

Multiple Sensor Fusion for Moving Object Detection and Tracking

Khan Rabiya Yunus¹, Prof : M. A. Mechkul²

^{1,2} Dept. Of Electronics and Telecommunication, SNJB's COE, Chandwad, Maharashtra, India

Abstract - This paper introduces a successful approach for Moving Object Detection and Tracking. It is an important part of advanced driver assistance system for moving object detection and classification. By using multiple sensor for object detection we can improve the perceived model of environment. In first step we describe the combined object representation. In second step we suggest a complete observation mixture design based on evidential structure to solve the detection and tracking problem by integrating the multiple image and doubt organization. Finally we add our mixture approach in real time application inside a vehicle.

Key Words: Intelligent vehicles with break, combination of accident protection sensor, vehicle detection to avoid accident, vehicle safety, object detection like dynamic or static, pedestrian detection for life safety.

1. INTRODUCTION

By various research and development vehicles have moved from being a robotic application of tomorrow to a current area. Intelligent vehicle system need to work gradually in free situations which are logical and dynamic. ADAS is useful for drivers to perform difficult tasks to keep away from risky situations. In help task: warning in danger driving situation is sent in the form of message, safety devices are activated to reduce sudden crash, independent conversion to avoid obstacles and warn to careless driver.

IVP consist of two main tasks:

1. Simultaneous Localization and Mapping (SLAM)
2. Detection And Tracking Moving Objects (DATMO)

SLAM deals with modeling static parts of environment whereas DATMO deals with modeling dynamic parts of environment. Intelligent Vehicles Applications like the advanced driving systems (ADAS) help drivers to perform difficult driving tasks and avoid risky situations. There are three main components of ADAS: Perception, Reasoning and Decision and Control.[1]

Once object recognition and tracking is completed a sorting step is needed in order to verify which class of objects are surrounding the vehicle. After getting the information about moving object near to vehicle, it can help to improve their tracking, and it is done by their behavior and according to their performance we can decide what to do. The Current state of the ability approaches for object classification focus only in one class of object (e.g. pedestrians, cars, trucks, etc.)

and they are depend on one type of sensor (active or passive) to perform such task. Including information from different type of sensors can improve the object classification and allow the classification of multiple class of objects. Individual object classification from specific sensors, like camera, proximity sensor have different reliability degrees according to the sensor advantages and drawbacks.

2. RELATED WORK

The main task in mobile robotics in field of intelligent vehicle is detection and tracking of moving objects. Here SLAM component perform low level fusion, similarly DATMO component perform detection and track level fusions. At Detection level, sensor detects and informs about moving object. After that this lists of moving object are combined and an improved list is prepared. At the tracking level, tracking list of moving object are combined to prepare an improved list of tracks. object representation can be improved by Combination at obstacle detection stage, allowing the tracking process to depend on this information to make enhanced organization decisions and accomplish improved object estimates. If we combine different sensor inputs, at that time we must know the classification accuracy of each sensor.

Here We can use all the detection information provided by the sensors (i.e., position, shape and class information) to build a mutual object representation. Given several lists of object detections, the proposed approach perform an evidential data association method to decide which detections are related and then combine their representations. We use proximity sensor and camera sensor to provide an approximate detection position; and we use shape, relative speed and visual appearance features to provide a preliminary confirmation distribution of the class of the detected objects. The proposed method includes uncertainty from the sensor detections without removal of non-associated objects. There are Multiple objects of interest are detected which include: pedestrian, bike, car and truck. Here Our method takes place at an early stage of DATMO component but we present it inside a complete real time perception solution.

3. FUSION DETECTION LEVEL

Our work proposes a sensor fusion framework placed at detection level. Although this approach is presented to work with three main sensors, it can be extended to work with

more sources of evidence. Figure 1 shows the general architecture of the proposed fusion approach. The inputs of this method are several lists of detected objects. Class information is obtained from the shape, relative speed and visual appearance of the detections. Here we get output of the fusion method comprises a fused list of object detections, which is represented by a combined representation that includes: position, shape and an facts distribution of class hypotheses.

3.1 Object Detection Level

Usually, we know the object detections are represented by their position and shape features. We think that class information can be important to consider at detection level. However, at this level there is not enough confidence about the class of the object.[4]

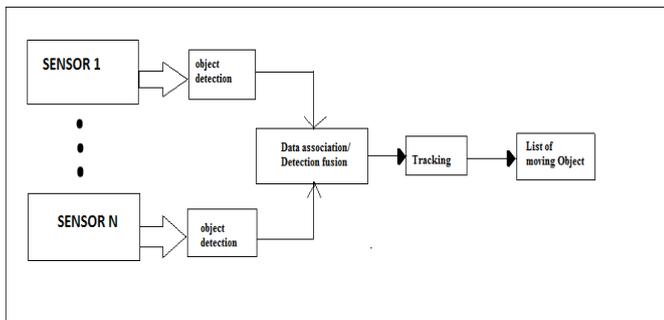


Fig.1. Schematic of proposed fusion architecture.[4]

A. PIR Based Detection

A very cheap, easy to assemble, easy to use Infrared sensor with a long detection distance and has less interference by visible light. The implementations of modulated IR signal immune the sensor to the interferences caused by the normal light of a light bulb or the sun light. This sensor has a screwdriver adjustment to set the appropriate detected distance to make it useful in many applications, and then gives a digital output when it senses something within that range. This sensor does not measure a distance value. It can be used for collision avoidance robot and machine automation. The sensor provides a non-contact detection.

B. Camera Images

We need to remove discriminative visual features, to obtain appearance information from images,

1) Visual Representation:

The Histograms of Oriented Gradients (HOG) descriptor has exposed promising results in vehicle and pedestrian detection. Here, we took this descriptors the core of our vehicle and pedestrian visual representation. The goal of this task is to generate visual descriptors of areas of the image to

be used in future stages to determine whether these areas contain an object of interest or not. We propose a sparse version of the HO descriptor (S-HOG) that focuses on specific areas of an image patch. This allows us to reduce the common high-dimensional HOG descriptor [12]. Fig.(a) illustrates some of the blocks we have selected to generate the descriptors for different object classes. These blocks correspond to meaningful regions of the object (e.g., head, shoulder and legs for pedestrians). HOGs are computed over these sparse blocks and concatenated to form S-HOG descriptors. To accelerate S-HOG feature computation, we followed an *integral image* scheme [1]

2) Object Classification

Due to performance constraints, we did not implement a visual based moving object detection. Instead, we used the regions of interest (ROI) provided by proximity sensor detection to focus on specific regions of the image. For each ROI, visual features are extracted, and a classifier is applied to decide if an object of interest is inside the ROI. The selection of the classifier has a large force on the resulting speed and quality. It combines many weak classifiers to form a powerful one, where weak classifiers are only required to perform better than option. For each class of interest (pedestrian, bike, car, truck), a binary classifier was trained off-line to identify object (positive) and non-object (negative) patches. For this training stage, positive images were collected from public (such as the Daimler dataset) and manually labeled datasets containing objects of interest from different object's viewpoints (frontal, rear, profile) Fig. b shows examples of the pedestrian and car detection results (green and red boxes respectively) before merging into the final objects. We guess the confidence of object classification for each possible object. Generally, the greater the number of positive areas (containing an object of interest), the higher the confidence that the object belongs to that specific class.

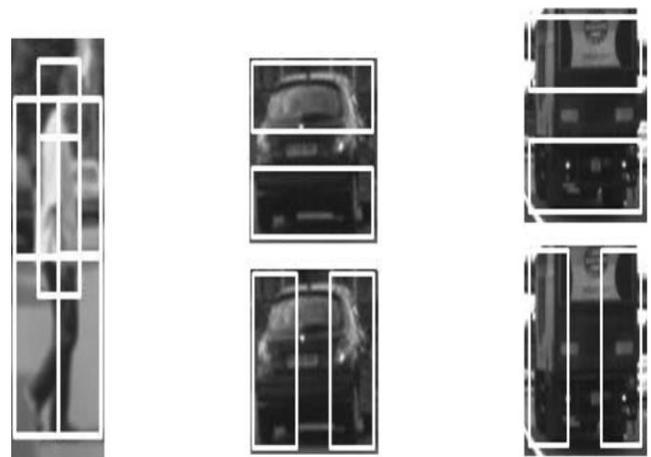


Fig. a. Informative blocks for each object class patch, from left to right: pedestrian, car and truck. Average size of the S-HOG descriptors for pedestrians, bikes, cars and trucks are 216, 216, 288 and 288.[1]

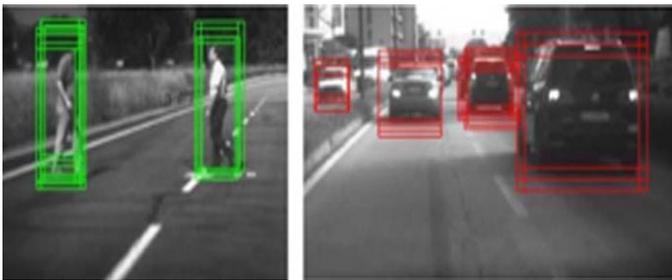


Fig. b. Examples of successful detections of pedestrians (left) and cars (right) from camera images.[1]

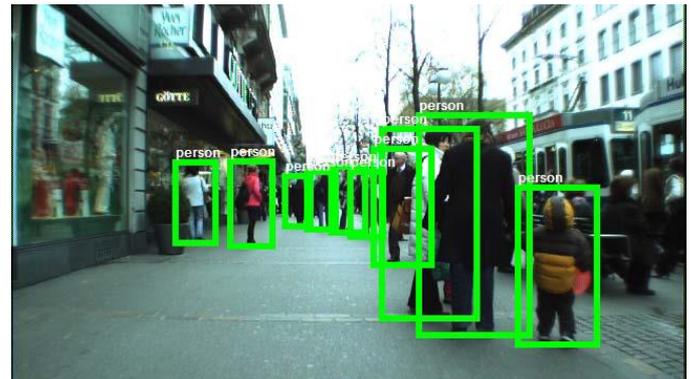
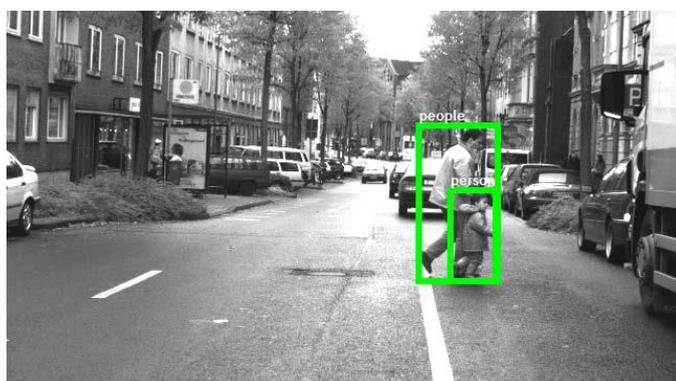
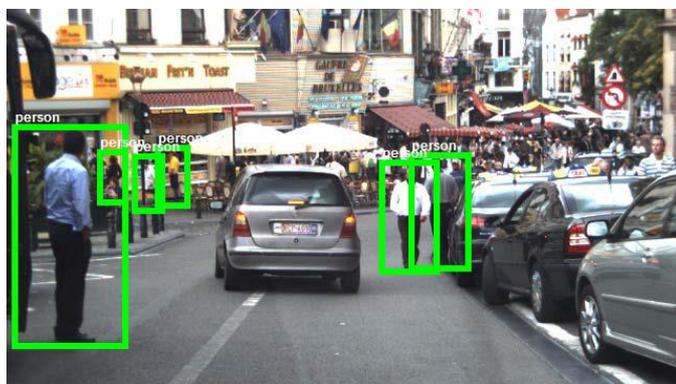


Fig.c. Examples of successful detections of pedestrians (left) and cars (right) from camera images.[3]

3.2. Camera based object classification

Proximity sensor processing detection provides a rough classification of the detected moving objects. This classification relies on the visible shape of the detection which is a strong assumption in a highly dynamic environment. We consider that an look based classification could provide more confidence about the class of the detected objects. Therefore we use the proximity sensor detections to generate regions of interest (ROIs) in the camera images. The ROIs are taken by vehicle and pedestrian classifiers to perform the camera based classification. A modified version of histogram of oriented gradients (called sparse-HOG) features, which focus on important areas of the samples, powers the pedestrian and vehicle visual descriptor at training and detection time.



4. ROPOSED METHODOLOGY

In this section, To improve the results of the perception task, a more reliable list of moving objects of interest represented by their kinematic state and appearance information. within a perception system different fusion levels are present in fig a. block diagram and explanation proposed system focuses on the fusion methods inside DATMO that use proximity sensor and camera. The different fusion levels within a perception system as follows.

4.1. System Arrangement

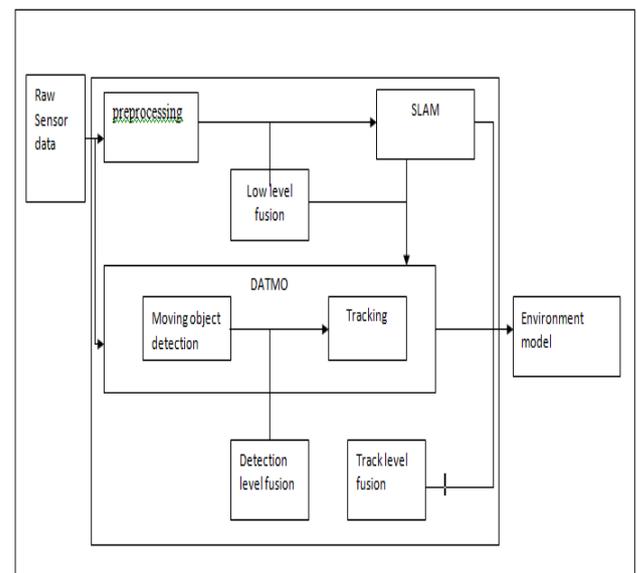


Fig.2. Fusion levels within the SLAM and DATMO components interaction.[1]

a) SIMULTANEOUS LOCALIZATION AND MAPPING (SLAM):

Simultaneous localization and mapping (SLAM) produce a map of the environment while continuously localizing the

vehicle within the map given all the measurements from sensors.

b)DATMO (Detects and Tracks the Moving Objects):

DATMO detects and tracks the moving objects surrounding the vehicle and estimates their future behavior. Within the DATMO task or as aggregate information for the final perception output classification is seen as a separate task. Knowing the class of objects surrounding the ego-vehicle provides a better understanding of driving situations.

5. Conclusion

In This paper we have reviewed the problem of intelligent vehicle perception .Specifically, we have focus on the DATMO component of the perception task. We have proposed the use of classification information as a key element of a composite object representation. We have analyzed the impact of our composite object description by performing multi-sensor fusion at detection level. We used main sensors to define, develop, test and evaluate our fusion approach: proximity sensor and camera.

6. REFERENCES

- [1] Ricardo Omar Chavez-Garcia and Olivier Aycard "Multiple Sensor Fusion and Classification for moving object detection and tracking" IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 17, NO. 2, FEBRUARY 2016
- [2] R. O. Chavez-Garcia, T. D. Vu, O. Aycard, and F. Tango, "Fusion framework for moving-object classification," in Information Fusion (FUSION), 2013 16th International Conference on, July 2013, pp. 1159–1166.
- [3] P. Dollar, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: An evaluation of the state of the art," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 4, pp. 743–761, April 2012.
- [4] R. O. Chavez-Garcia, T. D. Vu, and O. Aycard, "Fusion at detection level for frontal object perception," in 2014 IEEE Intelligent Vehicles Symposium Proceedings, Jun2014, pp. 1225–1230.
- [5] M. Perrollaz, R. Labayrade, C. Royere, N. Hautiere, and D. Aubert, "Long range obstacle detection using laser scanner and stereovision," in 2006 IEEE Intelligent Vehicles Symposium, 2006, pp. 182–187.
- [6] J. Civera, A. J. Davison, and J. M. M. Montiel, "Interacting multiple model monocular slam," in Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on, May 2008, pp. 3704–3709.
- [7] J.-X. Yu, Z.-X. Cai, and Z.-H. Duan, "Detection and tracking of moving object with a mobile robot using laser scanner," in 2008 International Conference on Machine Learning and Cybernetics, vol. 4, July 2008, pp. 1947–1952.