IJARCCE

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

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 7, July 2016

Pedestrian Detection Using Moving Camera

Sahil Dalal[1], Rajesh Birok[2], Ashutosh Goswami[3]

Department of ECE, Delhi Technological University, Delhi, India [1, 3]

Associate Professor, Department of ECE, Delhi Technological University, Delhi, India [2]

Abstract: In the paper, segmentation is performed using Sliding Window method. When the camera is moving, then in such cases, the prevalent methods for segmentation like background subtraction will get failed. Therefore, with the help of sliding window technique which scans through the frame through different scales, segmentation of the object of interest, which is the pedestrian in our case, is done. For this, due to the advantages offered by HOG such as better representation of human contour, invariance to illumination changes and small movements, and easy computation in constant time make it best suited for its application as a feature extractor. After feature extraction, Neural network classifier modelled on training and subsequent testing, decides whether the region cropped by sliding window is in actual containing a pedestrian or not. In addition to that, neural network is preferred because single neural classifier can be used for training and classification of multiple classes and also for large set of database, convergence is better in neural network based classifier. And from the results it is also proven that neural network is successful in pedestrian detection.

Keywords: ANN, Canny HOG, PDS, RD-HOG.

I. INTRODUCTION

A perfect on-board Pedestrian Detection System, referred to as PDS, must detect the presence of people, stationary or moving, on the way of the vehicle and react according to the risk of running over the pedestrian. The action performed by PDS in case a pedestrian appears right before the moving vehicle and is likely to get harmed includes: warning the driver in advance, or apply braking action, or deploy external airbags, perform an evasive manoeuvre or else. It is also necessary that the entire system works well without disturbing the driver needlessly in normal situations, if there is no risk at all. In addition to that, such a system should work satisfactorily well independent of the time, road, and weather conditions. Additionally, the cost of the pedestrian detection module should be relatively small compared to the total cost of the vehicle. Following are the various modules and techniques based on computer vision used that together constitute a Pedestrian Detection System, added to that various methods used to make it suitable for our application involving pedestrian detection using moving camera. For Pedestrian Segmentation, [1] sliding window based approach will be used. As the camera is moving [2-4], then in such moving systems case the prevalent methods for segmentation like background subtraction will get failed. Therefore, with the help of sliding window technique which scans through the frame through different scales, segmentation of the object of interest is done. For feature extraction, HOG (Histogram of Oriented Gradients) [5] based features descriptor is favourable because it is independent of intensity and based on change in intensity levels i.e. gradients or edges. Moreover, it works pretty well when it comes to dealing with human detection related applications in particular. After feature extraction, a suitable Neural network [6-7] classifier modelled on training and subsequent testing, decides whether the region cropped by sliding window is in actual containing a pedestrian or not. So here basically the classifier is trained for four classes. These four classes correspond to four different view of pedestrian like: front view, side view (left or right) and rear view. For training, features of positive and negative images are extracted. Positive images are pedestrian images and negative images are images of possible background objects. In addition to that, Neural network is preferred over SVM [8] classifier because single neural classifier can be used for training and classification of multiple classes and for large set of database convergence is better in neural network based classifier.

This research paper is organized in the following sections as: Section II tells about the overview of the complete work using a block diagram. Section III tells about the different types of ECG signals used and the proposed method. Section IV gives an idea about the experimental work exploited in the method proposed. Then in section V, results are discussed followed by conclusion in section VI.

II. BLOCK DIAGRAM

The following components are proposed for splitting the architectures of pedestrian detectors for ADAS [9], listed according to the processing pipeline order:

- Pre-Processing
- Canny Edge detection to obtain gradient image.
- Use of sliding window approach to scan the frame.
- Apply thresholding on the basis of gradient density.
- Compute the HOG features of the segmented image.

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 7, July 2016

• Apply the trained Neural network Classifier, to select the segmented HOG features and check if frame contains pedestrian or not.

It is also shown in the block diagram in Fig.1.

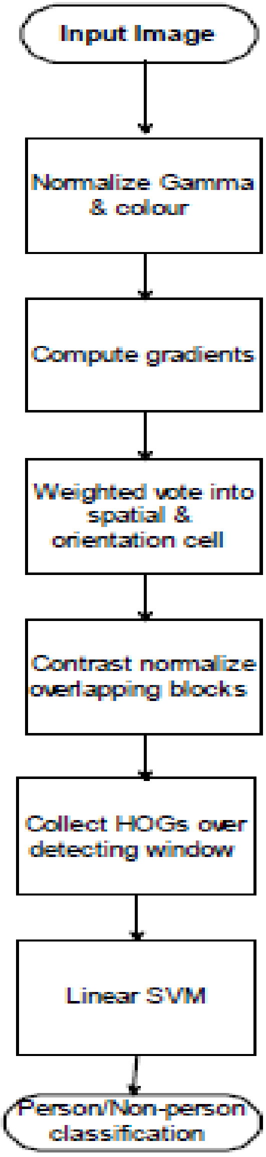


Fig.1. Block Diagram

III. PROPOSED METHOD

A. Database Formation

As we are using neural network classifier for object segmentation, so before recognize any image you need to do training for a classifier. For example, if object to be tracked is a human then we need to train our classifier for different views of human. So we need database for different views of object to be tracked [10]. Usually a database consists of positive and negative set of images. Positive set consist of images of object to be tracked and negative set consist of images of all other objects that could appear in background of the object to be tracked.

Fig.2. Sample of front view database

Fig.3. Sample images of back view database

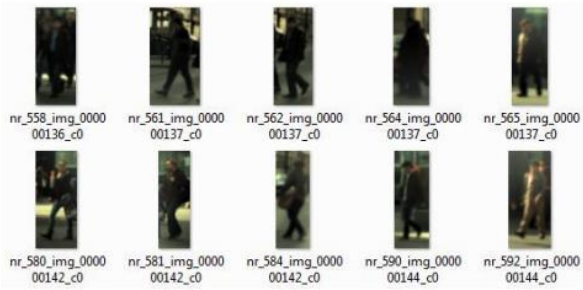


Fig.4. Sample images for left view database

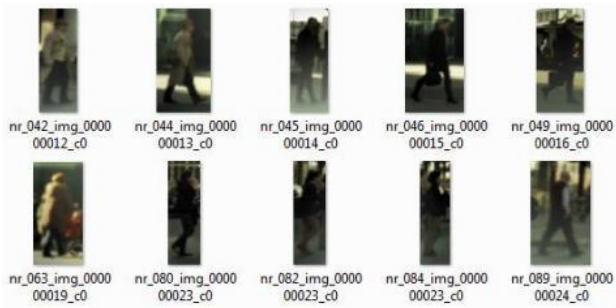


Fig.5. Sample images of right view database

B. HOG

Feature descriptors we have used to characterize an object are HOG [5] feature descriptors and its variant RDHOG [11] feature descriptors. These feature descriptors are well efficient and robust and also less computationally expensive in compare to features descriptors like SURF and SIFT. And these are also invariant to changes like scale, rotation and colour. HOG abbreviated as histogram of oriented gradients and RDHOG as relative discriminative HOG.

The Histogram of Oriented Gradients (HOG) is a feature descriptor used in computer vision and image processing for the purpose of object detection. The technique counts

Copyright to IJARCCE

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 7, July 2016

occurrences of gradient orientation in localized portions of an image.

The essential thought behind the histogram of oriented gradients descriptor is that local object appearance and shape within an image can be described by the distribution of intensity gradients or edge directions.

The image is divided into small connected regions called cells, and for the pixels within each cell, a histogram of gradient "directions" is compiled. The descriptor is then the concatenation of these histograms.

For improved accuracy, the local histograms can be contrast normalized by calculating a measure of the intensity across a larger region of the image, called a block, and then using this value to normalize all cells within the block. This normalization results in better invariance to changes in illumination and shadowing.

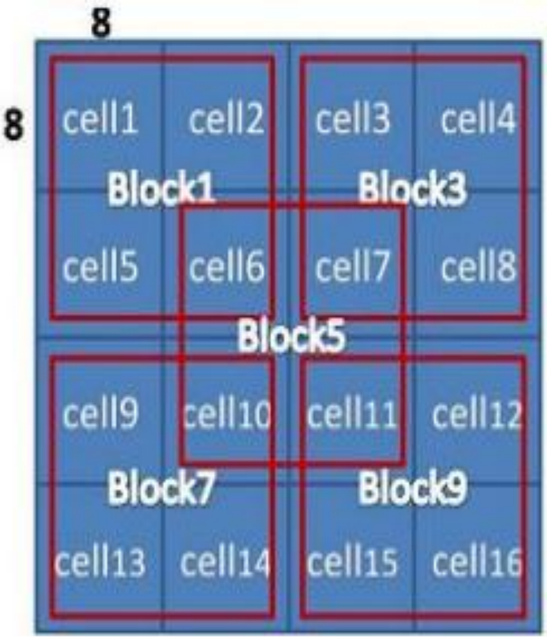


Fig.6. Block and Cell division in HOG and RDHOG feature descriptors

To calculate HOG feature descriptors first we resize our object to 32 by 32 pixels size. Then divide that in 9 overlapping blocks of 16 by 16 pixels size. Then each block is further divided into 4 cells of 8 by 8 pixels size. Then for each cell gradient is calculated, gradient id defined as change in intensity levels in particular direction. Then orientation for those gradients will be calculated and that orientation is further classified in to nine classes. That means we are constructing histogram of nine bins with a step size of 20 degree as range of orientation value 45 is 0 to 180 degree. So for each cell we are getting a vector of length 9, which means we will get a feature vector for 32 by 32 pixels image of length 324.

After getting 324 length HOG feature descriptor, now we will compute RDHOG feature descriptor by comparing HOG features of each block with the central block, as there are 9 blocks so 8 comparisons will be done. So we will get RDHOG feature vector of length 288.

The HOG descriptor has a few key advantages over other descriptors. Since it operates on local cells, it is invariant to geometric and photometric transformations, except for object orientation. Such changes would only appear in larger spatial regions. Moreover, as Dalal and Triggs (1) discovered, coarse spatial sampling, fine orientation sampling, and strong local photometric normalization permits the individual body movement of pedestrians to be ignored so long as they maintain a roughly upright position. The HOG descriptor is thus particularly suited for human detection in images.

C. ANN

Firstly, ECG classification is also done using Artificial Neural Network (ANN) [9]. In this also, train the network first by using some training data. A suitable training algorithm results in an ANN which is capable of generating a non-linear mapping function with the proficiency of demonstrating relationships between given ECG features and cardiac disorders.

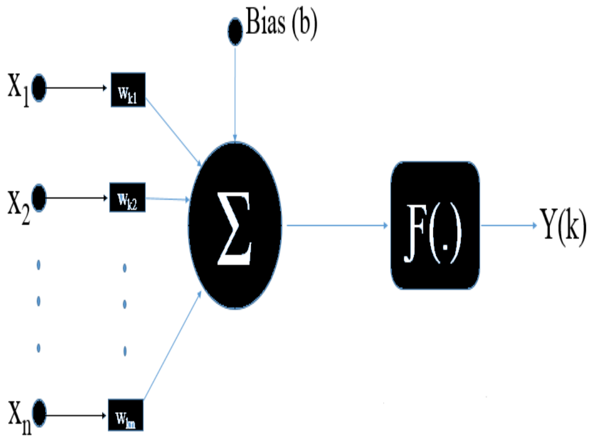


Fig.7. ANN Model

$$u_k = \sum_{j=1}^{n} x_j . w_{kj} \quad (10)$$

$$v_k = u_k + b_k \quad (11)$$

A well designed ANN will exhibit good generalization when a correct input output mapping is obtained even when the test input is slightly different from the data used to train the network.

IV. EXPERIMENTAL VIEW

a) Open the video in which pedestrian is to be detected, construct the video object of the video file and read number of frames in video. These frames now contain images run after another in a video sequence.
b) Now perform RGB to gray Conversion of the image.
c) Compute the Canny edge to obtain the Gradient Image.
d) Scan the image using Sliding Window concept column-wise and then row-wise in a loop fashion so as to cover entire frame.

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

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 7, July 2016

e) Now apply the threshold on the basis of gradient density, wherein the cropped segment having sufficient edge points are saved.
f) Compute the HoG features of the segmented image.
g) Apply the Neural Network Classifier, to select the segmented HOG features. The classifier is trained over a set of INRIA Pedestrian detection dataset images.
h) Now insert the bounding box wherever a pedestrian is detected and number the detected pedestrian in the output frame.

V. RESULTS & DISCUSSION

In pattern recognition and information retrieval with binary classification, Precision (also called positive predictive value) is the fraction of retrieved instances that are relevant, while recall (also known as sensitivity) is the fraction of relevant instances that are retrieved. Both precision and recall are therefore based on an understanding and measure of relevance, that is, precision is "how useful the search results are", and recall is "how complete the results are". In simple terms, high precision means that an algorithm returned substantially more relevant results than irrelevant, while high recall means that an algorithm returned most of the relevant results.
In a classification task, the precision for a class is the number of true positives (i.e. the number of items correctly labelled as belonging to the positive class) divided by the total number of elements labelled as belonging to the positive class (i.e. the sum of true positives and false positives, which are items incorrectly labelled as belonging to the class). Recall in a classification task is defined as the number of true positives divided by the total number of elements that actually belong to the positive class (i.e. the sum of true positives and false negatives, which are items which were not labelled as belonging to the positive class but should have been).

In information retrieval, a perfect precision score of 1.0 means that every result retrieved by a search was relevant (but says nothing about whether all relevant cases were retrieved) whereas a perfect recall score of 1.0 means that all relevant cases were retrieved by the search (but says nothing about how many irrelevant cases were also retrieved). Now finally applying the algorithm and using the codes run on MATLAB, the frames captured in real-time for the purpose of pedestrian detection using moving camera gives following results. The subsequent values of: True Positive (tp), False Negative (fn), False Positive (fp) help in determining the values of Recall and Precision, and comparative analysis for HoG based and Neural network based systems.

TABLE I COMPARISON OF HOG AND ANN BASED PARAMETERS

PARAMETER	HOG BASED	ANN BASED
True Positive	1437	1321
False Negative	220	336
False Positive	256	138
Recall	86.72	79.72
Precision	84.88	90.54

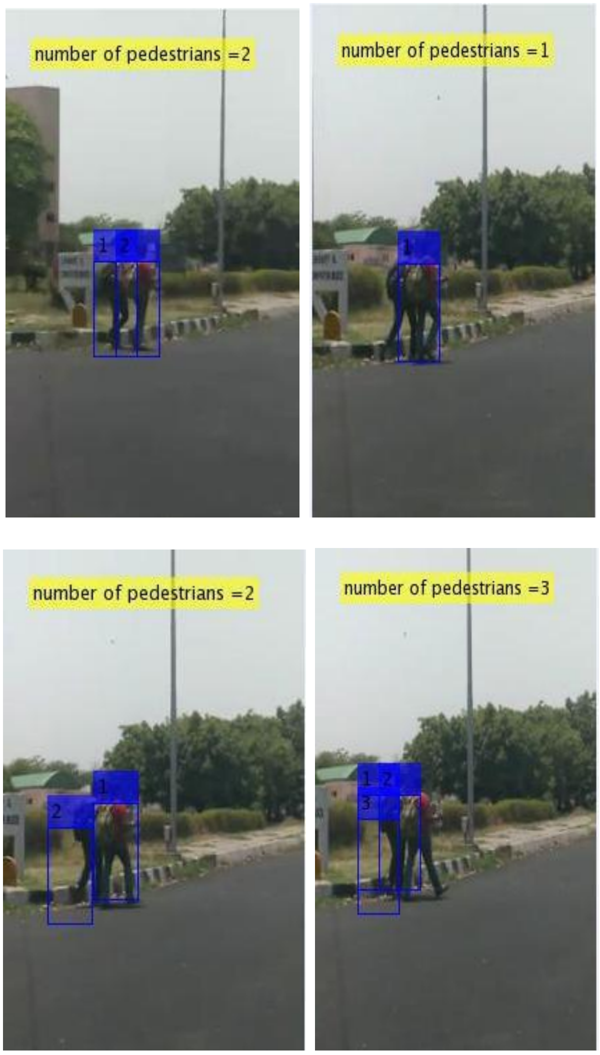
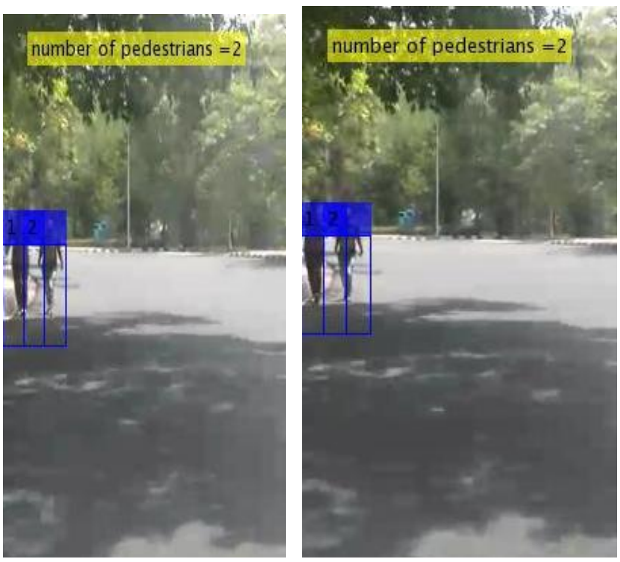


Fig.8. SCENE 1



Copyright to IJARCCE

IJARCCE

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

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified

Vol. 5, Issue 7, July 2016

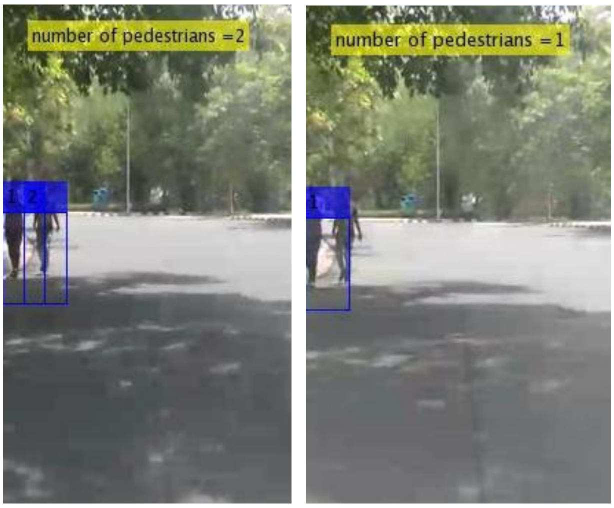


Fig.9. SCENE 2

VI. CONCLUSION

Object detection is a key technology for many future applications, among them automotive assistance systems. To achieve this goal, foreground objects are first detected through foreground/background segmentation based on our sliding window technique. The two key elements that make this system robust and real-time are motion-guided object detection, and neural network-based pedestrian detection with extracting features in several levels of complexity. Neural networks are trained on a large number of pedestrian and non-pedestrian data extracted from complex scenes; therefore, it is applicable to various real world situations. Our method gives good result but we will try other faster and reliable methods.

REFERENCES

[1] V. H. C. d. Melo, S. Leão, D. Menotti and W. R. Schwartz, "An Optimized Sliding Window Approach to Pedestrian Detection," Pattern Recognition (ICPR), 2014 22nd International Conference on, Stockholm, 2014, pp. 4346.

[2] B.Leibe, E. Seemann, and B. Schiele. "Pedestrian detection in crowded scenes"-2005.

[3] Fleuret et al. [Multi-Camera People Tracking with a Probabilistic Occupancy Map -2008] IEEE Transactions on Pattern Analysis and Machine Intelligence.

[4] P. Viola, M. Jones, and D. Snow, “Detecting Pedestrians Using Patterns of Motion and Appearance,” Int’l J. Computer Vision, vol. 63, no. 2, pp. 153-161, 2005.

[5] N. Dalal and B. Triggs, “Histograms of Oriented Gradients for Human Detection,” in IEEE Intl. Conference on Computer Vision and Pattern Recognition, 2005, pp. 886–893.

[6] C. M. Bishop, ‘Neural Networks for Pattern Recognition’, Oxford University Press, Indian Edition, 2003.

[7] L. Zhao and C. Thorpe, “Stereo and Neural Network-Based Pedestrian Detection,” IEEE Trans. on Intelligent Transportation Systems, 148-154, September 2000.

[8] Bo Wu and Ram Nevatia, "Detection of Multiple, Partially Occluded Humans in a Single Image by Bayesian Combination of Edgelet Part Detectors"- 2005.

[9] Adaboost -R.O.Duda, P.E.Hart and D.G.Stork, ‘Pattern Classification’, Johy Wiley, 2002.

[10] INRIA Pedestrian Dataset, MIT pedestrian database.

[11] Relative-Discriminative-Histogram-of-Oriented Gradients Based Particle Filter Approach to Vehicle Occlusion Handling and Tracking Bing-Fei Wu, Fellow, IEEE, Chih-Chung Kao, Student Member, IEEE, Cheng-Lung Jen, Student Member, IEEE, Yen-Feng Li, Ying-Han Chen, and Jhy-Hong Juang AUGUST 2014.

Copyright to IJARCCE
DOI 10.17148/IJARCCE.2016.5740