IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

TEMPLATE BASED CROWD ESTIMATION

[1]Alimineti Surya prakash Reddy, [2]Dr.V.Devendran

[1]Research Scholar, [2]Professor,

[1]Dept. Computer Science & Engineering, [2]School of Computer Science

Lovely Professional University, Phagwara, Punjab, India

Abstract: Monitoring and analyzing crowd in public places plays an vital role in corona pandemic situation in order to prevent the wild spreading of the virus and to ensure the public safety. For analyzing and estimating the crowd in specific zone through a video/stream produced by a statically fixed camera we used image-processing techniques for this research implementation. Considering the pandemic situation we have used specific techniques in this implementation background subtraction, shadow elimination, segmentation and pixel/texture based image processing methods. With this methodology we get a group wise classified crowd estimations such as low, medium and high crowd. The video source can be inputted by a fixed positioned camera for the best estimation. In this research we only implemented using few logical calculations and image pre-processing techniques in order to develop a cost efficient model and also to ensure the real time processing of a stream for estimating crowd with good accuracy even in low light condition and to produce on time results. This helps in minimizing the human engagement for monitoring and analyzing.

Index Terms - **Background subtraction, Shadow elimination, Gaussian model, Pixel differentiating, Image processing, Intelligent systems**

I. INTRODUCTION

In this pandemic situation crowded zones are hub for the spread of COVID. For the safety of people in this pandemic situation controlling crowd plays highly crucial role. To ensure the public safety the best way is to monitor the places using CCTV streams and make sure of the people density in that particular zone. As there are some places like metro stations, malls, stadiums, super markets etc., where crowd management should be implemented strorngly to avoid the places being highly crowded.

In this paper we proposed a methodology to automatically monitor and estimate the crowd density in the live stream without any human dependency. This methodology gives real time results and cost efficient as it can be used in a low configured systems. In this we input RGB video streams which are produced by a static cameras such as CCTV cameras. This methodology is also focused to produce the dynamic results in low lighting conditions.

The steps we followed for our methodology are like background subtraction, shadow elimination and pixel differentiating. We have reviewed and bought the best methodologies in order to overcome the challenges faced by the other researchers such as various lighting conditions, unable to produce real time results, getting false positive results, etc., we have catogarized our output crowd estimations into three types as comparing the people per area of zone as

- Highly crowded
- Medium crowded
- Low crowded.

If the area is covered over 70-75% in the size of the given zone then it is considered to be highly crowded or if the area is covered in the range of 40-70% then it is considered as medium crowded area and if it is below 40% then it is considered to be low crowded area.

II. RELATED WORK

In the paper of Zhen Tang, et al.,[1] the authors have shared the techniques to subtract the background and avoiding the shadow getting detected. In this authors have used Gausian filtering and distribution methods and named it as Gaussian Mixture Models(GMM). Here author used Gaussian filters to find out and identify the minute disturbances in the consecutive frame of an video The results may be poor in low light conditions. This technique works perfectly fine for real time purposes.

K. T.Park,et al.,[2] in there research on Unsupervised foreground segmentation using background elimination and graph cut techniques they have shared techniques to detect the foreground objects.To detect foreground regions,authors used image segmentation based on the JSEG algorithm .The JSEG algorithm is composed of two steps. In the first step, a colour quantisation based on peer group filtering (PGF) and in the second step, a spatial segmentation considering the class map is perform. The drawback in this is we cannot use it on videos.

K.Srinivasan, et al.,[3] in there research of improved Background Subtraction Techniques
for Security in Video Applications.Authors have shared the techniques to subtract the background for recurssive and non recurssive situations in video. Authors have tested multiple existing recursive and no recurssive algorithms for background elimination.Authors have developed an automatic threshold updation method for the filters to avoid excusing of lighting conditions and used differencing method after comparing with all existing techniques.The method they created may not give the perfect object segmentations and cannot distinguish the background and for ground color finitely.

Neha.S,et al.,[4] in there research on Adaptive Background Subtraction In Images. Author used segmentation masking techniques to achieve the goal and tried to separate the foreground part background part separately by segmenting and then differentiated on the basis of color and eliminated the selected background pixels of image.This method may not detect perfect foreground objects and cannot produce real time results.

In the paper of Rita Cucchiara.et al.[5], Authors have proposed techniques to identify moving objects, ghosts and shadow in video. Authors used segmentation masking and color saturation techniques in there work. Authors firstly identified the noise difference in the consecutive frames in a video and the used segmentation techniques to find the exact moving object and they have used color saturation HSV technique to differentiate the shadow and ghost from moving object.The complexity of the proposed algorithm is high so it may not give realtime results.

In the paper of CAO Lijun,et al.,[6] Authors have proposed a video-based crowd density analysis and prediction system for wide-area surveillance applications. Authors have applied Accumulated Mosaic Image Difference(AMID) technique to find the crowded areas by finding irregular motions The proposed method can adequately estimate specific number of people from the density of the crowded areas. Authors have used motion detection to find the path of individuals in a crowd to predict and analyse the crowd in next zones.The method which authors used is computationally complex but can work in long range detection.

In the paper of Sami Abdulla Mohsen Saleh,et al.,[7] proposed two techniques direct based and indirect based where both the methids have some issues. In direct based which is also know as detection based where we can proceed with model based analysis and clustering based analysis as per the author. And in indirect based which is also known as Feature based can be proceed by pixel based analysis, texture based analysis and corner points based analysis.

Xiaohang Xu, et al.,[8] in there research on Crowd Density Estimation of Scenic Spots Based on Multifeature Ensemble Learning They have used indirect methods and feature extraction methods like D-SIFT features extraction,ULBP features extraction and GIST features extraction and linked these all methods to ensemble learning to estimate the people count and the density. In the place of ensemble learning SVM's has been used.

Muhammed Anees V, et al.,[9] in there research on Deep Learning Framework for Density Estimation of Crowd Videos.They have firstly used Inception model for feature extraction and the they have used LSTM's for classifying. The proposed model is good for the accuracy but as coming to the real time output it fails beacuse of its highly complex methods like usage of Inception and LSTM combined.

III. METHODOLOGY

For crowd density estimation, we have used the following steps.

- Background Subtraction and Shadow Elimination technique.
- Feature extraction and object detection techniques.
- Feature based density estimation

a. Background Subtraction and Shadow Elimination technique.

We used Background subtraction to detect objects in the frame.There are few different methods to detect moving objects like,
> Optical flow method,
> Consecutive frame differencing
> Foreground extraction.

In optical-flow method, it extracts the features by clustering technique as it utilizes x-mean cluster for classifying the extracted feature points and segments the region as states as a moving object. But this technique is not suitable for real time processing as it has high complexity.

In consecutive frame differencing, difference between two consecutive frames is compared and eliminates the constant elements to provide the foreground elements. Here in this method we can obtain real time results as it is not a complex structure.

In foreground extraction, it is a technique where approaching outlines are contrasted and the foundation model and the moving objects will be distinguished. The drawback in this is that it give poor results where background is not constant.

As compared to all the existing Background subtraction methods we concluded to use frame differencing technique as it gives good results and process in real time which are the suitable regulations for the output we expected.

In the frame pre-processing step we have applied color conversion technique to change colored image(RGB) to grayscale image. For denoising the image we have applied binary threshold to the image.

In the output frame the detected moving objects are segmented and localized with the minimal distance to get a precise results.

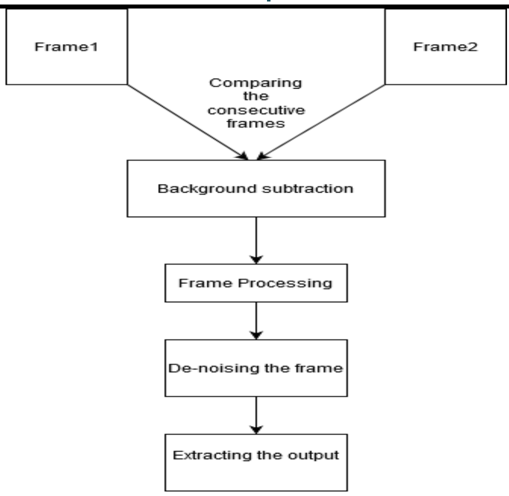

Fig1. Flowchart of Background subtraction

The challenging part while background subtraction was to avoid detecting shadows which were leading to false positive case. For the purpose of shadow elimination we have used gaussian models in the preprocessing steps.

Gaussian Model

We used Gaussian blurring technique to get the nearest summarized pixels and decrease the complexity of pixels of a frame. Therefore:

$$G_{i,j} = a * e^{-(x^2+y^2)/2*\sigma^2}$$

Eq1. Gaussian function for one dimension format

$$G_i = \alpha * e^{-(i-(ksize-1)/2)^2/(2*\sigma^2)}$$

Eq2. Gaussian function for two dimension format

Using the equations (Eq1/Eq2) we can calculate the weight of each pixel point of a matrix by using that weight matrix, we can calculate the gaussian blur value.As it is done by multiplying each point value with its weight value

Therefore:

$$B_{i,j} = w_{i,j} \times G_{i,j}$$

Where,

$w_{i,j}$ = point in matrix

$G_{i,j}$ = point in weigh tmatrix

Thresholding is a type of image segmentation, where we change the pixels of an image to make the image easier to analyze. We used binary thresholding to in order to get a smooth output.

$$\mathrm{dst}(x,y) = \begin{cases} \mathrm{maxval} & \text{if } \mathrm{src}(x,y) > \mathrm{thresh} \\ 0 & \text{otherwise} \end{cases}$$

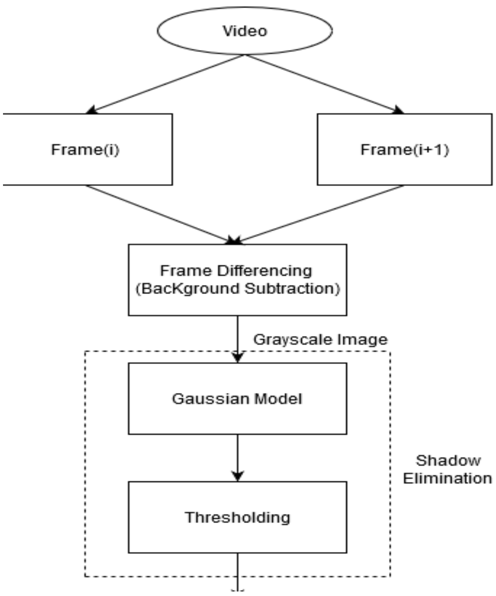

Fig2.Flowchart represents steps for shadow elimination

After thresholding we get a clean and denoised frames detecting moving substances. These detected substances are segmented and localized with minimal distance.

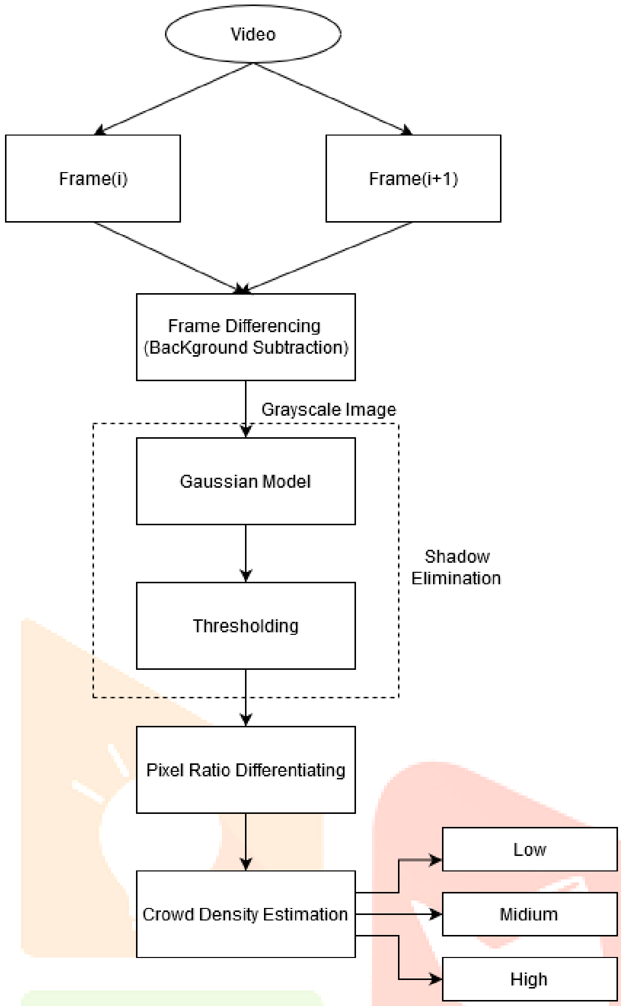

Fig3.Flowchart represents steps for crowd estimation into three categories.

After background subtraction and shadow elimination we have segmented the moving objects and localizes with minimal distance points.

Fig4. Image represents the threshold output which is produced after Background and shadow elimination process.

The threshold image shows 0's and 255's pixels as represented in Fig4. Only 0 and 255 valued pixels are bought up in the frame, where 0's represent blank pixels and 255's represent white pixels(objects).

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	255	255	255	0	0	0	0
0	0	255	255	255	255	0	255	255	0
0	0	0	255	255	255	0	255	255	255
0	0	0	255	255	255	0	255	255	0

Fig5. Image represents an example of threshold images which is obtained after Background and shadow elimination process.

On the basis of fig5 we have used pixel differentiating technique to obtain the area of density of that objects detected.

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	255	255	255	0	0	0	0
0	0	255	255	255	255	0	255	255	0
0	0	0	255	255	255	0	255	255	255
0	0	0	255	255	255	0	255	255	0

Fig6. Image represents the localized edges of object pixels

In fig6 the disturbed or object pixels are localized to calculate the difference between object and non object pixels to get the percentage of objects covered value compared with empty spaces.

As an example let us consider Fig5 and Fig6 the total matrix size is 6*10 i.e., 60 pixels in a frame. In that we differenced the blank and disturbed pixels separately to estimate the crowd percentage covered in the frame/zone.
Formula to calculate the percentage:

$$x_d / x_b \times 100$$

$x_b \rightarrow$ no.of blank pixels

$x_d \rightarrow$ no.of disturbed pixels

IV. Results and Disscussion

We used few cctv footage videos to evaluate the performance of our model
Background subtraction:
Initial step of the work begins with the background subtraction. Here are the threshold and RGB output screenshot results of absolute frame differencing.

Fig7. Threshold output after background subtraction

Fig8. Output after background subtraction

Gaussian model: To precisely detect the objects we used gaussian blurring method as to remove the noisy elements in order to eliminate the shadows. Here are the output screenshot results after applying gaussian model to the background subtracted frame.

Fig9. Threshold output after applying Gaussian model to Background subtraction frame

Fig10. Output after applying Gaussian model to Background subtraction frame

Here are some results of the proposed model on supermarket cctv footages. Where the crowd estimations are classified into Low, Medium and high.

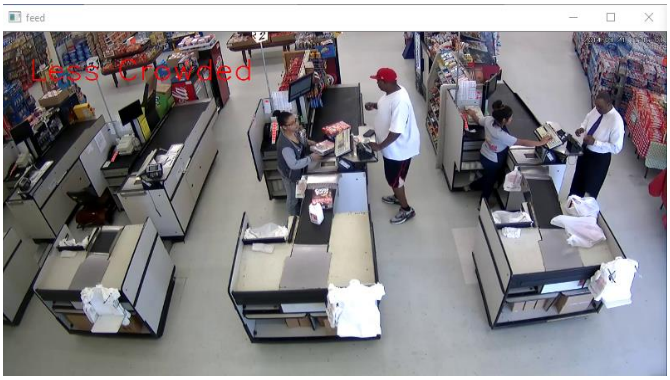

Fig11. Result of our model indicating the zone is less crowded

Fig12. Threshold image of Fig11

Fig13. Result of our model indicating the zone is highly crowded

Fig14. Threshold image of Fig13

IV. CONCLUSION AND SUMMARY

Automatic monitoring and real time crowd estimation is the main aim of this research.We have surveyed the existing methods and techniques[6]-[14] and added our methodologies to those techniques and made a model that produces real time frames processing. We have choose absolute frame differencing for background subtraction and binary threshold, guassian blurring model for shadow elimination. For density estimation we used pixel differencing method. All these techniques are low computational and leads in producing real time and precised results.

REFERENCES

[1] Zhen, T., & Zhenjiang, M. (2007). Fast background subtraction and shadow elimination using improved Gaussian Mixture Model. *HAVE 2007 - The 6th IEEE International Workshop on Haptic, Audio and Visual Environments and Games, Proceedings*, (October), 38–41.

[2] Park, K. T., Lee, J. H., & Moon, Y. S. (2009). Unsupervised foreground segmentation using background elimination and graph cut techniques. *Electronics Letters*, *45*(20), 1025–1027.

[3] Srinivasan, K., Porkumaran, K., & Sainarayanan, G. (2009). Improved background subtraction techniques for security in video applications. *2009 3rd International Conference on Anti-Counterfeiting, Security, and Identification in Communication, ASID 2009*, 114–117.

[4] Sakpal, N. S., & Sabnis, M. (2018). Adaptive Background Subtraction in Images. *2018 International Conference On Advances in Communication and Computing Technology, ICACCT 2018*, 439–444.

[5] Cucchiara, R., Grana, C., Piccardi, M., & Prati, A. (2003). Detecting moving objects, ghosts, and shadows in video streams. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*(10), 1337–1342.

[6] Cao, L., & Huang, K. (2013). Video-based crowd density estimation and prediction system for wide-area surveillance. *China Communications*, *10*(5), 79–88.

[7] Saleh, S. A. M., Suandi, S. A., & Ibrahim, H. (2015). Recent survey on crowd density estimation and counting for visual surveillance. *Engineering Applications of Artificial Intelligence*, *41*, 103–114.

[8] Xu, X., Zhang, D., & Zheng, H. (2017). Crowd Density Estimation of Scenic Spots Based on Multifeature Ensemble Learning. *Journal of Electrical and Computer Engineering*, *2017*(ii).

[9] Anees, M. V., & Kumar, S. G. (2018). Deep Learning Framework for Density Estimation of Crowd Videos. *Proceedings of the 2018 8th International Symposium on Embedded Computing and System Design, ISED 2018*, 16–20.

[10] Bhatia, S. K. (2018). *Advances in Intelligent Systems and Computing 924 Advances in Computer Communication and Computational Sciences*.

[11] Ahuja, K. R., & Charniya, N. N. (2019). A Survey of Recent Advances in Crowd Density Estimation using Image Processing. *Proceedings of the 4th International Conference on Communication and Electronics Systems, ICCES 2019*, (Icces), 1207–1213.

[12] Nemade, N. A., & Gohokar, V. V. (2017). A survey of video datasets for crowd density estimation. *Proceedings - International Conference on Global Trends in Signal Processing, Information Computing and Communication, ICGTSPICC 2016*, (December), 389–395. Kishore, P. V. V., Rahul, R., Sravya, K., & Sastry, A. S. C. S. (2015). Crowd Density Analysis and tracking. *2015 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2015*, 1209–1213.

[13] Khan, A., Shah, J. A., Kadir, K., Albattah, W., & Khan, F. (2020). Crowd monitoring and localization using deep convolutional neural network: A review. *Applied Sciences (Switzerland)*, *10*(14). https://doi.org/10.3390/app10144781