2014 IEEE 17th International Conference on
Intelligent Transportation Systems (ITSC)
October 8-11, 2014. Qingdao, China

Real-Time Obstacle Detection using Stereo Vision for Autonomous Ground Vehicles: A Survey

Nicola Bernini[1], Massimo Bertozzi[1], Luca Castangia[1], Marco Patander[1] and Mario Sabbatelli[1]

***Abstract*— One of the most important features for any intelligent ground vehicle is based on how is reliable and complete the perception of the environment and the capability to discriminate what an obstacle is. Obstacle Detection (OD) is one of the most widely discussed topics in literature. Many approaches have been presented for different application fields and scenarios; in last years most of them have been revisited using stereo vision or 2D/3D sensor technologies. In this paper we present a brief survey about Obstacle Detection techniques based on stereo vision for intelligent ground vehicles, describing and comparing the most interesting approaches. In order to provide a generic overview of these techniques, it has been decided to focus the study only on the algorithms that have provided a major contribution through real-time experiments in unsupervised scenarios.**

I. Introduction

Obstacle detection (OD) is one of the main control system components in autonomous vehicles [1] since a reliable perception of the real world is a key-feature for any obstacle detection system for dynamic environments. In last years, most of the historical approaches in literature have been readjusted in the framework of stereo vision and other 3D perception technologies (e.g. LIDAR) and important results have been provided by several experiments on autonomous ground vehicles (see Fig. 1). In order to achieve a good performance, most of the OD algorithms needs some assumptions about the ground [2] or about the approximated free space on it [3], [4], [5].

The obstacle detection field is a very broad one and a lot of obstacle detection systems have been developed in the last years in this domain [6]. An algorithm can be considered reliable and accurate if it provides: a real-time output, a stable and reliable tessellation of the environment, a robust state estimation of the obstacle detected and working regardless of lighting and weather conditions. Stereo vision errors and general performance have been widely discussed in literature [7],[8]; Matthies et al. [9] show a practical approach to evaluate the performance of an obstacle detection algorithm.

In this paper we present a brief survey on obstacle detection algorithms based on stereo vision and other 2D/3D sensors. Each obstacle detection system is focused on a specif tessellation or clustering strategy, hence they have been categorized into 4 main models:

1) *probabilistic occupancy map*

[1]N. Bernini, M. Bertozzi, L. Castangia, M. Patander and M. Sabbatelli are with VisLab-Dipartimento di Ingegneria dell'Informazione, Università degli Studi di Parma {bernini, bertozzi, castangia, patander, smario} @vislab.it

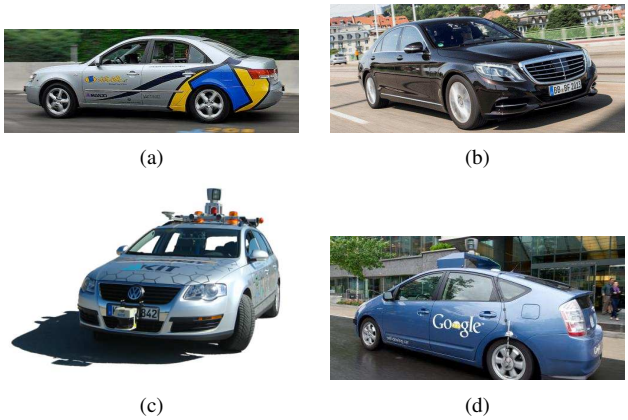


Fig. 1. Examples of fully autonomous ground vehicle: (a) *BRAiVE*- VisLab, University of Parma; (b) *Bertha* - Daimler; (c) *KITTI* - Karlsruhe Institute of Technology (KIT); (d) *Google Car* - Stanford Artificial Intelligence Laboratory (SAIL), Stanford University.

2) *digital elevation map*
3) *scene flow segmentation*
4) *geometry-based cluster*

II. Probabilistic Occupancy Map

The main model of the probabilistic occupancy map is proposed by Elfes [10]: *occupancy grid mapping*. It is one of the most famous approaches in literature.

The world is represented as a rigid grid of cells containing a random variable whose outcome can be free, occupied, or undefined (not mapped). For a proper formalization of the problem, let us consider a regular lattice D of X_i Random Variables with outcomes in a finite set of labels. The Elfes model requires 3 possible states: free, occupied, unknown. Let us call Ω the Phase Space and Z_t the measurements performed at t time.

The goal consists in computing $P(X, \{Z_\tau\}_{\tau=1,\dots,t})$ the associated joint distribution depending on a set of measurements carried out on a certain discrete set of time moments. Usually some assumptions are made in order to simplify the problem, namely *spatial conditional independence* and *temporal stochastic independence*.

In order to simplify the notation let us consider m as the random variable associated to a generic cell. The occupancy value of a cell m is determined using a probability density function given measurements z_t:

$$p(m|z_1, \dots, z_t) \tag{1}$$

978-1-4799-6077-4/14/$31.00 ©2014 IEEE

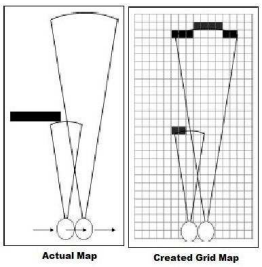


Fig. 2. Example of the generation of occupancy grid map

Equation 1 represents the state of the cell m given the measurements $z_1, \ldots, z_t$. Maps can defined over high-dimensional spaces. Assuming a 2D occupancy grid space and *static world*, namely the conditional independence among sensor reading given the knowledge of the map, the posterior density function in Equation 1 is reformulated in terms of log-odds as defined by Thrun [11]

$$p(m_{x,y}|z_t) = 1 - [e^{l^{(t)}_{x,y}}]^{-1} \tag{2}$$

with

$$l^{(t)}_{x,y} = l^{(t-1)}_{x,y} + \log \frac{p(m_{x,y}|z_t)}{1 - p(m_{x,y}|z_t)} - \log \frac{p(m_{x,y})}{1 - p(m_{x,y})}$$

and

$$l^{0}_{x,y} = \log \frac{p(m_{x,y})}{1 - p(m_{x,y})}$$

The log-odds regarding $p(m_{x,y}|z_t)$ in Equation 2 are recursively estimated through the Bayes rule, updating the cell value in different moments. More details are described in Thrun [11].

Fig. 2 shows the representation of a depth sensor measure in a 2D occupancy grid. Grey cells have unknown occupancy values, white cells are free and black cells are occupied. The main advantages of this method are the following ones: it is easy to construct and it can be as accurate as necessary.

A new set of stochastic occupancy grid models are detailed in Badino et al. [5]. Fig. 3 shows a representation of these probabilistic maps along with the corresponding projections in world coordinates. In this work the authors illustrate an innovative way to map the measurements computed by stereo vision. The disparity map represents the measurement processed by the algorithm to estimate an occupancy grid map. An estimation in world coordinates from the disparity map is implemented according to Equation 3 that is using a projection camera model based on the intrinsic and extrinsic parameters of the cameras.

$$p_k = P^{-1}(m_k) = \frac{baseline}{d} \cdot \begin{pmatrix} (u - u_0) \\ (v - v_0) \cdot \frac{fu}{fv} \\ fu \end{pmatrix} \tag{3}$$

where $m_k = (u, v, d)^T$ is a combination of (u, v) image coordinates and d the corresponding disparity computed by stereo, and $p_k = (x, y, z)^T$ the world point location of the m_k.

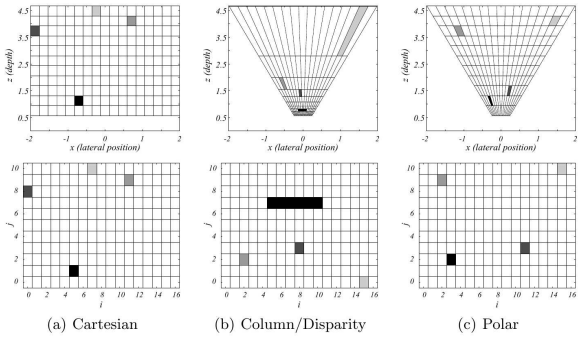
In the Badino's paper every cell of each grid maintains an *occupancy likelihood* $D(i, j)$ regarding the represented world area.

$$D(i, j) = \sum_{k=1}^{M} L_{ij}(m_k) \tag{4}$$

with M the number of measurements.

Equation 4 shows a definition of the function $D(i, j)$ where L_{ij} represents the occupancy likelihood for cell (i, j) given measurement m_k.

According to the Elfes model (see Equation 1) each occupancy likelihood function has been designed as a Gaussian probability density function G_{m_k}. A different function $L_{i,j}$ has been defined for each occupancy grid model.



(a) Cartesian (b) Column/Disparity (c) Polar

Fig. 3. Badino et al. occupancy grids. The figures on the top show how the world points are discretized into a top-view representation. The figures on the bottom show the corresponding occupancy grids. The cells have been marked with the same label of the corresponding world projections.

In [5], the authors present 3 occupancy grid maps to tessellate the measurements by stereo:

1) *Cartesian grid.* The world is represented by a cartesian grid and mapped linearly to a grid of fixed dimensions (see Fig. 3). Let us assume that cell (i, j) of the cartesian grid is centered at world coordinate (x_{ij}, z_{ij}). The likelihood function for cell (i, j) is represented in Equation 5.

$$L_{ij}(m_k) = G_{m_k}(P(p_{ij}) - m_k), \qquad p_{ij} = (x_{ij}, y, z_{ij})^T \tag{5}$$

For each point $p_{i,j}$, the y is the triangulated measurements height obtained with Equation 3.
The Gauss factor of the probability density function G_{m_k} is dependent on the difference between measurement and the reprojected cell position. Thus, the maximum likelihood factor is given to the cell which contains the triangulated measurement (see Fig. 4(a)). In a normal implementation the authors declare that updating every cell of the grid could be time consuming hence not suitable for a real-time application. They suggest to update only the cells significantly affected by the current measurement setting a proper distance threshold (e.g. Mahalanobis distance) between its projections.

2) *Column/Disparity grid*. The cells of the column/disparity grid correspond to discretized values of the u and d image coordinates. The occupancy grid criteria is based on mapping the measurements into (u, d) space assuming that a cell (i, j) corresponds to a coordinate $(u_{i,j}, d_{i,j})$ as shown in Fig. 3. In Equation 6 the likelihood function for the cell (i, j) is represented.

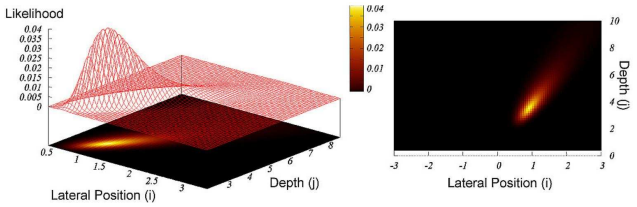
$$L_{ij}(m_k) = G_{m_k}((u_{ij} - u, 0, d_{ij} - d)^T) \quad (6)$$

Fig. 4(b) shows an example of the L_{ij} function. The disappearance of the v component in the measurement m_k regarding Equation 6 is due to the projection criteria onto the grid.

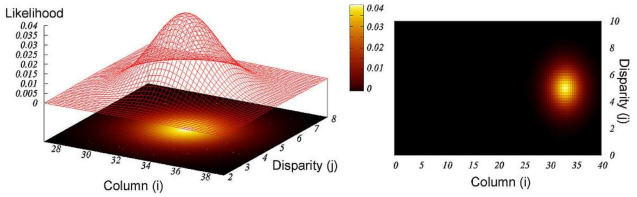
3) *Polar grid*. The mapping criteria of the polar occupancy grid is represented by the discretization of the values (u, z) where u corresponds to the column value in image space and z is the depth in the world coordinate system. The likelihood function for a generic cell (i, j) is detailed in Equation 7.

$$L_{ij}(m_k) = G_{m_k}((u_{ij} - u, 0, \frac{f_u \cdot baseline}{z_{ij}} - d)^T) \quad (7)$$

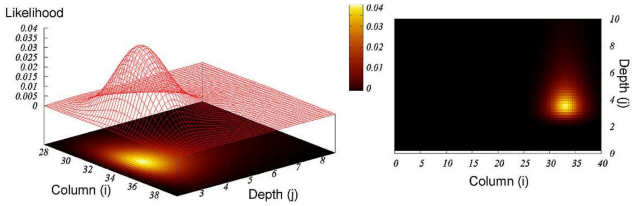
Fig. 4(c) shows an example of the L_{ij} function. As suggested by the authors, this solution overcomes the problem of the column/disparity approach: the decreasing resolution to distant points. The result can be easily evaluated comparing Fig. 3(b) and 3(c).



(a) Cartesian Occupancy Grid.



(b) Column/Disparity Occupancy Grid.



(c) Polar Occupancy Grid.

Fig. 4. Examples of the likelihood function Li, j related to the same measurement for each occupancy grid map.

When it is possible to make proper assumptions about the road model and the vehicle pose, successful solutions regarding the obstacle segmentation problem have been proposed, in some cases working in image coordinates by using v-disparity space [12].

Most of the recent obstacle detection algorithms have been developed in the occupancy grid maps framework. Recently one of the major contribution has been given by Badino et al. [13] with the *stixel* representation.

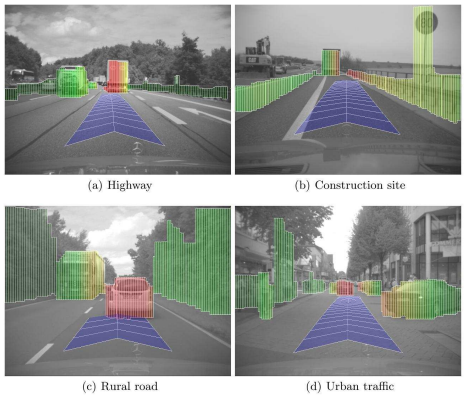
Stixel tessellation

The basic concept of this approach is the world representation into a set of rigid clusters called *stixels* by the authors, each obstacle is hence described as a union of these elements. A stixel based obstacle detector requires the following tasks to be performed:

1) a polar occupancy grid mapping of the measurements computed by stereo;
2) background and foreground of disparities using the previous polar grids;
3) height segmentation to estimate the heights of each stixels.

The last task is performed computing an upper boundary on the disparities cost image by means of dynamic programming. This part is detailed in Badino [13].

Real-time results are shown by the authors. They present a reliable obstacle detection algorithm that runs on an Intel Quad Core 3.00 GHz processor in 25ms. An evaluation of the stixel approach is shown in Fig. 5.



(a) Highway (b) Construction site

(c) Rural road (d) Urban traffic

Fig. 5. Evaluation of stixels in different real world road scenarios. The color encodes the lateral distance to the driving corridor.

III. Digital Elevation Map

A *Digital Elevation Map* (DEM) is a height-based representation of the measurements into a map like a cartesian occupancy grid (see Section II). This approach is widely applied mainly for terrain mapping [14]. A DEM can be computed with any 2D or 3D sensor (e.g. stereo vision, LiDAR and radar). Following the DEM-based approach, one of the major contributions for obstacle detection has been proposed by Oniga et al. [4]. The authors present a complete system for road surface estimation and obstacle detection in urban scenario.

In this work a DEM and two density maps are computed from the set of 3D points to obtain a compact representation with explicit connectivity between adjacent 3D locations. The road surface is fitted using a RANSAC approach to a small patch in front of the ego vehicle. To exploit this representation, the authors also propose an obstacle detection algorithm based on the density of 3D points per DEM cell (as a measure of the local slope). The density-based algorithm for obstacle detection is based on the density of 3D points: each DEM cell is classified as obstacle or road using a slope-based threshold criteria. Qualitative results are illustrated in Fig. 6.

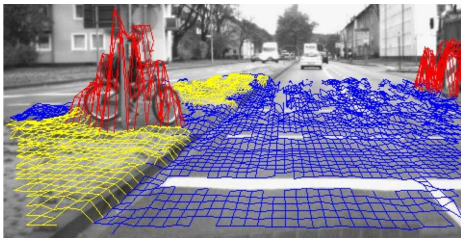


Fig. 6. An example of Oniga's DEM-based approach: road (blue), traffic isles (yellow) and obstacles (red).

The authors claim that, due to use of software-specific C optimizations and the DEM representation, an average processing time of 22ms has been achieved for the whole algorithm (on Pentium 4 at 2.6 GHz). Furthermore, with the image acquisition and the dense hardware reconstruction, a sustained processing frame rate of 23 frames/s has been obtained. From the performance point of view, a set of false positive and negative results is also presented.

An important contribution has been brought by Danescu et al. [15], consists in applying the particle filter strategy to perform DEM tracking. The authors define this approach *Dynamic DEM*.

IV. Scene Flow Segmentation

This technique, at first known as *optical flow*, is based on the temporal correlation to estimate the motion between two frames captured by camera at different times. In literature there are so many papers that show the implementation of an optical flow algorithm for obstacle detection but not many techniques guarantee a real-time processing [16], [17], [18]. Because of recent improvements in 3D reconstruction techniques by stereo, in the last years this approach has evolved into a new one: *scene flow estimation*.

A notable 3D approach is the so called 6D vision [19], where each 3D point (computed by an FPGA stereo system) is tracked by means of an efficient GPU Optical Flow implementation. An important study has been presented by Lenz et al. [20] where the scene flow estimation is computed by means of a temporal correlation regarding two couples of frames acquired by stereo, this is essentially a *visual odometry* based approach following Geiger [21]. Qualitative results for the Franke and Lenz approach are shown in corresponding Fig. 7 and 8.

Fig. 7. Scene flow estimation of 6D vision

Fig. 8. Output of Lenz's approach in urban scenarios: (a) Slowly moving and small objects such as the pedestrian in the first two frames are detected in a range up to 30 m. However, the track of such a small object is interrupted and it is not continuously tracked. Similarly moving groups of pedestrians are detected as one object since the scene flow difference is not unique and the number of detected interest points is to low. (b) Moving objects in this sequence are detected, but especially for the far range static objects are detected as well. (c) Turning cars and partly occluded objects, which where fully visible and observed in a previous frame, are detected.

In [20] the authors present an algorithm that runs in Matlab processing at least one frame per second on one core of an Intel Core2 Duo with 2.4 GHz and 4 GB RAM computing grayscale images with a resolution of 1392×512 pixels and at least 2000 interest points have been detected for each image. An hybrid approach regarding moving obstacles has been presented by Rabe et al. [22]. This work is also developed in the framework of 6D vision project. The core strategy in this application relies on the principle of fusing optical flow and stereo information given in [19]. The basic idea is hence to track points with depth information determined by stereo vision over two or more consecutive frames and to fuse the spatial and temporal information using Kalman Filters exploiting also an egomotion information. The algorithm is tested on a 3.2 GHz Pentium 4 computing 2000 image points with a cycle time of 40-80ms. Fig. 9 show an example of the approach described.

V. Geometry-based Cluster

In this generic category it has been decided to present only the solutions in literature that provided real-time results. The strategy that can best describe this category is Manduchi

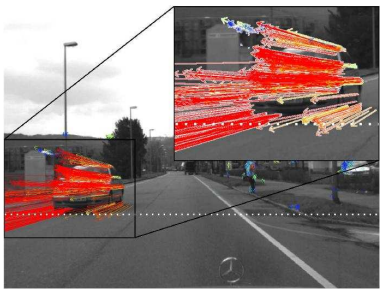


Fig. 9. Rabe's moving obstacle detector: velocity estimation results for an oncoming car

et al. [1]. In this work, the authors have postulated the first obstacle detection approach for any dimensional model (3D also). Indeed they designed an algorithm based on a search method that clusters the sensors measurements using a double cone model (see Fig. 10).

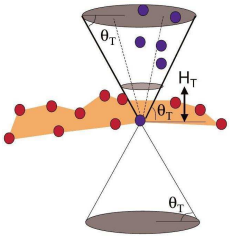


Fig. 10. 3-D obstacle search method using double cone that locates ground pixels (brown) and obstacle pixels (blue). The whole set of 3D points is projected on the frontal plane with respect to the camera (in order to simplify computations) then a scanning through the projected points set is performed. For each point, the double cone mask (projected it becomes a double triangle mask) is applied and if any other point is found in the region of interest they and the initial point are classified as an obtascle otherwise the initial point is classified as ground.

The performance of this approach are widely evaluated in [23].The authors described their stereo obstacle detection algorithm based on [1] showing the performance elaborated during an intercontinental experiment on their autonomous ground vehicle [24]. A similar approach was also used in PROUD test [25]. In order to achieve real-time processing, the authors have postulated a smart segmentation method along the disparities. The processing times are detailed in the following table:

PROCESSING TIMES

Algorithm step	Processing Time [ms]			
	Intel® Core™ i7 920		Intel® Core™ 2 Quad Q9100	
	SGM	SAD	SGM	SAD
Preproc.	2.9	2.9	4.9	5.0
DSI	21.5	5.8	60.2	7.8
DSI filt.	2.8	2.8	6.5	6.7
Obstacle det.	32.1	25.2	59.2	54.1
Total	59.3	36.7	130.8	73.6

Fig. 11 shows the qualitative results of the Broggi et at. [23].

The above approaches, however, do not exploit completely the 3D information provided by modern stereo matching algorithms and range finder devices. The complex cluster map shows full 3D information using adjacent stack of cells [26], or octrees connected cubes [27], and are able

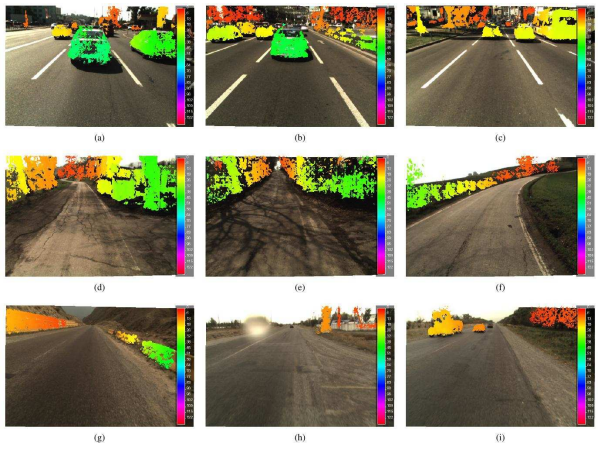


Fig. 11. Some sample outputs of Broggi et al. approach in different scenarios computed at 128 disparities. (a) - (c) a busy motorway in Kiev, (d) - (f) country roads with woods and uphill sections, (g) a deserted mountain motorway in Kazakhstan, (h) a raindrop on the right camera and (h) an upcoming tractor

to represent objects at multiple heights located at the same range and azimuth. Full 3D clusters show the great advantage of adequately represting obstacles with non conventional shapes, like *concave* ones. In [28] the authors designed an obstacles detector to generate a full 3D scene reconstruction to estimate both stationary and moving objects with minimum assumptions about the road, modeling the 3D point cloud, derived from a disparity image, to an accurate voxel reconstruction hence building complex clusters. These data structures contain the geometric and texture information to perform a segmentation following a flood fill approach. A vehicle pose estimation is carried out to determine the obstacles speed and position by means of an *egomotion* estimation, based on the *visual odometry* approach introduced in [21]. Through a temporal interpolation of previous 3D voxel reconstructions, the objects above the ground can be easily detected and their velocity and position can be estimated using a Kalman Filter. This algorithm has been for 640×480 pixel images on an Intel Core i7 at 10Hz (DSI=45ms, Voxel clustering and tracking=55ms). Some qualitative results are illustrated in Fig. 12.

VI. CONCLUSIONS

In this paper we presented a brief survey about the real-time approach for obstacle detection, mainly based on stereo vision. A classification of the obstacle detection methods is made in order to explain the different approaches presented in literature in the last years. Each work taken in consideration requires a good level of perception of the environment; dense disparity map and dense scene flow map are shown in works based on stereo.

This work is focused on the discrimination and selection of the OD techniques that gave a real contribution in terms of reliability, real-time processing capability and robustness. The presented approaches have proved effective but also showed some issues. DEM approach [4] and occupancy grid maps[13] are notably efficient but they are not able

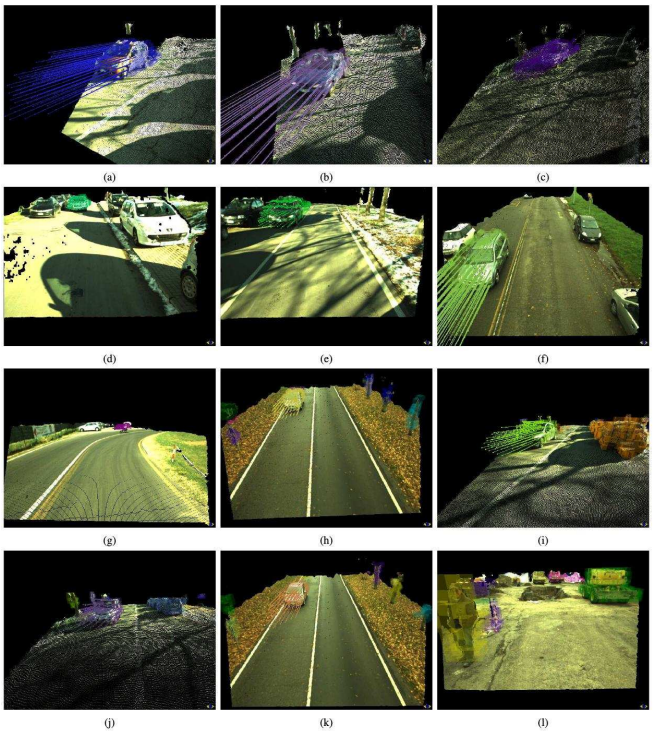


Fig. 12. Examples of voxel-based obstacle detection : (a)(b)(c)(d)(e)(f) moving cars without detection of stationary obstacles; (g) moving motorcycle without detection of stationary obstacles; (h)(i)(j)(k) moving cars and stationary obstacles; (l) stationary pedestrian and others obstacles in a construction site.

to represent concave or floating obstacles if another one is located at the same azimuth and position with different height. On the contrary, approaches based on geometry-based clusters [28] and scene flow segmentation [19], [20] provide a full 3D perception of the obstacles detected but are time-consuming. In terms of the obstacle state estimation each analyzed algorithm provides a reliable assessment combining a generic Kalman Filter and vehicle's egomotion information.

REFERENCES

[1] A. Talukder, R. Manduchi, A. Rankin, and L. Matthies, "Fast and reliable obstacle detection and segmentation for cross-country navigation," in *Intelligent Vehicle Symposium, 2002. IEEE*, vol. 2, June 2002, pp. 610–618 vol.2.

[2] Z. Zhang, R. Weiss, and A. R. Hanson, "Obstacle detection based on qualitative and quantitative 3d reconstruction," *IEEE Trans. on PAMI*, vol. 19, pp. 15–26, 1997.

[3] A. Broggi, E. Cardarelli, S. Cattani, and M. Sabbatelli, "Terrain mapping for off-road autonomous ground vehicles using rational b-spline surfaces and stereo vision," in *Intelligent Vehicles Symposium (IV), 2013 IEEE*, June 2013, pp. 648–653.

[4] F. Oniga and S. Nedevschi, "Processing dense stereo data using elevation maps: Road surface, traffic isle, and obstacle detection," *Vehicular Technology, IEEE Transactions on*, vol. 59, no. 3, pp. 1172 –1182, march 2010.

[5] H. Badino, U. Franke, and R. Mester, "Free space computation using stochastic occupancy grids and dynamic," in *Programming, Proc. Int'l Conf. Computer Vision, Workshop Dynamical Vision*, 2007.

[6] A. Discant, A. Rogozan, C. Rusu, and A. Bensrhair, "Sensors for obstacle detection - a survey," in *Electronics Technology, 30th International Spring Seminar on*, May 2007, pp. 100–105.

[7] H. P. Moravec, "Robot spatial perception by stereoscopic vision and 3d evidence grids," Tech. Rep., 1996.

[8] M. Felisa and P. Zani, "Incremental Disparity Space Image computation for automotive applications," in *Procs. IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems*, St.Louis, Missouri, USA, Oct. 2009.

[9] A. L. Rankin, A. Huertas, and L. H. Matthies, "Stereo vision based terrain mapping for off-road autonomous navigation," in *Proc. of SPIE, the International Society for Optical Engineering*, vol. 7332, 2009.

[10] A. Elfes, "Using occupancy grids for mobile robot perception and navigation," *Computer*, vol. 22, no. 6, pp. 46–57, June 1989.

[11] S. Thrun, "Learning occupancy grids with forward sensor models," *Autonomous Robots*, vol. 15, pp. 111–127, 2002.

[12] D. Pfeiffer and U. Franke, "Towards a global optimal multi-layer stixel representation of dense 3d data," in *Proceedings of the British Machine Vision Conference*. BMVA Press, 2011, pp. 51.1–51.12, http://dx.doi.org/10.5244/C.25.51.

[13] H. Badino, U. Franke, and D. Pfeiffer, "The stixel world - a compact medium level representation of the 3d-world," in *Proceedings of the 31st DAGM Symposium on Pattern Recognition*. Berlin, Heidelberg: Springer-Verlag, 2009, pp. 51–60.

[14] I. S. Kweon and T. Kanade, "High resolution terrain map from multiple sensor data," in *Intelligent Robots and Systems '90. 'Towards a New Frontier of Applications', Proceedings. IROS '90. IEEE International Workshop on*, Jul 1990, pp. 127–134 vol.1.

[15] R. Danescu and S. Nedevschi, "A particle-based solution for modeling and tracking dynamic digital elevation maps," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 15, no. 3, pp. 1002–1015, June 2014.

[16] W. Kruger, W. Enkelmann, and S. Rossle, "Real-time estimation and tracking of optical flow vectors for obstacle detection," in *Intelligent Vehicles '95 Symposium., Proceedings of the*, Sep 1995, pp. 304–309.

[17] C. Braillon, C. Pradalier, J. Crowley, and C. Laugier, "Real-time moving obstacle detection using optical flow models," in *Intelligent Vehicles Symposium, 2006 IEEE*, 2006, pp. 466–471.

[18] A. Wedel, A. Meißner, C. Rabe, U. Franke, and D. Cremers, "Detection and segmentation of independently moving objects from dense scene flow," in *Energy Minimization Methods in Computer Vision and Pattern Recognition*, ser. Lecture Notes in Computer Science, D. Cremers, Y. Boykov, A. Blake, and F. Schmidt, Eds. Springer Berlin Heidelberg, 2009, vol. 5681, pp. 14–27.

[19] U. Franke, C. Rabe, H. Badino, and S. K. Gehrig, "6D-Vision: Fusion of Stereo and Motion for Robust Environment Perception," in *Proceedings of the 27th DAGM Symposium*. Vienna, Austria: Springer, August 2005, pp. 216–223.

[20] P. Lenz, J. Ziegler, A. Geiger, and M. Roser, "Sparse scene flow segmentation for moving object detection in urban environments," in *Intelligent Vehicles Symposium (IV)*, 2011.

[21] A. Geiger, J. Ziegler, and C. Stiller, "Stereoscan: Dense 3d reconstruction in real-time," in *IEEE Intelligent Vehicles Symposium*, Baden-Baden, Germany, June 2011.

[22] C. Rabe, U. Franke, and S. Gehrig, "Fast Detection of Moving Objects in Complex Scenarios," in *Proceedings of the IEEE Intelligent Vehicles Symposium*, Istanbul, June 2007, pp. 398–403.

[23] A. Broggi, M. Buzzoni, M. Felisa, and P. Zani, "Stereo obstacle detection in challenging environments: the VIAC experience," in *Procs. IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems*, San Francisco, California, USA, Sept. 2011, pp. 1599–1604.

[24] M. Bertozzi, L. Bombini, A. Broggi, M. Buzzoni, E. Cardarelli, S. Cattani, P. Cerri, A. Coati, S. Debattisti, A. Falzoni, R. I. Fedriga, M. Felisa, L. Gatti, A. Giacomazzo, P. Grisleri, M. C. Laghi, L. Mazzei, P. Medici, M. Panciroli, P. P. Porta, P. Zani, and P. Versari, "VIAC: an Out of Ordinary Experiment," in *Procs. IEEE Intelligent Vehicles Symposium 2011*, Baden Baden, Germany, June 2011.

[25] A. Broggi, P. Cerri, S. Debattisti, M. C. Laghi, P. Medici, M. Panciroli, and A. Prioletti, "PROUD-public road urban driverless test: Architecture and result," in *Procs. IEEE Intelligent Vehicles Symposium 2014*, Dearborn (MI), United States, June 2014, pp. 648–654.

[26] H. Moracev, "Robot spatial perception by stereoscopic vision and 3D evidence grids," Technical Report CMU-RI-TR-96-34, CMU Robotics Institute, Tech. Rep., 1996.

[27] K. M. Wurm, A. Hornung, M. Bennewitz, C. Stachniss, and W. Burgard, "Octomap: A probabilistic, flexible, and compact 3d map representation for robotic systems," in *In Proc. of the ICRA 2010 workshop*.

[28] A. Broggi, S. Cattani, M. Patander, M. Sabbatelli, and P. Zani, "A full-3d voxel-based dynamic obstacle detection for urban scenario using stereo vision," in *Intelligent Transportation Systems - (ITSC), 2013 16th International IEEE Conference on*, Oct 2013, pp. 71–76.