



HAL
open science

Hybrid topological location and mapping for autonomous agricultural robots

L. Emmi, J. Dufour, Viviane Cadenat, M. Devy

► **To cite this version:**

L. Emmi, J. Dufour, Viviane Cadenat, M. Devy. Hybrid topological location and mapping for autonomous agricultural robots. European Conference on Precision Agriculture, Jul 2019, Montpellier, France. pp.767-774, 10.3920/978-90-8686-888-9_95 . hal-04909409

HAL Id: hal-04909409

<https://hal.science/hal-04909409v1>

Submitted on 23 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Hybrid topological localization and mapping for autonomous agricultural robots

L. Emmi, J. Dufour, V. Cadenat, M. Devy

LAAS, CNRS, 7 Avenue du Colonel Roche, 31400 Toulouse, France

luis.emmi@laas.fr

Abstract

This paper considers the problem of autonomous navigation in an agricultural field. It proposes a localization and mapping framework based on semantic place classification and key location estimation, which together build a hybrid topological map. The approach has been assessed through off-line data, recorded in real conditions, on diverse fields, at different seasons. The results have shown the interest of the approach, which allows: **(i)** simple and easy-to-update map to be obtained; **(ii)** the use of artificial landmarks to be avoided, thus **(iii)** autonomy of agricultural robots to be improved.

Keywords: Hybrid topological map, semantic identification, autonomous navigation.

Introduction

In the last twenty years, an important number of robots have been developed to help in daily tasks (Royakkers & van Est, 2015), including the management and care of crops. The development of research prototypes and commercial robots for agriculture has been recently growing (Ampatzidis et al., 2017). Their autonomy has also been improved, reducing the need for human intervention. Many of these developments focus on solving the problem of navigation within the cultivated areas of a farm, known as fields, where two main approaches stand (Bechar & Vigneault, 2016): GNSS-based navigation and sensor-based navigation. The first one gives the robot information about the path to be followed, although a precise map prior operation is required, commonly obtained through the use of Real-Time Kinematic GNSS (RTK-GNSS). Despite its advantages, some drawbacks can be identified: **(i)** its high cost, **(ii)** the need for up-to-date maps, and **(iii)** its vulnerability in certain areas (GNSS-denied zones). Sensor-based navigation relies on proximity sensors, such as cameras or LIDARs, for avoiding obstacles and following crop rows (Durand-Petiteville et al., 2018). And, although this topic has been deeply addressed in recent years, it still requires a prior knowledge of the beginning and end of the rows to ensure robust navigation. This crucial information is generally provided by the farmer, making most of the solutions specific to the considered field. Although some proposals have tackled this situation, there is still a need for general navigation strategies, less dependent on **(i)** human intervention, **(ii)** absolute localization, and **(iii)** controlled environment, allowing long-range navigation to be achieved.

This problem is at the core of this paper, motivated to improve the autonomy of Oz, a commercial weeding robot developed by Naïo Technologies (Anon., 2018), that requires the use of artificial landmarks for row boundary identification. Thus, this paper presents a localization and mapping framework that takes advantage of the particular structure of the field to extract natural landmarks for precise and autonomous navigation. The framework presented consists of a perception system (LIDAR and RGB camera) that acquires and labels the information, which then is analysed by a semantic

classifier to extract natural features. The perception system and the semantic classifier are accompanied by a map update process.

A field is generally structured by a sequence of rows connected through two headlands. This allows the working environment to be divided into a limited number of places, which are related to each other by specific rules. This type of problem can be described by a topological map, a commonly used strategy in in-door navigation to estimate the state (position and orientation) of a robot in a partitioned and semi-structured environment (Thrun & Bücken, 1996). In this sense, the framework presented includes a general modelling of the field expressed throughout a topological map. The semantic classifier benefits from this model to compare the acquired information and thus estimate the area where the robotic system is currently located, and the exact points where the transitions occur. Both elements together build a hybrid topological map. Some advantages can be highlighted for this type of map: **(i)** it captures the environment connectivity, while limiting the use of metric information; **(ii)** the map is 'lighter' and easier to update; **(iii)** absolute localization sensors and previous mapping are no longer necessary; and **(iv)** artificial landmarks are no longer required.

Related Work

Currently, there are numerous companies that offer diverse mapping services, from delimiting the farm zones, to obtaining in detail several vegetation and soil indices. The mapping task is primarily done by the use of UAVs, satellite images or manually. Most solutions use dedicated software to analyze the data, and mainly rely on RTK-GNSS. Moreover, most of the maps that are created in these solutions use the contour of the field given by an external source or acquired with their own system, and generate the path that the tractor driver should follow for optimal operation. These are called full coverage maps, and are independently created of the seeded crop. Most of the maps that could be available via this kind of service might provide some information to support mobile robotic systems to autonomously navigate on a farm, such as seeder-extracted maps. However, the acquired data needs to be manually processed to be adapted for: **(i)** a specific robotic system, **(ii)** a specific task, and **(iii)** a specific environment.

One alternative to absolute mapping is the fusion of proximity sensors (i.e. LIDAR) with odometry and/or camera systems, which has been a broad subject of research. Most of the developed works seek to solve the problem of row following, and only some of them address the detection of the beginning and end of the crop row. Normally, artificial landmarks are installed for the detection of these limits (Zhang et al., 2014). Although it is a suitable solution for perennial crops, manual work is required for their installation and removal in non-perennial crops. This restrains the autonomous capabilities of robotic systems. An alternative is using the information already provided by the natural environment, through the semantic extraction of key elements and an adequate modeling of the working environment. Related to this subject, a specific work has aroused interest, where semantic classification of the meaning of a field is proposed (Weiss & Biber, 2010). They have used a particle filter to determine the probability distribution of the robot state in the different locations where it is applied. It should be noted that in this approach **(i)** there are no limits on the meaning of open field, **(ii)** the row-start and row-end areas are not specific points, and **(iii)** it is not clear the difference between row-start and row-end areas, which hinders its use to accurately define the transition between the row and the open field.

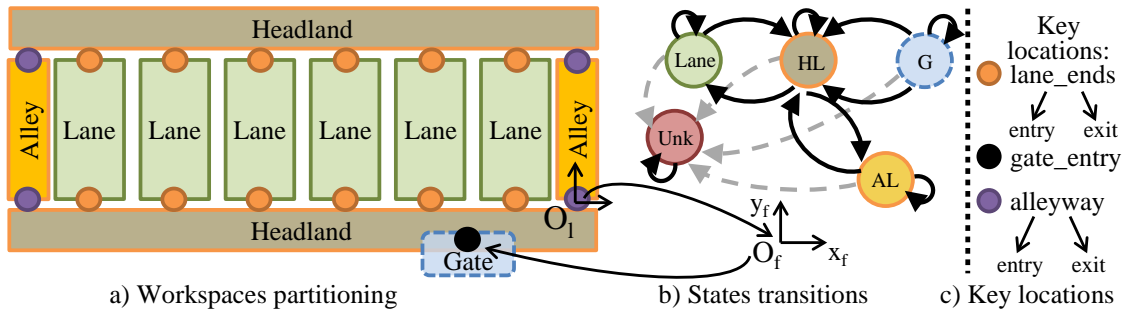


Figure 1. The proposed field model. a) The semantic field place partitioning. b) Topological representation of the workspaces. c) Key location types.

The field model

In general, a field is a cultivated area consisting of one type of crop, planted in parallel and at a relatively constant distance between them. A field can be divided into four places, named: headland, lane, alley, and gate (see Figure 1.a). The headland is where the robot switches between lanes or alleys, or exits the field. The alley and lane represents the cultivated area, although the alley corresponds to the lateral boundaries. Both could include more than one inter-row space. Moreover, their width corresponds to the width of the tool. Finally, the gate allows the robot to enter/quit the field.

These places represent the possible robot *states* (see Figure 1.b), i.e. a semantic representation of the localization of the robot in the field. A fifth *state* called ‘unknown’ has been included to handle the uncertainties when sensory data is not conclusive. At this step, the proposed topological map only allows the system to instantaneously distinguish in which type of place (*state*) the robot is present, but it is not able to predict future *states* or where transitions are located. This kind of map may not bring a sufficient accuracy to efficiently control the tools. Therefore, a metric map is included by estimating the exact coordinates of the transitions, called key locations (see Figure 1.c). The key locations are represented by points $p(x_p, y_p, z_p, \vec{v}) \in \mathbb{R}^3$, where \vec{v} is a unit vector that defines the orientation of p , related to $F_f = (O_f, \vec{x}_f, \vec{y}_f, \vec{z}_f)$. F_f is the field coordinate frame located at the gate. A temporary reference frame $F_l = (O_l, \vec{x}_l, \vec{y}_l, \vec{z}_l)$, related to F_f , serves to keep the perception system information referenced close to the last key location identified. Then, F_l is relocated whenever a *state* transition is detected. This procedure allows the drift caused by the odometry to be minimized, with the drawback that the previous key locations lose validity within F_f . On the other hand, in the map update process, the key locations can be referenced to a global reference frame. The fusion of the semantic identification and the estimation of the key locations builds the hybrid topological map.

Algorithm for semantic localization and mapping

Perception system: data acquisition and labelling

Figure 2 presents the algorithm developed. The algorithm looks for valid key locations by analysing the distribution of the crops, encompassed in clusters of laser point. The type of crop is the sole required configuration input, and the laser points are labelled depending on if they belong to that crop (see Figure 2.b). This allows the alleys to be identified. The laser points are placed in the same reference frame as the camera. Then,

the points that are within the positive detection are labelled (see Figure 2.b). Since the labelling process requires certain computational load, several epochs of laser data are labelled from a single camera image. The labelled points are referenced into F_l which is relocated each time the algorithm is initialized. Then, the points are clustered based on DBSCAN (density-based spatial clustering of application with noise).

Each cluster represents a crop row, or part of one, given that gaps can be found (see Figure 2.c). The clusters are kept stored until the system detects a change in the *state*. Therefore, every time new data is acquired, the stored clusters are updated. When a change in *state* is detected, the process is restarted, i.e. the clusters are discarded and a new F_l is defined. This procedure provides enough information to enrich the detection process of the key locations, even if the robot localization system is noisy.

For laser points labelling, an object identifier based on a classical convolutional neural network (CNN) called YOLOv3 (Redmon & Farhadi, 2016) has been developed. It has been trained to identify an important variety of market garden crops. The training was performed using approximately 14,000 examples (4/5 of the data for training, and 1/5 for testing). The samples were extracted from the on-board camera of the Oz robot, performing real operations, in different seasons, farms, light conditions, etc. (crops and number of samples: green lettuce: 8,258; red lettuce: 1,808; cabbages: 1,301; leeks/onions: 1,444; rapeseed: 1,135). Each sample was annotated by a hand-drawn bounding box and each annotation consists of an image where only the desired crop appears. To evaluate the selected network and the ability to correctly classify the crops, a comparative study has been performed with RetinaNet – Resnet152 (Lin et al., 2017), another common network for object classification, which was also been trained with the same data. Similar performances have been obtained on both networks, in terms of average precision (AP) and mean average precision (MAP). A threshold of about 0.5 of the intersection over union (IoU) was used for performance analysis. Nevertheless, YOLOv3 offers better results in terms of performance P (on a CPU and a GPU), which allows it to be implemented on the Oz robot without installing any extra or expensive hardware. Table 1 presents results in terms of detection capacity and processing time.

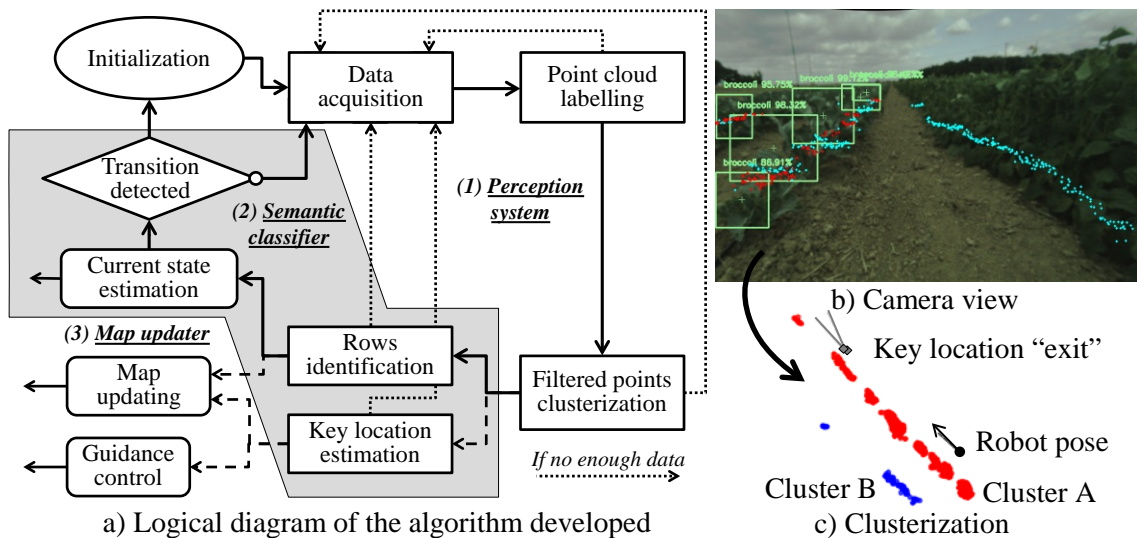


Figure 2. Algorithm for semantic localization and mapping. a) Logical scheme for state estimation and key location estimation. b) Laser points labelling based on crop detection. c) Laser points clusterization and key location estimation.

Table 1. Comparison of two CNNs trained for crop detection. Detection is considered true positive (TP) if the IoU between ground truth and detection is greater than 50%.

CNN	AP-.50					MAP-.50	P [ms]	
	Cabbage & Broccoli	Rape-seed	Green lettuce	Red lettuce	Onion /Leek		i7 6600U	GTX 1080
YOLOv3	0.871	0.676	0.966	0.957	0.499	0.794	1,460.7	56.9
RetinaNet	0.86	0.73	0.966	0.958	0.488	0.8	6,124.8	72.4

Semantic identification and key location estimation

For this work, only the identification of the headland ($n=H$), lane ($n=L$) and alley ($n=A$) are considered. For the two situations (see Figure 3), left and right points are estimated, placing them at a distance d of the nearest cluster. If both points are located, a key location is set at the midpoint between both ($lane_end$). Non-valid key locations are discarded, and only those that are more likely to represent the *state* transitions are maintained. They are named *valid key locations*, and are classified as entry ($t=En$) or exit ($t=Ex$). Two sorts of indices are calculated: key_idx_t , corresponding to a correlation between the same types of *valid key location*, and $angle_idx_t$, corresponding to the correlation between the robot orientation and the last *valid key location*. For every *state* studied, an index c_idx_n is calculated, given by (1) and (2):

$$\begin{cases} c_idx_H = \frac{key_idx_En}{angle_idx_En} \times cluster_idx & \text{if } left \cup right \\ & \text{alley_idx}=0 \\ & \text{lane_idx}=1 \\ c_idx_A = \frac{angle_idx_Ex}{cluster_idx} \times alley_idx, & \\ c_idx_L = \frac{angle_idx_Ex}{cluster_idx} \times lane_idx & \text{else} \\ & \text{alley_idx}=1 \\ & \text{lane_idx}=0 \end{cases} \quad (1)$$

$$cluster_idx = \frac{abs(B_{dist} - F_{dist})}{euc_dis(FC, BC)} \quad (2)$$

Equation (2) expresses how misaligned the robot is with respect to the nearest cluster, calculated by projecting a forward FR and backward BR points, and the nearby cluster points are located, designated as FC and BC respectively (see Figure 3.a). Then, the orthogonal distances (B_{dist}, F_{dist}) between the projected points and the cluster are calculated, where euc_dis represents the Euclidean distance between two points. These indices give an idea of how probable the robot is located in the topological map. Also, the *state* estimation is improved by taking into account the possible transitions allowed and the previous estimation. For this, a simplification of the state transition probability given a motion presented in Weiss & Biber (2010) is used, since the representation of the environment and the relationship between the different *states* is similar.

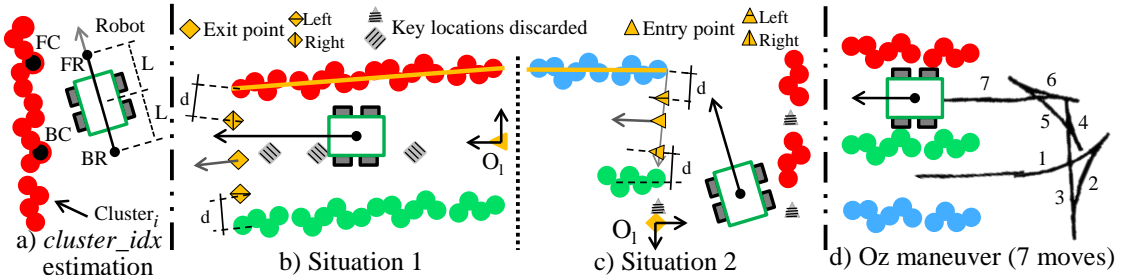


Figure 3. Example of key location extraction. a) a_index points definition ($L = 1$ m). b) Looking for exit. c) Looking for entry ($d = 0.1$ m). d) Example of Oz maneuvers.

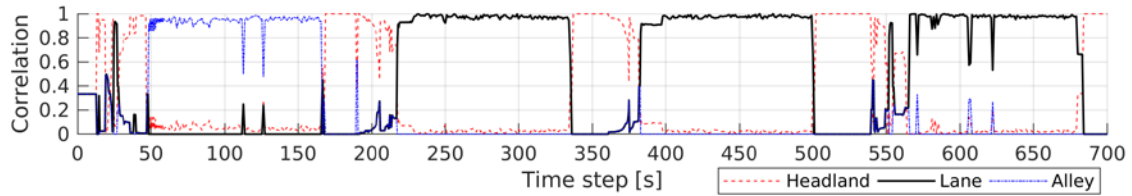


Figure 4. Correlation index c_idx_n for each state studied (dataset B1).

Guidance control and map update

Given the current *state* estimation and the *valid key location*, the navigation strategy can be adjusted to meet the current objective provided by the topological map and the general mission. The identification of the *valid key location* allows the exact coordinate where a *state* transition occurs to be predicted. This also allows the work within the field to be performed with greater efficiency and precision. The navigation strategy then relies on the search of possible lane candidates that require to be treated (traversed). Therefore, each time a *state* transition is identified, the last *valid key location* is stored into the metric map, referenced to F_f (see Figure 1.a). This is only possible if a global localization system is available (i.e. GNSS), even if the uncertainty of this measurement is bigger than the distance between two nearest key locations. This map can be used to give the robotic system an idea of the necessary orientation to find a specific node within the hybrid topological map. Although, once close to that location, the approach presented should be used for precise navigation.

Results and discussion

The approach proposed has been tested and validated off-line with seven diverse dataset extracted from the Oz robot, recorded in real working conditions. Each dataset contains a set of asynchronous raw images, as well as the laser point-cloud referenced with the odometry of the robot. Five were collected in the same farm, on two different days, for two different types of crop (cabbage and broccoli). The remaining set was obtained for different seasons in a field of leeks located in a greenhouse. Thus, these data have been acquired during various real situations: high changes in lighting conditions (different time of day), different crops, different locations, etc. Figure 4 shows the correlation indices c_idx_n for one of the passes (a broccoli field consisting of four crop rows with a length of about 50 m). The maxima of the indices are used to estimate the *state*, including previous estimation and transitions allowed. Figure 5 presents the *state* estimation based on c_idx_n and compared it with the ground truth.

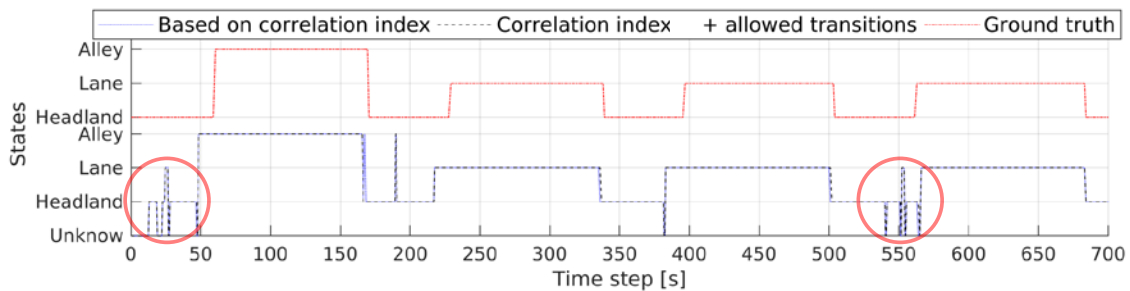


Figure 5. State estimation based on only the correlation index (continuous line) and on the previous estimation and transitions allowed (dotted line). Dataset B1.

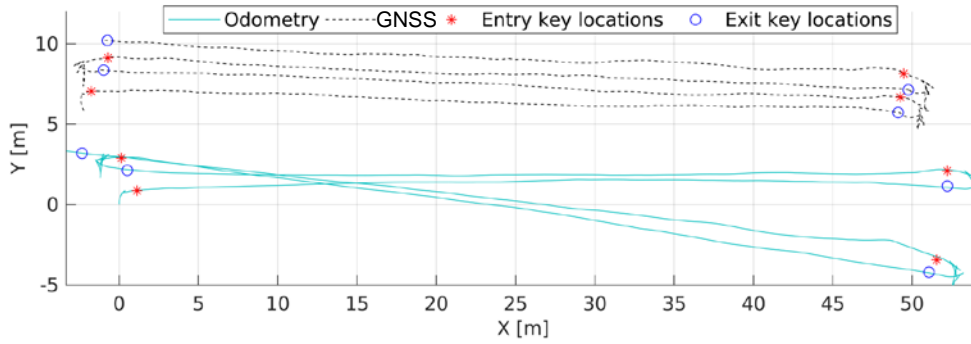


Figure 6. Odometry of the robot and GNSS with key locations estimated for dataset B1.

To define the ground truth, the positions of the artificial landmarks (red sticks) required by the Oz robot to navigate are used. It must be noted that, in some cases, these marks might not exactly be placed on the last crop, inducing a small error. In addition, every time the robot enters a lane, it performed a maneuver composed of several forward and backward movements (between five and seven, see Figure 3.c). Thus, in some cases, the perception system successively gets in and out of the lane, which generates false positive detections (see red circles on Figure 5). Indeed, this approach assumes forward motion only (there is no information about the motion direction). Despite these perturbations, the system is able to work properly. Along with the semantic robot localization, the key locations are also estimated. Figure 6 shows the odometry of the robot and the *valid key locations* for entry and exit. For the purpose of comparing the results, the *valid key locations* are also referenced with the GNSS sensor on-board the Oz robot, which may allow creating a map of the field to be used in the future. This map will not guarantee accurate and safe navigation given the GNSS accuracy limitations (± 0.6 m of positioning error), but in conjunction with the semantic localization in the topological map, the navigation strategy through the entire farm should be possible. Table 2 presents the analysis of the detection success rate (DSR), in addition to the mean distance error (MDE) between the detected key locations and the ground truth. Moreover, the orthogonal distance error (Orth_E) is analysed to identify if a key location was estimated in a different lane. Even with the error introduced in the definition of the ground truth, the results obtained are encouraging, demonstrating that it is possible to locate the rows boundaries without the use of artificial landmarks.

Table 2. State estimation and key locations results, for dataset: B for broccoli; C for cabbage; and L for leeks. For every key location type, a MSE of its position as well as the orthogonal projection error (Orth_E) are calculated compared with the ground truth.

Set	MDE [m]			Orth_E [m]			Transitions		DSR
	Entry	Exit	Mean	Entry	Exit	Mean	Entry	Exit	
B1	0.36	0.151	0.256	0.083	0.038	0.06	4	4	0.829
B2	0.829	0.533	0.681	0.3	0.187	0.244	4	4	0.875
C1	0.478	0.473	0.475	0.006	0.115	0.089	3	4	0.86
C2	0.655	1.153	0.904	0.052	0.137	0.095	3	4	0.923
C3	0.647	0.275	0.461	0.334	0.089	0.212	3	4	0.804
L1	0.921	0.868	0.895	0.148	0.458	0.301	8	7	0.805
L2	0.062	0.343	0.203	0.037	0.038	0.034	2	2	0.637

Conclusions

This paper considers the problem of autonomous navigation through an agricultural field. First, a model of the field is introduced to integrate it into a hybrid topological map. This map is part of the localization and mapping framework presented. This approach introduces an alternative for absolute localization or the use of artificial landmarks, which are not always reliable systems for navigation. The approach presented benefits from the structure of the field to identify indistinguishable characteristics that allows the instantaneous robotic pose to be estimated. Moreover, the hybrid attribute of the map also makes possible the state transitions to be predicted. This approach has been validated using off-line data, recorded in real working conditions, from diverse fields with different crops. The results have demonstrated the efficiency and the interest of the approach proposed, paving the way for low-cost and long-range navigation in agricultural environments.

Acknowledgements

The authors would like to thank colleagues from Naïo Technologies, Toulouse, for their participation to this work, through the collaborative project DESHERB'EUR funded by the program «Investment for the Future» of the French government.

References

- Ampatzidis, Y., De Bellis, L., and Luvisi, A. 2017. iPathology: robotic applications and management of plants and plant diseases. *Sustainability*, 9(6), 1010.
- Anon., 2018. Robots agricoles Naïo Technologies. <https://www.naio-technologies.com/> (last accessed 10/30/18).
- Bechar, A., and Vigneault, C. 2016. Agricultural robots for field operations: Concepts and components. *Biosystems Engineering*, 149, 94-111.
- Durand-Petiteville, A., Le Flecher, E., Cadenat, V., Sentenac, T., and Vougioukas, S. 2018. Tree Detection With Low-Cost Three-Dimensional Sensors for Autonomous Navigation in Orchards. *IEEE RAL*, 3(4), 3876-3883.
- Lin, T., Goyal, P., Girshick, R., He K., and Dollár, P. 2017. Focal Loss for Dense Object Detection. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2999-3007.
- Redmon, J., and Farhadi, A. 2016. “Yolo9000: Better, faster, stronger,” *arXiv preprint arXiv: 1612.08242*.
- Royakkers, L., and van Est, R. 2015. A literature review on new robotics: automation from love to war. *International Journal of Social Robotics*, 7(5), 549-570.
- Thrun, S., and Bücken, A., 1996. Integrating grid-based and topological maps for mobile robot navigation. In *Proceedings of the National Conference on Artificial Intelligence*, Portland, OR, AAAI Press/MIT Press, Menlo Park, CA, 944-951.
- Weiss, U., and Biber, P. 2010. Semantic place classification and mapping for autonomous agricultural robots. In *Proceedings of IROS Workshop on Semantic Mapping and Autonomous Knowledge Acquisition*, Taipei, Taiwan.
- Zhang, J., Maeta, S., Bergerman, M., and Singh, S. 2014. Mapping orchards for autonomous navigation. In *Proceedings of the American Society of Agricultural and Biological Engineers Annu. Int. Meeting*.