International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 06 | June 2019 www.irjet.net p-ISSN: 2395-0072

Estimation of Crowd Count in a Heavily Occulated Regions

Swathi D G [1], Jalaja G [2]

[1]Student, Department of Computer Science and Engineering, BNMIT, Karnataka, India
[2]Associate Professor, Department of Computer Science and Engineering, BNMIT, Karnataka, India

---

Abstract - *Crowd estimation is a challenging task of accurately estimating number of people in a crowd region. This paper aims to address crowd counting problem from the perspective of two models i.e, body part map and structural density map. The two models are created by combining the information of pedestrian, their head and context structure. Deep Convolutional neural networks and motion detection method is used to count the number of people in the crowd region, based on the pixel movement of the video frames. CNN technique improves the efficiency of counting people in videos and high accuracy is achieved.*

Key Words: **Crowd Counting, Deep Convolutional Neural Networks, Motion Detection, Pedestrian Detection, Crowd Estimation.**

1.INTRODUCTION

Crowd estimation is the task of efficiently estimating number of pedestrians in a dense region. Crowd counting has harassed much curiosity from scientist due to the practical stipulation like for controlling large number of pedestrians and public security. Detection of a human is a basic issue in video supervision systems. It is estimated that the world population will be 11.2 billion in 2100 years, which is double the current population of the world (7.4 billon, 2016). Due to rapidly growing population across the world, crowd analysis and crowd monitoring has become an important field for research. Manually counting people in the dense crowded areas user cannot estimate the accurate crowd count of the pedestrians present in the area. To overcome this, a system is developed to provide crowd count. Crowd count is any dense scene is provided based on three key factors: pedestrian, head and context structure, are planned as two scene models. The first model is body-parts map, which is obtained by finding the body parts of individual person in dense scene and merging the segmentation mask. The second model is structural-density map, which is created based on shape of individual persons obtained from body-parts map. Then result of two models are combined to provide crowd count of the dense scene. There are several applications of crowd counting some of them are listed below: -

- Safety monitoring: - Video surveillance camera used in public place for the safety and security of the people may break down due to limitation in the algorithm design of the system. In such scenarios, crowd counting system can used for event detection, congestion control and behavioral analysis.
- Intelligence gathering and analysis: - In malls and airport, depending on the number of people entering or length of queue the counters can be set up so that no human resource is wasted.
- Designing a public place: - Crowd counting system can be used to design public space like mall, stadium, rail tracks etc.

2. RELATED WORK

Cross scene crowd estimation is a difficult task, where no arduous data notations are required for estimating people count of dense crowd scene. Deep convolutional neural network (CNN) classifier is pre-trained to provide crowd count of the dense scene-based crowd density. A new dataset including 108 crowd images with 200000 head notations was introduced to better evaluate accuracy of cross-scene crowd estimation methods. To evaluate the efficiency and reliability of the method experiment was held on already existing datasets i.e, UCSD, UCF_CC_50 and WorldExpo'10 dataset. Cross-scene system fails to provide accurate count of the dense crowd scene [1]. Pedestrian analysis is challenging due to the gesture variation, obstruction, appearance and background clutters. Deep Decompositional network (DNN) classifier was used for parsing crowded images into different human parts such as face, hairs, hands, legs and body. Deep decompositional network together estimates obstructed regions and body parts of person by arranging three hidden layers: obstruction estimation layers, completion layers and decompositional layers. Pedestrian parsing method by DNN provides better accuracy than state-of-art method on crowded images with or without obstruction. The experiment was conducted on large benchmark PPSS dataset for evaluating the efficiency and reliability of pedestrian parsing method by DNN. The DNN system fails to work efficiently in heavy crowded scene [2]. Global regression methods are used for mapping low level features (texture, edge information and segmentation mask) of humans to provide crowd count of the dense scene. The system is evaluated over USCD dataset. The system ignores the spatial information and body structure information of pedestrian, thus fails to provide accurate crowd count of crowded scene [3]. The head is the most visible part from any crowded scene. The head detection is based on advance method of boosted essential features. To reduce a search region a novel point estimator base on gradient adjustment features to identify region similar to the head region from gray scale images. Head detector approach is evaluated on PETS 2012 and Turin metro station datasets. Experiment on these datasets gave good performance of head detector approach for crowd counting. The Head detector method fails to provide good performance in dense scene due to obstruction short people where not visible [4]. Human analysis and detection are the basic problem in any video surveillance system. Shape based pedestrian detection uses support vector machine (SVM) classifier for detecting pedestrian in the crowded scenes. Support vector machine classifier is pre-trained with few gestures of human to separate human and no-human patterns. If test data gesture matches in the train data then pedestrian will be included in crowd count or else not. The Shape based method is evaluated with three public datasets (INRIA, USC-B and MIT-CBCL) and two benchmark datasets (Caviar and Munich Airport). The Shape based system had many misdetections due to the human pose estimation failure [5].

3. METHODOLOGY

Estimation of crowd count of any crowd video taken from the crowded scene involves following image processing steps as show in below figure 1.

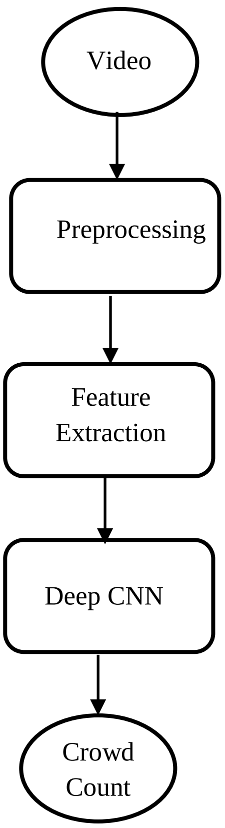

Figure 1 System Architecture of Crowd counting system

3.1 Preprocessing

The crowd video is uploaded, then sequence of frames are captured one by one from the video. The captured frame contains distortion caused by the camera positing relative to object or position of objects in frames. Gaussian blur technique is used to remove distortion and noise from the frames, which results in blurring an input frame by gaussian function.

3.2 Feature Extraction

Feature extraction is image processing techniques, where the input raw data is reduced to manageable features for processing. Motion detection is type of feature extraction method where changing position of object is detected relative to its surroundings. The preprocessed frames are taken for detection humans in crowded video. The humans are detected based on their movements relative to their surroundings.

The motion detector algorithm used to detect the human motion based on pixels movement includes 4 steps:

Step 1: Calculate the Difference between the background_Frame and the current_Frame.
Step 2: The threshold value of the frame calculated in (Step 1) is used to filter the areas of motion.
Step 3: Resulting frame from (Step 2) is then highlighted in the current_Frame to indicate areas of motion.
Step 4: Updated the background.

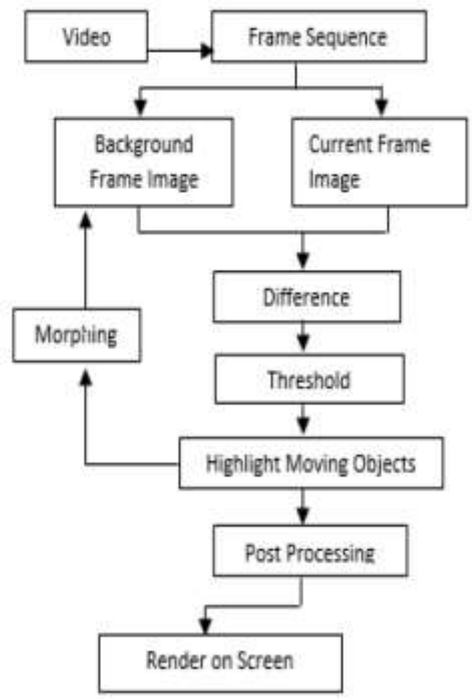

Figure 2 Motion Detector algorithm working

3.3 Deep Convolutional Neural Networks

Deep Convolutional neural network is type of neural network used to classify the human and non-human objects in the dense crowded scene. It is pre-trained with set of videos to detect only human in different pose, then test data is used to check the working of the neural network.

4. RESULT

Result of the proposed crowd counting system is shown in figures below. For crowd counting system crowded video is given as input and as the people starts moving in the video crowd is obtained. The crowd count is estimated in both direction i.e, number of people entering the in and out. The threshold is set, when the person moves from that threshold the crowd count is incremented by one else not. Proposed system works well in both densely and low occulated crowded regions.

Figure 3: Crowd counting method result for colored video

Figure 4: Crowd counting method result for grayscale image

Figure 5: Crowd counting method result for human legs (only)

5. CONCLUSION

In this paper, novel approach is presented to estimate the crowd count in the crowd videos. The input video is pre-processed to remove the distortion by using gaussian blur technique. Gaussian blur method acts as low pass filter, which remove the noise by blurring the input frame by using gaussian function. The motion detector algorithm is used to detect the motion of human with respect to background structure, based on the human motion crowd count of given video is estimated. CNNs classifier provides a better performance than other neural networks classifiers. The crowd count system provides a better performance even with different human pose and illumination conditions. In further, the crowd counting system can be implemented with face recognition algorithm, so that each person is counted only once.

REFERENCES

[1] Z. Lin and L. S. Davis, “Shape-based human detection and segmentation via hierarchical part-template matching,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 4, pp. 604–618, Apr 2010.
[2] V. B. Subburaman, A. Descamps, and C. Carincotte, “Counting people in the crowd using a generic head detector,” in Proc. IEEE Conf. AVSS, pp. 450-470, Sep 2012.
[3] A. B. Chan and N. Vasconcelos, “Counting people with low-level features and Bayesian regression,” IEEE Trans. Image Process., vol. 21, no. 4, pp. 2160–2177, Apr 2012.
[4] P. Luo, X. Wang, and X. Tang, “Pedestrian parsing via deep decompositional network,” in Proc. IEEE ICCV, pp. 2648–2655, Dec. 2013.
[5] C. Zhang, H. Li, X. Wang, and X. Yang, “Cross-scene crowd counting via deep convolutional neural networks,” in Proc. IEEE Conf. CVPR, pp. 833-841, Jun 2015.

[6] B. Wu and R. Nevatia, "Detection of multiple, partially occluded humans in a single image by Bayesian combination of edgelet part detectors," in Proc. IEEE ICCV, vol. 1. Oct. 2005, pp. 90–97.
[7] Y. Zhang, D. Zhou, S. Chen, S. Gao, and Y. Ma, "Single-image crowd counting via multi-column convolutional neural network," in Proc. IEEE Conf. CVPR, Jun. 2016, pp. 589–597.
[8] V. Lempitsky and A. Zisserman, "Learning to count objects in images," in Proc. Adv. NIPS, 2010, pp. 1324–1332.
[9] W. Ge and R. T. Collins, "Marked point processes for crowd counting," in Proc. IEEE Conf. CVPR, Jun. 2009, pp. 2913–2920.
[10] M. Wang and X. Wang, "Automatic adaptation of a generic pedestrian detector to a specific traffic scene," in Proc. IEEE Conf. CVPR, Jun. 2011, pp. 3401–3408.
[11] A. N. Marana, L. F. Costa, R. A. Lotufo, and S. A. Velastin, "On the efficacy of texture analysis for crowd monitoring," in Proc. IEEE SIBGRAPI, Oct. 1998, pp. 354–361.
[12] K. Chen, C. C. Loy, S. Gong, and T. Xiang, "Feature mining for localised crowd counting," in Proc. BMVC, 2012, vol. 1. no. 2, p. 3.
[13] A. B. Chan and N. Vasconcelos, "Counting people with low-level features and Bayesian regression," IEEE Trans. Image Process., vol. 21, no. 4, pp. 2160–2177, Apr. 2012.
[14] D. Kong, D. Gray, and H. Tao, "A viewpoint invariant approach for crowd counting," in Proc. IEEE ICPR, vol. 3. Aug. 2006, pp. 1187–1190.
[15] A. B. Chan, Z.-S. J. Liang, and N. Vasconcelos, "Privacy preserving crowd monitoring: Counting people without people models or tracking," in Proc. IEEE Conf. CVPR, Jun. 2008, pp. 1–7.
[16] D. Ryan, S. Denman, C. Fookes, and S. Sridharan, "Crowd counting using multiple local features," in Proc. IEEE DICTA, Dec. 2009, pp. 81–88.
[17] N. Paragios and V. Ramesh, "A MRF-based approach for real-time subway monitoring," in Proc. IEEE Conf. CVPR, vol. 1. Jun. 2001, pp. I-1034–I-1040.
[18] M. von Borstel, M. Kandemir, P. Schmidt, M. K. Rao, K. Rajamani, and F. A. Hamprecht, "Gaussian process density counting from weak supervision," in Proc. ECCV, 2016, pp. 365–380.
[19] M. Rodriguez, I. Laptev, J. Sivic, and J.-Y. Audibert, "Density-aware person detection and tracking in crowds," in Proc. IEEE ICCV, Nov. 2011, pp. 2423–2430.
[20] Y. Wang, Y. X. Zou, J. Chen, X. Huang, and C. Cai, "Example-based visual object counting with a sparsity constraint," in Proc. IEEE ICME, Jul. 2016, pp. 1–6.
[21] Y. Wang and Y. Zou, "Fast visual object counting via example-based
density estimation," in Proc. IEEE ICIP, Sep. 2016, pp. 3653–3657.
[22] C. Arteta, V. Lempitsky, and A. Zisserman, "Counting in the wild," in Proc. ECCV, 2016, pp. 483–49.