Garcia, F.; Prioletti, A.; Cerri, P.; Broggi, A.; Escalera, A.; Armingol, J. M. (2014). Visual feature tracking based on PHD filter for vehicle detection. *Information Fusion (FUSION), 2014 17th International Conference on*, Salamanca, 2014, pp. 1-6.

Visual Feature Tracking Based on PHD Filter for Vehicle Detection

F. Garcia*, A. Prioletti+, P. Cerri+, A. Broggi+, A. de la Escalera* and J.M. Armingol*

*Universidad Carlos III of Madrid
Madrid, Spain
{fegarcia,escalera,armingol}@ing.uc3m.es

+Vislab
Parma, Italy
{prio, cerri, broggi}@vislab.it

***Abstract*— Vehicle detection is one of the classical application among the Advance Driver Assistance Systems (ADAS). Applications like emergency braking or adaptive cruise control (ACC) require accurate and reliable vehicle detection. In latest years the improvements in vision detection have lead to the introduction of computer vision to detect vehicles by means of these more economical sensors, with high reliability.**

In the present paper, a novel algorithm for vehicle detection and tracking based on a probability hypothesis density (PHD) filter is presented. The first detection is based on a fast machine learning algorithm (Adaboost) and Haar-Like features. Later, the tracking is performed, by means features detected within the bounding box provided by the vehicle detection. The features, are tracked by a PHD filter. The results of the features being tracked are combined together in the last step, based on several different methods. Test provided show the performance of the PHD filter in public sequences using the different methods proposed.

Keywords—PHD Filter; Vehicle detection; Computer vision; Intelligent Transport Systems

I. INTRODUCTION

During latest years the recent advances in information technologies allow to create Advance Driver Assistant Systems (ADAS), technologies that help to detect and warn the driver in advance in case of hazardous situations. These applications require accurate and reliable detection of the vehicles in the road.

Classical sensing technologies (e.g. radar or laser technologies) have showed trustable results regarding to vehicle detection, but the high costs of these sensing technologies difficult the generalization of these kinds of applications and sensors. The modern computer vision technologies allow to detect and track vehicles with lower costs and wider sensing ranges. One of the main drawbacks of computer vision approaches is the trustrability and the reliability. Modern techniques try to overcome this difficulty by the use of modern tracking technologies.

Present work focuses in the use of a probability hypothesis density (PHD) filter to track the features belonging to a detected vehicle. By means of the use of the PHD filter the tracking of the detected vehicle allows to enhance the accuracy and the detection performance.

The paper is divided in the following sections:

Section II provides related works references, giving scientific context. Section III describes the general purpose of the application. Section IV introduces the machine learning algorithm used for vehicle and feature detection and section V details the feature tracking algorithm. Finally, test performed to the tracking algorithm are detailed in section VI and some conclusions are discussed in section VII.

II. STATE OF THE ART

Vision based vehicle detection is a common topic in the in Intelligent Transport Systems.

Sivaraman & Trivedi [1] provide wide state of the art of the current vision based vehicle detection. Regarding to monocular detection, works can be divided on those that use appearance features and optical flow based approaches. Among first, Histograms of Oriented Gradients (HOG), presented on [2] for pedestrian detection applications, are used on [3] and [4] for vehicle detection. Haar-Like features [5], are used in [6] for vehicle detection. In [7] HOG features are used together with laser scanner for vehicle detection and tracking. Optical flow based approaches take advantage of the motion of the vehicles to identify them. In [8] authors combine optical flow and symmetry, and in [4] optical flow is fused with radar for overtaking detection. Other approaches provide vehicle detections based on sensor fusion: In [9] obstacle detection and classification is performed based on radar and computer vision and [7] takes advantage of the laser scanner to enhance the vision detection.

In the latest years, the increasing popularity of the PHD filters in ITS community have lead to modern and useful applications where these estimation tools are used to enhance visual approaches. On [10] authors used PHD filters with Multiple Model for tracking road users (pedestrians and vehicles) in an intersection, based on multiple laser scanners. Authors in [11] used Cardinalized PHD to provide multisensor detection for vehicles and pedestrians.[12] and [13] used PHD together with feature detection to provide visual odometry.

III. GENERAL DESCRIPTION

On this paper, novel vehicle detection and tracking approach is presented. First the vehicle is detected based on machine learning methods. Once the vehicle is detected features are used to track the movement by means of a PHD filter. Finally, the information of the features are combined in the last stage. On this last stage several methods were tested.

Fig. 1. Braive test platform from Vislab.

The proposed method was developed within the BRAiVE (BRAin-drIVE) project (Fig. 1). BRAiVE is the prototype vehicle developed by Vislab for Advance Driver Assistance Systems (ADAS) and Autonomous Driving research. It is equipped with 10 cameras, 5 laser scanners, 1 GPS+IMU, 1 e-Stop system covering a 360° view around the vehicle. All the data supplied are available via the car's CAN bus. Four computers are connected to the sensors in order to equally split the processing load.

IV. VEHICLE DETECTION & FEATURE IDENTIFICATION

A soft-cascade approach based on Adaboost and Haar-like [5] features has been used to detect vehicles. Further reclassification techniques and calibration parameters were also introduced to remove false positives.

Once the vehicle is detected, and the bounding box identified, features are searched based on multiple local convolution, key point and descriptors, extracted from two different hash images as described in [14]. Stable feature locations are obtained by filtering the input images with 5x5 blob and corner masks and, Later it was applied a non-maximum-and non-minimum-suppression.

Bounding box movement is obtained from the optical flow of the features extracted in two consecutive frames.

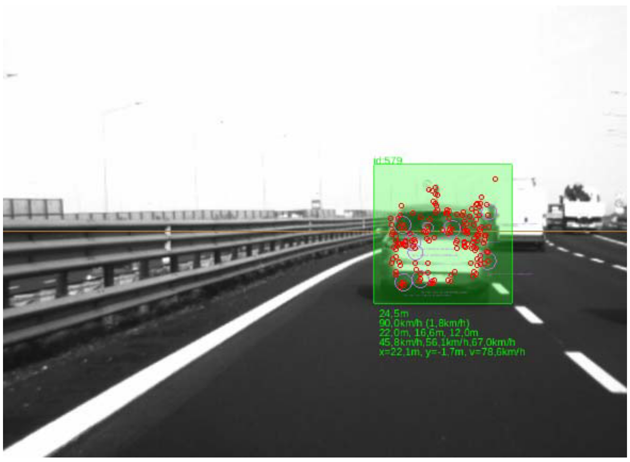

Fig. 2. Example of vehicle detection(green) and feature detection(red circles).

V. FEATURE BASED TRACKING

Once the features are identified and the new bounding box is selected, according to the movement of the features in subsequent scans, these features are introduced in the Gaussian Mixture PHD (GMPHD) filter for vehicle tracking explained next.

Once the vehicle is detected, the gaussians are created according to the estimation of the distance given by the extrapolation of the features (in pixels) to the real world. In order to provide accurate distance estimation, given the monocular approach, first one of the three world coordinates are fixed, and later the other two are obtained base on:

$$\begin{bmatrix} \mathrm{u_i} \\ \mathrm{v_i} \\ 1 \end{bmatrix} = K[R|\hat{t}_0] \begin{bmatrix} \mathrm{x_i} \\ \mathrm{y_i} \\ \mathrm{z_i} \\ 1 \end{bmatrix} = P \begin{bmatrix} \mathrm{x_i} \\ \mathrm{y_i} \\ \mathrm{z_i} \\ 1 \end{bmatrix} \tag{1}$$

where K is the intrinsic calibration parameters matrix, R and T are the extrinsic calibration parameters matrix (rotation and translation). x_i is the fixed coordinated (in meters) which represents the distance to the detected vehicle and y and z are the distance in meters to the feature, lateral and vertical respectively. Coordinates u,v are the pixels of the detected feature in the image.

As it is remarked, distance to the vehicle is obtained by a different approach, prior to the coordinate transformation. The approach for x detection is based on the bounding box created by the vehicle detector and the subsequent feature detection. The detected features are assumed to be all in the same vertical plane of the vehicle, thus they all share the same plane. Assuming that all the features share this coordinate, and that the vehicle is within the ground plane, the x coordinate is obtained taking the center bottom point of the bounding box obtained from the classifier. This approach is only used in the initialization of the PHD. Further observation, follows the pin-hole model, thus no assumptions regarding to the real coordinates should be done.

Once the features within the detected bounding box are identified in the real world, the set of gaussians are created according to the position of the features.

The tracking of the vehicle based on their features is based on the approach of Gaussian Mixture Probability Hypothesis Filter (GMPHD) [15]. This approach allows to estimate the state of the objects (features) based on a sequence of noisy and cluttered observations. These observations and theirs states can be defined as random variables with a state X_k, of a number of features N_k, at time k and with the measurements Z_k, all these variables are represented by a random finite set (RFS). The solution, based on the GMPHD approximation for the multi-object Bayes filter, gives a solution based on the propagation of the first moment (mean state $\hat{X}$) of the objects along time.

Thus the predicted and posteriori Probability Hypothesis Density intensity at a time k is approximated by a Gaussian Mixture and is given by (2) and (3) respectively:

$$v(x)_{k|k-1} = \sum_{i=1}^{J_{k|k-1}} w_{k|k-1}^{(i)} N\left(x; m_{k|k-1}^{(i)}, P_{k|k-1}^{(i)}\right) \quad (2)$$

$$v(x)_{k|k} = (1 - P_D) v_{k|k-1} + \sum_{i=1}^{J_{k|k}} w_k^{(i)} N\left(x; m_{k|k}^{(i)}, P_{k|k}^{(i)}\right) \quad (3)$$

where w is defined in (4) with P_D is the detection probability, thus $(1 - P_D) v_{k \vee k-1}$ factor, represents the possible non-detected features.

$$w_k = \frac{P_D w_{k|k-1} q_k(z)}{K_k + P_D \sum_{i=1}^{J_{k|k-1}} w_{k|k-1}^{(i)} q_k(z)} \quad (4)$$

where P_D is the detection probability and q_k is the probability of the feature to keep alive (not disappearing). K_k is intensity of the clutter. The q(z) is given in (5).

$$q_k(z) = N(z, Hm, R + HPH^T) \quad (5)$$

here N defines the normal distribution with z the observation vector for a given measurement, R is the measurement error, and P is the covariance matrix of the exponential.

This way, each object or feature, is defined by a normal distribution defined by the mean state m and the covariance matrix P, and the intensity factor W.

According to the definition of the GMPHD in [15], the $v(x)_{k|k}$ factor defined in (3) lacks of two factors i.e. the spawned features and the new births. As it is indicated in (6).

$$v(x)'_{k|k-1} = v(x)_{k|k-1} + \beta(x)_{k|k-1} + \gamma(x)_k \quad (6)$$

where $\beta(x)_{k|k-1}$ is the mixed Gaussians that represent the spawned features and $\gamma(x)_k$ are the new births. In this work, the probability of spawned is considered 0, since there is no possibility, by definition, for a feature to spawn. The birth probability is chosen to allow new births to last for some frames, as it is discussed subsequently. The GM for new births is defined according to :

$$\gamma(x)_k = \sum_{i=1}^{L} w_\gamma N\left(x; m^{(i)}, P^{(i)}\right) \quad (7)$$

where w_γ is the initial weight for the new births and m represents the state of the new feature detected and P is the initial Covariance Matrix.

Implementation issues:

Some problems and implementation issues has to be taken into account when designing a GMPHD for feature tracking. Some of them are related with the computational requirements, that are excessive for real time applications. Some other problems are also related with the quality of the measurements and the outliers, that should be discarded.

Number of features to track & pruning technique

One of main disadvantages of the PHD Filter is the high computational requirements, due to the exponential grow of the gaussians being tracked.

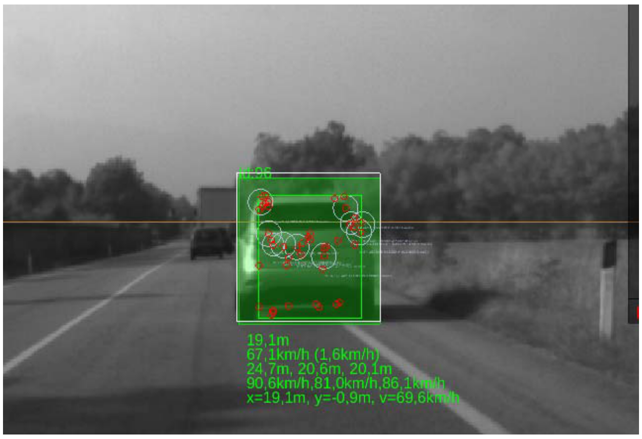

Fig. 3. Example of features being tracked (in white).

This issue is taken into account, by reducing the amount of features being tracked. This is performed by creating an adaptive threshold for gaussians pruning, following equation:

$$W_{th} = \begin{cases} W_{th} = W_{\min}, for J < J_{min} \\ W_{th} = min_{J^{th}}(W), for J > J_{min} \end{cases} \quad (8)$$

Thus this way only those features with a certain weigh, are tracked along time, those with low reliability are thus discarded. By means of this implementation, stable features i.e. those detected with higher frequency are tracked along time. The features with intermittent detection are discarded. Fig. 3 depicts an example of stable features being tracked.

Merging features

Before applying the pruning technique, the tracked exponentials are combined according to the features to which they are assigned. This is done according to the following equation:

$$\widehat{w}_k^{(j)} = \sum_{i=1}^{M} w^{(i)} \quad (9)$$

$$\widehat{m}_k^{(j)} = \frac{1}{\widehat{w}_k} \sum_{i=1}^{M} w^{(i)} x_k^{(i)} \quad (10)$$

$$\widehat{P}_k^{(j)} = \frac{1}{\widehat{w}_k} \sum_{i=1}^{M} w^{(i)} \left(P_k + (\widehat{m}_k - m_k^{(i)})(\widehat{m}_k - m_k^{(i)})^T\right) \quad (11)$$

where M is the set of features pointing to the same feature. And the new feature is defined according to $\widehat{P}_k^{(j)}$, $\widehat{m}_k^{(j)}$ and $\widehat{w}_k^{(j)}$.

Given the specific behavior of the features to track, that usually appears combined with several features closed. This previously presented combination algorithm was extended to the features closed in the image field. The combinations of these features, was performed according a circular gating in the image field, thus if the image was closer than a certain value, they were merged applying (9)-(11).

Vehicle Tracking & Outlier identification

The first idea, is to track the vehicle according to the number of features being tracked. But this configuration is incorrect in the present approach, since, as it was mentioned before, the number of features remains constant, due to the adaptive threshold.

Thus a given vehicle is considered disappeared when the sum of all the weights within a bounding box falls under the threshold for vehicle detection:

$$\acute{w}_k = \sum_{i=1}^{N} w^{(i)} \tag{12}$$

where N is the number of exponential within the bounding box that represents the vehicle.

When the number of features falls under the J_{min} described in (8), the new births features described in

(7) plays an important role, to allow the tracking of the vehicle, updating the features with the new ones. However, the initial value w_γ for new births allows not to interfere when the tracked features have enough stability given by their high weight. It is important to remark that not all the new features created are incorporated. This is due to the fact that some of them are considered clutters, and also to reduce the computational costs of the approach. The approach select randomly according to the probability created denominated $P_{clutter}$.

Finally, the movement of the vehicle has to be estimated based on the state of the features and taking into account the weight of each feature. Several approaches were tested and results are depicted in Test section, in order to chose the most suitable configuration for the vehicle state estimation. Approaches tested are depicted in (13) to (15).

$$\bar{\mathrm{m}}_w = \frac{1}{\bar{\mathrm{w}}_k} \sum_{i=1}^{F} w^{(i)} m_k{}^{(i)} \tag{13}$$

$$\bar{\mathrm{m}}_{Cov} = \frac{1}{\sum_1^F \frac{1}{|\mathrm{P_i}|}} \sum_{i=1}^{F} \frac{1}{|\mathrm{P_i}|} X_i \tag{14}$$

$$\mathrm{m}_{max} = X_i|_{\max(W_i)} \tag{15}$$

where F represents the number of features tracked that not discarded according to (12).

Movement model

The movement used was based in the pin-hole model for Observation and constant velocity model for transition matrix F. The movement was calculated by means of the Unscented transformation that avoids the non-linearity problems of the calculation of the observation

(18), as it is done by the Unscented Kalman Filter (UKF) [16].

$$\hat{X} = \begin{bmatrix} x \\ y \\ z \\ v_x \\ v_y \end{bmatrix} \tag{16}$$

$$\mathrm{Obs} = \begin{bmatrix} \mathrm{u} \\ \mathrm{v} \end{bmatrix} \tag{17}$$

$$Y = \begin{bmatrix} \dfrac{\mathrm{k_u x - u_0}}{\mathrm{z}} \\ \dfrac{\mathrm{k_v y - v_0}}{\mathrm{z}} \end{bmatrix} \tag{18}$$

$$\mathrm{F} = \begin{bmatrix} 1 & 0 & 0 & \mathrm{t} & 0 \\ 0 & 1 & 0 & 0 & t \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{19}$$

$$R = \begin{pmatrix} \sigma^2{}_{\epsilon,u} & 0 \\ 0 & \sigma^2{}_{\epsilon,v} \end{pmatrix} \tag{20}$$

$$Q = diagonal_{5x5}(\sigma^2{}_{\epsilon,i}) \tag{21}$$

F corresponds to the transition matrix, Q process noise, R measurement noise, Y is the prediction, calculated from the state vector X and Obs is the observation vector for the new features detected.

VI. TESTS

Test performed were based on the TME motorway dataset [17]. This dataset provide laser scanner ground truth allowing to compare the performance of the estimation algorithm.

Estimated distance to the vehicles [x] was recorded and compared with the available ground truth, results obtained by the tracking algorithm are depicted in table I. Three methods were tested for mixture of the information provided by the features being tracked. These four methods are depicted in equations (13) - (15).

It has to be remarked that only frames with features detected enough to allow the track were taken into account. Figures 4 to 9 show examples of vehicles being tracked in two representative situations, vehicle following and vehicle being overtaken.

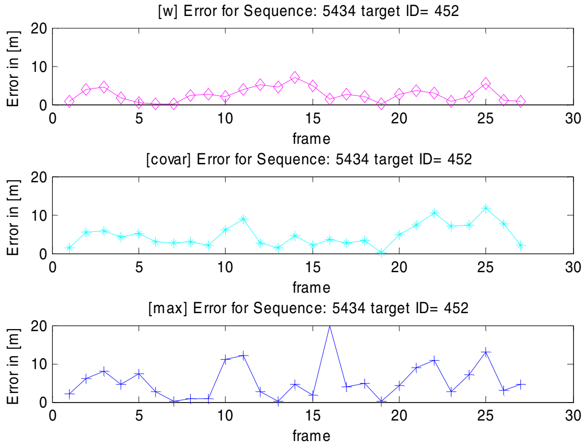

Fig. 4. Tracking example 1. Folowing a vehicle. error calculated for the location estimated using equations (13) top, (14) center and (15) down.

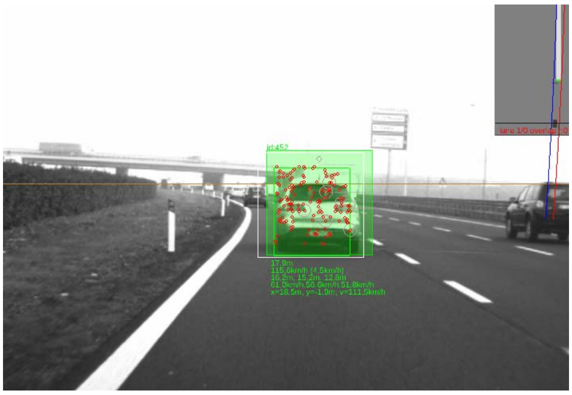

Fig. 5. Tracking example 1. Folowing a vehicle. Frame example.

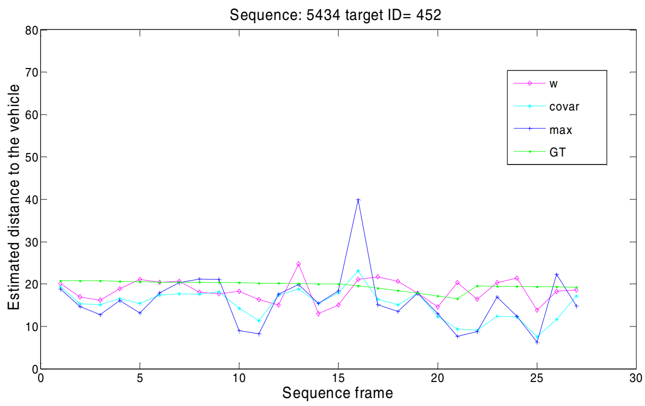

Fig. 6. Tracking example 1. Folowing a vehicle. Distance calculated using the three methods mentioned, magenta (13), cyan (14) and blue (15). Ground thruth is displayed in green.

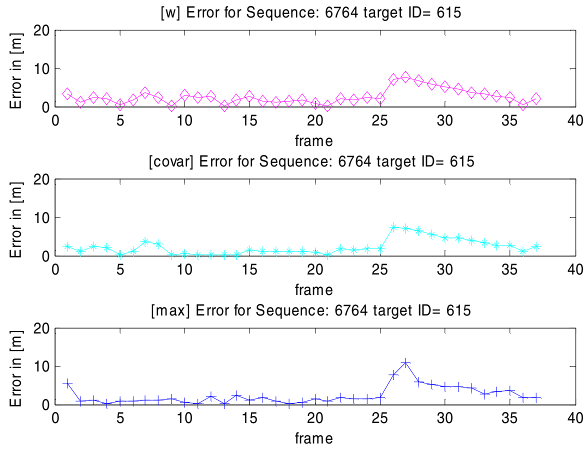

Fig. 7. Tracking example 2. Overtaking a vehicle. Mean error calculated for the elocation estimated using equations (13) top, (14) center and (15) down.

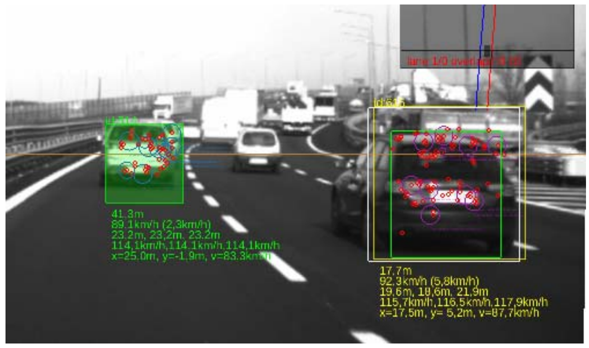

Fig. 8. Tracking example 2. Overtaking a vehicle. Frame example.

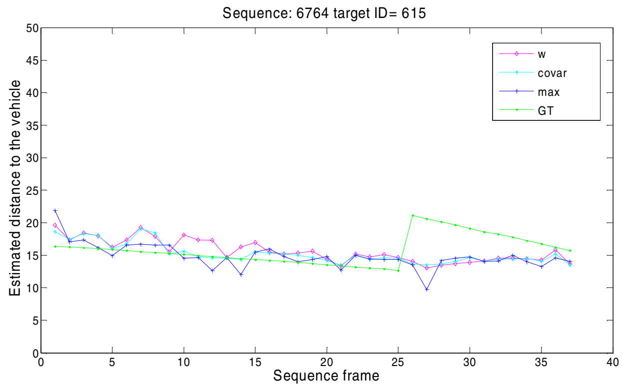

Fig. 9. Tracking example 2. Overtaking a vehicle. Distance calculated using the three methods mentioned, magenta (13), cyan (14) and blue (15). Ground thruth is displayed in green.

TABLE I. MEAN ERROR FOR THE TEST PERFORMED.

Algorithm	Mean error [m]
$\overline{\mathrm{m}}_{w}$	6.67
$\overline{\mathrm{m}}_{Cov}$	6.66
$\overline{\mathrm{m}}_{max}$	7.29

VII. CONCLUSIONS

Novel vehicle tracking algorithm based on GMMPHD filter applied to features was presented, and test results provided.

Results showed the execution of the tracking algorithm, based on the data set provided in [17]. Comparison of the different approaches for feature mixture, provide similar results for the use of the weights (13) and covariance (14). These results were considerably better than the use of the maximum weight method (15). The requirement of covariance determinant calculation, in the case of the covariance (14), makes the weights (13) the best alternative according to the results depicted in table , since they provide similar performance with lower computational costs.

Performance of the tracking algorithm presented makes it a suitable alternative to the available methods, such as KF and UKF, extended and widely used. However, the variability of the results is still an important drawback, new methods to identify and reduce the outliers and provide a smoother performance should be studied.

Future test will compare the performance of the detection algorithm and the enhancement of this detections by means of the use of PHD Filter which tracking performance is described in the present article.

ACKNOWLEDGMENT

This work was supported by the Spanish Government through the Cicyt projects (GRANT TRA2010-20225-C03-01) and (GRANT TRA 2011-29454-C03-02).

REFERENCES

[1] S. Sivaraman and M. M. Trivedi, "A review of recent developments in vision-based vehicle detection," in *Intelligent Vehicles Symposium (IV), 2013 IEEE*, 2013, pp. 310–315.

[2] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," *Comput. Vis. Pattern Recognition, 2005. CVPR 2005. IEEE Comput. Soc. Conf.*, vol. 1, pp. 886–893, 2005.

[3] M. Cheon, W. Lee, C. Yoon, and M. Park, "Vision-Based Vehicle Detection System With Consideration of the Detecting Location," *Intell. Transp. Syst. IEEE Trans.*, vol. 13, no. 3, pp. 1243–1252, Sep. 2012.

[4] S. S. Teoh and T. Bräunl, "Symmetry-based monocular vehicle detection system," *Mach. Vis. Appl.*, vol. 23, no. 5, pp. 831–842, 2012.

[5] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, 2001, vol. 1, pp. I–511.

[6] S. Sivaraman and M. M. Trivedi, "Active learning based robust monocular vehicle detection for on-road safety systems," *2009 IEEE Intell. Veh. Symp.*, pp. 399–404, 2009.

[7] F. Garcia, A. de la Escalera, J. M. Armingol, and F. Jimenez, "Context aided fusion procedure for road safety application," in *Multisensor Fusion and Integration for Intelligent Systems (MFI), 2012 IEEE Conference on*, 2012, pp. 407–412.

[8] J. Arróspide, L. Salgado, M. Nieto, and F. Jaureguizar, "On-board robust vehicle detection and tracking using adaptive quality evaluation," in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, 2008.

[9] M. Bertozzi, L. Bombini, P. Cerri, P. Medici, P. C. Antonello, and M. Miglietta, "Obstacle detection and classification fusing radar and vision," *2008 IEEE Intell. Veh. Symp.*, 2008.

[10] D. Meissner, S. Reuter, and K. Dietmayer, "Road User Tracking at Intersections Using a Multiple-Model PHD Filter," in *Intelligent Vehicles Symposium (IV), 2013 IEEE*, 2013.

[11] L. Lamard, R. Chapuis, and J.-P. Boyer, "CPHD filter addressing occlusions with pedestrians and vehicles tracking," in *Intelligent Vehicles Symposium (IV), 2013 IEEE*, 2013, pp. 1125–1130.

[12] F. Zhang, G. Chen, H. Stahle, C. Buckl, and A. Knoll, "Visual odometry based on random finite set statistics in urban environment," in *Intelligent Vehicles Symposium (IV), 2012 IEEE*, 2012, pp. 69–74.

[13] F. Zhang, H. Stahle, A. Gaschler, C. Buckl, and A. Knoll, "Single camera visual odometry based on Random Finite Set Statistics," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, 2012, pp. 559–566.

[14] A. Geiger, J. Ziegler, and C. Stiller, "StereoScan: Dense 3d reconstruction in real-time," *2011 IEEE Intell. Veh. Symp.*, pp. 963–968, 2011.

[15] B.-N. V. B.-N. Vo and W.-K. M. W.-K. Ma, "The Gaussian Mixture Probability Hypothesis Density Filter," *IEEE Trans. Signal Process.*, vol. 54, no. 11, 2006.

[16] E. A. Wan and R. Van Der Merwe, "The unscented Kalman filter for nonlinear estimation," *Technology*, vol. v, no. 3, pp. 153–158, 2000.

[17] C. Caraffi, T. Vojir, J. Trefny, J. Sochman, and J. Matas, "A System for Real-time Detection and Tracking of Vehicles from a Single Car-mounted Camera," in *ITS Conference*, 2012, pp. 975–982.