CROWD ANALYSIS USING VISUAL AND NON-VISUAL SENSORS, A SURVEY

Muhammad Irfan, Lucio Marcenaro

DITEN
University of Genova, Italy

Laurissa Tokarchuk

Centre for Intelligent Sensing
Queen Mary University of London

ABSTRACT

This paper proposes a critical survey of crowd analysis techniques using visual and non-visual sensors. Automatic crowd understanding has a massive impact on several applications including surveillance and security, situation awareness, crowd management, public space design, intelligent and virtual environments. In case of emergency, it enables practical safety applications by identifying crowd situational context information. This survey identifies different approaches as well as relevant work on crowd analysis by means of visual and non-visual techniques. Multidisciplinary research groups are addressing crowd phenomenon and its dynamics ranging from social, and psychological aspects to computational perspectives. The possibility to use smartphones as sensing devices and fuse this information with video sensors data, allows to better describe crowd dynamics and behaviors. Eventually, challenges and further research opportunities with reference to crowd analysis are exposed.

Index Terms— Crowd Analysis, Crowd Dynamics, Computer Vision, Crowd sensing, Smart-phones, Sensors

1. INTRODUCTION

A crowd is a deformable group of people occupying a particular area. In public places like airports, train stations and shopping malls, which are characterized by gatherings of thousands of people, there is necessary to detect, count, track individuals in order to identify any suspicious behavior. It is impossible for the human surveillance operators to track every individual and identify possible suspicious behaviors, manually, especially in crowded places. This issue motivates the researchers worldwide to design automated systems in order to facilitate these human operators for the aforementioned activities.

As an emerging field of research, crowd analysis is primarily motivated by the security and safety issues covering the surveillance of crowded environments. Target tracking [1], anomaly detection [2, 3], behavioural recognition [4, 5], emotions recognition [6] and crowd density estimation etc. [7, 8, 9] are possible ingredients of crowd analysis.

In the last decade, considerable research has been made to analyse crowds from signal processing community [10, 11]. Recently, research has focused mainly on automated crowd analysis and notable progress has been achieved [8]. However, several unsolved problems are still to be addressed in visual crowd analysis particularly dealing with accurate and robust target detection in crowd, tracking, crowd modeling, occlusion and clutter handling [11]. Owing to the aforementioned issues, such systems often fail to detect an event of interest and ultimately miss the required vital activity. A general framework [11] depicting the processes involved in the crowd analysis can be seen in Fig. 1. Several features such as objects locations and speed, density, etc. can be extracted (manually or automatically) from a guarded environment. These data can be used for creating a crowd model.
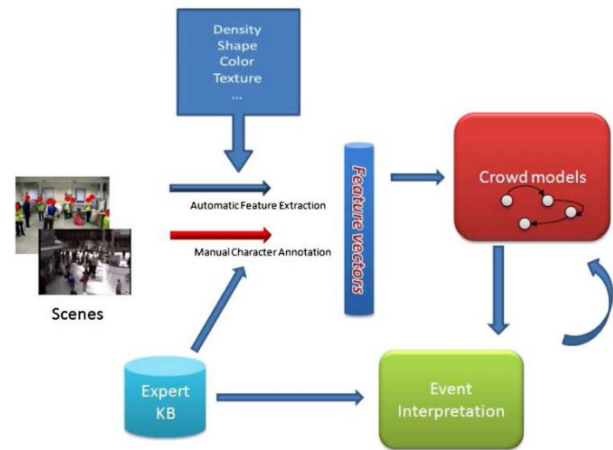

Fig. 1. A general framework for crowd analysis

Due to the widespread availability of smart phones and other non-visual sensors, the non-visual sensing based research community has been working recently on crowd analysis [12]. To analyse and understand crowd dynamics, non-visual sensing has remarkable advantages in terms of applications like density estimation and localization. Mobile phones facilitates on-body sensing which is not geographically restricted to a specific structured area. Nowadays mobile phones are already equipped with sensors like accelerometers, gyroscopes, Wi-Fi, GPS, proximity sensors etc. This rich set of sensors allows us to characterize the behaviour of individuals in a crowd [12]. While applications such as health monitoring are already proving the power and usefulness of smartphones [13], this paper will demonstrate that the sensors we already carry in our pocket are able to deeply improve security and safety of public spaces.

The main contribution of this paper is to analyze different methods applied with visual and non-visual sensors with particular attention towards tracking/detection and counting/density estimation of people in crowd. Different survey papers have been published recently regarding crowd analysis but these only targeted either visual analysis [8, 10, 11] or non-visual [14]. However, the aim of this paper is to provide a comparative survey for the crowd analysis that considers both visual and non-visual techniques.

The paper is organized as follows. Section 2 explores mobile sensing technologies and its usage in crowd analysis. In section 3 different applications and the relevant work to crowd analysis has been discussed. Section 4 demonstrates the different incentives offered to participants in crowd sensing. In section 5 the issues related

to users privacy have been discussed. In section 6 we show datasets usually used in visual analysis and discuss the software tools developed in order to acquire mobile phone based data. A theoretical comparison of both the major technologies, visual and non-visual, has been presented in tabular form in section 7. And finally section 8 concludes and explores the pros and cons of both visual and non-visual sensors and shows future opportunities in the domain.

2. MOBILE SENSING

The recent increase in the use of smart phones and other mobile devices has opened the opportunity to collectively sense and share information for common interests [15]. Smartphones and other wearable devices like smart watches, health monitors, pedometers, activity trackers now come equipped with an impressive array of sensors: multiple position sensors (GPS, WiFi, cellular radios), inertial sensors (accelerometers and gyroscopes), magnetic compass, microphone, light sensors, proximity sensors, and many more [16]. These provide such devices the ability to discover more about wearers and their activities [12]. Sensors available on smart phones can be categorized as inertial, positioning, and ambient sensors. Each sensor is capable of sensing different aspects of individual context and can be selected and configured based upon application requirements.

Smart phones can be used for collective sensing called mobile crowd sensing [14, 17]. Mobile crowd sensing refers to the wide variety of sensing models in which users with sensing and computing devices collect and contribute to required data for different application like object tracking, counting, behavior analysis and anomaly detection [17].

Mobile sensing based systems are recent trends in the crowd analysis research. The signal processing comunity has achieved substantial success in different aspects of crowd analysis. However, issues like occlusion, clutters, weather conditions etc. create hurdles in visual analysis of crowd. This compels us towards the exploitation of smart-phones and motivate us to investigate the sensor information extracted in order to tackle missed information by vision based sensors.

3. APPLICATIONS

Signal processing techniques have helped to understand and analyse crowd, based on macro and micro level modeling. Macro modeling refers to holistic approach of any crowd for analysis while micro relates to the distinct individual level study of crowd. According to [8], 22% studies have focused on micro-level modelling as compared to 71% at the macro-level, because the complexity increases while trying to analyse the individual level of crowd due to occlusion, computational complexity and the appearance similarities. On the other hand non-visual research has mostly targeted groups (people with similar behavior and objectives), rather than the whole crowd [18, 19]. Specific groups are detected and analysed for localization, tracking, density, behaviour and abnormality detection. Within crowd analysis there are five major areas of focus including localization, tracking, density estimation, behaviour recognition, emotions detection and abnormality detection. The following subsections will examine only detection/tracking and counting/density estimation approaches used by both visual and non-visual researchers before highlighting relevant issues that could be addressed by fusion techniques.

3.1. Tracking/Detection

Detection and tracking is the key feature in order to develop an automated surveillance and behavioural analysis system for crowd. Tracking is the problem of estimating trajectory of moving objects in a certain environment [20, 21]. In visual systems, the inherent complexities like occlusion, variations in viewpoints, environmental issues (such as rain, snow, shadows, etc.) or noise, light variations, crowd density etc. are the key challenges. Various motion models including optical flow or background modelling have been used for this application.

Different authors worked on the different aspects of the crowd ranging from individual to groups, from single camera to multi cameras [22] and from sparse to dense crowd. Ge et al. [23] worked on detecting and tracking small groups of pedestrians in crowd with limitation of noise as it was not working efficiently for dense crowd. To overcome this issue of noise for dense crowd, Cheriyadat et al. [24] proposed a method, considering estimated dominant motion in dense crowd by tracking low level features using optical flow. Also, Ozturk et al. [25] used the same dominant motion flows in crowd through a hierarchical clustering approach of SIFT(Scale-invariant feature transform) features. To get rid of the aforementioned noise and clutter, some authors have used body parts models for single and multiple people , which resulted in high accuracy [1, 26]. Social force models are also exploited for the detection and tracking of interacting groups of people in crowds [27]. The results were variable subject to density of crowd.

On the other hand, the non-visual based crowd analysis research has recently made remarkable advances. Researchers from non-visual systems are leveraging the mobile sensors embedded in smartphones to detect and track people in crowd. In [28] a mobile technology and a participatory Sensing approach has been proposed for Crowd analysis. Similarly, in [18] a group detection system, GruMon for dense, urban areas has been developed. It has the capability to monitor groups in a fast and accurate way within a dense and congested urban spaces. Although, the system developed has a high success rate of detection, it is only limited to small groups and structured areas like shopping malls. In other work, Barkhuus et al. in [29] proposed a multi modal method for indoor recognition of groups (flocks) using sensors like accelerometers, WiFi radios and compasses sensors. They used data fusion techniques to merge the multi sensor information in order to improve recognition accuracy against the former unimodal approaches. Although accurate detection and recognition of flocks was successfully achieved, complications like large group, Wi-Fi network availability and unstructured environments are affecting this work. In short, each system has common constraints in terms of frequency of individuals ranging from single to groups, structured to unstructured areas and sparse to dense crowd. The camera based systems have succeeded to detect and track individuals in crowd to some extent as accuracy decreases with respect to time in a video. Also, most of the systems are partially online to perform real time detection and tracking individuals in dense crowd. In contrast, mobile based systems are capable of detecting and tracking dense crowd in real time. Some challenges still persists in mobile based systems like privacy, application installation, energy constraint and infrastructure availability like Wi-Fi access points. However, studies have revealed that in terms of privacy, people can be encouraged to participate when they see personal benefit. Also, incentives can be given to motivate them to participate and install the required application. So over all scenario can be sum up in the way that visual systems can be effective in sparse crowd detection and tracking and mobile based system can insure progress

in dense and urban level crowd analysis.

3.2. Counting/ Density Estimation

Counting and detection of people in crowd is being used combined by different authors in their respective research activities. Estimating number of people in a particular area and space is of extreme importance in order to avoid any probable crowd disaster or any panic situation, and to support the subsequent emergency evacuations measures. The estimation of the density of a crowd may be required for different reasons but mainly in crowd management. In terms of individual applications, it supports in recommending where to go out based as, where many other people have already gone the same place being occupied.

Different types of sensors are used for the people counting systems which includes vision sensor, ultrasonic sensor, infrared sensor and depth cameras. Mostly, cameras are being used with different approaches at various locations with manifold angles. The camera may be used for both surveillance and people counting purposes as well. Researchers have worked on density measurement of crowd with respect to particular scenarios. Some have worked on public places like sports stadiums, metro stations, bus stops etc. while some have worked only for outdoor areas specifically.

Cong et al. [30] used velocity flow fields for people counting with an absolute error of 0.115–0.409. In another approach for people counting, proposed by Yogameena et al. [31] an error rate of only around 2.5% has been reported. Xing et al. [32] used the method of detection responses from video sequence, using detection flow, for crowd counting. A holistic approach was adopted by Garcia-Bunster et al. [33] for counting people waiting at routine bus stops. Fu et al. in [34] presented fusion approach of visual (RGB) and depth data for real time crowd counting.

On the non-visual based systems, authors have applied different approaches for counting crowd using Bluetooth sensors, Wi-Fi, RFID, magnetometer and inertial sensors etc. Weppner et al. [35] present a mobile phone based technique to estimate crowd density by using a pre-set scan of the environment for Bluetooth devices. A Wi-Fi technology based counting system has been developed by Saandeep et al. [36], where people moving between a pair of stationary transmitter/receiver antennas will be counted in a structured environment. A system, SmartEvacTrak, is developed by Ahmed et al. [37] characterized by a low-cost evacuation system which can count people using smartphones, while entering and exiting with accuracy of over 98%. Recently, Federico et al. [38] proposed a mechanism for quantifying crowd size by using mobile phones data and online service twitter where there is no need for users to install any application.

Considering the above works related to crowd measurement in terms of visual and non-visual methods, we can find that mobile phones can be a better choice. Video based systems still have many challenges to estimate the accurate number of participants due to occlusions, clutter and night vision issues. While, smartphones on the other side have advantage over camera based system as they do not require any deployment of infrastructure as in the case of camera based system.

4. INCENTIVES

In mobile crowd sensing, individuals usually avoid to participate and share their details. Various applications have been developed in order to gather reliable data from users to analyze crowd which do not guarantee to show good performance due to rare participation of users. To motivate the users to participate and share the data effectively the only possibility is to stimulate the individuals and grant incentives to participate.

Treasure [39] is a mobile game that is designed based on network infrastructure. In this game WiFi coverage maps are built in a given game area. Players carry GPS and WiFi enabled mobile devices and move in that particular space. In terms of geographical locations, geographic data is a type of sensed data which is collected and recorded by the mobile devices. The sounds and GPS traces can be collected in a specific region. GeoTicTacToe [40] is a location based version of the traditional game Tic Tac Toe which is based on location findings. The most intuitive incentive to participate and contribute to crowd sensing is to pay as it has mostly involved many smart phone users in it to sell their data. The participants receive money as rewards once they make contributions in the sensing activities.

So, in general we can conclude from the above statements that incentives play a very vital role in mobile crowd sensing systems as they attract a tremendous number of participants to contribute to the sensed data.

5. PRIVACY ISSUES

Privacy has been the most crucial aspect in the mobile crowd sensing systems. Users often refuse to share their personal data with respect to time and space [41]. This usually leads to a hindrance in a long term research activity involved. But in order to motivate the smart phone users to contribute to the crowd sensing, they can be asked to participate if the real situation, which is unwanted, and its possible controlling and avoidance ways, are shared with them. In [42, 41] it was shown that people are less concerned about their location privacy when being tracked in order to utilize beneficial services. However, challenges are faced as not every body is carrying smartphone and also it varies with different countries, cultures, gender etc. Thus, in terms of personal safety people often accept to share their personal data.

6. TOOLS AND DATASETS

Choices of datasets are the key components for a systematic evaluation of methods in crowd analysis. Visual datasets are frequently available on internet. Examples include UCSD, PET, UCF, CAVIAR, and UMN etc. [8]. In mobile crowd sensing datasets have been another major issue . As it varies upon situation to collect data and represent the data in tailored form for processing. Application tools have been the centric entity used to gather the crowd sensory data. Various applications have been developed and are used to collect crowd sense data using smartphones. A crowd sensing application SensingKit, [43], is an efficient, open-source, independent platform client-server system supported by both iOS and Android mobile devices. It has the capability of continuous sensing the devices motion (Accelerometer, Gyroscope, Magnetometer), location (GPS) and proximity to other smartphones (Bluetooth Smart). These tools are application specific like for music concerts, railway stations, shopping malls etc. Another application, MobiSens, [44] a versatile android based mobile sensing platform has been designed for real world applications. Social psychology studies based sensing program, EmotionSense, [45] was designed which can sense individual emotions along with activities, verbal and proximities interactions among social group members.

7. COMPARATIVE ANALYSIS

As discussed in section 1, different vision based tools for crowd analysis have been deployed but suffer from the poor scalability of crowded public spaces and also manually operating the systems is a tiresome job with unreliable results. Recently, multi-camera networks using fusion techniques have the capability to mitigate such issues in order to enable the systems automatic for monitoring crowds. Despite of the advances in visual based systems, it is still challenging to get information of global situation awareness during mass gathering from visual sensors. Again, thanks to the pervasiveness of mobile phones like smart phones, which are being used recently to analyze crowd. Mostly the state-of-the-arts, either visual or non-visual, are dependent on scope and type of crowd scene. In terms of scope the author varies holistic techniques with individual/object based methods using visual technologies. Crowd scenes may be structured and unstructured depending upon the crowd dynamics. As the names states in the holistic method, a top-down methodology is used and the crowd is considered as a single entity. Structured crowds have a structured flow of motions in a scene and vice versa. These methods use the global level information to estimate density as well as tracking of crowd as done in [46] using optical flow which is suitable both for structured and unstructured scenes. In individual based approaches, individuals are detected and tracked using energy function such as performed in [1].

On the other side, recently many non-visual sensors based approaches are used for crowd detection. Location is the crucial factor to infer the conditions of any crowd. To determine the location of any smart-phone, different approaches are used which are broadly categorized as: networked based localization and on-device based localization. The network based methods are based on cellular networks which can be exploited to find which smart-phone is connected to a particular cell tower. Since the location of each cell tower is already available so, position estimation is obtained easily. Couronne et al used the network bandwidth in [49] for detection of crowd in a particular location which is suitable both for structured and unstructured scenes. In the scope of on-device detection and tracking approaches the work of Rijurekha et al in [18] is a good example which helps to detect and track groups in public places like sopping malls. Table 1 and 2 summarizes our literature review regarding different technologies and approaches used in the crowd detection/tracking and density estimation. Consequently, this can be derived from the works listed in the tables above that non-visual based approaches are better alternatives to visual based systems as using on-device and network based approaches provide more accurate locations as compared to visual systems. Also, the systems are flexible with reference to structured and non-structured scenes. But, we cannot ignore the challenges faced by on-device based systems like privacy issues and participatory sensing as discussed in section 5.

8. CONCLUSION AND DISCUSSION

In this paper, we explored the different ways adopted for crowd analysis based on visual and non-visual sensors. Perspectives from computer vision and mainly smart phone based systems are presented as these have contributed to an enough in-depth study to crowd analysis applications like tracking/detection, counting, behaviour analysis and anomaly detection. Theoretical comparison is presented for tracking/detection and density estimation application area in order to choose the right technologies for crowd analysis.

In terms of tracking and detection visual systems can show performance to detect and track dense crowd in a structured environ-ment but at the macro level. Smart phone and social network based systems are efficient mostly on urban level as they have the capability to communicate with each individual in order to intimate in case of critical situation. Thus, non-visual system may be used both on macro and micro level tracking and detection. With reference to density estimation of crowd, visual systems have failed to find the accurate figure of crowd in a certain non-structured space. Due to high density of crowd, counting becomes difficult as the problems like occlusion, clutter, night vision and weather conditions may become hurdles. Recently it has been seen that the police fired tear gas on the supporters of football teams, resulting occlusion in camera based systems. Ultimately vision systems may not be able to decide where to direct the people to escape. Smart phone based systems can be accessed easily in emergency situations to show the right paths to the users. The non-visual research community has focused very rarely on the problem of detection/tracking and density estimation in crowd, which are based on smart phones. However, compromising in both technologies, we can infer that data fusion techniques [50] can help to combine the visual and non-visual information from sensors of different types to derive reliable patterns and data.

9. REFERENCES

[1] M. Rodriguez, I. Laptev, J. Sivic, and J. Y. Audibert, "Density-aware person detection and tracking in crowds," in *2011 International Conference on Computer Vision*, Nov 2011, pp. 2423–2430.

[2] D. Xu, R. Song, X. Wu, N. Li, W. Feng, and H. Qian, "Video anomaly detection based on a hierarchical activity discovery within spatio-temporal contexts," *Neurocomputing*, vol. 143, pp. 144 – 152, 2014.

[3] S. Chiappino, P. Morerio, L. Marcenaro, E. Fuiano, G. Repetto, and C. S. Regazzoni, "A multi-sensor cognitive approach for active security monitoring of abnormal overcrowding situations," in *15th International Conference on Information Fusion, FUSION 2012, Singapore, July 9-12, 2012*, 2012, pp. 2215–2222.

[4] Z. Cheng, L. Qin, Q. Huang, S. Yan, and Q. Tian, "Recognizing human group action by layered model with multiple cues," *Neurocomputing*, vol. 136, pp. 124 – 135, 2014.

[5] S. Chiappino, P. Morerio, L. Marcenaro, and C. S. Regazzoni, "Bio-inspired relevant interaction modelling in cognitive crowd management," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–22, 2014.

[6] M. W. Baig, M. S. Baig, V. Bastani, E. I. Barakova, L. Marcenaro, C. S. Regazzoni, and M. Rauterberg, "Perception of emotions from crowd dynamics," in *2015 IEEE International Conference on Digital Signal Processing (DSP)*, July 2015, pp. 703–707.

[7] R. Liang, Y. Zhu, and H. Wang, "Counting crowd flow based on feature points," *Neurocomputing*, vol. 133, pp. 377 – 384, 2014.

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

[9] P. Morerio, L. Marcenaro, and C. S Regazzoni, "People count estimation in small crowds," in *Advanced video and signal-based surveillance (AVSS), 2012 IEEE Ninth International Conference on*. IEEE, 2012, pp. 476–480.

Table 1. Overview of Technologies and Methods for Crowd Detection/Tracking Assessment

Sensor Modality	Scope	Technique	Structured	Unstructured	References
Visual	Holistic	Optical Flow	✓		[46]
		Dominant Motions	✓	✓	[24]
		Dominant Motions	✓		[25]
	Individual	Energy Function		✓	[1]
		Hierarchical Clustering		✓	[23]
		Synthetic Aperture Imaging	✓		[22]
		Baysian Approach	✓		[26]
		Social Force Model		✓	[27]
Non-Visual	Network	Bandwidth Usage	✓	✓	[47]
	On-device	GruMon	✓		[18]
		Tracking	✓	✓	[28]
		Data Fusion + Clustering	✓	✓	[29]

Table 2. Overview of Technologies and Methods for Crowd Counting/Density Assessment

Sensor Modality	Scope	Technique	Structured	Unstructured	References
Visual	Holistic	Velocity Estimation	✓		[30]
		Pixel Count	✓	✓	[48]
		Machine Learning + Thresholding	✓		[31]
		Linear Regression	✓		[33]
		RGB Plus Depth	✓	✓	[34]
		Optical Flow	✓		[46]
	Individual	Detection Flow	✓		[32]
Non-Visual	Network	Call Data Records	✓	✓	[47]
	On-device	Wi-Fi signals Measurement	✓		[36]
		SmartEvactrak	✓		[37]
		Social Network	✓	✓	[38]
		Bluetooth	✓	✓	[35]

[10] J. C. S. J. Junior, S. R. Musse, and C. R. Jung, "Crowd analysis using computer vision techniques," *IEEE Signal Processing Magazine*, vol. 27, no. 5, pp. 66–77, Sept 2010.

[11] B. Zhan, D. N. Monekosso, P. Remagnino, S. A. Velastin, and L. Q. Xu, "Crowd analysis: a survey," *Machine Vision and Applications*, vol. 19, no. 5-6, pp. 345–357, 2008.

[12] D. Roggen, M. Wirz, G. Troster, and D. Helbing, "Recognition of crowd behavior from mobile sensors with pattern analysis and graph clustering methods," *Networks and Heterogeneous Media*, vol. 6, no. 3, pp. 521–544, 2011.

[13] S. D Nanhore and M. M. Bartere, "Mobile phone sensing system for health monitoring," *International Journal of Science and Research (IJ SR)*, vol. 2, no. 4, pp. 252–255, 2013.

[14] B. Guo, Z. Wang, Z. Yu, Y. Wang, N.Y. Yen, R. Huang, and X. Zhou, "Mobile crowd sensing and computing: The review of an emerging human-powered sensing paradigm," New York, NY, USA, Aug. 2015, vol. 48, pp. 7:1–7:31, ACM.

[15] N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. T. Campbell, "A survey of mobile phone sensing," *IEEE Communications magazine*, vol. 48, no. 9, pp. 140–150, 2010.

[16] L. S. Ravindranath, A. Thiagarajan, H. Balakrishnan, and S. Maddenl, "Code in the air: Simplifying sensing and coordination tasks on smartphones," in *Proceedings of the Twelfth Workshop on Mobile Computing Systems & Applications*, New York, NY, USA, 2012, HotMobile '12, pp. 4:1–4:6, ACM.

[17] R. K. Ganti, F. Ye, and H. Lei, "Mobile crowdsensing: current state and future challenges.," *IEEE Communications Magazine*, vol. 49, no. 11, pp. 32–39, 2011.

[18] R. Sen, Y. Lee, K. Jayarajah, A. Misra, and R. K. Balan, "Grumon: Fast and accurate group monitoring for heterogeneous urban spaces," in *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems*, New York, NY, USA, 2014, SenSys '14, pp. 46–60, ACM.

[19] K. Jayarajah, Y. Leei, A. Misra, and R. K. Balan, "Need accurate user behaviour?: Pay attention to groups!," in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, New York, NY, USA, 2015, Ubi-Comp '15, pp. 855–866, ACM.

[20] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," New York, NY, USA, Dec. 2006, vol. 38, ACM.

[21] F. Castaldo, F. A. N. Palmieri, V. Bastani, L. Marcenaro, and C. S. Regazzoni, "Abnormal vessel behavior detection in port areas based on dynamic bayesian networks," in *17th International Conference on Information Fusion, FUSION 2014, Salamanca, Spain, July 7-10, 2014*, 2014, pp. 1–7.

[22] T. Yang, Y. Zhang, X. Tong, X. Zhang, and R. Yu, "A new hybrid synthetic aperture imaging model for tracking and seeing people through occlusion," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 9, pp. 1461–1475, Sept 2013.

[23] W. Ge, R. T. Collins, and R. B. Ruback, "Vision-based analysis of small groups in pedestrian crowds," *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 5, pp. 1003–1016, 2012.

[24] A. M. Cheriyadat and R. J. Radke, "Detecting dominant mo-

tions in dense crowds," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 4, pp. 568–581, 2008.

[25] O. Ozturk, T. Yamasaki, and K. Aizawa, "Detecting motion flows in unstructured/structured crowd scenes," in *Pattern Recognition (ICPR), 2010 20th International Conference on*, Aug 2010, pp. 3533–3536.

[26] M. Butenuth, F. Burkert, F. Schmidt, S. Hinz, D. Hartmann, A. Kneidl, A. Borrmann, and B. Sirmacek, "Integrating pedestrian simulation, tracking and event detection for crowd analysis," in *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*, Nov 2011, pp. 150–157.

[27] R. Mazzon, F. Poiesi, and A. Cavallaro, "Detection and tracking of groups in crowd," in *Advanced Video and Signal Based Surveillance (AVSS), 2013 10th IEEE International Conference on*, Aug 2013, pp. 202–207.

[28] M. Wirz, E. M. Kelly, T. Franke, V. Camilleri, M Montebello, D. Roggen, P. Lukowicz, and G. Troster, "Using mobile technology and a participatory sensing approach for crowd monitoring and management during large-scale mass gatherings," in *Co-evolution of Intelligent Socio-technical Systems*, pp. 61–77. Springer, 2013.

[29] M.B. Kjaergaard, M. Wirz, D. Roggen, and G. Troster, "Detecting pedestrian flocks by fusion of multi-modal sensors in mobile phones," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, New York, NY, USA, 2012, UbiComp '12, pp. 240–249, ACM.

[30] Y. Cong, H. Gong, S. C. Zhu, and Y. Tang, "Flow mosaicking: Real-time pedestrian counting without scene-specific learning," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, June 2009, pp. 1093–1100.

[31] B. Yogameena, S. S. Perumal, N. Packiyaraj, and P. Saravanan, "Ma-th algorithm for people count in a dense crowd and their behaviour classification," in *Machine Vision and Image Processing (MVIP), 2012 International Conference on*, Dec 2012, pp. 17–20.

[32] J. Xing, H. Ai, L. Liu, and S. Lao, "Robust crowd counting using detection flow," in *2011 18th IEEE International Conference on Image Processing*, Sept 2011, pp. 2061–2064.

[33] G. G., M. T. Torriti, and C. Oberli, "Crowded pedestrian counting at bus stops from perspective transformations of foreground areas," *IET Computer Vision*, vol. 6, no. 4, pp. 296–305, July 2012.

[34] H. Fu, H. Ma, and H. Xiao, "Real-time accurate crowd counting based on rgb-d information," in *2012 19th IEEE International Conference on Image Processing*, Sept 2012, pp. 2685–2688.

[35] J. Weppner and P. Lukowicz, "Bluetooth based collaborative crowd density estimation with mobile phones," in *Pervasive computing and communications (PerCom), 2013 IEEE international conference on*. IEEE, 2013, pp. 193–200.

[36] S. Depatla, A. Muralidharan, and Y. Mostofi, "Occupancy estimation using only wifi power measurements," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 7, pp. 1381–1393, July 2015.

[37] N. Ahmed, A. Ghose, A. K. Agrawal, C. Bhaumik, V. Chandel, and A. Kumar, "Smartevactrak: A people counting and coarse-level localization solution for efficient evacuation

of large buildings," in *Pervasive Computing and Communication Workshops (PerCom Workshops), 2015 IEEE International Conference on*, March 2015, pp. 372–377.

[38] F. Botta, H. S. Moat, and T. Preis, "Quantifying crowd size with mobile phone and Twitter data," *Royal Society Open Science*, vol. 2, no. 5, pp. 150162–150162, 2015.

[39] L. Barkhuus, M. Chalmers, P. Tennent, M. Hall, M. Bell, S. Sherwood, and B. Brown, "Picking pockets on the lawn: the development of tactics and strategies in a mobile game," in *International Conference on Ubiquitous Computing*. Springer, 2005, pp. 358–374.

[40] C. Schlieder, P. Kiefer, and S. Matyas, "Geogames: Designing location-based games from classic board games," *IEEE Intelligent Systems*, vol. 21, no. 5, pp. 40–46, 2006.

[41] L. Barkhuus and A. K. Dey, "Location-Based services for mobile telephony: a study of users' privacy concerns.," in *INTERACT*. Citeseer, 2003, vol. 3, pp. 702–712.

[42] M. Wirz, D. Roggen, and G. Troster, "User acceptance study of a mobile system for assistance during emergency situations at large-scale events," in *Human-Centric Computing (HumanCom), 2010 3rd International Conference on*, Aug 2010, pp. 1–6.

[43] K. Katevas, H. Haddadi, and l. Tokarchuk, "Poster: Sensingkit: A multi-platform mobile sensing framework for large-scale experiments," in *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*, New York, NY, USA, 2014, MobiCom '14, pp. 375–378, ACM.

[44] P. Wu, J. Zhu, and J. Y. Zhang, "Mobisens: A versatile mobile sensing platform for real-world applications," *Mobile Networks and Applications*, vol. 18, no. 1, pp. 60–80, 2013.

[45] K. R. Kiran, M. Musolesi, C. Mascolo, P.J. Rentfrow, C. Longworth, and A. Aucinas, "Emotionsense: A mobile phones based adaptive platform for experimental social psychology research," in *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*, New York, NY, USA, 2010, UbiComp '10, pp. 281–290, ACM.

[46] K. Barbara and C. Bauckhage, "Loveparade 2010 automatic video analysis of a crowd disaster," *Computer Vision and Image Understanding*, vol. 116, no. 3, pp. 307–319, 2012.

[47] F. Calabrese, M. Colonna, P. Lovisolo, D. Parata, and C. Ratti, "Real-time urban monitoring using cell phones: A case study in rome," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 1, pp. 141–151, 2011.

[48] R. Ma, L. Li, W. Huang, and Q. Tian, "On pixel count based crowd density estimation for visual surveillance," in *Cybernetics and Intelligent Systems, 2004 IEEE Conference on*, Dec 2004, vol. 1, pp. 170–173 vol.1.

[49] T. Couronne, A-M. O. Raimond, and Z. Smoreda, "Looking at spatiotemporal city dynamics through mobile phone lenses," in *2011 International Conference on the Network of the Future*, 2011.

[50] Y. N. Aye, S. Zhao, and C. Y. Shee, *Fusion of Inertial Measurements and Vision Feedback for Microsurgery*, pp. 341–349, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.