Design of Vision Based Pedestrian Detection System for Automobile Safety

Sujai T George Jacob[1], C Senthil Singh[2]

1. PG Student, ECE Department, Toc H Institute of Science and Technology / Cochin University, Cochin

2. Associate Professor, ECE Department, Toc H Institute of Science and Technology / Cochin University, Cochin

ABSTRACT

Driver Assistance Systems are becoming more common for safety systems in the automotive environment. With respect to road accident statistics, on-board pedestrian detection is a key task for future driver assistance systems. In this paper, we describes a system based on image processing to help the driver in these situations. Here a video camera is mounted in front of the vehicle and each frames from the video file is analyzed to take the proper decision. This system describes an image processing algorithm based on three modules - ROI generation, object classification based on HOG features and Kalman filter tracking cascaded together, and each module uses visual features to identify objects and classify each object as pedestrians from the cluttered background in the range of 20-50m. ROI generation is performed using adaptive thresholding technique based on the common fact that the gray images will have objects appearing brighter than the surrounding background. The suitable candidates are selected on the basis of various factors like height, width, aspect ratio etc. A two stage AdaBoost object classifier utilizing HAAR like and HOG feature extraction methods is described here. Adaboost is a learning algorithm building a stronger classifier combining many weak classifier with weighted majority vote. For tracking of selected pedestrians, a kalman filter based object tracking is employed. A template matching is also used in case of errors in kalman based object tracking. Matlab is used for the simulation part and the proposed algorithm works accurately with various lighting conditions and is suitable for practical applications.

Keywords:Region of Interest, Kalman Filter, Histogram of Oriented Gradient, AdaBoost

INTRODUCTION

Road safety is one of the growing concerns of our society. According to Indian Road Transport ministry statistics, 4.97 lakh road accident was reported in 2011 [18]. Pedestrians in the vehicle path are in danger of being hit causing severe injury both to the pedestrian and potentially also to the vehicle users. Road users generally ignore road safety rules due to inattentive driving though they are quite well aware of the general rules and safety measures. So the main cause of accidents and crashes seems only due to human errors. Pedestrian detection has been a focus of recent research due to its importance for practical applications such as automotive safety and visual surveillance. Accidents can be controlled with the implementation of various strategies like better policing, comprehensive driver training, and the implementation of on-vehicle driver assistance systems. Driver assistance systems aims to detect these potentially dangerous situations in advance and warn the driver to initiate protective measures in time. The main purpose of pedestrian detection system is to identify the pedestrians in dangerous conditions and to issue a warning signal back to driver. When a pedestrian is detected, risk analyzing algorithms are used to estimate the probability of a collision by taking into consideration both the predicted paths of the vehicle and the pedestrian.In this way, the collision between the vehicle and pedestrian can be avoided, or the vehicle's speed can reduced as much as possible before collision to an extent. This can reduce the risk of serious injury to a minimum possible level. Pedestrian detection brings many challenges. Pedestrian Detection is a demanding task due to continuous changes in various factors like high variability in appearance among pedestrians like pose, style and color of clothing as in Fig 1, illumination, cluttered background and weather conditions, and strict requirements in both speed and reliability.

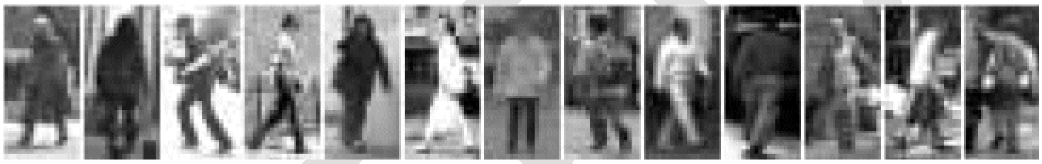

Fig 1: Pedestrian Images

Different types of methods are being used forpedestrian detection systems including various sensors likean ultrasonic sensor, a laser scanner, a microwave radar, and different kinds of cameras. Although active sensors like a laser scanner and radar can provide accurate information, it is difficult for them to distinguish pedestrians from other objects present in road.Thus, video cameras are more suitable for detecting pedestrians as they are similar to the human visual perception and provide necessary information for classifying the pedestrians from object by applying pattern recognition and feature extraction techniques [6]. The main advantage of using a video camera based system is that the huge amount of data that a single video camera can handle is much greater than what aseries of motion sensors. Therefore, there is a practical importance to develop a robust pedestrian detection system which can reduce and avoid accidents effectively.

The paper describes about a real time video processing based driver assistance system. In the experimental procedure, all the input information captured from the camera is analyzed effectively and processed for establishing the departure decision. This paper also describes an algorithm based on three modules - ROI generation, object classification and Kalman filter tracking of moving pedestrian. The rest of the paper is organised as follows. Section 2 discusses about the related works already performed in this particular field. Here the papers and different literatures referred and their main ideas are listed. In Section 3, the system architecture and the image processing algorithms are explained. All the Matlab simulation outputs and results discussions are provided in Section 4 and then Section 5 concludes the work and outlines the work performed.

RELATED WORK

The following section contains the technical papers and literatures that referred in the process of doing this project and for understanding the ROI generation technique, HOG Feature extraction technique and kalman filter based object tracking method.

Generally, a real time vision based pedestrian detection system can be divided into three module as object detection, classification of objects into pedestrian and non-pedestrian, and finally the pedestrian tracking to suppress the false alarms by analyzing when the pedestrian comes into same path as vehicle is travelling. Numerous approaches have been proposed despite various challenges, pedestrian detection and driver assistance devices still remains an active research area in recent years. Generally used methods include part based, edge detections, local features like HAAR, Histogram of oriented gradients etc for detecting multiple pedestrians. Junfeng Ge, Yupin Luo, and Gyomei Tei [6] proposes a monocular vision system for real time pedestrian detection and tracking during nighttime driving with a near infrared camera. This image processing system consists of ROI generation, object classification, and tracking integrated together to distinguish the objects from the cluttered background in the range of 20–80 m. The image segmentation step adopts two thresholds for each pixel to determine the foreground and background based on the fact that pedestrians appear brighter than the nearby background in NIR images. Gentle AdaBoost, is used for feature extraction and train the classifiers to identify each pedestrians. Xuewen Ma, Shuang Ma, MengyaoLic and Meiling Jin [16] suggests a Pedestrian recognition algorithm for real time applications using the characteristics of improved gradient direction combined with SVM classifier. This system extracts the region which may be including some person based on the shape of the image characteristics.

Sebastian Bauer, Ulrich Brunsmann and Stefan Schlotterbeck Macht [12] combines HOG features with SVM Classifiers. They suggests a sliding window scheme where a detection window is shifted on a regular lattice over the image at various scales. For each window, a feature set (HOG) is generated from the corresponding image patch and evaluated by a pretrained classifier that categorizes unknown samples into one of the predefined classes, pedestrian or non-pedestrian. The HOG feature is first used for pedestrian detection by Shashua [14]. They extract orientation histogram features from 13 fixed overlapping parts and use ridge regression to reduce the according to the different sub regions and training subsets to form compact features. Dalal and Triggs and Zhu et al [10] approach is to divide each normalized detection window of 64 × 128 into cells of size 8 × 8 pixels, and each group of 2 × 2 cells is integrated into a block with an overlap of one cell in both horizontal and vertical directions. A nine-bin HOG is constructed in each cell, and each block contains a concatenated vector of all its cells. Thus, each block is represented by a 36-D feature vector that is normalized to an L2 unit length, and each sample is represented by 7 × 15 blocks, that is, a feature vector of 3780 dimensions.

SYSTEM OVERVIEW

A real time vision based driver assist system can provide the driver with necessary warning which can reduce the risk of an accident. Here, the proposed system should work on real time vision and can provide a warning analyzing each real time instances happening in front of the vehicle. In the experimental procedure, all the input information captured in real time from the camera is analyzed effectively and warning device is responsible for launching alarm mechanism after the decision model detects any danger. The system should be able to detect the objects and should be robust against shadows, jerks, lightning conditions and any other changes in the road. All these circumstances must be efficiently handled in order to achieve an accurate vision system. An efficient adaptive threshold segmentation method is introduced for ROI generation under the common fact that the objects appear brighter than the background from the view of a horizontal scan line in gray images. [8] Haar-like features can capture different image features very quickly. But they are not very accurate in case of pedestrian detection applications because of high variety in pedestrian appearances whereas using Histogram of oriented gradients (HOG) features can produce higher performance but they are time consuming. To achieve high detection performance, a new methodology is proposed.The proposed system consists of the three modules - ROI generation, object classification, and tracking as in Fig 3. The ROI generation module contains two steps: image segmentation and candidate selection. The image segmentation step adopts adaptive thresholding technique where two thresholds are used for segmentation. The two thresholds for each pixel are calculated from the horizontal scan line to determine the foreground based on the fact that pedestrians appear brighter than the nearby background.In the candidate selection step, the regions that satisfy the scene related size and position constraints are selected as ROIs.

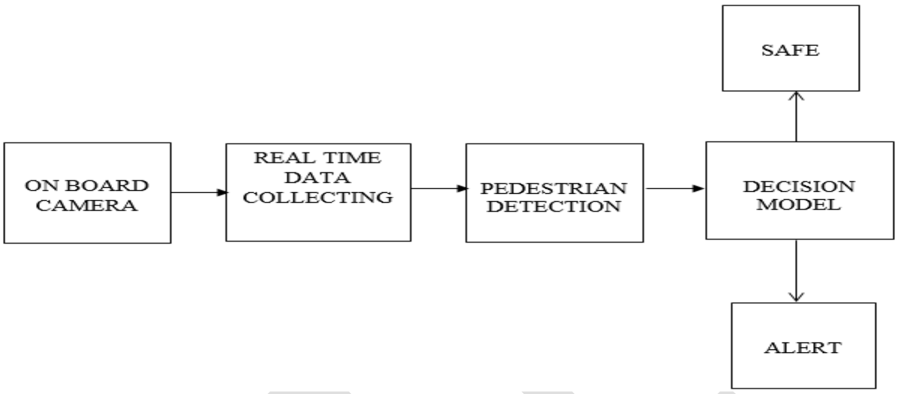

Fig 2: System Overview

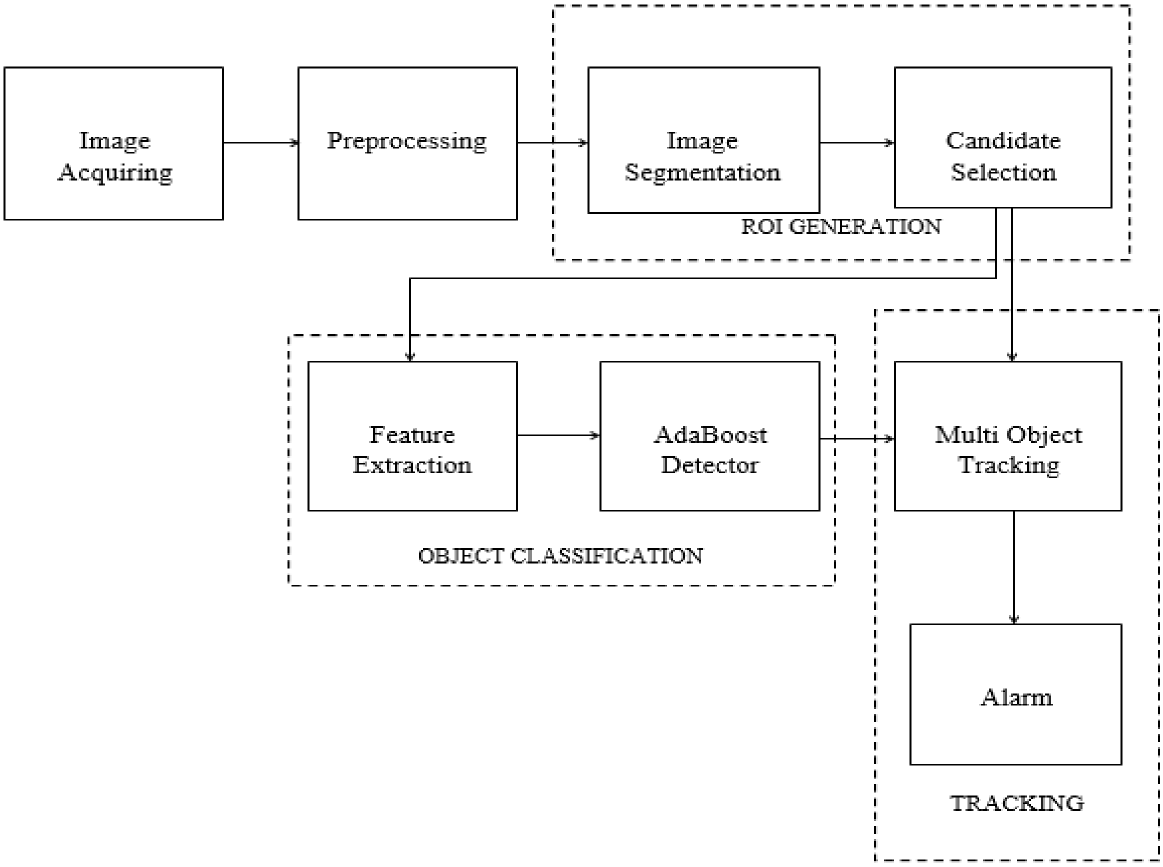

Fig 3: Pedestrian Detection System

ROI GENERATION

The ROI generation module tries to get regions that potentially contain pedestrians on fact that in the gray images pedestrians appear brighter than the surrounding background. Thresholding is the common way to divide a gray image into foreground and background. The preprocessing step reduces the noise by a 5 × 5 Gaussian filter. Erosion and dilatation operation can also be performed on the image frames to refine the segmentation results [6]. Under uneven lighting conditions, the popular solution is adaptive thresholding, where different thresholds are used for different pixels or sub regions in the image. If the pedestrians appear brighter than the surrounding background, the situation will keep the same from the view of the horizontal scan lines, even when the pedestrians have non-uniform brightness. So we are using a two threshold TL and TH for segmentation. For a given pixel I(i, j) in the image, let S(i, j) indicate the corresponding segmentation result

$$S(i,j) = \begin{cases} 1, \text{if } I(i,j) > TH(i,j) \\ 0, \text{if } I(i,j) < TL(i,j) \\ 1, I(i,j) \epsilon \ others \ and \ I(i-1,j) = 1 \\ 0, I(i,j) \epsilon \ others \ and \ I(i-1,j) = 0 \end{cases}$$

From the segmented candidates, most of the nonhuman regions are different from the pedestrian in size and position features such as height, width, aspect ratio, and the height limits at different distances. All these features can be used to reject the unwanted objects from passing to the object classification module, which can save computation time and boost system performance.

OBJECT CLASSIFICATION

The object classification module is a major step in pedestrian detection system. This step differentiate each detected objects into pedestrian and non-pedestrian classes. Learning based approaches like SVM Classifier [4], AdaBoost algorithm uses the pedestrian pattern more accurately than other methods. A collection of different visual features and combining different classifiers in cascade can be utilized for classification. We manually divides the pedestrian training set into non overlapping subsets, where each cluster represents a sample collection from a particular pose, view, or size.

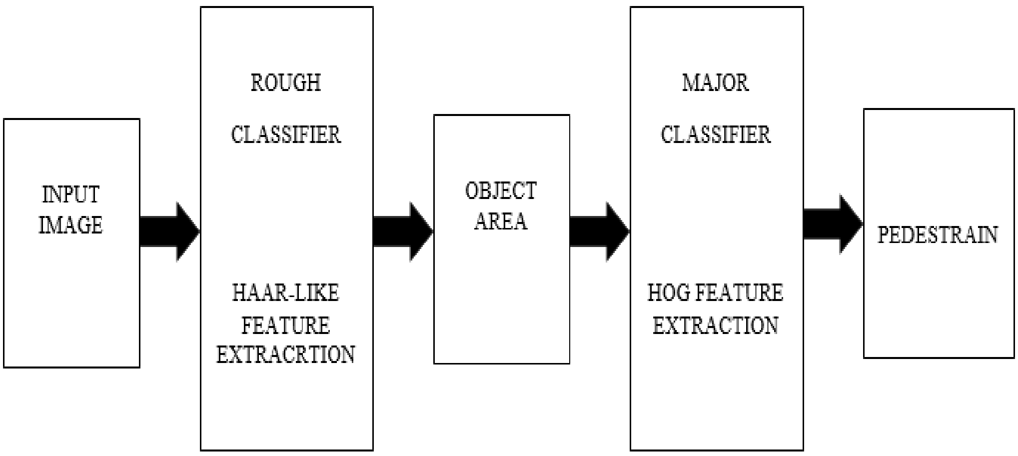

Fig 4: Two Stage Object classifier

The classification process includes two stages: In the first stage, a detector using Haar-like features scans the whole image over different positions. So we get several rectangular regions which may contain pedestrians. In the second stage, a classifier using HOG features classifies them into pedestrian or non-pedestrian. Here the classifier based on Haar-like features is used for rough classification, focusing on selecting well bounded pedestrian ROIs whereas classifiers based on HOG features are utilized to give precise determination [6].The object classification is trained through an AdaBoost algorithm. So this kind of cascaded structure can achieve increased detection performance while reducing computation time. If the pedestrian region is rejected in initial stage, it would be judged as non-pedestrian directly, while the false positives can be further classified in the second stage.

HAAR-Like Feature Extraction

Haar-like features are used for rapid object detections and widely used in many object detection applications like face, car, and pedestrian, because they effectively captures different appearance details of objects.The simple Haar-like features can be calculated by using integral images effectively. Adaboost classifiers in a cascade structure can also be employed along with Haar because of its high detection speed [8]. We employ five types of Haar-like features, as illustrated in Fig 5, including the edge and diagonal features and the line. A Haar-like feature can be defined as a block at any position and scale within the original image and value is calculated by difference of the sum of pixels of areas inside the rectangle i.e., black and white regions [6].

Fig 5: HAAR Features

HOG Feature Extraction

HOG is a grayscale image feature made by concatenating set of normalized gradient histograms. The simplest feature for object classification is image intensity, which can directly be extracted from raw data without or with histogram to reduce the influence of illumination [11], [12], [13]. In the HOG pedestrian feature extraction scheme, we extract HOG descriptors from the input image that the system takes with a camera. The computation of the HOG descriptor is done by

1) Calculate both the horizontal and vertical gradient of the input image.
2) Calculate magnitude and orientation of the gradient.
3) Split the image into 6×15 cells according to the reference size 24×60 and image size.
4) Compute a nine-bin histogram for each cell.
5) Normalize the histograms within a block of 2×2 cells.
6) Group all the normalized histograms into a single vector.

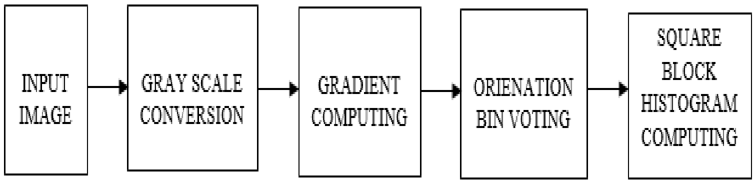

Fig6: HOG Features Extraction

Adaboost Learning

Adaboost is an iterated method, a strong learning algorithm for real-time applications. It combines a collection of simple weak classifiers on a small set of critical features to form a strong classifier using the weighted majority vote. The Adaboost algorithm is implemented by changing the data distribution by comparing with the positive and negative training sets to determine the correctness of samples and the accuracy of classification [8]. In Adaboost, every training sample is assigned with a weight, which indicates the probability of some kind of weak classifier to be selected into the training set. In our proposed system, Gentle AdaBoost, a type of the AdaBoost algorithm is used to select the relevant features and train the classifiers given the feature set and positive and negative training samples. In AdaBoost algorithm, we use classifier as the real valued weak learners that divide the instance space into a set of disjoint sub regions with the lowest error in each round. The real valued output in each sub region can be calculated by the sample weights falling into it [6].

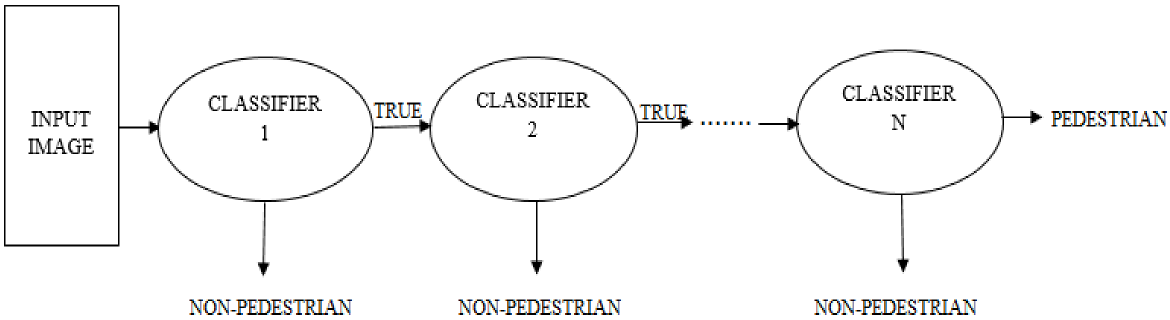

Fig 7: AdaBoost Classifier

1. From the data set Z = $(\boldsymbol{x_1}, y_1), (x_2, y_2) \ldots (\boldsymbol{x_m}, y_m)$, where $(\boldsymbol{x_i}, y_i) \in X \times \{-1, +1\}$, the number of weak classifiers to be selected T.
2. Initialize the sample weight distribution $D_1(i) = 1/m$.
3. For t = 1, . . . , T
    1) Learn a CART with lowest error as the best weak classifier to partition X into several disjoint sub regions X1, ...Xn.
    2) Under the weight distribution Dt, calculate $W_j^l = P(\mathbf{x}i \in Xj, yi = l) = \sum_{i:\mathbf{x}i \in Xj \wedge yi = l} Dt(i), \; l = \pm 1.$
    3) Set the output of h on each Xj as

$$h(x) = \frac{W_{+1}^j - W_{-1}^j}{W_{+1}^j + W_{-1}^j}$$

4. The final strong classifier H

$$H(\mathbf{x}) = \text{sign}[\sum_{t=1}^{T} ht(\mathbf{x}) - b]$$

OBJECT TRACKING

The pedestrian tracking module will fill the detection gap between frames and eliminate false detections, thus reduces errors once the pedestrians gets validated from classification module. The tracking algorithm depends on the detection results to initiate tracks and candidates to provide possible observations. If the tracked object fails to obtain observation from the pedestrians, template matching is used for complementary measurement. The tracking module has a pre-tracking stage where if a selected objects are detected more than a N_P times, then only it is moved to the tracking stage. For a pedestrian in the video, the state variables we consider are centroid position (X, Y), width (W), and height (H), as well as their differentials between two successive frames [6].The objects that we intend to track such as people and cars, have slowly changing velocities and these velocities are easy to observe [4].The Kalman filter is a framework for predicting a process's state, and using measurements to correct or 'update' predictions. Thus, the state vector for Kalman tracking is

$$x = (X, dX, Y, dY, W, dW, H, dH)^T$$

Updating Process State

Discrete time Kalman filters begin each iteration by predicting the process's state using a linear dynamics model [15]. For each time step k, a Kalman filter first makes a prediction $\hat{X}_K$ of the state at this time step

$$\hat{X}_K = Ax_{k-1} + Bu_k$$

where $\hat{X}_{k-1}$ is a vector representing process state at time k-1, A is a process transition matrix and u_k is a control vector at time k.

Updating Measurements

Kalman Gain: First, the Kalman filter computes a Kalman gain K_K, which is later used to correct the state estimate X_K

$$Kk = P_K^- H^T (H P_K^- H^T + R_K)^{-1}$$

where H converts state space into measurement space and R is covariance noise measurement [15]. Using the calculated Kalman gain k_k and measurements z_k from each time step k, we can update the state process by

$$\hat{x}_k = \hat{x}_k^- + k_k (x_k - H\hat{x}_k)$$

In addition to update of Kalman filter, another issue related to object tracking is the data association that tries to find the associated observation of the tracked object to correct the filter's prediction. Since the object is represented by its position and size, objects nearest neighbor can be regarded as the corresponding measurement of the tracked object.

$$D(X1, X2) = \sqrt{}((X1 - X2)^2 + (Y1 - Y2)2) + (|W1 - W2| + |H1 - H2|)$$

A candidate won't be considered as an observation if it's nearest neighbor and the overlap ratio is greater than 0.5. Then, template matching is introduced for improvement. The template is initiated at the beginning of the track and updated when the nearest-neighbor method works [6]. Once the nearest neighbor cannot be found in current frame, template matching is taken where if the best matching confidence is greater than 0.85, the resulting region is accepted as the measurement. The matching confidence is

$$\frac{\sum_{x,y}[(\mathrm{T}(x,y) - \bar{t}(\phi(x,y) - \bar{\phi})]}{\sqrt{\sum_{x,y}(\mathrm{T}(x,y)\ - \bar{t})\, 2 \sum_{x,y}(\phi(x,y) - \bar{\phi}) 2}}$$

Where T(x, y) is the template, Φ(x, y) is image signal of same size, $\bar{t}$ and $\bar{\phi}$ are mean values of T(x,y) and Φ(x, y).

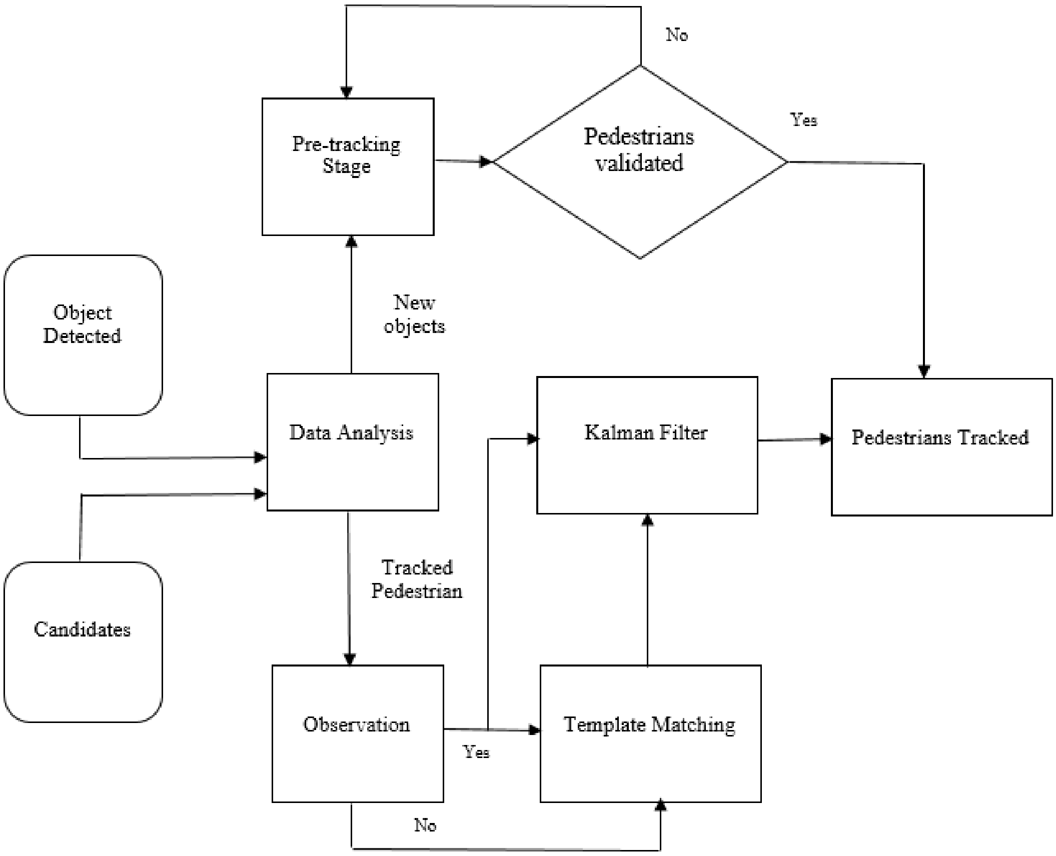

Fig 8:Tracking Module

RESULTS AND DISCUSSION

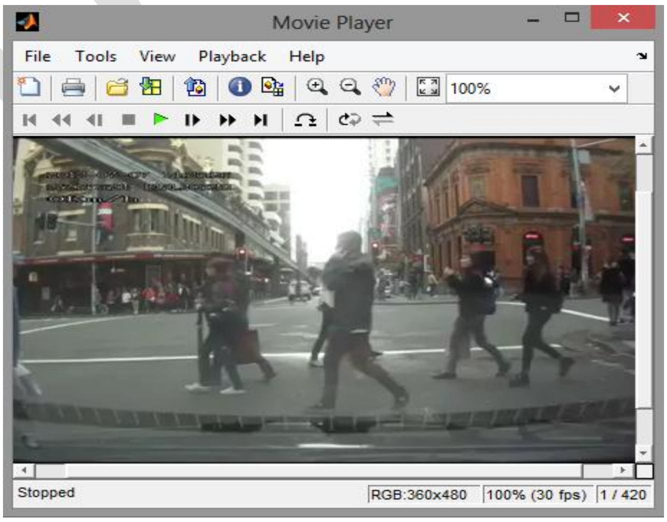

Fig 9: Image Frame

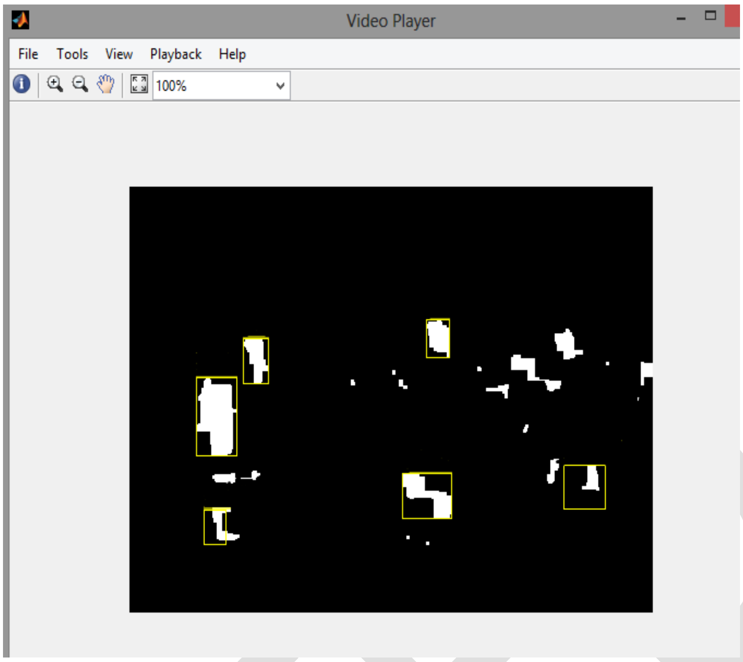

Fig 10: Threshold Image

Individual frames are extracted from the video file and each frames are processed. An image frame from the video file is shown in fig 9 and the image thresholding output is shown in Fig 10. Here any objects in current video frame are detected and then kalman tracking of object is performed subject to its motion as in Fig 11.

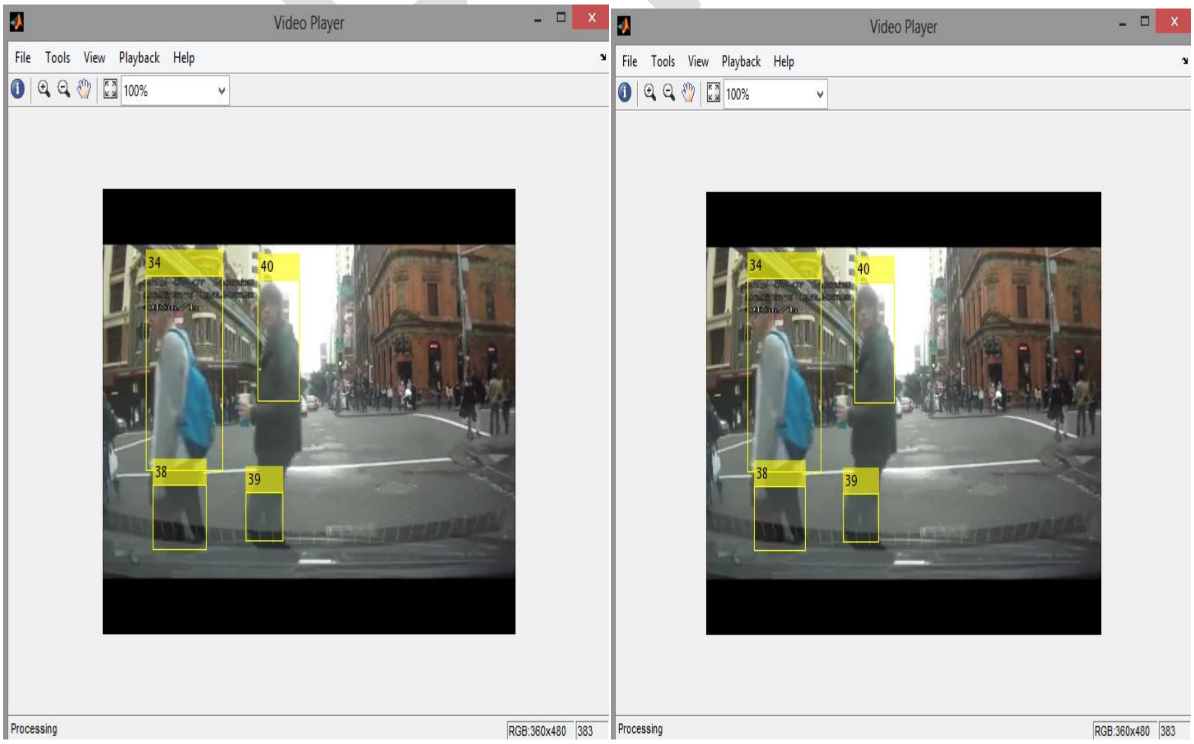

Fig 11:Simulation Output

CONCLUSION

In this paper, a video processing based driver assist system is developed. The system works well under different experimental conditions. It is able to detect various objects and classify it as pedestrians based on HOG and Haar-like feature extraction methods. The image processing is improved by ROI segmentation technique through adaptive thresholding. Object classification is performed using AdaBoost algorithm, where a stronger classifier is developed by combining a lot of weaker classifiers. A template matching kalman filter tracking also used to predict the path of pedestrian. The proposed algorithm efficiently detects the pedestrian within the minimum required computational time. For driver assistance applications it is necessary that the algorithm must be processed within short interval of time with higher accuracy because the driver should have sufficient time to take a proper decision. This proposed algorithm is suitable for real time application for Pedestrian detection and tracking systems.

REFERENCE

[1]Antonio Prioletti, Andreas Møgelmose, Paolo Grisleri, Mohan Manubhai Trivedi, Alberto Broggi, and Thomas B. Moeslund, "Part-Based Pedestrian Detection and Feature-Based Tracking for Driver Assistance: Real-Time, Robust Algorithms, and Evaluation" IEEE Transactions On Intelligent Transportation Systems

[2]A. Shashua, Y. Gdalyahu, and G. Hayun, "Pedestrian detection for driving assistance systems: Single-frame classification and system level performance," in Proc. IEEE Intell. Vehicles Symp., Jun. 2004, pp. 1–6.

[3]Christos Kyrkou, and Theocharis Theocharides"A Flexible Parallel Hardware Architecture for AdaBoost-Based Real-Time Object Detection" IEEE Transactions on Very Large Scale Integration Systems, Vol. 19, No. 6

[4]Fengliang Xu, Xia Liu, and Kikuo Fujimura "Pedestrian Detection and Tracking with Night Vision" IEEE Transactions on Intelligent Transportation Systems, Vol. 6, No. 1, March 2005

[5]G. R. Rakate, S. R. Borhade, P. S. Jadhav, M. S. Shah "Advanced Pedestrian Detection System Using Combination of Haar-like Features, Adaboost Algorithm and Edgelet-Shapelet" IEEE International Conference on Computational Intelligence and Computing Research

[6]Junfeng Ge, Yupin Luo, and Gyomei Tei "Real-Time Pedestrian Detection and Tracking at nighttime for Driver-Assistance Systems" IEEE Transactions on Intelligent Transportation Systems, Vol. 10, No. 2

[7]Kosuke Mizuno, Yosuke Terachi1, Kenta Takagi, Shintaro Izumi, Hiroshi Kawaguchi and Masahiko Yoshimoto "Architectural Study of Hog Feature Extraction Processor for Real-Time Object Detection" 2012 IEEE Workshop on Signal Processing Systems

[8]Masayuki Hiromoto Kentaro Nakahara Hiroki Sugano" A Specialized Processor Suitable for AdaBoost-Based Detection with Haar-like Features" 2007 IEEE

[9]Masayuki Hiromoto, Hiroki Sugano, and Ryusuke Miyamoto"Partially Parallel Architecture for AdaBoost-Based Detection with Haar-Like Features" IEEE Transactions on Circuits and Systems for Video Technology, Vol. 19, No. 1, January 2009

[10]N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recog., Jun. 2005, pp. 886–893.

[11]Q. Zhu, C. Yeh, T. Cheng, and S. Avidan, “Fast human detection using a cascade of histograms of oriented gradients,” in Proc. IEEE Conf. Comput. Vis. Pattern Recog., Jun. 2006, pp. 1491–1498.

[12]Sebastian Bauer, Ulrich Brunsmann, Stefan Schlotterbeck-Macht “FPGA Implementation of a HOG-based Pedestrian Recognition System” MPC July 2009

[13]Seung Eun Lee, Kyungwon Min, Taeweon Suh "Accelerating Histograms of Oriented Gradients descriptor extraction for pedestrian recognition" Computers and Electrical Engineering pp 1043–1048

[14]Shashua, Y. Gdalyahu, and G. Hayun, “Pedestrian detection for driving assistance systems: Single-frame classification and system level performance,” in Proc. IEEE Intell. Vehicles Symp. Jun. 2004, pp. 1–6.

[15]Sheldon Xu and Anthony Chang "Robust Object Tracking Using Kalman Filters with Dynamic Covariance"

[16]Xuewen Ma, Shuang Ma, Mengyao Li, Meiling Jin "Design and Implementation of Vehicle Pedestrian Detection System Based on FPGA "Advanced Engineering Forum Vols. 2-3 (2012) pp 495-500

[17]Yajun Fang, Keiichi Yamada, Yoshiki Ninomiya, Berthold K. P. Horn, and Ichiro Masaki “A Shape-Independent Method for Pedestrian detection With Far-Infrared Images” IEEE Transactions On Vehicular Technology, Vol. 53, No. 6, November 2004

[18]http://www.hindustantimes.com/indianews/newdelhi/every-3-minutes-an-indian-dies-in-a-road-accidentwarning_system