

# LIDAR and Stereo Imagery Integration for Safe Navigation in Outdoor Settings

Giulio Reina  
Department of Engineering  
for Innovation  
University of Salento  
Via Arnesano  
73100 Lecce, Italy  
giulio.reina@unisalento.it

Annalisa Milella  
Institute of Intelligent  
Systems for Automation  
National Research Council  
via G. Amendola 122/D  
70126 Bari, Italy  
milella@ba.issia.cnr.it

Werner Halft and  
Rainer Worst  
Fraunhofer IAIS  
Schloss Birlinghoven  
53757 Sankt Augustin  
Germany.  
{werner.halft, rainer.worst}@iais.fraunhofer.de

**Abstract**—Environment awareness through advanced sensing systems is a major requirement for a mobile robot to operate safely, particularly when the environment is unstructured, as in an outdoor setting. In this paper, a multi-sensory approach is proposed for automatic traversable ground detection using 3D range sensors. Specifically, two classifiers are presented, one based on laser data and one based on stereovision. Both classifiers rely on a self-learning scheme to detect the general class of ground and feature two main stages: an adaptive training stage and a classification stage. In the training stage, the classifier learns to associate geometric appearance of 3D data with class labels. Then, it makes predictions based on past observations. The output obtained from the single-sensor classifiers is statistically combined exploiting their individual advantages in order to reach an overall better performance than could be achieved by using each of them separately. Experimental results, obtained with a test bed platform operating in a rural environment, are presented to validate this approach, showing its effectiveness for autonomous safe navigation.

## I. INTRODUCTION

Safe autonomous driving in outdoor environments relies on the ability to distinguish drivable terrain from obstacles, including man-made artifacts, ruts, cliffs, trees, bushes, shrubs, and other vegetation that can obstruct or endanger the robot's motion. In the last few years, a large body of research has been devoted to address this issue through the use of different sensor modalities. Many researchers rely on stereovision to generate a 3D point cloud at relatively high frequency, [1], [2], [3], [4]. By applying geometric and statistical heuristics, the terrain surface and obstacles can be classified. However, the resulting map may present "holes" in some parts due to insufficient texture or noisy images, or it may be affected by lighting conditions. There are also many researchers using LIDAR sensors to detect the terrain surface for robot navigation, [5], [6]. In general, LIDAR sensors can return dense 3D point clouds; however, scanning LIDAR sensors often operate at a relatively low frequency (1 Hz or less) resulting in difficulties to capture dynamic obstacles. Expensive LIDARs like the Velodyne HDL-64E are an exception ([www.velodyne.com/lidar](http://www.velodyne.com/lidar)). Fixed (non-scanning) LIDAR sensors can operate at higher frequencies; however, they have to rely on the robot's self-localization system, whose accuracy generally decreases on rough terrains. Due to the mentioned limitations, various methods have been

proposed to combine vision and LIDAR into one system and mitigate the drawbacks of single sensor-based approaches. For example, in [7], an obstacle detection algorithm that can be used with both LIDAR or stereo with small modifications is proposed. New interest in this field has been generated by the application of probabilistic learning techniques and self-learning classification methods [8]. For instance, a self-supervised approach is proposed in [9], where monocular vision is combined with a laser range finder. Specifically, the laser is used to scan for a flat drivable surface area in the vicinity of the vehicle. Once identified, this area is used as training data for the computer vision algorithm. A similar scheme was applied in [10] to the specific case of a forested environment. Other sensors, including a radar [11], [12] and a stereo camera [3] have been also used as the supervising sensor to automatically train a ground classifier.

In this paper, a novel approach for terrain analysis is presented, which combines LIDAR and stereovision within a statistical framework. An adaptive self-learning scheme is proposed, whose basic principles can be applied to any 3D range sensor modality to develop single-sensor ground classifiers. Here, it is demonstrated for LIDAR and stereo data although they differ in resolution, accuracy and field of view. The LIDAR-based and stereo-based classifier comprises two main stages: a training stage and a classification stage. During the training phase, each classifier automatically learns to associate the geometric appearance of data with class labels. Then, it makes predictions based on past observations classifying new observations in two general categories, i.e. ground and non-ground. The ground class corresponds to observations from traversable terrain, whereas the non-ground class corresponds to all other data, including points from above ground objects (i.e., obstacles) or from occluded areas, and poor sensor reconstructions. Since the characteristics of the ground may change geographically and over time, each classifier is continuously retrained in every scan: new automatically labeled data are added to the ground model replacing the oldest labels in order to incorporate changes in the ground geometric appearance. Afterwards, the output obtained from the two single-sensor classifiers is statistically combined to improve the overall perception ability. It should be noted that most of the algorithms proposed in the literature generally rely on ground plane estimation, [13],

[14], and the obstacle detection task aims to identify objects that “stick out” of the ground. However, this assumption is of limited validity in outdoor scenarios. In this work, ground plane reasoning is not explicitly required, and the system automatically adapts to the changing geometry of the terrain. Furthermore, the proposed approach aims to detect scene regions that are traversable-safe for the vehicle rather than attempting to explicitly identify obstacles, [1], [15]. This is a subtle, but significant difference; only those regions where there is evidence that it is safe are labeled as traversable, thereby avoiding both positive and negative obstacles without explicitly detecting them. An additional advantage of the proposed obstacle detection scheme is that the output traversability map can be directly employed by grid-based planners [16].

### A. Experimental setup

This research was developed within the project Ambient Awareness for Autonomous Agricultural Vehicles (QUAD-AV) funded by the ERA-NET ICT-AGRI action and aimed to enable safe autonomous navigation in high-vegetated, off-road terrain. The proposed system was integrated with an off-road vehicle (see Fig. 1) that was made available by the partner National Research Institute of Science and Technology for Environment and Agriculture (IRSTEA) at the Montoldre farm facility, during an experimental campaign in October 2012. The test bed was equipped with a Point Grey Bumblebee XB3 camera and a scanning LIDAR SICK LMS111. The vision unit features two stereo configurations: a narrow stereo pair with a baseline of 0.12 m using the right and middle cameras, and a wide stereo pair with a baseline of 0.24 m using the left and right cameras. By employing the narrow baseline to reconstruct nearby points and the wide baseline to reconstruct more distant points, this stereo device takes the advantage of the small minimum range of the short baseline, while preserving the higher accuracy at each visible distance of the wide baseline [3]. Additional technical details of the stereo system are collected in Table I. The LIDAR generates single line scans covering an angle of 270 deg with an angular resolution of 0.5 deg at a rate of 50 Hz. It is mounted on a servo-controlled rotating stage that sweeps the scan plane through 360 deg around the longitudinal axis of the robot, generating a complete 3D point cloud with a size of 80,000



Fig. 1. The experimental platform used in this research provided by the partner IRSTEA and used for field validation in the QUAD-AV project.

TABLE I. TECHNICAL DETAILS OF THE STEREOVISION SYSTEM.

<b>Camera</b>	Trinocular
<b>Model</b>	Bumblebee XB3 Baseline: 12 cm and 24 cm
<b>Image size (pixels)</b>	1280 × 960 pixels at 15 FPS
<b>Field of view</b>	66deg(H) × 50deg(V)
<b>Optics</b>	focal length: 3.8 mm f2.0
<b>Range</b>	2 to 22 m

TABLE II. SPECIFICATIONS OF THE LIDAR SCANNER.

<b>LIDAR</b>	Rotating rangefinder
<b>Model</b>	3DSL with SICK LMS111
<b>Number of points</b>	~ 80,000
<b>Field of view</b>	360deg(H) × 270deg(V)
<b>Range</b>	0.5 to 17 m

points in about 3 s. Multiple scans obtained during the robot’s motion are aligned and integrated using ICP-based algorithms [17], to generate a dense 3D point cloud in a fixed reference frame. The salient technical details of the LIDAR system are collected in Table II.

The remainder of the paper is organized as follows. Section II describes the proposed self-learning classification scheme. In Section III, the system is validated in field tests performed with the test platform. Finally, Section IV concludes this paper.

## II. SYSTEM DESCRIPTION

A self-learning approach is proposed to detect traversable ground using 3D range data. The term “self-learning” refers to automatic training of a ground classifier. If in a traditional (i.e. manually) supervised classifier a user provides labeled examples to train each class of interest, in a self-learning scheme these training instances are automatically produced. In the context of this research, the use of a rolling training set is proposed. Initially, the robot has no knowledge of the ground class appearance. The training set is initialized at the beginning of the robot’s operation via a bootstrapping approach and progressively updated. The only underlying assumption is that the vehicle starts its operations from a clear (free of obstacles) area, so that each sensor initially “looks” at ground only. During this stage, features can be extracted from the sensor-generated 3D point cloud and they can be associated with the ground class. When sufficient data is accumulated, the ground classifier can be trained, and the ground labels are related with 3D point cloud properties. The task is that of generalizing from training data to unseen situations to identify single new observations as ground or non-ground. This allows the system to predict the presence of ground in successive scenes based on past observations. In order to account for ground changes, the model (i.e., the training set) is continuously updated using the most recent sensor readings. In every new-acquired scan, the latest training set is used to train the classifier. It is important to note that the proposed classification scheme can be used to characterize the scene structure obtained from any 3D range sensor. In this research, it is applied to develop two classifiers: one based on LIDAR data and one based on stereo data. Then, the single-sensor classifiers are statistically combined to obtain a unique classification result.

In the rest of this section, first the ground modeling and classification approach common to both the LIDAR and stereo classifier is described. Then, the classifier fusion scheme is

introduced.

### A. Ground modeling and classification

The goal is to classify a given terrain patch as being traversable or not. This problem can be formulated as a one-class classification [18]. In general, one-class classification methods are useful in two-class classification problems, where one class (the target class) is relatively well-sampled, while the other class (the outlier class) is relatively under-sampled or it is difficult to model. This is the case for our outdoor application, where most of the patches belong to the ground with sparse instances of non-ground. In addition, the variation of all possible non-ground classes is unlimited. That makes it difficult to model the non-ground class, whereas, although it changes geographically and over time, the ground class is generally less variable than random objects. To model the ground class, a feature-based representation using geometric features extracted by 3D range data is adopted. The ground feature set is then fitted with a multivariate Gaussian distribution and a Mahalanobis Distance (MhD)-based classification approach is adopted to recognize whether a new pattern is an instance of the ground class following an outlier detection strategy.

1) *Geometric features*: The appearance of ground is constructed upon a set of geometric features that can be extracted from 3D scene reconstruction. First, the point cloud generated by the sensor is divided into a grid of 0.4-m by 0.4-m terrain patches projected onto a horizontal plane. Geometric features are statistics obtained from the point coordinates associated with each terrain patch. This geometric information is used in the self-learning classification scheme to label all cells as traversable or non-traversable. Therefore, positive and negative obstacles, as well as, unknown regions can be implicitly detected and avoided. The first element of the geometric feature vector is the average slope of the terrain patch, i.e. the angle  $\theta$  between the least-squares-fit plane and the horizontal plane. The second component is the goodness of fit,  $E$ , measured as the mean-squared deviation of the points from the least-squares plane along its normal. This is the same as the minimum singular value of the points' covariance matrix. The third element is the variance in the height of the range data points with respect to the reference plane,  $\sigma_h^2$ . The fourth component is the mean of the height of the range data points,  $\bar{h}$ . Thus, the geometric properties of each patch is represented by a 4-element vector  $x = [\theta, E, \sigma_h^2, \bar{h}]$ .

2) *Ground class model*: Let  $X_t$  be an  $n \times m$  data table representing a sample of  $x_i$  feature vectors with  $i = 1, 2, \dots, n$ , each characterized by  $m$  traits ( $m = 4$ , in our case):  $X_t = \{x_1, \dots, x_n\}$ . These vectors constitute the training set at a given time  $t$ . If we compute the sample mean  $\mu_t$  and the sample covariance  $\Sigma_t$  of the data in  $X_t$ , we can denote the ground model at this time as  $M_t(\mu_t, \Sigma_t)$ , where  $\mu_t$  describes the location, and  $\Sigma_t$  the scale (shape) of the distribution. Then, in the next sensor scan acquired at time  $t + 1$ , the single new observation  $x_{new}$  can be classified by estimating its squared Mahalanobis distance from the ground model:

$$d^2 = (x_{new} - \mu_t)\Sigma_t^{-1}(x_{new} - \mu_t)^T \quad (1)$$

Assuming that the vectors  $x_i$  are independent and have Gaussian distribution, [3], it can be proved that the squared Mahalanobis distance is asymptotically distributed as the  $m$  degrees of freedom chi-square distribution  $\chi_m^2$ . Then, we can use the quantile  $\beta$  of the  $m$  degrees of freedom chi-square distribution as the delimiter (cutoff) for outlying observations, i.e.

$$L_\beta = \sqrt{\chi_{m;\beta}^2} \quad (2)$$

Any observation with Mahalanobis distance  $d$  satisfying the inequality  $d \geq L_\beta$  may be suspected to be an outlier.

It is worth to note that the ground model is continuously updated as the vehicle moves, using a rolling training window strategy: new ground feature vectors labeled in the most recent radar scans are incorporated, replacing an equal number of the oldest ground instances, thus keeping constant the size of the rolling window.

### B. Classifier fusion

In order to exploit the individual advantages of vision and LIDAR sensors and to reach an overall better performance, the single-sensor ground classifiers are combined. Combining classifiers aims at exploiting the complementary information residing in the single classifiers. Assume that we are given a set of classifiers, which have already been trained to provide as output the class *a posteriori* likelihood in the form of the Mahalanobis distance from the class center. For a one-class classification task, given an unknown observation  $x$ , the classifier  $i$  produces estimates of the *a posteriori* class likelihood, that is  $M_i(x)$ . In our case  $i = L, S$ , where  $L$  stands for LIDAR-based classifier and  $S$  for stereo-based classifier. Our goal is to devise a way to come up with an improved estimate of a final *a posteriori* likelihood  $M(x)$  based on all the resulting estimates from the individual classifiers. One way is to weight the individual output obtained from the classifiers with their prior probabilities that can be statistically quantified using ground-truth data. This analysis would provide various statistical quantities including for instance a confusion matrix. By appropriately normalizing true positives (TP), false positives (FP), true negatives (TN), false negatives (FN), obtained from the confusion matrix, we can construct (empirical) expected rates of positive predictive value or precision (P) and negative predictive value or rejection precision (RP),

$$P = \frac{TP}{TP + FP} \quad (3)$$

$$RP = \frac{TN}{TN + FN} \quad (4)$$

Being normalized, these rates are also probabilities. So now, we have values of uncertainty associated with the LIDAR and stereo only ground detection in the form of prior probabilities over these detections. These (prior) probabilities can be used as weights to combine statistically the decision of each classifier through a weighted sum and obtain a unique classification result,

$$M(x) = \sum_{i=L}^S \frac{W_i M_i(x)}{W_L + W_S} \quad (5)$$

where the weight,  $W_i$ , is equal to  $P_i$  or  $RP_i$  if the observation  $x$  is labeled as ground or non-ground, respectively, by

the classifier  $i$ . As a result, the uncertainty associated with each classifier is propagated throughout the final classification result.

### III. EXPERIMENTAL RESULTS

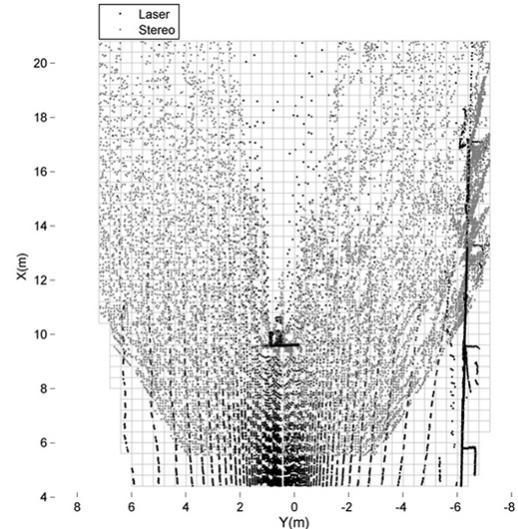
In this section, experimental results are described to validate the proposed approach for terrain analysis using LIDAR-stereo combination. The system was integrated with an experimental platform (see Fig. 1) and tested in rural settings. Various scenes were analyzed including positive obstacles (high-vegetated areas, trees, buildings, agricultural equipment), negative obstacles (ditches and other depressions), moving obstacles (vehicles, humans and animals), and highly-irregular terrain. In all experiments, the vehicle was driven by a human operator with a travel speed between 10 and 20 km/h, as the onboard sensors acquired data from the surrounding environment. Then, the proposed classification framework was applied offline. For each data set, the vehicle started its operations from an area that was clear of obstacles in order to initialize the ground model for both the LIDAR-based and stereo-based classifier. Few scans ( $s_0 = 3$ , in our case) were necessary to complete the bootstrap phase, requiring a short time interval (e.g., a 10 s window if a sampling rate of 0.3 Hz was used for the LIDAR scanner). After the system was initialized, the single-sensor classifiers were able to predict the presence of ground in successive acquisitions.

A subset of salient images taken from different data sets was hand-labeled to identify the ground-truth terrain class corresponding to each cell. In this manner, it was possible to provide a quantitative evaluation of the prior classification probabilities of the LIDAR and stereo-based classifier. Specifically, the precision, i.e. the number of true positives divided by the total number of elements labeled as belonging to the ground class, and the rejection precision, i.e. the number of true negatives divided by the total number of elements labeled as belonging to the non-ground class, were evaluated, as explained in Section II-B. A significance level of 5% (i.e.,  $\beta = 0.95$ ) was assumed for the cutoff threshold expressed by (2), in both the LIDAR-based and stereo-based classifiers. Based on the knowledge of the prior classification performance, it was possible to fuse the two single-sensor classifiers using (5).

In Fig. 2(a), a typical scene is shown acquired during the field testing of the system. Figure 2(b) shows the upper view of the 3D reconstruction as obtained by laser scanner (marked by black points) and stereovision (denoted by grey points). One first advantage of combining LIDAR with stereovision is that the overall field of view is increased with LIDAR and vision providing data mostly in the short and long range, respectively. In addition, the sparseness of LIDAR data is mitigated by dense stereo reconstruction. In Fig. 3, the results obtained from the LIDAR-based classifier are shown for the scene of Fig. 2(a) in terms of traversability map. Cells labeled as ground are marked in green, whereas cells that are labeled as non-ground are denoted in red. In Fig. 3(b), the same results are projected over the co-located image for visualization and comparison purposes. Pixels associated with ground and non-ground cells are marked using green and red, respectively. The LIDAR-based classifier correctly detected both the obstacle in front of the vehicle and the building on the right side, as well as the traversable ground. For this scene, the precision and rejection precision of the LIDAR-based classifier were 99.0%



(a)



(b)

Fig. 2. (a) Original visual image. (b) Reference grid divided into 0.4-m by 0.4-m cells. Points reconstructed by laser scanner are denoted in black. Points reconstructed by stereovision are marked in grey.

and 80.1%, respectively. Fig. 4 shows the results obtained from the stereo-based classifier applied to the running example of Fig. 2(a). Again, cells labeled as ground are marked in green, whereas cells that are labeled as non-ground are denoted in red. Classification results are projected over the original image as shown in Fig. 4(b), where pixels associated with ground and non-ground cells are marked using green and red, respectively. The stereo-based classifier was also correct in labeling obstacles and traversable ground. The precision and rejection precision of the stereo-based classifier resulted in 98.9% and 66.0%, respectively. When the two single-sensors classifiers are combined by weighting their results with the associated prior probability, the performance of the overall system resulted in a precision and a rejection precision of 99.3% and 76.5%, respectively. The traversability map obtained from the combined classifier is shown in Fig. 5. At a first glance, it may seem that the combined classifier performs worse than the LIDAR classifier in terms of rejection precision. However, it should be noted that the single-sensor and the combined classifiers differ in field of view, thus a direct comparison of their classification performance is possible only when considering the cells that are “seen” simultaneously by both sensors. As a matter of fact, if a given cell is being observed only by one sensor modality, it will be labeled by the single-sensor

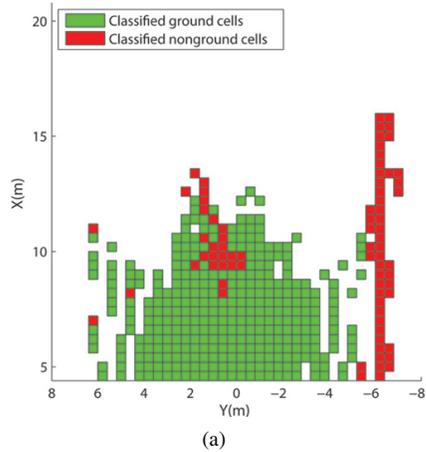


Fig. 3. (a) Traversability map obtained by the LIDAR-based ground classifier. (b) Results projected over the original image. Pixels associated with ground-(non ground-) labeled cells are marked using green (red).

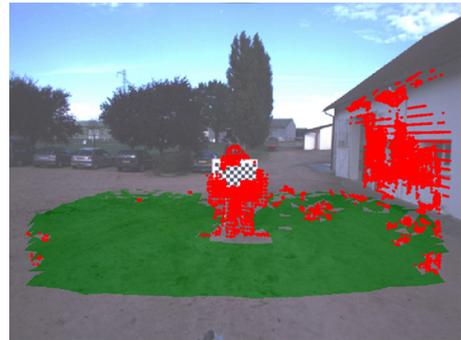
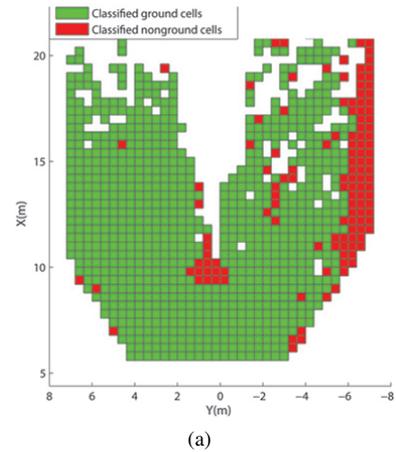


Fig. 4. (a) Traversability map obtained by the stereo-based ground classifier. (b) Results projected over the original image. Pixels associated with ground-(non-ground-) labeled cells are marked using green (red).

classifier and no sensor integration will be possible. Table III collects the results obtained from each classifier over the common cell subset that amounts to about the 35% of the total number of labeled cells. These results clearly demonstrate that the statistical fusion of the two classifiers helps in increasing the information content and accuracy of the output produced.

Figure 6 shows a different scenario where the vehicle travels on a country asphalt road. While the LIDAR-based module performs well, as demonstrated in Fig. 6(a)-(b), the stereo-based classifier produces many false negatives due to the presence of heavy shadowing on the road, as shown in Fig. 6(c)-(d). The traversability map, obtained fusing the two classifiers, is shown in Fig. 6(e). For this scene, the LIDAR-based classifier provided a precision and rejection precision of 99.0% and 98.3%, respectively, against values of 97.1% and 49.1%, respectively, for the stereo-based classifier. When the two systems are fused, P and RP resulted in 98.2% and 75.1%, respectively. Again, the combination of the two sensor modalities allowed the overall field of view to be increased,

TABLE III. CLASSIFICATION RESULTS OBTAINED FROM THE SINGLE-SENSOR AND COMBINED CLASSIFIERS OVER THE COMMON CELL SUBSET, I.E. THE CELLS THAT ARE LABELED SIMULTANEOUSLY BY BOTH SENSORS.

	LIDAR-based	Stereo-based	Combined
Precision	99.1%	96.5%	99.6%
Rejection Precision	87.5%	96.1%	98.3%

and the low rejection precision provided by the vision to be compensated while preserving at the same time high precision in detecting ground examples.

In summary, the combination of stereo and LIDAR data is useful in that 1) it allows to widen the overall field of view of the perception system, 2) vision can help to overcome limitations of LIDAR, such as sparseness of data and low acquisition frequency, by producing dense maps at relatively high frequency, 3) being not affected by lighting conditions,

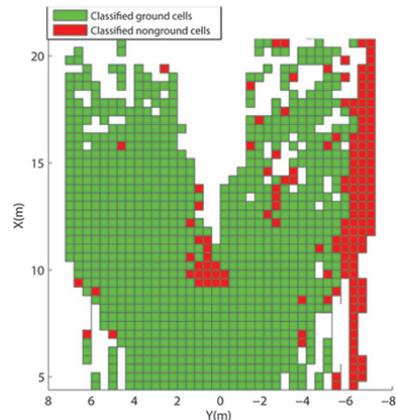


Fig. 5. Traversability map obtained by the combined LIDAR-stereo system for the scene of Fig. 2(a).

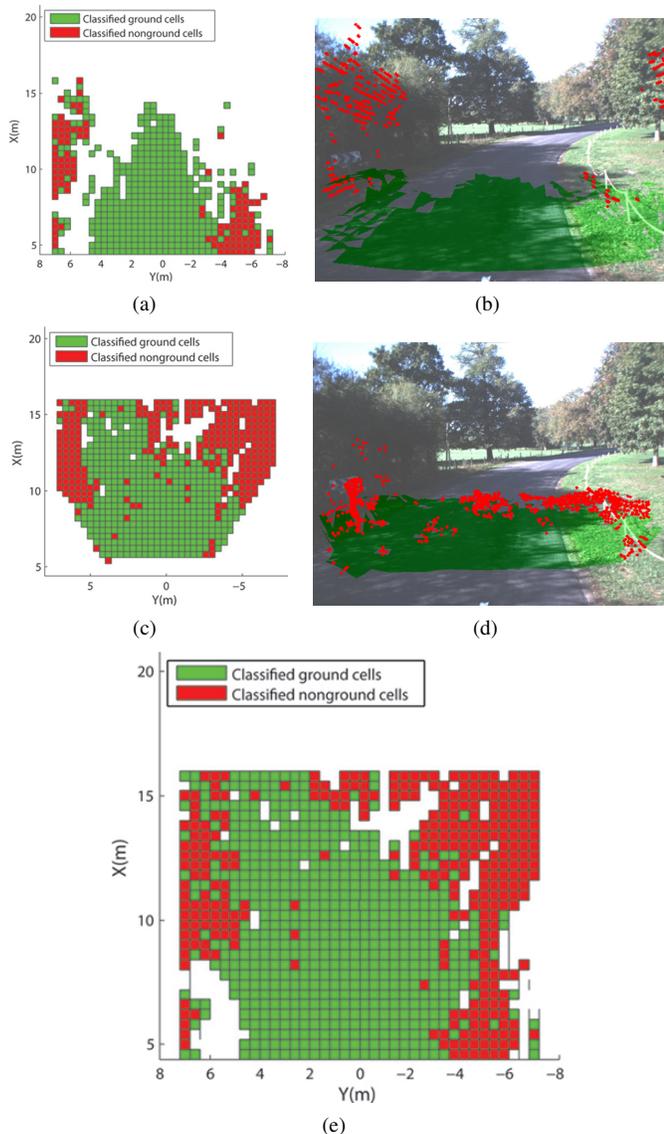


Fig. 6. (a) Traversability map obtained by the LIDAR-based ground classifier, (c) and the stereo-based ground classifier. (b, d) Results projected over the original visual image. (e) Traversability map obtained by the combined LIDAR/stereo system. Ground-labeled cells and associated pixels are marked using green. Non-ground labeled cells and associated pixels are denoted in red.

LIDAR can help to overcome limitations of vision, such as reconstruction errors due to poor lighting conditions, shadows and lack of texture.

#### IV. CONCLUSIONS

A multi-sensory approach combining LIDAR and stereo imagery within a statistical self-learning framework for traversable ground detection was presented. Field experiments obtained using a test platform in natural scenarios validated this approach showing good classification performance. The proposed system led to the following main advantages: (a) improvement of the perception performance of the combined LIDAR/stereo system due to complementarity of the two sensor modalities, (b) self-learning training of the system, where the sensors allow the vehicle to automatically acquire a set

of ground samples, eliminating the need for time-consuming manual labeling, (c) continuous updating of the system during the vehicle's operation, thus making it adaptive and feasible for long range and long duration navigation applications.

#### ACKNOWLEDGMENT

This work was supported by the ERA-NET ICT-AGRI through the grant Ambient Awareness for Autonomous Agricultural Vehicles (QUAD-AV).

#### REFERENCES

- [1] R. Manduchi, A. Castano, A. Talukder, and L. Matthies, "Obstacle detection and terrain classification for autonomous off-road navigation," *Autonomous Robot*, vol. 18, pp. 81–102, 2004.
- [2] A. Broggi, A. Cappalunga, C. Caraffi, S. Cattani, S. Ghidoni, P. G. P., P. Porta, M. Posterli, and P. Zani, "Terramax vision at the urban challenge 2007," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11(1), pp. 194–205, 2010.
- [3] G. Reina and A. Milella, "Towards autonomous agriculture: automatic ground detection using trinocular stereovision," *Sensors*, vol. 60(11), pp. 12 405–12 423, 2012.
- [4] A. Milella, G. Reina, and R. Siegwart, "Computer vision methods for improved mobile robot state estimation in challenging terrains," *Journal of Multimedia*, vol. 1(7), pp. 49–61, 2006.
- [5] C. Wellington and A. Stentz, "Online adaptive rough-terrain navigation in vegetation," in *International Conference on Robotics and Automation*, 2004, pp. 96–101.
- [6] J. Lalonde, N. Vandapel, D. Huber, and M. Hebert, "Natural terrain classification using three-dimensional lidar data for ground robot mobility," *J. of Field Robotics*, vol. 23(10), pp. 839–861, 2006.
- [7] S. Kuthirummal, A. Das, and S. Samarasakera, "A graph traversal based algorithm for obstacle detection using lidar or stereo," in *IEEE/RSJ Inter. Conf. on Intelligent Robots and Systems*, 2011, pp. 3874–3880.
- [8] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. Springer, 2003.
- [9] H. Dahlkamp, D. S. A. Kaehler, S. Thrun, and G. Bradski, "Self-supervised monocular road detection in desert terrain," in *Robotics Science and Systems Conference*, 2006.
- [10] S. Zhou, J. Xi, M. W. McDaniel, T. Nishihata, P. Salesses, and K. Iagnemma, "Self-supervised learning to visually detect terrain surfaces for autonomous robots operating in forested terrain," *J. of Field Robotics*, vol. 29(2), pp. 277–297, 2012.
- [11] A. Milella, G. Reina, J. Underwood, and B. Douillard, "Combining radar and vision for self-supervised ground segmentation in outdoor environments," in *IEEE/RSJ Inter. Conf. on Intelligent Robots and Systems (IROS)*, 2011, pp. 255–260.
- [12] G. Reina, A. Milella, and J. Underwood, "Self-learning classification of radar features for scene understanding," *Robotics and Autonomous Systems*, vol. 60(11), pp. 1377–1388, 2012.
- [13] R. Hadsell, P. Sermanet, J. Ben, A. Erkan, M. Scoffie, and K. Kavukcuoglu, "Learning long-range vision for autonomous off-road driving," *J. of Field Robotics*, vol. 26(2), pp. 120–144, 2009.
- [14] K. Konolige, M. Agrawal, M. R. Blas, R. C. Bolles, B. P. Gerkey, J. Solà, and A. Sundaresan, "Mapping, navigation, and learning for off-road traversal," *J. Field Robotics*, vol. 26(1), pp. 88–113, 2009.
- [15] P. Santana, M. Guedes, L. Correia, and J. Barata, "A saliency-based solution for robust off-road obstacle detection," in *IEEE Inter. Conf. on Robotics and Automation*, 2010, pp. 3096–3101.
- [16] E. Rohmer, G. Reina, and K. Yoshida, "Dynamic simulation-based action planner for a reconfigurable hybrid leg-wheel planetary exploration rover," *Advanced Robotics*, vol. 24(8-9), pp. 1219–1238, 2010.
- [17] A. Nüchter, K. Lingemann, J. Hertzberg, and H. Surmann, "6D slam-3D mapping outdoor environments," *J. Field Robotics*, vol. 24, pp. 699–722, 2007.
- [18] D. Tax, "One-class classification. concept learning in the absence of counter examples," Ph.D. dissertation, Delft University of Technology, Delft, Netherlands, 2001.