Clustering Based Head Count Using Hough Circle Transform

Avita Verlekar[1] Dr. H.G. Virani[2]
[1]Department of Electronics and Telecommunication Engineering, Goa College of Engineering
[2]Professors, Goa College of Engineering, India
[1]avita.verlekar@gmail.com, [2]hgvirani@gmail.com

Abstract—*In this paper we estimate the head count of people based on the Hough Circle Transform (HCT) algorithm .The paper also proposes the use of clustering to minimize the errors in head count. As opposed to frame differencing, this paper focuses on the use of background subtraction method which has the advantage of detecting stationary objects in addition to moving objects. Experimental results indicate that the proposed method for head count detection gives an average error of 0.220 per frame.*

Keywords-Hough Circle Transform (HCT), Human detection, Head Count,Clustering

I. INTRODUCTION

With increasing security awareness, cameras have become ubiquitous. Human detection and crowd counting is required for public safety and security. In large crowds, tragediesoften occur during religious, political and musical events. Thus many surveillance systems demand automated detection of humans in crowds.

The paper is organized as follows: In section II we discuss some of the related work in this field; in section III an overview of the method adopted is discussed giving a detailed description of the algorithm. In section IV the results of using the proposed method are indicated. Section V includes the conclusion and future work.

II. RELATED WORK

Different approaches have been proposed for estimating the human count from images and videos. Background subtraction, which is used to segment humans from background, usually forms an integral part of most approaches.

One approach of human detection and counting is by segmenting the foreground into individuals [4].Zhao et al. (2008, 2004) use previously discussedhead detection technique to generate initial human hypotheses. They use a Bayesian framework in which each person is localized by maximizing a posterior probability over location and shapes, to match 3D human shape models with foreground blobs. They handle the inter-object occlusion in 3D space.

In [6, 7], SURF features are extracted and only thosewhich are in motion are considered .The number of SURF features are mapped to the people count using a regression function. Another method involves exhaustively searching images using a scanning window. Human or non-human classification of each window depends on the colour, shape or motion features.Texture-based methods such as histograms of oriented gradient (HOG) use edge features and support vector machine (SVM) to detect humans [5]. In [8] a region based tracking approach using skeleton graph is used for head detection. The skeleton graph is extracted from the foreground mask obtained using background subtraction. An articulated appearance model is made which is used to identify people in complex scenarios. In [9], weights are given to feature response belonging to the object such that the sum of feature weights extracted froma full object sums to one. A partial count to the object is given if the object is visible partially.

III. OVERVIEW OF METHOD

Fig. 1: Process Flow

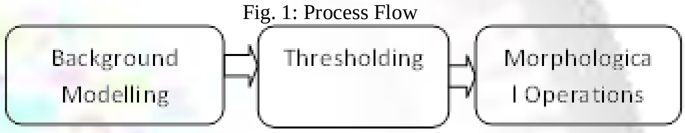

Fig. 2: Pre processing

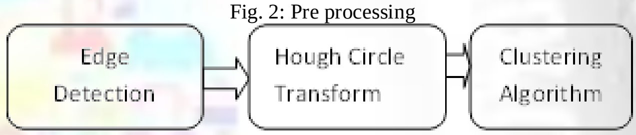

Fig. 3: Human Head Detection

The proposed system consists pre-processing of the video input followed by canny edge detection and clustering based HCT algorithm for head detection. Finally a count of the total number of humans present in the scene is obtained. The following subsections describe each of the blocks in detail.

A. *Pre processing*

The histogram equalized image is subjected to background subtraction followed by thresholding. Thus we can extract the foreground of the segmented image for further processing. Morphological operations are done on the frames to get quality input for edge detection.Parameters of Canny edge detector are adjusted for optimum results.

B. *Hough Circle Transform(HCT)*

Head detection is done using Hough Circle Transform (HCT) [10] with some modifications. Head count is used for detecting humans because head is the most distinguishable part of human body. Also, in further applications, in crowded scenes, head detection is one of the suitable methods for crowd counting.

HCT relies on 3 parameters namely x and y center coordinates and radius. For each edge point, a circle is drawn with that point as origin and radius r. The HCT uses a 3D array (accumulator) with the first two dimensions representing the coordinates of the circle and last third specifying the radii. The values in the accumulator array are increased every time a circle is drawn with desired radii over every edge point. The accumulator (which keeps count of how many circles pass through coordinates of each edge point) proceeds to a vote to find the highest count. The coordinates of the centre of the circles in the image are the coordinates with the highest count.

The Hough circle transform (HCT) works well even under partial occlusion, overlapping effects and breakages, as well as boundaries distorted by interference noise, ambient illumination, and object motion.

C. *Clustering Algorithm*

Initialization:

HCT performs head detection for a range of radii within each frame. All detected centres are stored in Znew.

Znew = [xnew, ynew]

z =first element of Znew

Initialise i =1

Loop until all i < length (Znew)

Iterate

a) Compare z with all elements in Znew. Find points that are within a certain distance from selected element and group them as one cluster.
b) Calculate the average value of a cluster and store value in first location of Znew.Delete all other values belonging to this cluster.
c) Repeat this till no further clusters can be formed.
d) i=i+1

Output

The number of clusters formed and their centre locations.

IV. EXPERIMENTAL RESULTS

Analysis was performed on self-created video of moving people using static camera placed at sufficient angle so that heads are visible. Video frame rate is 29frames/sec. Fig. 6 shows the edge image with the 'x' symbol denoting the head region. Results indicate correct detections for people in the frames.

Fig. 4: Input Frame

Fig. 5: Histogram equalized image

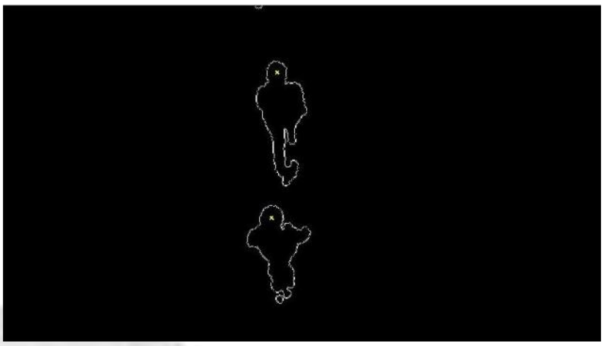

Fig. 6: Frame 37, Human Count=2

Presently some false detection of other body parts matching the Hough radius is contributing to the error (Fig 7). Thus results showed an average error of 0.22 per frame.

V. CONCLUSION AND FUTURE WORK

A method of clustering based Hough Circle Transform for human detection and counting has been proposed in this paper. Experiments performed on self created video validate the effectiveness of our approach. The proposed method achieves an average error rate of 0.220 per frame. Future work focuses on segmenting head region from human body in order to reduce false detections and improve accuracy of head count.

Fig. 7: Frame 6, Gray scale Image

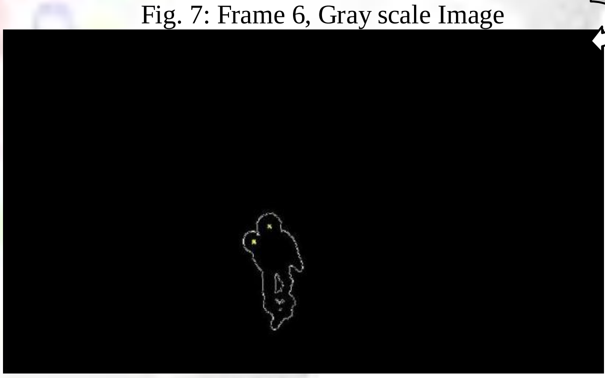

Fig. 8: Frame 6, False Detections: Human Count=2

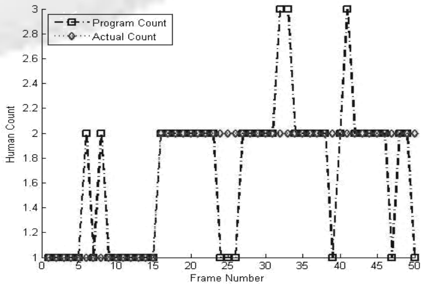

Fig. 9: Crowd Count v/s Frame Number
Average Error: 0.220 per frame

REFERENCES

[1] Detecting people using Hough Circle Transform in surveillance video, VISAPP 2010 - International Conference on Computer Vision Theory and Applications, Hong Liu, YueliangQian and Shouxun Lin.

[2] Yuen, H. K., Proncen, Illingworth, J., Kittler, J., 1990, Comparative Study of Hough Transform Methods for Circle Finding. Image and Vision Computing, Vol.8,No.1, 71-77.

[3] Peter Bone (2010) Matlab Central Exchange [Online] Available: www.mathworks.in/matlabcentral/fileexchange/9833-hough-transform-for-circle-detection/content/houghcircle.m

[4] T. Zhao, R. Nevatia, and B. Wu, "Segmentation and Tracking of Multiple Humans in Crowded Environments," *PAMI*, vol. 30, pp. 1198-1211, 2008.

[5] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *CVPR*, 2005, pp. 886893.

[6] D. Conte, P. Foggia, G. Percannella, F. Tufano, and M. Vento, "An effective method for counting people in video-surveillance applications." in VISAPP, 2011, pp. 67– 74.

[7] D. Conte, P. Foggia, G. Percannella, and M. Vento, "A method based on the indirect approach for counting people in crowded scenes," in AVSS, 2010, pp. 111–118.

[8] K. E. Aziz, D. Merad, B. Fertil, and N. Thome, "Pedestrian head detection and tracking using skeleton graph for people counting in crowded environments," in 2th IAPR Conference on Machine Vision Applications (MVA2011), 2011.

[9] V. Lempitsky and A. Zisserman, "Learning to count objects in images," NIPS, 2010.

[10] Neeti A. Ogale ,"A survey of techniques for human detection from video" .