IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 6, June 2016

Counting people by extracting low level features using Bayesian Regression

Rajshri S Vaswade[1], Prof. N.A. Dawande[2]

Student, Dept Electronics &Telecommunication, D.Y. Patil College of Engg, Ambi, Pune[1]

Associate Professor, Dept Electronics & Telecommunication, D.Y. Patil College of Engg, Ambi, Pune[2]

Abstract: Counting people is a crucial component in visual surveillance mainly for crowd monitoring and management. Now days, significant improvement has been made on the field by using features regression. On this context, perspective distortions have been frequently studied; however, crowded scenes remain particularly challenging and could extremely affect the count because of the partial occlusions that take place between individuals. To overcome this challenge this paper proposes Gaussian Mixture Model (GMM). It uses background subtraction to obtain highly accurate foreground segmentation for people counting approach. It leads to integrating an uniform motion model. The counting is established on foreground measurements, where perspective normalization and a crowd counting measures the density of a crowd with foreground pixel counts into a single feature. Afterwards, the correspondence between this frame-wise feature such as internal features, segment features, texture features are extracted and these extracted features are applied to the Bayesian regression. Then count is estimated in each frame.

Keywords: Bayesian regression, crowd analysis, Gaussian Processes

I. INTRODUCTION

Recently there is great interest in the video technology for observing all kinds of environment. It may have many purposes such as security, resource management. The application of this paper is to detect and track the people in a homogeneous crowd, which is composed of pedestrian that travels in a different direction without spending clear object separation or tracing is projected. In this the crowd is segmented into the sections of analogous motion by using the grouping of active texture motion model.

A staticentire low level features is removed from every separated region and maps the features into evaluates the variety of human beings according to section is found out with Bayesian Regression. In this two Bayesian regression fashions are located. It is a combination of Gaussian method regression with a compound kernel which explains the both the global and nearby developments of count.Mapping is restricted by real valued outputs that do not healthy with a discrete counts.

There is no need for pedestrian detection, object monitoring, or object-based image. Primitives to achieve the pedestrian counting aim, even when the gang is great and inhomogeneous, e.g., has subcomponents with exclusive dynamics and appears loose in out of doors environments, as proven in fig 1.In reality, based on crowd centric approach the hassle seems easier.
On this virtually crowd segment into the vicinity of interest, extract a fixed of complete features from the every section and calculates the group length with appropriate regression function. By means of warding off in-between processing degrees together with human beings detection or monitoring, which are prone to occlusion trouble.

Fig.1 Scenes containing a sizable crowd with inhomogeneous dynamics due to pedestrian motion in different directions.

II. BLOCK DIAGRAM OF PROPOSED SYSTEM

The block diagram of proposed system is as shown in figure 2 below.

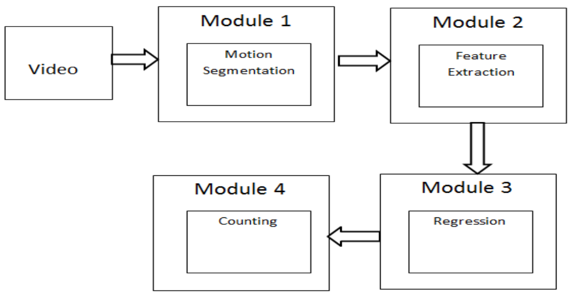

Fig.2 Block diagram of proposed system

This block diagram is consisting of four modules they are motion segmentation, feature extraction, regression and counting respectively. An outline of the proposed crowd counting device is proven in the video is first segmented into crowd regions shifting in different directions. Features

 DOI 10.17148/IJARCCE.2016.56121

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 6, June 2016

are then extracted from every crowd phase; ultimately, the quantity of human beings in step with phase is predicted from the characteristic vector, the use of the BPR module.

A. MODULE 1:
Input is from which the total number of people is counted s in motion. And the output is the video is segmented into the various frames.

B. MODULE 2:
Input is motion segmented frames.After applying canny edge map we get the output is edge pixel of extraction image.

C. MODULE 3:
Input is extracted image; apply the Gabor filter for removing the noise. Output is various features such as entropy, homogeneity, energy is calculated.

D. MODULE 4:
Input to this module is the features are extracted from the image frame .Output is count of people is estimated by using Bayesian regression.

III.IMPLEMENTATION METHODOLOGY

The input to the system is video, after the reading the video this video is converted into the number of frames with the help of motion segmentation, perspective normalization and Gaussian mixture model. After converting the frame conversion apply morphological operation such as dilation, erosion, opening and closing due to these operations image quality is improved. Simultaneously by applying canny edge detector, shape descriptor and features are extracted which includes segment, internal edge and texture features. Once the features are extracted apply Gabor filter due to this noise is removed from the image frame. After that all these features are applied to the Bayesian regression which is based on bays theorem by calculating probability finally count is estimated which is explained in the following sections. The algorithm of proposed system is as shown in figure 3 shown below.

A. Motion Segmentation
This is the first step in the system in which it segments the scene into the crowd subcomponents of interest. The main aim is to count people moving in different direction with different speed. It consists of mixture of dynamic texture to segment the crowd into subcomponents of different motion flow. The video is characterized as group of spatiotemporal areas, which are displayed as independent samples from a combination of dynamic textures. Video locations are then scanned serially; a patch is removed at each location and allocated to the mixture component of the biggest subsequent probability.

B. Perspective Normalization
The effect of perspective is taken into consideration for the extraction features from the crowd segment due to the fact that objects closer to the camera look like larger, any pixels related to a near foreground object account for a smaller portion of it than those of an item farther away. This could be compensated by normalizing for attitude in the course of feature extraction in this document; each pixel is weighted in keeping with attitude normalization map, that's primarily based on the predicted intensity of the object that generated the pixel. Pixel weights encode the relative size of an item at distinctive depths, with large weights given to far gadgets.

C. Gaussian Mixture Model
It is common method which is used for real time segmentation of moving regions in image sequences includes background subtraction, or thresholding the error between an estimate of the image without moving objects and the current image.

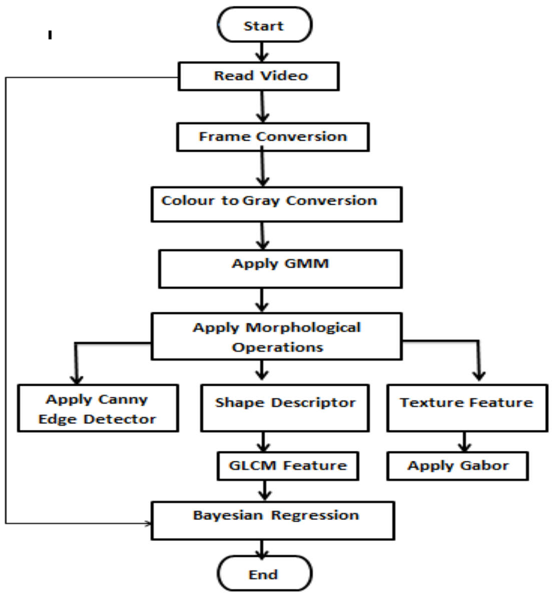

Fig3Flowchart of proposed system

D. Morphological Operations
In this step in order to improve the image quality apply the various morphological operations such as dilation, erosion, opening and closing.These operations are applied only for the binary images only.

E. Features Extraction
In this section it includes various features such as segment area it may vary linearly with the wide variety of human beings in the scene. Whereas the general trend is certainly linear, neighbourhood nonlinearities rise up from a variety of factors, inclusive of occlusion, segmentation mistakes, and pedestrian configuration. To model those nonlinearities, additional features, which can be mostly based on segment form, edge data, and texture, are extracted from the video. When computing capabilities primarily based on area or length, every pixel is weighted by using the corresponding value inside the perspective map. When the features are based totally on edges every edge pixel is weighted via the square root of the angle map cost.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 6, June 2016

1) Segment Features
In order to capture the segment properties such as shape and size these features are extracted. The features such as segment perimeter, area, perimeter orientation etc.are also extracted.

2) Internal Edge Features
The edges inside a crowd segment are a robust clue about the quantity of humans in it. The output is masked to form the internal edge image and numbers of features are extracted by applying Canny edge detector to the image. It includes the features such as edge length, edge orientation.

3) Texture Features
Texture features, which can be primarily based on the grey-level co-occurrence matrix,features such as energy, entropy, homogeneity, contrast etc. are extracted from the image is calculated with the help of GLCM features. The image is first quantized into eight gray degrees and masked by means of the segment.

F. Canny Edge Detector
Canny edge detector is used to find out edges of the images which is significantly reduces the data amount of data and filters out useless information, while preserving the important structural properties in an image. The optimal edge detector is another name of Canny edge detection algorithm.

G. Bayesian Regression
Various features are extracted such as segment features, internal texture features are applied to the Bayesian regression by calculating the probability of all these features finally based on bays theorem, the count is estimated from the image frame.

IV.RESULT AND DISCUSSION

Counting people with low level features with regression the results obtained are shown below. The video is first segmented into crowd regions moving in different directions. Features are then extracted from each crowd segment, after the application of a perspective map. Based on the features extracted from each image frame,finallythe count is estimated.

A. Crowd Segmentation
Step one of the gadget is to segment the scene into the gang subcomponents of interest. That is achieved via first using a mixture of dynamic textures to phase the crowd into subcomponents of distinct movement glide. The video is represented as series of spatiotemporal patches, which are modelled as independent samples from an aggregate of dynamic textures.

Video places are then scanned sequentially; a patch is extracted at each location and assigned to the mixture issue of the most important posterior opportunity. The region is declared to belong to the segmentation region associated with that component. The figure 4.1 shows result for movement segmentation of the image frame containing three peoples in the figure shown below.

B. Feature Extraction
Features such as segment area should vary linearly with the number of people in the scene.

Fig.4.1 Motion segmentation of the image frame containing three people.

a) Segment Features
Features are extracted to capture section residences along with form and length. Functions also are extracted from the segment perimeter, i.e., computed with the aid of morphological erosion. It calculate the section functions such as area, perimeter etc. The figure 4.2 shows the result of finding perimeter of image frame containing three people in the frame is shown below.

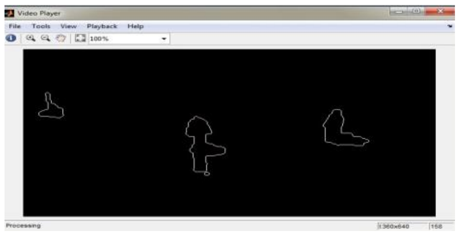

Fig. 4.2 Perimeter of the image frame of the containing three people in the frame.

b) Internal Edge Features:
The aims within a crowd segment are a robust clue approximately the number of people in it. A Canny edge detector is carried out to the image; the output is masked to shape the internal edge length, edge length orientation which estimates the degree of area-filling of the edges.The figure 4.3 shows the result of internal image features for the input image frame containing three peoples in the image is shown below.

Fig.4.3. Internal edge of the image frameof containingthree people in image frame image frame

c) Texture Features
Texture features, which can be based totally on the gray-level co-occurrence matrix. It calculates the functions

 DOI 10.17148/IJARCCE.2016.56121

inclusive of homogeneity, energy, entropy and many others. As an end result, thevalues of homogeneity, energy and entropy are calculated as 1.4309, 0.0507 and 0.9086 respectively.The figure 4.4 shows the calculation for the homogeneity, energy, entropy etc.

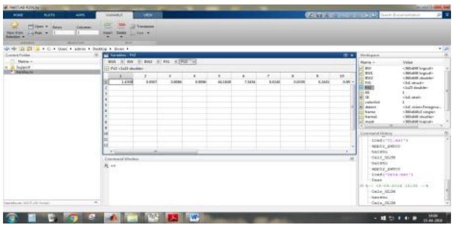

Fig. 4.4 Values for the homogeneity, energy, entropy etc. of the image frame.

The nature of graph energy, entropy and homogeneity vs number of frames is as shown in figure below.

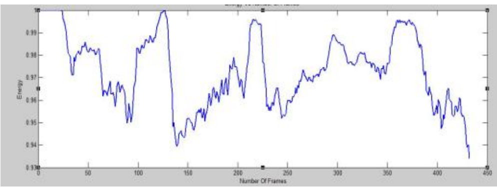

Fig 4.5 Graph of Energy vs Number of Frames.

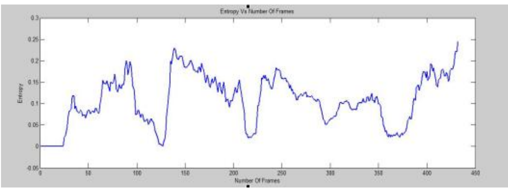

Fig. 4.6 Graph of Entropy vs Number of Frames.

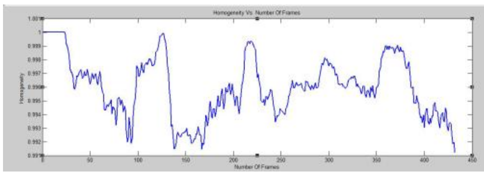

Fig. 4.7 Graph of Homogeneity vs Number of Frames.

C. Bayesian Regression

By applying all the features extracted from the image frame such as segment features, internal edge features and texture features to the regression technique finally the count of people is estimated from the video. As the image frame containing three peoples in the image frame then count is estimated as a result is three and two people in the image frame is shown in the figure 4.8,4.9 respectively shown below.

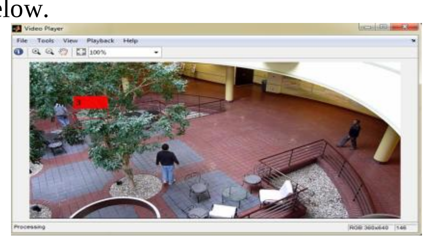

Figure 4.8: Count is estimated from the image frame containing three people in image frame.

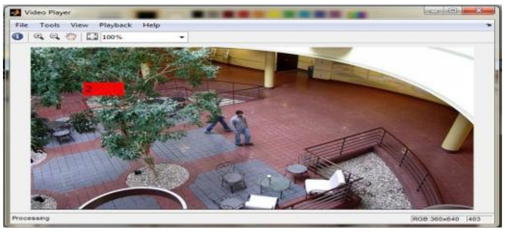

Fig.4.9. Count is estimated from the image frame containing two people in image frame

V. CONCLUSION

This paper depicts by using Bayesian regression technique to calculate the size of inhomogeneous crowds, which are poised of pedestrians traveling in different directions, without using midway vision operations, such as object detection or feature tracking.The count estimated by this method is accurate even if the occlusion occurs. Bayesian regression method is found more accurate for denser crowd whereas Gaussian regression method is suitable for the less dense crowd. Finally it is suggested that this proposed system is suitable for real world environments for long periods of time. Future scope includes, the further improvement to the performance of Bayesian counting from sparse crowds should also be possible.It is also possible study noise models without this restriction.

ACKNOWLEDGMENT

I would like to thank **Prof. N.A Dawande** for his valuable guidance and deep insight.I would like to thank my family members forgiving me support and confidence at each and every stage.

REFERENCES

[1] Antoni B. Chan, and Nuno Vasconcelos,” Counting People With Low-Level Features and Bayesian Regression” vol. 21, no. 4, april 2012.

[2] P. Viola, M. Jones, and D. Snow, “Detecting pedestrians using patterns of motion and appearance,” Int. J. Comput. Vis., vol. 63, no. 2, pp.153–161, 2005.

[3] T. Zhao and R. Nevatia, “Bayesian human segmentation in crowdedsituations,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2003, vol. 2, pp. 459–466.

[4] A. B. Chan and N. Vasconcelos, “Bayesian Poisson regression for Crowd counting,” in Proc. IEEE Int. Conf. Comput. Vis., 2009, pp.545–551.

[5] T. Zhaoand R. Nevatia, “Tracking multiple humans in crowded environment,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2004, pp. II-406–II-413.

[6] Z. Zivkovic, “Improved adaptive Gaussian mixture model for background subtraction,” in Proc. ICVR, 2004, pp. 28–31.

[7] S.-F. Lin, J.-Y.Chen and H.-X. Chao, “Estimation of number of People in crowded scenes using perspective transformation,” IEEE Trans. Syst., Man, Cybern., vol. 31, no. 6, pp. 645–654, Nov. 2001.

[8] B. Leibe, E. Seemann, and B. Schiele, “Pedestrian detection in crowded scenes,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2005, vol.1, pp. 875–885.

[9] A. B. Chan andN.Vasconcelos, “Modelling, clustering, and segmenting video with mixtures of dynamic textures,” IEEE Trans. Pattern Anal.Mach. Intell., vol. 30, no. 5, pp. 909–926, May 2008.