The QUAD-AV Project: multi-sensory approach for obstacle detection in agricultural autonomous robotics

Raphaël Rouveure[1]*, Michael Nielsen[2], Anders Petersen[2], Giulio Reina[3], Mario M. Foglia[4], Rainer Worst[5], Sami Seyed-Sadri[5], Morten R. Blas[6], Patrice Faure[1], Annalisa Milella[7], Kasper Lykkegård[6]

[1]Cemagref TSCF, 24 avenue des Landais, BP 50085, 63172 Aubière, France
[2]Danish Technological Institute, Forskerparken 10, 5230 Odense, Denmark
[3]University of Salento, Via Arnesano, 73100 Lecce, Italy
[4]Politecnico di Bari, Viale Japigia 128, 70125 Bari, Italy
[5]Fraunhofer IAIS, Schloss Birlinghoven, 53754 Sankt Augustin, Germany
[6]CLAAS Agrosystems, Bøgeskovvej 6, 3490 Kvistgaard, Denmark
[7]CNR-ISSIA, via Amendola 122/D, 70125 Bari, Italy
**Corresponding author. E-mail: raphael.rouveure@irstea.fr*

Abstract

Autonomous vehicles are being increasingly adopted in agriculture to improve productivity and efficiency. For an autonomous agricultural vehicle to operate safely, environment perception and interpretation capabilities are fundamental requirements. The Ambient Awareness for Autonomous Agricultural Vehicles (QUAD-AV) project explores a multi-sensory approach to provide an autonomous agricultural vehicle with such ambient awareness. The proposed methods and systems will aim at increasing the overall level of safety of an autonomous agricultural vehicle with respect to itself, to people and animals as well as to property. The “obstacle detection” problem is specifically addressed within the QUAD-AV project. The paper focuses on the presentation of the different selected technologies (vision/stereovision, thermography, ladar, microwave radar) through the presentation of preliminary results.

Key words: Robotics, obstacle detection, sensors, fusion

1. Introduction

The idea is that of using different sensor modalities and multi-algorithm approaches to detect the various kinds of obstacles and to build an obstacle database that can be used for vehicle control. For instance, bearing and distance to the nearest collision can be estimated and used by the path planner to change route or speed if an obstacle is in close proximity to the vehicle’s planned path.

The obstacles that might be encountered in the field are separated into four overall categories that should be detected and handled in different ways: positive obstacles such as trees, metallic poles, buildings; negative obstacles (holes, ditches); moving obstacles, including vehicles, people and animals; and difficult terrain (significant slope, water, etc.). Further, obstacles may vary greatly from situation to situation, depending on type of crop, fruit, vegetable or plant grown, curvature of landscape as well as weather conditions. Owing to the variety of situations and problems that may be encountered, no sensor exists that can guarantee reliable results in every case. Any candidate sensor has its strengths and drawbacks. Therefore, a complementary sensor suite should be used to gain the best performance.

The QUAD-AV project (funded by the ICT-AGRI European Research Program) investigates the potential of four technologies: vision/stereovision, ladar, thermography and microwave radar. Existing state-of-the-art sensors, some developed by the partners, have been modified and interfaced in order to be demonstrated in an agricultural context. Using data provided by the different sensors, research focuses on the development of novel sensor processing and sensor fusion techniques to detect and classify obstacles in an agricultural environment. One important challenge is to improve state-of-the art approaches or to introduce innovative methods, taking into account real time processing constraints and challenging environmental conditions, such as lighting variations, dust, fog, presence of living beings.

All sensors were mounted on a test bed tractor, and a data acquisition campaign was conducted in a farm facility near Helsingør, Denmark, in September 2011. Various agriculturally relevant scenarios were tested and we present preliminary results obtained in obstacle detection considering the different sensors individually. Fig. 1 shows the positioning of the sensors on the CLAAS AXION 840 tractor.

FIGURE 1: Positioning of the sensors on the tractor

2. Relevance of various sensors

For this first step, sensors are evaluated independently considering the various obstacles encountered in the field during the experiments. These obstacles are realistic, representative of the obstacles that can be found during an agricultural work. Within the following paragraphs, the four sensor modalities are briefly described, and examples of results are presented.

2.1. Stereovision

Vision is our most powerful sense through which we can get knowledge of the environment and interact intelligently with our surroundings. Similarly, mobile robots can take advantage of visual capabilities. Video sensors supply contact-free, precise measurements and are flexible devices that can be easily integrated with multi-sensor robotic platforms. Hence, they represent a potential answer to the need of new and improved perception capabilities for autonomous vehicles. Specifically, stereovision allows three-dimensional scene reconstruction using two or more cameras placed at different fixed locations in space. By matching corresponding points between images captured by the cameras and assuming the stereo calibration parameters to be known, it is possible to reconstruct the three-dimensional scene. In turns, this allows to implement tasks such as obstacle detection. Regions in the scene can be labelled as either traversable or non-traversable. Alternatively, they can be assigned a traversability cost (e.g., regions containing no hazards, mild hazards, or severe hazards can be assigned low, mild, and high traversability costs, respectively). In this research, the potential of a multi baseline trinocular system is investigated by the University of Salento.

Multi baseline stereo

The choice of the optimal stereo baseline depends on two opposing, although equally important factors: field of view and maximum range. A narrow baseline increases the shared field of view of the two cameras, while yielding to shorter maximum range. Conversely, a larger baseline decreases the common field of view, but leads to higher maximum range and precision at each visible distance. The use of a trinocular configuration in place of a binocular one allows it to combine the advantages of two different baselines by the addition of one camera (Broggi et al., 2010).

In the proposed trinocular system, the narrow pair has a baseline of 12 cm and uses the left and middle cameras. The wide pair has a baseline of 24 cm and uses the left and right cameras. By employing the narrow baseline to reconstruct nearby points and the wide baseline for more distant points, this trinocular stereo system takes the advantage of the small minimum range of the narrow baseline, while preserving, at the same time, the higher accuracy and maximum range of the wide baseline configuration (see left plot in Fig. 2).

Obstacles can be detected based on geometric properties using a classifier that features two main stages: an adaptive training stage and a classification stage. During the training stage, the system automatically learns to associate appearance of geometric data with class labels. Then, it makes predictions based on past observations. The training set is continuously updated online using the latest stereo readings, thus making it feasible to use the system for long range and long duration navigation, over changing environments (Reina et al., 2012).

A typical result is shown in the right plot of Fig. 2, where points with high probability of belonging to the ground are denoted in green and points with high likelihood of belonging to obstacles are marked in red.

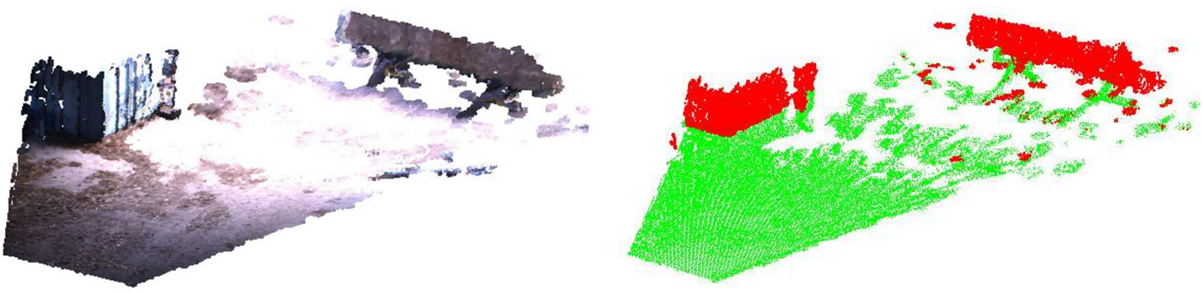

FIGURE 2: Example of three-dimensional scene. Left: reconstructed 3D point cloud using stereo multiplexing of the wide and narrow baseline. Right: scene segmentation using the stereo data. Green: classified ground; red: classified non-ground.

2.2. Ladar

Continuously rotating laser range finders (also known as ladar: laser detection and ranging) are popular sensors for autonomous vehicles. By measuring the distance at which a rotating laser beam is reflected, information about the environment is collected without any dependency on lighting conditions. Such devices are available in a broad price range, where the price of a single sensor could easily exceed the price of the entire tractor. To keep the application economically reasonable, Fraunhofer’s 3D Laser Scanner (3DLS) can be used as a best price option to create 3D point clouds of the environment. The 3DLS consists of a rotational unit with one or two standard 2D laser range finders mounted on it, such that the scanning plane can be rotated. Two different types of sensors are used for the experiments: Sick LMS291 and Sick LMS111, which are both suitable for outdoor operation. The 3DLS is mounted at the front of the tractor with a constant rotation of 0.582 rad/sec. Therefore a rotation of 360° lasts about 10.8 seconds.

During the field experiment, different kinds of interesting hardware and software issues appeared. If one relies only on the ladar sensor, the safe velocity of the vehicle is restricted by the scan duration of the 3DLS. In the case that the vehicle is moving, it is not possible to detect moving obstacles in an acceptable time frame. The 3DLS is well applicable for medium distances, e.g., to generate a map of static obstacles. Positive obstacles like trees, metallic poles, buildings, fences, etc. must be distinguished from vegetation, hedges, and harvest areas. The laser scans of flexible vegetation are partially penetrable, whereas solid obstacles shadow large parts, as it is shown in Fig. 3. To make the best use of the information delivered by the ladar sensor, we are going to implement an algorithm that performs a traversability analysis under off-road conditions, similar to the approach presented by (Kuthirummal et al., 2011).

FIGURE 3: Photo composition of a point cloud created by the ladar sensor and a camera image of the corresponding scene.

2.3. Thermography

The test was performed with FLIR A615 VGA thermal camera. Thermal images were loosely synchronized through software, because the camera did not allow synching. It ran a 50 Hz capture clock and all you can do is skipping frames until you want an image. NUC calibration halts the camera for a few seconds, so this setting must be left on manual mode; otherwise there is the risk that it starts calibration while something dire is happening.

The temperature signatures increase on living things, vehicles, water, but also stones and poles that integrate heat in the sun. This means that it is not a simple task of thresholding the temperatures.

A better segmentation was obtained through region growing and filtering of small and cold particles. See Fig. 4 where a dog and a child was hiding in the woods. A basic temperature threshold would only respond to the bare skin and eyes of the living beings.

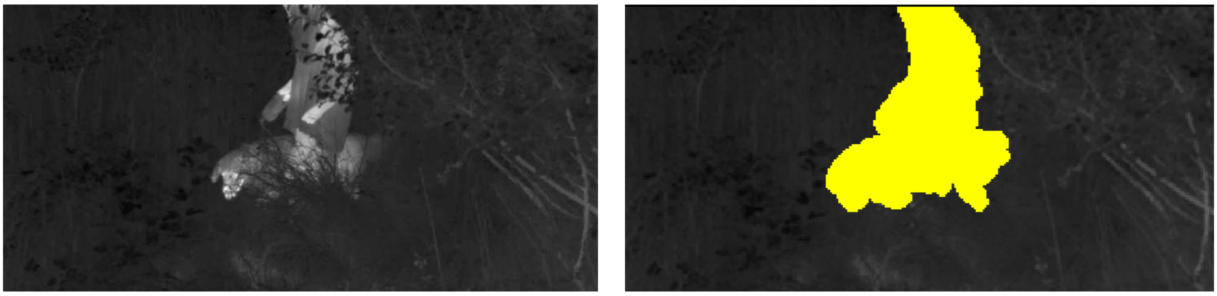

FIGURE 4: Human being with a dog behind vegetation, visible in thermography. Right: result of scene segmentation.

2.4. Microwave radar

The microwave radar technology has been reserved during a long period for military or space applications. But today civil applications are interested by some specific characteristics of radar technology such as (i) robustness in harsh environmental conditions (using millimeter or centimeter wavelength, the radar is not disturbed by dust, rain, light variations, etc.) and (ii) ability to achieve measurements over long range distances.

The K2Pi radar used within the project has been developed by IRSTEA institute. It is a Frequency Modulated Continuous Wave (FMCW) radar (Skolnik, 1980) which is well intended for short and medium distances (up to several hundred of meters). The radar is using a rotating antenna (horizontal plane) in order to provide a panoramic view (360°) of the environment. The main characteristics of the K2Pi radar are presented in Table 1.

TABLE 1: K2Pi radar main characteristics

Carrier frequency	24.125 GHz	Transmitted power (EIRP)	18 dBm
Size (length-width-height)	27-24-30 cm	Weight	10 kg

Each second, the radar provides a panoramic image of the environment in the range 3-100 m. The image uses a PPI (Plan Position Indicator) format, in which each element is localized in polar coordinates. The value of each point of the image indicates the amplitude of the backscattered signal (proportional to the Radar Cross Section). Using panoramic images provided by the radar, a Simultaneous Localization And Mapping (R-SLAM) algorithm developed by IRSTEA is used to build a map of the environment and to compute the localization of the tractor within the map (Rouveure et al., 2009). An example of radar map construction is presented in Fig. 5. The overall covered distance is about 6 km. The mean velocity of the tractor is 2.9 m/s. The radar is on the roof, 3.5 m above the soil. The aerial view (left plot) shows the main elements of the environment (trees, hedges, lake). The gray level of the radar map (middle plot) indicates the amplitude of the backscattered signal. The red dots in the radar map show the computed trajectory. An example of scene segmentation is presented in the right plot. Three different zones are distinguished: ground (red), non-ground or obstacle (red), and low reflective or electromagnetic shadow areas (blue).

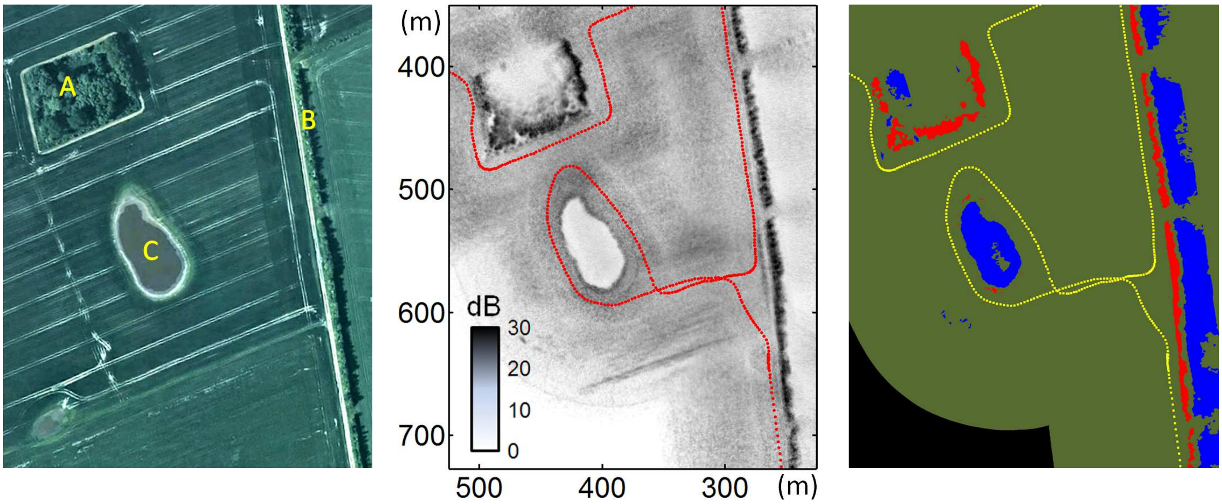

FIGURE 5: Microwave radar result. Left: aerial view of the test zone (Google Earth, June 2011). A: trees; B: hedge; C: small lake. Middle: corresponding radar map. The red dots indicate the computed trajectory of the tractor. Right: scene segmentation result. Green colour indicates ground zone and red colour obstacles. Blue colour shows low reflective targets such as water or electromagnetic shadow zones.

3. Conclusion

The objective of the QUAD-AV project is to explore a multi-sensory approach in order to develop perception and interpretation capabilities within the general framework of agricultural robotics. Considering that one technology cannot solve all situations encountered in agricultural environments, a set of sensor technologies is evaluated within this project: stereovision, ladar, thermography and microwave radar. A first experimental campaign has been achieved in real conditions in order to build a database. Within this first phase, the four technologies are tested individually and their ability to separate ground and non-ground obstacles is evaluated. Further work will consider: the evaluation of different kind of segmentation processes; the integration of sensors data in the same spatial reference; data merging in order to increase the robustness and the relevance of the segmentation processes, including obstacles classification.

Reference list

Broggi, A., Cappalunga, A., Caraffi, C., Cattani, S., Ghidoni, S., Grisleri, P., Porta, P., Posterli, M., & Zani, P. (2007). TerraMax Vision at the Urban Challenge 2010. IEEE Transactions on Intelligent Transportation Systems, 11(1), 194-205.

Kuthirummal, S., Das, A., Samarasekera, S. (2011). A graph traversal based algorithm for obstacle detection using lidar or stereo. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 1579-1584.

Reina, G., Milella, A., & Underwood, J. (2012). Self-learning classification of radar features for scene understanding. Robotics and Autonomous Systems, Available online 4 April 2012, ISSN 0921-8890, 10.1016/j.robot.2012.03.002.

Rouveure, R., Monod, M.O., & Faure, P (2009). High Resolution Mapping of the Environment with a Ground-Based Radar Imager. IEEE International Radar Conference (RADAR'09), 6 p.

Skolnik, M.I. (1980). Introduction to radar systems. In Electrical Engineering Series, McGraw-Hill International Editions.