PEDESTRIAN DETECTION IN A VIDEO SEQUENCE USING HOG AND COVARIANCE METHOD

Bhuvanarjun. K. M[1], T. C. Mahalingesh[2]

[1]*M. Tech Student, Siddaganga Institute of Technology, Tumakuru, Karnataka (India)*

[2]*Associate Professor, Dept. of Electronics and Communication, Siddaganga Institute of Technology, Tumakuru, Karnataka (India)*

ABSTRACT

Visual surveillance is one of the most frequently admired areas of research in the field of image processing and computer vision. The concept of object detection and recognition has been explored rigorously in the last few decades. Object detection is a key step in any smart visual system because it provides the primitive information for logical perceiving of the video footage. The significance of object detection and recognition has been comprehended well in the fields of Visual Surveillance, Human-Computer Interaction, Advanced Driver Recognition System, Medical Imaging and many more. There exists a great deal of problems for which there exists no optimal solution, so we have made an effective effortin the same direction for obtaining the acceptable result. To develop a detection algorithm using various image processing techniques is the main aim of this paper. A method is proposed to improve the detection rate in videos using combination of algorithms. The algorithms used here are Histogram of Oriented Gradients (HOG) and Covariance based method. The proposed method shows improvement in detection time andthe detection rate as well.

Keywords*: *Covariance, Histogram Of Oriented Gradients (HOG), Image Processing,Pedestrian Detection, Visual Surveillance.

I. INTRODUCTION

Object detection is a key component for many video surveillance applications such as Advanced Driver Assistance System (ADAS), Human- Computer Interaction (for gesture recognition to implement smart home and offices), and Visual Surveillance (for recognition of cautious and distrusting activities), Medical Imaging (tumor detection and analysis of patients) and many more. It is necessary to make video surveillance system smart by making them fast, robust and reliable by developing suitable algorithms which provides the highest degree of solution. Also the latest innovations in the field of computing power, data storage and processing have made the arena of visual surveillance most interesting. There are number of problems which are associated with object and/or pedestrian detection such as non-rigid object structure, changing appearance of the object and the background, occlusions occurring due to other objects as well as itself and motion of the camera itself. Also the variation in appearance which is because of geographic variability, various poses, illumination and scaling of objects causes the additional challenges. The primitive factor which needs to be considered in such scenarios is the computation and detection time. Thus understanding and comprehending the scenes in a video is challenging problem which requires scientific innovation and statistical analysis and is a very fruitful domain with serious real world applications.

Object detection is fundamentally recognition of non – stationary objects in a video sequence i.e., basically identifying a moving object. This paper addresses the issue of object detection in occlusion and low illumination scenarios. The system adopted by combining HOG and Covariance based detection gives satisfactory results in pedestrian detection. The proposed method of detection is evaluated on standard database as well as the database obtained locally.

The rest of the paper is organized in the following manner: Section 2 presents the previous work carried out in the field of object detection using various algorithms. Section 3 provides information on the proposed method using combination of algorithms and Section 4 gives the detailed description of feature extraction, HOG and Covariance method. Experimental results are presented in Section 5.

II. LITERATURE SURVEY

Several algorithms have been published for the purpose of pedestrian detection in the past decades, which have provided the roadmap in the visual surveillance domain.

In [1] Na Shou et al., have proposed a region of interest based pedestrian detection system, which reduces the overall size of the image thus reducing the detection time and improving accuracy. Navneet Dalal et al.,[2] have proposed a method called as Histogram of Oriented Gradients(HOG) for pedestrian detection. HOG features are computationally efficient and no feature has been show to outperform HOG in pedestrian detection [3, 4, 5]. The combination of ROIs based HOG makes the algorithm efficient computationally and real time as well. Oncel Tuzel et al. have proposed region descriptor method which can be applied for object detection and classification, known as Covariance based method[6]. These covariance matrices are combination of various image features which consists of statistical and spatial information [7, 8]. Pedro Cortez-Cargill et al.,[9] have discussed the accuracies of various feature vectors used for calculating the covariance matrix and have suggested the features which reduces the computation time and thus resulting in better detection system. In [10], the authors have combined HOG and Optical flow for detection and tracking of humans thus emphasizing the scope of combination of algorithms.

III. OVERVIEW OF THE PROPOSED DETECTION SYSTEM

The overview of the proposed pedestrian detection system using HOG and Covariance based method is described in Fig. 1.In the beginning, to identify the foreground objects, background subtraction along with some morphological operations are performed. Once we separate background and foreground objects, dynamic region of interest is extracted. For the detection window obtained, HOG features and Covariance matrix are calculated. HOG features are calculated and are fed to the linear Support Vector Machine (SVM). The SVM classifies the given input as either human or non-human. For the same detection window, Covariance features are calculated simultaneously. Covariance matrix is generated by concatenating various features such as coordinates, gradients, norms etc. For the obtained Covariance matrix dissimilarity measure is calculated using the template of the human image already stored in the database and the given image. The dissimilarity measure indicates whether there exists the desired object or not. If both the algorithms give the output as human, only then the algorithm proceeds for detection. Since we are detecting the object until it is disappearing from the scene there is no need for separate tracking algorithm thus reducing the complexity.

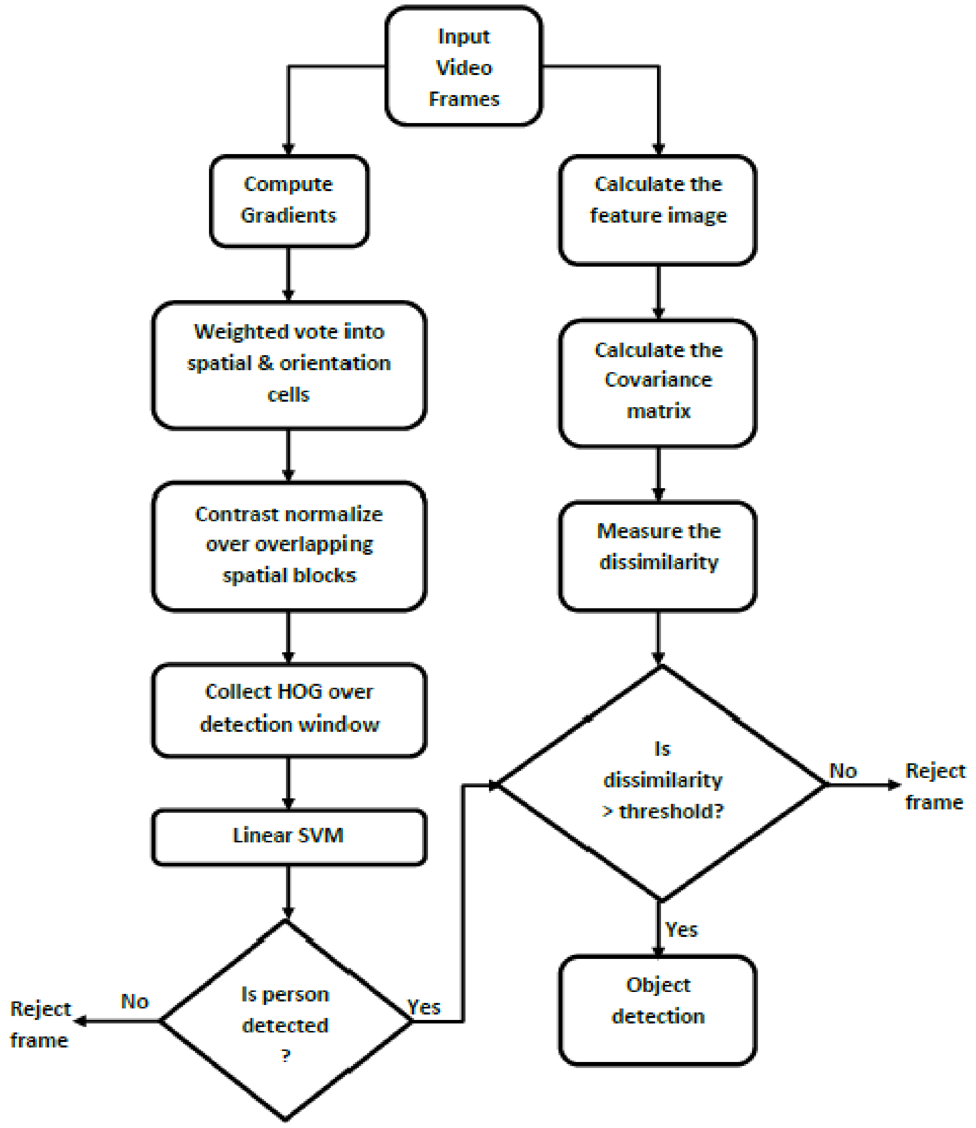

Figure 1. Overview of The Proposed Pedestrian Detection System Using HOG and Covariance Method

IV. DETECTION SYSTEM

In order to obtain the region of interest, the object in the foreground has to be localized as the work carried out in this methodology is for fixed camera. The foreground can be localized using the background subtraction algorithm. Here the background is initialized by considering a reference frame which does not have any object in it. Frame difference is calculated between the reference frame and the current frame which is shown in (1):

$$Current\_frame - Reference\_frame > Threshold \quad (1)$$

Only foreground objects will be available once we perform the background subtraction. By performing some morphological operations and using image thresholding, the background subtracted image will be converted into a binary image. The relevant figures are shown in Fig. 2.

Figure 2. Overview of Background subtraction stage (a) Reference frame (b) Current frame (c) Binarized image

For the image obtained after the background subtraction and thresholding, which is a binary image, region of interest is calculated. These interest points are nothing but the corners in both horizontal and vertical directions, which indeed corresponds to the variations in brightness levels. If the minimum of the two Eigen values of autocorrelation matrix of the second derivative of the image is greater than some predefined threshold then it will result in good corners, which is given by (2):

$$R = Min(|\lambda_1|, |\lambda_2|) > \lambda \quad (2)$$

Dynamic region of interest is created by plotting the interest points over a blank frame and a rectangle is drawn which bounds these interest points. The coordinates of the region of interest is mapped to the initial original gray image only if average intensity value is greater than the threshold. The extracted coordinates of region of interest is shown on Fig. 3. The detection algorithm is applied on this window.

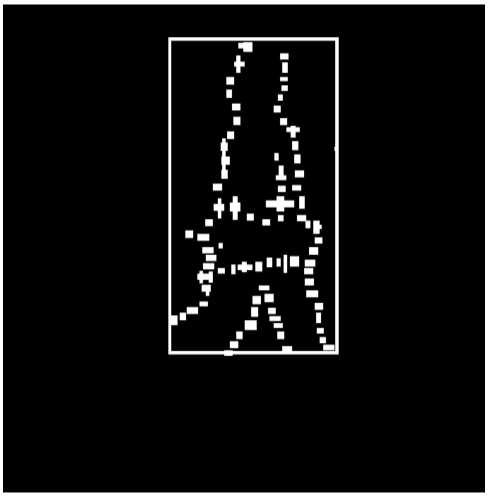

Figure 3. Dynamic Region of Interest Around Corner Points

4.1 Histogram of Oriented Gradients

The HOG features are very suitable for pedestrian detection as they are robust and give efficient performance. The HOG features are computed for the window size of 64 x 128 pixels, known as detection window. The detection window will be shifted throughout the entire image using full comprehensive scanning. To calculate the HOG features for each detection window, derivatives are calculated along both x and y directions using the 1-D derivative masks given by $M_x=[-1,0,1]$ and $M_y=[-1,0,1]^T$. The derivative masks are convolved with the original image I which results in gradients in x and y direction is given by:

$$G_x = I * M_x \quad (3)$$

$$G_y = I * M_y \quad (4)$$

The gradient magnitude and orientation angle are calculated as:

$$|G(x, y)| = \sqrt{(G_x^2 + G_x^2)} \quad (5)$$

$$\varphi(x, y) = \arctan \frac{G_y}{G_x} \quad (6)$$

The orientation bins are spaced evenly over 0^0 - 360^0 or 0^0 – 180^0 based on signed or unsigned gradient values respectively. Here 9 bins are created with 20^0 for each bin with 0^0 – 180^0 unsigned gradient value.The 64 x 128 image is divided into a closely compacted grid. In this dense grid each unit of 8 x 8 pixels is called a cell. A group of 4 adjacent cells is called a block. Next step is to create cell histograms. Weighted votes are calculated for each orientation angle among all the pixels present in a cell. This orientation gives the direction of the edges passing through that particular pixel. A 9-D histogram is created for each cell. This in turn will result in a 9 x 4 = 36-D histogram for each block. Next block is considered for HOG feature calculation by extending 50% over the adjacent blocks. Thus in total 7 horizontal blocks and 15 vertical blocks, this sums to 7 x 15 = 105 blocks and is depicted in Fig. 4. The computed histogram is normalized to take care of illumination abnormalities. Finally the HOG feature vector is created by concatenating all the histograms in the entire image for all the blocks i.e., 105 x 36 = 3780 length feature vector. The histograms are normalized using L2-norm.

4.2 Covariance Method

Covariance based object detection is a simple and efficient method. An object is represented in terms of its statistical and spatial information. The Covariance matrix provides a way for combining various kind of information such as the modalities and the features of the given image.

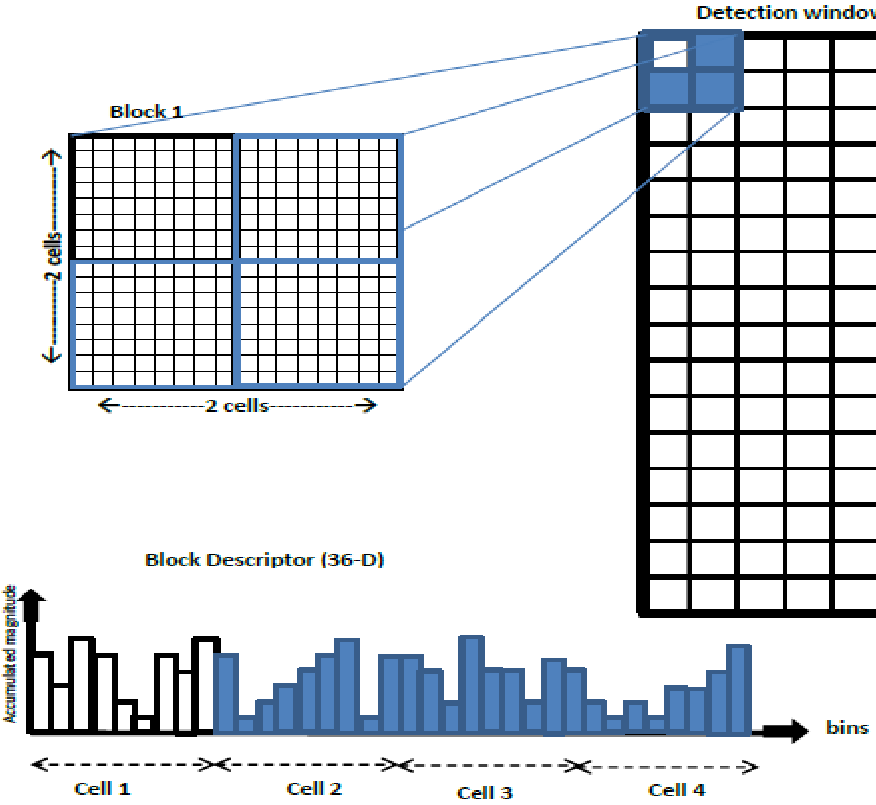

Figure 4. HOG Computation Block Descriptor

Let I be the given input image. Let W x H x d dimensional feature image be F, extracted from the input image I, and F is given by:

$$F(x, y) = \phi(I, x, y) \tag{7}$$

where ϕ can be any mapping such as color, gradient, norm, filter responses, coordinates etc.

For a given region of rectangularshape $R \subset F$, let $\{z_k\}_{k=1,2\ldots n}$ be the d-dimensional feature points inside R.

The region R with d x d matrix of covariance of the feature points is given as:

$$C_R = \frac{1}{n-1}\sum_{k=1}^{n}(z_k - \mu)(z_k - \mu)^T \tag{8}$$

where μ represents the mean of the points.

The frequently used machine learning methods do not hold good for Covariance matrices as they do not lie in Euclidean space. The distance between the feature vectors is calculated using the nearest neighbor algorithm. The distance measure used to measure the dissimilarity between the two covariance matrices is given by:

$$\rho(C_1, C_2) = \sqrt{\sum_{i=1}^{n} \ln^2 \lambda_i (C_1, C_2)} \tag{9}$$

where $\{\lambda_i(C_1, C_2)\}_{i=1,2\ldots n}$ are the generalized eigen values of C_1, C_2, computed from

$$\lambda_i C_1 x_i - C_2 x_i = 0 \quad i = 1\ldots d \tag{10}$$

and $x_i \neq 0$ are the generalized eigen vectors.

The dimension of covariance matrices are very low i.e., it has only $(d^2 + d) / 2$ different values. The diagonal values present in the covariance matrix are the variance of the each feature and the correlations is represented by non-diagonal values.

In this methodology, a 7-D feature vector using pixel locations (x, y) and norm of the first and second order derivatives of the intensities with respect to x and y is calculated using (7).

V. EXPERIMENTAL RESULTS

In this section we have presented the experimental results of the proposed methodology. The proposed algorithm is implemented using OpenCV 2.4 on Microsoft Visual Studio 2010 IDE. The work is carried out in PC Intel Core i3 second generation CPU @ 2.13GHz processor and 3GB RAM. The size of the video frame is 320 x 240. The results obtained using the proposed method is satisfactory. The average time taken to detect the object using the proposed method is only 120 ms which is enough for real-time implementation.

The result of the proposed system is shown in Fig. 5. Also the detection rate can be improved in low illumination conditions by applying a new 1-D kernel [-1, 1, 2]. This kernel enhances the edge pixels while giving a fade-out or washed out look on the given image. The result of this method is shown in Fig. 6.

Figure 5. Pedestrian Detection using HOG and Covariance Method

Figure 6. Pedestrian Detection Using HOG and Covariance Method and Using 1-D Kernel

VI. CONCLUSION

In this paper we have proposed pedestrian detection system in a video sequence using HOG and Covariance based method. The pedestrian detection system is developed with less complexity and yields good detection rate with reduced detection time. The main contribution of this paper is combination of algorithms which reduces the false detection rate and also solves the problem of occlusion thus improving the efficiency. The conducted experiments have shown satisfying results in detection.

Further work of detection can be extended for multiple cameras in densely occluded and cluttered scenarios.

REFERENCES

[1] Na Shoul, Hong Peng, Hui Wang, Li-Min Meng and K.-L. Du, "An ROIs Based Pedestrian Detection System for Single Images", IEEE 5th International Congress on Image and Signal Processing, 2012

[2] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection", IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 886 - 893, 2005.

[3] M. Ishikawa et aI, "Selection of Histograms of Oriented Gradients Features for Pedestrian Detection", ICONIP 2007, Part II, LNCS 4985, pp. 598-607,springer-Verlag Berlin Heidelberg, 2008.

[4] P. Dollar, C. Wojek, B. Schiele, P. Perona, "Pedestrian detection: an evaluation of the state of the art," IEEE Trans. Pattern Anal. Machine Intell., vol. 34, pp. 743-761, April ,2012.

[5] Daimeng Wei, Yong Zhao, Ruzhong Cheng and Guoliang Li, "An Enhanced Histogram of Oriented Gradient for Pedestrian Detection", Fourth International Conference on Intelligent Control and Information Processing (ICICIP), Beijing, China, June 2013.

[6] O. Tuzel, F. Porikli, and P. Meer, "Region covariance: A fast descriptor for detection and classification", In Proc. 9th European Conf. on Computer Vision, Graz, Austria, volume 2, pages 589-600, 2006.

[7] Yaping Liu, Jian Yao, Renping Xie, and Sa Zhu, "Pedestrian Detection from Still Images Based on Multi-Feature Covariances", Proceeding of the IEEE International Conference on Information and AutomationYinchuan, China, August ,2013

[8] Prajna Parimita Dash, Santhosh Aitha and Dipti Patra," Ohta Based Covariance Technique for Tracking Object in Video Scene", IEEE Students' Conference on Electrical, Electronics and Computer Science (SCEECS), Bhopal, India, Pages 1-4, March,2012

[9] Cargill P. C, Rius C. U, Domingo Mery and Alvaro Soto, "Performance Evaluation of the Covariance Descriptor for Target Detection", International Conference of the Chilean Computer Science Soceity (SCCC), Santiago, Chile, , pages 133 – 141, Nov 10-12, 2009.

[10] Dayanand Kumar N. C, K. V Suresh, "HOG-PCA Descriptor with Optical Flow based Human Detection and Tracking", International Conf. on Communication and Signal Processing, April 3-5, 2014.