A Review of Crowd Counting Techniques

Jeevitha S* & Rajeswari R*
*Assistant Professor, Department of Computer Applications, Bharathiar University, Coimbatore-641046

Received: May 22, 2018 Accepted: July 03, 2018

ABSTRACT *Crowd counting application plays a vital role in the field of security system. In recent years video surveillance has typically advanced in technology for outdoor security. This survey paper describes the challenges of counting the crowd in a crowded environment. It also provides a survey of various work carried out in crowd counting. The paper also implemented a crowd counting application based on Histogram of Oriented gradients with Support Vector Machine (HOG/SVM) techniques.The implemented proposed method was tested by using the Mall dataset and the results are provided.*

Keywords: *crowd density, crowd counting, surveillance systems, HOG/SVM.*

I. Introduction:

India has the world's second largest population. Recently India is facing many crimes due to over population, poverty, ethnicities, and multiple religions. The increase of population leads to security threat due to this reason now a days application for counting the crowd are more and more important for public security affairs. Security related applications allows greater networking of cameras, greater field of vision, cheaper access and come with a host of tools like facial recognition and as well as vehicle tracking. Video surveillance market is witnessing extremely large growth in areas including manufacturing, BPO [Business Process Outsourcing] retail, airport security, hospitality, city surveillance, college campus, companies, and shopping malls. To avoid overcrowding the safety plans should be managed by monitoring the occupancy level.

One of the challenging tasks is to estimate crowd from using video surveillance. The greatest difficulty in estimating crowd from video surveillance is due to the fact that: (1) Individual objects contain occlusion. (2) Continuous change in the crowd over time (3) Live Background captured with low-Resolution. (4) The inherent difficulty in accurately modeling the crowd behavior [1]. A lot of work has been carried out for counting people in nature scenes and these methods are divided into two categories: detection based and measurement based crowd counting techniques. In detection based method individual human face and shapes are detected using background subtraction, the system will detect the foreground regions and it will match the edges of the foreground region.

This paper provides a survey of various crowd counting techniques. The rest of this paper is organized as follows. A summary of the existing crowd counting techniques is given in section II. Implementation of HOG and SVM based crowd counting technique is given in section III. Section III also provides the results of the implemented method using mall dataset. The conclusion of this paper is given in section IV.

II. Overview of Existing crowd counting methods:

A lot of research has been done on crowd counting and analysis. Most of the techniques depend on the fact that if the instances of people are detected or if the motion of people is detected, then it is possible to estimate the density of the crowd. In order to detect people, segmentation is used. There are different approaches for image segmentation like threshold based, edge-based, cluster-based, pixel-based and neural network based [5]. Trajectory clustering based approach presents a simple unsupervised Bayesian clustering framework to detect the individual's movements in crowds [11]. Clustering was performed based on the density of the crowd and used for detecting motion flow patterns in video crowds [12].

Clustering the input image defines computer vision clustering is known as image segmentation which changes the image into something meaningful format and easier to analyze by partitioning the digital image into a set of a pixel or also known as the super-pixel. The fastest pedestrian detector in the west (FPDW), extended this approach to fast multi-scale detection after it was demonstrated how feature computed at a single scale can be used to approximate features at nearby scales. Achieving a 5fps on 640* 480 images as this method does not need to compute a fine scale-space-pyramid [7][13]. The main drawback of this feature is it will easily block the noisy edge [9] from a cluttered background. Region level

edge-based features are up with edge-based features at a pixel level, edge features obtained from local image regions have also been explored [9].

Hashemzadeh et al proposed a technique to count the crowd by detecting in each and every individual scene the shape of a human body and then tracking the independent motion through motion trajectories. This technique typically works by removing few holistic features like foreground area, edge count, texture etc. These methods are followed to estimate crowd [2].

Ullah et al used Density independent hydrodynamics model (DIHM) to evaluate the performance of the DHIM compares the results with 10-state-of-the-art coherency detection methods including the Lagrangian [20] particle dynamics (LPD), the mixtures of dynamic textures (MDT), the motion segmentation in crowds (MSC), the spatio-temporal model (STM), the local translation domain (LTD), the detection of coherent motion (DCM), the collective motion detection (CMD) the segmentation based on dynamic system (SDS), the trajectory clustering approach (TCA), and the thermal diffusion process (TDP) [26]. Saleh et al estimated crowd by two main approaches which are direct approach i.e. object based target detection and indirect approach e.g. pixel-based, texture-based, and corner points based analysis [11].

Gao et al proposed an effective head detection based people counting method. Three modules are used to detect head: Histogram oriented gradient (HOG) features for detecting, for head classification SVM classifier are used and they have constructed a new dataset which is trained using CNN model [22]. Chang et al designed a count-net based on background and this method focuses more on people's heads by separating crowd images into patches which belongs to one image which are then summed as the output of the whole framework [23].

Ryan et al designed multiple cameras to count people in the crowd using the invariant crowd counting algorithm to compensate the overlapping of the viewpoints. They also investigated the size, shape, edges, and key points. Size refers to the interesting segments from an image. Shape detects the descriptors in an image. Edges measures pixel intensities across an image. Key points detect corners in an image [21]. Dalal et al used HOG features to detect the person and on-person classification by image gradient orientation in a dense grid to contrast normalization overlapping spatial blocks. A test case is used for detecting object linear SVM based human detection [17].

Zhou et al proposed two methods viz., higher-order singular value decomposition (HOSVI) and support vector machine (SVM) to estimate crowd density. They have used geometrical connection for crowd density estimation to reduce the perspective effects and Kalman filter to optimize the estimation [37].

Shen et al define Light Effect Supervisor Model (LESM) to reduce the foreground illumination and to detect the background texture] [3]. Liu et al present an approach that utilizes multiple exemplar agent-based motion models (AMMs) to extract motion features from the captured crowd trajectories which have been demonstrated to be effective at modeling crowd interactions, in order to bridge the gap between AMMs and crowd trajectories. It also recognize different types of crowd motion that are not robust. The fact that most AMMs are non-linear models controlled by several parameters further increases the difficulty of training a robust model [8].

Mousse et al present a novel system to count the crowd scene with overlapping cameras. By using codebook algorithm foreground pixel of every single view in a real-time object are fused. In order to find the visual hull of a foreground, convex hull is obtained by computing each camera's visual hull into ground plane through homograph projection. Then the fuse will obtain polygons by using geometric properties [20].

Sheng et al proposed conventional holistic features to solve the failure of semantic attributes and spatial cues of the images. Dense attribute feature is mapped from the original pixel space where each dimension of a pixel will indicate the strength of the probabilistic of a certain semantic class [13]. Lida et al presented the Congestion based pedestrian method for controlling the crowd in a high-density situation, different movements of pedestrian could happen. On the one hand, people may stay a long time in a jam they could push each other and manage to move if the size of the crowd is high enough, cohesive movements are observable so that velocity entropy is used as an indicator for congestion detection and also tested by the simulation database. Anylogic software is commercial multi-method simulation modeling software which simulates pedestrian flows in a certain environment and it can also be applied to a real-world case [6].

Abnormality detection is explored both from a micro-analysis and macro-analysis points of view [13, 1]. Various anomaly detecting techniques have been defined [2, 6]. Dynamic model (DIHM) for coherency detection [6].Zitouni et al analysis in three distinct phases: crowd modeling, crowd monitoring and crowd management using Lagrangain methods target the detection through motion analysis [12][6]. Zhao et al proposed a novel method of counting the crowd by a unified hierarchical clustering metric to the similarity of original tracklets by using gestalt laws. [10]. Pu et al proposed deep convolutional neural

network (ConNet) for crowd estimation with two-fold: First, Deep network and second 31 scenes of subway-carriage with a dataset over 160k density [14].

Laibe et al handled this problem, without any priors. Their method begins with automatic identification of prominent individuals from the crowd and then uses neighborhood motion concurrence to model the behaviors of individuals to predict motion while leveraging fire-frame instantaneous flow at the time of dynamic changing flow and anomalies [27]. Xiang et al investigate three spatiotemporal statistical properties of the pedestrian which obeys length of a path using power law speeds follow Gaussian distribution and it should maintain pedestrian speed lower in entrance/exit using fuzzy c-means (FCM) algorithm, it will cluster the motion patterns. Since error will occur at the time of tracking, outliers are eliminated using the local factor (LOF) algorithm [15].

Leibe et al presented object segmentation by influencing the new algorithm to evidence multiple stages and from different sources with the combination of local information with global cues of an object silhouette [16]. Setti et al present Spectators Hockey (S-Hock) dataset to describe the high level and low-level features to perform many different tasks like people counting and head pose estimation like excitement and automatic summarization [18].

Lu et al proposed an extended floor field cellular automaton (CA) model which is formulated by leveraging the leader-follower behavior rule that is evident in pedestrian group behaviors. [19]. Fu et al proposed convolutional neural network (ConvNet) which contributes twofold: CNN is introduced to estimate crowd density. The speed of estimation is accelerated by removing some network connections. Later two ConvNet classifiers have been designed to improve both accuracy and speed [24]. Hu et al investigate the use of deep –learning to estimate each individual in a single image from mid-level or high –level crowd. First, a ConvNet structure is used to extract crowd features. Second supervisory signals are used to estimate the crowd density and crowd counting employs crowd features [24]. Zeyad et al designed Collaboration of Gaussian Process Model (GP) method to count people by taking different kernels since it is used to handle and measure the next level of occlusion in a frame [25]. Table 2 provides a comparison of existing approaches for crowd counting.

Table: 2 Existing approaches for crowd counting along their models and dataset

Models	Approaches	Dataset	Challenges	Authors
Macro-level techniques	Motion flow model	PETS 2009	People counting	Zitouni et al
Unsupervised approach	Social Network Model (SNM)	UCSD	Anomaly Detection	Chaker et al
Key-point and segment based approach	SURF	PETS 2009	Crowd density estimation	Hashemzadeh et al
Oriented Inner Edges (OIEs)	Light Effect Suppression (LESM)	PETS 2009	Occlusion problem	Shen et al
Gaussian Process Method	Gaussian Regression Model	Mall & UCSD	Occlusion	Q.H et al
Convolutional Neural Network (CNN)	Count-net based Model	UCF & AHU - CROWD	Focusing background and people's head	Zhang et al
Convolutional Neural Network (CNN)	ConvNet Classifier	PETS_2009 & Subway	Crowd Density Estimation	Fu et al
Spectator crowd analysis	Sociological classifiers	S-HOCK	Crowd analysis	Setti et al
Gestalt law to measure similarity of tracklets	_	CUHK , CASIA crowd	Human crowd analysis	Zhao et al
Agent- based motion flow	Crowd motion flow	Real world crowd dataset	Crowd behavior	Liu et al

Figure 1: Target Image from Mall Dataset

Figure 2: HOG/SVM classification for crowd counting

III. Implementation and Result:

Histogram Oriented Gradient and Support Vector machine (HOG/SVM) based crowd counting technique is implemented using MATLAB. Fig.1 shows the sample image with the locations of the people

which is used as a ground truth image. Mall dataset is used in our work. 10 video sequence images from Mall dataset are used for our experiments images and the results obtained from these images are shown in Fig.2. A graph showing the actual number of people in the image and the number of people counted using the implemented method is shown in fig 3.

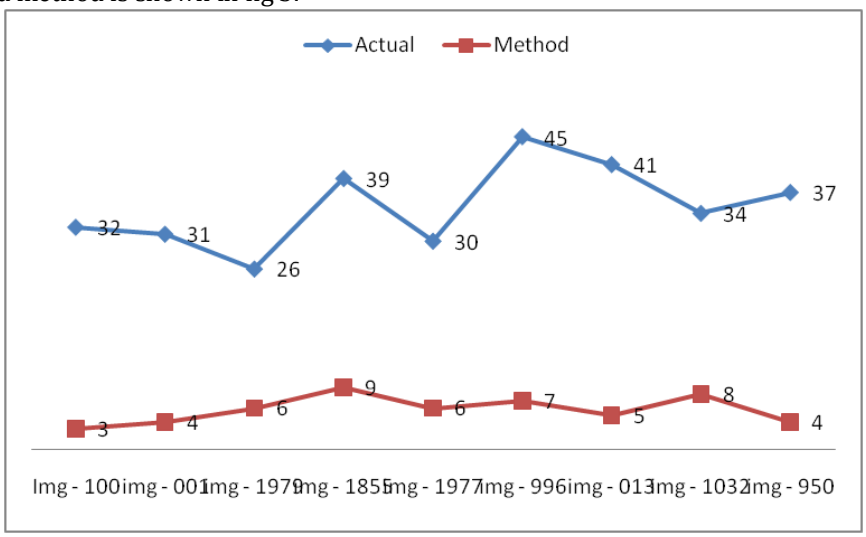

Figure 3: Comparison graph for people count actual and HOG/SVM method

IV. Conclusion:

In the recent years, research on crowd counting is given importance as it helps in detecting people misbehaving in video sequences. This paper presented a study of various crowd counting techniques. Counting crowd is difficult due to reasons such as illumination changes in each image scene. The results obtained using an implemented HOG/SVM based crowd counting method is also given in this paper. Mall dataset is considered for evaluating the implemented method.

References:

1. Rima Chaker, Zaher Al Aghbari, ImaranN.Junejo, (2017) "Social network model for crowd anomaly detection and localization", Pattern Recognition, Vol. 61, pp 266-281.
2. Mahdi Hashemzadeh, NacerFarajzadeh, (2016),"Combining keypoint- based and segment- based features for counting people in crowded scenes", Information Sciences, Vol. 345, pp 199-216,
3. MayingShen, Tanfeng Sun, xingahao Jiang, KeXu, "Crowd counting Estimation in Video Survelliance Based on Linear Regression Function" 9th international Congress on Image and Signal Processing, Vol. 978, 201.
4. Alexey V.Kurilkin and Sergey V.Ivanov, (2016), "A Comparison of methods to detect people flow using video processing", Procedia Computer Science, 5th International young scientist conference on Computation Science, vol.101, pp 125-134, .
5. Lida Huang, Tao Chaen, Yan Wang, Hongyong Yuan, (2015),"Congestion detection of pedestrian using the velocity entropy: A case study of Love Parade 2010 disaster", Physica , Vol. 440, pp 200-209.
6. Jase Macros Rodriguez Fernandez, (2014), "Computer vision of pedestrian detection using Histogram of Oriented Gradients".
7. Wenxi Liu, Rynson W.H. Lau, Dinesh Manocha, (2016), "Robust individual and holistic features for crowd scene classification", Patter Recognition, Vol. 58, pp 110-120.
8. Welqi Zhao, Zhang Zhang, Kaiqi Huang, (2018),"Gestalt laws based tracklets analysis for human crowd understanding", Patter Recognition, Vol. 75, pp 112-127 .
9. Biyun Sheng, ChunhuaShen, Guosheng Lin, Wankau Yang, Changyin Sun, (2016), "Crowd counting via Weighted VLAD on Dense Attribute Feature Maps", Vol. 29.
10. Priyansh Sharma and JenkinSuji, (2016), "A Review on image segmentation with its Clustering techniques", International Journal of Signal processing, Image Processing and Pattern Recognition, Vol.9, No.5 (2016), PP.209-218.
11. Vishwanath A. Sindagi, Vishal M.Patel, "A Survey of recent advances in CNN- based single image crowd counting and density estimation", Pattern Recognition Letters, Vol. 000, pp 1-14.
12. HaroonIdrees, Nolan Warner, Mubarak Shah,(2014), "Tracking in dense crowds using prominence and neighborhood motion concurrence", Image and vision Computing, Vol. 32, pp 14-26.
13. Wei Lu, Xiang Wei, Weiwei Xing, Weibin Liu, (2017),"Trajectory- Based motion pattern analysis of crowds", Neurocomputing, Vol. 247,pp 213-223.

14. VishwanathA.Sindagi, Vishal M.Patel,(2017), "A Survey of recent advances in CNN-based single image crowd counting and density estimation", Patter Recognition Letters.
15. NavneetDalala and Bill Triggs, (2005), "Histograms of Oriented Gradients for Human Detection", Proceedings of the 2005 IEEE Computer Society Conferences on Computer Vision and Pattern Recognition.
16. Lili Lu, Ching-yao Chan, Jian Wang, EWei Wang, (2017)," A study of pedestrian group behaviors in crowd estimation based on an extended floor field cellular automaton model", Transportation Research Part C, Vol. 81, pp 317-329.
17. Mikael A. Mousse, CinaMotamed. Eugene C.Ezin, (2017), "People counting via multiple views using a fast information fusion approach", Springer science + Business Media New York, Vol. 76, pp 6801-6819.
18. David Ryan, Simon Denman, Clinton Fookes, SridhaSridharan, (2014),"Scene invariant multi camera crowd counting", Pattern recognition letters, Vol. 44, pp 98-112.
19. ChenqiangGao, Pei Li Yajun Zhang, Jiang Liu , Lan Wang, (2016),"People counting based on head detection combining Adaboost and CNN in crowded surveillance environment", Neurocomputing ,Vol. 208, pp 108-116.
20. Youmei Zhang, Failang Chang, Mengdi Wang, Fulei Zhang, Cho Han,(2018), "Auxiliary learning for crowd counting via count-net", Neurocomputing, Vol.273, pp 190-198.
21. Min Fu, Pei Xu, Xudong Li, Qihe Liu, Mao Ye, Ce Zhu,(2015), "Fast crowd density estimation with convolutional neural networks", Engineering Application of Artificial Intelligence, Vol. 43, pp 81-88.
22. Yaocong Hu, Huan Chang, FudingNian, Yan Wang, Tengi Li, (2016), "Dense crowd counting from still images with convolutional neural networks", J.Vis. Commun. Image R., Vol. 38, pp 530-539.
23. Zeyad Q.H. Al-Zaydi, David L.Ndzi, Yanyan Yang, Munirah L. Kamarudin, (2016), "AN adaptive people counting system with dynamic features selection and occlusion handling", J. Vis. Commun. Image R., Vol. 39, pp 218-225.
24. Bingyin Zhou, Fan Zhang, LizhongPeng, (2012), "Higher-order SVD analysis for crowd density estimation", ELSEIVER Computer vision and Image Understanding, Vol. 116, pp 1014-1021.
25. HabibUllah, Muhammad Uzair, MohibUllah, Asif Khan, Ayaz Ahmad, Wilayat Khan,(2017), "Density independent hydrodynamics model for crowd coherency detection", Neurocomputing, Vol. 242, pp 28-39.
26. Bastian Leibe, Edgar Seeman, and BerntSchiele, "Pedestrain Detection in Crowded Scenes",(2005), Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Patter Recognition pp.1063-6919.