

The passive sensing suite of the TerraMax autonomous vehicle

Alberto Broggi*, Andrea Cappalunga*, Claudio Caraffi*, Stefano Cattani*, Stefano Ghidoni*,
Paolo Grisleri*, Pier Paolo Porta*, Matteo Posterli*, Paolo Zani*, and John Beck†

*VisLab – Dipartimento di Ingegneria dell'Informazione
Università degli Studi di Parma, ITALY

<http://vislab.it>

{broggi, kappa, caraffi, cattani, ghidoni, grisleri, portap, posterli, zani}@vislab.it

†Oshkosh Corporation

Oshkosh, WI, USA

JBeck@oshtruck.com

Abstract—This paper presents the TerraMax autonomous vehicle, which competed in the DARPA Urban Challenge 2007. The sensing system is mainly based on passive sensors, in particular four vision subsystems are used to cover a 360° area around the vehicle, and to cope with the problems related to complex traffic scenes navigation. A trinocular system derived from the one used during the 2005 Grand Challenge performs obstacle and lane detection, twin stereo systems (one in the front and one in the back) monitor the area close to the truck, two lateral cameras detect oncoming vehicles at intersections, and a rear view system monitors the lanes next to the truck looking for overtaking vehicles. Data fusion between laserscanners and vision will be discussed, focusing on the benefits of this approach.

I. INTRODUCTION

In this paper a prototype for an autonomous vehicle is presented. TerraMax, shown in Figure 1, took part to the DARPA Urban Challenge in Victorville, CA on November 3rd, 2007. After having successfully completed the National Qualification Event, TerraMax, together with 10 other vehicles, participated to the final race; unfortunately, TerraMax was unable to complete the race due to an unrecoverable software fault while exiting a parking lot.

The new vehicle derives from the one that participated and completed the DARPA Grand Challenge 2005 [1]; during that race, the fully autonomous truck was successful in demonstrating GPS waypoints following, obstacle avoidance, and tunnels, narrow roads and cliffs negotiation, in 28 hours of nonstop continuous operation.

The differences between the current edition of the race and the previous ones are deep, starting from the dynamic environment. The vehicle has to handle many complex traffic situations and deal with both human-driven and other autonomous vehicles. Development efforts have been targeted not only at meeting the Urban Challenge requirements, but also at making the final system robust enough to be considered a top contender.

TerraMax is being developed to produce an autonomous truck capable of providing logistics support, although its size brings additional hurdles, like the higher manoeuvring



Fig. 1. The TerraMax vehicle.

precision – compared to other smaller vehicles – needed to follow lanes barely wider than the truck itself. The final goal, beyond the Urban Challenge competition, is to engineer a system that can be supplied in kit form.

TerraMax must be able to sense its surroundings, understand situations, and react quickly; a service-based system architecture coupled with efficient path planning, trajectory algorithms and behavior modes ensures that sensor data is processed in a timely manner, while robust low level control and X-by-Wire components enable TerraMax to manage situations.

A complete allround-view is of paramount importance because the vehicle has to be very precise in movements, and take into account the environment and its evolution before making any decision: to meet these requirements a sensors belt has been set up around the truck. The choice to primarily use passive sensors is dictated by the will of producing a vehicle suitable for mass production: active ones are likely to interfere with each other, thus potentially degrading the performance of multiple autonomous vehicles operating in close proximity.

TerraMax features 4 vision systems, shown in Figure 2, each one devoted to monitor a particular area next to the truck; section IV explains in detail their purpose. Three laserscanners (LIDARs) are also mounted onboard, two in the front and one in the back, as shown in Figure 3, together with all the other sensors.

The integration of a vision system with active sensors is useful because the two technologies are complementary, and in some ways redundant: while the problem of lane detection is mainly addressed using vision [2] [3] [4], obstacle detection can get many benefits from an intelligent fusion with scan data. This topic will be discussed in section III.

II. TERRAMAX VISION SYSTEM

There are four vision systems onboard: trinocular, stereo, rearview, and lateral; each of them is composed of a computer connected to a number of cameras and LIDARs, depending on the application. Computers are equipped with a 2.00 GHz, dual core CPU, 2 GB of RAM and dual GBit ethernet. A 4 GB flash disk, chosen for robustness reasons, allows to load the operating system and the vision software. The computers are placed into the cab together with the rest of the autonomous driving system, in a box under the passengers seats, as shown in Figure 4. The vision PCs are cooled by fans, and the temperature in the box is kept low thanks to an air conditioning system. Vision computers are powered through solid state switches, so that they can be turned on an off by using a manual switch operated by a human from the truck dashboard, or through a CAN message sent by another system.

Each computer is connected through an 800 Mbps, FireWire B link to a subset of the 11 cameras mounted on the truck, depending on the system purpose (see Table I for details). Polarizing films and polarizers have been mounted in front of each lens to reduce reflections and increase contrast: reflections, especially due to sunlight may cause problems like smear.

Cameras for the trinocular system have been mounted inside the cab. Other cameras have been mounted outdoor, into appropriately sealed camera housings from APG. These camera mounts also serve as a protection against direct sunlight, but – despite this – in sunny and non windy days the temperature inside the housings can reach 45°C.

Each vision computer is connected to the truck network through a 100 Mbps switch. In particular, this connection

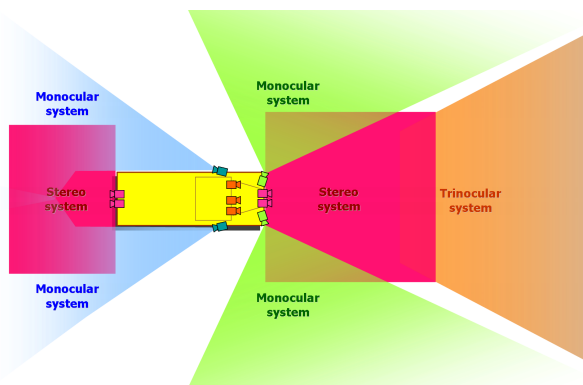


Fig. 2. All the Terramax vision subsystems and their placements. Color coding: red – stereo systems; orange – trinocular system; blue – rear view systems; green – lateral systems.



Fig. 3. All the Terramax sensors highlighted. Color coding: red – cameras; green – LIDARs; blue – GPS antennas.

gives access to one or more LIDARs, and to the other systems on the truck such as the World Perception Server (responsible for populating a persistent world map with the high-level data coming from the various sensors), GPS, INS, and the Vehicle Manager. LIDARs supply both raw (scan points) and processed (objects) data. Vision systems decode raw LIDAR data in order to perform a low-level fusion, as explained in sections III and IV.

All the cameras and LIDARs are synchronized at 12.5 Hz by a specific sync box. The time of all the system is kept by an NTP server running on the trinocular PC.

All the vision computers run the same software framework, and the various applications are implemented as separate plugins. This architecture allows to abstract hardware, while making a common processing library available to the applications, thus making algorithms development independent of the underlying system. Each vision system can be modeled as a blackbox receiving as input a number of image streams, along with several additional data from other auxiliary sensors (such as LIDARs, GPS, and INS units), performing a specific processing on this data, and supplying a set of results. Typically the outputs are high level data such as lists of classified objects, lanes descriptions, information on obstacles.

Each system controls its cameras using a selective auto-exposure feature. Each image is quickly analyzed in order to measure the light in a specific region of interest. The exposure values for each camera (shutter and gain) are then adjusted between fixed and independent ranges of shutter and gain using a PID controller. The controller acts to keep the difference between the image average and the setpoint under a fixed error value. Moreover, to reduce the amount of noise in the image in normal light conditions, the shutter has the precedence, and is changed keeping the gain to its minimum value. The gain is incremented in low light conditions, once

TABLE I
SUBSYSTEMS SUMMARY

	TRINOCULAR	STEREO	LATERAL	REARVIEW
<i>Cameras</i>	3 cameras 1024×768	4 cameras 1024×768	2 cameras 1920×1080	2 cameras 1024×768
<i>Cameras Position</i>	Upper part of the windshield, inside the cab	Two on the front camera-bar, two on the back of the truck, all looking downwards	On the sides of the front camera-bar	External, on top of the cab, looking backwards and downwards, rotated by 90°
<i>Optics</i>	focal length 4.16 mm, f2.16, FoV: 60°×42°	focal length 1.8 mm, f1.4, FoV: 154°×115°	focal length 8 mm, f1.4, FoV: 80°×63°, for HiRes sensors	focal length 4.16 mm, f2.16, FoV: 60°×42°
<i>LIDAR</i>	Front	Front, back	Not used	Back
<i>Algorithms</i>	Lane detection, stereo obstacle detection	Lane detection, stop line detection, curb detection, short-range stereo obstacle detection	Moving traffic detection	Overtaking vehicles detection
<i>Range</i>	7 to 40 m	0 to 10 m and -6.5 to -16.5 m	10 to 130 m	-4 to -50 m
<i>Notes</i>	3 stereo systems with baselines: 1.156 m, 0.572 m, 1.728 m	2 stereo systems (front and rear)	Enabled when the truck stops at crossings	Monitors the adjacent lanes to allow/inhibit lane change manoeuvres

the shutter has reached its maximum value. In the trinocular and stereo systems, where multiple cameras are used at once for processing, having the intelligent autoexposure active on all of them would be a problem, given the inevitable drift in shutter and gain values applied to each sensor; in order to obtain uniform acquisition conditions each subsystem features a master-slave architecture, where a single camera computes the desired values, and broadcasts them to the others.

All the vision systems are fault tolerant with respect to one or more, temporary or permanent, sensor failure events. The software is able to cope with FireWire bus resets or LIDAR communication problems, and to reconfigure itself to manage the remaining sensors, providing one uniform interface to handle this kind of events. Each specific application is responsible to react to the framework sensor failure messages depending on its purpose. For example, the trinocular may still work if one camera or the LIDAR become unavailable, but cannot if two cameras are broken. Likewise, the lateral and rearview systems might lose one camera, but still keep on performing the required processing on the remaining one. The stereo system can still function if the LIDAR is not sending updates, but it cannot if one of the cameras fails.

III. DATA FUSION

The vehicle is equipped with three LIDARs, capable of providing a 4-planes scan of the surrounding environment. The sensors output both processed data (such as detected obstacles) and raw points: while the former is useful for the high-level path planner, the latter can be exploited to improve the performance of the various vision systems.

Vision and LIDAR systems are complementary in many ways: for example, when performing a tridimensional reconstruction of the world LIDAR points offer a very accurate, yet sparse representation of the world, while vision provides dense, but less accurate measurements; some tasks, like lane and path detection, are also better accomplished through the use of vision. Integrating the data coming from the two

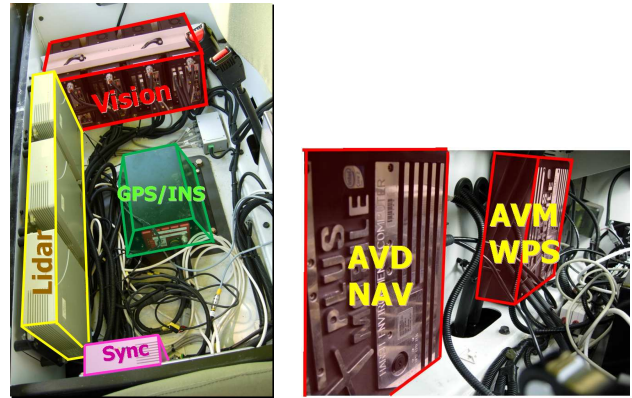


Fig. 4. PC location inside the passenger seat. On the left image PCs managing all different sensors and the synchronization device are highlighted. On the right image are shown Autonomous Vehicle Driver (AVD) and Navigation System PC (NAV) and World Perception Server (WPS) and Autonomous Vehicle Manager (AVM) PC.

sensors in useful, non-trivial ways is a challenging operation, but while it poses additional problems, like multi-system calibration (which must be performed accurately in order to collect consistent data, e.g. obstacle positions), it can also significantly improve the overall system behavior. Each vision subsystem implements a unique kind of fusion, which will be detailed in the following paragraphs, but all adhere to the common design guideline of being capable of working independently of the presence of laser data, to be robust against LIDAR sensor failures.

The final goal has been to achieve a vision system set that can reliably perceive all round the vehicle even without any fusion with other sensors. Data fusion helps the overall system to manage some particular situations and to increase the information and accuracy of the output produced.

IV. VISION SUBSYSTEMS

In this section all the four vision subsystems are described in order to define their specific tasks. Output examples for each system are shown in Figures 5, 6, 7, and 8.

A. Trinocular

The sensing of obstacles up to a distance of 50 m is of paramount importance when driving in an urban environment since other vehicles traveling in either direction, or fixed obstacles may be on the vehicle's trajectory. Stereoscopic systems offer great advantages, since they reconstruct the 3D environment and therefore help to detect anything above the road surface. The innovative solution successfully designed and fielded in the 2005 DARPA Grand Challenge [5] is used again: an asymmetrical trinocular system forms 3 stereo systems with 3 different baselines. According to the needed depth of perception (linked to the vehicle speed) the software selects the pair of cameras to use. The two image streams are fed to the "trino" PC which detects the current vehicle pitch thanks to the use of the V-Disparity approach [6]; it then performs a 3D reconstruction of the scene and detects any kind of obstacle (defined as anything raising up from the road surface), finally providing this information to the World Perception Server.

The system merges raw LIDAR scan data with the filtered stereovision 3D data; the resulting set of values is used as a starting point for a flood-filling step on the disparity space image values, followed by a segmentation, to obtain the final obstacles. When LIDAR data is available the results become more dense in poorly textured areas and near the truck, where

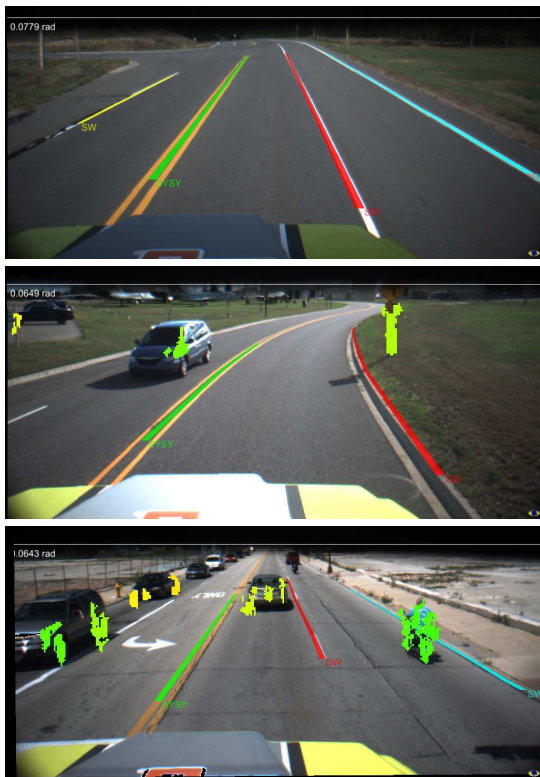


Fig. 5. Examples of Trinocular output: both lanes and obstacles are highlighted. Different colors for obstacles encode different distances, different colors for lanes encode different positions related to the truck. On the top left of the image also the instantaneous pitch of the vehicle and the estimated horizon line is shown.

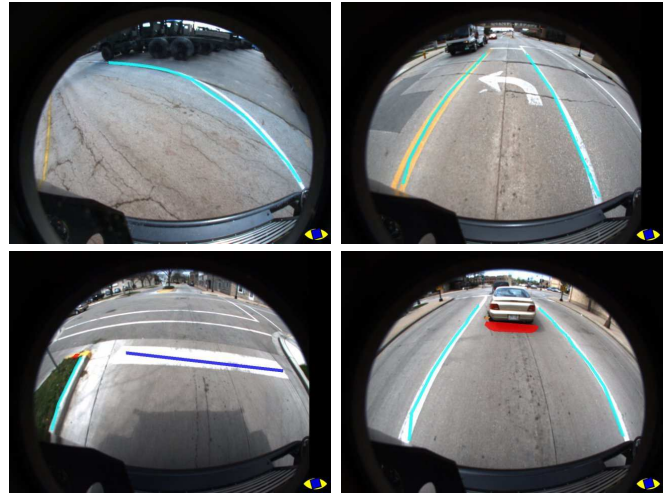


Fig. 6. Examples of Stereo output. From the top-left clockwise: sharp curves detection, lane detection, obstacle and lane detection, and stop line and lane detection.

the stereo matching is usually noisy, and many points are being rejected, as obstacles are seen under quite different angles by the cameras. Being mounted on top of the cab, the trinocular system is able to detect queued vehicles, like those found at crossings or parked along the road, making it easier for the path planner to obey traffic rules.

Along with obstacle detection, the trino system is also responsible of performing lane detection: the implemented algorithm exploits the available stereo information in order to lower the possibility of false positives, such as those introduced by poles or vehicle parts. The lane detector receives as input a full-resolution color image, where the areas corresponding to the obstacles detected in the previous steps are removed; since the detection of both white and yellow lines is necessary, two images are extracted from the color image: a grey-scale one and a yellow-highlighting one, and all undergo the same low level processing. In order to remove the perspective effect and obtain a bird's eye view of the scene, the images are remapped with an inverse perspective mapping transformation, computed taking into account the current pitch value, which is made available by the obstacle detector. The image is then processed searching for local horizontal luminance variations, typical of lane markings. Different line types are supported, namely single white, dashed white, single yellow, dashed yellow and double yellow. Tracking is used to improve outputs reliability and stability, and detection is possible even in presence of faded or worn lines.

B. Stereo

The navigation in an urban environment must be sharp and very precise. The trinocular system described in the previous section can only be used for driving at medium-to-high speeds, since it covers a (7-50 m) range. TerraMax includes two stereo systems (one in the front and one in the back) which provide a precise sensing in the close proximity

of the vehicle. The two stereo systems are identical: the 4 cameras are connected to the same PC which selects which system to use based on driving direction (forward or backward). Thanks to wide angle (fisheye, about 160°) lenses, these sensors gather information about an extended area of the immediate surroundings; the software is designed to improve the detection of the trinocular system: the stereo systems (based on software included into other intelligent vehicles prototypes [7]) are able to detect obstacles and lane markings with a high confidence on the detection probability and position accuracy, moreover the system is able to detect stop lines to allow a precise positioning of the truck in intersections and lane markings in sharp curves invisible to the trinocular system. The cameras setup selected for this system has the drawback of making it difficult to locate the exact boundaries of the detected items; to address this shortcoming, laser data is clustered and matched against the detected elements in the field of view, and when a correspondence is found the precise shape is sent to the World Perception Server. If no laser data is available, or the matching phase produces poor results, the system provides only the position of the obstacles, with no additional shape information.

C. Lateral

Lateral perception is important when facing an intersection. During a traffic merge manoeuvre, the vehicle is allowed



Fig. 7. Examples of Lateral output. A red box frames each moving object.



Fig. 8. Examples of Rear view output. In the upper row, images coming from the left side are shown, while those acquired on the right side are visible in the lower row. Overtaking vehicles are marked with a red line.

to pull into traffic only when oncoming traffic leaves a gap of at least 10 seconds. Hence, the vehicle needs to perceive the presence of oncoming traffic, as well as estimate cars speed from long distance. The vehicle needs to perceive the presence of oncoming traffic, regardless of the junction layout. In other words, the intersecting road might be seen from an angle, which in general may be different from 90° ; the lateral system must therefore be able to perceive traffic coming from different angular directions.

The system, described in [8], is equipped with two high resolution cameras, in order to achieve a working range of up to 100 meters yet maintaining a wide angular coverage, thus obeying DARPA specifications on traffic merge requirements. The use of such kind of cameras requires the development of algorithms capable of handling large amounts of data. Timing constraints are less strict than in other systems since the lateral one is active only when the vehicle is standing at crossroads. The computational load is further reduced by focusing the attention on a specific part of the image, while the rest is processed at a lower resolution, and merging the results of both elaborations before sending the output to the World Perception Server.

D. Rear view

When driving on a road with multiple lanes, the lane change manoeuvre is a common task in order to overtake slower traffic. Two lateral cameras installed on top of the driving cabin acquire images of the road behind and on the vehicle side. The cameras are installed rotated by 90° and frame the scene in portrait mode so that, despite being installed at 2.8m, they can frame the area close to the vehicle and extend their vision over the horizon. This system can overcome some LIDAR limitations, like its inability to provide useful results when moving along dusty roads, where

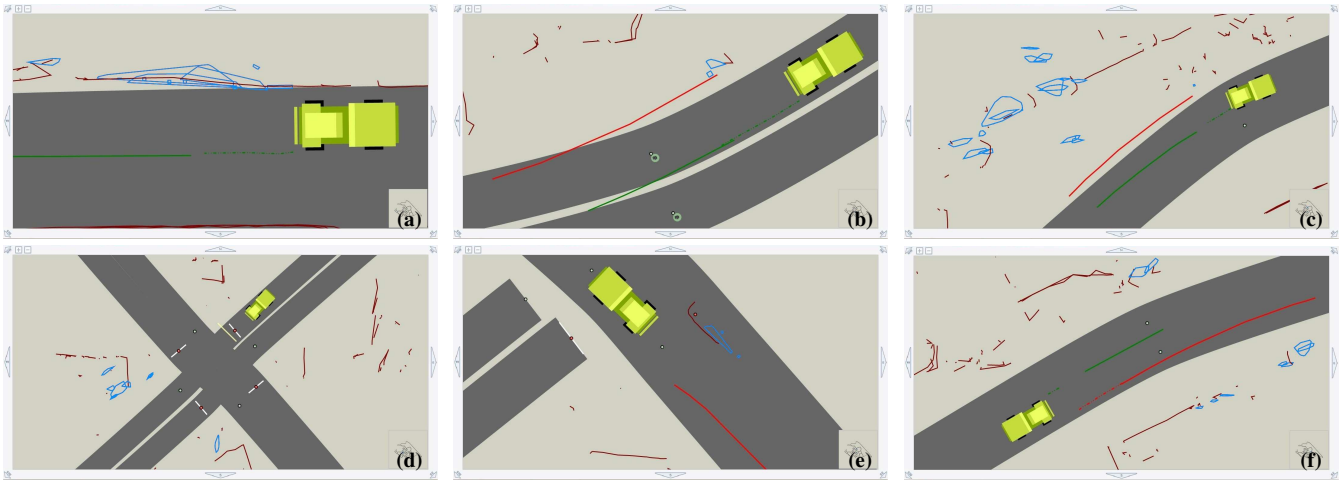


Fig. 9. All sensor data merged in the course visualizer. Dashed lines (red and green) represent Stereo lane updates, while solid lines (red and green) represent Trinocular lane updates as shown in figures *a*, *b*, *c*, *e*, and *f*. Detected stop lines are colored in off-white as in figure *d*; stop lines in bright white comes from the map provided by DARPA. Obstacles are shown with purple lines if they come from LIDAR and with blue polygons if they come from vision systems; a clear example is shown in figures *a* and *e*.

the clouds produced by the truck itself negatively affect the laser beams. This is a problem mainly of the rear LIDAR since dust clouds are produced by the vehicle in motion; sometimes it affects also the front ones especially during sudden stops when dust clouds overtake the truck. Despite this problem, LIDAR data has been used to refine the detected obstacles position measurement, thus reducing the errors introduced by vehicle pitching, since distance estimation is performed by a monocular system.

V. FINAL CONSIDERATIONS

The TerraMax autonomous vehicle features active and passive sensors, leveraging the strengths of both approaches, and overcoming their limits through ad-hoc fusion algorithms; some examples of the complete description of the truck surroundings are shown in Figure 9. This solution allows to correctly detect features hard to manage by the LIDAR, like short obstacles in the close proximity of the truck and reflective ones, lane detection or vehicles queued at crossings (thanks to the trinocular vision system); it also allows to cope with walls, side barriers and poorly textured objects in general, which are easily recognized by the LIDAR, but not by vision. The Urban Challenge experience has shown that vision is capable of covering a 360° area around the truck, providing the information needed for an effective path-planning. With no moving parts, reduced cost, and few integration requirements the developed system is suitable for use in different automotive applications, where a thorough understanding of the road environment is required.

Active technologies like high definition LIDARs still outperform vision in dense tridimensional reconstruction of the surroundings, and have been the main data source for many top teams, definitely becoming a success factor in the Urban Challenge final event. Despite these achievements, such sensors do not seem a viable solution for integration on commercial vehicles, given their mechanical characteristics

(considerable size and weight, heavy rotating parts), installation requirements (the sensing unit must be placed high on vehicle top to harness its full potential), and high production cost.

ACKNOWLEDGMENTS

The authors want to thank all Team Oshkosh for their precious support and for the endless effort put on the Urban Challenge project.

REFERENCES

- [1] D. Braid, A. Broggi, and G. Schmiedel, "The TerraMax Autonomous Vehicle," *Journal of Field Robotics*, vol. 23, no. 9, pp. 693–708, Sep. 2006.
- [2] J. McCall and M. Trivedi, "Video-based lane estimation and tracking for driver assistance: survey, system, and evaluation," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 7, no. 1, pp. 20–37, Mar. 2006.
- [3] S. Sehestedt, S. Kodagoda, A. Alempijevic, and G. Dissanayake, "Efficient lane detection and tracking in urban environments," in *3rd European Conference on Mobile Robots (EMCR 07)*, Freiburg, Germany, Sep. 2007.
- [4] S. Nedeveschi, F. Oniga, R. Danescu, T. Graf, and R. Schmidt, "Increased accuracy stereo approach for 3d lane detection," in *IEEE Intelligent Vehicles Symposium (IV2006)*, Tokyo, Japan, Jun. 2006, pp. 42–49.
- [5] A. Broggi, C. Caraffi, P. P. Porta, and P. Zani, "The Single Frame Stereo Vision System for Reliable Obstacle Detection used during the 2005 Darpa Grand Challenge on TerraMax," in *Procs. IEEE Intl. Conf. on Intelligent Transportation Systems 2006*, Toronto, Canada, Sep. 2006, pp. 745–752.
- [6] R. Labayrade and D. Aubert, "A Single Framework for Vehicle Roll, Pitch, Yaw Estimation and Obstacles Detection by Stereovision," in *Procs. IEEE Intelligent Vehicles Symposium 2003*, Columbus, USA, Jun. 2003, pp. 31–36.
- [7] A. Broggi, P. Medici, and P. P. Porta, "StereoBox: a Robust and Efficient Solution for Automotive Short Range Obstacle Detection," *EURASIP Journal on Embedded Systems – Special Issue on Embedded Systems for Intelligent Vehicles*, Jun. 2007, iSSN 1687-3955.
- [8] A. Broggi, A. Cappelunga, S. Cattani, and P. Zani, "Lateral Vehicles Detection Using Monocular High Resolution Cameras on TerraMax," in *Procs. IEEE Intelligent Vehicles Symposium 2008*, Eindhoven, Netherlands, Jun. 2008, in Press.