IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 1, Issue 5, Oct-Nov, 2013
ISSN: 2320 - 8791
www.ijreat.org

Human Mass Density Estimation by Using Bayesian Regression

Jyoti R Maranur[1], Rekha B Venkatapur[2]

[1]**Computer Science & Engineering, VTU**
Gulbarga, 585105/Karnataka ,India

[2] **Computer Science & Engineering, KS Institute of Technology**
Banglore,062/Karnataka ,India

Abstract

Human mass density estimation is a system used to estimate the people density walking through a door or specific area. Most of the time, this system is used at the entrance of a building so that the total number of visitors can be recorded. A set of holistic low-level features is extracted from each segmented region, and a function that maps features into estimates the number of people per segment is learned with Bayesian Method. In this paper a system proposed based on face detection using video camera, extract all 256 features, detect and track faces and stored in database and count all extracted faces to avoid multiple faces of same person by using decision tree. People counting system is important for marketing research or in security application. The aim of this work is to make a prototype of a real-time counting people system video based. Experimental results show that the counting accuracy can be achieved above 95% on average.

Keywords: *Face Detection, Bayesian Method, Binary Decision Tree, Motion Detection.*

1. Introduction

There is currently a great interest in vision technology for monitoring all types of environments. This could have many goals, e.g., security, resource management, urban planning, or advertising. From a technological standpoint, computer vision solutions typically focus on detecting, tracking, and analyzing individuals (e.g., finding and tracking a person walking in a parking lot or identifying the interaction between two people). While there has been some success with this type of "individual-centric" surveillance, it is not scalable to scenes with large crowds, where each person is depicted by a few image pixels, people occlude each other in complex ways, and the number of targets to track is overwhelming. Nonetheless, there are many problems in monitoring that can be solved without explicit tracking of individuals. These are problems where all the information required to perform the task can be gathered by analyzing the environment holistically or globally, e.g., monitoring of traffic flows, detection of disturbances in public spaces, detection of highway speeding, or estimation of crowd sizes.

By definition, these tasks are based on either properties of the crowd as a whole or an individual's deviation from the crowd. In both cases, to accomplish the task, it should suffice to build good models for the patterns of crowd behavior. Events could then be de1tected as variations in these patterns, and abnormal individual actions could be detected as outliers with respect to the crowd behavior. we show that there is no need for pedestrian detection, object tracking, or object-based image primitives to accomplish the pedestrian counting goal, even when the crowd is sizable and inhomogeneous, e.g., has subcomponents with different dynamics and appears in unconstrained outdoor environments, such as that of Fig.1

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

In fact, we argue that when a "crowd-centric" approach is considered, the problem appears to become simpler. We simply segment the crowd into subparts of interest (e.g., groups of people moving in different directions), extract a set of holistic features from each segment, and estimate the crowd size with a suitable regression function. By bypassing intermediate processing stages, Face detection or tracking, which are susceptible to occlusion problems, the proposed approach produces robust and accurate crowd counts, even when the crowd is large and dense. In fact, we argue that when a "crowd-centric" approach is considered, the problem appears to become simpler. We simply segment the crowd into subparts of interest (e.g., groups of people moving in different directions), extract a set of holistic features from each segment, and estimate the crowd size with a suitable regression function. By bypassing intermediate processing stages, Face detection or tracking, which are susceptible to occlusion problems, the proposed approach produces robust and accurate crowd counts, even when the crowd is large and dense.

The remainder of this paper is organized as follows, Section 2 contains the related literature on people counting, in section 3 methodology, Section 4 contains face detection and tracking, In section 5 experimental results are presented. Conclusion is presented in section 6.

2. Literature review

In the literature, large amount of techniques can be found with respect to the people counting. Some of the articles related are provided in this section.

P. Viola, M. Jones, and D. Snow, et al (2005), proposed an Detecting pedestrians using patterns of motion and appearance. The detector is trained to take advantage of both motion and appearance information to detect a walking person.

T. Zhao, R. Nevatia, and B. Wu, et al (2008), proposed on Segmentation and tracking of multiple humans in crowded environments. Present a method that can track humans in crowded environments, with significant and persistent occlusion by making use of human shape models in addition to camera models, the assumption that humans walk on a plane and acquired appearance models.

Chan et al (2008) present a system to estimate dynamic crowd size. They segment the crowd into different parts based on motion, after segmentation they extract features from every segmented part. They find correspondences between crowd size and segmented area, and use Gaussian process regression to learn this correspondence.

Chao-Ho Chen1, Tsong -Yi Chen ,et al (2012), proposed on , A Cost-Effective People-Counter for a Crowd of Moving People Based on Two-Stage Segmentation. To obtain a more accurate people count, the two-stage segmentation is developed for extracting each person from a crowd. Firstly, a crowd is segmented by frame-difference technique, followed by morphological processing and region growing. Then, the connected-component labeling method is used to generate many individual people patterns from the segmented crowd. People-image features, such
As the area, heights, and width of each people-pattern, are analyzed in order to correctly segment each person from each individual people-pattern. Finally, each person segmented is tracked till touching the base-line and then is counted. Experimental results show that the counting accuracy can be achieved above 91% on average if the crowd moves normally

3. Methodology

3.1Pre-processing

Preprocessing characteristically deals with techniques for enhancing contrast, segregating regions, and eliminating or suppressing noise. Preprocessing herein includes normalizing the images by bringing them to a common resolution; performing histogram equalization and applying the Gaussian filter to remove small distortions without reducing the sharpness of the image.

3.1.1 Algorithm

1. Input video clip

   Start pre-processing:

2. Obtain digitized image.

$$a(x,y) \rightarrow a[m,n]$$

3. Convert the RGB image to gray scale image.
4. Resize the frame.
5. Filter the Gray scale image using Laplacian of Gaussian.

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x^2+y^2)}{2\sigma^2}} \quad (1)$$

3.2 Frame Acquisition

Acquiring images at video rates of 25–30 frames per second is sometimes referred to as "real time" imaging, but of course this term should properly be reserved for any imaging rate that is adequate to reveal temporal changes in a particular application. For some situations, time lapse photography may only require one frame to be taken at periods of many minutes, hours, or even days. For others, very short exposures and high rates are needed .The main objectives of pre-processing are:

Noise reduction.
Normalization of the data.
Compression in the amount of information to be retained.

3.3 Filtering

The aim of filtering is to remove noise and diminish spurious points usually introduced by uneven text and/or poor sampling rate of the image acquisition device. Various spatial and frequency domain filters can be designed for this purpose. The basic idea is to convolve a predefined mask with the image to assign a value to a pixel as a function of the gray values of its neighboring pixels.

3.3.1 Gaussian Filter

As an effective noise minimization technique, the Gaussian filter is often referred to as a shaped smoothing filter, although the blurring achieved is much less pronounced than the other filters. A typical 3 X 3 Gaussian filter is

• 1	• 2	• 1
• 2	• 4	• 2
• 1	• 2	• 1

Figure 2 Gaussian Filter

Choosing a set of integers to implement a Gaussian is difficult, since the goal is to approximate the smooth analytical curve of the Gaussian. One of the procedures would be to enter kernel weights as real numbers, and generate Gaussian values as needed, when the user enters the standard deviation value.
The smoothing operation can be speeded up considerably in the case of Gaussian filters by separating the operation into two simpler ones. Instead of using the entire square array of weights, which for a 3 × 3 kernel would require 9 multiplications and additions, the filter can be separated into a vertical Gaussian blur with a linear array of weights (3 multiplications and additions) followed by a horizontal Gaussian blur (another 3 for a total of 6). Equation is used to calculate the values for the weights, which again may be scaled and rounded as integers.
1. Dimension Gaussian function

$$G\ (x, y) = \frac{1}{\sqrt{2\pi}\ \sigma} e^{-\frac{(x^2)}{(2\sigma^2)}} \qquad (2)$$

2. Dimension Gaussian function

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x^2+y^2)}{2\sigma^2}} \qquad (3)$$

3.4 People counting Algorithm

Figure describes the proposed people counting algorithm which mainly includes crowd segmentation, person segmentation, and person counting and tracking.

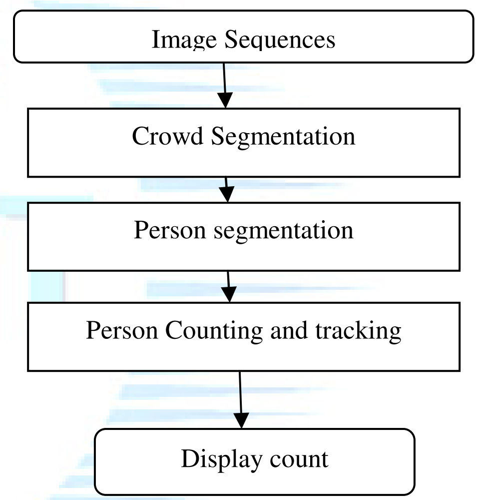

Fig 3 face detection procedure

3.5 Feature Extraction

In principle, features such as segment area should vary linearly with the number of people in the scene. While the overall trend is indeed linear, local nonlinearities arise from a variety of factors, including occlusion, segmentation errors, and pedestrian configuration to model these nonlinearities, an additional 29 features, which are based on segment shape, edge information, and texture, are extracted from the video. When computing features based on area or size, each pixel is weighted by the corresponding value in the perspective map. When the features are based on edges (e.g., edge histogram), each edge pixel is weighted by the square root of the perspective map value.

Features are extracted to capture segment properties such as shape and size. Features are also extracted from the segment perimeter, i.e., computed by morphological Erosion with a disk of radius 1.
3.5.1 Segment feature

1. Area—number of pixels in the segment.
2. Perimeter—number of pixels on the segment perimeter.
3. Perimeter-area ratio—ratio between the segment perimeter and area. This feature measures the complexity of the segment shape: segments of high ratio contain irregular perimeters, which may be indicative of the number of people contained within.
4. “Blob” count—number of connected components, with more than 10 pixels, in the segment.

5.3.2 Edge Features

The edges within a crowd segment are a strong clue about the number of people in it. A Canny edge detector is applied to the image, the output is masked to form the internal edge image, and a number of features are extracted.
1. Edge length—number of edge pixels in the segment.
2. Edge orientation—6-bin histogram of edge orientations.

3.6 Face detection and tracking

Face Detection Procedure as shown Fig 4

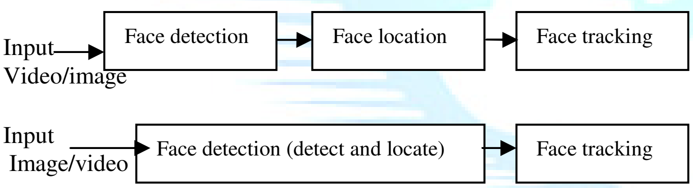

Fig 4. Face detection procedure

Depending on the scenario.

Controlled environment. It’s the most straightforward case. Photographs are taken under controlled light, background, etc. Simple edge detection techniques can be used to detect faces.

Color images. The typical skin colors can be used to find faces. They can be weak if light conditions change. Moreover, human skin color changes a lot, from nearly white to almost black. But, several studies show that the major difference lies between their intensity, so chrominance is a good feature . It’s not easy to establish a solid human skin color representation. However, there are attempts to build robust face detection algorithms based on skin color.

Images in motion. Real time video gives the chance to use motion detection to localize faces. Nowadays, most commercial systems must locate faces in videos. There is a continuing challenge to achieve the best detecting results with the best possible performance. Another approach based on motion is eye blink detection, which has many uses aside from face detection.

Face location is a simplified approach of face detection. Its goal is to determine the location of a face in an image where there’s only one face.

5. Experimental results

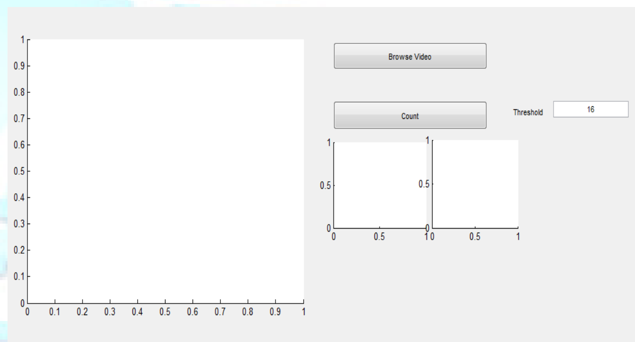

Fig 5 Graphical User Interface

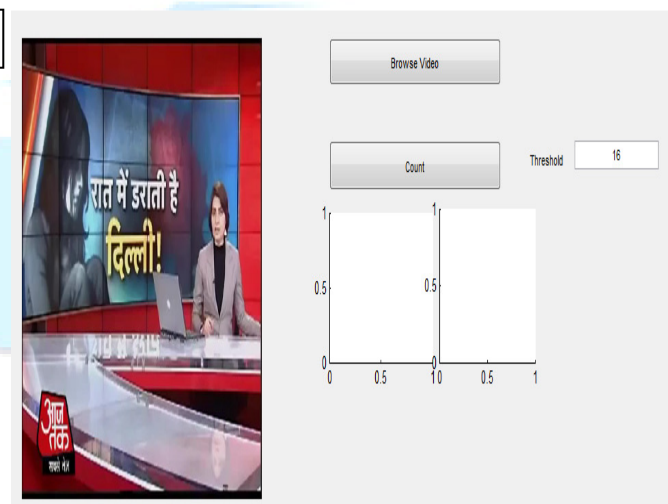

Fig 6 Browse Video/image

Fig 7 Total number of face display on the message box

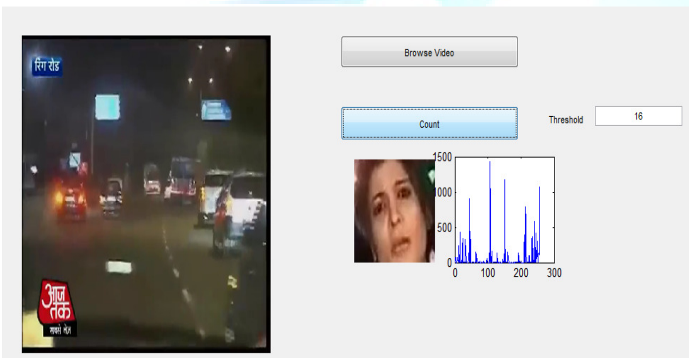

Fig 8 count the total number of faces

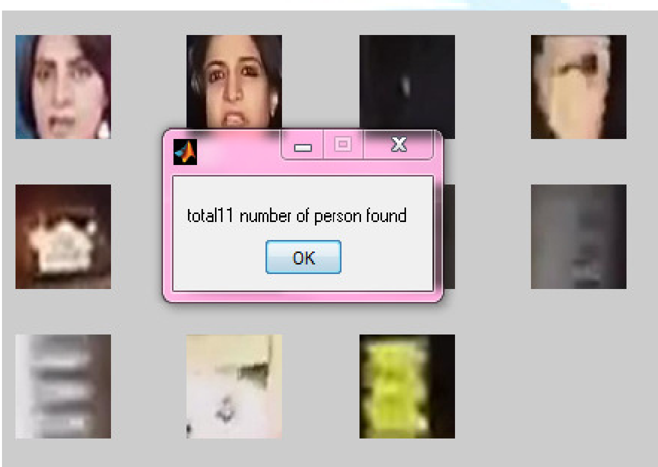

Fig 9 Total number of faces displayed on message box

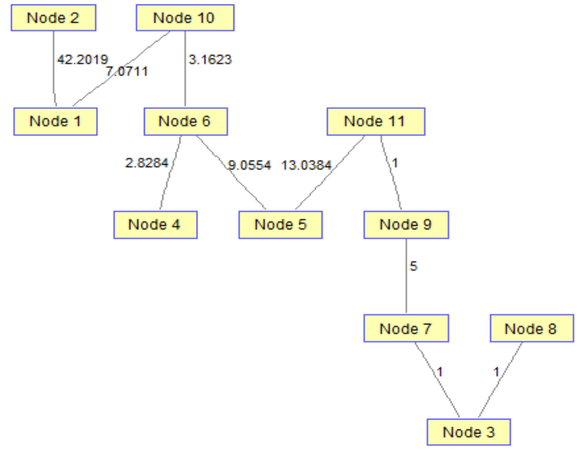

Fig 10 Generate decision tree based on number of faces with distance

6. Conclusions

We have proposed the use of Bayesian method to estimate the size of inhomogeneous crowds, which are composed of pedestrians traveling in different directions, without using intermediate vision operations, such as object detection or feature tracking. The System include image processing, 256 feature extraction, segmentation, pre-processing, face detection. future work is more accuracy then the existing system.

References

[1] A. C. Davies, J. H. Yin, and S. A. Velastin, "Crowd monitoring using image processing," *Electron.* Common. *Eng. J.*, vol. 7, no. 1, pp. 37–47, Feb.1995

[2] N. Paragios and V. Ramesh, "A MRF-based approach for real-time subway monitoring," 2001.

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

[4] B. Wu and R. Nevatia, has proposed work on "Detection of multiple, partially occluded humans in a single image by Bayesian combination of Edge let part detectors," in the year 2005.

[5] L. Dong, V. Parameswaran, V. Ramesh, and I. Zoghlami, "Fast crowd segmentation using shape indexing," in Proc. IEEE Int. Conf. Compute. Vis., 2007, pp. 1–8.

[6] A. B. Chan and N.Vasconcelos, has proposed work on "Modeling, clustering, and segmenting video with mixtures of dynamic textures," in the year May 2008.

[7] T. Zhao, R. Nevatia, and B. Wu, has proposed work on "Segmentation and tracking of multiple humans in crowded environments," in the year Jul. 2008.

[8] Tarek Abd El-Hafeez "A New System for Extracting and Detecting Skin Color Regions from PDF Documents" Vol. 02, No. 09, 2010,

[9] Manuel Grana, Proyecto Fin de Carrera, Face Recognition Algorithms, June 16, 2010

[10]AntoniB.Chan, Member, IEEE, andNunoVasconcelos, SeniorMember, IEEE, Counting people with low-level features and Bayesian regression, 2012