Garcia, F., Martin, D., de la Escalera, A. & Armingol, J. M. (2017). Sensor Fusion Methodology for Vehicle Detection. *IEEE Intelligent Transportation Systems Magazine, 9*(1), 123–133.

DOI: 10.1109/mits.2016.2620398

# Sensor Fusion Methodology for Vehicle Detection

**Fernando García, David Martín, Arturo de la Escalera, and José María Armingol**
*IEEE Member, Intelligent Systems Lab,*
*Universidad Carlos III de Madrid,*
*C/ Butarque 15, 28911 Leganés, Madrid, Spain*
*E-mail: fegarcia@ing.uc3m.es*

***Abstract*—**A novel sensor fusion methodology is presented, which provides intelligent vehicles with augmented environment information and knowledge, enabled by vision-based system, laser sensor and global positioning system. The presented approach achieves safer roads by data fusion techniques, especially in single-lane carriageways where casualties are higher than in other road classes, and focuses on the interplay between vehicle drivers and intelligent vehicles. The system is based on the reliability of laser scanner for obstacle detection, the use of camera based identification techniques and advanced tracking and data association algorithms i.e. Unscented Kalman Filter and Joint Probabilistic Data Association. The achieved results foster the implementation of the sensor fusion methodology in forthcoming Intelligent Transportation Systems.

## I. Introduction

Nowadays the safety on interurban roads is still a challenge for vehicle manufacturers, drivers and forthcoming vehicles, because of the elevated cost of the accurate sensors and the limitations of the road safety applications. The presented work in this paper deals with sensor fusion robustness in interurban roads for safety applications. Sensor fusion algorithms are the core of this work, which is aimed at trustable and robust vehicle detection. Furthermore, perception algorithms have been analyzed under the framework of intelligent methods, to allow safe real-time detection of vehicles, estimate dangerous maneuvers and reliable identification of potential risks. This novel sensor fusion in-vehicle methodology will foster the development of safer roads, by means of forthcoming Advanced Driver Assistance Systems (ADAS) applications at high velocities in single-carriageways roads, such as: safe overtaking with evasive maneuvers, safe imminent collision, and safe avoidance maneuvers.

The tests performed allowed to find the optimal approaches for the fusion algorithm, i.e., tracking and data association methods for optimal high-level fusion.

The rest of the article is organized as follows: section II describes state of the art, section III shows the experimental vehicle IVVI 2.0 emphasizing the availability of in-vehicle low-level devices that are used in this work. Section IV presents obstacle detection and classification. Section V explains the tracking and data association fusion procedures for road environments that are the foundation of safe vehicle detection. Section VI shows the obtained results. Finally, conclusions and future works are presented in Section VII.

## II. State of the Art

Different approaches in the intelligent transportation systems related with data fusion are generally divided according to the abstraction level where fusion is performed:

In **low-level** applications, a new set of raw information is created from different sources. One of the classical fusion applications in computer vision is the stereo vision, where information from two cameras is combined, creating the new information i.e. the disparity map. In [1] and [2] stereovision is used to perform pedestrian detection and in [3] this information is used together with laser scanner for this purpose. Other low-level approaches provide sensor in order to enhance the global position systems (GPS) [4].

In **medium-level** approaches, a first pre-processing stage provides a set of features for each sensor. These sets are combined to perform the obstacle detection and classification, in [5] and [6] authors present different approaches whether combining or not the different features of the different sensors comparing results.

In **high-level** fusion approaches detection and classification is performed for each sensor independently, and a final stage combines the detections according to the certainty of the detections and sensors accuracy. In [7] Adaboost vision is used for pedestrian detection and Gaussian Mixture Model Classifier (GMMC) for laser scanner based pedestrian detections. Existence probability based on Dempster-Shafer is used in [8] and [9].

Other approaches among Intelligent Vehicles researches, use data from a laser scanner to detect regions of interest (ROI) in images, and computer vision to classify among different obstacles included in these ROIs. In [10], raw images with SVM machine learning method is utilized. Authors in [11] use Histogram of Oriented Gradients (HOG) features and Support Vector Machine (SVM) classification approach. Following, in [12] Invariant features and SVM are used to perform the vision-based pedestrian detections. These approaches take advantage of the trustability of the laser scanner for obstacle detection, however, fusion is limited to speed up the process by detecting robust ROIs.

Moreover, other sensor fusion approaches take advantage of the properties from various sensors in different way. In [13] information from a stereovision camera and a laser scanner is combined. First, the application uses stereovision information to locate the road. Then it uses this information to remove those obstacles that are irrelevant for the application (i.e. outside the road). Next, it constructs a set of obstacles using the information from both sensors and tracking is performed using a Kalman Filter approach. In [14] it is used information from laser scanner, to search particular zones of the environment where pedestrians could be located and visibility is reduced, such as the space between two vehicles, and performs detections using vision approach. Finally, in [15] laser scanner and radar approach are used for obstacle detection and tracking as well as camera to show the results. Obstacle classification differentiates between moving and non-moving obstacles through computing Mahalanobis distance among the clusters given by the laser scanner.

The presented work in this paper, utilizes monocular based vehicle detection, combined with laser scanner detection. Regarding to the first, works can be divided in those which use appearance features, and optical flow based algorithms. Appearance features such as HOG is presented in [16] for pedestrian detection applications, and in [17] and [18] for vehicle detection. A different approach, is based on Haar-Like features (presented first time on [19] for face detection) and in [20] is used for vehicle detection. Optical flow based approaches take advantage of the movement of the vehicles, in [21] optical flow is fused with radar for overtaking vehicles detection, in [22] authors combine optical flow and visual symmetry information. Other approaches provide vehicle detection with advance tracking approaches, such as [23] which takes advantage of a Probability Hypothesis Density Filter (PHD filter), to enhance monocular detection.

## III. General Description of the IVVI Experimental Platform

The proposed algorithms provide vehicle detection based on laser scanner and computer vision, taking advantage of context and online information provided by the GPS with inertial sensor. The tests performed allowed to find the optimal approaches for the fusion algorithm i.e. tracking and data association methods for optimal high-level fusion. The algorithms were tested in the platform IVVI 2.0 [24] (Fig 1).

The laser scanner provides region of interest (ROI) detection which is later used for both laser scanner and computer vision vehicle detection. These detections are performed based solely on the information provided by each sensor. Later both detections are combined in the fusion stage (high level fusion), creating the fused vehicle tracks.

Regarding data association techniques, two approaches were tested, Global Nearest Neighbors and Joint Probabilistic Data Association.

The work proposes a novel approach for high level computer vision and laser scanner fusion. Based on original laser scanner detection algorithm, presented in [25], which takes advantage of the behavior of the laser scanner and a well-known computer vision approach. For data fusion, several solutions were tested, which were based on advance tracking and data association techniques. Specifically, three tracking techniques were analyzed, Kalman Filter (KF), Unscented Kalman Filter (UKF) and Particle Filter (PF). Regarding to data association techniques, two approaches were tested, Global Nearest Neighbors and Joint Probabilistic Data Association. The tests performed allowed to find the right configuration which provided an advance vehicle detection technology with high trustability.

Three sensors were used in this research, laser scanner, computer vision, as well as a GPS with inertial measurements:

- X-sens MTi-G. It is a GPS device, with an inertial measurement unit included. By the combination of sensors it can be measured, not only the velocity of the vehicle, but also other information, relevant for the applications e.g. Euler angles and angular velocity. This sensing device allowed vehicle movement reconstruction, essential in vehicle shape reconstruction [25].
- Laser Scanner from SICK, model LMS-291-S05. Which provides a field of view of 100 degrees with 0.25 degrees of resolution and a maximum distance of 80 meters.
- Point Grey CCD firewire camera, a digital camera with a resolution 640×480.

## IV. Obstacle Detection and Classification

In this section, low-level detection and classification algorithms used for single sensor vehicle detection are introduced, as well as the calibration techniques used for sensor fusion.

### *A. Laser Scanner Obstacle Detection and Classification*

Laser scanner have an important advantage in comparison with other widely spread sensors, the reliability of their detections. On the other hand, its main disadvantage is the relatively small amount of information provided. However, enough to provide estimation of the obstacle's detected shape, and in specific situation such as vehicle detection, it can provide classification, as depicted in [25] for vehicles or even provide pedestrian detection as explained in [26], combined with computer vision approaches.

The algorithm for vehicle detection, as presented in [25], is composed by 2 stages, at low-level the data is received every 52 milliseconds (19 Hz approximate) and a first shape identification is performed. In the second stage, time integration is performed, providing a higher-level classification. This integration is performed by correlating obstacles in subsequent scans.

### *B. Data Alignment*

Sensors involved in data fusion applications usually do not measure the same physical phenomena; they are not located in the same place nor do they measure using the same coordinate systems, and also they cannot be synchronized. Data alignment tries to give solution to this problem, providing common reference system.

Although both sensors provide 2D information, the coordinate system used does not match, thus coordinate change is mandatory, consequently rotation and translation of the coordinate axis is mandatory: work presented in [26] depicts the specific solution provided for this sensor fusion, based on an on-line calibration solution. Fig. 2 shows example of the results of on line calibration, with the detections of the laser scanner shown in the image plane.


**FIG 1** Test platform IVVI 2.0 (Intelligent Vehicle based on Visual Information 2.0) with the sensors visible.

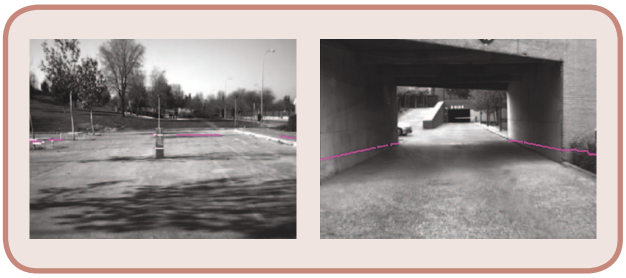

**FIG 2** Example of two calibration sequences where the Euler angles are corrected, varying the extrinsic parameters (yaw, pitch, roll angles and *x,y,z* distances between coordinates axes).

### *C. Vision Based Classification*

Vehicle classification system, used for the present approach, was based on the Haar-Like features. It was presented by Viola and Jones [27]. This approach is based on the use of fast Adaboost classifiers and simple features allowing to classify obstacles in a fast and reliable way.

## V. Tracking and Data Association Fusion Procedures for Road Environments

This section depicts the algorithms constructed in order to provide accurate and trustable tracking procedure and data association techniques based on Multiple Target Tracking (MTT) techniques.

Two main steps are related with tracking: estimation methods and data association techniques. The first one refers to the estimation of the movement along time, the later deals with the matching of new detections with the previous ones. Three estimation methods have been implemented and tested, KF, UKF and PF. These methods have been tested based on the Global Nearest Neighbors (GNN) approach for data association technique in order to provide a test bench able to accurately measure the differences. Regarding to the association methods, two different algorithms have been developed and tested. The first is based on the aforementioned GNN, which is the simplest, and represents the basis of the subsequent procedures. Later Joint Probabilistic Data Association (JPDA) method was adapted and tested to check the advantages or disadvantages of the different procedures by giving an extent comparison of the different possibilities available. Besides, Multiple Hypothesis Tracker [28] and [29] was considered, however this approach requires a high computation costs, and an exponential grow of the number of tracks, thus it was discarded for real time limitations.

### *A. Movement Estimation: Target Model*

The high updating frequency of the detections (around 19 detections per second) allows modelling vehicle movement using constant velocity model. This model is defined in equations (1) and (2), which present the system error $Q$ and measurement error $R$ covariance matrixes, and the constant velocity model equations are depicted in equations (3) to (6) used in the UKF formulation.

$$Q = \begin{bmatrix} \frac{a_x^2 t^3}{3} & \frac{a_x^2 t^2}{2} & 0 & 0 \\ \frac{a_x^2 t^2}{2} & a_x^2 & 0 & 0 \\ 0 & 0 & \frac{a_y^2 t^3}{3} & \frac{a_y^2 t^2}{2} \\ 0 & 0 & \frac{a_y^2 t^2}{2} & a_y^2 \end{bmatrix} \quad (1)$$

$$R = \begin{pmatrix} \sigma^2_{\epsilon,x} & 0 \\ 0 & \sigma^2_{\epsilon,y} \end{pmatrix} \quad (2)$$

where $\sigma_{\epsilon,x}$ y $\sigma_{\epsilon,y}$ is the standard deviation for the measures in $x, y$ coordinates. These deviations have been calculated using test sequences as presented in [26] for pedestrians. As both systems share the ROI coordinates, the deviation in the measurements is considered equal for both detections. The values $a_x$ and $a_y$ in equation (1(2) is the maximum value of the acceleration in every axis.

$$\hat{X} = \begin{bmatrix} x \\ y \\ v_x \\ v_y \end{bmatrix} \quad (3)$$

$$\hat{Y} = \begin{bmatrix} x \\ y \end{bmatrix} \quad (4)$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (5)$$

$$A = \begin{bmatrix} 1 & 0 & t & 0 \\ 0 & 1 & 0 & t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)$$

where $v_x$ and $v_y$ are the speed of the vehicles and $t$ is the time elapsed, $\hat{X}$ is the state vector, $\hat{Y}$ the measurements vector, $H$ the transition matrix and $A$ the state transition model, used in the prediction of the movement of the vehicles using the constant velocity model.

Estimation methods used in the p resent work included the following three ones: KF, UKF and PF. The next section depicts the results provided for each method in the different tests performed.

The KF approach was created using the linear model used (3-6). The UKF, based on the same model, allows modelling the linearity errors thanks to the unscented solution. Finally, the particle filter formulation allows a more general solution, in exchange of computational costs.

### *B. Association Techniques*

#### 1. Global Nearest Neighbor Solution for Data Fusion in Road Environments

The simplest solution when dealing with any fusion applications is a GNN approach. It consists, in the selection of the most suitable solution at a given time and using them in the subsequent scans, discarding the less likely ones.

This approach represented the basis of subsequent approaches that use more complexes solutions (i.e. JPDA) also created in the scope of the present work. Besides, the simplicity of the algorithm for both developing and configuring made it the best solution for testing the estimation filters. Thus, the GNN approach was used to test the estimation filters presented before, for the target model presented.

The use of consolidated and non-consolidated tracks helps the system to add reliability. Only consolidated tracks are considered trustable detections.

### a. Gating and Data Association

First step in the association process, is to discard those non-likely association, to do so, a gating procedure is created. This Gating procedure creates a region where the most likely combinations are included (i.e. only detections closer to a track than its respective gate, are considered for pairing). The associations that do not fall within the gate are considered non-likely and thus discarded. Gating is performed using a square approach (7):

$$\text{Gate} = K\,\sigma_r \tag{7}$$

where $\sigma_r$ is the residual standard deviation and $K$ is a constant that was empirically chosen (according to [28] and [29] typically it is chosen $K \geq 3$).

After gating, association is performed using Euclidean normalized distance and a stability factor, giving less priority to less stable measures, as presented in (8):

$$d^2 = \frac{(x_i - \bar{x})^2}{\sigma_x^2} + \frac{(y_i - \bar{y})^2}{\sigma_y^2} + \ln(\sigma_x \sigma_y) \tag{8}$$

where $d$ is the computed distance between previous and presented tracks to be associated, and ($\sigma_x$, $\sigma_y$) the appropriate values of covariance matrix of Kalman Filter.

Assignment Matrix was used to track association, following a least overall cost assignment [28] and [29].

### b. Track Creation and Deletion Logic

Track creation and deletion policy is based on the definition of two different kind of tracks: consolidated and non-consolidated. First refers to those tracks that were confirmed by both sensors whether concurrently or in subsequent scans. Later refers to tracks detected by a single sensor, thus it is not trustable enough since the other sensor have not corroborated it. So, it is important to define the policy for creation and detection as well as to define when a track is updated i.e. it is considered to have a match with the new detections. This policy was empirically obtained and is depicted in table 1.

The use of consolidated and non-consolidated tracks helps the system to add reliability. Only consolidated tracks are considered trustable detections. Hence, only them are reported. This way false positives, mainly from the laser scanner, are discarded because detections that are not corroborated by the vision system are not reported. Furthermore, the use of both sensors, once the track is consolidated, to update the tracks, allows that once a pedestrian is detected it can be tracked, even if it is not detected by one of the subsystems, e.g. when it goes out of the camera field of view.

### c. Estimation

As it was explained before, the estimation methods used for the GNN solution were three, KF, UKF and PF. The solution to the different tests performed, and the conclusions are given in results section.

## 2. Joint Probabilistic Data Fusion Association Filter for Road Environments

Data fusion approach using the JPDAF (JPDA Filter) was proposed to overcome the limitations of the basic GNN approach, obtained in the different tests and which are explained in test and results section.

Table 1. Track management logic for GNN solution, according to the sensors that detects it in the updating process.

| Track vs New Observation | Single Sensor | Both Sensors | No Match |
|---|---|---|---|
| Non-consolidated | If sensor 1 = detected & sensor 2 = detected then track consolidated. Otherwise non-consolidated. Track updated. | Track consolidated. Track updated. | If #consecutive_no_detections > 4 then Track eliminated. |
| Consolidated | Track updated | Track updated | If #consecutive_no_detections > 5 then Track eliminated. |
| No match | New non-consolidated track | New consolidated track. | – |

Only the assigned observation can contribute to change the current status of a single track from non-consolidated to consolidated or to compute for deletion of nondetected.

Three steps were created in the association algorithm. Following a process similar to the presented in approaches the three steps to follow are:

- Assignation process: where all the joint probabilities are calculated, and a single assignation is performed for each track.
- Filter updating: each track filter is updated with information of all the observation within the gate of the track.
- Track management: new tracks are created and deleted following the logic given by the fusion procedure.

a. Assignation Process

In this step, the probabilities for all possible hypotheses are computed and the joined probabilities calculated. The joint association probability (9), represents the probability of a joint association event $\theta_{kj}^{m}$ that associated measurement $m$ to track $j$ at a time $k$.

$$P(\theta \mid Z_k) = P_D{}^{M-n}\,(1-P_D)^n P_{FA}{}^{m_k-(1-M)} \prod_{j=1}^{m_k} g_{i,j} \qquad (9)$$

where $P_D$ is the detection probability and $P_{FA}$ is the false alarm probability, $M$ is the number of targets being tracked and n is the tracks assigned to the clutter (no associations). Finally $g_{i,j}$ (association likelihood) is defined assuming a 2 dimensional Gaussian association likelihood, for all the measurements to the target:

$$g_{i,j} = \frac{1}{(2\pi)^{N/2}\sqrt{|S_{ij}|}} e^{-\frac{d_{i,j}^2}{2}} \qquad (10)$$

where $d_{i,j}$ is the Euclidean distance between the prediction and the observation. $S_{i,j}$ is the residual covariance matrix. Assuming independence of errors in Cartesian coordinates, we would have $\sqrt{|S_{ij}|} = \sigma_x \sigma_y$ and $N = 2$.

Thus, an assignation matrix is created, where each row represents an observation, and each column a track. As a result, probabilities for all the combinations are computed. The assignation is performed according to this matrix. The assignation is performed following a 1/1 assignment. It means that only one track is assigned to a given observation. In this way, an observation only can be assigned to a single track, all tracks with no assignment would increase their counter for non-detection, and if the counter reaches to a given value, they are eliminated.

b. Filter Update

All the possible associations for a given track are taken into account in the actualization process; thus, the correction is performed using the joined probabilities calculated. All the association hypotheses are weighted in the updating stage of the estimation filter. The innovation of the filter is calculated with all possible combinations, weighted accordingly to the association likelihood values obtained.

$$I_k = \sum_{i=1}^{m} [P(\theta \mid Z_k)(Z_{i,k} - H_k \hat{X}_{k|k-1})] \qquad (11)$$

where $I_k$ is the innovation for the estimation filter of a given track, vectors $Z$, $H$ and $\hat{X}$ are corresponds to the matrixes described in (3-6) for a given target $i$ in a time $k$.

This process of updating stage takes into account all the detections within the gate. It is possible that a given observation is used in the update stage of the filter for more than one track, although it is considered to belong to a single track for the track management process. Thus, for track management, a single assignation policy is followed, a given observation is used only to update the track logic of a single track. This dual behavior is one of the main differences to the classic JPDA filter applications.

c. Track Management

Track management logic follows GNN logic presented before, here a single observation can only update the status of a single track, although in practice, it can be used to update the filter of more than one track. Thus only the assigned observation can contribute to change the current status of a single track from non-consolidated to consolidated or compute for deletion of non-detected.

Track creation logic is different in this case. Only when an observation is out of the gate of any track, a new track is created. This solution is interesting to avoid false positives related with more than one camera positive detection of the same vehicle as it will be detailed in next section.

## VI. Results

Test were performed in real driving conditions to check the viability of each algorithm independently as well as the fusion system. These tests included a single vehicle performing all kinds of maneuvers, such as overtaking maneuvers, being overtaken, driving in front of the test vehicle and performing turnings (including roundabouts). All those tests were mainly performed in inter-urban scenarios with 28 different sequences used in each tests, recorded with IVVI 2.0, over a total of more than 4000 detections.

Two association methods were implemented and compared for the vehicle detection approach, GNN and JPDA. All of them were implemented using UKF.

### *A. Low-Level Algorithms Performance for Vehicle Detection*

A first test of the laser scanner based vehicle detection algorithm performance was conducted, results are depicted in [25]. Results showed that in the best conditions (static vehicle) and with direct vision, the algorithm can reach to a 100% of detection up to 35 meters in approaching movements and 62 meters in moving away movements. As it was explained in the paper, these lower results in approaching movements were due to the lower reflectivity of front parts of vehicles.

The performance of the single sensor approaches in the test performed are depicted in Table 2 and 3 for camera and laser respectively. The percentage represents the number of detections among the available targets, as well as the percentage of misdetections per frame.

Some details should be highlighted before analyzing the above results:

Laser scanner detection presents a high amount of false positives. Here, only those that are in the camera field of view were taken into account. These errors were frequent in movements involving lateral movements, strong braking or acceleration movements. The inertial device resulted insufficient to avoid these errors due to wrong polyline construction. As a result, it is in these situations where the fusion process has special importance. Thus, the high positive rate of the laser scanner is important. However, special attention should be taken to the amount of false positives: in one of every ten frames a false positive is returned. Consequently, it is clear that fusion approach is necessary to overcome these problems related to this technology.

Vision approach, on the other hand, has a small positive detection rate (approx. 50%). It is mainly because the training stage was performed keeping in mind further fusion stages. It was rather interesting having a small positive rate which provides a small false positive rate. It also has to be remarked that this system lacks of tracking that could help to provide better performances.

The results showed that if the camera is able to give a positive detection of a vehicle every 2 or 3 frames, with enough reliability due to the low false positives, it should help to create a trustable and reliable fusion system.

After analyzing the results, the conclusions obtained from these tests are summarized in the next points:

- Laser scanner detection algorithms provide a high amount of positives as well as false positives, that later should be avoided in subsequent fusion stage.
- Vision algorithms can be used to add trustability to these laser scanner detections due to its low false positive rate. However, the low positive rate means that this algorithm is not robust enough, and fusion is necessary to add robustness.

### *B. Tracking Algorithm Performance*

Again, the three different tracking algorithms were tested in the mentioned sequences, using the basic GNN approach and the results are provided in Table 4. The difference between the predicted position and the observation provided by the laser scanner are checked and the standard deviation calculated.

The results provided were very similar for the three algorithms, although UKF and PF obtained better results than the KF approach. The similar accuracy of PF and UKF, together with the less computational costs of the latest makes it the most suitable solution.

### *C. Association Methods Performance*

Two association methods were implemented and compared for the vehicle detection approach, GNN and JPDA. All of them were implemented using UKF.

Several tests were performed to check the differences in the algorithms presented. In the present section each one of them is going to be compared with the classic GNN method. Although overall results were similar, specific

Table 2. Vision based vehicle detection performance.

| | % of Positive Detection | % of Misdetections (per Frame) |
|---|---|---|
| Total (camera): | 47.72 | 1.13 |

Table 3. Laser scanner based vehicle detection performance.

| | % of Positive Detection | % of Misdetections (per frame) |
|---|---|---|
| Total (laser): | 91.03 | 8.19 |

Table 4. Estimation filter performance comparison.

| σ KF | σ UKF | σ PF |
|---|---|---|
| 0.45 [m] | 0.44 [m] | 0.44 [m] |

σ represents standard deviation from the prediction to the laser scanner measurements.

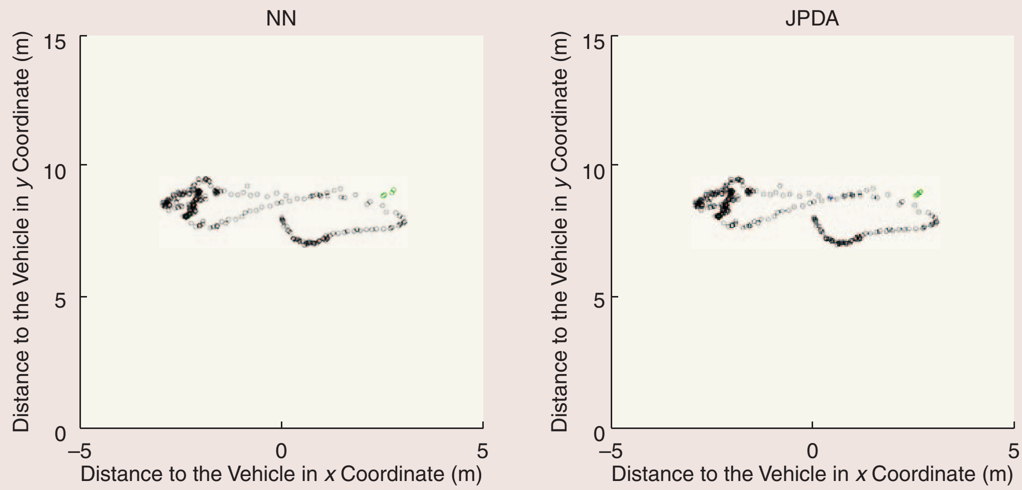


**FIG 3** Results of the tracking of a single vehicle after several movements (distance coordinates in meters), including a roundabout. JDPA results are shown in right and GNN in left image. Green detections represent estimation with no match, black are matching observations.

tests, showed important differences. To give an idea of the differences in performances of the algorithms, three different sequences are going to be detailed:

- The first consists of basic movements, where the IVVI 2.0 was moving following another vehicle.
- The second sequence involves a more complex movement. IVVI 2.0 performs an overtaking maneuver over two cars.
- The third sequence is the most difficult one, with two vehicles entering in the field of view, one of them performing incorporation to a road, and the other moving inside this road.

### a. Test 1

Behavior in this test, depicted in Fig. 3 and 4, gives an interesting result. As can be observed in the figure 3, both sequences' results are identical, as well as the time consumed to perform the detection. This situation was interesting to demonstrate that both algorithms can be considered equivalent when there is a single obstacle in the environment. Some false positives were given from the single scanner algorithms, mainly from laser scanner, but both algorithms were able to eliminate them.

**FIG 4** Frame examples of the test performed with the IVVI 2.0 following a vehicle. The movement involved turnings inside a roundabout.

### b. Test 2

In this sequence, the test platform (IVVI 2.0) performs an overtaking maneuver over two vehicles, the results depicted in Fig. 5 shows similar results for both subsystems. Detection performed by the JPDA approach was able to follow the vehicle even outside the field of view of the camera (red circle). It is important to point that the delays to perform the assignations in this test were similar.

### c. Test 3

This sequence presents more difficulties, for having more than one vehicle in the sequence with occlusions included (Fig. 6). Also, it is important to notice that the IVVI 2.0 platform is performing a high turn that leads to a high number of false positives in the laser scanner algorithm, but thanks to the vision system, these false positives are discarded. However, the problem is that all those false positives introduce new tracks to perform the algorithm, which lead to certain delays in the assignment. Besides this problem, the algorithm, as it is shown in Fig. 7, gives better results, even though the difference between both algorithms is small. The delays presented could easily be overcome with more powerful computers or a parallel programming approach.

The conclusion of the present test was that the best algorithm is the JPDA in both providing better results, and also giving smoother behavior in the tracking process. According to the test, the latter is the cause that allows keeping the track, even in these situations where the GNN approach loses it. On the other hand, it is important to point out that this algorithm has the drawback of the computation costs,

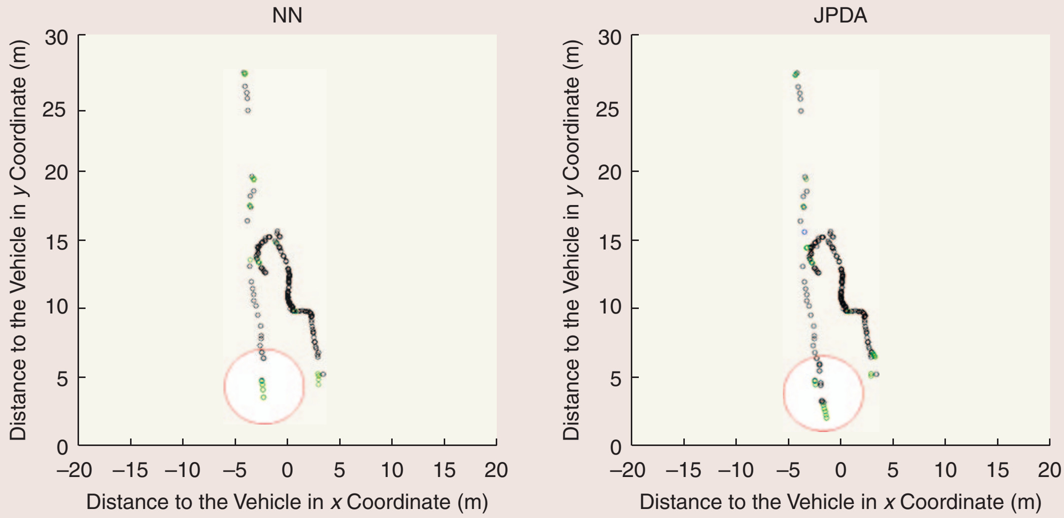


**FIG 5** Results of the test sequence with the vehicle (IVVI 2.0) performing an overtaking maneuver (distance coordinates in meters). JDPA and GNN results are shown (right and left respectively). Green detections represent estimation with no match, black are matching observations.

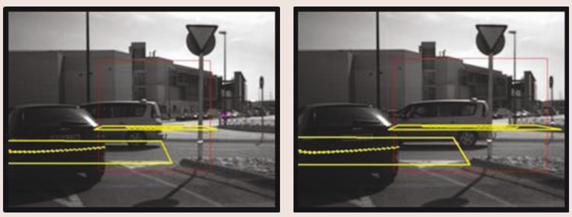

**FIG 6** Frames of the test sequence with two vehicles crossing.

thus special care should be taken to avoid situations with lots of tracks in case of the use of JDPA approach. It should also be remarked that the GNN approach showed good performances; hence, this algorithm should also be taken into account in case of the necessity of a less cost demanding algorithm with good performances.

### *D. Fusion System Performance*

Results obtained by the whole set of tests are depicted in Table 5. The performance of the different subsystems

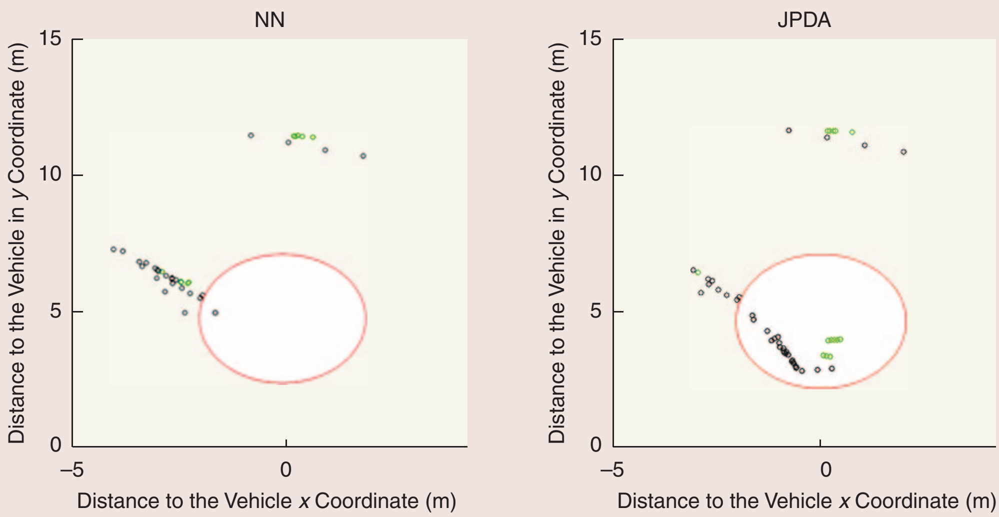


**FIG 7** Results of the test sequence with two vehicles in a crossing (distance coordinates in meters). JDPA results are shown at right and GNN at left in the image. Green detections represent estimation with no match, black are matching observations.

Table 5. Overall results for the test of the vehicle algorithms.

| | Camera | | Laser Scanner | | Fusion | |
|---|---|---|---|---|---|---|
| | % of Positive Detections | % Miss (Per Frame) | % of Positive Detections | % Miss (per Frame) | % Positive Detections | % Miss (per Frame) |
| Total: | 47.72 | 1.13 | 91.03 | 8.19 | 92.03 | 0.59 |

independently is included to allow contrasting the performances of the whole system and each system independently. The fusion algorithm consisted of a JPDA approach with UKF estimation.

Table 5 depicts the overall results of the complete system. The following points summarize the results obtained.

The main goal for the fusion system in this approach was to maintain the good results of the laser scanner system providing reliability to the detection by reducing the amount of false positives. As it is depicted in the table 5, it was possible to accomplish these requirements due to the fusion procedure.

Given the high positive rates of the laser scanner, the task of increasing the positive rate of the overall system in comparison to the laser scanner system was very difficult. Even though it proved extremely difficult, it was slightly increased.

## VII. Conclusions

The main goal of the presented work was to provide sensor fusion methodology for intelligent vehicles, able to overcome the limitations of each sensor, providing a robust and reliable safety application for road environment. The results provided show that by fusing the information of the computer camera and a laser scanner, and using other information sources (i.e. context and inertial system) it was possible to accomplish the complex task of safe vehicle detection in inter-urban scenarios. Furthermore, the presented systems give the possibility to increase the set of sensors thanks to its scalability.

This work represents a step forward on sensor fusion for road safety applications. First the GNN fusion based approach is presented, based on UKF estimation filter and both computer vision and laser scanner technology to detect and track vehicles. However it was improved by the JPDA algorithm, this approach proved to be a reliable application able to work in real time in complex environments. The tests proved that this approach provides better performances than the GNN approach, also tested on the work. All these proved that the main purpose of this work was fulfilled, such as to enhance the capacities of basic sensors in road environment by using data fusion techniques.

Future works will try to enhance vehicle detection with new tracking techniques, such as Probability Hypotheses Density Filters (PHD), and compare the possible results with the presented techniques. Moreover, context is a new issue in data fusion. The present approach adds context information as a new information source. Future works should go forward in the use of context information to enhance classical detection algorithms e.g. accurate GPS localization and digital map with relevant information such as zebra crossings, strong turns or possible occlusion detection.

## About the Authors

***Fernando Garcia*** was born in Badajoz, Spain City, in 1982. He received the Eng. degree in Telecommunications in 2007, M.S. in Robotics and Automatics in 2007 and, PhD in Robotics in 2012, all in Universidad Carlos III de Madrid.

He is with Intelligent System Lab since 2007. Since 2009, he has been an Assistant Professor with the Systems and Automatics department in Universidad Carlos III de Madrid. His research interests include data fusion, computer vision, intelligent vehicles and human factors in intelligent vehicles.

Dr. Garcia was a recipient of the VII Barreiros Foundation award to the best research in the automotive field, he was also awarded with the award to the best PhD dissertation 2013-2015, by the Spanish chapter ITSS.

***David Martín***, graduated in Industrial Physics (Automation) from the National University of Distance Education (UNED, 2002) and Ph.D. degree in Computer Science from the Spanish Council for Scientific Research (CSIC) and UNED, Spain 2008.

He was Ph.D. student at CSIC from 2002 to 2006. He was fellow at the European Organization for Nuclear Research (CERN, Switzerland, 2006–2008) and Post-Doc researcher in Robotics at CSIC (2008–2011). Currently, he is Professor and Post-Doc researcher at Carlos III University of Madrid and member of the Intelligent Systems Lab since 2011. His research interests are Computer Vision, Sensor Fusion, Intelligent Transportation Systems, Advanced Driver Assistance Systems, Autonomous Ground Vehicles, Unmanned Aerial Vehicles, and Vehicle Positioning and Navigation.

In 2014, he was awarded with the VII Barreiros Foundation award to the best research in the automotive field. In 2015, the IEEE Society has awarded Dr. Martín as the best reviewer of the 18th IEEE International Conference on Intelligent Transportation Systems.

***Arturo de la Escalera*** engineering degree from the Universidad Politecnica de Madrid (Madrid, Spain) in Automation and Electronics Engineering in 1989, where he also obtained his Ph.D. degree in Robotics in 1995.

He founded the Intelligent System Laboratory in the year 2000. His current research interests include Advanced Robotics and Intelligent Transportation Systems with special emphasis on vision sensor systems and image data processing methods for environment perception and real-time pattern recognition.

Dr. Escalera is heading the Intelligent Systems Lab since 2005. In 2014, Arturo was awarded with the VII Barreiros Foundation award to the best research in the automotive field.

***José María Armingol*** graduated from the Universidad Politécnica de Madrid (Spain) in Automation and Electronics Engineering in 1992. He received his Ph.D. in Robotics and Automation from the Universidad Carlos III de Madrid in 1997.

He is Full Professor in the Systems Engineering and Automation Department at Universidad Carlos III de Madrid since the year 2012 and co-founded and member of the Intelligent Systems Lab since the year 2000. His research interests focus on Computer Vision, Image Processing and Real-Time Systems applied to Autonomous Vehicles and Advanced Driver Assistance Systems.

In 2014, he was awarded with the VII Barreiros Foundation award to the best research in the automotive field.

## References


[1] M. Bertozzi, A. Broggi, M. Felisa, S. Ghidoni, P. Grisleri, G. Vezzoni, C. Hilario Gómez, and M. Del Rose, "Multi stereo-based pedestrian detection by means of daylight and far infrared cameras," in *Object Tracking and Classification Beyond the Visible Spectrum*, R. I. Hammoud, Ed. *New York*: Springer-Verlag, 2009, pp. 371–401.

[2] C. Hilario, J. Collado, J. Armingol, and A. La Escalera, "Pedestrian detection for intelligent vehicles based on active contour models and stereo vision," *Proc. Computer Aided Systems Theory Conf.*, 2005, pp. 537–542.

[3] B. Musleh, F. García, J. Otamendi, J. M. Armingol, and A. De La Escalera, "Identifying and tracking pedestrians based on sensor fusion and motion stability predictions," *Sensors*, vol. 10, no. 9, pp. 8028–8053, 2010.

[4] E. D. Martí, D. Martín, J. García, A. de la Escalera, J. M. Molina, and J. M. Armingol, "Context-aided sensor fusion for enhanced urban navigation," *Sensors (Basel)*, vol. 12, no. 12, pp. 16802–16837, Jan. 2012.

[5] C. Premebida, O. Ludwig, M. Silva, and U. Nunes, "A cascade classifier applied in pedestrian detection using laser and image-based features," in *Proc. IEEE Intelligent Transportation Systems Conf.*, 2010, pp. 1153–1159.

[6] C. Premebida, O. Ludwig, and U. Nunes, "LIDAR and vision-based pedestrian detection system," *J. F. Robot*, vol. 26, no. IV, pp. 696–711, 2009.

[7] C. Premebida, G. Monteiro, U. Nunes, and P. Peixoto, "A lidar and vision-based approach for pedestrian and vehicle detection and tracking," in *Proc. IEEE Intelligent Transportation Systems Conf.*, 2007, pp. 1044–1049.

[8] M. Aeberhard, S. Paul, N. Kaempchen, and T. Bertram, "Object existence probability fusion using Dempster-Shafer theory in a high-level sensor data fusion architecture," in *Proc. IEEE Intelligent Vehicles Symp.*, 2011, pp. 770–775.

[9] J. P. Hwang, S. E. Cho, K. J. Ryu, S. Park, and E. Kim, "Multi-classifier based LIDAR and camera fusion," in *Proc. IEEE Intelligent Transportation Systems Conf.*, 2007, pp. 467–472.

[10] M. Aeberhard, S. Paul, N. Kaempchen, and T. Bertram, "Object existence probability fusion using Dempster-Shafer theory in a high-level sensor data fusion architecture," in *Proc. IEEE Intelligent Vehicles Symp.*, 2011, pp. 770–775.

[12] T. Denoux, N. El Zoghby, V. Cherfaoui, and A. Jouglet, "Optimal object association in the dempster-shafer framework," *IEEE Trans. Cybern.*, vol. 44, no. 12, pp. 2521–2531, 2014.

[11] O. Ludwig, C. Premebida, U. Nunes, and R. Ara, "Evaluation of boosting-SVM and SRM-SVM cascade classifiers in laser and vision-based pedestrian detection," in *Proc. IEEE Intelligent Transportation Systems Conf.*, 2011, pp. 1574–1579.

[12] A. Pérez Grassi, V. Frolov, and F. Puente León, "Information fusion to detect and classify pedestrians using invariant features," *Inf. Fusion*, vol. 12, no. 4, pp. 284–292, 2010.

[13] R. Labayrade, C. Royere, D. Gruyer, and D. Aubert, "Cooperative fusion for multi-obstacles detection with use of stereovision and laser scanner," *Auton. Robots*, vol. 19, no. 2, pp. 117–140, 2005.

[14] A. Broggi, P. Cerri, S. Ghidoni, P. Grisleri, and H. G. Jung, "Localization and analysis of critical areas in urban scenarios," in *Proc. IEEE Intelligent Vehicles Symp.*, 2008, pp. 1074–1079.

[15] H. Cheng, N. Zheng, X. Zhang, J. Qin, and H. Van De Wetering, "Interactive road situation analysis for driver assistance and safety warning systems: framework and algorithms," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 1, pp. 157–167, 2007.

[16] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Computer Vision Pattern Recognition Conf.*, 2005, vol. 1, pp. 886–893.

[17] M. Cheon, W. Lee, C. Yoon, and M. Park, "Vision-based vehicle detection system with consideration of the detecting location," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 3, pp. 1243–1252, Sept. 2012.

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

[19] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. IEEE Computer Vision and Pattern Recognition Conf.*, 2001, vol. 1, pp. I–511.

[20] S. Sivaraman and M. M. Trivedi, "Active learning based robust monocular vehicle detection for on-road safety systems," in *Proc. IEEE Intelligent Vehicles Symp.*, 2009, pp. 399–404.

[21] F. Garcia, P. Cerri, A. Broggi, A. de la Escalera, and J. M. Armingol, "Data fusion for overtaking vehicle detection based on radar and optical flow," in *Proc. Intelligent Vehicles Symp.*, 2012, pp. 494–499.

[22] J. Arróspide, L. Salgado, M. Nieto, and F. Jaureguizar, "On-board robust vehicle detection and tracking using adaptive quality evaluation," in *Proc. 15th IEEE Int. Conf. Image Processing*, 2008.

[23] F. Garcia, A. Prioletti, P. Cerri, A. Broggi, A. de la Escalera, and J. M. Armingol, "Visual feature tracking based on PHD filter for vehicle detection," in *Proc. IEEE Int. Conf. Information Fusion*, 2014, pp. 1–6.

[24] D. Martín, F. García, B. Musleh, D. Olmeda, G. Peláez, P. Marín, A. Ponz, C. Rodríguez, A. Al-Kaff, A. de la Escalera, and J. M. Armingol, "IVVI 2.0: An intelligent vehicle based on computational perception."

[25] F. García, F. Jiménez, J. E. Naranjo, J. G. Zato, F. Aparicio, J. M. Armingol, and A. de la Escalera, "Environment perception based on LIDAR sensors for real road applications," *Robotica*, vol. 30, no. 02, pp. 185–193, 2012

[26] F. García, J. García, A. Ponz, A. de la Escalera, and J. M. Armingol, "Context aided pedestrian detection for danger estimation based on laser scanner and computer vision," *Expert Syst. Appl.*, vol. 41, no. 15, pp. 6646–6661, 2014.

[27] P. Viola, and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. IEEE Computer Vision Pattern Recognition Conf.*, 2001, vol. 1, pp. I–511–I–518.

[28] S. S. Blackman, *Multiple-Target Tracking with Radar Application. Norwood, MA:* Artech House, 1986, vol. 1.

[29] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems. Norwood, MA:* Artech House, 1999.


ITS