

# Vehicle Detection and Localization in Infra-Red Images

L. Andreone<sup>†</sup>, P. C. Antonello<sup>†</sup>, M. Bertozzi<sup>§</sup>, A. Broggi<sup>§</sup>, A. Fascioli<sup>§</sup>, and D. Ranzato<sup>‡</sup>.

<sup>†</sup>Centro Ricerche FIAT, Italy

<sup>§</sup>Dip. di Ingegneria dell'Informazione, Università di Parma, Italy

<sup>‡</sup>Dip. di Informatica e Sistemistica, Università di Pavia, Italy

**Abstract**—This paper presents an algorithm for detecting vehicles in IR images. Initially the attention is focused on portions of the image that contains hot objects only. These areas are then selected and refined using aspect ratio and size constraints about vehicles; even situations with overlapping vehicles are considered. The result is further investigated exploiting specific vehicle thermal characteristics. A simple tracking phase is performed to improve the detection results.

Thanks to the knowledge of camera intrinsic parameters the distance of vehicles is computed using an assumption about vehicles width.

The system proved to be effective in different scenarios, but further tests are required to validate it in a wider range of weather conditions. It is able to detect vehicles in front of the vision system in the range 25 m÷100 m at a 12 Hz processing rate.

**Keywords**—Infra-Red Images, Vehicle Detection, Machine Vision.

## I. INTRODUCTION

Several research groups have been working on on-board Vehicle Detection by means of vision [1]. This task is particularly challenging to accomplish, due to the wide range of situations that must be taken in account, i.e. moving camera, cluttered background, partially occluded vehicles, different vehicles colors and textures. . .

Different approaches have been investigated such as the search for a specific shape [2,3], the use of a *a-priori* knowledge about vehicle symmetry [4,5,6], the detection of texture [7], the use of approximant contours, or the use of a vehicle model [8,9].

Most of research groups base their approaches on the processing of monocular images; others exploit stereo-vision [10,6,11] or the analysis of optical flow fields [2,12].

Only recently, thanks to the availability of Infra-Red (IR) devices, the analysis of IR images is becoming more and more common. The processing of IR images presents some advantages with respect to the processing of images in the visible domain. In fact, vehicles emit heat from tyres, engine, and muffler; this can be used for focusing the attention on few reduced portions of images that display hot objects. IR sensors are to a lesser extent dependent on different weather and illumination conditions than standard ones; even day or night snapshots of the same scene are very similar, thus reducing the range of situations to be taken into account. In addition, the use of IR images can be a solution to the problem of detecting vehicles that feature different colors or textures and to the problem of avoiding noisy patterns such as shadows. Unfortunately, IR images also present some disadvantages such as a more difficult processing during hot and sunny weather.

Anyway, the most interesting feature of IR devices is the fact that they represent an enhancement with respect to devices in the visible domain. In fact, while the latter simply mimic human

vision, IR cameras are able to extend vision beyond the usual limitations. On the other hand, current IR technology provides sensors with limited resolution.

While some projects simply investigated the feasibility and advantages of using IR sensors for driving assistance [13] or navigation [14], others developed systems for Vehicle Detection based on a fuzzy classification method [15] or eigenwindow approach [16,17]. The former detects the presence of vehicles but does not compute parameters needed for driving assistance such as vehicles distance or speed; the latter is aimed at the detection of a vehicle featuring specific patterns.

This paper presents an algorithm for detecting the preceding vehicles in IR images. The aim is the localization of vehicles moving on the road. No interest is given to parked vehicles or obstacles. Initially the attention is focused on portions of the image that contain hot objects. These areas are then processed to validate the presence of vehicles; detected vehicles are then tracked and their distance and speed are computed as well.

The system was installed on a prototype vehicle equipped with IR devices by the *Centro Ricerche FIAT* (CRF) and tested inside the CRF test track as well as on public roads (see figure 1). IR sensors with a small angular aperture were used, in order to place the vehicle detection in an area far away from the on-board vision system.



Fig. 1 The IR device installed on a prototype vehicle.

This paper is organized as follows: the next paragraph describes the algorithm, section 3 illustrates the results obtained so far, while section 4 ends the paper with some final remarks.

## II. THE ALGORITHM FOR VEHICLE DETECTION

IR images convey information about the temperature of objects. Conversely, objects do not exhibit features typical of the visible domain such as texture, color, or shadows. The advantage of a simplified information is anyway traded for a reduced

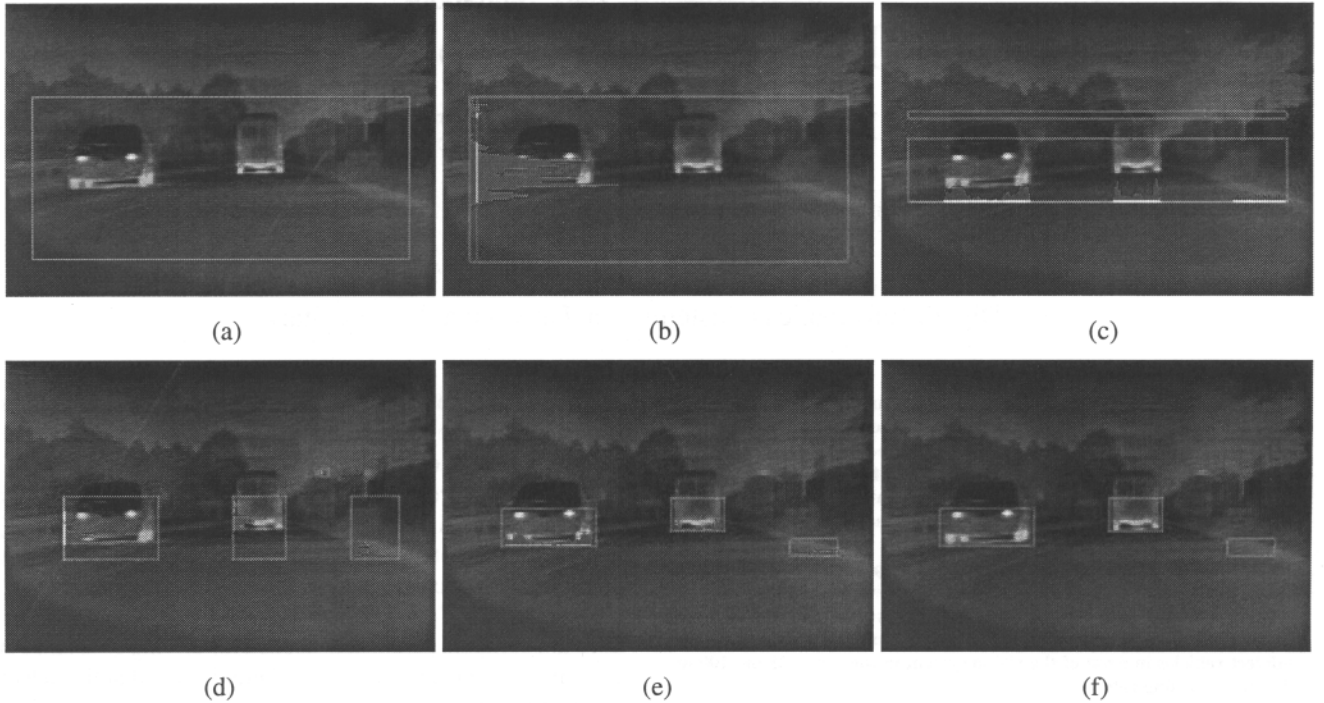


Fig. 2 Iterative selection of the areas of interest: (a) Original image and region of interest, (b) row-wise histogram, together with its threshold, (c) horizontal stripes and their vertical histograms, (d) rectangular bounding boxes and their row-wise histograms at the second iteration, (e) new rectangular bounding boxes and their column-wise histograms at the third iteration, (f) final selection.

capability to discriminate among different objects. Inevitably, recognition has to strongly rely on shape features. Nevertheless, the temperature of the vehicle is mainly concentrated on the wheels, engine and muffler, thus the complete shape of the chassis cannot often be distinguished from the background. Therefore, the detection has to be focused on the lower part of the target (wheels and muffler), which represents an invariant feature for preceding vehicles. The different temperature of vehicles, which mainly depends on travel time, has to be taken into account as well.

The algorithm is based on a first phase aimed at focusing the attention on hot spots in the image, followed by a selection of bounding boxes which are considered as possible candidates. Vehicles are then identified among candidate objects on the ground of specific features related to their shape. A tracking phase follows, allowing to confirm the hypotheses through the analysis of the temporal behavior of the object. The distance to the vehicle is also computed.

#### A. Attentive phase

The first phase of the processing is aimed at focusing the attention on the areas of the image with a high intensity value, representing hot objects.

A rectangular region of interest is considered which encloses the area where vehicles may appear and are completely visible, discarding the borders of the image (see figure 2.a). In this region a first low threshold is applied on the pixel values to get rid of cold areas, preserving only pixels with a grey-level value higher than this threshold.

A row-wise histogram is then computed on the grey-level values. In order to select horizontal stripes containing hot regions, the histogram is filtered with an adaptive threshold whose value is a fraction of the average value of the whole histogram (see figure 2.b). Obviously, multiple hot objects may happen to be horizontally aligned in the image, so that their contributions add up in the histogram. Nevertheless, hot spots belonging to the same horizontal stripe can be distinguished by computing a new column-wise histogram of the grey-levels for each stripe, as shown in figure 2.c. This procedure yields rectangular bounding boxes framing interesting areas. To refine these bounding boxes, the row-wise and column-wise histogram procedure is iteratively applied to each rectangular box until its size is no longer reduced (see figure 2.d, e and f).

#### B. Selection and refinement of candidates

The iterative histogram-based procedure focuses on rectangular boxes framing hot spots, regardless of their size, shape factor, and content distribution. Therefore, the interesting regions detected by the attentive phase are to be further examined to understand if they can represent good candidates for the presence of a vehicle.

Only the bounding boxes satisfying specific aspect ratio constraints are selected as candidates. Since the detection is based on the bottom part of the vehicle, the aspect ratio used in the filter discards tall and slim boxes, while keeping short and large ones. Also too small and too large bounding boxes are dropped. The presence of a minimum portion of sufficiently high intensity values is required as well, since it generally characterizes tyres

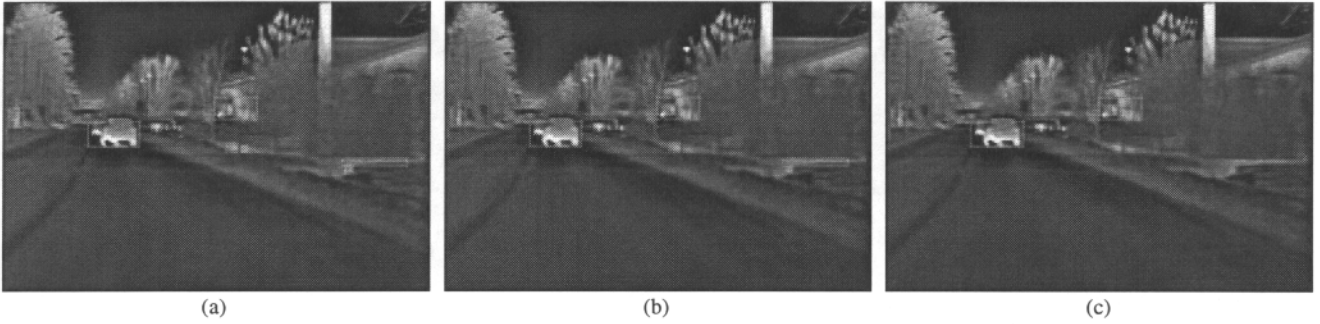


Fig. 3 Selection of candidates: (a) interesting areas yielded by the attentive phase, (b) removal of wrong bounding boxes based on size and aspect ratio, (c) bounding boxes with low intensity are dropped.

and muffler. The selection phase is shown in figure 3.

Besides getting rid of the boxes which do not satisfy such minimal conditions to represent vehicles, a refinement is performed to improve the framing of objects. In fact, different objects which overlap in the image may be grouped into a single box as if they were a unique object. This frequently happens for vehicles at different distances, the closest partly occluding the others. A specific procedure has been devised to decompose bounding boxes in multiple subrectangles each possibly framing a single object. Starting from the assumption that hot wheels must appear on the lower sides of the vehicle, this procedure searches for low intensity zones placed at the bottom corners of the bounding box. If such areas are detected and possess a specific size and aspect ratio compared to the box, the box is decomposed. In particular, if the height ratio between the empty rectangle and the box is greater than the width ratio, the box is vertically subdivided, otherwise it is horizontally subdivided. This choice tends to maximize one of the resulting subrectangles, i.e. the one which more likely encloses the foreground object. Figure 4 shows an example of decomposition.

### C. Vehicle Identification

Taken the assumption that each selected candidate represents a single object, the purpose of the next step is the identification of vehicles among candidate objects on the ground of specific features related to their shape. Two elements have been individuated as characterizing the information content of vehicles: the strong edges originated by the tyres and the typical pattern produced by the bottom of the vehicle (wheels and muffler)

In fact, tyres represent a common distinctive feature for vehicles with respect to other objects. In particular, while on the internal side the tire image tends to merge with the muffler and the heat coming from the engine, external and bottom edges of the tire are strongly visible (see figure 5.a). Therefore for each candidate edges are extracted and binarized through a Sobel operator. Then an edge pattern forming a  $90^\circ$  angle is searched for in each bottom corner of the box, as shown in figure 5. This approach follows the algorithm already developed in the visible domain and extensively tested on the ARGO vehicle [18]. In addition, in the IR-based processing the gradient direction of the edges is taken into account as well, since the object is expected to be brighter than the background. If two sufficiently long and

strong corners lying on the same horizontal line are recognized, the bounding box is labelled as valid and is assigned a vote proportional to the quality of the corners.

In order to achieve a robust identification, a second element characterizing vehicles has been individuated through the analysis of a high number of bounding boxes. It has been observed that the wheels and muffler (which are common to all kind of vehicles) generate a typical pattern in the bottom of the vehicle

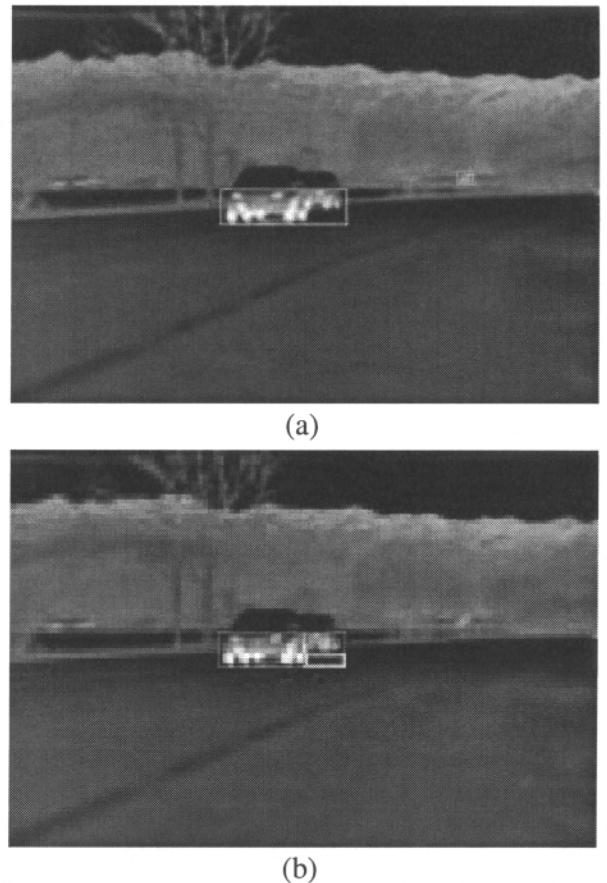


Fig. 4 (a) Overlapping vehicles and (b) result after the bounding box decomposition.

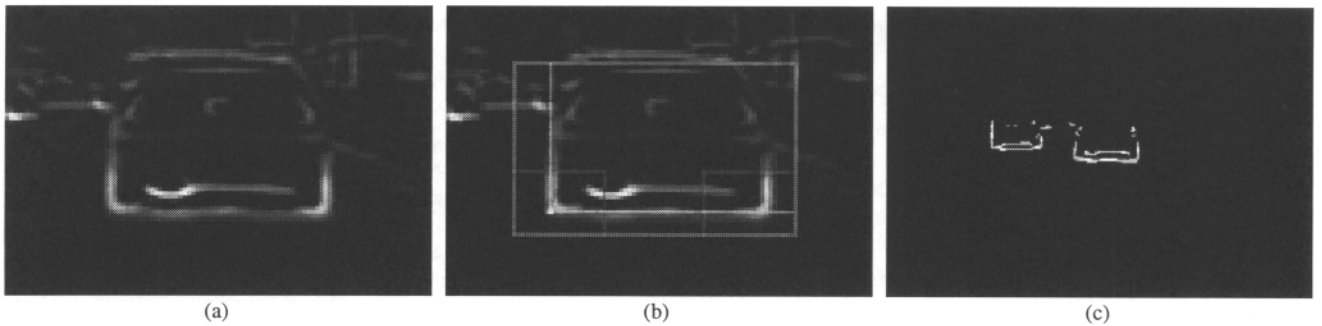


Fig. 5 Identification of corners in the edge image: (a) magnified edge image, (b) search areas and detection of left corner, (c) localized corners.

image. The histogram computed on the columns of the bounding box shows a common behavior for vehicles. More precisely, it exhibits high peaks in correspondence to the wheels and a valley in the central area. The shape of the histogram (that will be called *signature*) conveys information about the intensity distribution typical of vehicles.

The signature is even more representative when the histogram is computed on the edge image, since edge patterns are more similar from vehicle to vehicle (see figure 6). The signature is obtained from the analysis of a high number of images, averaging the histograms obtained from bounding boxes framing vehicles. Normalization and linear interpolation are performed to take into account different sized rectangles.

To identify vehicles, the histogram computed for each of the candidate bounding box is compared to the signature. Many different measures of similarity have been considered, the simpler being the computation of the area of the difference between the histogram and the signature. This method may fail due to different positions of the muffler and different shapes of vehicles. Therefore a new approach to the similarity measure has been evaluated. It is based on the comparison of the histogram shape to a synthetic template defined by rules. The rules under evaluation are:

- presence of peaks on the sides,
- valley in the middle with height less than one third of the peaks height.

A vote is assigned to the boxes with characteristic which sufficiently match the signature. The vote is averaged with the other votes attributed by preceding steps of the algorithm. A simple threshold discards low confidence results.

#### D. Vehicle Tracking

The result of the detection of the same vehicle in subsequent frames can move within the image or present different size and aspect ratio. Moreover, bounding boxes due to noise and the possibility that a vehicle can be missed in a frame must be considered as well.

In order to cope with these problems, a simple state machine is used. After the processing of the first frame a list of bounding boxes is produced. At this stage these bounding box are marked as *potential vehicles*, thus they do not appear in the current result.

The processing of subsequent frames produces new results that are compared against this list: new bounding boxes are sim-

ply added to the list while, if the system recognizes the match to already detected bounding boxes, the correspondent records are updated.

The position is used for matching two bounding boxes and determining if they represent the same vehicle.

When a bounding box representing the same vehicle has been tracked in a number of subsequent frames, it is marked as *actual vehicle*. Bounding boxes are removed from the list once they are no longer detected in several subsequent frames. The number of frames in the tracking window depends on the processing rate.

Obviously, this mechanism allows to discard bounding boxes due to noise, since they appear in a reduced number of subsequent frames (generally 1 or 2). On the other hand it permits to successfully track vehicles even if the detection is not continuous in each frame.

#### E. Distance Estimation

An estimation of vehicle distance based on the calibration of intrinsic and extrinsic camera parameters has been tested. The calibration procedure was carried out acquiring images of hot objects at known distances. Unfortunately, as shown in [6], a mere monocular calibration is too sensitive to vehicle (and thus camera) pitch and roll and to a non-flat road slope.

In the visible domain, stereoscopic vision can make the distance estimation more robust [6], but it is not a affordable solution with IR sensors due to their prohibitive cost.

Therefore, a different approach was adopted based on the assumption of a fixed vehicle width together with the knowledge of intrinsic calibration parameters. This assumption allows to estimate vehicle distance as a function of bounding boxes width. Figure 7 shows a comparison between this method for distance estimation against the calibration-based approach; it shows that vehicle width-based algorithm is less affected by vehicle movements or varying road slopes.

### III. RESULTS AND CRITICAL ANALYSIS

The system was installed on prototype vehicles and some preliminary tests have been carried out in the *Centro Ricerche FIAT* facilities and in extra-urban scenarios. Two different IR sensors were alternatively evaluated, both installed in the front part of the car at a height of 60 cm and 100 cm from the ground. Both devices feature a  $320 \times 240$  FPA format, while they differ in camera aperture and technology ( $15.5^\circ \times 11.6^\circ$  and uncooled

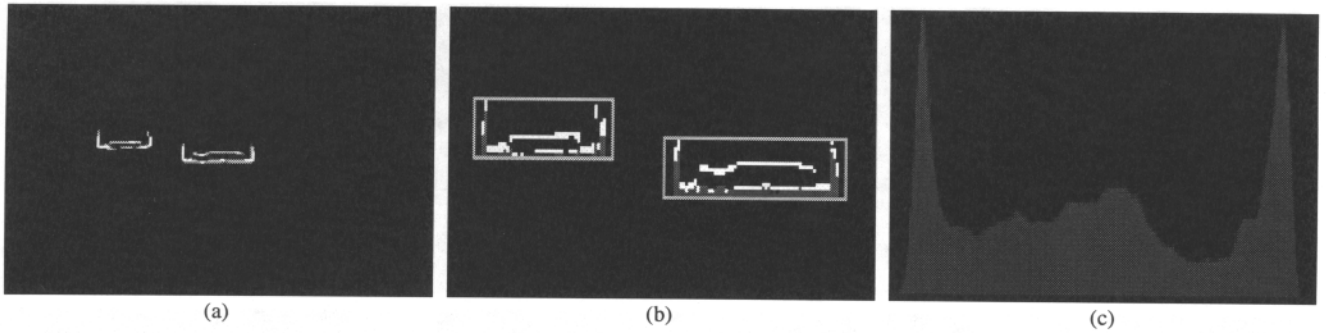


Fig. 6 Computation of signature: (a) binary edge image, (b) magnified edge vertical histogram, (c) signature obtained averaging 30 samples.

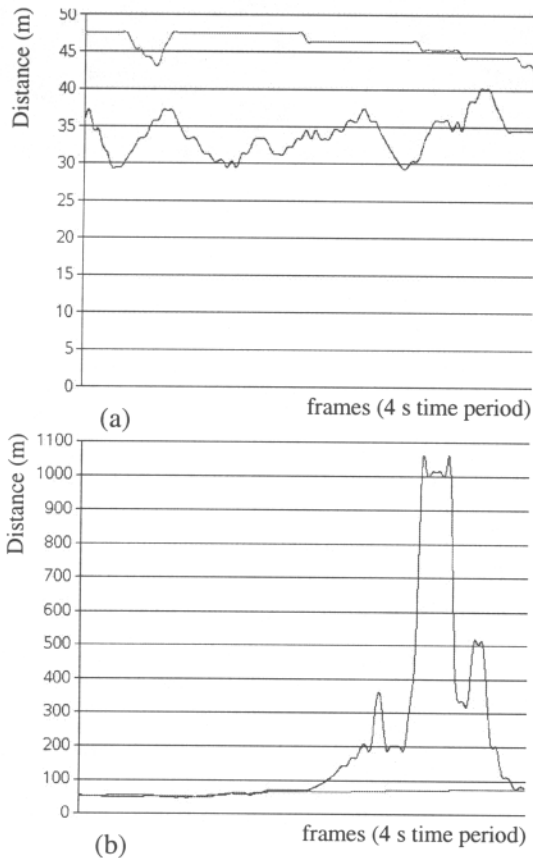


Fig. 7 Comparison between calibration-based (blue) and vehicle width-based (red) methods for distance estimation: (a) effect induced by vehicle pitch movements and (b) behavior in a hill.

microbolometer against  $18^\circ \times 13.5^\circ$  and pyroelectric and dielectric effects). The main appreciable difference is the improved image quality provided by the one based on pyroelectric and dielectric effects.

The system proved to be effective in different conditions: with or without vehicles, in different light conditions, in extra-urban scenarios, and able to detect vehicles up to 100 m. This distance limit is due to camera optics. The test period did not extend to different seasons (winter/summer) and did not include

appreciable weather changes (fog/rain/snow). Figure 8 shows the results in a few situations.

The system is able to successfully detect vehicles even when they overlap; when a large group of overlapping vehicles is present in the scene, only the closest to the vision system is correctly detected, while the others are sometimes grouped. Even if the algorithm is targeted to the detection of preceding vehicles, also vehicles moving in the opposite direction are generally detected though they emit a low amount of heat. Analogously, the developed method is sufficiently general to detect road participant without muffler such as trailers.

On the other hand, the most critical behavior corresponds to the presence of vehicles too close to the vision system; in these situations, vehicles feature large bright spots which may be interpreted as potential far away vehicles. This limits a reliable detection to vehicles at least 25 m far away from the vision system. Figure 9 shows the partial result of the attentive phase when vehicles are too close to the vision system.

The current not-optimized system is able to process images at a 12 Hz on a 1 GHz Intel architecture.

#### IV. CONCLUSIONS

This paper presents an algorithm for detecting vehicles in IR images. Initially the attention is focused on portions of the image that contains hot objects only. These areas are then selected and refined using aspect ratio and size constraints about vehicles; even situations with overlapping vehicles are considered. The result is further investigated exploiting specific vehicle thermal characteristics. A simple tracking phase is performed to improve the detection results.

Thanks to the knowledge of camera intrinsic parameters the distance of vehicles is computed using an assumption about vehicles width.

The system proved to be effective in different scenarios, but further tests are required to validate it in a wider range of weather conditions. It is able to detect vehicles in front of the vision system in the range 25 m ÷ 100 m at a 12 Hz processing rate.

#### REFERENCES

- [1] M. Bertozzi, A. Broggi, and A. Fascioli, "Vision-based Intelligent Vehicles: state of the art and perspectives," *Journal of Robotics and Autonomous Systems*, vol. 32, pp. 1–16, June 2000.
- [2] S. M. Smith, "ASSET-2: Real-time motion segmentation and object tracking," *Real Time Imaging Journal*, vol. 4, pp. 21–40, Feb. 1998.



Fig. 8 Results of the algorithm in different situations.

- [3] G. S. K. Fung, N. H. C. Yung, and G. K. H. Pang, "Vehicle Shape Approximation from Motion for Visual Traffic Surveillance," in *Procs. IEEE Intl. Conf. on Intelligent Transportation Systems 2001*, (Oakland, USA), pp. 610–615, Aug. 2001.
- [4] S. Kyo, T. Koga, K. Sakurai, and S. Okazaki, "A robust Vehicle Detecting and Tracking System for Wet Weather Conditions using the IMAP-VISION Image Processing Board," in *Procs. IEEE Intl. Conf. on Intelligent Transportation Systems'99*, (Tokyo, Japan), pp. 423–428, Oct. 1999.
- [5] A. Kuehne, "Symmetry-based vehicle location for AHS," in *Procs. SPIE - Transportation Sensors and Controls: Collision Avoidance, Traffic Management, and ITS*, vol. 2902, (Orlando, USA), pp. 19–27, Nov. 1998.
- [6] M. Bertozzi, A. Broggi, A. Fascioli, and S. Nichele, "Stereo Vision-based Vehicle Detection," in *Procs. IEEE Intelligent Vehicles Symposium 2000*, (Detroit, USA), pp. 39–44, Oct. 2000.
- [7] T. Kalinke, C. Tzomakas, and W. von Seelen, "A Texture-based Object Detection and an Adaptive Model-based Classification," in *Procs. IEEE Intelligent Vehicles Symposium'98*, (Stuttgart, Germany), pp. 341–346, Oct. 1998.
- [8] S. Denasi and G. Quaglia, "Obstacle Detection Using a Deformable Model of Vehicles," in *Procs. IEEE Intelligent Vehicles Symposium 2001*, (Tokyo, Japan), pp. 145–150, May 2001.
- [9] M. Lützel and E. D. Dickmanns, "Road Recognition with MarVEye," in *Procs. IEEE Intelligent Vehicles Symposium'98*, (Stuttgart, Germany), pp. 341–346, Oct. 1998.
- [10] M. Hariyama, T. Takeuchi, and M. Kameyama, "Reliable Stereo Matching for Higly-Safe Intelligent Vehicles and Its VLSI Implementation," in *Procs. IEEE Intelligent Vehicles Symposium 2000*, (Detroit, USA), pp. 128–132, Oct. 2000.
- [11] U. Franke, D. Gavril, S. Görzig, F. Lindner, F. Paetzold, and C. Wöhler, "Autonomous Driving Goes Downtown," *IEEE Intelligent Systems*, vol. 13, pp. 40–48, Nov.–Dec. 1998.
- [12] Z. Hu and K. Uchimura, "Tracking Cycle: A New Concept for Simultaneously Tracking of Multiple Moving Objects in a Typical Traffic Scene," in *Procs. IEEE Intelligent Vehicles Symposium 2000*, (Detroit, USA), pp. 233–239, Oct. 2000.
- [13] P. Barham, X. H. Zhang, L. Andreone, and M. Vache, "The Development of a Driver Vision Support System using Far Infrared Technology: Progress to Date on the DARWIN Project," in *Procs. IEEE Intelligent Vehicles Symposium 2000*, (Detroit, USA), pp. 545–549, Oct. 2000.
- [14] K. Owens and L. Matthies, "Passive Night Vision Sensor Comparison for Unmanned Ground Vehicle Stereo Vision Navigation," in *Procs. IEEE Intl. Conf. on Robotics and Automation*, vol. 1, (San Francisco, USA), pp. 122–131, Apr. 2000.
- [15] B. N. Nelson, "Automatic Vehicle Detection in Infrared Imagery Using a Fuzzy Inference-Based Classification System," *IEEE Trans. on Fuzzy Systems*, vol. 9, pp. 53–61, Feb. 2001.
- [16] M. Kagesawa, S. Ueno, K. Ikeuchi, and H. Kashiwagi, "Recognizing Vehicles in Infrared Images Using IMAP Parallel Vision Board," *IEEE Trans. on Intelligent Transportation Systems*, vol. 2, pp. 10–17, Mar. 2001.
- [17] M. Kagesawa, A. Nakamura, K. Ikeuchi, and H. Saito, "Local-Feature Based Vehicle Class Recognition in Infra-red Images using IMAP Parallel Vision Board," in *Procs. IEEE Intelligent Vehicles Symposium 2000*, (Detroit, USA), pp. 334–339, Oct. 2000.
- [18] A. Broggi, M. Bertozzi, A. Fascioli, C. Guarino Lo Bianco, and A. Piazzi, "Visual Perception of Obstacles and Vehicles for Platooning," *IEEE Trans. on Intelligent Transportation Systems*, vol. 1, pp. 164–176, Sept. 2000.



Fig. 9 The result of the attentive phase in a situation where two vehicles are too close to the vision system.