

Real Time Pedestrian Detection and Tracking for Driver Assistance Systems

Anjali Patil, Arunkumar G.



Abstract- In Autonomous driving technology detecting pedestrians and vehicles should be fast and efficient in order to avoid accidents. Pedestrian detection and tracking is challenging for complex real world scenes. In proposed system Kalman filter has been used to detect and track the pedestrians. From three frames initially eigen object is computed in video sequences for detection of moving objects, then shape information is used to classify humans and other objects. Moreover with the help of continues multiple frames occlusion between objects get calculated. In the proposed system an application is developed which gives automatic warning in case of doubtful activities performed by pedestrian of zone monitoring which can be used in various domains like defence and traffic monitoring. Proposed algorithm gives accurate moving object detection and advanced sensors are used to detect human movements ahead and alert the driver by using buzzer, result does not affect by body pose of individual.

It's not applicable for moving real time object detection



Fig 1: Pedestrian detection using sensor

Indexed Terms- Pedestrian detection, Tracking, ACF

I. INTRODUCTION

In past few years number of road accidents has been increased dramatically. According to Road safety report of 2017, 1.2 million deaths happened because of road accidents and there is record of 50 million injuries due to road accidents. Report says that more than half accidents happened with pedestrians which force us to focus on safety of pedestrians. Advanced Driver Assistance system (ADAS) focuses on safety of pedestrians by avoiding accidents. ADAS uses sensors like Lidar, Radar to perform pedestrian detection. Most widely used sensors are cameras due to its low cost and rich information. Vision based algorithms are developed for pedestrian detection like Histogram Oriented Gradient (HOG) with Support Vector Machine (SVM). Surveillance Systems are used to detect doubtful and suspicious activities, Large amount of videos been collected which helps surveillance system. System should automated to extract needed information about suspicious behavior. The kalman filter which we have used in proposed system predicts the location and size of each pedestrian and forms a bounding box around all pedestrians. We have developed an application which provides automatic warning in case if any suspicious activity is observed which can be used in defense and traffic monitoring. We as a driver can't always predict what a pedestrian on the road will do, especially children. It is difficult to guess for computer also but sensors can do this work easily. Vision of pedestrians and drivers is greatly reduced during dusky days. Sensors make it simple for us.

II. LITERATURE SURVEY

A wide range of literature is available on object detection, mostly on pedestrian detection. Pedestrian detection can be grouped into two types, first is Image based detection and second is Video based detection. In first type only static information is used for detection. In second type that is video based detection static as well as dynamic information is taken in consideration like motion f data, depth of data, it improves performance and accuracy. Pre-processing is helpful to improve the accuracy.

Dalal proved that the Histograms of Oriented Gradient (HOG) descriptors performs better human detection than other existing algorithms[1]. Dalal has used first group for pedestrian detection i.e. Image based pedestrian detection.

Human detection in images is challenging task due to their variable appearance and possibility of poses they might adopt. Dalal has shown this using normalized HOG oriented features. To improve the performance of human detection, the appearance features of HOG are used. Viola has developed a system which detects walking pedestrians. Many features have been developed for edge information for detecting pedestrians. HOG and SVM together gives better detection results. Though HOG and SVM together gives excellent detection results, but this approach is not so good to deal with moving object recognition in real scenes. It is limited to image based detection system only it's not applicable for moving real time object detection. Hong-Son Vu et.al.,[2] proposed fusion method which consists of detection and classification modules, where detection shows sizes and location of moving objects and classification is used which classifies the detected moving objects into different categories.

Manuscript published on 30 July 2019.

* Correspondence Author (s)

Anjali Patil, Mtech from VIT University, Vellore in Computer Science and Engineering.

G.Arunkumar, Asst. Prof (Senior) VIT University , vellore, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Background Subtraction (BS) method is used to detect dynamic objects. AdaBoost algorithm classifies detected dynamic objects in their respective categories. Detecting moving object like pedestrians is very difficult task in real time because of illumination variations, articulation and cluttered background. Advances of Background Subtraction techniques are used to minimize search space and detection time says Hong-Son Vu. BS technique is adopted to extract moving objects on complete scene image. The observed processing time is 4.2 times faster than that of AdaBoost algorithm.

Previously 2 main methods were used by people for object detection i.e.

- 1) Sliding window based upon HOG with SVM.
- 2) Viola- Jones detector.

The mentioned methods consume high computational cost. Hong-Son Vu et.al, proposed new method which consist of two steps i.e.

- 1) Extraction of moving objects proposals
- 2) Test detection performance on datasets.

Finally new method is proposed to find bounding boxes generated around foreground objects and for classification of objects AdaBoost algorithm is used

Though background subtraction algorithm has been used in this paper, it's not suitable for identifying moving objects in complex environment.

Wren et. al [3] proposed background subtraction technique by modeling the static background at each pixel location with a single 3D Gaussian distribution. After the background is modeled, every pixel of the input frame is then considered as the foreground pixels. Pfinder is a real time system for tracking and guessing the individual's behaviour. Multi-class statistical model of color and shape is used in system to get a 2D representation of head and hands in a wide amount of viewing conditions. These representations are useful for many applications like video database, low bandwidth coding and wireless interfaces.

Pfinder expects scenes which are less dynamic. Pfinder is capable to take small changes in the scenes but it becomes difficult when comes to sudden, large changes in scenes. Another limitation to the dynamic scene change is, only one user is expected in the scene at a time. There is no problem in low level segmentation or blob tracking algorithms but it causes problems with primitive gesture recognition system.

Chien et. al [4] proposed a video object recognition and tracking for smart cameras in surveillance networks with robust threshold decision algorithm for video object segmentation with a multi-background model. Another is video object tracking based upon particle filter. Proposed model can track no rigid objects.

The proposed algorithm uses a mechanism which is different from per pixel background subtraction to prevent possible errors. There is only one factor which affects threshold i.e. camera noise is to be zero mean Gaussian distribution.

Assumption made in this paper is not practical in real life scenario so we cannot consider this effective.

S Azmat et. al [5] proposed an algorithm which focuses on utilizing fine grain data parallelism and optimizing memory access pattern to target low cost

adaptive background modelling algorithm on GPU's with low power. Background modelling techniques are required in basic surveillance applications but their accuracy is not sufficient for computer vision problems. Gaussian Mixture Model (GMM) are robust to dynamic background but they are slow. Adaptive background modelling implemented on GPU's due to its fine grained parallelism with significant improvements of performance over CPUs. The main focus of this paper is to optimizing memory access patterns and exploiting fine grained pixel level parallelism to target low cost adaptive background modelling algorithm.

Though Computational and memory cost is less but this algorithm has a comparable accuracy with GMM.

Y.W.Xu et.al [6] proposed pedestrian localization technique using extended kalman filter. The localization model sets up relation between observation and state vector and then choose correct state for filter using perspective projection principle. This technique gives location of pedestrian quickly and accurately. This is used to avoid collisions between pedestrians and vehicles. For this only single camera is required which makes system low cost with good performance. There is no limitation about height of pedestrian. Height of pedestrian does not affect the system.

Pedestrian's location information is predicted in this paper but it cannot predict direction of pedestrian.

Anjana Das K M and O V Ramana Murthy [7] have proposed a method to monitor the driver in order to assist him. Input to the system is video and optical flow between video frames is calculated in later stage. The next step performed is separation of orientation and magnitude field. Super pixel segmentation is performed and Histogram Descriptor is defined and this is used to train one class SVM classifier.

From the results we can see that the latency is high. In order to improve the accuracy and to reduce the latency multiple classifiers can be used so that more efficient model can be build.

III PEDESTRIAN DETECTION SYSTEM

Now a days Assistance systems are used which offer object detection, mainly pedestrian detection to avoid the collisions and offer safety on the roads to individual. The systems are designed for safety purposes by tracking smaller and complex dynamic objects in the path of cars and warning drivers about the same.

The proposed pedestrian detection system in this paper accurately detects and track pedestrians until they are in our Region of Interest. The architecture has been shown below in the architecture diagram.

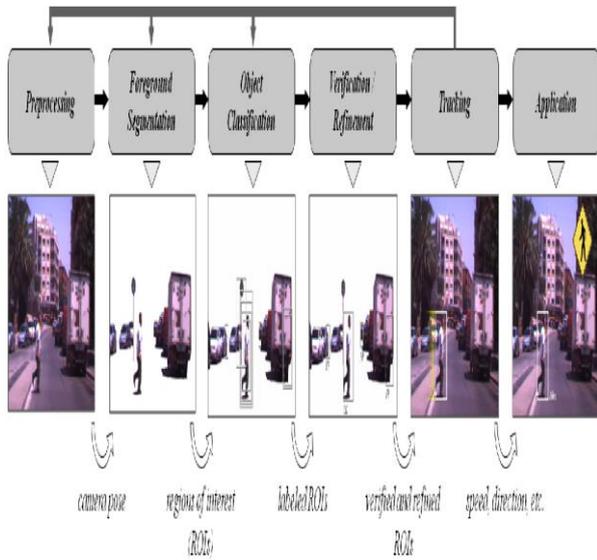


Fig 2: Pedestrian Detection System

The system works in the steps mentioned below.

- 1) Image Preprocessing
- 2) Frame Conversion
- 3) Feature Extraction
- 4) Moving object Detection
- 5) Object Tracking
- 6) Driver alert

Kalman filter is most widely used technique for pedestrian detection and tracking. It provides object's future location precisely. Video frames taken as an input from user and then gets preprocessed, after that we apply kalman filter to detect the pedestrian and track him in future until he is in our region of interest.

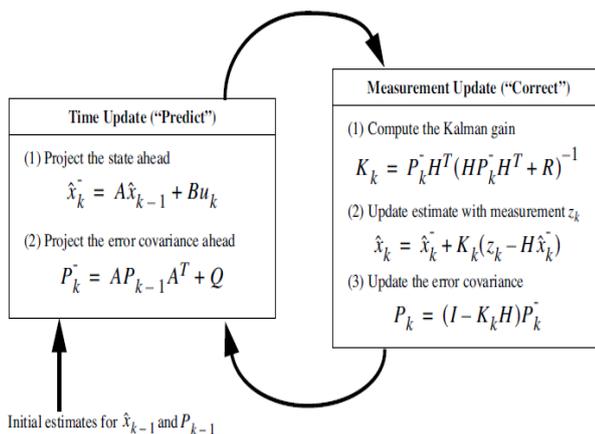


Fig 3 : Kalman filter formula

```

1. Input: Edge detection
2. Find contours of the detected objects
3. Calculate area, width and height of bounding rectangles box for the detected objects
4. for (size_t i = 0; i < contours.size(); ++i) do
5.   if ((area > threshold1) && (box.width > threshold1) && (box.height > threshold1)) then
6.     for (size_t j = 0; j < contours.size(); ++j) do
7.       if ((area > threshold2) && (box.width > threshold2) && (box.height > threshold2)) then
8.         - Calculate the distance d from the defined location to neighboring locations.
9.         if (d < threshold3) then
10.          - Cluster the neighboring locations round the defined location
11.        end if
12.      end if
13.    end for
14.  end if
15. end for
    
```

Fig 4: Algorithm- Kalman filter

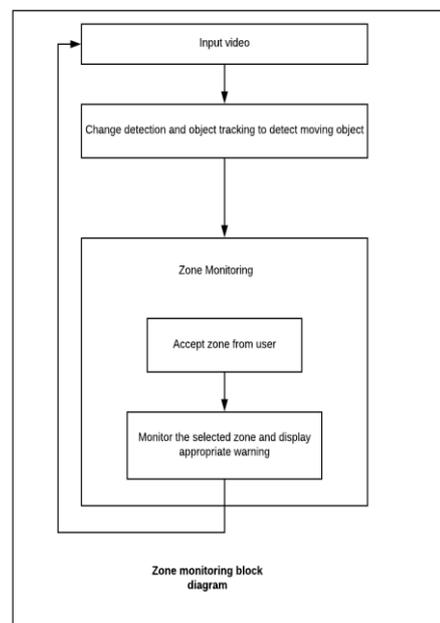


Fig 5: Zone monitoring application

In this model Bhattacharya distance measure is used to compute the similarity between the histogram of any given change in the current and previous state.

$$Bhattacharya\ distance\ (D) = \sum_{u=1}^m \sqrt{p^{(u)} * q^{(u)}}$$

$$weight = (1 / \sqrt{2\pi}\sigma) * e^{-D^2 / 2\sigma^2}$$

Here,

p = Histogram of current state

q = Histogram of previous state

σ^2 = variance of previous state of a change

For each change, weighted mean of centers of its corresponding new probable states is considered as center of its new state. Histogram of the new state is computed using following equation where α is the learning rate.

$$\text{histogram of new state} = [\alpha * \text{histogram of previous state}] + [(1 - \alpha) * \text{histogram of new state}]$$

IV RESULTS

As shown in the following figures kalman filter is used in proposed system which predicts the location and size of each pedestrian and forms a bounding box around all pedestrians. We have also developed an application which gives automatic warning in case of any doubtful activity which can be used in defence and traffic monitoring. Proposed algorithm gives accurate moving object detection and advanced sensors like ultrasonic are used to detect human movements ahead and alert the driver by using buzzer, result does not affect by body pose of individual.

A. Screenshots of output

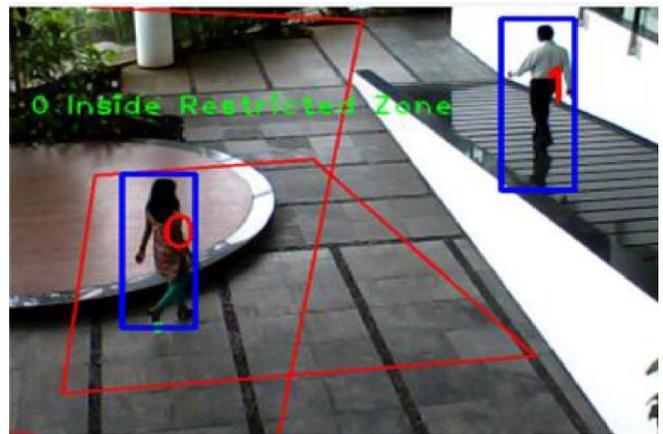
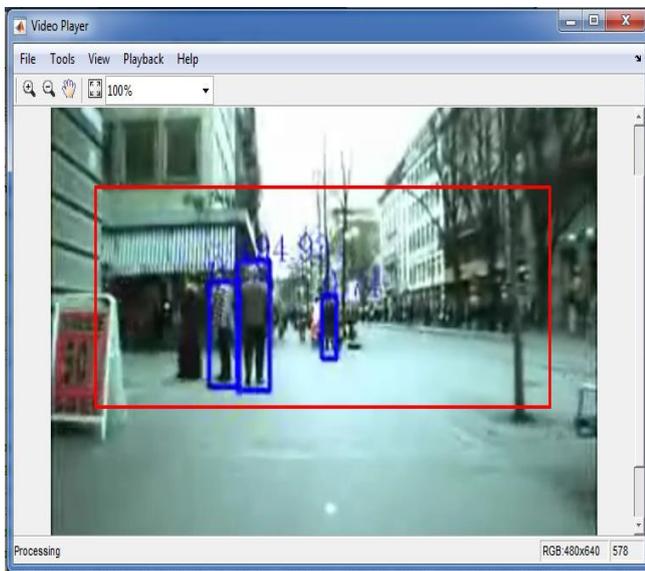


Fig 6: Person inside restricted zone

B. Precision Graph

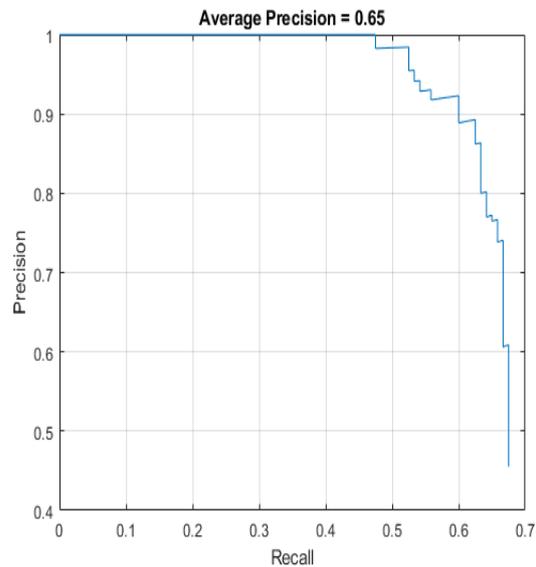
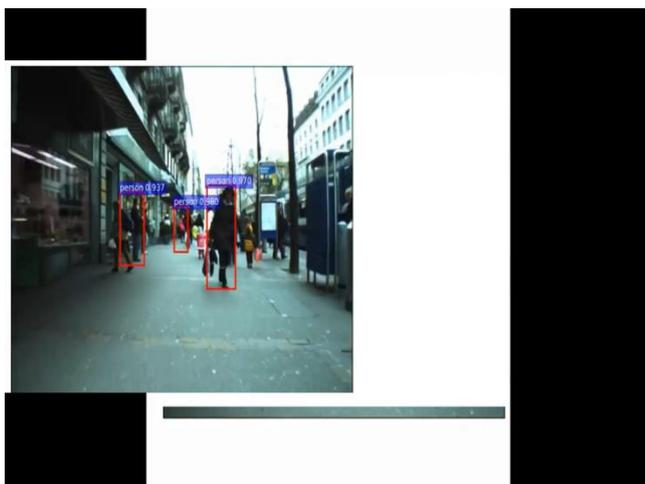


Fig 7: Precision curve



C. Test Cases

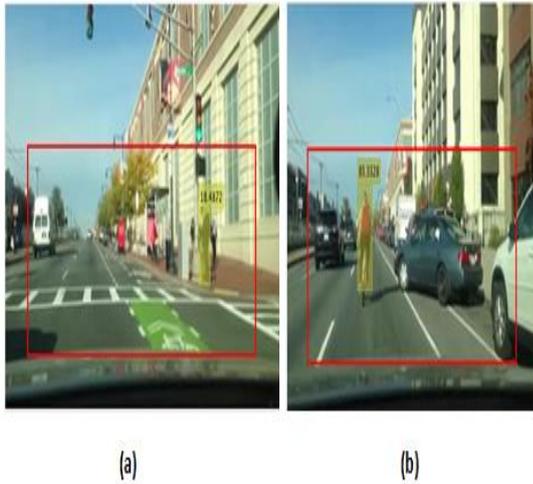


Fig 8: Detection of approach on input dataset a) True False positive: Correct detection but not annotated b) Bounding box error: Bounding box is small.

V. CONCLUSION

Tracking pedestrians in crowded area is a real life challenge. In this research I compared two strong algorithms to prove that Adaptive GMM performs better than ACF algorithm. There are cases where pedestrians are not getting detected in ACF algorithms. GMM is detecting and tracking all pedestrians.

‘x’ and ‘y’ coordinates are considered in case of Adaptive GMM algorithm. It gives better result.

In case of ACF algorithm pedestrian will get detected only when the distance between car and pedestrian is in certain range. The number of false positives are more in this case.

Lastly hardware is an alternative to this software system. There are some scenarios when due to presence of fog driver is not able to see clearly, hardware performs better in this case.

REFERENCES

1. Dalal, Navneet, and Bill Triggs. "Histograms of oriented gradients for human detection." Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. Vol. 1. IEEE, 2005.
2. Hong-Son Vu, Jia-Xian Guo et al., "A Real time Moving objects Detection and Classification Approach for Static cameras, 2016. International Conference on Consumer Electronics-Taiwan
3. Wren, Christopher Richard, et al. "Pfinder: Real-time tracking of the human body." IEEE Transactions on pattern analysis and machine intelligence 19.7, pp. 780-785 1997
4. Chien, Shao-Yi, et al. "Video object segmentation and tracking framework with improved threshold decision and diffusion distance." IEEE Transactions on Circuits and Systems for Video Technology 23.6, pp. 921-934, 2013.
5. S. Azmat, Linda Wills, Scott Wills, "Parallelizing Multimodel Background Modelling on Low Power Integrated", Journal of Signal Processing System, pp. 1-11, February 2016.
6. y.w. Xu , X.B. Cao and T. Li "Extended Kalman Filter Based Pedestrian Localization for Collision Avoidance", Proceedings of the 2009 IEEE International Conference on Mechatronics and Automation August 9 - 12, Changchun, China.
7. Anjana Das K M and O V Ramana Murthy "Optical Flow Based Anomaly Detection in Traffic Scenes", 2017 IEEE International Conference on Computational Intelligence and Computing Research.
8. Gurunathan, A., & Viswanatham, V. M. (2017) "Autonomic Performance Enhancement Environment for Websphere Application Server. International Journal of Pure and Applied Mathematics, 116(23), 719-731".

9. Arunkumar, G. and Madhu Viswanatham, V. (2017). "Autonomic Performance Management Approach for Business Product: Action Planner for Internal Change Management", Journal of Advanced Research in Dynamical and Control Systems, 113(13), 482—493.

AUTHORS PROFILE



processing.

Anjali Patil has completed Mtech from VIT University, Vellore in Computer Science and Engineering. Anjali has done her Mtech project in the field of Machine learning for Pedestrian detection system. She has experience of working as an Intern with Tata Consultancy Services, Pune in the field of image



Bigdata.

G. Arunkumar received his P.hD Computer science and engineering at VIT University Vellore India and working as a Asst. Prof (Senior) VIT University , vellore, India. His area of research is autonomic computing for web based business applications. Also other areas he is working in cloud computing, machine learning and