Radar-vision fusion for vehicle detection

Luca Bombini, Pietro Cerri, Paolo Medici
Dipartimento di Ingegneria dell'Informazione
Università di Parma
Parma, I-43100, Italy
{bombini,cerri,medici}@ce.unipr.it

Giancarlo Alessandretti
Innovative Technologies
Centro Ricerche FIAT
Orbassano (TO), I-10043, Italy
giancarlo.alessandretti@crf.it

***Abstract*— This paper describes a vehicle detection system fusing radar and vision data.**

Radar data are used to locate areas of interest on images. Vehicle search in these areas is based on vertical symmetry. All vehicles found in different image areas are mixed together and a series of filters are applied in order to delete false positives. The algorithm analyzes images on a frame by frame basis, without any temporal correlation. Two different statistics, frame-based and event-based, are computed to evaluate the method efficiency. Results and problems are discussed, and some ideas for possible enhancements are provided.

I. Introduction

DIFFERENT preventive safety functions are now introduced on road vehicles to assist the driver and to reduce the risk of accidents. Key points for improved operation are the effectiveness and the information content in the perception of the surrounding scenario, including road features and obstacles. Radar-vision fusion is an interesting approach, based on often complementary devices, which can provide several advantages: in particular improved reliability from multiple detections and the merging of position measures with good longitudinal and lateral accuracies. Therefore a purpose of the present study has been to investigate how different parameters can be obtained according to the best performance of each sensor.

The advantages and the problems of fusing radar and camera data for vehicle detection are well known [1]; methods differ mainly for the fusion level: low level fusion, intermediate level fusion and high level fusion have all proved to reach good results. Low level fusion combines several sources of raw data to produce new raw data that is expected to be more informative and synthetic than the inputs [2]; in intermediate level fusion various features such as edges, corners, lines, texture parameters, etc are combined into a feature map that is then used by further processing stages; while in high level fusion each source of input yields a decision and the decisions are fused [3]. This work is developed using high level fusion and focuses on validation of radar targets, as shown by Sole [4].

Some methods use models to identify a vehicle; many different models have been used, ranging from trivial models to complex ones, like deformable models that add details of the car approaching the camera [5], or 3D models that take into consideration vehicle misalignment with the camera. [6]. All these methods need models that match different vehicle types.

The search for vehicle features provides a simplified way of localizing vehicles. For example symmetry is a characteristic that is common to most vehicles. Some groups have already used symmetry to localize vehicle [7] [8], they have tried out a lot of methods to find symmetry on images: using edges, pixel intensity, and other features.

The vehicle detection algorithm used in this work is based on symmetry [9] and uses radar data in order to localize areas of interest. Data fusion operates at high level; the vision system is used to validate radar data and to increase their accuracy.

In the current setup a progressive gray-scale camera is mounted inside the cabin close to the rear-view mirror; the camera aperture is 45 °, the image resolution is 640×480 but only 640×300 pixels are used by the vision system. A scanned radar with a 77 GHz frequency is mounted above the front bumper: obstacles up to 50 meters can be detected.

The next section presents the algorithm details, section III shows some results obtained and presents two different statistical method used in order to evaluate the system efficiency, and finally in section IV conclusions and future developments are presented.

II. Algorithm

The first step of the algorithm converts radar objects into the image reference system, using a perspective mapping transformation: the radar point is shown on the object base. This transformation is performed using calibration data achieved by fine intrinsic and extrinsic camera parameters measurement. Because the parameters measurement is performed once, at system setup, and no stabilization is applied, errors may occur when extrinsic parameters change (mainly vehicle pitch) due to road roughness or vehicle acceleration (see fig.1). Moreover radar may provide an incorrect lateral position: points my not be centered onto the obstacle shape or even fall outside it, as shown in figure 2. The definition of search area building needs to take care of these possible errors.

Fig. 1. Camera miscalibration caused by pitch variation: the radar point is not correctly represented on the vehicle base

Fig. 2. Lateral radar error

Wide margins are used both on the left and right sides of the search area: values between 2.5 and 4 meters have been tested. The area height is defined to be half of its width; the bottom of the area is positioned at a fixed percentage of height below the radar points (30% of area height), in a way that the vehicle should be included even in case of strong pitch variation. Only radar data that refer to points inside the image are considered; since the sensors have been chosen to have approximately the same horizontal angular field of view, almost all radar points can be remapped into the image. Nonetheless a radar point can be remapped to an image point very near to left or right margin: in such a case a part of the search area can lay outside the image. In order to solve this problem two different approaches have been tested. The first one is based on moving the search area horizontally until the whole area is inside the image: this solution is not very efficient because this new search area contains a part of the image very distant from the radar point; this area is generally useless for the detection and can also cause errors. The second approach is based on the search area cropping: in this case only the useful part (the visible) of the image is treated.

In order to simplify and speed up the following steps of the algorithm and to delete details of too close vehicles, all the areas are resampled to a fixed size. The reduced search areas do not have the same aspect ratio of the original ones, resampling them to a fixed size causes image deformations and altered information; in order to avoid this problem, these areas are reduced preserving their aspect ratio, in such a way these search areas are smaller than the normal ones.

The binarized vertical edges image of the area is used to compute the symmetry. The symmetry is computed considering every column of the image as a possible symmetry axis, on different sized bounding boxes whose height matches the image height and with a variable width ranging from 1 to a predetermined maximum value. Symmetry χ is computed as the ratio between the square of symmetric edges (s) versus all edges present in the considered area(n).

$$\chi = \frac{s^2}{n}$$

Two vertical edges are considered symmetrical if their orientation is opposite.

χ values are stored into a matrix whose columns match to the symmetry axis and whose rows refer to the bounding box width. A possible vehicle is localized if χ is low for narrow area widths and becomes high over a certain value. The width of the symmetry search area over which χ becomes higher than a fixed threshold represents the actual vehicle width (fig. 3).

To be sure that a vehicle is present, its bottom is searched, looking for horizontal edges. The bottom search is based on the idea that a dark area (the shadow) is often present just beneath the vehicle. The vehicle top is searched too, but the the box is validated even if the top is not found because sometimes it can be very hard, or actually impossible, to find it: truck tops are often outside the search area or even outside the image. If the top can not be found a box with a fixed ratio between width and height is used (see fig. 7.b). This search is not very refined: sometimes image scaling problems can cause a false top (see fig. 7.d).

The algorithm is designed to return only one vehicle for each radar point, but more than one possi-

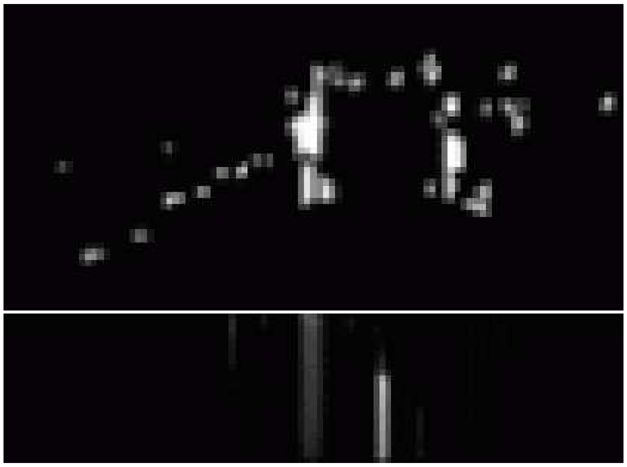

Fig. 3. Vertical edges and symmetry matrix. In the symmetry matrix, two columns are very bright: the first one refers to the vehicle left border, the second one refers to the vehicle center. The first column is entirely bright because symmetry is due to edges close to each others; the second column is only partially bright, the upper rows are dark because they correspond to narrow symmetry areas, where no symmetric edges are present. Lower rows, that correspond to large symmetry areas, are bright.

ble vehicles can be detected in a single area: a filter that determines which vehicle has to be validated is mandatory. This filter simply choses the most central box (see fig.4).

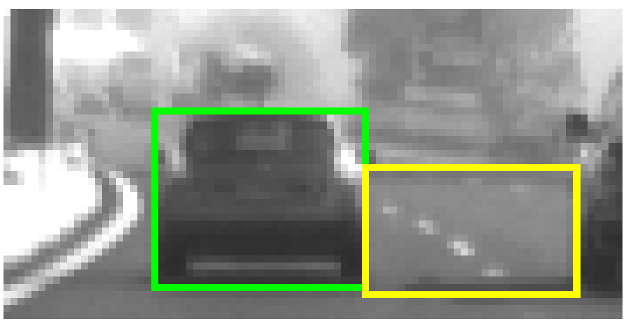

Fig. 4. The most central box (the green one) is chosen.

When all radar data have been examined, all the vehicles found are resampled to the original size and mixed together; overlapping vehicles are also managed.

Using an inverse perspective mapping transformation, real width and position of vehicles can be computed. In the computation of these values, radar provides distance while vision provides position and width so that the radar precision on distance measurement and the vision refinement ability are capitalised together.

Unfortunately not all detected boxes are correct: some false positives caused by road signs or other objects in the scene can be present as well. A filter is used to discard some false positives: it removes too large or too small boxes that probably do not identify a vehicle.

It is also possible that a car is detected in more than one search area (it happens when a single object returns multiple radar points), so overlapping results may be present. Only one box per vehicle is expected, so a further step is required to merge similar boxes and eliminate redundant ones: the largest box is preferred, since often the smallest box is generated by vehicle parts. Furthermore when two rectangles with similar size have their base at the same height an average is computed, and this new box is considered.

Figure 5 shows intermediate results: the red cross represents the radar point, the blue rectangle represents the interest area, the green rectangle represents the vehicle that was detected and cyan crosses represent border points provided by radar.

Fig. 5. Intermediate results

III. Results

The system was tested in extraurban and highway environments with good results. To evaluate the system performance, ground truth was manually collected by annotating the presence of each vehicle in more than 12000 images. The annotated sequences represent a mix of all possible scenarios: high and low traffic, extraurban and highway, fast, slow, and stopped vehicles, sunshine, cloudy and rain.

The percentage of refined detection is near 50%. This number must not be considered as low, because it represents the percentage of vehicles, whose position and size are detected with an extremely high precision (with an error lower than 40 centimeters) independently in every frame. This number also reflects the sensors problems, such as radar misses and vision misses due to bad environmental conditions (such as rain or darkness). Five different performance indexes were defined for these statistics: refined detections (RD), false negatives (FN), radar false negatives (RFN), false positives (FP), not-vehicle obstacle (NVO).

SEQUENCE	EVENTS	RADAR MISS	NOT REFINED	REFINED BY VISION
highway	5	0	1	4
highway with traffic	9	3	1	5
highway with shadow problems	10	0	3	7
trucks on highway	2	0	0	2
highway junction	14	5	4	5
extraurban with curves	1	0	0	1
extraurban with approaching vehicles	8	1	2	5
extraurban with strong shadows	10	0	2	8
complex environment	10	0	3	7
total	**69**	**9 (13%)**	**16 (23%)**	**44 (64%)**

TABLE I

RESULTS ON DIFFERENT SEQUENCES.

- RD vehicles detected by radar and validated and refined by vision;
- FN vehicles detected by radar but not validated or not correctly refined by vision;
- RFN vehicles not detected by radar or detected with a too low precision;
- FP not-vehicle obstacle detected by radar and validated by vision;
- NVO not-vehicle obstacle detected by radar and not validated by vision;

Altought the definition of refined detections is straight forward, the other values need an explanation. Radar false negatives are defined as vehicles not contained, or not entirely contained, in any search area: this value obviously depends on search area size. The chosen interest area width is 2.5 meters; and the RFN is 39% of framed vehicles; this value can be decreased by raising the area width, but this change will increase false positives. More than the half of radar false positives refer to vehicles partially contained in a search areas. The same consideration can be made for false negatives: the false negatives density is about 13% but only 5% is due to actually missed vehicles, remaining 8% is due to vehicles detected with a low precision. The number of false positives is low: in all the test sequences only one persistent false positive is present (see fig.6).

Radar supplies not only vehicles, but also others obstacles; about 15 not-vehicle obstacles are detected every 100 frames. Vision is not able to differentiate 4 obstacles, out of these 15, every 100 frames.

Event-based statistics were computed as well, considering an event as a vehicle present in more than 10 frames. Radar completely misses 13% of the events (mainly due to traffic), while vision is not able to refine 23% of the events at all. According to this data 64% of events are correctly detected and refined, 73% of object supplied to vision are correctly refined.

Fig. 6. Persistent false positives in complex scenario

Figures 7 and 8 show respectively good results and errors obtained in different scenarios. Figures 8.*a* and 8.*b* show traffic cases; in the first image a single radar point is generated by vehicles close to each other: its position is not suitable to detect any vehicle. The second image shows a delivery van individuated by two radar points and some other vehicles not detected.

In table I event based statistics obtained on different sequences are proposed. Radar misses are present only in traffic or complex scenarios. The main issue are traffic and environment complexity together with shadow or general illumination problems.

As already mentioned, not-precise detections , due to traffic or low visibility conditions, are unfortunately frequent; these cases in our performance analysis are classified as false negatives, since the final goal of this system is to estimate the vehicle shape with a high accuracy.

Fig. 7. Examples of correct results: the algorithm works reliably in simple cases (*a*); it detects both vehicles moving away and approaching (*b*); it works even in hard cases, such as rain (*c*) and noisy scenario (note the double radar detection) (*d*); it can detect cars (*e*) and truck (*f*).

IV. Conclusions

In this paper a method to fuse radar data and vision has been described. This method reaches good results both in extraurban and highway environments. Even if not all vehicles are detected in all images, the system is promising to be used for safety application, because all the closest and most dangerous vehicles are anyway detected.

A tracking algorithm might be very helpful to increase the robustness of the system and the detection persistence.

The use of other methods to generate areas of interest and the search for more than one vehicle in an interest area may solve some radar sensor problems, such as its inability to detect all vehicles and to distinguish vehicles close to each other.

Acknowledgments

The work described in this paper has been developed in the framework of the Integrated Project APALACI - PReVENT, a European research activity funded by the European Commission to contribute to road safety by developing and demonstrating preventive safety applications and technologies.

The authors gratefully thank Alberto Broggi for his support and review of this project.

References

[1] J. Laneurit, C. Blanc, R. Chapuis and L. Trassoudaine, Multisensorial data fusion for global vehicle and obstacles absolute positioning, *IEEE Intelligent Vehicles Symposium*, pp. 138-143, June 2003, Columbus, OH.

[2] B. Steux, C. Laurgeau, L. Salesse and D. Wautier, Fade: a vehicle detection and tracking system featuring monocular color vision and radar data fusion, *IEEE Intelligent Vehicles Symposium Proceedings*, 2002

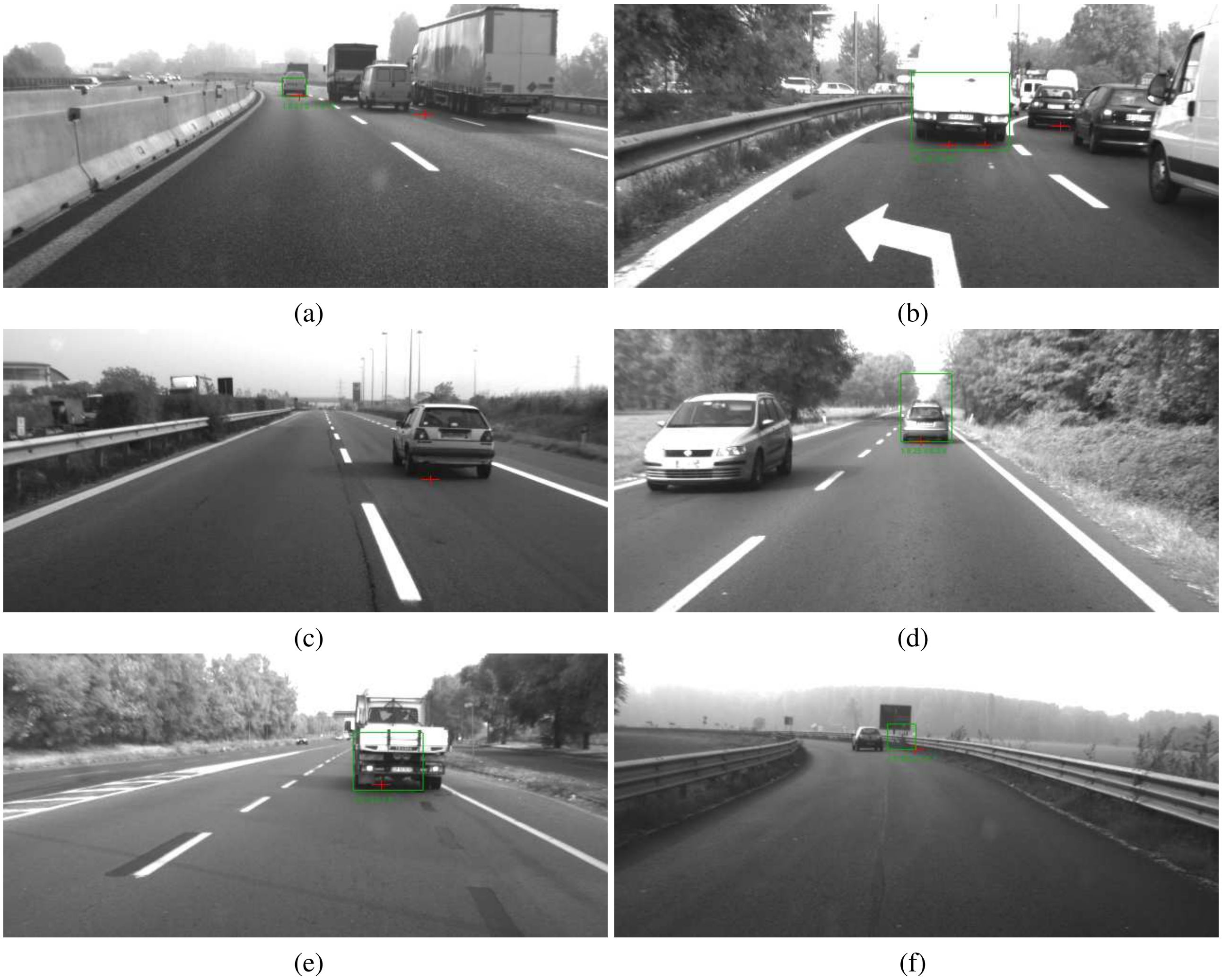

Fig. 8. Examples of errors: the most frequent errors are due to traffic (*a* and *b*), but some errors may also happen in simple cases: figure *c* shows a vision false negative due to low contrast on left border, figure *d* shows a radar false negative (the oncoming vehicle), figure *e* shows an inaccurate detection and figure *f* shows a false positive due both to vision and radar.

[3] U. Hofmann, A. Rieder and E. D. Dickmanns, Radar and vision data fusion for hybrid adaptive cruise control on highways, *International Conference on Computer Vision System*, pp.125-138, 2001, Vancouver, Canada

[4] A. Sole, O Mano, G.P. Stein, H. Kumon, J. Tamatsu and A. Shashua, Solid or not solid: vision for radar target validation, *IEEE Intelligent Vehicles Symposium Procedings*, pp. 819 824, June 2004, Parma, Italy

[5] S. Denasi and G. Quaglia, Obstacle Detection Using a Deformable Model of Vehicles, *IEEE Intelligent Vehicles Symposium Proceedings*, pp. 145-150, May 2001, Tokyo, Japan

[6] K. Fleischer, H. H. Nagel and T. M. Rath, 3D-Model-based-Vision for Innercity Driving Scenes, *IEEE Intelligent Vehicles Symposium Proceedings*, June 2002, Paris, France

[7] A. Kuehnle, Symmetry-based vehicle location for AHS, *SPIE - Transportation Sensors and Controls: Collision Avoidance, Traffic Management, and ITS Proceedings*, pp.19-27, November 1998, Orlando, FL

[8] T. Zielke, M. Brauckmann and W. von Seelen, Intensity and Edge-based Symmetry Detection with an Application to Car-Following, *CVGIP: Image Understanding*, pp. 177-190, vol.58, 1993

[9] A. Broggi, P. Cerri and P.C. Antonello, Multi resolution vehicle detection using artificial vision, *IEEE IntelligentVehicles Symposium Procedings*, pp. 310-314, June 2004, Parma, Italy.