High Level Sensor Data Fusion for Automotive Applications using Occupancy Grids

Ruben Garcia, Olivier Aycard, Trung-Dung Vu
LIG and INRIA Rhônes Alpes
Grenoble, France
Email: firstname.lastname@inrialpes.fr

Malte Ahrholdt
Volvo Technology Corporation
Goteborg, Sweden
Email: malte.ahrholdt@volvo.com

***Abstract*—We describe a general architecture of vehicle perception system developed in the framework of the European project PReVENT-ProFusion [1]. Our system consists of two main parts: the first part where the vehicle environment is mapped and moving objects are detected; and the second part where previously detected moving objects are verified and tracked. In this paper, we focus on the first part, using occupancy grid to model the vehicle environment, perform sensor data fusion and detect moving objects. Experimental results on a Volvo Truck equipped with laser scanner and radars show the effectiveness of our approach.**

***Index Terms*—vehicle perception, sensor data fusion, sensor model, laser scanner, radar, occupancy grid.**

I. INTRODUCTION

Perceiving or understanding the environment surrounding a vehicle is a very important step in driving assistance systems or autonomous vehicles [18][17]. In the previous work [2], we designed and developed a generic architecture of vehicle perception. The architecture (Fig. 1) is divided into two main parts: the first part where the vehicle environment is mapped and moving objects are detected; and the second part where previously detected moving objects are verified and tracked.

In the first part of the architecture, to model the environment surrounding the vehicle, we use the Occupancy Grid framework developed by Elfes [7]. Compared with feature-based approaches [10], grid maps can represent any environment and are specially suitable for noisy sensors in outdoor environments where features are hard to define and extract. Grid-based approaches also provide an interesting mechanism to integrate different kinds of sensors in the same framework taking the inherent uncertainty of each sensor reading into account. On the contrary of a feature based environment model, the only requirement for an OG building is a bayesian sensor model for each cell of the grid and each sensor. This sensor model is the description of the probabilistic relation that links sensor measurement to space state, that OG necessitates to make the sensor integration. Fortunately it is possible for a wide class of sensors to factorise this amount of data by taking advantage of the characteristics of the sensor. Regarding telemetric sensors, sensor model for sonar [19] and laser range finders [15] have been defined and used to map the environment. 3D occupancy grids have been built using stereo vision[12] and a set of camera [8].

[1] www.prevent-ip.org/profusion

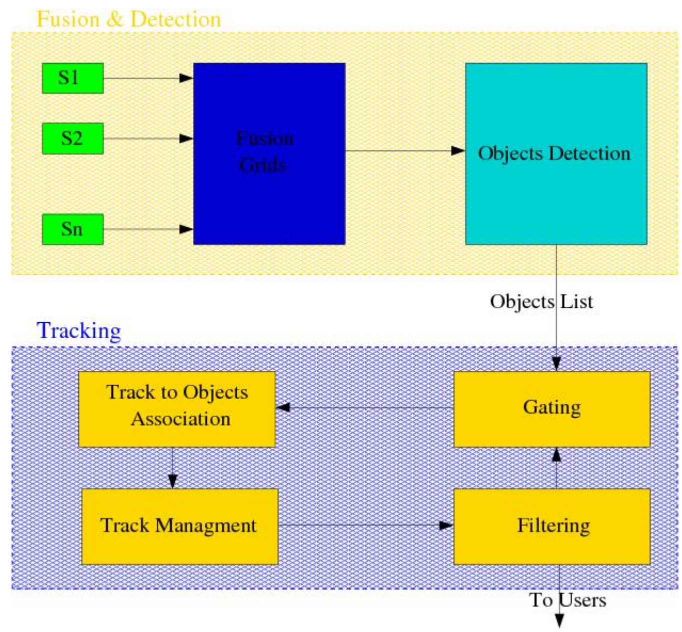

Fig. 1. Architecture of the perception system

In the second part, detected moving objects in the vehicle environment are tracked. Since some objects may be occluded or some are false alarms, multi object tracking helps to identify occluded objects, recognize false alarms and reduce miss-detections. In general, the multi objects tracking problem is complex: it includes the definition of filtering methods, but also association methods and maintenance of the list of objects currently present in the environment [3][16]. Regarding tracking techniques, Kalman filters [9] or particle filters [1] are generally used. These filters require the definition of a specific dynamic model of tracked objects. However, defining a suitable motion model is a real difficulty. To deal with this problem, Interacting Multiple Models [11][14] have been successfully applied in several applications. In the previous work [6], we have developed a fast method to adapt online IMM according to trajectories of detected objects and so we obtain a suitable and robust tracker. To deal with the association and maintenance problem, we extend our approach to multiple objects tracking using the Multiple Hypothesis Tracker [4][5].

This architecture has been used in past projects [2], and is currently used in the framework of the European project PReVENT-ProFusion. The goal of this project is to design and develop a general architecture to perform perception tasks (ie, mapping of the environment, localization of the vehicle in the map, and detection and tracking of moving objects). In this context, our architecture has been integrated and tested on two demonstrators: a Daimler-Mercedes demonstrator and a Volvo Truck demonstrator. The main difference between these 2 demonstrators is the level of abstraction of data provided by the different sensors on each demonstrator: raw data for the Daimler-Mercedes demonstrator (ie, low level of abstraction) and preprocessed data for the Volvo Truck demonstrator (ie, high level of abstraction). To deal with the difference, we design and implement specific procedure of the first part of the architecture: specific sensor models and also specific techniques to detect moving objects using the occupancy grid. Moving object tracking in the second part of the architecture remains the same for both demonstrators.

In [17], a detail description of the first part for the Daimler-Mercedes demonstrator is reported: building specific sensor models and designing specific techniques for detecting moving objects. Moreover, results and comparison with other perception system for Pre-Crash applications is described in [13]. In this paper, we describe the specificities of our architecture for the Volvo Truck demonstrator: designing and implementing sensor models for preprocessed data and specific techniques for detection of moving objects.

The rest of the paper is organized as follows. In the next section, we present the Volvo Truck demonstrator. A brief overview of Environment Mapping with Occupancy Grid (including Sensor Data Fusion) is given in section III. Specific sensor models for the Volvo Truck are described in Section IV. Detection of moving objects using occupancy grid previously built is detailed in section V. Experimental results are given in Section VI and finally in Section VII conclusions and future works are discussed.

II. THE VOLVO TRUCK DEMONSTRATOR

The test vehicle platform is based on a Volvo FH12 420 Globetrotter truck. The main components of the perception system are:

- An IBEO laser scanner, mounted in the front left corner of the truck. This sensor has a field of view of 210 ° and a range of 80 meters;
- A lane camera and vision system;
- A long range radar (LRR) system. This sensor has a field of view of 12 ° and a range of 200 meters;
- A short-range radar (SRR) system;
- An automotive PC hosting the data fusion platform.

In this paper, we only use the laser scanner and the LRR as inputs of the perception system. Moreover, data of each sensor are processed, and each sensor delivers a list of moving objects present in the environment. The perception system provides the decision system with the objects detected ahead of the own vehicle. With help of the lane camera an improved object-to-lane assignment is investigated. Based on the fusion system a decision is made if the collision is likely or unavoidable. If a collision is detected to be likely, the driver is warned, if an impact has become unavoidable, the vehicle is braked automatically to mitigate the consequences of the collision.

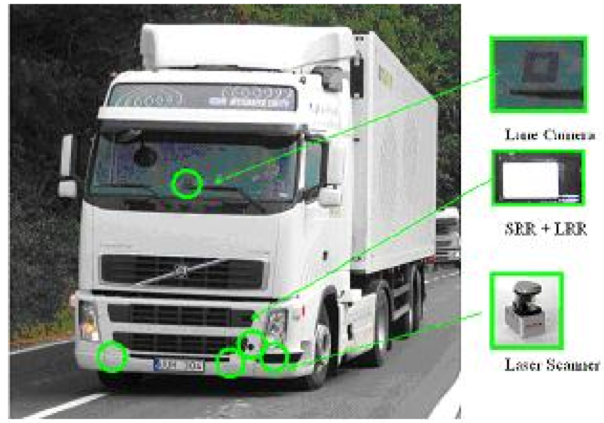

Fig. 2. The Volvo Truck demonstrator.

III. OCCUPANCY GRID

Since the two kind of sensors (laser scanner and LRR) provide pre-filtered data at object level as lists of target points, so to perform mapping and sensor data fusion, we have to design and implement a sensor model for these two kind of sensors.

Before describing our approach in detail, we introduce some notations used in the paper. We denote the observation from vehicle at time t by the variable $z = \{z^1, ..., z^K\}$ including K individual measurements corresponding to K observations from one or several sensors.

In this representation, the vehicle environment is divided into a two-dimensional lattice M of rectangular cells and each cell is associated with a measure taking a real value in $[0, 1]$ indicating the probability that the cell is occupied by an obstacle. A high value of occupancy grid indicates the cell is occupied and a low value means the cell is free. Assuming that occupancy states of individual grid cells are independent, the objective of a mapping algorithm is to estimate the posterior probability of occupancy $P(m \mid z^{1:K})$ for each cell m of the grid, given observations $z^{1:K} = \{z^1, ..., z^K\}$.

Using Bayes theorem, this probability is determined by:

$$P(m \mid z^{1:K}) = \frac{P(z^{1:K} \mid m) \cdot P(m)}{P(z^{1:K})} \tag{1}$$

If we assume that each measurement z^i is independent from the other given we know m, $P(z^{1:K} \mid m) = \prod_{k=1}^{K} P(z^k \mid m)$. Then equation (1) becomes:

$$P(m \mid z^{1:K}) = \frac{P(m)}{P(z^{1:K})} \cdot \prod_{k=1}^{K} P(z^k \mid m) \tag{2}$$

Applying again Bayes theorem to $P(z^k \mid m)$, we obtain:

$$P(m|z^{1:K}) = \frac{P(m)}{P(z^{1:K})} \cdot \prod_{k=1}^{K} \frac{P(m|z^k) \cdot P(z^k)}{P(m)} \quad (3)$$

Equation (3) gives the probability for an occupied cell. By analogy, equation (4) gives the probability for a free cell:

$$P(\overline{m}|z^{1:K}) = \frac{P(\overline{m})}{P(z^{1:K})} \cdot \prod_{k=1}^{K} \frac{P(\overline{m}|z^k) \cdot P(z^k)}{P(\overline{m})} \quad (4)$$

By dividing equation (3) by (4), we obtain:

$$\frac{P(m|z^{1:K})}{P(\overline{m}|z^{1:K})} = \frac{P(m)}{P(\overline{m})} \cdot \frac{P(\overline{m})^K}{P(m)^K} \cdot \prod_{k=1}^{K} \frac{P(m|z^k)}{P(\overline{m}|z^k)} \quad (5)$$

If we define $Odds(x) = \frac{P(x)}{P(\overline{x})} = \frac{P(x)}{1-P(x)}$, equation (5) turns into:

$$Odds(m|z^{1:K}) = Odds(m) \cdot Odds(m)^{-K} \cdot \prod_{k=1}^{K} Odds(m|z^k) \quad (6)$$

The corresponding *log Odds* representation of equation (6) is:

$$\begin{aligned} &log\,Odds(m|z^{1:K}) \\ &= log\,Odds(m) - K \cdot log\,Odds(m) + \sum_{k=1}^{K} log\,Odds(m|z^k) \quad (7) \end{aligned}$$

In (7), what we need to know are two probability densities, $P(m|z^k)$ and $P(m)$. $P(m)$ is the prior occupancy probability of the map cell which is set to 0.5 representing an unknown state, that makes this component disappear. So, the final equation is:

$$log\,Odds(m|z^{1:K}) = \sum_{k=1}^{K} log\,Odds(m|z^k) \quad (8)$$

From the *log Odds* representation, the desired probability of occupancy $P(m|z^{1:K})$ can be easily recovered.

The remaining probability $P(m|z^k)$, is called the *inverse sensor model*. It specifies the probability that a grid cell m is occupied based on a sensor measurement z^k. So, for each sensor, we have to define this model.

Regarding (8), to perform fusion between observations given by different sensors:

1) We firstly design an inverse sensor model for each sensor;
2) We use this model to build an occupancy grid for each sensor;
3) Finally, since probabilities in occupancy grids are represented with *log Odds* representation and since cells in a given occupancy grid are mutually independent, fusion of different occupancy grids is made by adding *log Odds* representation for a given cell in each occupancy grid (8). The resulting grid of this sum represents the sensors' fusion.

IV. BUILDING THE TWO INVERSE SENSOR MODELS

The inputs of the environment modeling procedure are outputs of laser scanner and LRR. Both sensors provide information in polar coordinates, with the following characteristics: the laser scanner sends a list of impacts points for each detected moving object while the LRR provides for each moving object its position in polar coordinates. Assuming that occupancy states of individual grid cells are independent, the objective is to estimate the posterior probability of occupancy $P(m|z^k)$ for each cell of the grid, given the observation z^k. To compute this probability, we need to design an inverse sensor model for each sensor. In the two next subsections, we describe how we design these models.

A. Laser scanner inverse sensor model

Fig. 3. Laser Scanner Data

The laser scanner provides a list of detected moving objects, this list contains for each object a list of hit points; with these hit points, we build a bounding box representing occupied space by each object. Figure 3 shows two moving objects by means of the bounding boxes, and also the field of view of the laser scanner.

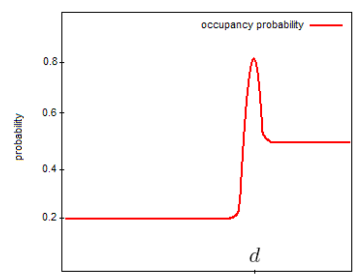

Fig. 4. Profile of an inverse sensor model illustrates the occupancy probability along a laser beam measuring a distance of d.

For each hit of the laser z^i, we firstly consider a 1D model (figure 4) to estimate the posterior probability of occupancy $P(m|z^i)$ for each cell m. Figure 4 shows the profile of this 1D inverse sensor model. This model corresponds to our knowledge about a point of impact at distance d along a laser beam:

1) before the point of impact there is high probability that there is any object. So cells before the point of impact have low probability to be occupied;

2) close to the point of impact there is high probability that there is an object. So cells located close to the point of impact have high probability to be occupied;
3) after the point of impact, we have any information about the presence of an object. So cells located after the point of impact have the same probability to be occupied than to be empty.

To build this 1D model of a laser beam, we have to find the 3 numerical values corresponding to our knowledge of a laser beam and a point of impact at distance d. This can be done by several ways like: by means of learning from examples or from characteristics provided by the manufacturer or by means of values given by hand from the experience with the use of that sensor. On the basis of collected data, we saw that the laser scanner is precise in relation to the position of the moving object reason for which, the following probabilities for the inverse model are considered: $\{0.2; 0.8; 0.5\}$ where:

- 0.2 represents the probability that a cell before the point of impact is occupied, we can say that the cell has a high probability of being empty,
- 0.8 represents the probability that cells to the impact are occupied;
- 0.5 represents our ignorance on the present state of the cell, that is the case for the cells behind the point of impact,

To go from 1D model to 2D occupancy grid, the cells of the 2D occupancy grid are initialized with a value of 0.5 (no information about the presence of an object) while the cells within the field of view of the laser are initialized with 0.2. Since the laser sensor provides pre-treated data where, by each moving object it provides a list of impacts from which a bounding box is constructed, to bind those data to the occupancy grid, all the cells corresponding to the bounding box are considered cells of impact, meaning that those cells have occupied probability of 0.8, while the cells behind the bounding box have a 0.5 probability which represents our ignorance on its state of occupation.

B. LRR inverse sensor model

The LRR sensor provides a list of moving objects, where for each object, it provides its position in polar coordinates as well as its Doppler speed; the figure 5 shows an example of this data where there are three moving objects and the field of view of the radar.

In the case of the LRR, we consider a 2D model where data consist of an angle θ and a distance r for each hit point in the sensor's field of view. Moreover, we consider that angle θ and distance r are mutually independent according to the knowledge of the cell m. The equation (1) to obtain the probability of cell m to be occupied becomes:

$$P(m|\theta^k, r^k) = \frac{P(m) \cdot P(\theta^k|m) \cdot P(r^k|m)}{P(m)P(\theta^k|m)P(r^k|m) + P(\overline{m})P(\theta^k|\overline{m})P(r^k|\overline{m})} \quad (9)$$

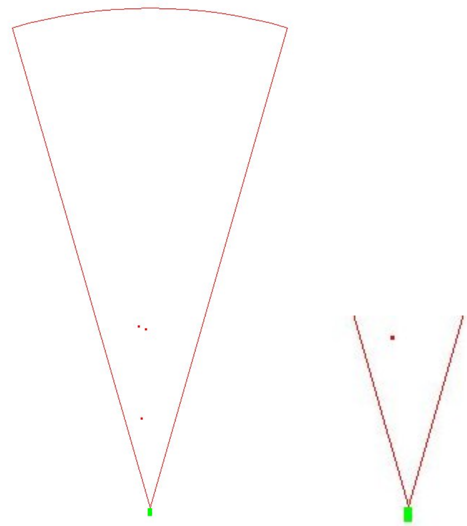

Fig. 5. Long Range Radar Data

The four conditional distributions $P(\theta^k|m)$, $P(r^k|m)$, $P(\theta^k|\overline{m})$ and $P(r^k|\overline{m})$ must be specified and they represent the definition of the LRR sensor model which can be defined of similar form to the laser scanner sensor model, nevertheless the LRR has less precision than the laser scanner reason for which we considered it like a gaussian sensor in 2D. Additionally, we considered that θ and r are mutually independent on the basis of that cell m is known. With these considerations; the LRR sensor model can be represented by:

$$\begin{aligned} P(\theta^k|m) &= \frac{1}{\sqrt{2\pi}\sigma_\theta} e^{-\frac{1}{2}(\frac{\theta_t - \mu_\theta}{\sigma_\theta})^2} \\ P(r^k|m) &= \frac{1}{\sqrt{2\pi}\sigma_r} e^{-\frac{1}{2}(\frac{r_t - \mu_r}{\sigma_r})^2} \end{aligned} \quad (10)$$

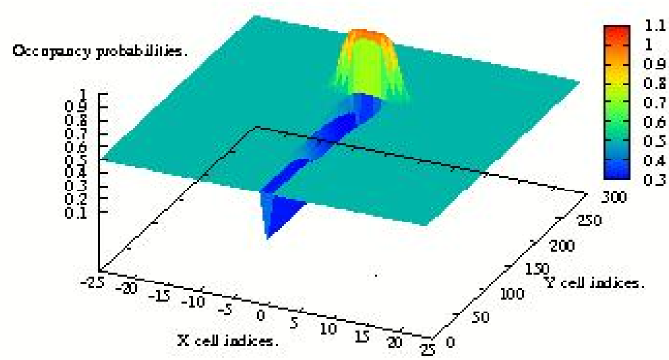

Fig. 6. Profile of an inverse sensor model 2D.

The two-dimensional occupancy grid profile shown in the figure 6 corresponds to a single long range radar measurement with the radar sensor being modelled as having gaussian uncertainty in both range and angle.

To build the 2D occupancy grid of the LRR:

1) cells are initialized with a 0.5 probability, which indicates the ignorance on their state of occupation;
2) in order to calculate the occupied probability of the cells around the impact, the inverse model of the sensor (Eq (9), (10)), is used with $\sigma_\theta = 0.013$ and $\sigma_r = 1.5$. We take into accounts the cells around the point of impact. If they are within $\theta \pm 2\sigma_\theta$ and $r \pm 2\sigma_r$.
3) the cells behind the zone of impact have occupied probability of 0.5 because their occupancy state is unknown.

V. DETECTION OF MOVING OBJECTS

Once an occupancy grid is obtained, we want to extract the moving objects which are likely located in regions with high occupation probability. Object-regions may have arbitrary shapes and are generally discriminant from background. From these characteristics, we apply a threshold segmentation method.

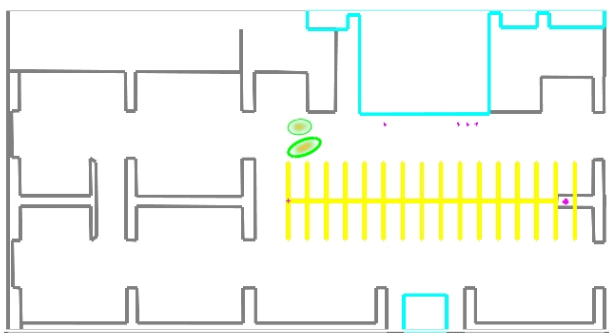

Fig. 7. detection of objects approximated with ellipses

First, an adaptive threshold is computed based on a discrete histogram of cell occupation probability values and the threshold is chosen as the mean value of the histogram. We use this threshold to transform the grid into a binary image where positive pixels represent occupied areas. In the next step a two pass segmentation algorithm is applied to extract all 4-connected groups of cells. Each connected group corresponding to a possible object is finally approximated by an ellipse represented by mean value and covariance matrix of the corresponding region (see Figure 7).

VI. EXPERIMENTAL RESULTS

The algorithm proposed for the sensors' fusion is tested on database collected with the Volvo truck demonstrator, the vehicle was driven through different kinds of scenarios such country roads and highways. In this implementation, the width and height of occupancy grid are set to 80 m and 270 m respectively, and the cell size is set to 50 cm. White color corresponds to low probability of occupancy and black color corresponds to high probability of occupancy. Intermediate level of gray color corresponds to intermediate probability of occupancy. The figure 8 shows some results of the fusion process. Outside of the field of view, the cells have a probability of occupation of 0.5 corresponding to apriori of the occupancy state of the cell, within the field of view, if an object exists, the cells contain a high probability of occupation, then the color is more dark to these cells, the cells behind a occupied cell have a 0.5 probability indicating the incapacity of the sensor to detect objects in that zone because they are hidden. In the empty zone of the field of view, the cells have a low probability of occupation, indicating this with a gray clearly color. The figure 8 shows four images; from left to right we show:

1) data provided by the two sensors and the field of view of each one of them. The red points represent the LRR data while the blue rectangles correspond to the bounding box created from the data provided by the laser scanner.
2) the second image represents the occupancy grid corresponding to the laser scanner, The two black rectangles represent detected objects, the cells behind these rectangles have a gray color indicating that its state of occupation is unknown. Outside the field of view, the cells are also gray and remaining cells within the vision field are almost white indicating their high probability of being empty;
3) the third image is similar to the second one, representing the occupancy grid for the LRR, the set of cells corresponding to an occupied area have different probabilities, this depends to the inverse sensor model, this fact is shown with diverse levels of gray for the set of cells that represent an object. In this image there are three objects the central part of the zone of occupation is black, since these cells have major occupation probability.
4) the last image corresponds to the fusion between the laser scanner and the LRR, the white zone corresponds to the fusion between the free zones of both fields of view of the sensors. The only object detected by both sensors has a high probability of occupancy and the area behind this object corresponds to occluded area. Other objects (only detected by one sensor) have lower probabilities of occupancy than the object detected by both sensors.

VII. CONCLUSION AND FUTURE WORKS

In this paper, we detail the first part of sensor fusion and moving object detection in our generic architecture for vehicle perception and its implementation on a Volvo Truck Demonstrator. We used occupancy grid framework for environment mapping and perform sensor data fusion. We described sensor models delivering preprocessed data and demonstrated our approach on real-life data. In the second part, the detected objects (outputs of the first part) are tracked using Interacting Multiple Models and Multiple Hypothesis Tracker.

ACKNOWLEDGMENT

The work is supported by the European project PReVENT-ProFusion.

REFERENCES

[1] S. Arulampalam, S Maskell, N Gordon, and T Clapp. A tutorial on particle filter for online nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2), 2002.

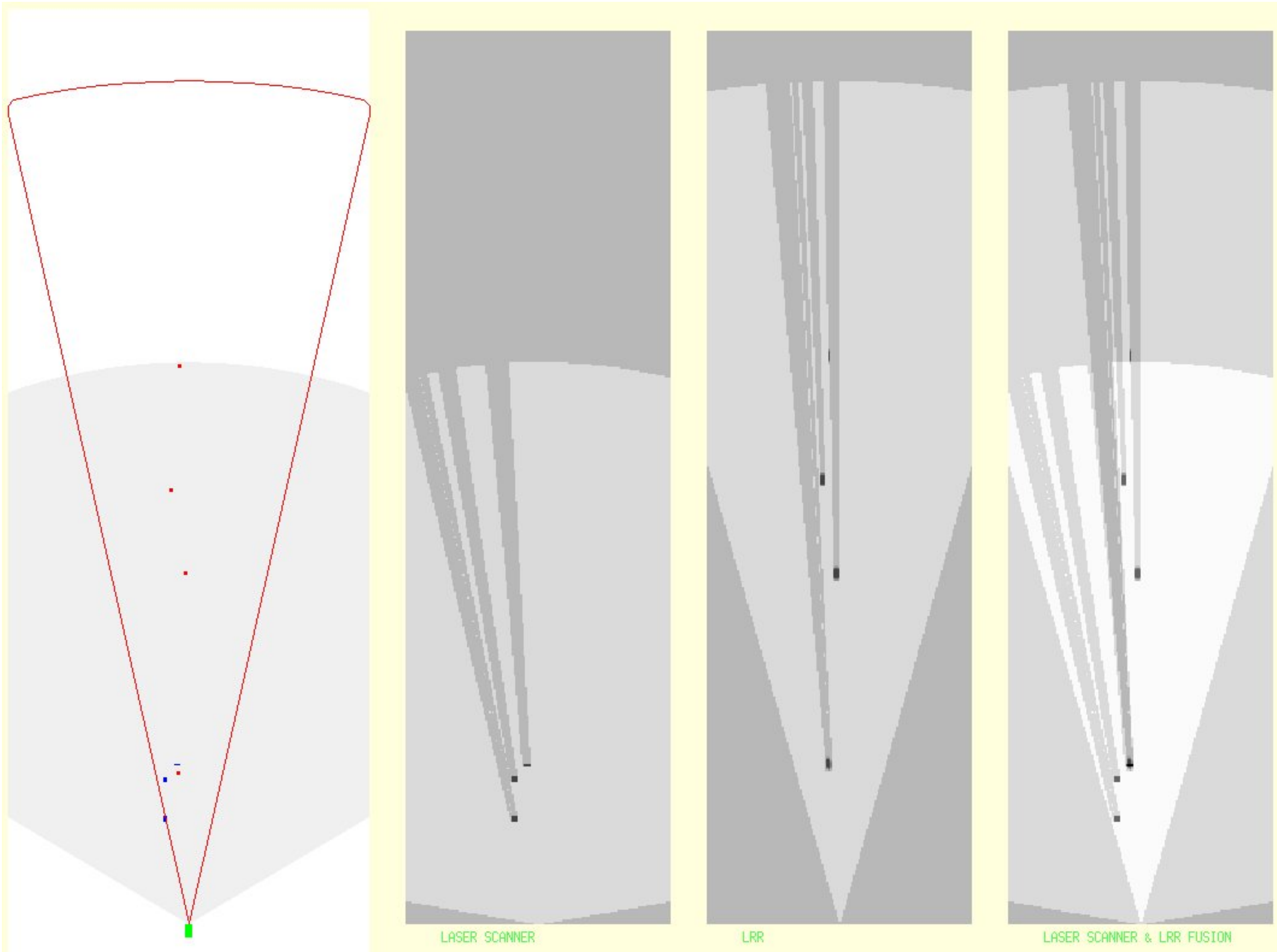

Fig. 8. Fusion of laser and radar data at object level. See text for more detail.

[2] Olivier Aycard, Anne Spalanzani, Manuel Yguel, Julien Burlet, Thierry Fraichard, Christian Laugier, and David Raulo. Puvame - new french approach for vulnerable road users safety. In *Proc. of the IEEE Intelligent Vehicle Symp.*, Tokyo (JP), June 2006.

[3] Y. Bar-Shalom and T.E. Fortman. *Tracking and Data Association*. Academic Press, 1988.

[4] Y. Bar-Shalom and X. Li. *Multitarget Multisensor Tracking : Principles and Techniques*. YBS Publishing, 1995.

[5] S. S. Blackman. Multiple hypothesis tracking for multiple target tracking. *Aerospace and Electronic Systems Magazine, IEEE*, 19(1):5–18, 2004.

[6] J. Burlet, O. Aycard, A. Spalanzani, and C. Laugier. Adaptive interactive multiple models applied on pedestrian tracking in car parks. In *Proc. of the IEEE-RSJ Int. Conf. on Intelligent Robots and Systems*, Beijing (CN), 2006.

[7] A. Elfes. *Occupancy grids: a probabilistic framework for robot perception and navigation*. PhD thesis, Carnegie Mellon University, 1989.

[8] J. S. Franco and E. Boyer. Fusion of multi-view silhouette cues using a space occupancy grid. In *IEEE International Conference on Computer Vision*, 2005.

[9] R.E. Kalman. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 35, Mars 1960.

[10] J. Leonard and H. Durrant-Whyte. Simultaneous map building and localization for an autonomous mobile robot. volume 3, 1991.

[11] E. Mazor, A. Averbuch, Y. Bar-Shalom, and J. Dayan. Interacting multiple model methods in target tracking: a survey. *Aerospace and Electronic Systems, IEEE Transactions on*, 34(1):103–123, 1998.

[12] H.P. Moravec and M.C. Martin. Robot navigation by 3d spatial evidence grids. Technical report, Mobile Robots Laboratory, Robotics Institute, Carnegie Mellon University, 1994.

[13] S. Pietzsch, O. Aycard, J. Burlet, TD. Vu, T. Hackbarth, N. Appenrodt, J. Dickmann, and B. Radig. Results of a precrash application based on laserscanner and short range radars. In *IEEE International Conference on Intelligent Vehicles*, 2008.

[14] X. Rong Li and Vesselin P.Jilkov. A survey of maneuvering target tracking-part v: Multiple-model methods. *IEEE Transactions on Aerospace and Electronic Systems*, 2003.

[15] S. Thrun. Learning metric-topological maps for indoor mobile robot navigation. *Artificial Intelligence*, 1(99):21–71, 1999.

[16] S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press, September 2005.

[17] Trung-Dung Vu, Olivier Aycard, and Nils Appenrodt. Online localization and mapping with moving object tracking. In *Proc. of the IEEE Intelligent Vehicle Symp.*, 2007.

[18] C.-C. Wang. *Simultaneous Localization, Mapping and Moving Object Tracking*. PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, April 2004.

[19] B. Yamauchi. A frontier based approach for autonomous exploration. In *IEEE International Symposium on Computational Intelligence in Robotics and Automation*, pages 146–151, 1997.