

Stereo Vision-Based Start-Inhibit for Heavy Goods Vehicles

Massimo Bertozzi, Alberto Broggi, Paolo Medici, Pier Paolo Porta
Dipartimento di Ingegneria dell'Informazione
Università di Parma
Parma, I-43100, Italy
{bertozzi, broggi, medici, portap}@ce.unipr.it

Agneta Sjögren
Volvo Technology Corporation
Dept 6420, M1:6
SE-405 08 Göteborg, Sweden
Agneta.Sjogren@volvo.com

Abstract—This paper presents a robust method for obstacle detection with stereo cameras. Arbitrarily aligned cameras are calibrated using a dense grid; a direct mapping between image pixels and world points is made to remove lens distortion and perspective in the same pass. Cubic splines are used to recover unknown points not present in the grid. After the transformation phase, left and right images are compared and the differences are analyzed using a polar histogram to detect vertical structures and to reject noise and small objects. World coordinates of detected objects are recovered and fed to the sub-system for further processing and to take appropriate actions.

According to experimental results, the proposed algorithm can be useful in different automotive applications, requiring real-time segmentation without any assumption on background. In particular the system has been tested to investigate presence of obstacles in blind spot areas around heavy goods vehicles (HGV).

The system presented in this paper is currently installed in a Volvo truck.

I. INTRODUCTION

Statistics indicate that heavy trucks are over-represented in traffic accidents resulting in fatalities. For instance figures from Great Britain [1] indicate that in year 1998 heavy goods vehicles (HGV) were involved in 17% of all road accidents resulting in fatalities, despite making up just 7% of the traffic.

Many accidents involving trucks are related to the limited field of view of the driver. There are large blind spots all around the vehicle (figure 1). Some of these blind areas can be at least partly covered by additional mirrors. However this is not always an optimal solution considering the aerodynamic effects and also the resulting complex driver interface.

Examples of traffic situations (see figure 2) where the limited field of view can result in conflicts are:

- lane change and turn situations to the passenger side,
- starting from stationary at cross-walks or other places where a person or an object can be close in front of the vehicle,
- situations with cross road traffic sideways,
- and back-up situations especially when ranging up to a loading dock.

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

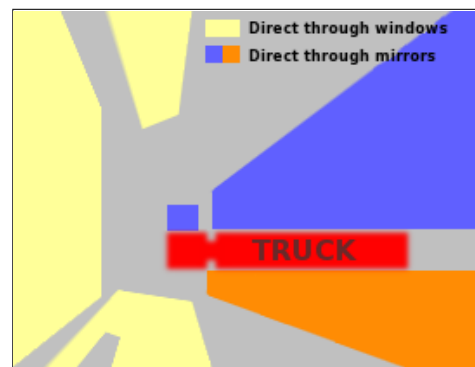


Fig. 1. Typical field of view at ground level for a truck driver.

The truck application developed within APALACI concentrates on covering the forward blind spot. This type of accidents accounts for approximately 10% of all accidents between trucks and unprotected road users and about 20% of all fatal accidents between trucks and unprotected road users.

Different countermeasures that could reduce the fatalities in accidents involving heavy goods vehicles have been identified in [2]. Concerning the category of unprotected road users, very few countermeasures, that could substantially reduce the toll of fatalities, were identified (see table I). The most effective single measure would be to improve the forward vision from HGV cabs so that an average size pedestrians could be seen even when standing right up against the front of the vehicle. This, according to the report, would have been likely to save the lives of 12% of the pedestrians killed by HGVs. Changing the design of the front of a truck in this way is not an easy task. Similar benefits can be achieved by using sensors to detect the presence of a pedestrian or an obstacle and to warn the driver and also to prevent the vehicle from taking off when something is present in the forward blind spot.

In the initial phase of the APALACI project, a study was performed in order to get input from professional drivers on how they view the problem of blind zones and what type of solution they would like to see (also other new safety functions were considered in this study). The participants were well aware of the problem with the blind zone in front of the

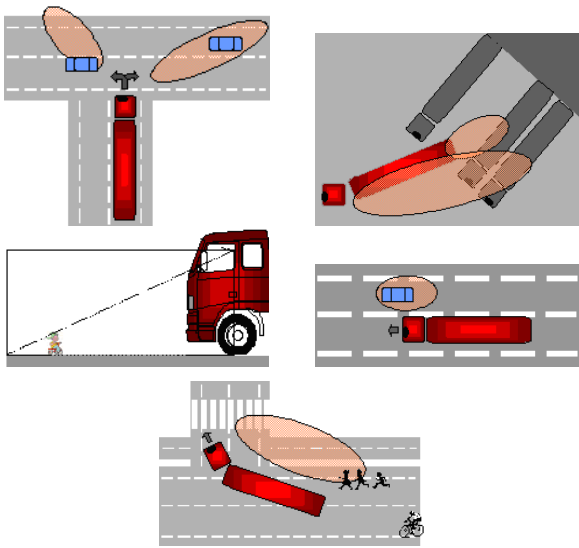


Fig. 2. Examples of dangerous traffic situations due to blind spots.

TABLE I
POTENTIAL BENEFITS FOR UNPROTECTED ROAD USERS FROM VARIOUS
COUNTERMEASURES.

| Options | Estimate of savings (% of fatalities in impacts with HGVs) |
|---|--|
| Improve forward vision from HGV cab | 12% of pedestrian |
| Fit current sideguards to all HGVs > 3500 kg (thus ending exemption) | 4% of pedal cyclist, 1% of pedestrian |
| Fit sideguards with increased strength and lower ground clearance to all HGVs | 5% of pedal cyclist, 3% of motorcyclist, 2% of pedestrian |
| Fit solid "side skirts" to all HGVs | 9% of pedal cyclist, 5% of motorcyclist, 3% of pedestrian |

vehicle. They thought the idea of giving information about blind spots in front of the vehicle and to give information about vulnerable road users was good. The participants were in general skeptic to systems intervening with the driving task. However, a system inhibiting the driver from starting when something is present in the blind spot, could be acceptable according to the participants, but only if it is easy to override. In particular, the participants mentioned situations like entering a ferry where you need to get really close to the vehicle in front. Considering human machine interface (HMI) the participants favored a camera display as the main driver interface. They wanted to see by themselves what was obstructing them.

The results from the focus group study show that the expected acceptance is high for the Start-inhibit and pedestrian recognition application. The drivers thought that the system would be useful, effective, assisting and raising alertness. The drivers were however concerned with problems related to technological maturity such as reliability, durability, cost and maintenance. The result from the focus group discussions

has influenced the system solution in such a way that a camera system is used as one of the object sensors, in order to be able to display an image of the blind spot to the driver and that special considerations have been taken to situations where the driver needs to drive close to the vehicle in front. In order to achieve a robust system two short range radars are also added for the purpose of obstacle detection. Figure 3 shows an overview of the system solution in the demonstrator/experimental vehicle.

A short form of the final functional description of the Start-inhibit function is: the Start-inhibit application will assist the driver by automatically checking the proximity of the vehicle and in case of an obstacle or pedestrian in the specified area the system will inform the driver before take-off. If the driver neglects the warning and there is a risk of collision, the system will also prevent the vehicle from taking-off. The system shall be possible to override through a driver/operator manoeuvre. Acceleration from stand still will be prohibited when there is a vulnerable road user in the area close in front of the truck. If the vehicle is starting from parked position (gear stick in parked position), acceleration will also be prohibited if any obstacle is detected in front of the vehicle. If the vehicle has been driving and is stationary (with gear in drive-position) acceleration shall only be inhibited if a vulnerable road user is detected in front of the vehicle. The presence of other objects shall only result in a warning to the driver. The reason for this separation between vulnerable road users and other objects is to avoid the system from intervening in queue driving or other dense traffic situations.

The warnings presented to the driver will be of either visual or acoustic character depending on the situation [2]. A display on an on-board monitor was evaluated and discarded according to the opinion of drivers focus group. In fact, drivers are already required to check many mirrors and several other visual information.

In this paper a stereo vision-based system for Start-inhibit is presented. The paper is organized as follows. A first section introduces the system architecture and the stereo vision system dividing between the calibration problem and the algorithms developed, then a section on results and conclusion is proposed.

II. THE SYSTEM

The system is composed of two cameras with spheric lenses to get a wide field of view, but introducing a strong distortion on images. They are placed right below the front window, as in figure 4, and are arbitrary aligned. In particular only the driver blind spot area is framed by the cameras.

Two well known approaches for stereo obstacle detection have been considered:

- the computation of the disparity of each pixel [3],
- the use of Stereo Inverse Perspective Mapping [4].

An obstacle detection algorithm for offroad autonomous driving is presented in [3]. The dominant surface (e.g. the ground) is found through a v -disparity image [5] computation, while the obstacles come from a Disparity Space Image (DSI)

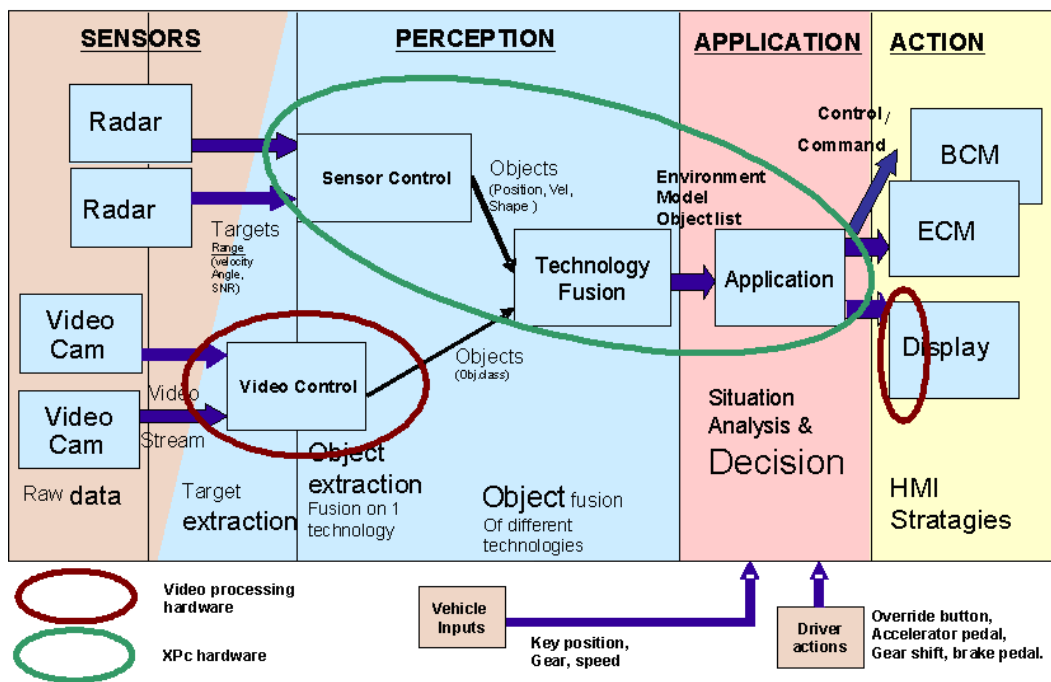


Fig. 3. Start-inhibit in the scope of the APALACI architecture.



Fig. 4. Position of the cameras.

analysis. In this case the cameras' axis of the stereo system are almost parallel to the ground. Unfortunately, this approach is not suitable for Start-Inhibit, because one of the most important design issues is to not force a specific cameras' alignment. Infact the approach described in [3] requires a perfect calibration and constraints on cameras orientation.

Therefore a stereo inverse perspective mapping-based approach has been considered. The whole processing is performed by means of two main steps:

- distortion and perspective removal from the stereo images,
- obstacle detection.

Concerning the first step, the problems of distortion removal and inverse perspective mapping without the knowledge of the intrinsic and extrinsic parameters of cameras have to be solved. Lens distortion is usually modelled as polynomial radial distortion [6,7] and it is removed by recovering coefficients of this polynomial. After the distortion removal phase, extrinsic parameters are obtained [8].

Nevertheless, the highly complex mathematical model of the sferic lens would have affected the computational time; therefore, a graphic interface to manually associate the grid points of the source image to their homologous points on a square grid on the *IPM* image has been developed [4]. This pre-processing is performed offline.

In order to detect obstacles, two different approaches have been tested: the first searches for connected blocks on the thresholded image generated from the difference between left and right images after distortion removal and inverse perspective mapping; the other is based on the use of a polar histogram (see [9,10]). These two approaches have been fused into one algorithm to get the best from both. The whole algorithm flow chart is described in figure 5 and discussed in the following.

A. Calibration

Camera calibration is one of the most important topics for vision systems.

In our case, highly distorting cameras are used without any knowledge about the intrinsic and extrinsic camera parameters. An analytic approach to calibration would be computationally prohibitive: the equations that are normally used to model

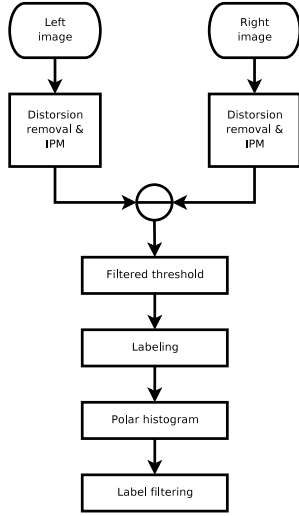


Fig. 5. Algorithm's block diagram.

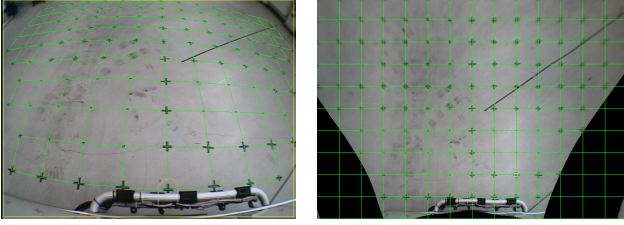


Fig. 6. Original and undistorted image of the grid.

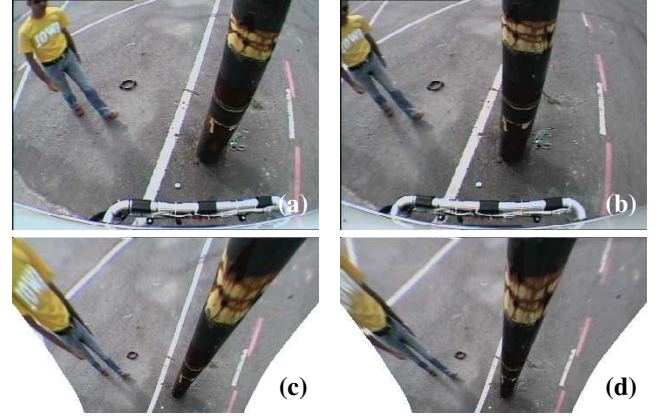
spheric lenses have an order too low to be applied in these circumstances because of the wide angle of the camera.

Therefore an empiric strategy has been used: during an offline preprocessing a look-up table that allows a fast pixel remapping is generated; namely each pixel of the distorted image is associated to its corresponding pixel on the undistorted image. Images of a grid, painted on a stretch of flat road in front of the truck, are used to compute the look-up table (see figure 6). A manual system to mark all the crossing points on the source image is used.

Thanks to the knowledge of the relative position of the truck compared to the grid itself and to the assumption that the road can be considered nearly flat in the proximity of the vehicle, it is possible to compute a new image (the *IPM* image) removing both the perspective effect and camera distortion at once. A non-linear interpolation function is used to remap the pixel of the source image that are not cross points.

The process to find out the coordinates pair (x,y) of the source image from the (i,j) pixel of the *IPM* image is divided in two steps.

Let's assume to have a grid with N vertical lines and M horizontal lines. For each vertical line of the grid, a function f_n is defined, where $n \in [1,N]$ is the line number. The spline creation is constrained by the correspondences between the crossing points of each line in the source image and in the *IPM*


 Fig. 7. Perspective and distortion removal: (a) left source image. (b) right source image. (c) left *IPM* image. (d) right *IPM* image.

image; see equation (2) as an example, assuming x_1, y_1, x_2, y_2 , etc. as the coordinates of the crosspoints on the source image.

$$f_n(j) : \mathfrak{R} \longrightarrow \mathfrak{R}^2 \quad f_n(j) = \begin{cases} f_n^x(j) \rightarrow x \\ f_n^y(j) \rightarrow y \end{cases} \quad (1)$$

$$\begin{cases} f_1^x(0) = x_1 \\ f_1^y(0) = y_1 \\ f_1^x(1) = x_2 \\ f_1^y(1) = y_2 \\ \vdots \\ f_1^x(N) = x_N \\ f_1^y(N) = y_N \end{cases} \quad (2)$$

Using functions $f_1(j), f_2(j), \dots, f_N(j)$ another class of functions can be built, called $g_j(i)$ and defined as described in equation (3) with equation (4) as constraint.

$$g_j(i) : \mathfrak{R} \longrightarrow \mathfrak{R}^2 \quad g_j(i) = \begin{cases} g_j^x(i) \rightarrow x \\ g_j^y(i) \rightarrow y \end{cases} \quad (3)$$

$$\begin{cases} g_j(1) = f_1(j) \\ g_j(2) = f_2(j) \\ \vdots \\ g_j(N) = f_N(j) \end{cases} \quad (4)$$

In this way all the pixels of the *IPM* image have a correspondence to a pixel of the source image and the cubic spline interpolation method allows to get the best match between the two sets of pixels. An example of the resulting images obtained using these equations is shown in figure 7.

Being the system based on stereo vision, two tables, one for each camera and both fixed under the same reference frame, are calculated with this procedure. Look-up table generation is a time-consuming step, anyway it is computed only once when the cameras are installed or when their position is changed.

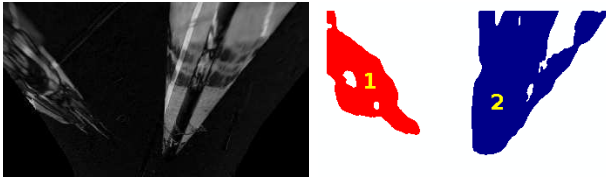


Fig. 8. Difference image and labeling.

B. Algorithm

Starting from the *IPM* images, a difference image D is generated comparing every pixel i of left image (L) from the right one (R) and computing an absolute distance value.

$$D_i = |L_i - R_i| \quad (5)$$

In particular, working on RGB color images, the distance used is the average of absolute differences on color channels.

Then a particular threshold filter is applied on the resulting image D . In particular for each pixel is defined a square area A centered on it; the average value m of all the pixels in that area is computed and a threshold is applied on m . The resulting value is assigned to the pixel (see equation (6)).

$$\forall i \in D \quad m = \frac{\sum_{j \in A} D_j}{\sum_A 1} \quad T_i = \begin{cases} 0 & \text{if } m < \gamma \\ 1 & \text{if } m > \gamma \end{cases} \quad (6)$$

This is a kind of low-pass filtering and is useful to find the most significant differences in these images. Compared to similar methods like a thresholding followed by an opening, it is faster because easy to be optimized and, nevertheless, works on the whole scope of values of grey images

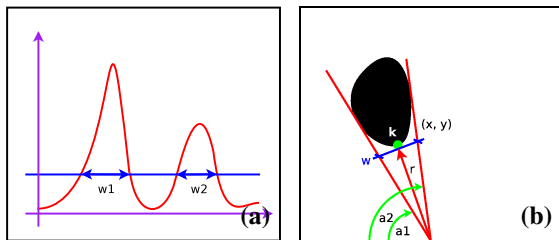


Fig. 9. (a) Polar histogram thresholding and filtering and (b) information extracted from the detected obstacle.

On the resulting image, connected areas are searched for and labeled: a progressive number is assigned to each label for further identification (as shown in figure 8). A polar histogram is computed for each region. The focus used to compute the polar histogram is the projection of the mid point between the two cameras onto the road plane. These regions produce peaks on the polar histogram. Thus the presence of strong peaks can be used to detect obstacles.

Some specific configurations of this histogram have to be considered, due to regions that are weakly connected or too

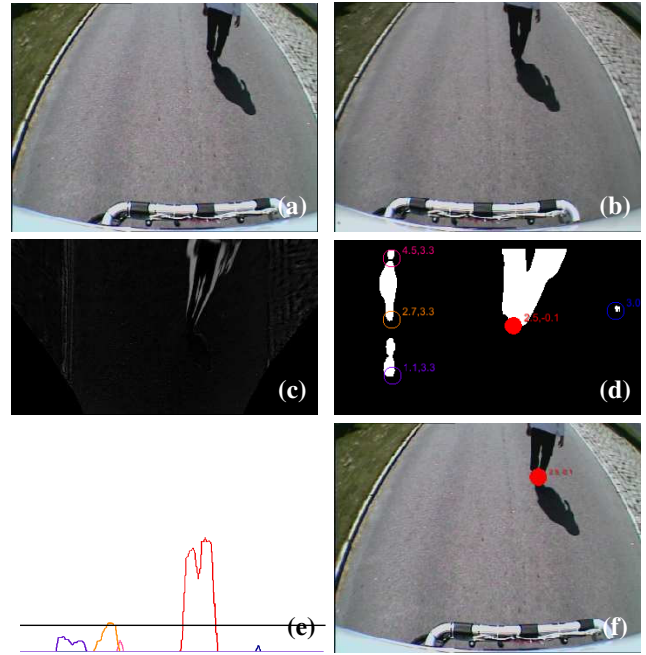


Fig. 10. (a) left source image (b) right source image (c) difference image (d) labeling (e) polar histograms (f) result image.

thin to be a real obstacle. Therefore it is necessary to further filter the polar histograms to remove regions that can not be considered as obstacles.

This filtering is performed considering the width of the histogram for the region of interest. The width of the histogram is computed in correspondence to a given threshold. When a polar histogram features more peaks, several values of width (w_1, w_2 , etc.) are generated (see figure 9.a). If $\max\{w_1, w_2, \dots, w_n\} > w_{min}$ (where w_{min} is a width threshold) then the region previously labeled is maintained, otherwise it is discarded.

For each resulting region the point k closest to the origin of the polar reference system and the angles of view (a_1, a_2) under which the region is seen are computed (see figure 9.b).

A rough width (w) of the detected object is computed as well, applying the following equation and considering r as the distance of k from the focus.

$$w = 2r \cdot \tan\left(\frac{a_2 - a_1}{2}\right) \quad (7)$$

Working on the *IPM* image, the location in world coordinates of point k can be estimated through the same look-up table previously used.

Figure 10 shows the complete set of intermediate results starting from the left end right original images; the difference and labeled images; the polar histogram whose filtering allows detecting one obstacle and discarding the small road curb; and finally the left original image with red marker indicating the obstacle.

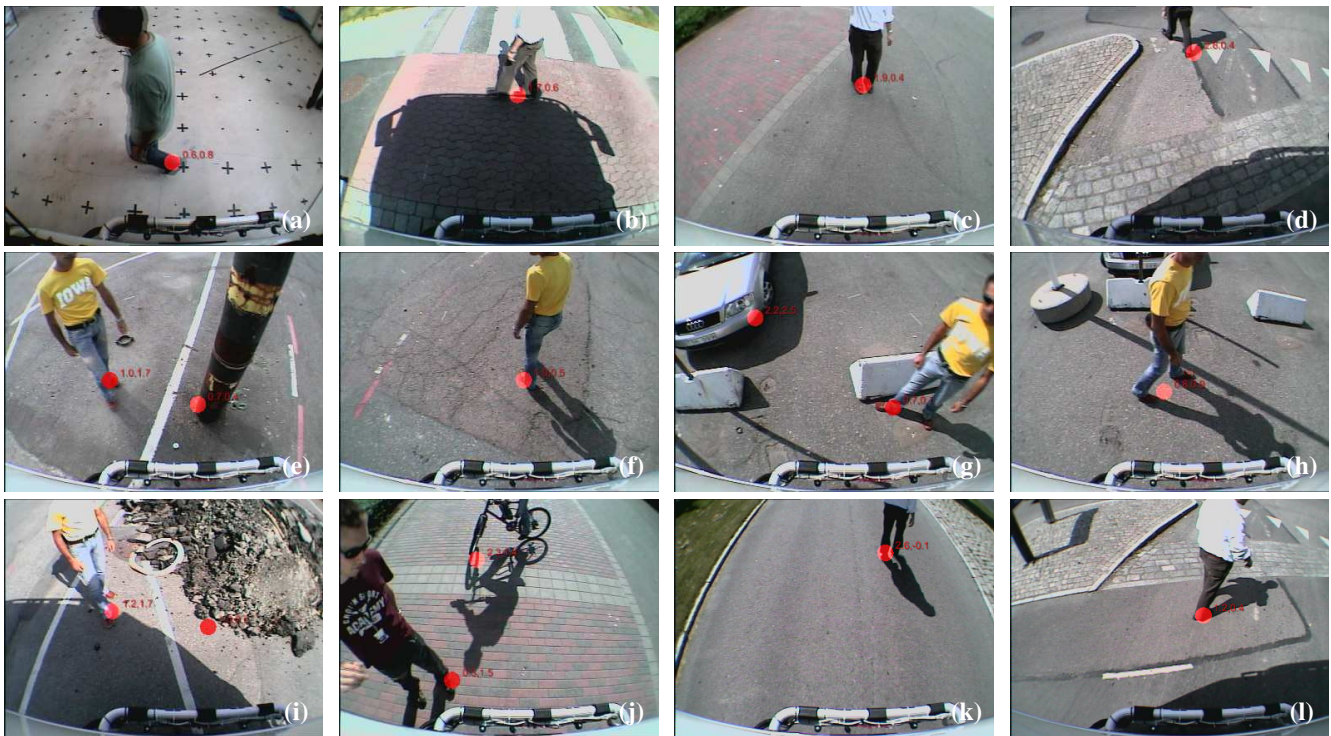


Fig. 11. Result images, showing typical algorithm output. A red dot shows the closest point of contact of each obstacle on the ground.

III. DISCUSSION

This paper presents an easy, fast, and reliable stereo obstacle detection technique for a start inhibit system. Cameras mounted on a vehicle are arbitrary aligned, meaning that no special alignment is required by specialists or IT professionals. The algorithm works on stereo image pairs in PAL resolution (768×288 merging one field with the other). At this resolution, the algorithm takes an average of 24 ms to process one frame on a 3 GHz Intel Pentium 4.

Tests were made in several environmental conditions considering different kinds of road and obstacles. Figure 11 shows some examples of the algorithm output remapped onto the original image. However in the final application, the output is provided to the vehicle driver on undistorted images [7] through a look-up table built on the same grid used to generate the bird-view. Red circles are used to mark the positions of obstacles.

Taking advantage of the stereo approach, the road texture, road markings, and shadow are successfully filtered out. Moreover, the algorithm easily detects large obstacles, rejecting most of smaller ones, like sidewalk borders. In general, due to the particular configuration of the system, any vertical object is correctly detected, thus the use of image tracking or temporal comparisons seems not mandatory.

Experimental results show that cameras movements caused by variation of pitch and roll of vehicle at low speed driving are not large enough to generate error and are efficiently removed.

Overlapping objects are detected as a single and large obstacle, as shown in figure 11.g.

REFERENCES

- [1] L. Thatcher, "Fatalities from Accidents Involving Heavy Goods Vehicles. Trends, Causes and Countermeasures." TRL Ltd, Tech. Rep. PR/SE/026/99, June 1999.
- [2] APALACI, "Selection of sensors and component configuration on the basis of data on cost and performance. Revision of specs for the experimental set-up." PreVENT, Tech. Rep. D50.51, 2005.
- [3] A. Broggi, C. Caraffi, R. I. Fedriga, and P. Grisleri, "Obstacle Detection with Stereo Vision for off-road Vehicle Navigation," in *Procs. Intl. IEEE Wks. on Machine Vision for Intelligent Vehicles*, San Diego, USA, June 2005.
- [4] M. Bertozzi, A. Broggi, and A. Fascioli, "Stereo Inverse Perspective Mapping: Theory and Applications," *Image and Vision Computing Journal*, vol. 8, no. 16, pp. 585–590, 1998.
- [5] R. Labayrade, D. Aubert, and J.-P. Tarel, "Real time obstacle detection on non flat road geometry through V-disparity representation," in *IEEE Intelligent Vehicles Symposium, Versailles*, June 2002, pp. 646–651.
- [6] D. Claus and A. Fitzgibbon, "A rational function lens distortion model for general cameras," in *Procs. Conf. on Computer Vision and Pattern Recognition*, vol. 1, San Diego, USA, June 2005, pp. 213–219.
- [7] F. Devernay and O. D. Faugeras, "Straight lines have to be straight," *Machine Vision and Applications*, vol. 13, no. 1, pp. 14–24, 2001.
- [8] Roger Y. Tsai, "A Versatile Camera Calibration Technique for High-accuracy 3D Machine Vision Metrology Using off-the-shelf TV Cameras and Lenses," *IEEE Journal of Robotics and Automation*, vol. 3, pp. 323–344, Aug. 1987.
- [9] M. Bertozzi and A. Broggi, "GOLD: a Parallel Real-Time Stereo Vision System for Generic Obstacle and Lane Detection," *IEEE Trans. on Image Processing*, vol. 7, no. 1, pp. 62–81, Jan. 1998.
- [10] K. Lee and J. Lee, "Generic obstacle detection on roads by dynamic programming for remapped stereo images to an overhead view," in *IEEE International Conference on Networking, Sensing and Control*, vol. 2, Taipei, Taiwan, Mar. 2004, pp. 897–902.