, V., Ramesh, V., Zoghlami., I.: Fast Crowd Segmentation Using Shape Indexing. Rio de Janeiro, Brazil (2007)

[32] Marana, A., Velastin, S., Costa, L., Lotufo, R.: Estimation of crowd density using image processing. In: IEE Colloquium on Image Processing for Security Applications (Digest No: 1997/074), p. 11 (1997)

[33] Cho SY, Chow TW, Leung CT. A neural-based crowd estimation by hybrid global learning algorithm. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics). 1999 Aug;29(4):535-41.

[34] Boltes, M., Zhang, J., Tordeux, A., Schadschneider, A. and Seyfried, A., 2018. Empirical results of pedestrian and evacuation dynamics. *Encyclopedia of Complexity and Systems Science. Springer, Berlin, Heidelberg*.

[35] S. Gwynne and K. E. Boyce, “Engineering Data,” in SFPE handbook of fire protection engineering, M. J. Hurley, Ed., New York: Springer, 2016, pp. 2429–2551.

[36] Swets, D., Punch, B.: Genetic algorithms for object localization in a complex scene. In: IEEE International Conference on Image Processing, pp. 595–598 (1995)

[37] Li, S.Z., Zhu, L., Zhang, Z., Blake, A., Zhang, H., Shum, H.: Statistical learning of multi-view face detection. In: Proceedings of the 7th European Conference on Computer Vision-Part IV, pp. 67–81. Springer, London (2002)

[38] Jones, M., Viola, P.: Fast multi-view face detection. Mitsubishi Electric Research Lab TR-20003-96 (2003)

[39] Huang, C., Ai, H., Li, Y., Lao, S.: Vector boosting for rotation invariant multi-view face detection. In: Tenth IEEE International Conference on Computer Vision, vol. 1, pp. 446–453 (2005)

[40] Huang X., Li L., Sim T. (2004) Stereo-based human head detection from crowd scenes. In: International Conference on Image Processing, 2004. ICIP’04, Vol. 2, pp. 1353–1356

[41] Wu, B., Nevatia, R.: Detection of Multiple, Partially Occluded Humans in a Single Image by Bayesian Combination of Edgelet Part Detectors. In: Tenth IEEE International Conference on Computer Vision, 2005. ICCV 2005, vol. 1, 90–97 (2005)

[42] Elgammal, A., Davis, L.: Probabilistic framework for segmenting people under occlusion. In: Eighth IEEE International Conference on Computer Vision, 2001. Proceedings of ICCV 2001, vol. 2, pp. 145–152 (2001)

[43] Leibe, B., Seemann, E., Schiele, B.: Pedestrian detection in crowded scenes. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recogn. CVPR 2005. 1 (2005)

[44] Broggi, A., Bertozzi, M., Fascioli, A., Sechi, M.: Shape-based pedestrian detection. In: Proceedings of the IEEE Intelligent Vehicles Symposium 2000. Dearbon (MI), USA (2000)

[45] Heisele B., Woehler C. (1998) Motion-based recognition of pedestrians. In: Proceedings of Fourteenth International Conference on Pattern Recognition, 1998, vol. 2, pp. 1325–1330 (1998)

[46] Shashua, A., Gdalyahu, Y., Hayun, G.: Pedestrian detection for driving assistance systems: single-frame classification and system level performance. Intelligent Vehicles Symposium, 2004 IEEE pp. 1–6 (2004)

[47] Brostow, G., Cipolla, R.: Unsupervised Bayesian detection of independent motion in crowds. In: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 594–601. IEEE Computer Society Washington, DC, USA (2006)

[48] Reisman, P., Mano, O., Avidan, S., Shashua, A., Ltd, M., Jerusalem, I.: Crowd detection in video sequences. Intell. Vehicles Symp. IEEE, 66–71 (2004)

[49] Liu, S., Zhao, Y., Xue, F., Chen, B. and Chen, X., 2019. DeepCount: Crowd counting with WiFi via deep learning. *arXiv preprint arXiv:1903.05316*.

[50] Depatla, S., Muralidharan, A. and Mostofi, Y., 2015. Occupancy estimation using only WiFi power measurements. *IEEE Journal on Selected Areas in Communications*, *33*(7), pp.1381-1393.

[51] C. Groba, "Demonstrations and people-counting based on Wifi probe requests," In 2019 IEEE 5th World Forum on Internet of Things (WF-IoT). IEEE, 2019, pp. 596-600.

[52] M. Uras, R. Cossu, and L, Atzori, "PmA: a solution for people mobility monitoring and analysis based on WiFi probes," In 2019 4th International Conference on Smart and Sustainable Technologies (SpliTech). IEEE, 2019, pp. 1-6.

[53] J. Weppner, B. Bischke, and P. Lukowicz, "Monitoring crowd condition in public spaces by tracking mobile consumer devices with wifi interface," in Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct. ACM, 2016, pp. 1363-1371.

[54] M. Vanhoef, C. Matte, M. Cunche, L. S. Cardoso, and F. Piessens, "Why MAC address randomization is not enough: An analysis of Wi-Fi network discovery mechanisms," in Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security. ACM, 2016, pp. 413–424.

[55] I. K. Tan, O. B. Yaik, and O. B. Sheng, "Predicting shopper volume using ARIMA on public Wi-Fi signals," International Information Institute (Tokyo). Information, vol. 19, no. 8A, p. 3295, 2016.

[56] H. Shu, C. Song, T. Pei, L. Xu, Y. Ou, L. Zhang, and T. Li, "Queuing time prediction using WiFi positioning data in an indoor scenario," Sensors, vol. 16, no. 11, p. 1958, 2016.

[57] Singh U, Determe JF, Horlin F, De Doncker P. Crowd Forecasting based on WiFi Sensors and LSTM Neural Networks. IEEE Transactions on Instrumentation and Measurement. 2020 Jan 27.

[58] Gabriel, P., Verly, J., Piater, J., Genon, A.: The state of the art in multiple object tracking under occlusion in video sequences. Advanced Concepts for Intell. Vis. Syst., pp. 166–173 (2003)

[59] Mathes, T., Piater, J.: Robust non-rigid object tracking using point distribution models. Bri. Mach. Vis. Conf. 2 (2005)

[60] Zhao, T., Nevatia, R.: Tracking multiple humans in complex situations. Pattern Anal. Mach. Intell. IEEE Trans. 26(9), 1208– 1221 (2004)

[61] Zhao, T., Nevatia, R.: Tracking multiple humans in crowded environment. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition vol. 2, II–406–II–413 (2004)

[62] Wu, B., Nevatia, R.: Tracking of multiple, partially occluded humans based on static body part detection. In: CVPR '06: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 951–958 (2006)

[63] Smith, K., Gatica-Perez, D., Odobez, J.: Using particles to track varying numbers of interacting people. In: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol. 1–01, pp. 962–969 (2005)

[64] Khan, Z., Balch, T., Dellaert, F.: MCMC-based particle filtering for tracking a variable number of interacting targets. IEEE Trans. Pattern Anal. Mach. Intell. 27(11), 1805–1819 (2005)

[65] Marques, J., Jorge, P., Abrantes, A., Lemos, J.: Tracking groups of Pedestrians in video sequences. IEEE 2003 Conf. Comput. Vis. Pattern Recogn. Workshop, 9, 101 (2003)

[66] McKenna, S., Jabri, S., Duric, Z., Rosenfeld, A., Wechsler, H.: Tracking groups of people. Comput. Vis. Image Understanding 80(1), 42–56 (2000)

[67] Sullivan, J., Carlsson, S.: Tracking and labelling of interacting multiple targets. In: European Conference on Computer Vision, LNCS, vol. 3953, pp. 619–632. Springer, Heidelberg (2006)

[68] Mittal, A., Davis, L.: M2 tracker: a multi-view approach to segmenting and tracking people in a cluttered scene. Int. J. Comput. Vis. 51(3), 189–203 (2003)

[69] Chang, T., Gong, S., Ong, E.: Tracking multiple people under occlusion using multiple cameras. In: British Machine Vision Conference, pp. 566–575 (2000)

[70] Khan, S.M., Shah, M.: A multiview approach to tracking people in crowded scenes using a planar homography constraint. In: 9th European Conference on Computer Vision, LNCS, vol. 3954, pp. 133–146. Springer, Heidelberg (2006)

[71] Kim, K., Davis, L.S.: Multi-camera tracking and segmentation of occluded people on ground plane using search-guided particle filtering. In: European Conference on Computer Vision, LNCS, vol. 3953, pp. 98–109. Springer, Heidelberg (2006)

[72] Boltes, M. and Seyfried, A., 2013. Collecting pedestrian trajectories. *Neurocomputing*, *100*, pp.127-133.

[73] Kang, D., Ma, Z. and Chan, A.B., 2018. Beyond Counting: Comparisons of Density Maps for Crowd Analysis Tasks—Counting, Detection, and Tracking. *IEEE Transactions on Circuits and Systems for Video Technology*, *29*(5), pp.1408-1422.

[74] Franchi, G., Aldea, E., Dubuisson, S. and Bloch, I., 2019, September. Crowd Behavior Characterization for Scene Tracking. In *2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)* (pp. 1-8). IEEE.

[75] Li, K., Yuen, C., Kanhere, S.S., Hu, K., Zhang, W., Jiang, F. and Liu, X., 2018. An experimental study for tracking crowd in smart cities. *IEEE Systems Journal*, *13*(3), pp.2966-2977.

[76] Zhan, B., Remagnino, P., Velastin, S.: Analysing Crowd Intelligence. Second AIxIA Workshop on Ambient Intelligence (2005)

[77] Zhan, B., Remagnino, P., Velastin, S.: Mining paths of complex crowd scenes. Lecture Notes in Computer Science pp. 126–133 (2005).

[78] Zhan, B., Remagnino, P., Velastin, S.: Visual analysis of crowded pedestrain scenes. XLIII Congresso Annuale AICA, pp. 549–555 (2005)

[79] Andrade, E., Fisher, R.: Hidden Markov models for optical flow analysis in crowds. In: Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06), vol. 01, pp. 460–463. IEEE Computer Society Washington, (2006)

[80] Andrade, E., Fisher, R.: Modelling crowd scenes for event detection. In: Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06), vol. 01, pp. 175–178. IEEE Computer Society Washington, DC (2006)

[81] Andrade, E.L., Blunsden, S., Fisher, R.B.: Performance analysis of event detection models in crowded scenes. In: Proceedings of Workshop on Towards Robust Visual Surveillance Techniques and Systems at Visual Information Engineering 2006, pp. 427– 432. Bangalore, India (2006)

[82] Davies, A., Yin, J., Velastin, S.: Crowd monitoring using image processing. Electron. Commun. Eng. J. 7(1), 37–47 (1995)

[83] Boghossian, B., Velastin, S.: Motion-based machine vision techniques for the management of large crowds. In: The 6th IEEE International Conference on Electronics, Circuits and Systems, vol. 2 (1999)

[84] Maurin, B., Masoud, O., Papanikolopoulos, N.: Monitoring crowded traffic scenes. In: Proceedings of The IEEE 5th International Conference on Intelligent Transportation Systems, 2002. pp. 19–24 (2002)

[85] Cupillard, F., Bremond, F., Thonnat, M.: Behaviour recognition for individuals, groups of people and crowd. IEE Seminar Digests 7 (2003)

[86] Cupillard, F., Bremond, F., Thonnat, M., INRIA, F.: Group behavior recognition with multiple cameras. Applications of Computer Vision, 2002 (WACV 2002). In: Proceedings of Sixth IEEE Workshop, pp. 177–183 (2002)

[87] Chan, M.T., Hoogs, A., Bhotika, R., Perera, A., Schmiederer, J., Doretto, G.: Joint recognition of complex events and track matching. In: CVPR '06: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1615–1622. IEEE Computer Society, Washington, DC, USA (2006)

[88] FHWA.: Traffic analysis tools primer,traffic analysis toolbox (1) (2004)

[89] Helbing, D.: Models for pedestrian behavior (1992).

[90] Helbing, D., Molnár, P.: Social force model for pedestrian dynamics. Phys. Rev. E 51(5), 4282–4286 (1995)

[91] Helbing, D., Molnar, P.: Self-organization phenomena in pedestrian crowds (1997).

[92] Chu, J., Li, J., Xu, M., Zhao, L.: Simulating escape panic based on the mechanism of asymmetric information distribution. In: Complex Systems Summer School Final Project Papers. Santa Fe Institute, Santa Fe (2005).

[93] Hughes, R.: A continuum theory for the flow of pedestrians. Trans. Res. B Methodol. 36(6), 507–535 (2002)

[94] Kirkland, J., Maciejewski, A.: A simulation of attempts to influence crowd dynamics. IEEE Int. Conf. Syst. Man Cybernet. 4328– 4333 (2003)

[95] R.R., C., Hughes, R.: Mathematical modelling of a mediaeval battle: the battle of agincourt. Math. Compute. Simul. 64(2), 259–269 (2004)

[96] Musse, S., Thalmann, D.: A model of human crowd behavior: group inter-relationship and collision detection analysis. Proc. Workshop Comput. Anim. Simul. Eurograph. 97, 39–51 (1997)

[97] Pan, X., Han, C., Dauber, K., Law, K.: Human and social behavior in computational modeling and analysis of egress. Automation in Construction 15(4), 448–461 (2006)

[98] Brenner, M., Wijermans, N., Nussle, T., de Boer, B.: Simulating and controlling civilian crowds in robocup rescue. In: Proceedings of RoboCup 2005: Robot Soccer World Cup IX, Osaka (2005)

[99] Crowd, Dynamics: http://www.crowddynamics.com/

[100] Legion: http://www.legion.biz/about/index.html

[101] AEA, Techology: a technical summary of the aea egress code. Technical Report 1 (2002)

[102] Kretz, T., Schreckenberg, M.: F.a.s.t. — floor field — and agentbased simulation tool (2006)

[103] Banarjee, S., Grosan, C., Abarha, A.: Emotional ant based modeling of crowd dynamics. In: Seventh International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC'05), pp. 279–286 (2005)

[104] Kirchner, A., Schadschneider, A.: Simulation of evacuation processes using a bionics-inspired cellular automaton model for pedestrian dynamics. Phys. A Stat. Mech. Appl. 312(1–2), 260– 276 (2002)

[105] Duives, D.C., Daamen, W. and Hoogendoorn, S.P., 2013. State-of-the-art crowd motion simulation models. *Transportation research part C: emerging technologies*, *37*, pp.193-209.

[106] Kuligowski, E.D., 2016. Computer evacuation models for buildings. In *SFPE Handbook of Fire Protection Engineering* (pp. 2152-2180). Springer, New York, NY.

[107] Kumar, Santosh, Deepanwita Datta, Sanjay Kumar Singh, and Arun Kumar Sangaiah. "An intelligent decision computing paradigm for crowd monitoring in the smart city." *Journal of Parallel and Distributed Computing* 118 (2018): 344-358.

[108] Huang, Lida, Tao Chen, Yan Wang, and Hongyong Yuan. "Forecasting Daily Pedestrian Flows in the Tiananmen Square Based on Historical Data and Weather Conditions." In *ISCRAM*. 2015.

[109] Mautz, R.: Indoor positioning technologies. Habilitation Thesis. ETH Zurich, Switzerland, 2012.

[110] Padgette J, Scarfone K, Chen L. NIST Special Publication 800-121 Revision 2, Guide to Bluetooth Security. National Institute of Standards and Technology (NIST). 2017.

[111] UCSD dataset: www.svcl.ucsd.edu/projects/peoplecnt/

[112] Mall dataset: personal.ie.cuhk.edu.hk/~ccloy/downloads_mall_dataset.html

[113] BEHAVE datasets: groups.inf.ed.ac.uk/vision/BEHAVEDATA/

[114] CBE dataset: graphics.cs.ucy.ac.cy/research/downloads/crowd-data

[115] Crowd data repository, The Chinese University of Hong Kong: www.ee.cuhk.edu.hk/~xgwang/datasets.html

[116] Crowd data repository, City University of Hong Kong: visal.cs.cityu.edu.hk/downloads/

[117] Yet Another Computer Vision Index To Datasets (YACVID): riemenschneider.hayko.at/vision/dataset/index.php?filter=+crowd. Alternative link, if original link is inaccessible: https://blog.csdn.net/GarfieldEr007/article/details/49951125

[118] TRANCOS Dataset, Universidad de Alcalá: agamenon.tsc.uah.es/Personales/rlopez/data/trancos/

[119] Daimler Pedestrian Benchmark Datasets: http://www.gavrila.net/Datasets/Daimler_Pedestrian_Benchmark_D/daimler_pedestrian_benchmark_d.html

[120] WiFi-based activity recognition dataset: github.com/linteresa/WiAR

[121] Determe JF, Singh U, Horlin F, De Doncker P. Forecasting Crowd Counts With Wi-Fi Systems: Univariate, Non-Seasonal Models. IEEE Transactions on Intelligent Transportation Systems. 2020 May 12.

[122] ShanghaiTech Dataset: https://github.com/desenzhou/ShanghaiTechDataset

[123] UCF-QNRF Dataset: https://www.crcv.ucf.edu/research/data-sets/ucf-qnrf/

[124] Monitoring human activity dataset by the University of Minnesota: http://mha.cs.umn.edu/

[125] GTA5 Crowd Counting (GCC) dataset: https://arxiv.org/pdf/1903.03303.pdf

[126] Fussey, P. and Murray, D.: Independent Report on the London Metropolitan Police Service's Trial of Live Facial Recognition Technology. Human Rights Centre, University of Essex, 2019.

AUTHOR INFORMATION

Utkarsh Singh received a bachelor's degree in electrical and electronics engineering from Uttar Pradesh Technical University, India, in 2012. He completed his master's degree with AICTE fellowship in Power Systems from Thapar University, India, in 2014. He completed his PhD with MHRD fellowship in Power Quality from the Indian Institute of Technology Roorkee in February 2018. From June 2018 to January 2020, he worked as a postdoctoral researcher in the OPERA-Wireless Communications Group at Université libre de Bruxelles, Belgium, on project 'MUFINS' funded by INNOVIRIS. Since February 2020, he is working as a postdoctoral researcher in the Artificial Intelligence Lab at Vrije Universiteit Brussels. His research interests include applied machine learning, optimization, signal processing and power systems.

Jean-François Determe received the bachelor's degree in electrical engineering (Master en ingénieur civil électricien) from Université libre de Bruxelles (ULB) in 2013.He also jointly received the PhD in Engineering from ULB and Université Catholique de Louvain in 2018. From 2013 to 2017, he was an FNRS research fellow, funded by the Belgian FRS-FNRS. Since 2018, he has been a postdoctoral researcher with the OPERA-WCG department at ULB and is currently funded by Innoviris. His research interests focus on applied time series analysis and on sparse recovery algorithms.

François Horlin received the Ph.D. degree from the Université Catholique de Louvain (UCL) in 2002. He specialized in the field of signal processing for digital communications. After his Ph.D., he joined the Inter-university Micro-Electronics Center (IMEC). He led the project aiming at developing a 4G cellular communication system in collaboration with Samsung Korea. In 2007, François Horlin became professor at the Université libre de Bruxelles (ULB). He is currently supervising a research team working on next generation communication systems. Localization based on 5G signals, filter bank-based modulations, massive MIMO and dynamic spectrum access are examples of currently investigated research topics.

Philippe De Doncker received the M.Sc. degree in physics engineering and the Ph.D. degree in science engineering from the Université Libre de Bruxelles (ULB), Brussels, Belgium, in 1996 and 2001, respectively. He is currently a Professor with the ULB, where he leads the research activities on wireless channel modeling and electromagnetics.