

POLITECNICO DI MILANO
Facoltà di Ingegneria
Scuola di Ingegneria Industriale e dell'Informazione
Dipartimento di Elettronica, Informazione e Bioingegneria
Corso di Laurea Magistrale in
Ingegneria Informatica



**DESIGN, DEVELOPMENT AND DEPLOYMENT
OF A MOBILE MANIPULATOR
FOR VINEYARD MONITORING AND PROTECTION**

Relatore:

PROF. MATTEO MATTEUCCI

Correlatore:

PROF. LUCA BASSETTA

Tesi di Laurea Magistrale di:

GIOVANNI BERI
Matricola n. 852984

Anno Accademico 2016-2017

*"Il mio falegname con 30 000 lire
la faceva meglio"*

— Giovanni Storti

RINGRAZIAMENTI

Per prima cosa voglio ringraziare il professor Matteucci, per la disponibilità e umanità con cui mi ha seguito in questo percorso di tesi; inoltre lo ringrazio di avermi dato la possibilità di lavorare sulle problematiche d'insieme del robot sviluppato, e non solo sui singoli dettagli.

Insieme a lui, vorrei poi ringraziare particolarmente le altre persone che hanno contribuito direttamente a questo lavoro di tesi: il professor Bascetta, per le preziose competenze e per l'umiltà che per me hanno contribuito molto a creare un clima di collaborazione, soprattutto nelle trasferte; Gianluca Bardaro, per i centratissimi consigli tecnici, la pazienza mentre mi ingarbugliavo nel C++, e la compagnia nelle trasferte; Giulio Fontana, per il supporto tecnico e la compagnia soprattutto nelle prime disorientate settimane nell'AIRLab; Federico Polito, per l'amicizia e le battaglie contro il Kinova.

Per il supporto durante tutti gli anni al Poli vorrei ringraziare la mia famiglia: i miei genitori Wally e Marco, mio fratello Ciltro e mia sorella Maria Sara (che non è un granché, è vero, ma ho soltanto quella), e i nonni, zii e i cugini che mi hanno aiutato e fatto compagnia anche mentre ero a Milano.

Desidero davvero ringraziare tutti gli amici, del Poli e non, con cui ho condiviso gli anni dell'università. In modi diversi sono stato aiutato da tante persone, ma dato che non riesco né a scriverle né a ricordarle tutte, esordisco con un ringraziamento plenario a tutti quelli che sanno o pensano di essere in questo elenco. Ci tengo però anche a specificare pochi ringraziamenti particolari. Primo fra tutti, all'indissolubile gruppo degli informatici, senza cui veramente non so se ci sarei arrivato in fondo: Sid, Fatte, Beppe, Biagio, Zure e Carlos. Poi, agli equipaggi di AppaRolle, AppaPirata (vol. I-IV) e AppaHonest, perché l'unico vantaggio dell'abitare da solo sarebbe stata la pizza tutte le sere. Agli amici della Bassa, che mi hanno sempre riaccolto anche quando ho iniziato a usare intercalari poco emiliani: Magri, Bendolo, Baraldi, Celeste, Laura, Giacomo, Co' (anche se migra!!).

Infine, un grazie particolare all'Anna, che mi aiuta ad andare al cuore delle cose.

Grazie,

Giovanni

CONTENTS

Abstract	XII
Sommario	XIII
1 INTRODUCTION	1
1.1 Thesis contribution	2
1.2 Structure of the thesis	2
2 BACKGROUND AND TOOLS	4
2.1 Robot Operating System	4
2.2 TF: The Transform Library	9
2.3 Odometry	11
2.3.1 Differential drive robot	12
2.3.2 Skid-steering kinematics	15
2.4 Sensor fusion	16
2.4.1 Robot Localization	18
2.5 Navigation Stack	20
2.6 Motion Planning	21
3 THE GRAPE PROJECT	27
3.1 Project Description and Goals	27
3.2 Contextualization of thesis work in the project	29
3.3 Related Projects	32
4 GRAPE HARDWARE AND SOFTWARE ARCHITECTURE	34
4.1 GRAPE hardware architecture	34
4.1.1 Robotic Base	34
4.1.2 Robotic Arm	35
4.1.3 LIDARs	37
4.1.4 RGB-D camera	39
4.1.5 Tools for dispensers grasping	39
4.1.6 Other sensors	40
4.2 GRAPE software architecture	40
4.2.1 Evaluation of vineyard navigation	41
4.2.2 Evaluation of vineyard monitoring task	42
4.2.3 Evaluation of dispenser application task	42
4.2.4 Details of the actions interfaces	45
4.2.5 The supervisor	47
4.2.6 Software modules treated in this thesis	48
5 ODOMETRY AND NAVIGATION SYSTEMS IN GRAPE	49
5.1 Sensor Fusion system	50
5.2 Navigation system	53
5.2.1 Navigation in mapped environment with AMCL	53
5.2.2 Insertion of prohibition layer	57
5.2.3 Switch to mapless navigation	57
5.2.4 Switch to Timed Elastic Band as local planner .	59
5.2.5 AMCL as input to sensor fusion	61

6	MANIPULATION	63
6.1	Kinova ROS API	63
6.2	Motion execution methods for Jaco ² Arm	64
6.2.1	<i>MoveIt!</i> framework	64
6.2.2	Direct publication of joint velocities	65
6.2.3	Inversion of differential kinematics	66
6.2.4	Image Based Visual Servoing	68
6.3	Scan motion action server	70
6.4	Dispenser application action server	71
6.4.1	Image based validation	72
6.4.2	Visual servoing tasks	76
7	EXPERIMENTAL RESULTS	83
7.1	Sensor fusion	84
7.2	Navigation	85
7.3	Manipulation	85
8	THE ROBÌ PROJECT	92
8.1	Robì mobile manipulator	92
8.2	Robì as GRAPE robotic base	94
8.2.1	Hardware configuration	95
8.2.2	Software architecture	97
9	CONCLUSIONS	100
	BIBLIOGRAPHY	102

LIST OF FIGURES

Figure 2.1	<i>Three phases of the setup of communication of nodes through topics.</i>	6
Figure 2.2	<i>An example of the Robot Operating System (ROS) Computation Graph, visualized with rqt: nodes are represented with circles, rectangles represent topics, and arrows go from a node to a topic it publishes on, or from a topic to a node subscribed to it. It's easy to recognize the many-to-many relationship.</i> . . .	7
Figure 2.3	<i>Visualization of the ROS Computation Graph with rqt in the context of a data bag being played: there is a single node (GRAPE_robot) that simultaneously plays recorded topics messages.</i> . . .	8
Figure 2.4	<i>Sketch of the implementation of actionlib, through ROS topics and a callback system.</i> . . .	8
Figure 2.5	<i>A robot in 2 different positions, with tf frames in evidence: x-axis is red, y-axis is green, z-axis is blue. The frames are the same in both configuration, but the transformations (i.e., rototranslations) between them are different.</i> . . .	10
Figure 2.6	<i>An example of tf tree</i> . . .	11
Figure 2.7	<i>If the axis of all the wheels intersect in a single point, it's called ICC and the robot can move without slipping</i> . . .	12
Figure 2.8	<i>On Figure 2.8a, the scheme of a differential drive motion model; in Figure 2.8b, an example of a differential drive robot (Pioneer 3DX).</i> . . .	12
Figure 2.9	<i>Differential drive: $(x, y, \theta) \rightarrow (x', y', \theta')$</i> . . .	14
Figure 2.10	<i>Husky platform from Clearpath Robotics is the platform used for the development of Ground Robot for vineyArd Monitoring and ProtEction (GRAPE) project.</i> . . .	15
Figure 2.11	<i>On Figure 2.11a, the scheme of a skid steering motion model; in Figure 2.11b, an example of a skid steering vehicle.</i> . . .	16
Figure 2.12	<i>The equivalence of differential drive motion model and skid steering model according to Wang et al., 2015</i> . . .	17
Figure 2.13	<i>Sensor fusion 2.13a vs multi-sensor integration 2.13b</i> . . .	18
Figure 2.14	<i>A scheme of Robot Localization sensor fusion</i> . . .	19

Figure 2.15	<i>Local and global costmap visualized with Rviz. As you can see, the local costmap (in brighter colors) is built around the robot, and moves together with it.</i>	21
Figure 2.16	<i>A scheme representation of the Navigation Stack; see the color legend for classification in provided, optional provided, and platform specific nodes.</i>	22
Figure 2.17	<i>A graphical representation of the trajectory planned for moving a robotic arm from start (P₁) to goal pose (P₂), satisfying geometrical constraints of the robot.</i>	23
Figure 2.18	<i>The graphical construction of C-space from workspace, by sliding robot shape along the borders of obstacle regions. In Figure 2.18a you can see the procedure in 2D, in Figure 2.18a in 3D. Note that in 3D you only need to compute 2D procedure for each θ, and then stack all obtained images.</i>	24
Figure 2.19	<i>Graphical representation of the ways of proceeding of a sample-based planner (2.19a), and of a grid-based planner (2.19b)</i>	25
Figure 2.20	<i>A block scheme representing the high-level MoveIt! architecture.</i>	26
Figure 3.1	<i>A vine before (3.1a) and after (3.1b) having leafed out.</i>	28
Figure 3.2	<i>Different shapes of pheromone dispensers available on the market. The model used in GRAPE is the one depicted in image 3.2b</i>	30
Figure 3.3	<i>Detail of the dispenser type used in GRAPE.</i>	30
Figure 3.4	<i>Simulation of the Husky robot in the vineyard environment. Simulation built using Gazebo simulation software.</i>	31
Figure 3.5	<i>The 3D printed vine tree mockup, used in validation phase of the dispenser application task algorithm.</i>	32
Figure 3.6	<i>Other field robots developed in EU projects in the last 6 years: 3.6a Rhea, 3.6b CROPS.</i>	32
Figure 3.7	<i>Other field robots developed in EU projects in the last 6 years: 3.7a Vinerobot, 3.7b Vinbot.</i>	33
Figure 4.1	<i>Husky platform after a navigation session in environment. You can easily note the amount of stones and mud stuck into the wheels.</i>	35
Figure 4.2	<i>Robotic arm Jaco² with 3 fingers, from Kinova Robotics (4.2a) and the same arm mounted on a wheelchair 4.2b</i>	37
Figure 4.3	<i>The final configuration of the sensors (Laser Imaging Detection and Ranging (LIDAR) and RGB-D camera) mounted on top of the end effector.</i>	37

Figure 4.4	<i>Three LIDARs mounted on our unmanned ground vehicle: Tim561 (4.4b) from SICK, Puck Lite (4.4a) from Velodyne, UTM-30LX-EW (4.4c) from Hokuyo.</i>	38
Figure 4.5	<i>The final version of the feeder, holding three dispensers.</i>	40
Figure 4.6	<i>The final version of the 3D-printed nails (4.6b), and their perspective drawing (4.6a)</i>	40
Figure 4.7	<i>The complete robotic platform, with sensors/actuators final arrangement.</i>	41
Figure 4.8	<i>Examples of visual binary fiducial markers (more specifically, ArUco markers) as the ones we used in visual servoing applications.</i>	43
Figure 4.9	<i>Examples of laser scans collected in Casciana Terme during the integration week in January 2018, with the LIDAR mounted on top of the end effector of the arm. You can recognize the vine plant, and part of the Husky base. The colors are only for visualization purpose.</i>	44
Figure 4.10	<i>Example of a dispenser application task, correctly executed on a mockup plant.</i>	44
Figure 5.1	<i>Computation graph of the cascade of Robot Localization nodes; in blue, the input topics, and in green the output topics, of the nodes highlighted in red.</i>	52
Figure 5.2	<i>Evolution of the belief computed by Adaptive Monte Carlo Localization (AMCL), taking into account the motion update only (ignoring the sensor update).</i>	55
Figure 5.3	<i>Map of the vineyard in Garriguella (ES), computed during the last integration week using Robot Localization and Gmapping. In 5.3b, a magnified detail.</i>	55
Figure 5.4	<i>Belief about the robot estimate using AMCL, through a particles cloud where each particle represents a hypothesis about the real pose.</i>	56
Figure 5.5	<i>Example of global plan that takes into account the map obstacles.</i>	59
Figure 5.6	<i>RTK stations provided by SmartNet Italy.</i>	60
Figure 5.7	<i>Example of local costmap, in the final configuration: you can distinguish the virtual fences (blue), the obstacles detected by the LIDAR (pink), and the inflation layer (light blue). The green polygon represents the robot footprint.</i>	60
Figure 5.8	<i>Example of local plan computed by DWA that get stuck in the jagged border of the obstacles.</i>	61
Figure 5.9	<i>Example of conflict between the pose estimated from AMCL (displayed by the particles positions), and the final Robot Localization output (displayed by tf frames).</i>	62

Figure 6.1	<i>Visual representation with Rviz of the obstacles on top of the Husky (red shape) used by MoveIt! planner: the Velodyne sensor (light blue shape), and the physical support for camera, GPS and Inertial Measurement Unit (IMU) (green shapes).</i>	66
Figure 6.2	<i>Representation of the reflection of light on a glossy surface: a ray incident on the surface is both scattered at many angles (diffuse reflection component), and reflected at one single angle (specular reflection component).</i>	74
Figure 6.3	<i>Convex hull computation of the edges of a sample image.</i>	74
Figure 6.4	<i>Different outcomes of the deployment validation algorithm: success in plant mode (Figure 6.4a), failure in plant mode (Figure 6.4b; the dispenser in the background does not induce a false positive because it does not satisfy depth requirements), success in nail mode (Figure 6.4c).</i>	76
Figure 6.5	<i>Visual representation of the dispenser availability check algorithm.</i>	77
Figure 6.6	<i>Visual representation of the dispenser availability check algorithm.</i>	77
Figure 6.7	<i>Different outcomes of the grasp validation algorithm: true positive (6.8a), true negative (6.8b), false positive (6.7c), error situation that however brings to a true positive (6.7d)</i>	78
Figure 6.8	<i>Two markers used for visual servo task in grasping phase. In Figure 6.8b both the markers are detected, while in Figure 6.8a the dispenser with ID= 0 is not visible due to specular reflection on marker surface, but the other marker is still detected so the visual servo task can proceed.</i>	80
Figure 6.9	<i>Example of the computed wings midpoint, used to compute the target position of the features (marker corners)</i>	80
Figure 6.10	<i>Some of the steps computed to identify the feature point tracked for dispenser application on the vine tree: switch from RGB to grayscale (Figure 6.10a), selection of pixels similar to the pixel corresponding to deployment point (Figure 6.10b), identification of branch area close to the deployment point (Figure 6.10c), identification of the circle including the selected points (Figure 6.10d).</i>	82
Figure 7.1	<i>Pose estimation in time using raw wheels odometry (Figure 7.1a), and using GPS (Figure 7.1b).</i>	86

Figure 7.2	<i>The odometry estimated by Robot Localization sensor fusion framework, fusing (among others) the pose estimated by the wheels (see Figure 7.1a) and by the GPS (see Figure 7.1b)</i>	87
Figure 7.3	<i>Output of mapping phase, using Gmapping as Simultaneous Localization And Mapping (SLAM) algorithm and Robot Localization as sensor fusion framework. The map of Garriguella vineyard in Figure 7.3a, the Casciana Terme one in Figure 7.3b.</i>	87
Figure 7.4	<i>Visualization through Rviz of the Husky performing autonomous navigation, with the local costmap overlaid to the satellite images of Garriguella vineyard. You can see the global (green line) and local (red line) navigation plans, the tf tree of the robot and the Husky footprint used by Move Base for the interaction with the costmap.</i>	88
Figure 7.5	<i>Evolution in time of the image coordinates of the corners of the tracked marker (see Figure 7.5a) during visual servoing driven deployment of a dispenser on a nail. In Figure 7.5b, the evolution in time of the module of the error on the tracked features.</i>	89
Figure 7.6	<i>Evolution in time of the image coordinates of the corners of the tracked marker (Figure 7.6a) during visual servoing driven grasping of the dispenser from the feeder. In Figure 7.6b, the evolution in time of the module of the error on the tracked features,</i>	90
Figure 7.7	<i>Evolution of the 3D (x, y, z) position of the end effector of the Jaco² during dispenser deployment using MoveIt!</i>	90
Figure 7.8	<i>Evolution of the z coordinate of the end effector, during grasping operation; visual servo driven movement is the central one in the graph.</i>	91
Figure 8.1	<i>The Robì mobile manipulator chassis.</i>	92
Figure 8.2	<i>3D rendering of different configuration of Robì base, obtained thanks to the chassis mechanical design.</i>	93
Figure 8.3	<i>HUB10GL, the in-wheel motor model adopted in Robì.</i>	94
Figure 8.4	<i>Some of the sensors Robì is equipped with: LMS291 (8.4a), and LD-MRS400001 (8.4b), both laser scanners from SICK; Trimble 5700 GPS transceiver (8.4c), from Trimble</i>	95
Figure 8.5	<i>The Nucleo board we used to communicate with the motors control board (Figure 8.5a), and the shield mounted on it to integrate IMU sensor (see Figure 8.5b).</i>	97

Figure 8.6	<i>The Robì platform, equipped with the sensors required by GRAPE project.</i>	98
------------	--	----

LIST OF TABLES

Table 4.1	Kinova Jaco ² product specification	36
Table 4.2	Mounted LIDARs comparison	38
Table 4.3	Move Base action specification	46
Table 4.4	Scan motion action specification	46
Table 4.5	Process point cloud action specification	47
Table 4.6	Dispenser Deployment action specification . .	47
Table 5.1	Robot localization configuration vectors . . .	52
Table 8.1	Ranges of Robì geometrical characteristics . .	93
Table 8.2	Robì LIDARs comparison	96

ACRONYMS

GRAPE	Ground Robot for vineyArd Monitoring and ProtEction
ROS	Robot Operating System
LIDAR	Laser Imaging Detection and Ranging
IMU	Inertial Measurement Unit
ECHORD++	European Coordinator Hub for Open Robotics Development Plus Plus
PCL	Point Cloud Library
API	Application Programming Interface
EKF	Extended Kalman Filter
SLAM	Simultaneous Localization And Mapping
AMCL	Adaptive Monte Carlo Localization

ABSTRACT

This thesis work is placed in the context of the [GRAPE](#) project, an agricultural robotics project launched by [ECHORD++](#) that aims to the creation of an autonomous mobile manipulator for automatic deployment of synthetic pheromones dispensers in the vineyards. These devices are used in vineyards for pest management techniques, as *mating disruption*. The team that carried out the project development included people from both Politecnico and Eurecat, a Catalan research center. The first part of the work focused on the design of the overall software architecture, together with the identification of subtasks and the interfaces between the software modules that implement them. The thesis work proceeded with the development of odometry system, autonomous navigation, and part of the manipulation components for the target platform. The challenges induced by the unstructured nature of the vineyard environment led, after several refining steps, to the creation of a sensor fusion system that exploits [IMU](#), GPS, and wheels velocities to estimate the robot pose in a mapless autonomous navigation system. On the other hand, the development of the manipulation part was complicated by the control flaws of the robotic arm we had at disposal, that we solved using techniques based on inversion of differential kinematics. The system underwent two one-week validation sessions on the field, carried out in vineyards in Casciana Terme (IT) and Garriguella (ES); during these sessions, we were able to both make the system more robust with respect to unexpected issues, and measure the performance of the system in its actual environment.

SOMMARIO

Questo lavoro di tesi è da collocare nel contesto del progetto **GRAPE**, un progetto di robotica agricola finanziato da **ECHORD++** che mira alla creazione di un manipolatore mobile autonomo per la distribuzione automatica di erogatori di feromoni sintetici all'interno dei vigneti. Erogatori di questo tipo vengono impiegati in alcune tecniche di disinfestazione, dette di *confusione sessuale*. Il team di sviluppo del progetto **GRAPE** era composto sia da membri del Politecnico di Milano, sia da membri di Eurecat, un centro tecnologico catalano. Inizialmente il lavoro di tesi si è focalizzato sulla progettazione dell'architettura software complessiva del sistema, attraverso l'identificazione di sotto-problemi e la definizione delle interfacce tra i diversi moduli software corrispondenti. Il passo successivo è stato la realizzazione del sistema odometrico, del sistema di navigazione autonoma, e di parte della logica del manipolatore per la piattaforma in oggetto. La struttura geometricamente indefinita del terreno e degli ostacoli nell'ambiente del vigneto ha portato, dopo svariati perfezionamenti, a un sistema di fusione sensoriale basato sull'utilizzo di **IMU**, **GPS**, e velocità delle ruote per la stima di posizione e orientamento del robot nel tempo, e un sistema di navigazione autonoma svincolato dall'uso di mappe. Lo sviluppo della logica del manipolatore, invece, è stato fortemente influenzato dagli oggettivi limiti di funzionalità del braccio robotico in dotazione, per aggirare i quali sono state sviluppate diverse tecniche basate sull'inversione della cinematica differenziale. Il sistema è stato anche sottoposto a due sessioni di validazione sul campo, da una settimana l'una: la prima a Casciana Terme (IT), la seconda a Garriguella (ES), in cui il sistema è stato reso più robusto rispetto a condizioni di lavoro non previste in laboratorio, e le performance della piattaforma sono state testate nell'effettivo ambiente di utilizzo.

INTRODUCTION

In this thesis we worked on the design and development of the software components of a mobile manipulator for vineyard monitoring and protection, and on extensive validation of the developed platform during two *on-the-field* sessions. The research and technology field this work belongs to is called *agricultural robotics*, and it's oriented to the introduction of Information and Communication Technologies in an innovative way as solution for agricultural challenges. The successful introduction of these technique could gradually turn traditional farming into precision farming, with consequently decrease of the chemical load in food and environment, and improvement of profits and yield for farmers.

GRAPE is an experimental project funded by European Coordinator Hub for Open Robotics Development Plus Plus (**ECHORD++**), a robotic research project which aims to promote the interaction between robot manufacturers, researchers and users. More specifically, the goal of the platform developed in **GRAPE** is to accomplish automatic distribution of synthetic pheromones for integrated pest control in vineyards. The teams that carried out the whole system design, development and testing across the 18 months project duration was composed by both people from Politecnico, and people from Eurecat, a Catalan research center.

The first, fundamental problem this thesis deals with is the development of odometry and navigation systems for the Unmanned Ground Vehicle at issue; even if those problems could seem not so complex, they are significantly different to solve with respect to the correspondent indoor problems, because of some highly destabilizing factors due to the vineyard environment: bumpiness, steepness and consistency of the soil, time-varying weather conditions and unpredictable vegetation. The presence of these factors makes the **LIDAR** sensors less reliable, and the pose estimation through wheel velocities almost untrustworthy.

The other main component of the platform treated in this thesis is the on-board manipulator, which main objective is to realize grasping and actual deployment of the pheromone dispensers on the vines. Additionally, it implements other tasks related to the validation of the aforementioned operations, and to the selection of the most suitable point for deployment. Main difficulties on this topic were due to the motion limits of the manipulator itself, in terms of both accuracy and mobility.

THESIS CONTRIBUTION

To solve the problem of odometry estimation in such a complex environment like the vineyard, that presents large part of typical issues of outdoor robot navigation, we created a sensor fusion system based on the Robot Localization package. The developed system exploits GPS, [IMU](#), and wheels sensed velocities to provide a very robust real-time estimate of the geolocalized pose and orientation of the platform. The traditional configuration of [ROS](#) navigation stack for indoor navigation (*i.e.*, localization through [AMCL](#) in a mapped environment) worked quite poorly in our context, mostly because of map repetitiveness, then we crafted our own navigation configuration. In our settings, the obstacles considered by the robot are the ones, added by hand in a *prohibition layer*, which represents the vineyard rows in both local and global costmap, and the ones detected by a frontal [LIDAR](#) in the local costmap only. This configuration, tested with *Timed Elastic Band* algorithm as local planner, demonstrated really good performances during *on-the-field* sessions.

In our work we also contributed to the development of the manipulation system, that essentially consisted in three different parts. Firstly, the smooth motion in front of the vine selected for deployment, in order to acquire the 3D point cloud of the vine with the [LIDAR](#) mounted on top of the end effector. Secondly, the grasping of the dispenser from the feeder, and its validation through computer vision techniques. Lastly, three different dispenser deployment modes, each followed by a validation procedure through computer vision techniques: an open-loop procedure to deploy on the vine, a visual servoing driven procedure to deploy on an expressly placed nail with the help of a fiducial marker, and a visual servoing driven procedure to deploy directly on the vine, with no help of fiducial markers.

STRUCTURE OF THE THESIS

This document is structured as follows.

- In Chapter 2, we provide an overview of the main technological tools we adopted to develop our solution, together with the theoretical concepts required to better understand them. In particular, we give an overview of the functioning of [ROS](#), the framework we used to develop the entire platform, and of some *off-the-shelf* packages we included in our project to implement the solution of sensor fusion, navigation, and manipulation problems. From the theoretical point of view, we explain basic concepts about pose estimation of the robot in the environment, and motion planning.

- In Chapter 3 we present a precise description of the scope and the involved partners of **GRAPE**, the experimental project this thesis is part of. Moreover, we present a collection of other European projects in agricultural robotics field in the last years.
- In Chapter 4 we describe the hardware and software architecture of the system implemented in the context of the **GRAPE** project. In the section dedicated to hardware architecture, we provide technical specifications and justification for the choice of mobile base, robotic arm, and sensors mounted on the robot, while in the section dedicated to software architecture we describe the main software modules of the implementation, and the interfaces between them. Finally, we precisely specify the contribute of this thesis to the overall project.
- In Chapter 5, we present our solution to the problem of odometry estimation and autonomous navigation in the unmanned ground vehicle at issue. Since both of them went through several stages, for each significant step we highlight the weaknesses that pushed for searching for a better solution, till the final configurations.
- In Chapter 6 we firstly list all the identified criticality of the used robotic arm, and the different motion techniques we used to overcome them. Secondly, we provide a fine-grained description of the manipulator specific tasks: the execution of the motion to properly scan the vine selected for deployment, different methods for the actual deployment of the dispenser, and validation of dispensers graping and deployment.
- In Chapter 7 we provide the results of extensive *on-the-field* validation of sensor fusion, navigation, and manipulation systems, executed during the *integration weeks* envisaged by the project schedule.
- In Chapter 8 we describe the configuration of Robì, a prototype small-sized mobile manipulator for agricultural application, as an alternative to Clearpath Robotics Husky as unmanned ground vehicle base for **GRAPE** project.
- In Chapter 9 we draw our conclusions about our work, and we propose some possible future improvements to the platform.

2

BACKGROUND AND TOOLS

In this chapter we are going to describe the general concepts this thesis deals with, together with the main tools we used to address the project. Since this thesis is in the frame of **GRAPE** project (see Chapter 3), most of them are typical of the robotic field and, more specifically, of the agricultural robotics. This last field should be seen in the wider context of the so-called *E-agriculture*; to give a precise definition of this term, we make reference to the FAO (Food and Agriculture Organization) definition¹:

E-agriculture, or ICTs in agriculture, is about designing, developing and applying innovative ways to use ICTs with a primary focus on agriculture. E-agriculture offers a wide range of solutions to agricultural challenges and has great potential in promoting sustainable agriculture while protecting the environment.

The **GRAPE** project, described with further details in Chapter 3, is about the design and realization of an unmanned ground vehicle with control and operative tasks in a vineyard environment, so in this chapter we deal with topics concerning software development in robotics, estimation of the state of a robot, autonomous navigation, and motion planning.

ROBOT OPERATING SYSTEM

ROS is the *robotic middleware* we used to develop the software components of the system described in this thesis. We decide to use it because of its great modularity, the availability of a very large number of packages, well documented Application Programming Interface (**API**)s and an active community. Moreover, **ROS** is a very widespread system, so its power and versatility are well known in the field of software development for robotics. Citing words from its official website², these are **ROS** main features:

It provides the services you would expect from an operating system, including hardware abstraction, low-level device control, implementation of commonly-used functionality, message-passing between processes, and package management. It also pro-

¹ http://www.fao.org/fileadmin/templates/rap/files/uploads/E-agriculture_Solutions_Forum.pdf

² <http://wiki.ros.org/ROS/Introduction>

vides tools and libraries for obtaining, building, writing, and running code across multiple computers.

ROS is actually a *meta-operating system*, that is, it's not an operating system in the traditional sense (it requires to be run on top of an another operating system; currently, the only officially supported OS is Linux Ubuntu), but it provides a peer-to-peer network that processes can use to create and process data together. This network is implemented through TCP, and it's called *Computation Graph*. In this section, we're going to describe **ROS** with more detail, with particular emphasis on the different techniques that nodes can use to communicate among them. Keep in mind that, even if these techniques differ significantly, they all are strongly typed, *i.e.*, in order to define a channel (with *channel* we now mean one of the technique that we are going to describe; it's not the name of a specific communication tool) you also need to define the types of message that are going to be exchanged through it.

ROS embeds a lot of useful message types (*e.g.*, *LaserScan.msg*, *PoseWithCovarianceStamped.msg*), grouped by domain (*e.g.*, *Sensor_msgs*, *Geometry_msgs*). However a simple message definition language is provided, and users are encouraged to define their own message types to make them as self-explanatory as possible. We provide now a list of the components of the Computation Graph, *i.e.*, **ROS** processing units, and possible tools at disposal to make them communicate.

ROS MASTER Even if the Computation Graph is a peer-to-peer network, a central process, called **ROS Master**, is required to exist, to provide naming and registration services to all the user processes in it. Once the processes have located each other through the services offered by the Master, they can communicate peer-to-peer without involving the central entity.

NODES The processes in the Computation Graph are called **nodes**, and they are the atomic units of the computational graph. The **ROS API** are available in C++, Python and Lisp, but C++ is the most widely used. One of the aims of **ROS** is to be modular at a fine-grained scale, so a complex task should be achieved through cooperation of several different nodes, each with quite narrow tasks, rather than one large node that include all the functionalities. Nodes can use different techniques for communication, depending whether the message is a part of data stream or it is a request message (*i.e.*, a response message is expected) and, in this last case, on the (expected) duration and complexity of the computation of the response.

TOPICS Topics implements a *publish-subscribe* paradigm, and they are the easiest way that nodes can use to communicate with each other. They basically are named channels, characterized by the

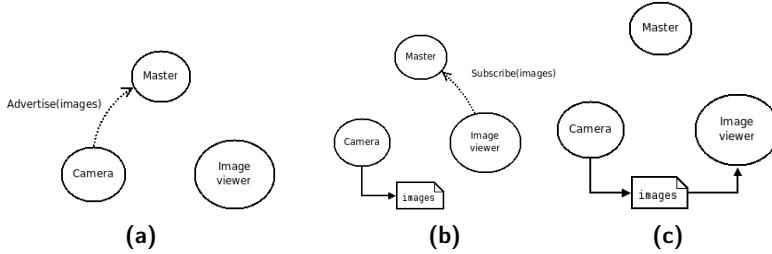


Figure 2.1: Three phases of the setup of communication of nodes through topics.

type of the messages that are sent through it. When a node *publishes* a message on a certain topic, the message is read by all the nodes that previously *subscribed* to that topic, interfacing with the Master. Note that:

- this technique leads to a strong decoupling between publishers and subscribers to a topic, because a publisher node is, from an high-level perspective³, not even aware of the presence of subscribers, and vice versa.
- the relationship between publishers and subscribers is *many-to-many*, i.e., multiple nodes can publish on a topic, and multiple nodes can subscribe to a topic.

We can easily conclude that this method is suitable for passing streams of data (e.g., the handler of a **LIDAR** streams its measurements over the network, or a node publishes the velocities commands for the wheels of a robot), but there is no notion of a *response* to a message, so it's not suitable for *request-response* communication.

SERVICES Services are defined by a name, and a couple of message types that describe the *request* type and the *response* type. Each service is offered by a Service server to any Service client that perform a call. So, Services implement an inter-node communication that is very similar to traditional function calling in most common programming languages (e.g., C++, Java), in the sense that:

- Service calls are blocking
- using Services, the inter-node communication is *one-to-one*

These properties make Services suitable for punctual (in opposition to data stream) inter-node communication, such as:

- request for parameters values
- request for frame acquisition, to a node that handles an RGB camera

³ Actually, publisher nodes always know the list of nodes subscribed to their topics. But this is only used in connection phase, and to avoid a situation where a node publish on a topic with no subscribers, for the sake of efficiency.

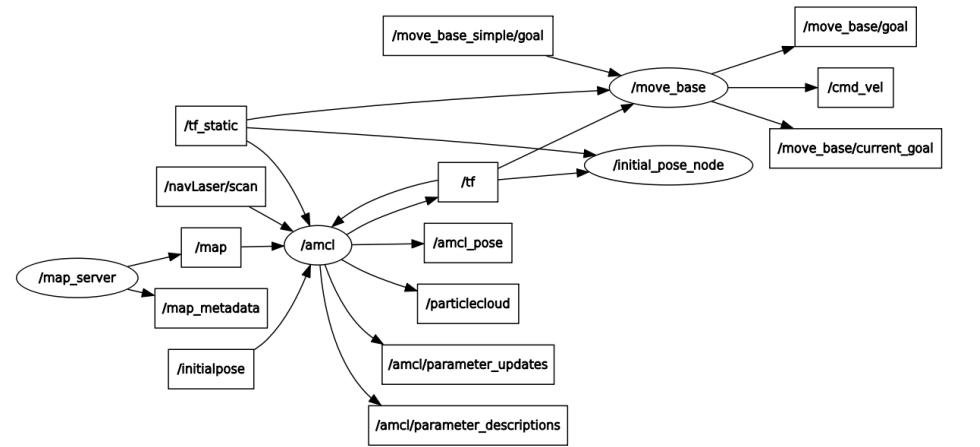


Figure 2.2: An example of the [ROS](#) Computation Graph, visualized with `rqt`: nodes are represented with circles, rectangles represent topics, and arrows go from a node to a topic it publishes on, or from a topic to a node subscribed to it. It's easy to recognize the many-to-many relationship.

- request for map cleaning, to a node that handles robot navigation

ACTIONS While Services, with their resemblance to traditional function calls, can address pretty well the problem of *one-to-one* inter-node communication, they can be quite unsatisfying if the computation required to produce the response is demanding in term of execution time (*e.g.*, navigation of a robot from one point to another in an environment), because the caller is stuck at the line with the Service invocation until the end of the procedure. Services show their weaknesses also in situations where it could be useful to observe the intermediate results of the computation triggered by the request (*e.g.*, a very complex manipulation procedure). **Actions** are very suitable in this context because, at the cost of a more complex implementation, provide an asynchronous and fully preemptable remote procedure call, with the possibility of monitoring intermediate results if needed, and a native exit status to check the state (active, rejected, preempted, succeeded, aborted...) of the execution. Differently from Topics and Services, Actions are not native in [ROS](#), and their functionalities are built on top of the other [ROS](#) messaging systems (See Figure 2.4). Asynchronicity is provided by the use of callbacks.

A final consideration about the [ROS](#) mechanism for recording, and playing back, data published on topics. This software module is called *rosbag*, and takes advantage of the fact that [ROS](#) is aware of all the messages exchanged through its messaging system. This tool gets very useful in a project like [GRAPE](#), where is not trivial at all to simulate or replicate in a laboratory environment the physical context of a vineyard, and of course meaningful data are required in order to

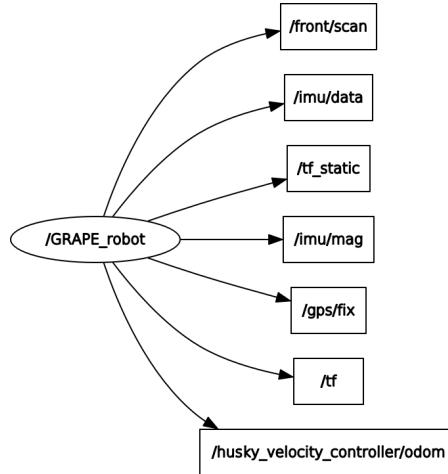


Figure 2.3: Visualization of the [ROS](#) Computation Graph with [rqt](#) in the context of a data bag being played: there is a single node (GRAPE_robot) that simultaneously plays recorded topics messages.

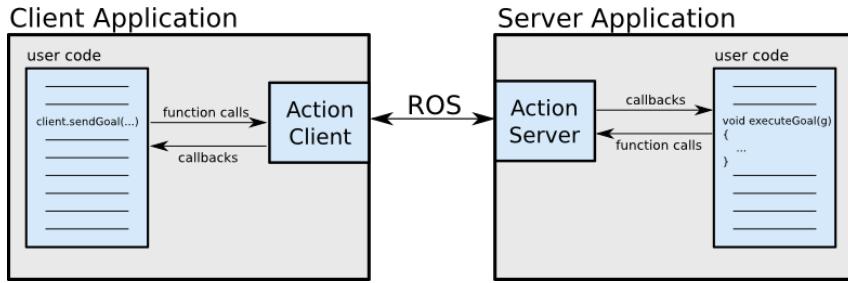


Figure 2.4: Sketch of the implementation of [actionlib](#), through [ROS](#) topics and a callback system.

correctly validate algorithms and tools.

The paradigm is very simple: in every moment you can invoke a `rosbag record` command, that records every single message exchanged on the [ROS](#) topics; data are logged in a single file with `.bag` extension, that can be lately filtered and/or compressed. Bag files can be played back in a very easy way invoking `rosbag play` command; the playing of the bag is implemented with a single node that streams messages on all the topics existing in the bag (see Figure 2.3). The module also provides an option to use *simulated time*, i.e., playing the bag exactly with the same message timing as in the real recorded session. In this situation, the decoupling induced by *publish-subscribe* is a strong advantage, since you can substitute all publishing nodes with a single one, with no consequences⁴.

⁴ Except in some pathological cases: for example, a node can require the list of active nodes in the computational graph, and the response would be different when the data are recorded and when the corresponding bag is played.

TF: THE TRANSFORM LIBRARY

tf is a ROS library, designed in the early phases of ROS development, to keep track of reference frames in a robotic system. Our system takes great advantage of *tf* usage to accomplish both navigation and manipulation tasks. The goal of *tf* is:

" [...] provide a standard way to keep track of coordinate frames and transform data within an entire system such that individual component users can be confident that the data is in the coordinate frame that they want without requiring knowledge of all the coordinate frames in the system" (Foote, 2013)

The utility of such a component is straightforward, even in quite simple robotic systems. We now describe a situation we stumbled upon exactly in the development of the GRAPE project, where of the utility of *tf* is very easy to understand; however, the context of this example is described in Chapter 6. In this example a LIDAR, mounted on top of the final joint of a robotic arm, acquires data while the arm is moving in order to create a point cloud that is going to be processed later. To get a meaningful point cloud, it's mandatory to keep track of the movement of the LIDAR with respect to a point with speed equal to zero (e.g., the base link of the arm, or the base link of the whole robot), and this gets even more difficult because of the multiple (6 in our specific case) joints of the arm; but this problem can be easily addressed by means of *tf*.

Frames are very useful for two main tasks:

- represent the configuration of the robot, by assigning frames to the significant physical elements of the robot (e.g., joints, sensors, wheels). Since *tf* graph is a tree, a root frame of the robot should exist; in typical configurations, the root frame name is *base_link*, and it's placed in the geometrical center of the robot. All other frames associated to physical components of the robot belong to *base_link* subtree.
- represent the position of the robot in an environment. A typical example is the frame *odom*, that is typically centered in the point where the robot is located when it's switched on. Another example is the frame *map*, that is used as a reference in case the robot is localized not only with respect to its initial point, but in a given map.

tf implementation relies on ROS topics (see Section 2.1) and achieves its goal by building an oriented graph where vertices are reference frames, and edges are transformations (rototranslations) between frames. *tf* does not assume a constant structure and, if a path exists between two reference frames in the graph, the direct transformation between

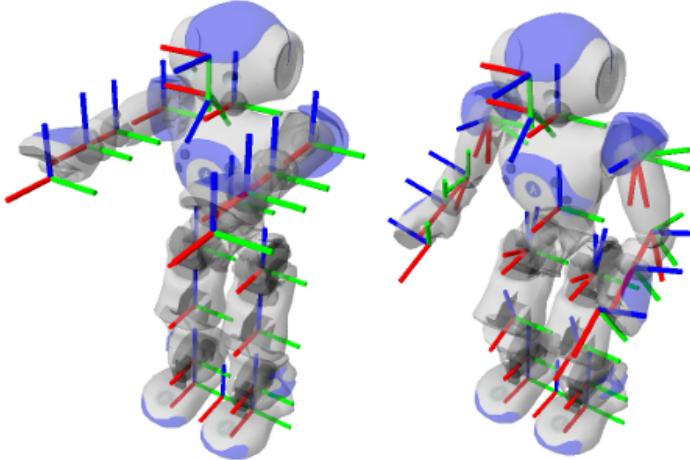


Figure 2.5: A robot in 2 different positions, with tf frames in evidence: x-axis is red, y-axis is green, z-axis is blue. The frames are the same in both configuration, but the transformations (i.e., rototranslations) between them are different.

them can be computed by composition of transformation. Since, in general, multiple paths between 2 vertices can exists in a directed graph and this could lead to ambiguity in computing the transformation between two reference frames, the graph is forced to be a tree, *i.e.*, acyclic. Disconnected subgraphs are allowed, but of course transformation between vertices that belong to different subgraphs cannot be computed. The main components of the library are:

- **tf broadcasters:** they are simple software components, that publish a transformation between two reference frames every time an update is available. Different broadcasters do not sync together the publishing phase.
- **tf listeners:** they are more complex components, because they take into account that broadcasters are not synced. Since both transformations and queries to *tf* graph are time stamped, listeners make use of queues to store the most recent transformations, and they interpolate old values using SLERP (Spherical Linear intERPolation) to return a transformation for which there is no measured value at the requested timestamp.

In Figure 2.5 you can see a graphical representation of the reference frames tracked in a Nao Robot, while Figure 2.6 shows an example of *tf* graph visualized with visualization framework *rqt*.

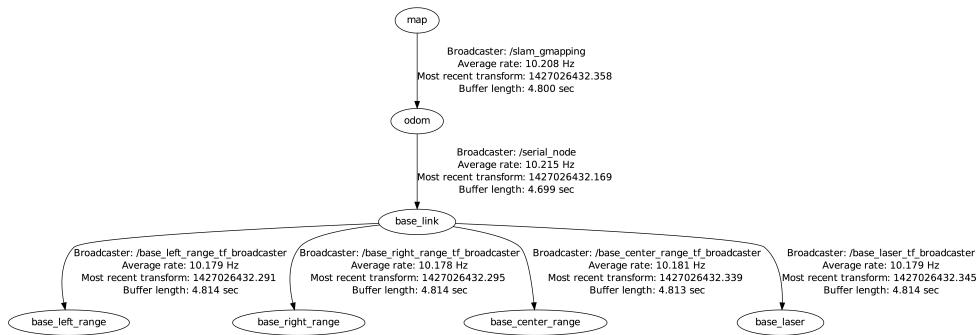


Figure 2.6: An example of tf tree

ODOMETRY

The problem of odometry, *i.e.*, estimation of the position of a robot in an environment, is harder than it could seem. Formally, odometry estimation is the problem of estimating over time the tuple:

$$\langle x, y, z, \theta, \dot{x}, \dot{y}, \dot{z}, \dot{\theta} \rangle \quad (2.1)$$

given the measurement of some motion sensors. To better understand the complexity of the problem, let's analyze an extremely simple model: a robot with a single, freely rotating wheel. In this frame, assuming rotary encoders on the wheel, we can think about measuring directly the wheel speed and integrate these measurements to get the traveled distance, and measure the variation in the orientation of the wheel to get the position. Unfortunately, a lot of imperfections exist, that may lead to errors in the outcome, for example:

- the wheel can be non perfectly perpendicular to the ground
- the friction between the floor and the wheel might not be enough to avoid slippage
- there is no such thing as a perfect sensor, so the use of encoders introduce an error

Even if all these concurrent causes seem negligible, you have to take into account that the errors sum up over time, so an error of a few millimeters per meter might become significant over time.

Moreover, the probability of slippage gets higher in systems with more than one wheel, because for a system with multiple wheels to move without slippage, a point must exists around which all the wheels can move along a circular path. This point is called Instantaneous Center of Curvature (see Figure 2.7), and can be easily identified by looking for the intersection of the axis of all wheels. If the wheels axis does not intersect in a single point, the system will not

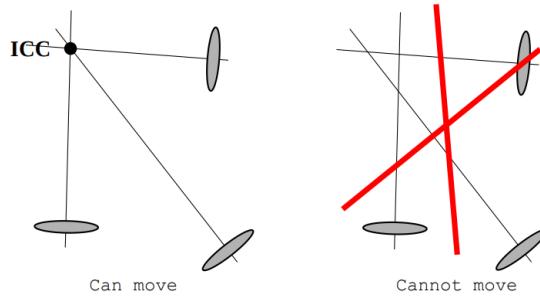


Figure 2.7: If the axis of all the wheels intersect in a single point, it's called ICC and the robot can move without slipping

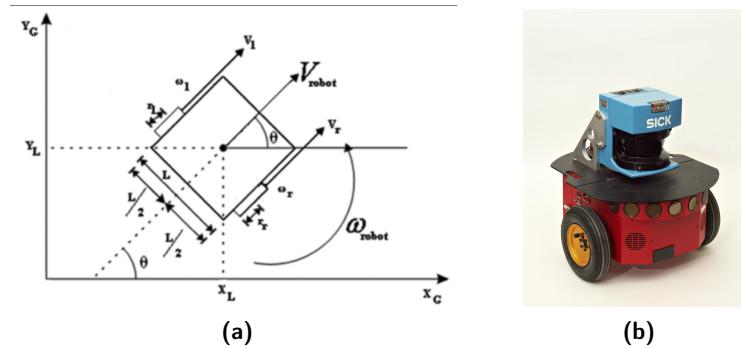


Figure 2.8: On Figure 2.8a, the scheme of a differential drive motion model; in Figure 2.8b, an example of a differential drive robot (Pioneer 3DX).

be able to move without slipping. However, even if the odometry estimated from the wheels is not a good solution if used alone, it can be used as a starting point for other, more complex, methods. For this reason we are now going to describe how to estimate the odometry starting from the wheels encoders in our specific robot. This computation is different according to the **motion model** of the considered robot. The robot we used is the Husky platform (see Figure 2.10) from Clearpath Robotics. It moves with a *skid steering* kinematics, that is a derivative of *differential drive* kinematics. Thus, we're going to describe these two motion model with more details.

Differential drive robot

In a differential drive system, the movement of the robot is only based on two separately driven wheels, placed on either side of the robot, on the same axis (see Figure 2.8), and optionally a central, non-actuated caster wheel for stability. The two side wheels are not steerable, so the changes of direction are realized through application of different speed to the two wheels. For example, intuitively, if the wheels move at the same speed and in the same direction, the robot will move straight; if the wheels move at the same speed but different

directions, the robot rotates in place. By recalling the definition of Instantaneous Center of Curvature, we observe that, if the wheels are correctly aligned, a differential drive robot always has a well-defined Instantaneous Center of Curvature and the slippage of the wheels is not very accentuated.

At each instant in time, since the Instantaneous Center of Curvature is well-defined, both the left and right wheel follow a path that moves around Instantaneous Center of Curvature at the same angular speed ω , and thus:

$$\begin{cases} \omega(R + \frac{L}{2}) = v_r \\ \omega(R - \frac{L}{2}) = v_l \end{cases} \quad (2.2)$$

$$(2.3)$$

where L is the distance between the center of the two wheels, v_r and v_l are, respectively, the linear velocity of the right and left wheel, R is the signed distance between the Instantaneous Center of Curvature and the midpoint of the wheels. Note that the only parameter constant through time is L , since it's a physical property of the robot structure, while all other parameters evolve during the movement.

By combining 2.2 and 2.3, we get:

$$R = \frac{L}{2} \frac{(v_r + v_l)}{(v_r - v_l)}, \quad \omega = \frac{(v_r - v_l)}{L} \quad (2.4)$$

Observing these results we can validate the intuitive impressions about particular cases made a few lines above:

- if $v_r = v_l$, the curvature radius is infinite, because the robot is moving straight.
- if $v_r = -v_l$, the robot is moving around the midpoint of the wheels.

We give now some details about odometry computation. Let's assume, in a certain moment $t = t_0$, that the robot pose is (x, y, θ) . We assume that in the time interval $t_0 \rightarrow (t_0 + \delta t)$ the values v_r and v_l are constant; if we observe Figure 2.9 under these condition, we have:

$$ICC = (x - R\sin\theta, y + R\cos\theta) \quad (2.5)$$

Write now the expressions for (x', y', θ') :

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \cos(\omega\delta t) & -\sin(\omega\delta t) & 0 \\ \sin(\omega\delta t) & \cos(\omega\delta t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x - ICC_x \\ y - ICC_y \\ \theta \end{bmatrix} + \begin{bmatrix} ICC_x \\ ICC_y \\ \omega\delta t \end{bmatrix}$$

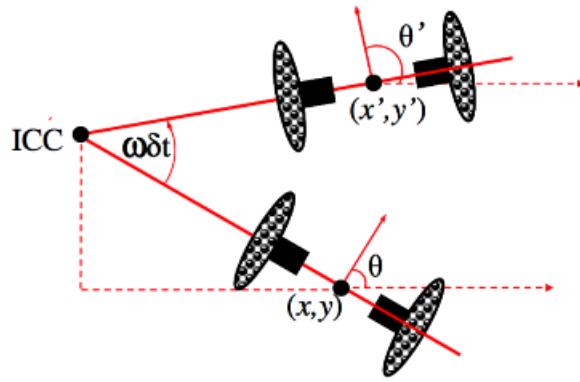


Figure 2.9: Differential drive: $(x, y, \theta) \rightarrow (x', y', \theta')$

With this procedure we have identified 3 of the elements of the target tuple (see Expression 2.1), but we still have to retrieve x, y and θ . For this reason we consider that by assuming an initial pose (x_0, y_0, θ_0) , knowing that:

$$V(t) = \frac{(v_r + v_l)}{2} \quad (2.6)$$

where $V(t)$ represents the overall speed of the robot, and assuming to know the functions $v_r(t)$ and $v_l(t)$, i.e., the linear speed of the wheel in time, we can calculate (x, y, θ) by integrating the speed of the robot over time, that is:

$$x(t) = \int_0^t V(t) \cos(\theta(t)) dt \quad (2.7)$$

$$y(t) = \int_0^t V(t) \sin(\theta(t)) dt \quad (2.8)$$

$$\theta(t) = \int_0^t \omega(t) dt \quad (2.9)$$

and, in our specific case we can write it as function of v_l and v_r , that are the quantities that are directly measured on the wheels.

$$x(t) = \frac{1}{2} \int_0^t (v_r(t) + v_l(t)) \cos(\theta(t)) dt \quad (2.10)$$

$$y(t) = \frac{1}{2} \int_0^t (v_r(t) + v_l(t)) \sin(\theta(t)) dt \quad (2.11)$$

$$\theta(t) = \frac{1}{L} \int_0^t (v_r(t) - v_l(t)) dt \quad (2.12)$$

So the tuple required by the odometry calculation is now complete.



Figure 2.10: Husky platform from Clearpath Robotics is the platform used for the development of [GRAPE](#) project.

Skid-steering kinematics

The Husky platform (see Figure 2.10) belongs to a different kinematic model, that relies on four (instead of two) actuated wheels. However, the motion model of the Husky is similar to differential drive because:

- wheels are not steerable
- being v_{f*} the speeds of the front wheels, v_{r*} the speeds of the rear wheels, v_{*r} the speeds of the right wheels, v_{*l} the speeds of the left wheels, we always have:

$$v_{fr} = v_{rr} \quad (2.13)$$

$$v_{fl} = v_{rl} \quad (2.14)$$

This type of motion model is called *skid steering*, and is often used in real-world applications (see Figure 2.11b) because of its simple and robust mechanical structure, which leaves more room in the vehicle for the mission equipment. In addition, it has good mobility on a variety of terrains, which makes it suitable for all-terrain missions. This motion model also presents a variety of problems and weaknesses. For example, if you recall what we told about Instantaneous Center of Curvature in Section 2.3, it's clear that the Instantaneous Center of Curvature of such a model will always be at the infinite, since wheels are organized in 2 parallel rows that cannot steer. Thus, skid steering robots can't turn without slipping of the wheels! From a theoretical point of view, the main consequence is a complexity in obtaining an accurate kinematics and dynamic model of *skid steering* (Yi et al., 2007). In the context of [GRAPE](#) project this leads to two major practical problems:

- the slippage of the robot leads to higher power consumption of the electric engines with respect to a system with explicit steering (Shamah, 1999). This is something to be taken into account in the sizing phase of the power system of the robot.

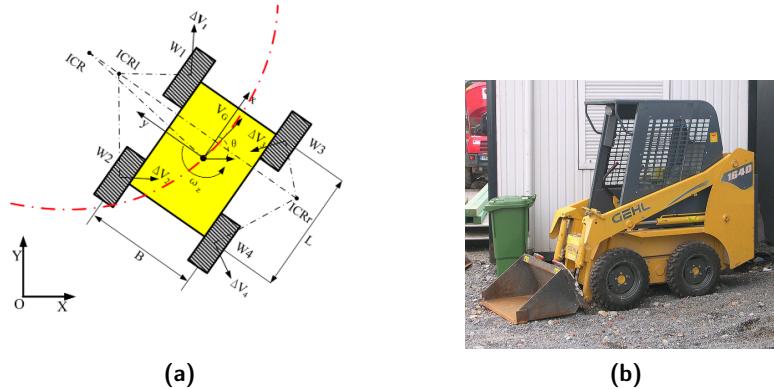


Figure 2.11: On Figure 2.11a, the scheme of a skid steering motion model; in Figure 2.11b, an example of a skid steering vehicle.

- the estimation of the odometry is going to be much more imprecise than in the differential drive case, for the error introduced by the slipping of the wheels.

Actually, there is a quite simple way (Wang et al., 2015) to model with an acceptable approximation the skid steering as an equivalent differential drive system, where the distance between the wheels is obtained multiplying the original distance for a factor χ that is function of the physical structure of the robot. This approximation relies on some assumptions:

- the mass center of the robot is located at the geometric center of the body frame.
- the two wheels of each side rotate at the same speed.
- **the robot is running on a firm ground surface, and four wheels are always in contact with the ground surface.**

We cannot guarantee that conditions 1) and 2) will hold in our context, but we are almost sure that condition 3) is **not** going to be verified, given the condition in which our robot is going to operate (vineyard terrain), so this is another reason for the wheel odometry not to be very precise. Thus, the integration of several sensors beyond the wheels encoders is essential to correct the wheel odometry.

SENSOR FUSION

First of all, clarify what we mean for sensor fusion, following the definition of Elmenreich, 2002:

"Sensor fusion is the combining of sensory data or data derived from sensory data in order to produce enhanced data in form of an internal representation of the process environment. The

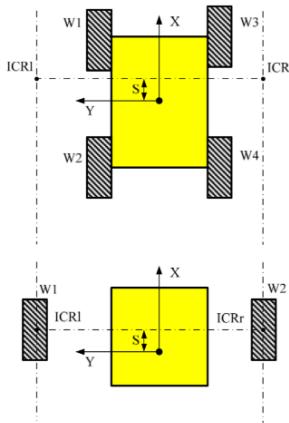


Figure 2.12: The equivalence of differential drive motion model and skid steering model according to Wang et al., 2015

achievements of sensor fusion are robustness, extended spatial and temporal coverage, increased confidence, reduced ambiguity and uncertainty, and improved resolution."

Let's analyze this last list, element by element, recalling the analysis of Remagnino, Monekosso, and Jain, 2011:

- *robustness*: redundancy generated by the presence of multiple sensors make the robot more resistant to partial failures.
- *extended spatial and temporal coverage*: disseminate sensors are more likely to measure the dimension concerned with temporal continuity and from several positions.
- *increased confidence*: measures confirmed by more than one sensor are given a larger weight.
- *reduced ambiguity and uncertainty*: joint information reduces the set of ambiguous interpretations of the measured value.
- *robustness against interference*: by increasing the dimensionality of the measurement space (e.g., measuring the desired quantity with optical sensors and ultrasonic sensors) the system becomes less vulnerable against interference.

Even if the name could be misleading, sensor fusion is different from *multi-sensor integration* in the sense that multi-sensor integration only consists in the simultaneous use of disparate sensor sources in order to accomplish a goal task. Sensor fusion techniques make a step further, and aim to the construction of a single common representation, using all the available sensor sources.

The difference between multi-sensor integration and sensor fusion is described graphically in Figure 2.13, to underline that sensor fusion

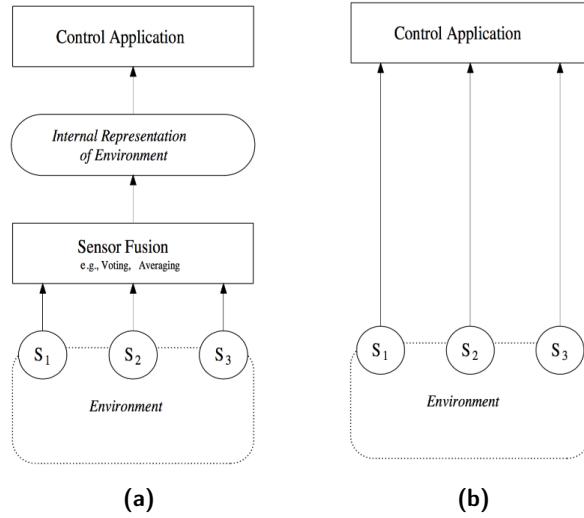


Figure 2.13: Sensor fusion 2.13a vs multi-sensor integration 2.13b

actually provides one single representation of the state of the environment that is used as a single input stream by the control application, while in multi-sensor integration the application uses each of the sensor data streams as direct input. In this context, we are assuming the definition of environment, from Thrun, 2002: pose of the robot, velocity of the robot, and velocity of the joints, configuration of actuators, location and features of surrounding objects, location and velocity of moving object and people.

In ROS ecosystem, several open source solutions exist that implements *sensor fusion-based* odometry; we tried two of them, which gave proof of good functioning in the past: **ROAMFREE** (Cucci and Matteucci, 2013, Calabrese, 2014, Cucci and Matteucci, 2014), a graph-based algorithm, and **Robot Localization** (Moore and Stouch, 2014, Dudek, Szymkiewicz, and Winiarski, 2016, Mohanty et al., 2016), based on Kalman filters.

After the experimental campaign in Casciana Terme, ROAMFREE turned out to be harder to configure for our specific problem, while Robot Localization had some specific documentation about its usage in the required context⁵, so we opted for this last one. Further details about our usage of Robot Localization are given in Chapter 5, while we are now giving a few hints about its general functioning.

Robot Localization

Citing Robot Localization documentation⁶:

⁵ http://docs.ros.org/kinetic/api/robot_localization/html/integrating_gps.html

⁶ http://docs.ros.org/kinetic/api/robot_localization/html/

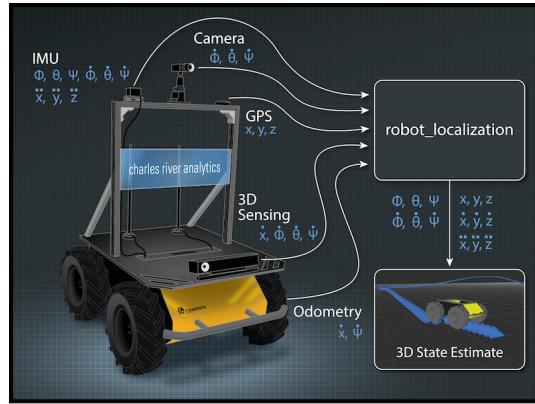


Figure 2.14: A scheme of Robot Localization sensor fusion

"Robot_localization is a collection of state estimation nodes, each of which is an implementation of a nonlinear state estimator for robots moving in 3D space. It contains two state estimation nodes, `ekf_localization_node` and `ukf_localization_node`". In addition, `robot_localization` provides `navsat_transform_node`, which aids in the integration of GPS data."

The Robot Localization algorithm is implemented specifically for [ROS](#), and it's widely adopted by [ROS](#) community for its documentation, and its easiness of use and configuration. It fuses an unlimited number of sensor sources, for each of which you can specify a configuration vector given in the frame id of the input message, *i.e.*, you can specify the message fields to be fused in the global estimate. For example, since GPS output is not so precise about altitude estimation, you can specify that only latitude and longitude parameters to be fused in the global estimate.

The state variables considered by Robot Localization are:

$$\langle x, y, z, \psi, \theta, \phi, \dot{x}, \dot{y}, \dot{z}, \dot{\psi}, \dot{\theta}, \dot{\phi}, \ddot{x}, \ddot{y}, \ddot{z} \rangle \quad (2.15)$$

where ψ, θ, ϕ correspond to the Euler angles *roll*, *pitch*, *yaw*. The node accepts as inputs several types of [ROS](#) messages:

- *nav_msgs/Odometry*: the odometry estimation output by another node can be used entirely as input; this is the case of the odometry estimated only using wheels encoders, and it also allows for cascaded localization nodes.
- *sensor_msgs/Imu*: the output of an [IMU](#) sensor, *i.e.*, a device composed by accelerometer, gyroscope, and magnetometer.
- *geometry_msgs/PoseWithCovarianceStamped*: an estimate of the position and orientation of the robot, together with the timestamp of the measure and its covariance matrix.

- *geometry_msgs/TwistWithCovarianceStamped*: an estimate of the linear and angular velocities of the robot, together with the timestamp of the measure and its covariance matrix.

Moreover, the Kalman filter implemented in Robot Localization is capable of *continuous estimation*: if for some reason no data are received from the various sensors for a large enough amount of time, the filter keeps producing odometry estimation exploiting an internal motion model. Robot Localization can be configured also to publish the *tf* transformation associated to the estimated odometry. This is one of the usage of *tf* described in Section 2.2.

NAVIGATION STACK

In Computer Science, a software stack is a set of programs that collaborate, in order to achieve a common goal. Since we are talking about ROS ecosystem, the programs are ROS nodes, and in the case of Navigation Stack, the common goal is to provide a modular out-of-the-box navigation system for unmanned ground vehicles. It was originally developed for Willow Garage's PR2 robot, but its usage can be easily extended to differential drive and holonomic wheeled robots. The planner, Move Base, is known to be a very good solution for indoor navigation (Marder-Eppstein et al., 2010), so we were not sure about its performance in an outdoor environment as required in GRAPE project. Luckily, it turned out to be a very good solution even in our situation, so it was integrated in the final navigation system. We are now giving a brief description of the main building blocks of the Navigation Stack (Figure 2.16):

ODOMETRY SOURCE A topic from which read the pose estimated from the odometry system, e.g., simple wheels odometry, output of a sensor fusion node as Robot Localization or ROAM-FREE.

SENSOR SOURCE A topic from which read the data coming from the laser sensors (LIDAR) that are probing the environment, for obstacle avoidance, localization and possibly mapping tasks.

AMCL Implementation of a probabilistic localization system, based on the Monte Carlo localization, as described in Fox et al., 1999.

MAP SERVER A ROS node where the map is published as a single topic message. This block is not used if Move Base is used in mapless mode.

BASE CONTROLLER This is the only output topic of the Navigation Stack, and of course it contains speed commands for the base of the robot.

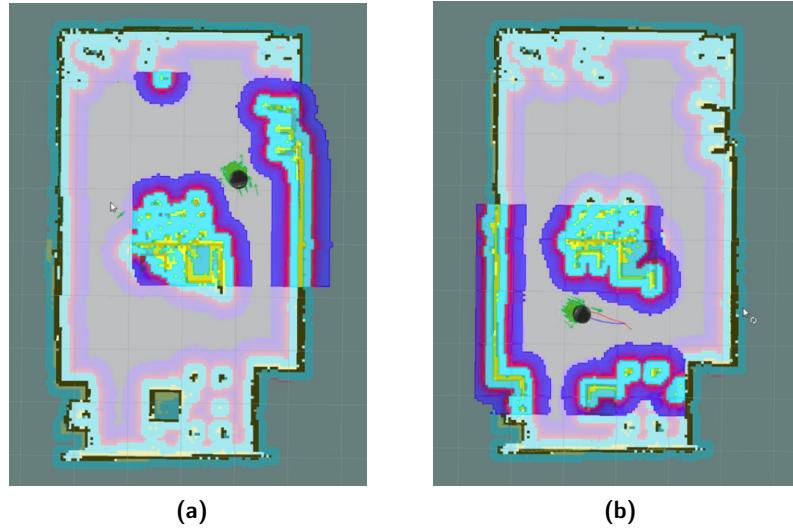


Figure 2.15: Local and global costmap visualized with Rviz. As you can see, the local costmap (in brighter colors) is built around the robot, and moves together with it.

LOCAL AND GLOBAL COSTMAP They represent the information about the obstacles on the 2D plane of the ground, with a certain inflation radius which represent the size of the robot base. Each cell of the gridmap is associated to a certain cost which measures "how hard is it" to traverse that cell of the gridmap; possible values to represent the severity of obstacles are *Lethal obstacle*, *Inscribed*, *Possibly circumscribed*, *Freespace*, *Unknown*. The **global** costmap represent whole environment as is build from a known map, while the **local** costmap is built using only incoming laser scans. Local map is, in general, a scrolling window that moves in the global costmap in relation to robot current pose, and it's continuously updated.

LOCAL AND GLOBAL PLANNERS Global planner takes as input the global costmap, and traces the path with lowest cost from current position of the goal position; the local planners instead takes care of continuously update (if required) the global plan in the light of the incoming laser measurements. This combination is essential in case of unexpected obstacles.

MOTION PLANNING

Motion planning is a term used in robotics for the process of breaking down a desired movement task into discrete motions that satisfy movement constraints, and possibly optimize some aspect of the movement. This process is essential for autonomous systems, that accepts high-level description of tasks and should require no further

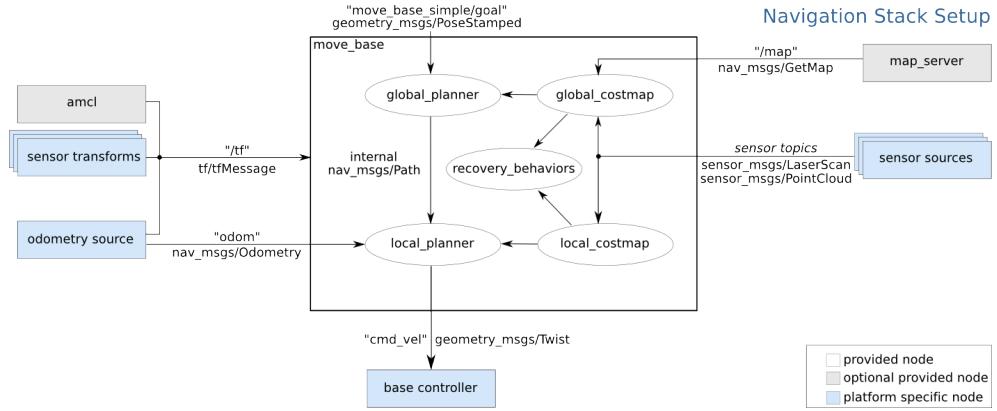


Figure 2.16: A scheme representation of the Navigation Stack; see the color legend for classification in provided, optional provided, and platform specific nodes.

human intervention; in this way the user can focus on *what* he wants done, instead of *how* to do it (Latombe, 2012). Motion planning can be applied to any kind of mechanical device equipped with actuators and sensors under the control of a computing system (e.g., a robotic arm that perform *pick-and-place*, a differential drive robot that has to reach a goal in an environment). Motion planning includes (among other): obstacle avoidance, computation of collision-free paths, building reliable sensory-based motion strategies.

More concisely, the basic problem of motion planning is, given:

- a start pose of the robot
- a desired goal pose (position, orientation)
- a geometric description of the world
- a geometric description of the robot (both as geometric shape, and as kinematic constraints due to kinematic motion model of the robot)

finding a path that moves the robot gradually from start to goal while never touching any obstacle.

Even if the movement of the robot has to be executed in the real world, the planning of the motion is typically computed in the *configuration space*, or *C-space* (Lozano-Perez, 1983). Formally, let: the robot A , at a certain position and orientation, be described as a compact subset of $W = \mathbb{R}^n$, $N = 2$ or 3 , and the obstacles B_1, \dots, B_n be closed subsets of W . In addition, let F_a and F_w be Cartesian frames embedded in A and W , respectively. F_a is a moving frame, while F_w is fixed. A **configuration** q of A is a specification of the position T and the orientation Θ of F_a with respect to F_w . The **configuration space** of A is the space C of all the configurations of A . Usually, configuration space is high dimensional; a configuration is expressed as a vector of

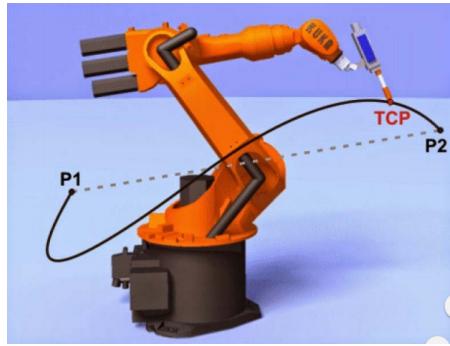


Figure 2.17: A graphical representation of the trajectory planned for moving a robotic arm from start (P_1) to goal pose (P_2), satisfying geometrical constraints of the robot.

positions and orientations, so the robot in configuration space is always represented as a point. We distinguish the configuration space from the *workspace*, that is the physical environment in which the robot move. A few examples:

- the robot is a single point and the workspace is a 2-dimensional plane; C is a plane, and a configuration can be represented using two parameters (x, y) .
- the robot is a 2D shape that can translate and rotate, the workspace is a 2-dimensional plane; C is 3-dimensional and a configuration can be represented using 3 parameters (x, y, θ) .
- the robot is a solid 3D shape that can translate and rotate, the workspace is 3-dimensional; C is 6-dimensional, and a configuration requires 6 parameters: (x, y, z) for translation, and Euler angles (ψ, θ, ϕ) .
- the robot is a fixed-base manipulator with N revolute joints and no closed-loops; C is N -dimensional.

Note that, as physical workspace, the configuration space is splitted in free space (C_{free}), and obstacle space (C_{obs}). From a practical point of view, C -space can be drawn by virtually sliding the robot shape along the edge of the obstacle regions, inflating them of the area covered by the robot. You can see some graphical examples of this procedure in Figure 2.18.

In simple configuration spaces, *grid-based search* algorithms can be used. With this approach, a grid is overlaid to configuration space, and assume that each configuration is identified with a grid point. With this approach, the problem is reduced to a graph search and exact algorithms like Dijkstra and A^* , or suboptimal algorithms like A^* , ARA^* , AD^* can be used.

But this approach gets easily unfeasible (for example, A^* has complexity $O(b^d)$ (Hart, Nilsson, and Raphael, 1968), with b branching

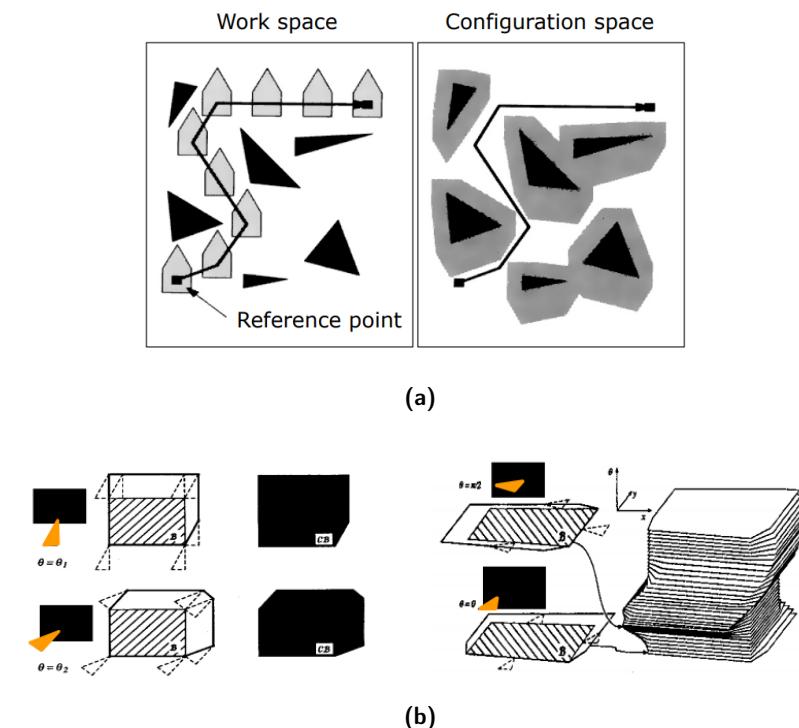


Figure 2.18: The graphical construction of C-space from workspace, by sliding robot shape along the borders of obstacle regions. In Figure 2.18a you can see the procedure in 2D, in Figure 2.18a in 3D. Note that in 3D you only need to compute 2D procedure for each θ , and then stack all obtained images.

factor and d depth of shortest path), so **sample-based approaches** are currently state-of-the-art planning algorithms. Sampling bases approaches, in a nutshell:

- are more efficient in most practical problems but offer weaker guarantees.
- are probabilistically complete: increasing computing time, the probability tends to 1 that a solution is found if one exists (otherwise it may still run forever).
- performance degrades in problems with narrow passages.

In sample-based planning there isn't an explicit characterizing of C_{free} and C_{obs} , but only a collision detection algorithm that probes C to see whether a certain configuration lies in C_{free} or not. An example of sample-based planning algorithm is **Rapidly exploring random trees** (LaValle, 1998); its pseudocode is shown in Algorithm 1.

Algorithm 1 RapidlyExploringRandomTrees(q_{goal})

```

1: G.init(q0)
2: iterNumber ← 0
3: N ← 100                                ▷ for example
4: repeat
5:   if iterNumber mod N ≠ 0 then
6:     qrand ← RANDOM_CONFIG(C)
7:   else
8:     qrand ← qgoal
9:   qnear ← NEAREST(G, qrand)
10:  G.add_edge(qnear, qrand)
11:  iterNumber ← iterNumber + 1
12: until qnear = qgoal

```

In the context of **GRAPE** project, motion planning is an essential component in two different modules:

1. Navigation module: **Move Base**, described in Section 2.5 is a motion planner that makes use of grid-based planning algorithm, and is used for the navigation of the robotic base in the vineyard environment
2. Manipulation module: **MoveIt!**⁷ is a state-of-the-art sample-based planner, and is used to plan the motion of our robotic arm in manipulation tasks (see Figure 2.20 for a block scheme of *MoveIt!* architecture).

⁷ <http://moveit.ros.org/>

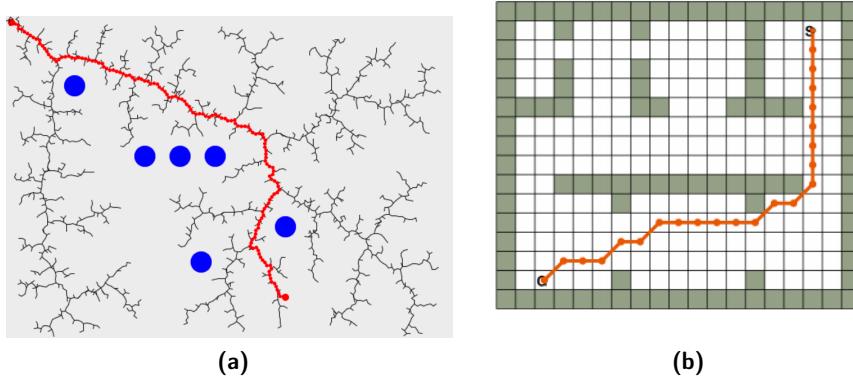


Figure 2.19: Graphical representation of the ways of proceeding of a sample-based planner (2.19a), and of a grid-based planner (2.19b)

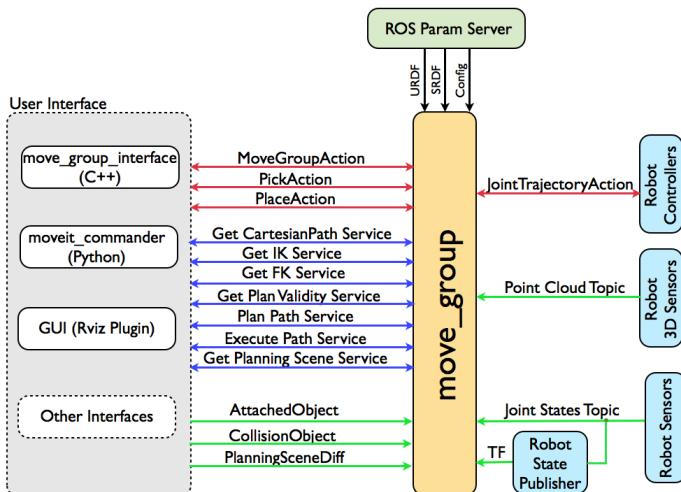


Figure 2.20: A block scheme representing the high-level MoveIt! architecture.

THE GRAPE PROJECT

In this chapter, we are going to give a description of the project this thesis work is part of, with particular emphasis on the parts that were specifically addressed in the thesis work. For the description of the project, we make reference to its official proposal.

PROJECT DESCRIPTION AND GOALS

GRAPE is an experimental project of **ECHORD++** and, as hinted in the first part of Chapter 2, it focuses on vineyard farming activities and aims at setting up a robotic manipulation platform able to support lead users to develop a variety of farming applications. The main goal of **GRAPE** is not the complete development of an industrial platform, but, coherently with the research scenario, the primary objectives are:

1. the development of example applications in the pre-mentioned context, exploiting and improving the so-called *key enabling technologies*¹ in robot navigation, perception and manipulation of the project partners. The resulting capabilities are likely to allow to make it easier for small and medium enterprises working in the fields of agricultural robots and, more in general, plant protection.
2. increase of robot acceptance by farmers and agronomists. Since the very last goal of this project is not solely about pure research but about the enterprise world too, particular attention is given into the realization of a robotic platform that could be accepted by potential end users. For this reason, a constant interaction between the project partners and the potential stakeholders (*i.e.*, vinegrowers) is maintained also in development phases.

The example applications have been selected in order to challenge perception and action capabilities, and to achieve vineyard monitoring, navigation and manipulation tasks.

This decision is taken in the scope of turning traditional farming into precision farming; the realization of such a turn would allow for both the decrease of the chemical load in food and environment, and

¹ Key Enabling Technologies are a group of six technologies: micro and nanoelectronics, nanotechnology, industrial biotechnology, advanced materials, photonics, and advanced manufacturing technologies. They have applications in multiple industries and help tackle societal challenges; they are considered crucial in the creation of advanced and sustainable economies.



Figure 3.1: A vine before (3.1a) and after (3.1b) having leafed out.

an improvement of profits and yield for farmers, that would get a return for their investment (Herring, 2001). The introduction of precision farming techniques leads indeed to a lot of advantages, for example early detection of plants diseases, or application of pesticides and fungicides with high precision and only when/where needed.

It's clear that some of these high-precision tasks are still too complex to be automated, and current state-of-the-art in research and technology have been proved to be not yet mature to give rise to a commercial product able to compete with a skilled agricultural agent. The goal of **GRAPE** is creating the enabling technologies to allow agricultural companies to develop vineyard robots to keep filling the gap that exists with respect to traditional methods.

GRAPE project specification identifies these two application examples:

- **Vineyard monitoring and autonomous navigation:** the developed robotic platform must be able to autonomously navigate the vineyard and localize itself into a map of the environment (downstream a mapping phase), to be able to monitor the state of the vineyard (*e.g.*, foliage and grapes inspection). The localization and navigation parts, however, are a key point for any kind of tasks for an unmanned ground vehicle field robot, because of the probability of high slope ground, rugged terrain morphology and unstructured map. We'll see the characterization and analysis of the problem of navigation in a rugged outdoor environment in chapter 5.
- **Autonomous application of pheromone dispensers:** pheromone dispensers (see Figures 3.3, 3.2) are used for *mating disruption* techniques, to protect grapevines from grape moths, by disrupting the bugs' reproductive cycle making use of synthesized sex pheromones. Also in this case, the most relevant challenge is brought by the unstructured environment, that makes the sensing and manipulation tasks significantly harder. However, note that pheromones deployment task gets easier if you think that the timing of the reproductive cycle (on which we

have no control at all) forces this operation to be performed in a season where grape plants are pruned, and have not leafed out yet (Carde, 1995). Of course, the absence of foliage and grapes makes this task significantly easier (see Figure 3.1).

For what concern the second goal of the project, *i.e.*, a major commitment about the acceptance of the robotic system in a field that's still strongly led by traditional methods, its realization consists mostly in the creation of an user-friendly interface (possibly for smartphones or tablets) for the GRAPE system, that allows an user to:

- constantly and easily visualize the position and activity of the robot.
- perform supervisory control on the platform, *e.g.*, selection of the plants for the dispensers deployment.
- in case of a failure of the unmanned ground vehicle in an autonomous task, exploit a teleoperation interface to carry out the operation.

The whole project also has to be designed with an adequate degree of modularity, in order to support a possible extension to a fleet of robot able to operate in parallel. Note that, as described in Section 2.1, this goal gets easier by the usage of ROS framework.

The GRAPE projects involved three partners, each with different assignments, and different responsibilities in the project context: PoliMi, Eurecat (a technology center located in Catalonia, that operates in many industrial and research fields, including robotics), and Vitirover (an agricultural robots company, located in France). To make the different teams interact and test the compatibility of the different parts of the system, some sessions of integration and tests, called *integration weeks*, have been carried out within the duration of the whole project. In this thesis we are going to focus mostly on the experiments and results carried out during the last two GRAPE integration weeks, held in January and March 2018.

CONTEXTUALIZATION OF THESIS WORK IN THE PROJECT

Since the timespan of the GRAPE project as a whole covers 18 months, we took advantage of the preceding work described in Serrano et al., 2017, that is a description of the ongoing work around the GRAPE project. As stated in the paper, some results were already accomplished in a few subjects:

- in-depth analysis of the main challenge and obstacles in the project, and requirement extraction



Figure 3.2: Different shapes of pheromone dispensers available on the market. The model used in GRAPE is the one depicted in image 3.2b



Figure 3.3: Detail of the dispenser type used in GRAPE.

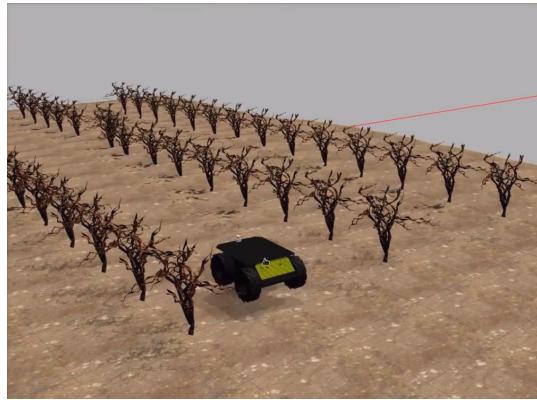


Figure 3.4: Simulation of the Husky robot in the vineyard environment. Simulation built using Gazebo simulation software.

- identification and obtainment of the specific hardware components (mobile platform, robotic arm, LIDARs, IMUs, cameras, network equipment, computational unit)
- segmentation of the macro objectives into smaller subtasks
- results of intermediate experiments carried out during the first integration week in Garriguella, Spain. The experiments mostly concern the mapping, localization and navigation tasks
- creation of a simulated vineyard and Husky environment using Gazebo simulation tool (see Figure 3.4)
- collection of data generated during the same integration week, making use of ROS data logging tool (as described in Chapter 2.1)
- creation of a 3D printed mockup vine tree (see Figure 3.5), to make the laboratory work more efficient and meaningful for the development and validation of the parts that would interact directly with the vine trees (*i.e.*, dispenser application task). The mockup was created from the laser scans of a real vine tree in Figueres (ES).

The problems tackled in this thesis are a subset of the ones assigned to PoliMi in the project proposal. The problems we faced are the following, and will be described in depth in next chapters:

- design of the overall software architecture in the ROS framework (see Chapter 4)
- revision and enhancement of the mapping, localization and navigation systems (see Chapter 5)
- design and part of the implementation of the pheromone dispenser application (see Chapter 6)



Figure 3.5: The 3D printed vine tree mockup, used in validation phase of the dispenser application task algorithm.



Figure 3.6: Other field robots developed in EU projects in the last 6 years: [3.6a](#) *Rhea*, [3.6b](#) *CROPS*.

RELATED PROJECTS

We are now presenting a short list of EU projects related to robotics for precision farming, with direct or indirect application to vineyards. All these projects have been developed in the last 6 years, and demonstrate the growing interest of the scientific community and potential end users to this application and research field.

VINEROBOT Automated measurement and monitoring of parameters such as grape yield, vegetative growth, water status and grape composition in vineyards by use of an unmanned ground vehicle (see Figure [3.7a](#)) endowed with artificial intelligence with sensing technologies. Chlorophyll-based fluorescence, RGB machine vision and thermography technologies are used to monitor vineyard on-the-go (Diago, Tardaguila, et al., [2015](#)).

VINBOT Creation of an all-terrain autonomous mobile robot (see Figure [3.7b](#)) with a set of sensors capable of capturing and analysing vineyard images together with 3D data, by means of cloud computing applications, to determine the yield of vineyards and to share information with the winegrowers, in order to mix

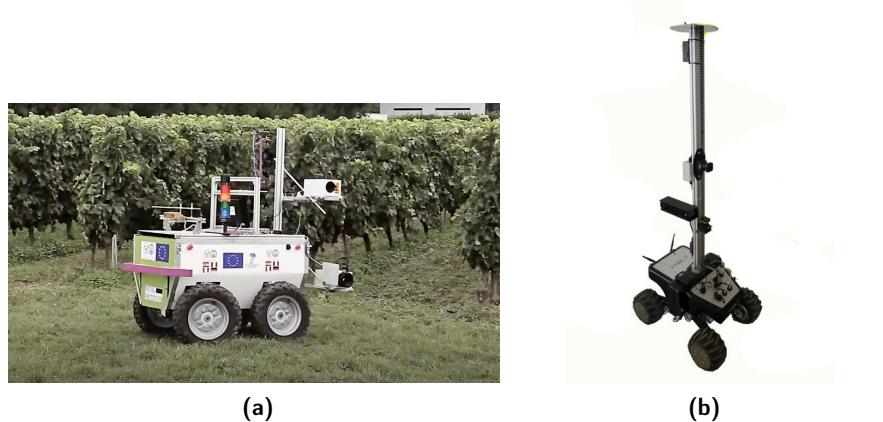


Figure 3.7: Other field robots developed in EU projects in the last 6 years: 3.7a Vinerobot, 3.7b Vinbot.

the grapes of homogeneous quality to efficiently market a range of wines by quality and price. (Lopes et al., 2017).

RHEA Design, development, and testing of a new generation of automatic and robotic systems for both chemical, mechanical and thermal effective weed management focused on both agriculture and forestry, and covering a large variety of European products including agriculture wide row crops and forestry woody perennials. RHEA aims at diminishing the use of agricultural chemical inputs in a 75%, improving crop quality, health and safety for humans, and reducing production costs by means of sustainable crop management using a fleet (see Figure 3.6a) of heterogeneous robots, both ground and aerial, equipped with advanced sensors, enhanced end effectors and improved decision control algorithms (Santos, Ribeiro, and Fernandez-Quitanilla, 2012).

CROPS Design and development of an highly configurable, modular and clever carrier platform that includes modular parallel manipulators and intelligent tools (sensors, algorithms, sprayers, grippers) for high value crops like greenhouse vegetables, fruits in orchards, and grapes for premium wines. The CROPS robotic platform (see Figure 3.6b) is capable of site-specific spraying (targets spray only towards foliage and selective targets), selective harvesting of fruit (detects the fruit, determines its ripeness, moves towards the fruit, grasps it and softly detaches it), reliable detect and classificate obstacles and other objects to enable successful autonomous navigation and operation in plantations and forests (Oberti et al., 2014). Further development in CROPS projects also includes Sweeper project (Ringdahl et al., 2016).

4

GRAPE HARDWARE AND SOFTWARE ARCHITECTURE

In this chapter we are going to describe in a detailed way the architecture of the robotic platform we used to implement the system to fulfill the requirements described in Section 3.1. First, in Section 4.1, we give an overview of the hardware components, from the robot base to all the actuators/sensors present on board. In Section 4.2, we analyze the software architecture of the system, explaining what the main modules are, and the relationships between them.

GRAPE HARDWARE ARCHITECTURE

Robotic Base

The choice of the robotic base is, of course, a crucial point in the development of our system, for two reasons:

- without a well-functioning robotic base, no other objective can be achieved.
- the vineyard environment is particularly challenging in nature (this particular aspect is going to be well-justified in Chapter 5), thus a non-prudent choice could be very dangerous.

As anticipated in Section 2.3, the choice fell on the **Husky** unmanned ground vehicle from Clearpath Robotics¹, (see Figure 2.10). There are several reasons for the adoption of this specific model:

- it's a widely used platform (e.g.: Madhavan et al., 2015; Hinkel et al., 2015; Ostafew, Schoellig, and Barfoot, 2013).
- it has a *skid steering* kinematics that, as already seen in Section 2.3, can be essentially reduced to a *differential drive* kinematics, that is very simple.
- it's specifically designed for outdoor use, so robustness and ability to deal with unstructured terrain are problems addressed by its rugged construction and high-torque drivetrain.
- it offers a large payload capacity, to host the multitude of sensors, actuators and computational units we need.
- it's fully supported in ROS, with open source drivers, configuration files and examples.



Figure 4.1: Husky platform after a navigation session in environment. You can easily note the amount of stones and mud stuck into the wheels.

This choice turned out to be substantially good, since the ROS drivers and configuration files were actually pretty easy to use and modify. The only major weakness identified in the Husky is related to its behavior on muddy ground, and is due to the wheels shape. After a few minutes of navigation in the mud, the unmanned ground vehicle tends to collect a significant amount of earth and stones in the wheels grooves (see Figure 4.1) and, consequently, lose grip on the ground. This problem could be attenuated or solved by use of tracks instead of wheels.

Robotic Arm

By observing how pheromone dispensers must be deployed on the vine plant in Figure 3.2 (in Figure 3.2b, you can see the dispenser shape that we actually used for GRAPE project; in Figure 3.3, a close-up of our dispenser type), it's easy to understand that we require a very flexible robotic arm, with advanced movement capacity and capable of precise movement. Additional constraints also come from the limited space available on the Husky base, and from the limited power supply at disposal aboard. The choice fell on *Jaco²* arm from Kinova Robotics² (see Figure 4.2a, Table 4.1 for product specifications), for these reasons:

- 6 DOF provide the required flexibility
- 3 fingers can be used for the dispensers deployment.
- while programmatic control is necessary for GRAPE purpose, the joystick control makes it easy to perform quick tests and analyze the robot capabilities

¹ <https://www.clearpathrobotics.com/husky-unmanned-ground-vehicle-robot/>

² <http://www.kinovarobotics.com/wp-content/uploads/2017/06/JACO%C2%80B2-User-Guide-Assitive-Robotics-April-2017.pdf>

PARAMETER	VALUE
Degrees of Freedom	6
Weight	5.2 Kg
Payload	1.3 Kg
Reach	90 cm
Maximum Linear arm speed	20 cm/s
Power supply	18 ÷ 29 VDC
Average Power	25W
Peak power	100W
Communication Protocol	RS485
Material	Carbon fiber

Table 4.1: Product specifications of Jaco² from Kinova Robotics.

- unlimited joints rotations.
- provided mechanical support for sensors mounting on top of the arm
- reduced weight and power consumption
- small overall dimensions
- [ROS](#) open source drivers already provided by Kinova Robotics

Some of the nice features of this device (weight, dimensions) actually derive from the original purpose of the arm, that is in the context of assistive devices for disabled people, to improve quality of life and independence of people on electric powered wheelchairs (see Figure 4.2b). But this fact also brought some unexpected weaknesses, that introduced a series of quite serious problems. The identified weaknesses are:

- a very large imprecision in the movements piloted via programmatic control: this is probably due to joystick control seen as a "main control mode" since, as previously stated, the arm is born as an assistive device. This weakness was compensated partially modifying the open source drivers of the arm, and partially by the introduction of different control strategies. The magnitude of the error, measured as distance between the desired cartesian position of the end effector and the actual one, is 0 ÷ 3 cm after the changes to the [ROS](#) driver of the arm, and it was 0 ÷ 15 before it.

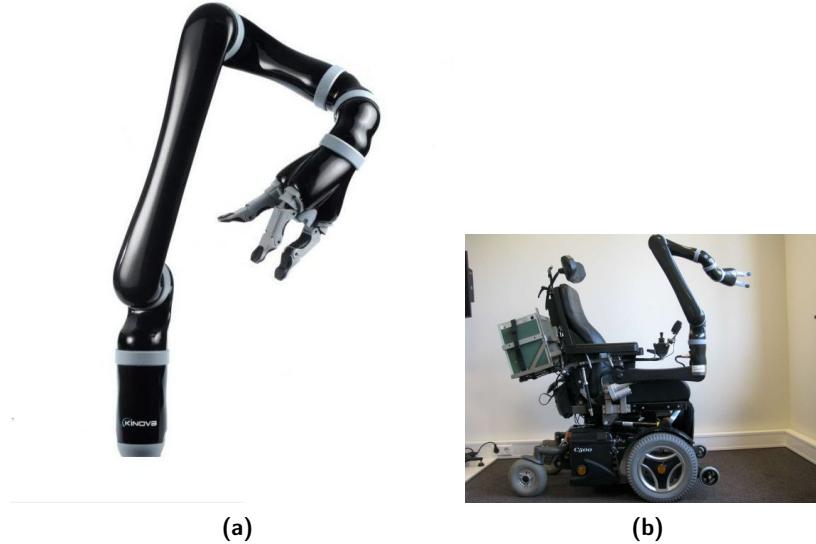


Figure 4.2: Robotic arm Jaco² with 3 fingers, from Kinova Robotics ([4.2a](#)) and the same arm mounted on a wheelchair [4.2b](#)



Figure 4.3: The final configuration of the sensors (LIDAR and RGB-D camera) mounted on top of the end effector.

- buggy implementation of joints constraints, probably due to the same cause of the previous point. This created a major problem mostly because of the wiring used for the sensors (described in next sections) mounted on the arm, that kept entangling because of the unlimited joints rotation capability.

After the analysis of the macro goals of **GRAPE** project, it was decided to use the arm also as a support for other sensors for the dispenser deployment task: a **LIDAR** and an RGB-D camera (see Figure [4.3](#)).

LIDARs

The final robotic system mounts three different **LIDARs**: two in front of the Husky for visualization and navigation purposes, and the other one mounted on a metal support on top of the arm end effector. The

	VLP-16	TIM561	UTM-30LX-EW
Number of Channels	16	1	1
Scan Angle	360°	270°	270°
Rotation rate	5Hz 10Hz 20Hz	15 Hz	40Hz
Angular Resolution	0.1° 0.2° 0.4°	0.33°	0.25°
Range	100m	10m	30m
Power Consumption	8W	4W	<8W
Weight	590g	250g	210g

Table 4.2: Comparison of Husky on-board LIDARs



Figure 4.4: Three LIDARs mounted on our unmanned ground vehicle: Tim561 (4.4b) from SICK, Puck Lite (4.4a) from Velodyne, UTM-30LX-EW (4.4c) from Hokuyo.

purpose of this last LIDAR is explained in Chapter 6; for now, all you need to know is that it's used in the procedure for the identification of the most suitable point for dispenser deployment. The LIDARs come from different producers, but they are quite similar being designed for outdoor use. The two used for navigation are **Velodyne VLP-16**³ and **Hokuyo UTM-30LX-EW**⁴, the one on top of the arm is **SICK Tim561**⁵. Main differences and similarities between these models are highlighted in Table 4.2

³ <http://velodynelidar.com/vlp-16-lite.html>

⁴ <https://www.hokuyo-aut.jp/search/single.php?serial=170>

⁵ <https://www.sick.com/de/en/detection-and-ranging-solutions/2d-lidar-sensors/tim5xx/tim561-2050101/p/p369446>

RGB-D camera

The other sensor mounted on top of the arm end effector is a **Realsense** RGB-D camera from Intel. This camera is able to sense the depth of the acquired images, by means of an active infrared stereo technology. The main usage of the camera is to provide images used for the feedback loop of the visual servo control of the Jaco² arm. Visual servoing is a technique which uses feedback information extracted from a vision sensor to control the motion of a robot (Hutchinson, Hager, and Corke, 1996), and it's used in this project to compensate the lack of precision in the arm motion operations (see Section 4.1.2). This camera was also used in validation operations (validation of dispenser grasping, validation of dispenser deployment on the plant) and identification of exceptional cases (no dispensers left).

The positive aspect of this camera model comes from its contained size and weight. We spotted only one significant weak point, that is a quite high minimum depth distance (20 cm), due to an intrinsic limit of the embedded depth technology. This problem created a few issues in dispenser application tasks, but they were easily bypassed.

Tools for dispensers grasping

Additional hardware was placed on the unmanned ground vehicle specifically for the accomplishment of the grasping of the dispenser, for two separate reasons:

- the shape of the dispensers (see Figure 3.3) was not directly compatible with the grasping capability of the end effector of Jaco² arm (see the Jaco² fingers in Figure 4.2a).
- the dispensers have to be stored on the Husky platform, in such a way to make them graspable.

We required custom hardware to accomplish these objectives so, given the prototype nature of the project, 3D stamped hardware was used. For what concern the storage of the pheromone dispensers, the final version of the dispenser feeder was built out of two 3D-printed small towers (see Figure 4.5), with a sequence of very smooth aligned grooves, parallel to the horizontal plane, where the dispensers are slotted in. The smoothness of the grooves is the critical part of the design, indeed the previous versions of the towers, with sharper grooves, highlighted a difficulty of the grasping phase if the dispensers got caught on the feeder's corners.

The fingers of the robot are not all the same, because one of them act as a sort of "opposable thumb", that is useless in our application. So, two "nails" were printed to be applied to the two symmetric fingers, with a side dispenser-sized groove each, in order to grasp the dispenser by opening the two symmetric fingers.

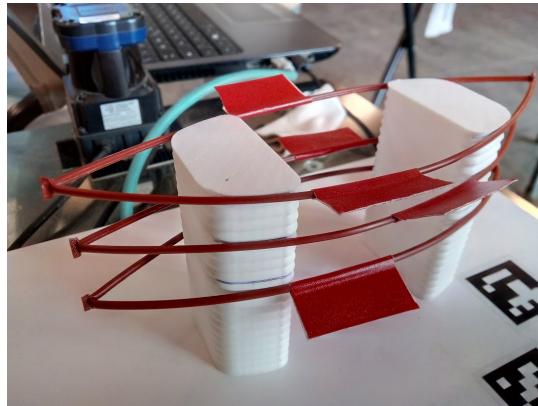


Figure 4.5: The final version of the feeder, holding three dispensers.

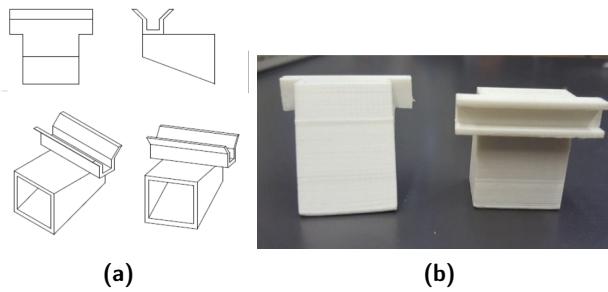


Figure 4.6: The final version of the 3D-printed nails (4.6b), and their perspective drawing (4.6a)

Other sensors

The unmanned ground vehicle holds onboard a collection of other sensors, useful for the accomplishment of the established tasks. In this section we give a short overview of the few remaining ones:

IMU SENSOR Device including a magnetometer, accelerometer, gyroscope. These sensors were used in the odometry estimation system through sensor fusion algorithms (see Chapter 5).

GPS Also this sensor data stream was included in the odometry estimation system.

CAMERAS Additional camera mounted in a slightly lifted position with respect to the robot base; they are useful in case of teleoperation of the Husky or the robotic arm, and for visualization and monitoring purposes.

You can see the final sensor/actuators arrangement in Figure 4.7.

GRAPE SOFTWARE ARCHITECTURE

Now that the hardware configuration is clear, we focus on the *behavior* of the unmanned ground vehicle. We start by a description of the sub-



Figure 4.7: The complete robotic platform, with sensors/actuators final arrangement.

problems that we identified starting from the macro goals we talked about in Section 3.1, and then move to the software architecture.

Evaluation of vineyard navigation

The problem of autonomous navigation of unmanned ground vehicles is a very well-known problem, studied at least from the nineties (Gage, 1995), because it's the first problem to deal with in almost all systems involving mobile robots. For this reason, and being our robotic system a prototype, the problem of vineyard navigation reduced to the following subproblems:

- identification of a sensor fusion framework in [ROS](#), to provide an odometry estimate, and configuration of such a framework, with particular attention to which sensors fuse together.
- identification of a navigation framework in [ROS](#), to provide motion planning and speed control of the mobile base, and configuration of such a framework to fit our problem.

Even if these tasks could seem trivial, two major problems arose:

- complex framework like these ones have tenths of parameter to tune, and the documentation is not always very precise.
- the *off-the-shelf* navigation systems are usually born for indoor environments, that differ a lot from the very unstructured nature of the vineyards.

As hinted in Section 2.5 and 2.4.1, our choices fell on **Move Base** package for the navigation, and **Robot Localization** as odometry system. More detail of the subproblems mentioned above are provided in Chapter 5.

Evaluation of vineyard monitoring task

For the vineyard monitoring task, as foreseen by the project proposal, we opted for a semi-autonomous monitoring: we provided the unmanned ground vehicle with a camera with video stream capability, that allows human users with competences in the domain of plants health to observe in person the live data streamed by the robot for assessment of crop condition. Human operators are also provided teleoperation interface for navigation and manipulation, for supervision and emergency handling function. This choice came from both the absence in the project proposal of precise objectives on this topic, and the complexity of feature extraction and evaluation in this context.

Evaluation of dispenser application task

While the navigation task is a very common problem and *off-the-shelf* solutions are easy to find, the dispenser application is a completely non-standard task, so we could only rely on libraries and packages for low-level subtasks and we had to design and implement the whole procedure from scratch. We divided the procedure in 3 subproblems:

1. **Creation of a point cloud of the tree:** the robot is sitting still, with a vine on the same side of the Jaco² arm. When the procedure is triggered, the arm moves in order to produce a point cloud of the trees close to it, using the **LIDAR** mounted onto it.
2. **Deployment points identification:** once the point cloud has been produced, an algorithm based on Point Cloud Library (**PCL**) trims the point cloud from the points that don't belong to the tree, process it, and outputs a sequence of 3D points (x, y, z) corresponding to points on the tree suitable for the deployment of the dispenser, typically a protuberance of the tree, ordered from the best one to the worst one according to a suitable metric.
3. **Deployment of the dispenser:** the arm starts the dispenser deployment procedure, taking as input the sequence of points got in the previous step:
 - a) it grasps the dispenser (if any) from the dedicated support.
 - b) it applies the dispenser on the best feasible point from the input sequence.

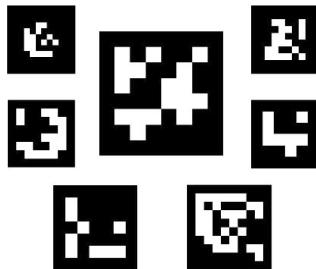


Figure 4.8: Examples of visual binary fiducial markers (more specifically, ArUco markers) as the ones we used in visual servoing applications.

- c) it goes back to a retired position, to achieve better stability during the movement of unmanned ground vehicle.

In Figure 4.10, you can see an example of successful dispenser application, using the mockup plant. **NOTE:** the integration weeks were not all held in the same place (two of them in Garriguella (ES), and one in Casciana Terme (IT)), so we could observe that the vine trees shape is very different in the different locations (for example, the trees in Garriguella tend to have small almost vertical branches on the top of the pruned tree, while the Italian ones had smaller protuberances with variable orientation).

This factor makes our system much harder to validate, because it's even more difficult to identify useful features in the unstructured vineyard environment. Thus, we introduced a further deployment mode, to be used in case our deployment procedure on the plants turned out not versatile enough to be efficient on a new tree type. Since we wanted to bypass the problems due to different tree structures, in this mode we assume to have some appropriately shaped nails, expressly arranged around the vineyard for dispenser applications. As described in Chapter 6, this deployment mode also requires the application of *binary square fiducial markers* (see Figure 4.8), so in this way our application pushes a bit more towards a vineyard that modifies in order to be more suitable to robotic applications.

The steps of this deployment mode are slightly different from the previous ones:

- a) it grasps the dispenser (if any) from the dedicated support.
- b) it applies the dispenser on the nail, regardless of the input sequence (which become useless).
- c) it goes back to a retired position, to achieve better stability during the movement of unmanned ground vehicle.

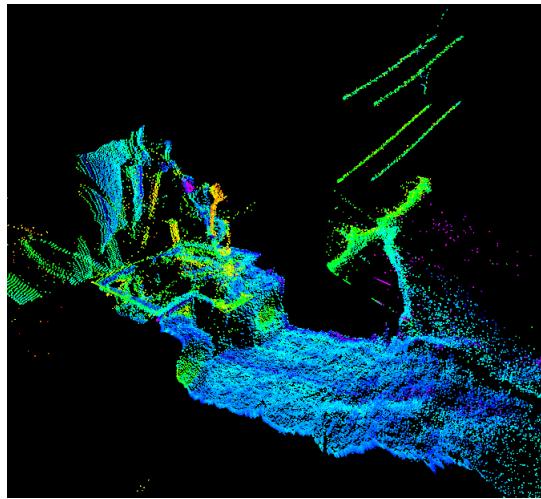


Figure 4.9: Examples of laser scans collected in Casciana Terme during the integration week in January 2018, with the LIDAR mounted on top of the end effector of the arm. You can recognize the vine plant, and part of the Husky base. The colors are only for visualization purpose.

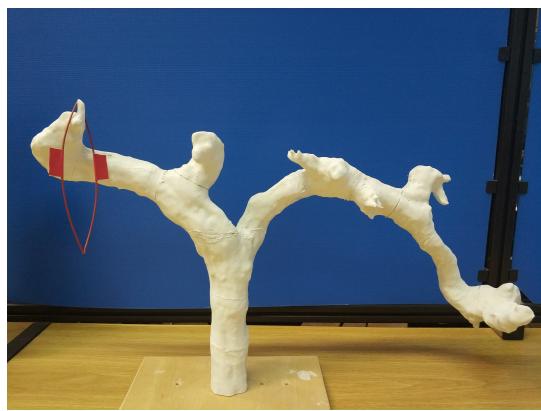


Figure 4.10: Example of a dispenser application task, correctly executed on a mockup plant.

All these procedures are quite long and complex, we wanted the user to have a high control on them, including the monitoring of intermediate results. For these reasons, these procedure are implemented as [ROS](#) actions.

Details of the actions interfaces

In this section we are giving a detailed description of the action interfaces, to better understand their scope and goals. Note that none of them includes a flag to check for correctness of execution, because we decided to use the internal status variable embedded in [ROS](#) actions.

Move Base action interface

- **Goal Message:**
 - *target_pose*: the pose (position, orientation) we want our robot to reach
- **Result Message:**
 - *base_position*: the final pose (position, orientation) of the robot; in case of success, it should be very close to the specified *target_pose*
- **Feedback Message:** *n.a.*

Scan motion action interface

- **Goal Message:** *n.a.*
- **Result Message:** *n.a.*
- **Feedback Message:** *n.a.*

Note that even an empty [ROS](#) action like this makes sense, because of the length (~30 seconds) of the procedure triggered by an action call.

Point cloud processing action interface

- **Goal Message:** *n.a.*
- **Result Message:**
 - *deployment_pose*: contain an array (x, y, z) of possible deployment poses, identified from the plant point cloud; it's used as input to the dispenser deployment action server (see Table 4.5)
- **Feedback Message:** *n.a.*

Goal Message	geometry_msgs/PoseStamped target_pose
Result Message	geometry_msgs/PoseStamped base_position
Feedback Message	---

Table 4.3: Move Base action specification

Goal Message	---
Result Message	---
Feedback Message	---

Table 4.4: Scan motion action specification*Dispenser deployment action interface*

- **Goal Message:**

- *deployment_poses*: sequence of feasible deployment points, ordered from the best one to the worst one. These are the points output by *Point cloud processing action*
- *deploy_mode*: an integer, used to select the technique to be used for the dispenser deployment; in fact, three different mode has been developed, in order to fit the different condition that the unmanned ground vehicle could face. In short, the modes available for selection are:
 - * Closed loop control, deployment target is an expressly placed nail, hammered into a pole.
 - * Closed loop control, deployment target is a protuberance of the vine tree.
 - * Open loop control, deployment target is a protuberance of the vine tree.

Details about the deployment mode are given in Chapter [6](#).

- **Result Message:** *n.a.*

- **Feedback Message:**

- *ManipulationStatus status*: a message of proprietary type, which contains structures used to communicate the action caller the current execution state (*e.g.*, dispenser fallen, arm moving in front of the target point).

Goal Message	- - -
Result Message	DeploymentPose[] deployment_poses
Feedback Message	int8 execution_percent range [0,100]

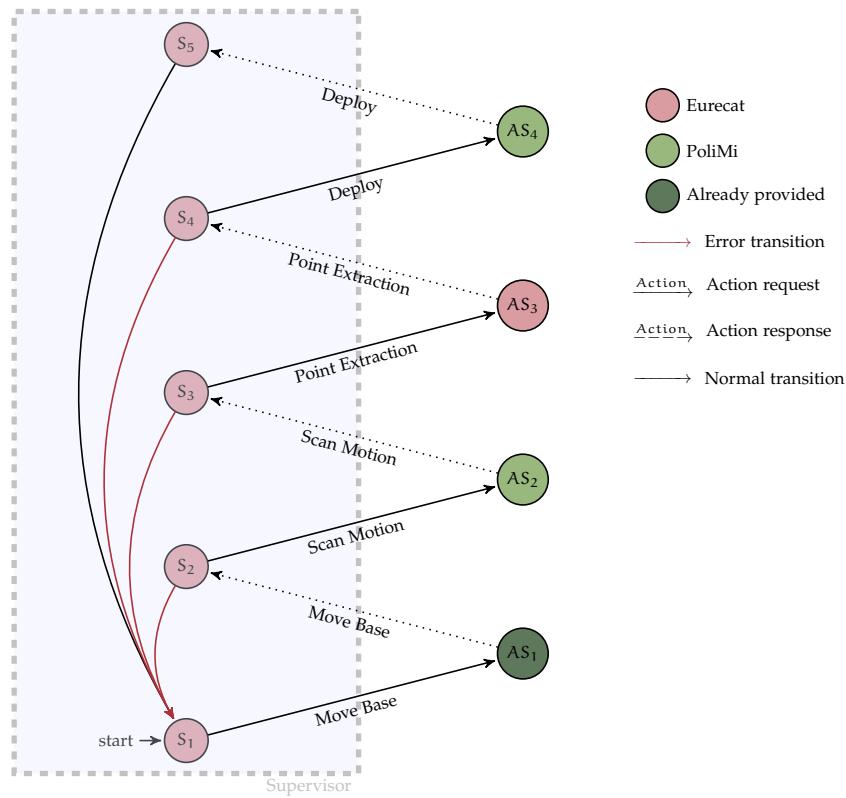
Table 4.5: Process point cloud action specification

Goal Message	DeploymentPose[] deployment_poses int32 deployMode
Result Message	- - -
Feedback Message	ManipulationStatus status

Table 4.6: Dispenser Deployment action specification

The supervisor

Since all main operations in the project structure are implemented through action servers, a list of action clients must exists. For our implementation, we opted for a single software module that acts as an action client for all the actions mentioned above. We call this module *Supervisor*, and its implementation uses behavior trees (Colledanchise and Ögren, 2017) as an overall model. The objective of the supervisor is to implement the global finite state automata of the system, where basically **nodes** are action servers or action clients.



As pointed out by Figure 4.2.5, all action clients belong to the supervisor software module, and major tasks are implemented through action servers. In that graph we also specify the division of labor between the different teams:

- Software modules assigned to **PoliMi**
 - AS₂: Scan motion action server
 - AS₄: Dispenser deployment action server
- Software modules assigned to **Eurecat**
 - AS₂: Point cloud processing action server
 - S_i : Supervisor (*i.e.*, all action clients and finite state automata implementation)

Software modules treated in this thesis

After this global overview of the hardware and software components of the **GRAPE** project as a whole, we can finally specify which of the aforementioned topics has been directly treated in this thesis work. They are:

- unmanned ground vehicle navigation system (see Chapter 5)
- development of Scan motion action server (see Chapter 6)
- development of part of Dispenser deployment actions server (see Chapter 6)

ODOMETRY AND NAVIGATION SYSTEMS IN GRAPE

The studies of Singhal, 1997, that analyze main issues in autonomous mobile robot navigation, identify three main problems related to outdoor autonomous navigation:

- the unstructured environment where the actions take place, in opposition to the clear and definite one that is typical in indoor navigation (flat floor, right angles, smooth surfaces).
- multiple sensors required, to be aware enough of the surrounding environment, in order to take decisions.
- moving obstacles, like for example pedestrians or cars in a city road.

The system developed in the context of **GRAPE** project suffers critically mostly of the first two point listed above. Indeed, the vineyard environment, even if characterized by a high degree of regularity in the global structure, due to the presence of parallel rows of trees, presents strong difficulties regarding lack of structure in the terrain configuration (steep, bumpy, clay-rich, muddy, according to the weather conditions and the intrinsic composition and morphology of the terrain) and in the obstacles (non-straight surfaces, different reflectivity and lighting conditions).

The problem of the fusion of the data coming from the disparate sensors mounted on board of the robot was already addressed from a theoretically point of view in Section 2.4, and it's the major component in the localization system of the robot. Note that, being our robot an agricultural unmanned ground vehicle, the problem of moving obstacles is not particularly relevant, since they are most likely to be humans, aware of the robot presence and functionalities, or small wild animals (*e.g.*, foxes, hares, cats) in remote cases, that however have a very low probability of approaching the Husky base during movement.

In the next section we are going to analyze with detail the configuration of both the odometry and navigation systems running on the unmanned ground vehicle. It will be clear that both the problems, even if eventually solved using *off-the-shelf* solutions, were all but simple problems and required a real understanding of the numerous problems that arose.

SENSOR FUSION SYSTEM

The configuration of a robust odometry system for vineyard navigation has been a workpoint since the beginning of the project, because of the expected (and actual) strong instability of the wheel odometry. To the reasons mentioned at the beginning of the chapter, we recall also the intrinsic lack of precision due to skid steering kinematics of the Husky base. We can describe the search for a robust odometry system for the unmanned ground vehicle in three main phases:

1. As hinted in Section 2.4, in the early phases of **GRAPE** project, the designated sensor fusion framework was **ROAMFREE**, that gave proof of good functioning in the past (see Section 2.4 for precise references). ROAMFREE platform provides multi-sensors pose tracking and it is designed to be flexible and to adapt to every kind of mobile robotic platform. The tracking module is based on Gauss-Newton minimization of the error functions associated to sensor readings. In order to improve generality, physical sensors are abstracted with *logical sensors*, which are characterized only by the type of the proprioceptive measurement they produce.

Another ROAMFREE module provides instead self-calibration modules, using error models for each sensor category to provide on-line correction of common sources of distortion, bias and noise (*e.g.*, hard and soft magnetic distortion, sensor displacement or misalignment). The usage of ROAMFREE framework was specified in the project proposal, thus it was of course the default choice for the sensor fusion framework in the first *integration week* in Garriguella (ES), before the beginning of this thesis work.

The **ROS** implementation of ROAMFREE is still experimental, anyhow a configuration was identified to fit the requirements. Unfortunately, the same configuration turned out to perform poorly during the second integration week in Casciana Terme (IT). This was caused by concurrence of:

- presence of a significant slope in the vineyard, in opposition to the flat (even if bumpy) terrain in Garriguella
- difficulty to handle inaccurate configuration of magnetometer and **IMU**
- overall trickiness in the configuration procedure, being ROAMFREE integration with **ROS** still experimental

The low performances of this framework in this context were pretty clear, so we opted for a more tested and consolidated solution, **Robot Localization**.

2. The configuration of Robot Localization requires, for each fused sensor, a configuration vector, which specifies which compo-

nents of the input estimate should be fused into the final pose estimate.

Note that the configuration vector is given in the *frame_id* of the input message: for example, consider a velocity sensor that produces a *geometry_msgs/TwistWithCovarianceStamped* message with *velocity_sensor_frame* as *frame_id*. We now assume that the transform would convert \dot{X} in *velocity_sensor_frame* to \dot{Z} in the *base_link_frame*. To include the \dot{X} data from the sensor into the filter, the configuration vector should set the \dot{X} velocity value to true, and not the \dot{Z} velocity value.

In Table 5.1 you can see the configuration vectors for the sensors taken as input from Robot Localization. Note that, for example, GPS only gives a contribute about x and y , since the altitude estimation of GPS is often not precise, and it provides no information about the robot orientation. Our configuration of Robot Localization relies on **two** different Extended Kalman Filter (**EKF**) nodes in cascade:

- a *local* localization node, that fuses continuous data, and estimates the position and orientation of the *base_link* frame of the robot inside the *odom* frame, *i.e.*, the frame fixed in the initial position of the robot. In other words, the task of the node is to publish *odom* \rightarrow *base_link* *tf* transform (see Section 2.2 for recall).
- a *global* localization node, that also fuses low frequency sensors, *i.e.*, GPS, and estimates the position and orientation of *odom* frame inside *map* frame. The *map* \rightarrow *odom* transformation "adjusts" the position of frame *odom* in order to contextualize the position of the unmanned ground vehicle inside the map. Thus, *odom* \rightarrow *base_link* *tf* transformation tracks the robot position with respect to its initial pose frame *odom*, and *map* \rightarrow *odom* transformation moves the initial pose frame to localize the robot inside a map.

In Figure 5.1 you can see the computation graph of the aforementioned odometry system: you can observe that:

- both nodes write on *tf* topic the respective *tf* transforms.
- the local node takes **IMU** (*/imu/data* topic) and wheels (*/husky_velocity_controller/odom* topic) estimate for the fusion task.
- the global node takes as input the pose estimate from GPS (*/odometry_gps* topic), and the pose estimate from the local node (*/odometry/ekf_local_filtered* topic), to implement the cascade scheme.

	x	y	z	ψ	θ	ϕ	\dot{x}	\dot{y}	\dot{z}	$\dot{\psi}$	$\dot{\theta}$	$\dot{\phi}$	\ddot{x}	\ddot{y}	\ddot{z}
Wheels	✓	✓	✓				✓	✓	✓	✓					
IMU				✓	✓	✓					✓	✓	✓	✓	✓
GPS points	✓	✓													

Table 5.1: Configuration vectors for the sensors fused for odometry estimation, using Robot Localization.

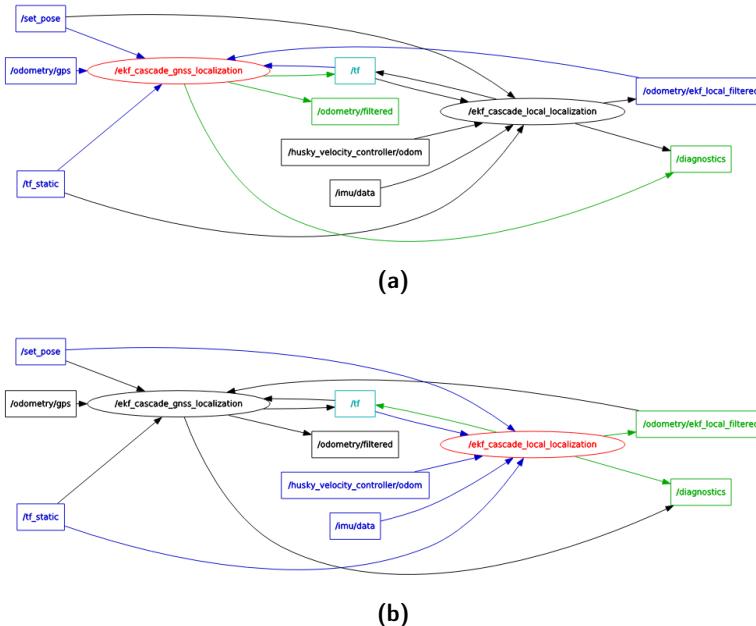


Figure 5.1: Computation graph of the cascade of Robot Localization nodes; in blue, the input topics, and in green the output topics, of the nodes highlighted in red.

3. Even if the cascade produced a pretty precise odometry estimate, the odometry computation was later shrunk in a single **EKF** node for two main reasons:
 - as we'll describe in Section 5.2, the navigation system switched to a mapless configuration, so the *tf* transform *map* → *odom* became useless.
 - the **EKF**s cascade configuration led to a short delay in the odometry computation (~200-300 ms). This caused some problems with the obstacles detection implemented through **LIDAR** data, which were processed with negligible delay, and especially when the robot moved from the end of a row of vines to the beginning of another one.

NAVIGATION SYSTEM

We already anticipated in Section 2.5 that the reference navigation framework is **Move Base**, that is an *off-the-shelf* ROS software stack for autonomous navigation (see Figure 2.16 for a graphical representation of the main software modules of Move Base). Since, as already stated, this framework usually provides very stable autonomous navigation in indoor mapped environment, the first attempt was to shift the same configuration in the vineyard environment. In this section we'll explain the steps from this initial naive solution to the final accepted one.

Navigation in mapped environment with AMCL

Before the robot is able to navigate the vineyard, the navigation in mapped environments requires a preparatory *mapping* phase; in particular, this is a **SLAM** problem, that is the computational problem of constructing or updating a map of an unknown environment, while simultaneously keeping track of an agent's (the unmanned ground vehicle) location within it. Note that *localization* is a different task with respect to *odometry estimate*, because the first aims to estimate the position of the robot inside a map, while the latter aims to estimate the position of the robot with respect to its initial pose. More technically, in the **ROS** framework *odometry estimate* algorithms publish the *tf* transform from *odom* frame to *base_link* frame, while *localization* algorithms publish the one from *map* to *odom*.

Even if **SLAM** is a *chicken-and-egg* problem (localization would be easy in a known map, under some assumptions, and mapping would be easy if the robot position is assumed known), several algorithms exist that solve, even if approximately, the **SLAM** problem. In the early phases of **GRAPE**, previous to this thesis work and described in Serrano et al., 2017, three different **SLAM** algorithms were tested in our specific environment: *Gmapping*, *Google's Cartographer* and *KartoSLAM*, and **Gmapping** algorithm was selected because of its performance and efficiency in terms of computational load. In Figure 5.3 you can see the map created in Garriguella during the last integration week, using Gmapping. The accomplished accuracy of the maps was satisfying, but actually a lot of problems occurred in the main phase, that is the autonomous navigation of the robot in the mapped environment. The most relevant problems that emerged were the following ones.

AMCL IN THE VINEYARD MAP

As you can see in Figure 2.16, an optional node called **amcl** exists, that provides a **ROS** implementation of **AMCL** algorithm (Thrun et al., 2001). **AMCL** is a localization algorithm, that uses a particle filter to provide a *belief*, i.e., the robot's estimate of its

current state, through a probability density function distributed over the state space. A single particle represents an hypothesis about the real pose of the robot in the environment (see Figure 5.4); thus, regions in the state space with many particles correspond to a greater probability that the robot will be there, and regions with few particles are unlikely to be where the robot is.

In **AMCL** each particle is assigned a weight which corresponds to the probability that, had the robot been at the state of the particle, it would perceive what its sensors have actually sensed (in particular, *amcl ROS* package is only capable to deal with laser scans as sensor input); this weight is used in the algorithm to select the particles that are going to survive through iterations. For this reason, the algorithm requires both a sensor model of the used sensors, and a motion model of the involved robot, in order to update the *belief*. In Figure 5.2 you can see several steps of *belief* update, only based on the motion model of the robot. Pseudo code of Monte Carlo Localization is provided in Algorithm 2.

Algorithm 2 Monte Carlo Localization(X_{t-1}, u_t, z_t)

```

1:  $X_t = \emptyset$                                  $\triangleright X_t$  is belief about the robot state
2:  $\bar{X}_t = \emptyset$ 
3: for  $m = 1$  to  $M$ :                          $\triangleright M$  is number of particles
4:    $x_t^{[m]} = \text{motion\_update}(u_t, x_{t-1}^{[m]})$      $\triangleright u_z$  is actuation command
5:    $w_t^{[m]} = \text{sensor\_update}(z_t, x_t^{[m]})$        $\triangleright z_t$  are sensed data
6:    $\bar{X}_t = X_t + \langle x_t^{[m]}, w_t^{[m]} \rangle$ 
7: endfor
8: for  $m = 1$  to  $M$ :
9:   draw  $x_t^{[m]}$  from  $\bar{X}_t$  with probability  $\propto w_t^{[m]}$ 
10:   $X_t = X_t + x_t^{[m]}$ 
11: endfor
12: return  $X_t$ 

```

AMCL with laser scanners provides a very effective localization in feature-rich environments, like in typical indoor environments with very sharp corners and well defined obstacles, as the ones in Figure 5.4. However, as hinted at the beginning of this section, a few tests were sufficient to show the weakness of this localization algorithm in our context, due to the high repetitiveness of the vineyard map. In Figure 5.3b you can see a magnified detail of the map of Garriguella vineyard, computed using Robot Localization as odometry system, and Gmapping for **SLAM**; most of the obstacles detected by the frontal **LIDAR** are the trunks of the trees, together with poles and high weeds. It's intuitive that if the map in which the robot moves presents

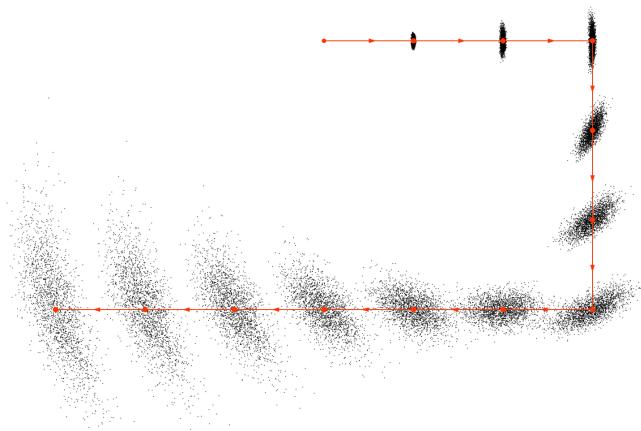


Figure 5.2: Evolution of the belief computed by AMCL, taking into account the motion update only (ignoring the sensor update).

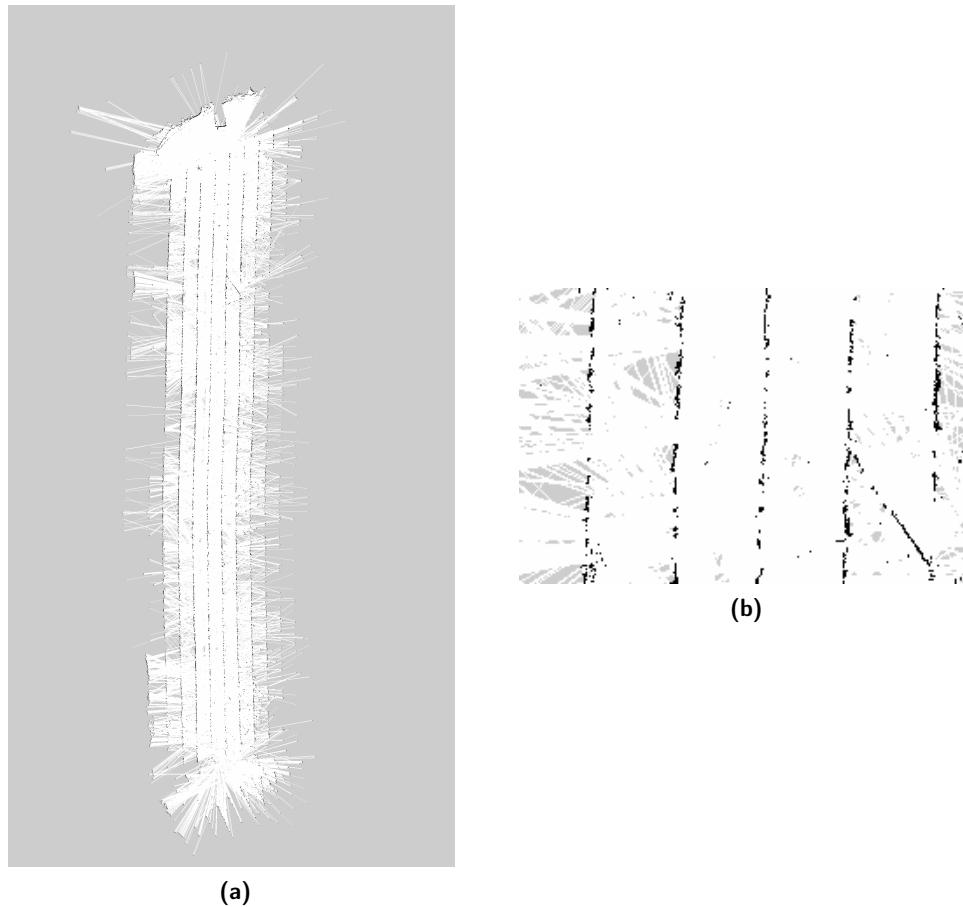


Figure 5.3: Map of the vineyard in Garriguella (ES), computed during the last integration week using Robot Localization and Gmapping. In 5.3b, a magnified detail.

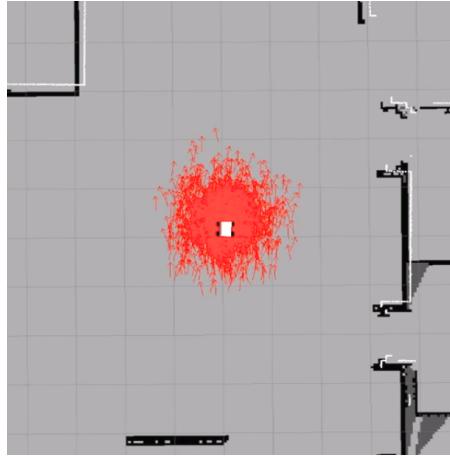


Figure 5.4: Belief about the robot estimate using [AMCL](#), through a particles cloud where each particle represents a hypothesis about the real pose.

a lot of repeated patterns, the *survival of the fittest* paradigm that the *sensor update* step of [AMCL](#) should implement is not very effective. Moreover, the bumpiness of the terrain leads the plan that the [LIDAR](#) scans to change significantly its orientation in time. So, for example, given a position in the map, the sensed obstacles over time in this location could be different portions of the same vine, then the shape is likely not to be constant.

ROW CROSSING PROBLEM

The vineyard environment also presented another criticality, due to the average distance between the vines belonging to the same row; indeed, because of the height of the [LIDAR](#) used for obstacles detection, the trunks of the plants are detected while the horizontal branches of the vines often are not. If this situation occurs, the space between the trees is sensed as free of obstacles, even if the horizontal branches of the vines are an actual obstacle to the unmanned ground vehicle. The first attempt to solve this problem was to simply provide an higher inflation radius (see Section 2.6) to the obstacles, in order to fill the gaps between a plant and the adjacent ones. However, this naive solution couldn't find a trade-off to the extreme situations:

- a too small inflation radius is not able to provide a solution to the problems, *i.e.*, fill the gaps between adjacent trees.
- an inflation radius large enough to fill the aforementioned gaps, makes the navigation almost impossible, because even a small obstacle among two rows (*e.g.*, some weed) completely occludes the path.

Insertion of prohibition layer

Given the impossibility to solve the problem of row crossing with the standard inflation radius of the obstacle layer, we found a very effective solution that relies on the very simple global structure of the vineyard. Indeed, looking at the map in Figure 5.3a, it's easy to notice that the obstacles that cause the row crossing problem are of course aligned, and parallel to each other, so the idea was to implement a series of "virtual fences" corresponding to the vine rows. The implementation was possible thanks to a feature of `costmap_2d`, the package of ROS responsible of the creation of the occupancy grid maps. Indeed, this package implements a **layered** costmap: this means that, for example, the global and local costmaps (see Section 2.6) live in two different layers, that are computed independently:

- the global costmap is computed from the obstacles described in the occupancy grid map, possibly inflated of a chosen radius.
- the local costmap is computed in a spacial window around the robot position, from the obstacles detected by the laser scans. The obstacles possibly are inflated of a chosen radius.

Since an arbitrary number of custom layers can be added by implementing C++ base class `costmap_2d::Layer` of Move Base package, two additional layers were introduced, in order to include the virtual obstacles of the rows both in global and local costmap:

- **Global prohibition layer:** this includes the virtual fences in the global costmap, and it's useful for the global planner to take them into account for the creation of the global navigation plan.
- **Local prohibition layer:** this includes the virtual fences in the local costmap, and it's useful for the local planner to take them into account for the creation of the local navigation plan. This is required because otherwise the local planner might over-optimize the global plan and, however take advantages of the gap between the plants. This situation is comparable to an indoor planning problem in which a local planner detects an open door that was shut at mapping time: even if the global path doesn't take advantage of the door gap, the local planner might choose to traverse it if it reduces the motion execution cost.

Switch to mapless navigation

After several effort in Move Base parameter tuning, the role of **AMCL** in the poor performance in navigation tasks was clear. Since the map in navigation task is useful in terms of:

- **global planning:** it provides a set of obstacles that helps the robot to optimize the global path in planning phase. See for example Figure 5.5: if no map was provided, the global plan would have been a straight line between the start and the goal states, and all the obstacle avoidance lies on local planner. This of course can lead to a performance degradation, because the knowledge of some obstacles in advance often brings to much more efficient global plans.
- **localization:** in order to provide an estimate of the robot pose in the environment (for example, through AMCL).
- **visualization:** to better contextualize the position of the robot for human users.

Given the problems caused by AMCL, we decided to switch to a **mapless** navigation configuration, because we were able to compensate all the advantages of the mapped navigation using other tools or methods:

- **global planning:** as we already pointed out in Subsection 5.2.2, even if the vineyard environment is strongly unstructured at fine-grained level (for the reasons explained at the beginning of this section), it has a fixed and very simple global structure given by the parallel vine rows, as you can see in vineyard maps (Figure 5.3). Thus, the inclusion of the virtual fences layers only in the global costmap is a simple, but very effective approximation of the vineyard global map.
- **localization:** in both our field test locations (Casciana Terme (IT) and Garriguella (ES)), we could rely on high-precision RTK GPS (Langley, 1998) as input for the unmanned ground vehicle odometry system. Even if RTK correction worldwide coverage is not guaranteed, several companies exist that provide this service. For example, see Figure 5.6 for a map of RTK stations provided in Italy by the sole SmartNet Italy¹ company. Moreover, since our navigation tasks are executed outdoor, the assumption of an high-precision GPS system is reasonable and we only need to specify the virtual fences positions through the GPS coordinates of their endpoints to navigate the virtual fences costmap layer.
- **visualization:** in this configuration our unmanned ground vehicle is always geolocalized; so, we can contextualize the robot position for human users by overlaying to the robot position the corresponding satellite images, using for example the *Rviz_satellite* plugin for Rviz.

¹ <http://it.smartnet-eu.com/>



Figure 5.5: Example of global plan that takes into account the map obstacles.

Since we were able to compensate for all the advantages of the navigation with a map, we decided to switch to a mapless configuration, at the cost of adding the assumption of high-precision GPS and geolocalized virtual fences at disposal. Note that, in this configuration, the *localization* task is not relevant anymore since there is not an actual map to localize into; then the *tf* transform from *map* to *odom* is reduced to the identity transformation.

Switch to Timed Elastic Band as local planner

Also the standard local planner of Move Base, a [ROS](#) implementation of Dynamic Window Approach (Fox, Burgard, and Thrun, 1997), showed some weak points in the vineyard context. In Figure 5.7 you can see an example of local costmap, computed during an autonomous navigation tasks in Garriguella: the virtual fences are highlighted in blue, the obstacles detected by the frontal [LIDAR](#) in pink, and the inflation layer in light blue. Since the detected obstacles are mainly tree trunks, they are detected as small, non continuous obstacles. The inflation of such obstacles leads to a jagged border of the resulting obstacle cells set. If, for example, the detected obstacle was a straight wall, the resulting obstacle layer would have been much more straight and sharp. Dynamic Window Approach had some problems to deal with such a large amount of dead ends, and often got stuck in the obstacle contours (see Figure 5.8). Thus, we switched to another *off-the-shelf* local planner, compatible with the Move Base local planner interface: an implementation for [ROS](#) of the *Timed Elastic Band* algorithm (Rösmann et al., 2012). This algorithm, at the cost of an higher computational load, provided much more robust navigation.



Figure 5.6: RTK stations provided by SmartNet Italy.

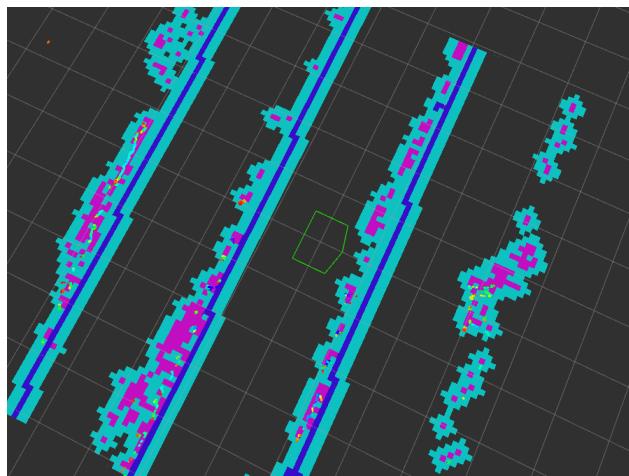


Figure 5.7: Example of local costmap, in the final configuration: you can distinguish the virtual fences (blue), the obstacles detected by the LIDAR (pink), and the inflation layer (light blue). The green polygon represents the robot footprint.

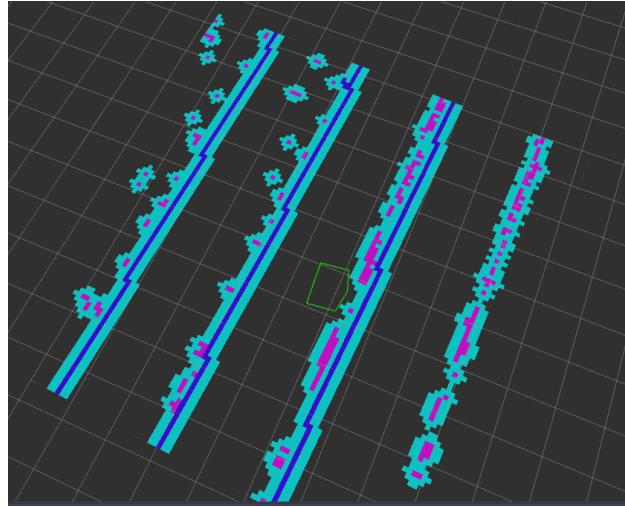


Figure 5.8: Example of local plan computed by DWA that get stuck in the jagged border of the obstacles.

AMCL as input to sensor fusion

In the late phases of the [GRAPE](#) project, some efforts were made to, however, take advantage in some way of [AMCL](#) algorithm potential. In particular, the alignment with the sensed laser scans could lead to the following benefits:

- better estimation of the robot orientation
- compensation of the measurement errors of the GPS, since the virtual fences are specified through GPS coordinates of the end-points.

We already highlighted the strong drawbacks due to using [AMCL](#) as localization algorithm, so the attempt was to disable its publication of the *tf* transform from *map* to *odom*, and fuse its estimated pose together with the other sensor inputs in the Robot Localization node. In this way the weaknesses of [AMCL](#) alone are likely to be compensated from the other sensors' estimates. If this configuration had resulted effective, switching back to a navigation with map would have been a reasonable choice. Unfortunately, due to the lack of time (final review of [GRAPE](#) project was held in March, 21st, 2018), this configuration was not sufficiently tested, and mapless navigation was the only one that was actually validated on the field.

However, the tests that were performed were sufficient to highlight some difficulties in the coexistence of GPS and [AMCL](#) as input sources in the Robot Localization nodes, probably for these reasons:

- Low publishing rate of both [AMCL](#) and GPS (~ 1 second).
- Pose estimates always discordant between them, maybe for some misconfiguration or bias of GPS module.

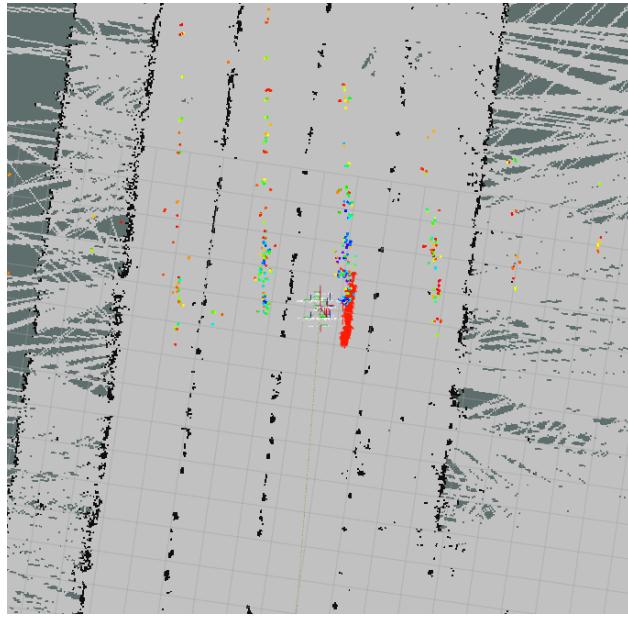


Figure 5.9: Example of conflict between the pose estimated from [AMCL](#) (displayed by the particles positions), and the final Robot Localization output (displayed by [tf frames](#)).

For this reason, the output pose of Robot Localization suffered of constant jumps between the estimate of GPS and the one of [AMCL](#), instead of keeping both of them into account. In Figure 5.9 an example is provided of discordance between the estimate from [AMCL](#) (represented by the particles position) and the one from GPS (corresponding in the figure to the actual estimated pose).

6

MANIPULATION

In this chapter, we'll explain our implementation of the action servers described in Chapter 4 related to the scan motion for point cloud registration, and to the pheromone dispenser application. But before it, we'll see a few details about the Kinova API, and the different techniques we used for the motion planning and execution for the Jaco² arm.

KINOVA ROS API

ROS drivers of Kinova devices are very complex, and provide via ROS topics a large amount of API useful for both monitoring the arm state, and controlling it. Besides it, *jaco_arm_driver* node also publishes the *tf* subtree that describes the position of all the arm joints. In this section we do not want to provide a deep analysis of the drivers software architecture, but we limit ourselves to list the interfaces offered as ROS topics, since they are the ones that we used in GRAPE project. The published topics are, used to output information about the arm state, are:

- */j2n6s300_driver/out/cartesian_command*
- */j2n6s300_driver/out/finger_position*
- */j2n6s300_driver/out/joint_angles*
- */j2n6s300_driver/out/joint_command*
- */j2n6s300_driver/out/joint_state*
- */j2n6s300_driver/out/joint_torques*
- */j2n6s300_driver/out/tool_pose*
- */j2n6s300_driver/out/tool_wrench*

They allow to observe the position of the end effector both in Cartesian and C-space, together with position of the arm in C-space, the position of the fingers, the angular velocity of the joints, the torque and forces applied to the joints.

The subscribed topics, used to control the arm movement, are:

- */j2n6s300_driver/in/cartesian_force*
- */j2n6s300_driver/in/cartesian_velocity*

- */j2n6s300_driver/in/joint_torque*
- */j2n6s300_driver/in/joint_velocity*

They allow to control the end effector both in force and velocity, and the joints in both torque and velocity.

MOTION EXECUTION METHODS FOR JACO² ARM

MoveIt! framework

The first, and easier, motion execution method analyzed is simply to use an *off-the-shelf* motion planning and execution framework, **MoveIt!**. However, in Section 4.1.2 we revealed in advance that the movement of the chosen arm, even after a bug fixing session in **ROS** drivers of Jaco² arm, presented a series of major weakness using the arm **API** through *MoveIt!*. In detail, the problems we identified were:

- non-repeatability of the movements; this was a problem because in situations with plenty of obstacles like the dispenser application, we could not rely on a few test movement to be sure that the movements did not interfere with obstacles.
- lack of precision in the movements; if a target position of the arm was given, the actual position of the end effector at the end of the execution was likely to be several centimeters away from the expected one, about $0 \div 5\text{cm}$. This led to huge problems for precision tasks (dispenser grasping, dispenser deployment) where the maximum tolerable error is of the order of 1mm because of the dispenser narrowness.
- inability in following a trajectory, maintaining the end effector orientation; this capability is mostly required during the scan motion, because it guarantees a uniform density of points collected on the plant, and this is useful with a view to the post-processing operations for the identification of the deployment point. If we impose such a director trajectory, only a small percentage of it is actually executed.
- normalization of the joint angular positions in $[0; 2\pi]$ planning phase, with resulting loss of information about the absolute angular position of the joints of the arm. Even if in the vast majority of situations this is not a problem (since all joints of the Jaco² are capable of unlimited rotations), we needed to keep this factor into account because of the wiring for power supply and data interface of the sensors (**LIDAR** and RGB-D camera) mounted on top of the arm end effector. Indeed, even if the end effector joint is capable of infinite rotation, there is only a subset

of this interval, that we approximate to $[0; 2\pi]$, that is compatible with the wiring setup of the system, to avoid the wires being pulled. But since the planner has no access to the absolute angular position of the joints, it doesn't exist a smooth solution¹ to the problem of imposing one or more complete revolutions of the end effector joint.

Beyond all the flaws of our specific arm, an intrinsic weakness of this control type is that it can be very precise in reaching a specific target pose, but they cannot handle situations where there is a degree of uncertainty about the target pose, and we are in this situation in two different moments:

- when we need to grasp a dispenser, because the position of the dispenser in the dispenser feeder depends on how many dispensers have already been removed and however we do not want to have very strict constraints on the positioning of the dispensers in the feeder.
- when we need to deploy the dispenser on the plant, because the deployment points output from the point cloud processing are likely not to be perfectly precise.

On the other hand, one big advantage of *MoveIt!* framework is that it supports, in planning phase, the specification in URDF² of user-specified static obstacles, to be introduced as constraints in *C-space*. For example, in Figure 6.1 you can see the volumes of sensors and other physical obstacles on top of the real Husky base, modeled as parallelepipeds; motion planning with *MoveIt!* will never produce a motion plan containing configurations that intersect these obstacles.

For all the aforementioned reasons, *MoveIt!* framework is suitable for the motions where:

- extreme precision is not a requirement.
- constant orientation of the end effector is not a requirement.
- no constraints about the absolute angular position of the arm joints.
- the goal position is well known.

Direct publication of joint velocities

We introduced the method of publication of raw joint velocities to bypass the problem of the normalization of the joints angular position.

¹ a non-smooth solution using *MoveIt!* is to consider the revolutions as concatenation of a sequence of rotations of an angle $\theta \in [0; 2\pi)$ of the end effector joint.

² Unified Robot Description Format; it's an XML format used for representing physical components of a robot model

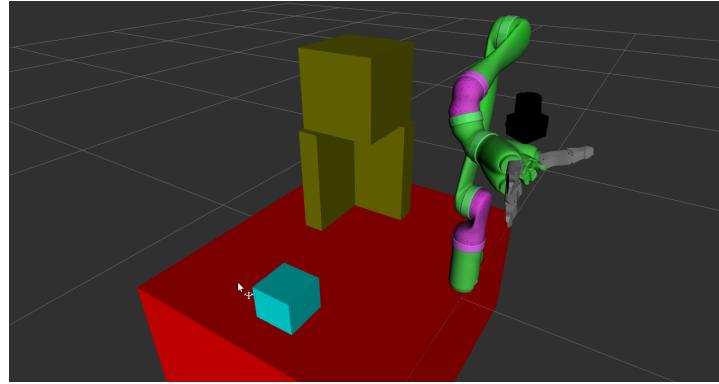


Figure 6.1: Visual representation with Rviz of the obstacles on top of the Husky (red shape) used by MoveIt! planner: the Velodyne sensor (light blue shape), and the physical support for camera, GPS and IMU (green shapes).

This method consists in a very simple closed-loop control that compares the target position of a joint with the actual one (obtained through Jaco² ROS API), and applies a velocity to the concerned joint if the error is still too high. This method is feasible because Jaco² ROS API provides (among other) a topic for controlling speeds in joint space (*i.e.*, the configuration space, see Section 2.6) of the arm, thus we could control the velocity of each of the joints by publishing at a given frequency (100Hz) a message with the desired velocity for each joint. The stop condition in movements executed with this method is given by the distance of the joint position from the target angular position for the controlled joint being less than a given threshold. The actual joint position is monitored on the specific topic, published by Jaco² ROS API.

This method applies well for simple movements, with no sophisticated control on velocity or position. In particular, we were interested in applying a constant velocity to the end effector joint, in order to solve the problem of the wires entanglement. As already stated in Section 6.2.1, this problem was not solvable by means of MoveIt! planner.

Inversion of differential kinematics

We introduced the method of *inversion of differential kinematics* to bypass the problem of moving the arm maintaining the end effector pose constant. This method, like the previous one, relies on ROS API for the publication of velocities in joint speed, but is suitable for more complex movements than the previous one. Our goal is to express the velocity of the end effector frame in terms of joint velocities, *i.e.*, given the desired velocity of the end effector, calculate the 6 joints velocities to be published on the suitable topic in order to achieve the desired

end effector velocity. This is a *inversion of differential kinematics* problem, formally: being q a coordinate vector that represent the position of the robot in joint space, v the vector of linear and angular velocity of the end effector frame with respect to the fixed base frame, and $J(q)$ the geometrical Jacobian matrix of the manipulator, composed by the stacking of $J_P(q)$, related to the end effector position, and $J_o(q)$, related to the end effector orientation:

$$v = \begin{bmatrix} \dot{p} \\ \omega \end{bmatrix} \quad q = \begin{bmatrix} q_1 \\ \vdots \\ q_n \end{bmatrix} \quad J(q) = \begin{bmatrix} J_P(q) \\ J_o(q) \end{bmatrix}$$

The following relationships hold:

$$\begin{aligned} \dot{p} &= J_P(q) \cdot \dot{q} \\ \omega &= J_o(q) \cdot \dot{q} \end{aligned}$$

In compact form:

$$v = \begin{bmatrix} \dot{p} \\ \omega \end{bmatrix} = J(q) \cdot \dot{q}$$

Thus, we can solve the initial problem with:

$$\dot{q} = J^{-1}(q) \cdot v \quad (6.1)$$

Systematic algorithm exists to compute the geometric Jacobian matrix (Joshi and Tsai, 2002); in our case, it's simply provided by *MoveIt! ROS API*. In our implementation, the computation (with expression 6.1) and publication of the joint speed is a periodic task executed with frequency 100Hz. This is required because the Jacobian varies in time, according to the positions of the joints. Since this motion technique allows for very smooth and uniform motion of the end effector, maintaining its orientation, we used it to execute the arm scan motion that requires exactly the aforementioned characteristics.

Note that, even if this method is very effective, it has some limitations for what concerns the *exceptional behavior*, since it does not embed dangerous situations detection (that is instead provided by more complex frameworks like *MoveIt!*). After a few tests, it was clear that we had to introduce some checks during movements executed with this method, since the poses required to perform the scan motion were dangerously close to arm singularities. The check we introduced acts like a stop condition for the movement execution, and it's based on the assumption that, in a non-pathological context, the Cartesian

distance between the end effector pose and the target pose is monotonically decreasing in time. The value $d(t)$ computed as:

$$\begin{aligned} p_{\text{endEffector}}(t) &= (x_c(t), y_c(t), z_c(t)) \\ p_{\text{target}} &= (x_t, y_t, z_t) \\ d(t) &= \sqrt{(x_c(t) - x_t)^2 + (y_c(t) - y_t)^2 + (z_c(t) - z_t)^2} \end{aligned}$$

is monitored in time using the values of $p_{\text{endEffector}}$ provided by Jaco² drivers, and the arm is stopped with consequent abortion of the corresponding action if $d(t)$ is observed to have grown in time instead of shrinking.

In order to avoid false positives due to sensor noise when the arm still has to actually start to move in the very first moments of the action execution (the arm is not moving, but of course the sensed end effector pose is not 100% stable over time), we apply this error detection measure only after the arm has moved a few centimeters (e.g., 2-3 cm) from the initial pose. The other stop condition for the motion execution is that the value of $d(t)$ shrinks under a certain threshold (e.g., 2-3cm). Keep in mind that the focus in this motion technique is on smoothness of the movement and constancy in the end effector orientation, rather than the precision in reaching the target pose.

Image Based Visual Servoing

This technique allows for much more precise positioning of all the other described in previous sections; it is a closed-loop control and makes use of the RGB-D camera mounted on top of the end effector. Now, we'll give a formal description of image-based visual servoing algorithm (Chaumette and Hutchinson, 2006).

Define a feature point as the vector:

$$\mathbf{f} = \begin{bmatrix} u \\ v \end{bmatrix}$$

where (u, v) are the coordinates of the point on the image plane representing the image of the considered point in 3D space. In case more than one 3D point, and thus more than one feature point, a feature vector can be defined as:

$$\mathbf{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}$$

In image-based visual servoing, the goal configuration is defined by a desired configuration f_d of the image features. Therefore, the image error function is given by

$$f_d = \begin{bmatrix} f_u \\ f_v \end{bmatrix}$$

$$e(t) = f(t) - f_d$$

The image-based control problem aims at finding a mapping from this error function to a commanded camera motion, under the following assumptions:

- the manipulator is a kinematic positioning device, *i.e.*, manipulator dynamics are neglected.
- the trajectories generated by the IBVS controller can be tracked by a lower level manipulator controller.
- the target is fixed and the desired configuration f_d is constant.

The most common approach to solve the aforementioned problem is to compute a desired camera velocity and use this as the control input, assuming that the robot can be commanded by a set of linear and rotational Cartesian velocities. The camera velocity can be computed using the relation, represented by the *interaction matrix*, between the velocity of the camera frame and the velocity of the features on the image plane:

$$\dot{f}(t) = L(f, q)v_{cam}$$

Where the interaction matrix for a feature point is given by:

$$L(f) = \begin{bmatrix} -\frac{f_u}{Z} & 0 & \frac{u}{Z} & \frac{uv}{f_u} & -\frac{f_u^2+u^2}{f_u} & v \\ 0 & -\frac{f_v}{Z} & \frac{v}{Z} & \frac{f_v^2+v^2}{f_v} & -\frac{uv}{f_v} & -u \end{bmatrix}$$

and v_{cam} is the vector of linear and rotational Cartesian velocities of the camera. In the case of a vector of feature points (*i.e.*, $n > 1$), the interaction matrix can be obtained stacking the interaction matrices associated to each point. Assuming that the number of selected features is such that the whole interaction matrix has more rows than columns ($n \geq 4$), the previous relation can be inverted using a least squares solution as follows:

$$v_{cam} = (L^T L)^{-1} L^T \dot{f}(t) \quad (6.2)$$

Since the derivative of the error function is given by

$$\dot{e}(t) = \frac{d}{dt}(f(t) - f_d) = \dot{f}(t) = L(f)v_{cam}$$

the following control rule is obtained:

$$v_{cam} = (L^T L)^{-1} L^T \dot{e}(t)$$

Since the Jaco² arm drivers provide a [ROS](#) topic to control the end effector in Cartesian velocities, the assumption holds. In this motion technique, the stop condition is given by the error between the target features position and the actual feature position $e(t)$, computed starting from the desired and actual position for each feature f :

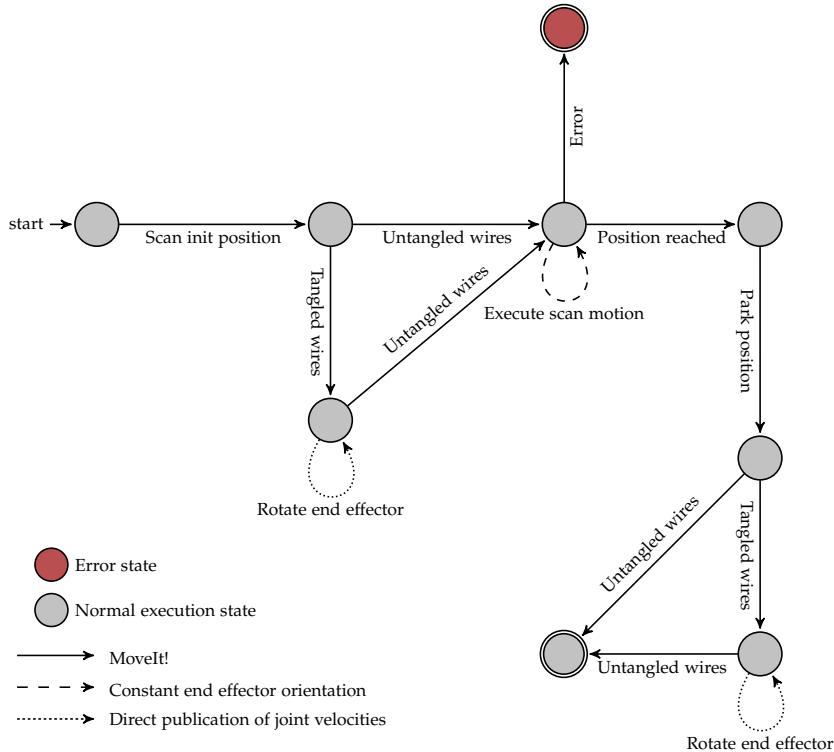
$$\begin{aligned} e_{fx}(t) &= (x_{actual} - x_{target}) \\ e_{fy}(t) &= (y_{actual} - y_{target}) \\ e(t) &= \sqrt{\sum_f e_{fx}^2(t) + \sum_f e_{fy}^2(t)} \end{aligned}$$

being below a predetermined threshold. The usage of visual servoing techniques in our implementation will be more clear in Section [6.4](#).

SCAN MOTION ACTION SERVER

The scan motion is a quite simple procedure; in Figure [6.3](#) you can see the finite state automata triggered by the action call, with highlight of the different motion methods used for each of the motions. As you can see, most of the movements are implemented through *MoveIt!* framework, except for:

- the disentanglement of the wires after critical movements. The wires problem was mostly accentuated in the scan movement because some of the joints were more stretched, so a check is made before and after the scan motion. The motion is implemented with direct publication of end effector joint velocity.
- the scan motion itself, that requires constant orientation of the [LIDAR](#) and so of the end effector, too. The motion is implemented through inversion of differential kinematic to maintain the end effector orientation.



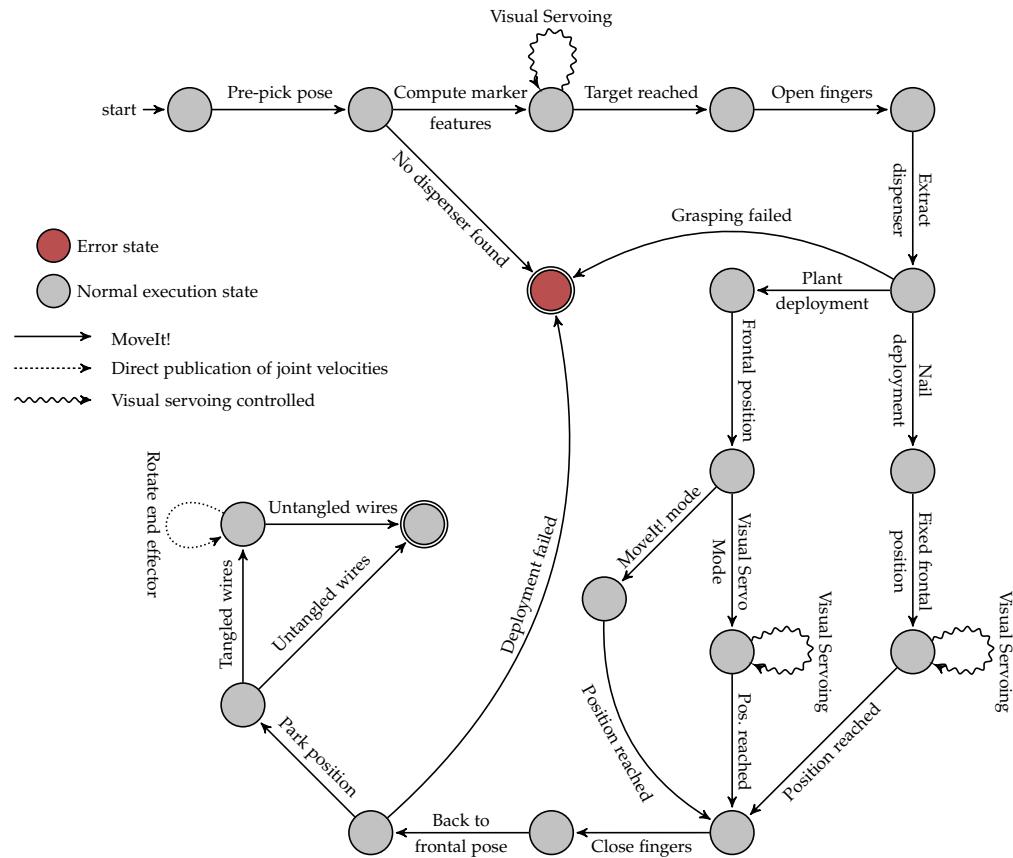
DISPENSER APPLICATION ACTION SERVER

In Figure 8.2.2 you can see the finite state automata triggered by the action call, with highlight of the different motion methods used for each of the motions. As in the previous action, most of the movements are executed with *MoveIt!*, except for:

- the disentanglement of the wires, as in scan motion action server.
- the grasping of the dispenser from the feeder, executed with visual servoing control.
- the deployment of the dispenser in nail mode, executed with visual servoing control.
- one of the dispenser deployment mode on the plant, executed with visual servoing.

This action is composed of two main parts:

1. Grasping of the dispenser, and image-based grasping validation.
2. Deployment of the dispenser, and image-based deployment validation.



Our image processing tasks are based on C++ API of **OpenCV library**³. Implementation detail are provided in Sections 6.4.1 and 6.4.2.

Image based validation

In order to check for the presence of dispensers in the feeder, and validate the grasp and deploy operations, we make use of the Intel RealSense RGB-D camera mounted on top of the arm end effector. The thinness of dispensers with respect to the camera resolution make them hard to detect, especially if, after a deployment operation, a rotation of the dispenser occurs and it's not possible framing it from a frontal position. For this reason, all our dispensers are equipped with two small red wings built out of tape, one on each side, as showed in Figure 4.10. Note that the red color of the wings was chosen because it stands out with respect to the typical colors of the vineyard during winter or spring months, and it's an important factor in the image processing algorithms that follow.

CHECK DISPENSER AVAILABILITY In this phase, the camera is placed on the vertical of the feeder and, to check for dispenser availability, it looks for two red wings in a central rectangular area (*valid*

³ <https://opencv.org/>

area) that should contain the feeder. In Figure 6.5, you can see some intermediate results of the algorithm execution:

1. Color segmentation of the image, to find pixel areas with color similar to the wings (see 6.5a).
2. Morphological operations (*dilatation* and *erosion*) on the selected pixel regions, for noise reduction and discard of too small regions (see Figure 6.5b).
3. Discard of the regions outside a predefined rectangular area in the center of the image (see Figure 6.5c).

In Figure 6.5d you can see the outcome of the aforementioned analysis; the wings identification is successful and their detected contour highlighted. The whole procedure is declared *failed* if no marker is detected in the image plane for a given amount of time (e.g., 5 seconds).

GRASPING VALIDATION As you can see in Figure 8.2.2, this check is performed after the fingers has opened and the end effector has already moved from the grasping pose but it still is on the vertical of the feeder, because we want to detect a possible failure as soon as possible. In case of successful grasping, we expect to see the red wing of the dispenser in the bottom right corner of the image (*valid area*). The copious amount of tests showed that that the grasping procedure couldn't lead to a deterministic position of the wing in the image plane, thus we had to take additional measure for the validation procedure to be effective. In particular, in the laboratory environment where our procedure were developed, problems came from:

- the color of the test bench similar shade of the dispenser wings color, especially under direct enlightenment.
- the *specular component* (see Figure 6.2) of the reflection from the desk surface causes noise in the depth sensing when using stereoscopic infrared sensors (Bhat and Nayar, 1998), in particular the depth of pixels belonging to highly reflective surfaces is always very low. For this reason, discarding pixel regions by thresholding over the distance of the regions from the camera is useful, but not sufficient to discard all the noise regions.

Because of the non-deterministic position of the dispenser in the image, we could not simply discard these noise by cropping a narrower valid area or setting a stricter threshold value for the region area. We noted that noise regions that survive the first steps of the filtering procedure often have a jagged contour, while we know that the dispenser wings have very sharp edges in the image. Basing on this observation, we introduced

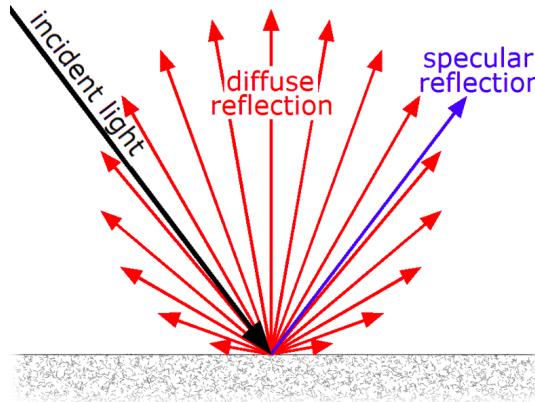


Figure 6.2: Representation of the reflection of light on a glossy surface: a ray incident on the surface is both scattered at many angles (**diffuse reflection component**), and reflected at one single angle (**specular reflection component**).

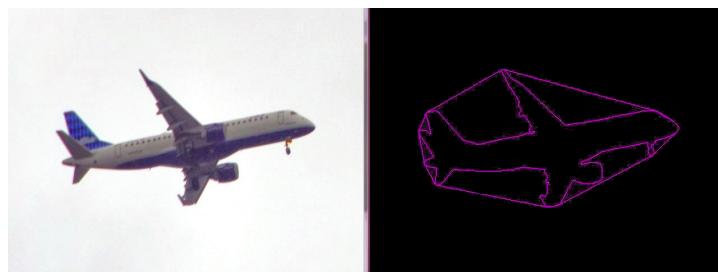


Figure 6.3: Convex hull computation of the edges of a sample image.

an additional filtering step, where for each of the pixel regions survived to previous filtering steps:

- compute the convex hull of that region, *i.e.*, the smallest convex set of pixels that contains it (see Figure 6.3).
- compute the matching between the region itself, and its convex hull; the OpenCV method for calculation of the match between different images are based on the Hu invariants (Hu, 1962).
- discard of the regions with matching value below a fixed threshold.

The idea behind this filtering procedure is that the trapezoid region of the dispenser is very regular and convex so it will be very similar to its convex hull, while the jagged contour of a noise region will be quite different from its convex hull.

Even if these problems were due to elements in the test environment and not in the actual vineyard, we decided to introduce the additional measure also in the final version of the procedure to make it more robust to possible noise we were not aware of in development phase.

The algorithm steps for the validation of the dispenser grasping are:

1. Color segmentation of the image, to find pixel areas with color similar to the wings (see Figure 6.6a).
2. Morphological operations (*dilatation* and *erosion*) on the selected pixel regions, for noise reduction and discard of too small regions (see Figure 6.6b).
3. Discard regions with too high depth value, *e.g.*, wings of dispensers still in the feeder, that are not took into account (see Figure 6.8b).
4. Discard regions outside the valid area (bottom right corner).
5. Discard region with a too low degree of similarity with their convex hull.

If at least one region survives the whole filtering procedure, this region is likely to be the one identified by the dispenser and the grasping is validated. NOTE: we check for *at least* one region instead of *exactly* one, because the procedure is much more likely to select noise regions as correct than discard the dispenser region, so false negatives are very unlikely and if we checked for exactly one region we would be more likely to produce false positives. At each grasp validation request, the whole filtering procedure is executed on a fixed number of consecutive frames (*e.g.* 50 frames), and the final output is given by the majority vote of the cumulated executions. This procedure is made to reduce the effect of random noise in the measurement.

DEPLOYMENT VALIDATION This operation is similar to the previous one, because we assume that, if the deployment task has been successful, by moving the camera back and a few centimeters downward maintaining the deployment orientation, we expect to detect the dispenser in the captured image. In our implementation we do not deal with cases where the dispenser rotates and, from the side, the wings are not visible. There is a slightly difference between the case of deployment on the nail and on the plant, because:

- in nail mode, the camera position from which search for the dispenser wings is fixed.
- in plant mode, the camera position from which search for the dispensers wings is computed as fixed offset from the deployment point took as an input by the action server, while in this mode it's a fixed position.

Apart from that, the validation procedure is the same in both plant and nail mode, and it's composed by the following steps:

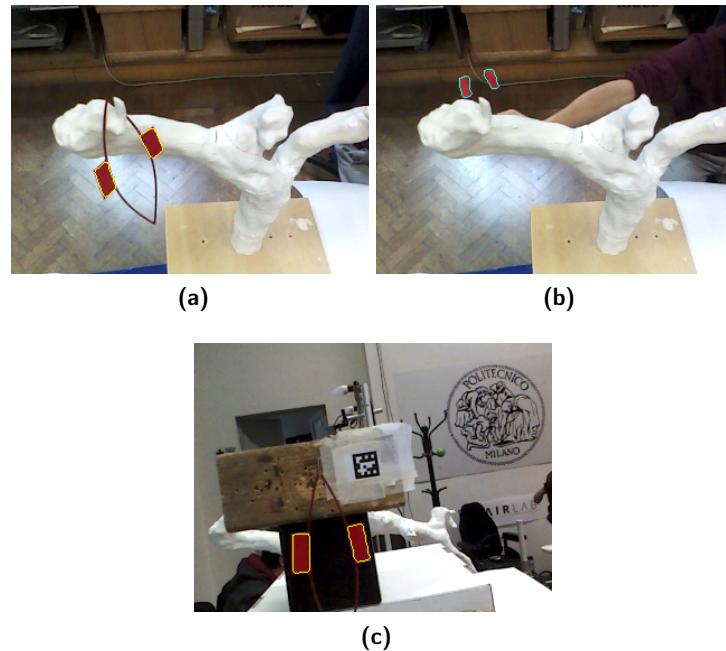


Figure 6.4: Different outcomes of the deployment validation algorithm: success in plant mode (Figure 6.4a), failure in plant mode (Figure 6.4b; the dispenser in the background does not induce a false positive because it does not satisfy depth requirements), success in nail mode (Figure 6.4c).

1. Color segmentation of the image, to find pixel areas with color similar to the wings.
2. Morphological operations (*dilatation* and *erosion*) on the selected pixel regions, for noise reduction and discard of too small regions.
3. Discard regions with too high depth value, *e.g.*, wings of dispenser deployed on other plants, in adjacent rows.
4. Discard region with a too low degree of similarity with their convex hull.

Like grasping validation, also the deployment validation procedure is executed on a fixed number of consecutive frames, that produce a cumulated result given by their majority vote. In Figure 6.4 you can see possible outcomes of the deployment validation.

Visual servoing tasks

In this section we'll describe with more detail each of the operations executed with visual servoing control. However, this is done mostly for the sake of completeness, because the design and implementation

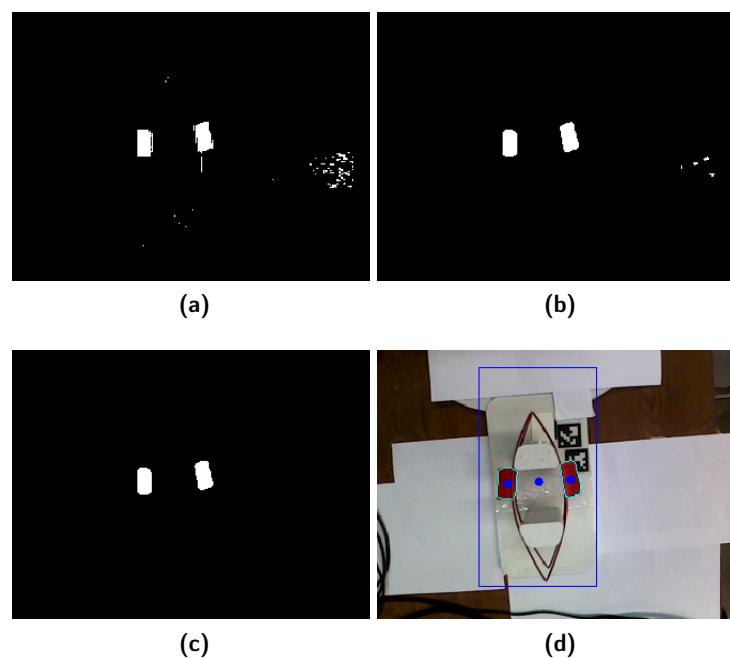


Figure 6.5: Visual representation of the dispenser availability check algorithm.

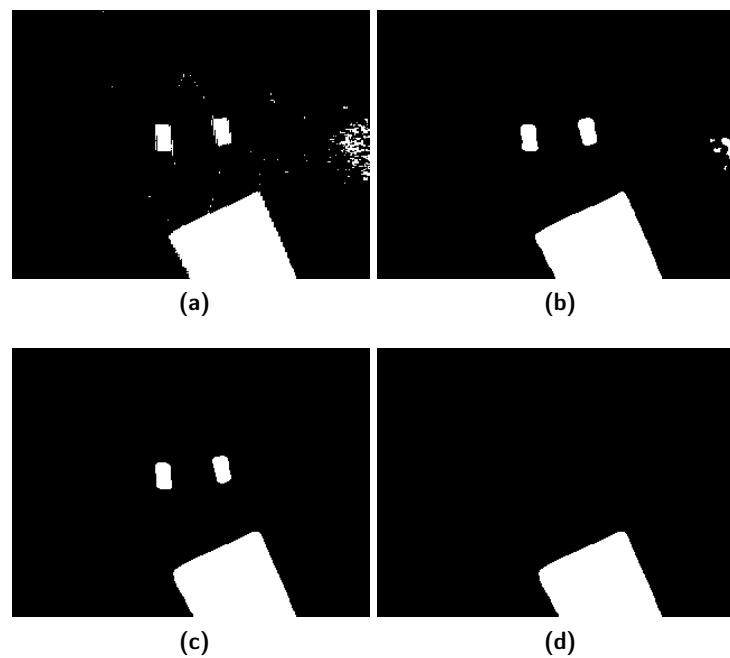


Figure 6.6: Visual representation of the dispenser availability check algorithm.

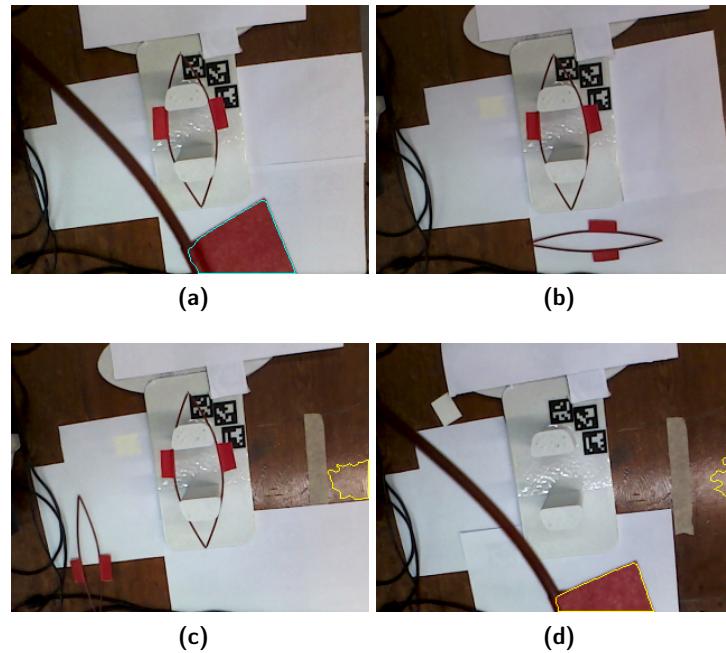


Figure 6.7: Different outcomes of the grasp validation algorithm: true positive (6.8a), true negative (6.8b), false positive (6.7c), error situation that however brings to a true positive (6.7d)

of the automatic control was not directly part of this thesis work, so technical details will be omitted.

As mentioned in Section 6.2.4, for the execution of image based visual servoing control, some features must be tracked in order to calculate the Cartesian velocities for the end effector. For this reason, all the arm motions controlled with visual servoing are preceded by a motion, executed through *MoveIt!*, that brings the end effector in a position that is frontal to the features that are tracked for the visual servoing control, such that they are present in image plane. For each case, we'll see which features are used and how. We'll see that some of the operations rely on ArUco markers (Munoz-Salinas, 2012); they are a type of binary square fiducial markers (see Figure 4.8), with some characteristics designed to make them particularly suitable for obtainment of the camera pose for computer vision applications:

- a black wide black border that facilitates its fast detection in the image.
- an inner binary matrix which determines its identifier and allows for the application of error detection and correction techniques.
- the marker size determines the size of the internal matrix, *e.g.*, a marker size of 4x4 is composed by 16 bits.

Make reference to the finite state automata of Figure 8.2.2 for better understanding of the operations in the whole context.

DISPENSER GRASPING

For this task, the tracked features are the 4 corners of an ArUco marker placed on the side of the dispenser feeder, in a fixed position with respect to the feeder center. For this task, the desired target positions of the marker corners (the tracked features) are computed starting from the detected position (in image frame) of the midpoint of the two dispenser wings, and the sensed depth of the wings pixel regions. This last information is required because the desired target of the end effector varies according to the position of the highest dispenser in the feeder.

In the first implementation of this operation problems occurred when, during the descent of the end effector toward the feeder, the *specular component* of the reflection from the marker surface pointed directly the camera eye. In that case the noise in the image was too high to detect the marker features. For this reason, a second marker with different ID (see Figure 6.8) was placed next to the previous one, on the same side of the feeder, in order to be always visible from the camera. Before visual servoing control kicks in, the target positions of the features of both markers are computed independently; the procedure has a "main" reference marker that, if detected, provides the features for the visual control; if the primary marker is not visible (see Figure 6.8a), the control procedure smoothly switches to use the features of the second marker for the visual control.

As explained in Section 6.2.4, if the number of features is enough, we can compute a full set of linear and angular Cartesian velocities

$$<\dot{x}, \dot{y}, \dot{z}, \dot{\psi}, \dot{\theta}, \dot{\phi}>$$

for the end effector. Since we have at disposal 4 features in this task, the end effector speed can be fully determined by visual servoing control.

DEPLOYMENT IN NAIL MODE

This task is a simpler version of the previous one, because:

- also in this case, the features are given by the corners of an ArUco marker, placed this time at a small fixed offset from the target nail (see Figure 6.4c).
- a single marker is used, because the specular reflection effect is much less relevant on a vertical surface.
- the target positions of the features are a fixed value, because there is no such a thing like different dispenser height in grasping phase.



Figure 6.8: Two markers used for visual servo task in grasping phase. In Figure 6.8b both the markers are detected, while in Figure 6.8a the dispenser with ID= 0 is not visible due to specular reflection on marker surface, but the other marker is still detected so the visual servo task can proceed.

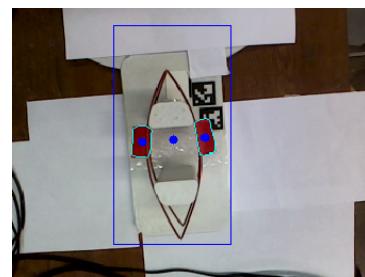


Figure 6.9: Example of the computed wings midpoint, used to compute the target position of the features (marker corners)

- the initial frontal position is not function of the deployment point passed in the action server goal, but is a fixed value.

As highlighted in the last point, this deployment procedure is 100% independent from the input deployment point. Thus, if this deployment mode is selected, scan phase is not required. Note that, also in this phase, we can rely on 4 different features (the marker corners) to be tracked in the control phase, so all 6 linear and angular velocities can be computed.

DEPLOYMENT IN PLANT MODE

This task is quite different from the previous ones, because of course we cannot take advantage of markers since our deployment point has been selected on the vine plant through the scan procedure described in Section 6.3, so we must use a different type of features. The visual feature used for this control task is a single point in the image, corresponding to the center of a circle, which contains points close to the deployment point in image plane, and with similar color. The idea behind it is to select a internal point of the same small branch that the input deployment point belongs to. In Figure 6.10 you can see some of the image processing steps that leads to the identification of the feature point. The target position of the feature in image frame, instead, is hardcoded to a specific pixel, corresponding to the position of the dispenser center. Note that, since we only have one feature at disposal, matrix $L(f)$ in this case is not large enough to provide a solution for all linear and angular velocities (see Equation 6.2). Indeed, in this deployment mode:

- $\dot{\psi} = \dot{\theta} = \dot{\phi} = 0$, i.e., the end effector orientation is constant
- \dot{z} is proportional to the distance between the end effector pose and the deployment point
- \dot{x} and \dot{y} are controlled by the visual servoing controller.

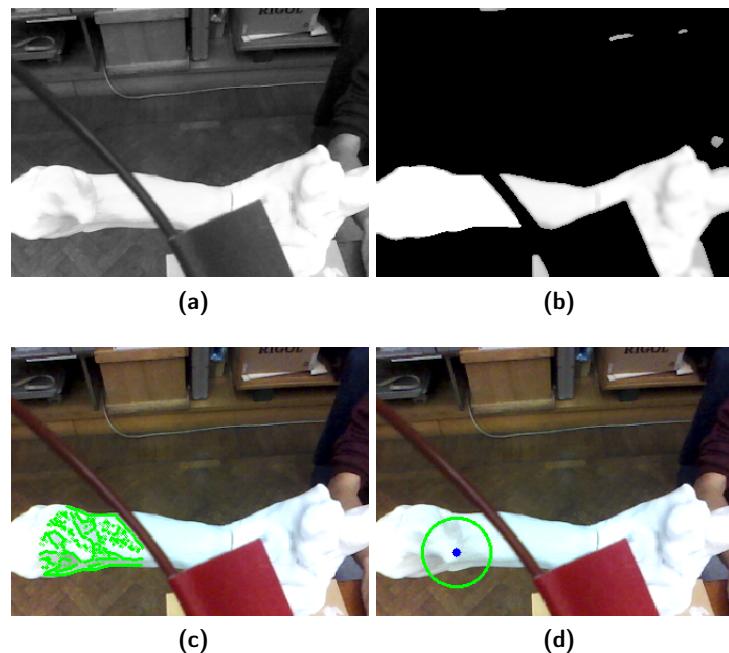


Figure 6.10: Some of the steps computed to identify the feature point tracked for dispenser application on the vine tree: switch from RGB to grayscale (Figure 6.10a), selection of pixels similar to the pixel corresponding to deployment point (Figure 6.10b), identification of branch area close to the deployment point (Figure 6.10c), identification of the circle including the selected points (Figure 6.10d).

EXPERIMENTAL RESULTS

On the field sessions were a crucial point in the validation of **GRAPE** system, for two main reasons:

1. Robotics is a field that, by definition, requires interaction with the physical world: both the components of the robot itself, and the environment. Thus, simulation could strongly speed up the development in certain phases, but the creation of a robotic system cannot opt out of experimental sessions. This consideration is especially true in agricultural robotics field, because a simulated environment or a laboratory environment are extremely likely to be too different from the actual robot environment to be significant, *e.g.* in terms of light conditions, terrain structure, sensor noise.
2. As explained in Section 3.1, the research team included members from both Politecnico di Milano and Eurecat, that are located in two different countries. Even if this should not have been a big problem, the division of labor created significant inefficiencies because:
 - even if the sensor fusion and navigation parts were assigned to Politecnico, the Husky unmanned ground vehicle was property of Eurecat, so it was in Barcellona.
 - even if the analysis of the vine scan in order to detect the most suitable deployment point was in charge of Eurecat, the arm was property of Politecnico so it was in Milano.

This problem was mitigated by the logging tool of **ROS** (*rosbag* package we mentioned in Section 2.1), that demonstrated its potential in such a situation: *bags* recorded during (non-autonomous) navigation sessions of the Husky were very useful to test the sensor fusion system, while *bags* of the laser scans sensed during execution of the scan motion in front of the mockup vine (owned by Politecnico team, see Figure 3.5), were used by Eurecat team to test the point cloud creation and analysis.

However, the usage of *bags* in our context showed two main weaknesses, the first of which is specific of our context, while the second is intrinsic to the use of recorded data:

- autonomous navigation algorithms cannot be tested by means of recorded data *bags* because, during an autonomous navigation session, the data that are sensed while using a specific navigation algorithm depend on the choices of the al-

gorithm itself; the same reasoning does not apply to neither sensor fusion nor point cloud analysis tasks. For this reason, the only way to test a navigation algorithm except for *on-the-field* experiments is via simulation by means of a physical simulator. Some effort were spent in this direction before the beginning of this thesis work (see Section 3.2), but very few *on-the-field* tests have been enough to show that the simulated environment was way too different from the real one for the test to be significant.

- validation of algorithms by means of recorded data significantly increases the risk of overfitting, *i.e.*, adapting the model to perform properly only on the data in the (few) *bags* at disposal, and not on new data coming, for example, from an *on-the-field* experiment. The best ways to decrease the probability of overfitting are:
 - use as many bags as possible to test the algorithms during development.
 - test the system in the real environment as often as you can. Remember, however, that sometimes overfitting can occur also in real environment, in less obvious way. For example, the ROAMFREE configuration executed in Garriguella during the first integration week was absolutely ineffective when applied in Casciana Terme vineyard, because it fitted too tightly the features of Garriguella vineyard thus it didn't represent a general solution.

SENSOR FUSION

Since the final sensor fusion system included data from GPS, we can exploit satellite imagery to better visualize and evaluate the sensor fusion output. For this purpose, we used the *Rviz_satellite* plugin for Rviz. In this section we'll show data recorded during a 13 minutes long autonomous navigation session of the Husky, executed across three different rows in the Garriguella vineyard, during the third integration week, in March 2018. In our case, satellite images also provide a very coarse-grained ground truth about the path, because we know the rows that the robot has actually traversed during navigation.

As explained in Section 5.1, we expected a very untrustworthy wheels odometry, and the experimental results actually proved it. In Figure 7.1a you can see the pose estimated by the raw wheels velocities only. It's easy to notice that there's no way to use it as unique odometry source: the wheels slippage, especially during rotational movements, makes the estimated path completely different from the

actual one. The data from GPS, shown in Figure 7.1b, are instead really accurate, since RTK correction was used.

Indeed, the final estimated odometry, that you can see in Figure 7.2, is very close to the plotted GPS data in Figure 7.1b. However, a sensor fusion based on GPS only would not be a real solution, because:

- the GPS update frequency is low (~ 1 Hz), so other sensors are required to improve the estimate between two GPS measures
- the GPS does not provide information about the orientation
- GPS availability, especially with RTK correction, is not guaranteed all over the world, so other sensors must be available to compensate if required

NAVIGATION

In the odometry system, the satellite imagery provided, as already stated, a sort of coarse-grained ground truth. The same reasoning cannot be applied to navigation operations, thus the evaluation of a navigation system by means of a quantitative metric is not easy.

Because of the tight deadlines in the GRAPE project, no particular effort was put in this topic, so we limited ourselves to observe high-level qualitative features of the local planner behavior (e.g. number of times the robot got stuck in obstacles during navigation, smoothness of the movement) to assess both the inadequacy of *Dynamic Window Approach* planner, and the proper functionality *Timed Elastic Band* planner.

In Figure 7.4 you can see a frame of the same autonomous navigation session in Garriguella mentioned in Section 7.1. The global navigation plan that the robot is following is drawn as a green line, while the shorter red line represents the local plan computed by *Timed Elastic Band*. The ability of this algorithm to deal with winding paths is clear in this image.

In Figure 7.3, you can see the result of mapping process executed with Gmapping as SLAM algorithm and Robot Localization as sensor fusion system. In both of them the vineyard structure is clearly visible.

MANIPULATION

As expected, the execution of arm movements through *MoveIt!* framework, when correctly embedded in the unmanned ground vehicle system, gave no surprises in *on-the-field* experiments. In Figure 7.7, you can see the 3D position of the end effector during a grasping and deployment operation, completely executed through *MoveIt!* except, of



(a)



(b)

Figure 7.1: Pose estimation in time using raw wheels odometry (Figure 7.1a), and using GPS (Figure 7.1b).

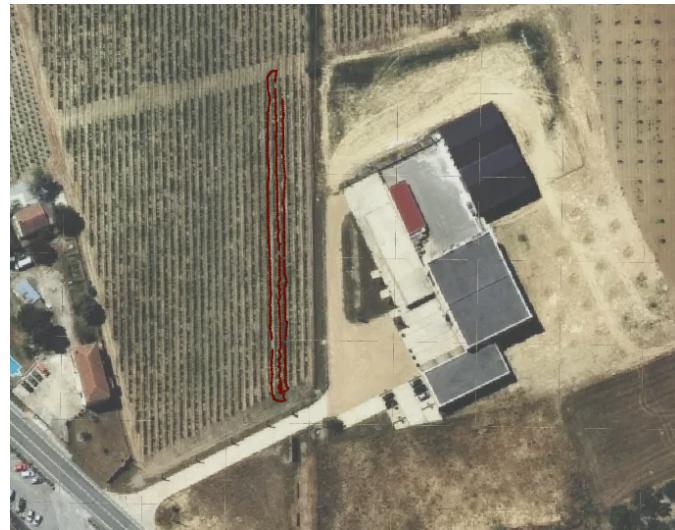


Figure 7.2: The odometry estimated by Robot Localization sensor fusion framework, fusing (among others) the pose estimated by the wheels (see Figure 7.1a) and by the GPS (see Figure 7.1b)

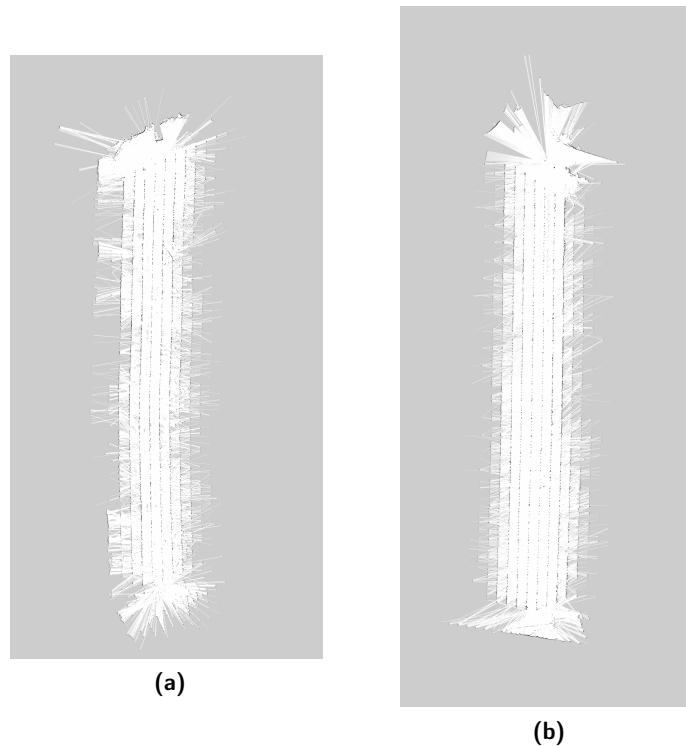


Figure 7.3: Output of mapping phase, using Gmapping as SLAM algorithm and Robot Localization as sensor fusion framework. The map of Garriguella vineyard in Figure 7.3a, the Casciana Terme one in Figure 7.3b.

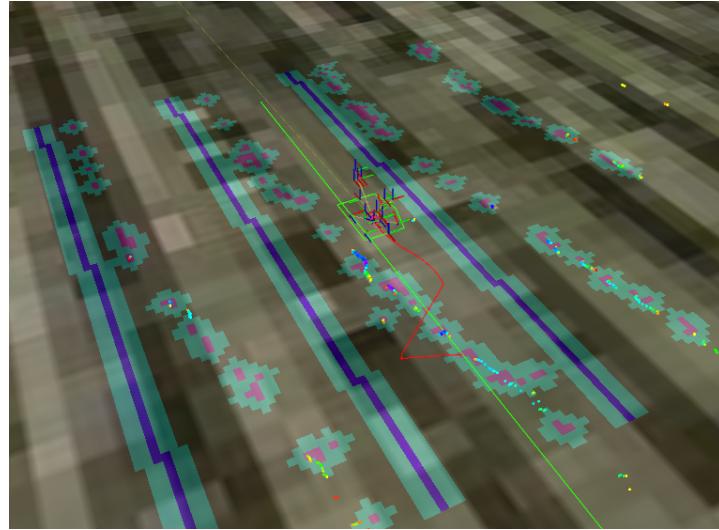


Figure 7.4: Visualization through Rviz of the Husky performing autonomous navigation, with the local costmap overlaid to the satellite images of Gariguella vineyard. You can see the global (green line) and local (red line) navigation plans, the tf tree of the robot and the Husky footprint used by Move Base for the interaction with the costmap.

course, for the core movement of the grasping, the movement 3→4 in the Figure. The evolution of the z coordinates of the end effector during segment 3→4 is shown in Figure 7.8.

On the other hand, manipulation sub-operations that relied on computer vision techniques were expected to have problems due to the unpredictable time varying light conditions in outdoor context. A large amount of tests showed two different outcome to *on-the-field* validation:

- computer vision tasks based on tracking of ArUco markers (dispenser grasping from the feeder, dispenser application on the nail) turned out to be robust even in presence of the aforementioned light conditions. This is mainly due to the robustness of OpenCV library algorithms in the detection of the markers.
- computer vision tasks based on the detection of the red wings of the dispensers (check for dispenser availability in the feeder, grasping and deployment validation) were significantly influenced by the light changes, and they were weaker in the hours about noon, when the sun was on the vertical.

In Figure 7.5 and 7.6, quantitative analysis of visual servo driven movements are provided, through the evolution in time of both the 8 tracked features ((x, y) coordinates for each of the marker corners) and the module of the overall error of the corners with respect to their desired positions. Figure 7.5 data comes from the deployment of a dispenser on a nail, while data of Figure 7.6 comes from the grasping of a dispenser from the feeder. In the graphs related to the feature

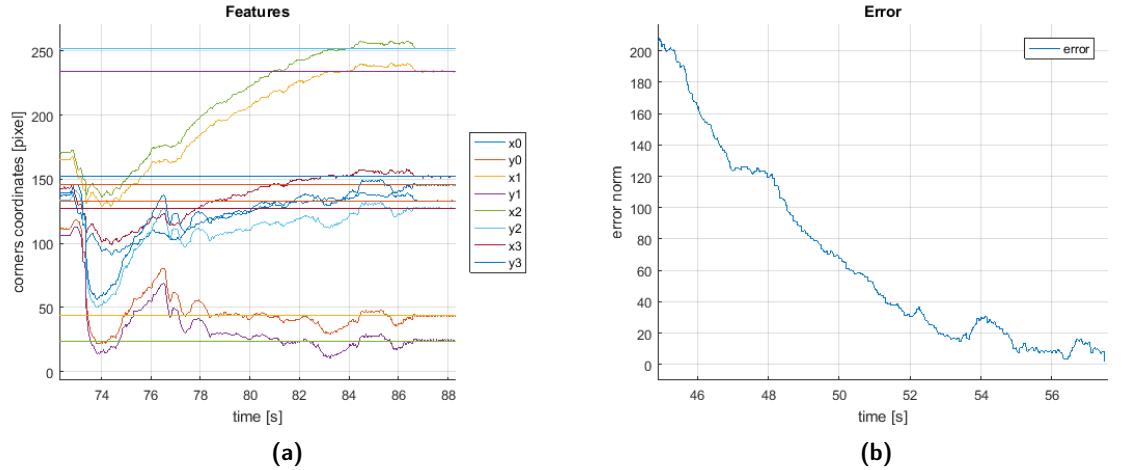


Figure 7.5: Evolution in time of the image coordinates of the corners of the tracked marker (see Figure 7.5a) during visual servoing driven deployment of a dispenser on a nail. In Figure 7.5b, the evolution in time of the module of the error on the tracked features.

tracking, the desired value for each of the feature is represented with an horizontal line; notice that, at the end of the execution, each of the features is almost equal to the its target value.

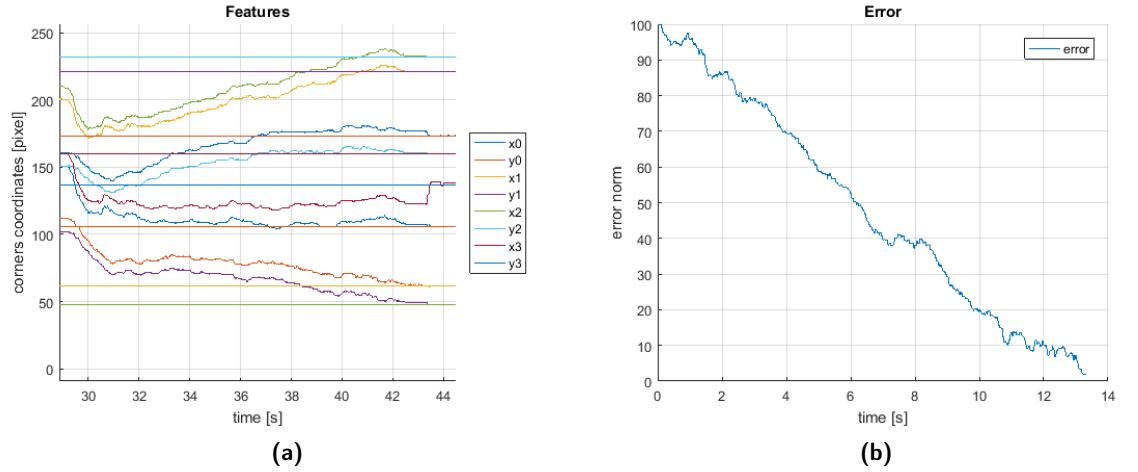


Figure 7.6: Evolution in time of the image coordinates of the corners of the tracked marker (Figure 7.6a) during visual servoing driven grasping of the dispenser from the feeder. In Figure 7.6b, the evolution in time of the module of the error on the tracked features,

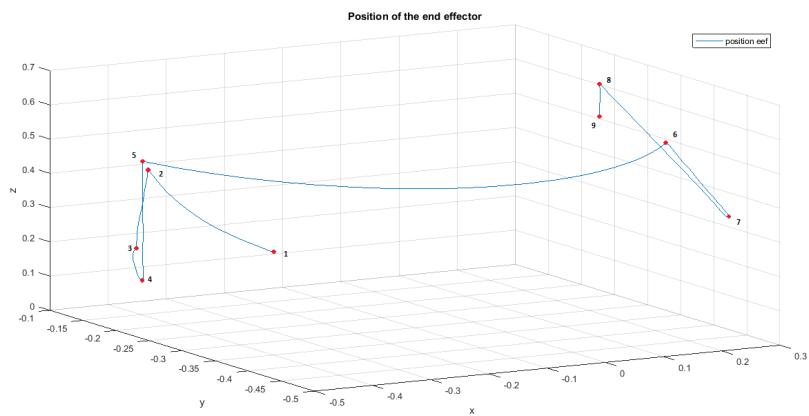


Figure 7.7: Evolution of the 3D (x , y , z) position of the end effector of the Jaco² during dispenser deployment using MoveIt!

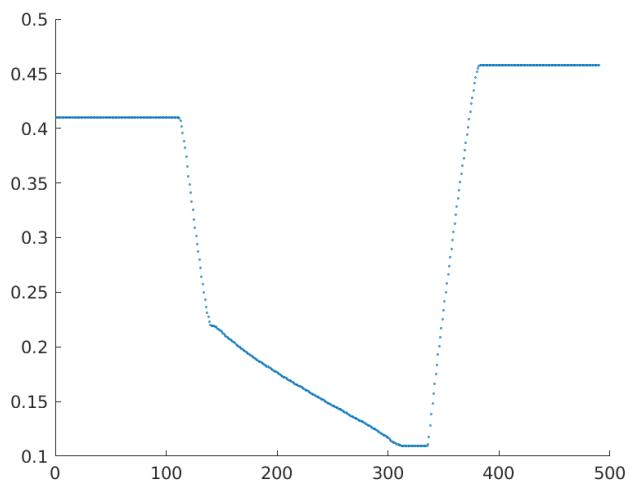


Figure 7.8: Evolution of the z coordinate of the end effector, during grasping operation; visual servo driven movement is the central one in the graph.

THE ROBÌ PROJECT

A branch of this thesis work involved the customization of the prototype mobile manipulator Robì (Bascetta, Baur, and Gruosso, 2017) to fit the requirements of the mobile base requested by the **GRAPE** project. The configuration of Robì platform, unfortunately, didn't get to the end because the tight deadlines of **GRAPE** project left no room for working on it, thus the only complete and tested platform was the one with the commercial unmanned ground vehicle, Husky from Clearpath Robotics. However, since significant steps were made into the configuration of Robì prototype for an actual agricultural robotics task, it's worth describing the developed platform in this Chapter.

ROBÌ MOBILE MANIPULATOR

Robì (see Figure 8.1) is a prototype small-sized mobile manipulator for agricultural applications, whose aim is to support the development and testing of innovative perception and control algorithms. Both the mechanical structure and the motion control system are designed to be simple, flexible, low-cost and low-weight, to create a system that can act as an open source base for project in agricultural robotics, contrary to the large amount of task-specific platforms in agricultural robotics (*e.g.*, for asparagus (Chatzimichali, Georgilas, and Tourassis, 2009) or tomato (Yaguchi et al., 2016) harvesting).

Since the platform is thought as a versatile mobile base to be adapted to several fields (*e.g.*, vineyards, herbaceous plants) its chassis is designed to have flexible geometrical characteristics (ground clearance, track, wheelbase).



Figure 8.1: The Robì mobile manipulator chassis.

Ground clearance	[0.25, 0.35] m
Track	[0.75, 1.2] m
Wheelbase	[0.6, 1] m

Table 8.1: Ranges of Robì geometrical characteristics.

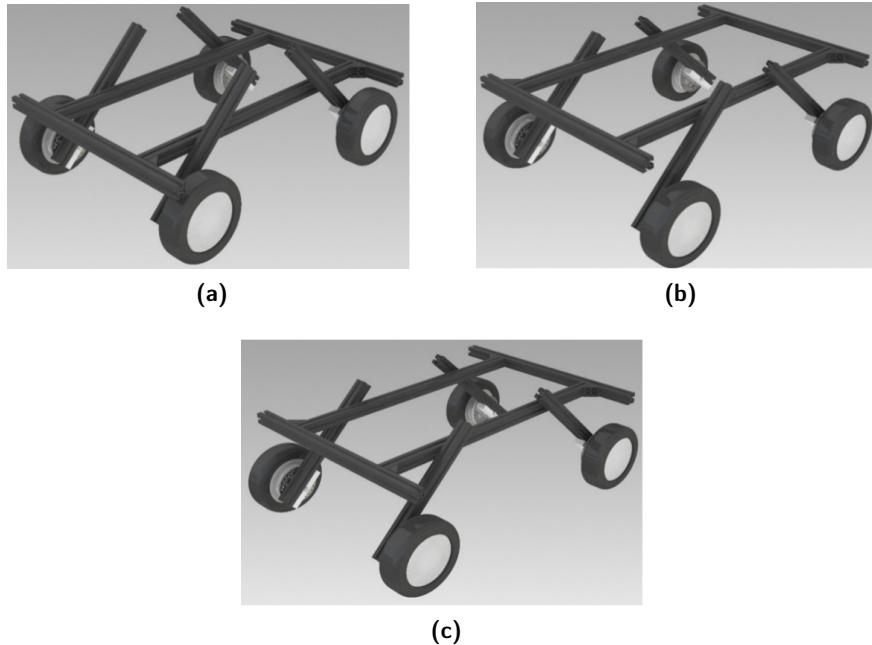


Figure 8.2: 3D rendering of different configuration of Robì base, obtained thanks to the chassis mechanical design.

More generally, the design principles of Robì are:

- **low cost**, to make it easier to use it for example applications and, possibly, fleet-based applications.
- **low weight**, to increase the battery life and, consequently, allow for more long and versatile missions,
- **simple mechanical design**, to make it easy to build it out of a mounting kit.

The aforementioned principles are implemented through the following design choices:

CHASSIS

The chassis is made out of ITEM® aluminium bars, that presents lightweight but strong section; the slide rails embedded in the ITEM® bars, together with four rotational joints applied to the bars supporting the wheels, allow for multiple configuration of the robot geometrical characteristics to vary in the ranges listed in Table 8.1. In Figure 8.2 you can see 3D rendering of some



Figure 8.3: HUB10GL, the in-wheel motor model adopted in Robì.

of the available Robì configurations. The bars that support the wheel, however, can't rotate during the robot movement, thus Robì has a skid steering kinematic model.

IN-WHEEL MOTORS

Robì is powered by four in-wheel electric DC brushless motors; the model is HUB10GL (see Figure 8.3). In-wheel motors have been chosen mostly because they allow for much simpler mechanical design, with respect to classical electrical powertrain, because:

- no transmission is required.
- the weight is reduced, and the available space on the chassis increases.
- high level of maneuverability, without the need to introduce Ackermann kinematics.

The motors are independent, so a suitable control board is required to control each of them; the board used in Robì are the ones described in VESC project¹, an open source brushless motor control project. These boards offers several interfaces to control the motors: PPM signal, analog, UART, I²C, USB or CAN-bus.

ROBÌ AS GRAPE ROBOTIC BASE

The features mentioned in the previous section made Robì a convincing candidate as the mobile platform for the **GRAPE** project, because of its flexibility as a mobile manipulator for agricultural robotics field. Unfortunately, as stated at the beginning of this Chapter, for mere lack of time the Robì platform didn't get mature enough to be actually employed in the context of **GRAPE**. In this section, we describe

¹ <http://vedder.se/2015/01/vesc-open-source-esc/>

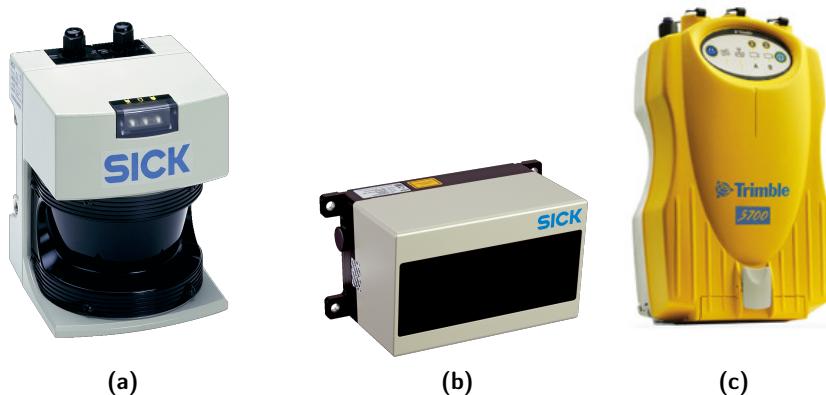


Figure 8.4: Some of the sensors Robì is equipped with: LMS291 (8.4a), and LD-MRS400001 (8.4b), both laser scanners from SICK; Trimble 5700 GPS transceiver (8.4c), from Trimble

the steps taken towards the use of Robì as a mobile manipulator in vineyard environment.

Hardware configuration

Since our goal was to create the system that was later developed on the Husky robot, similarity occurs in the sensors choice; on the other side, differently from Husky configuration, we had to dedicate significant efforts in the creation of an hardware interface between the on-board computer (a commercial HP laptop) and the motors control boards. Given the flexibility of Robì physical structure, we decided to add a simple structure in the frontal part of Robì as a support for the bulkier sensors (see Figure 8.6), built out of ITEM® bars for compatibility. Note that the Jaco² was still the designated arm in this configuration (see Figure 8.6b).

LIDARS

In section 8.1 we pointed out that Robì, as the Husky robot, comes up with a skid steering kinematic model. Thus, we chose to use two frontal LIDARs for obstacles avoidance during navigation tasks. The LIDAR models were:

- **Sick LMS291:** this sensor² only has a single measuring plane, maintained parallel to the ground, in order to detect long-range obstacles during navigation.
- **Sick LD-MRS400001:** this sensor³ is a multi-range laser scanner, with 4 measuring planes with narrow aperture

² <https://www.sick.com/ag/en/detection-and-ranging-solutions/2d-lidar-sensors/lms2xx/lms291-s05/p/p109849>

³ <https://www.sick.com/de/en/detection-and-ranging-solutions/3d-lidar-sensors/ld-mrs/ld-mrs400001/p/p112355>

LMS291 LD-MRS400001		
Number of Channels	1	4
Scan Angle	180°	85°
Rotation rate	75Hz	12.5 Hz
Angular Resolution	0.25°	0.125°
Range	80m	300m
Power Consumption	30W	8W
Weight	4.5kg	1kg

Table 8.2: Comparison of Robì on-board LIDARs

angles. Unlike *LMS291* scanner, its arrangement in the sensors support bar is slightly tilted forward, to be able to scan closer obstacles.

GPS

We opted for a very reliable solution, the **Trimbe 5700** GPS receiver (see Figure 8.4c) from Trimble, together with its dedicated antenna.

RADIO

In order to easily control and monitor the activity of the unmanned ground vehicle especially during outdoor tasks, we also installed on Robì the required hardware to create a *point-to-point* radio bridge with another base station (*e.g.*, another laptop in a more comfortable place). Two base stations were chosen, which performance were measured in urban environment (because, as already stated, the development of Robì as *GRAPE* robotic base was aborted before *on-the-field* validation):

- **Rocket M5** station⁴ from Ubiquiti, together with an omnidirectional antenna, mounted directly on Robì.
- **NanoStation M5** station⁵ from Ubiquiti, to be connected to the external device.

This couple of devices were able to durably create a radio link of ~340 meters.

MOTORS INTERFACE

In order to communicate with the motors control boards, we had to introduce a single hardware interface among them and the on-board PC. We decided to use **NUCLEO-F746ZG** development board⁶ from ST (see Figure 8.5a). It provides:

⁴ https://dl.ubnt.com/datasheets/rocketm/RocketM_DS.pdf

⁵ https://dl.ubnt.com/datasheets/nanostationm/nsm_ds_web.pdf

⁶ <http://www.st.com/en/evaluation-tools/nucleo-f746zg.html>

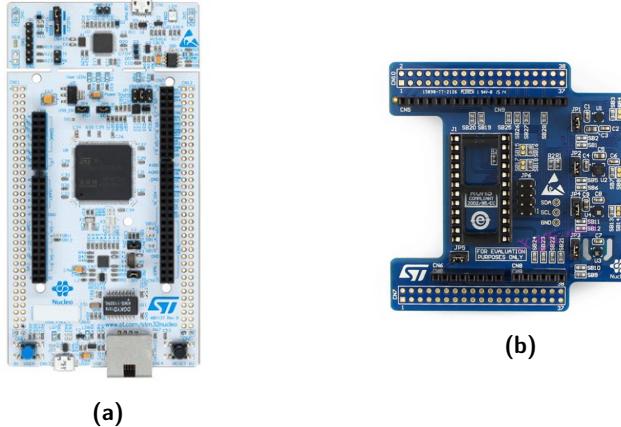


Figure 8.5: The Nucleo board we used to communicate with the motors control board (Figure 8.5a), and the shield mounted on it to integrate IMU sensor (see Figure 8.5b).

- RJ-45 Ethernet interface, very useful to communicate with the on-board PC.
- CAN protocol interface, to communicate with the motors control boards.
- ST Zio connector, which extends the Arduino Uno V3 connectivity and provides an easy means of expanding the functionality of the Nucleo platform with a wide choice of specialized shields.

In particular, the interface with the on-board PC was implemented as a bidirectional UDP connection. This will be cleared out in Section 8.2.2.

IMU

For the choice of IMU, we decided to take advantage of the Zio interface of the Nucleo board, so we inserted in the system the X-NUCLEO-IKS01A1 shield⁷ (see Figure 8.5b). This shield includes, among others:

- LSM6DS0 MEMS 3D accelerometer and 3D gyroscope
- LIS3MDL MEMS 3D magnetometer

Software architecture

We couldn't test any business logic on Robì platform, again because it didn't reach the main phases of GRAPE project as the chosen platform. However, a possible platform switch would take great advantage from ROS intrinsic modularity. In the ROS computational graph, the only components influenced by the underlying hardware are:

⁷ <http://www.st.com/en/ecosystems/x-nucleo-iks01a1.html>



(a)



(b)

Figure 8.6: The Robì platform, equipped with the sensors required by [GRAPE](#) project.

- the topic where to publish the target linear and angular velocities for the unmanned ground vehicle (typically, `cmd_vel` topic).
- the topic where to read the odometry estimated only from the wheel movement (in the Husky, `/husky_velocity_controller/odom` topic).
- the topics where to read the disparate sensors measurements.
- `tf` topic, where the complete `tf` tree of the unmanned ground vehicle is published.

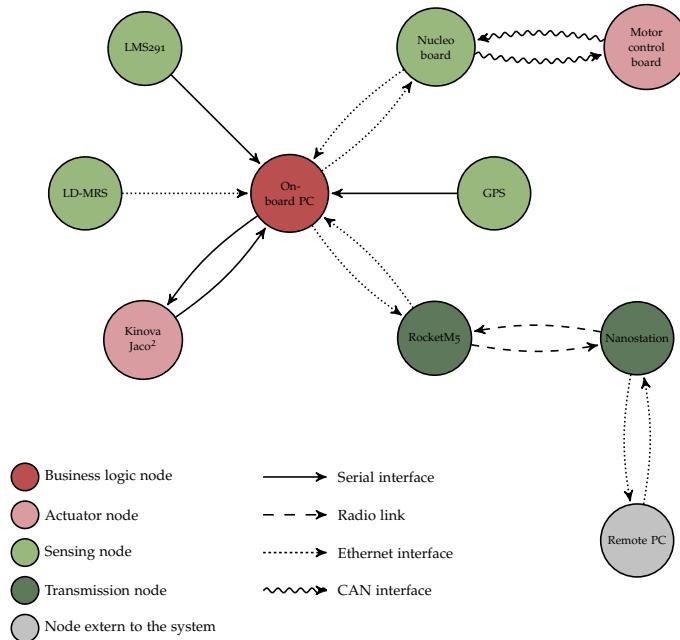
For these reason, we are quite confident that, excluding the unavoidable changes at least in Move Base configuration, the platform switch *should* not be too challenging.

Anyway, we are going to show the low-level software architecture of the system, by means of a graph where:

- the nodes are hardware components, that generate or process data.
- the (oriented) edges represent the data flow, from the node that generates data (tail) to the node that process them (head); different arrow types represent different communication strategies.

Note that:

- actuator nodes both receive and generate data, because they both accept commands and generate feedbacks.
- Nucleo board both provides data from IMU shield and wheels feedback (outgoing arrow), and accepts velocity commands for the motors (ingoing arrow). Both communications are implemented via UDP sockets.



CONCLUSIONS

In this work we designed, developed and validated the software components of an autonomous mobile manipulator. The system is equipped with an odometry system that robustly provides an accurate estimation of the robot geolocalized pose and orientation, exploiting data from GPS, IMU, and wheels velocities. The pose estimation is used by the platform to achieve autonomous navigation on steep, rough, and very bumpy terrain and, by means of a single-plane frontal LIDAR, the vehicle is able to deal with irregular obstacles such as the vines trunks and other accidental vegetation (*e.g.*, high weeds) in the vineyard. AMCL localization algorithm was not able to deal with the high degree of repetitiveness of the vineyard map, so we opted for the substitution of the global map with an handcrafted *prohibition layer* to represent the vineyard rows as obstacles. Even if this operation has to be made by hand, we consider it more efficient solution than the map usage because:

- the vineyard rows are very simple and regular obstacles to be specified by hand, since they are identified by only two geolocalized endpoints each; thus, GPS coordinates of endpoints can also be extracted via satellite imagery.
- the *prohibition layer* specification is easier and faster than traditional mapping phase, and in our case fully substitutes it.
- the inclusion of *prohibition layers* in both local and global costmaps avoids row crossing problem.

The main goal of GRAPE project, the automatic deployment of pheromones dispensers in the vineyard, is executed by means of the on-board manipulator, and validated via computer vision techniques designed to detect the expressly-placed dispenser's red wings in the image frame. We designed two different execution modes for the dispenser deployment operation:

1. application of the dispenser in a point output by the processing of a 3D reconstruction of the target vine.
2. application of the dispenser on a target nail, via visual servoing control with the help of specifically placed binary fiducial markers.

These solutions can be selected according to the willingness of the wine growers not to modify the vineyard structure at all (solution 1), or to carry out a simple preparation of the vineyard in order to enjoy

an easier and more robust dispenser deployment procedure (solution 2).

After two *on-the-field* sessions, we can certify the robustness of sensor fusion, localization, and manipulation systems in the actual vineyard environment. However, we were also able to identify some weaknesses in the resulting platform, that could be addressed in future work:

- during the system development, no particular emphasis to the power consumption problem was given; thus, the current power supply system of the Husky can support the platform for about 4 hours. Even if the change of batteries is a simple and fast operation, it's carried out by humans and it largely decreases the platform potential. A power-aware restyling of both software and hardware components of the system would be one of the main improvements of the platform.
- the sensor fusion system strongly relies on high-precision GPS, that actually is not available worldwide. Navigation system is the component that would mainly suffer from faulty localization, so a possible solution to disengage it in some measure from high-precision GPS would be to introduce a (maybe partially) image-driven navigation system, that exploits the regularity of vine rows to select the navigation direction.
- even if we had difficulties to include [AMCL](#) algorithm in the final sensor fusion system (see Section 5.2.5), we think that a successful inclusion could be very useful for the development of a sensor fusion system that, given the well known potential of [AMCL](#) in indoor localization, could effectively estimate the robot pose both in indoor and outdoor environment.
- the computer vision operations that do not rely on binary fiducial markers (*i.e.*, validation of grasping and deployment of the dispenser, detection of dispenser availability in the feeder) turned out to be weaker than expected when executed in strong and direct sunlight conditions, but for mere lack of time we could not directly address this problem. A more fine-grained configuration of the procedures might be sufficient to significantly improve their performance.

BIBLIOGRAPHY

- Bascetta, Luca, Marco Baur, and Giambattista Gruosso (2017). "ROBI': A Prototype Mobile Manipulator for Agricultural Applications." In: *Electronics* 6.2, p. 39 (cit. on p. 92).
- Bhat, Dinkar N and Shree K Nayar (1998). "Stereo and specular reflection." In: *International Journal of Computer Vision* 26.2, pp. 91–106 (cit. on p. 73).
- Calabrese, Luca (2014). "Robust odometry, localization and autonomous navigation on a robotic wheelchair." In: (cit. on p. 18).
- Carde Ring T Minks, Albert K (1995). "Control of moth pests by mating disruption: successes and constraints." In: *Annual review of entomology* 40.1, pp. 559–585 (cit. on p. 29).
- Chatzimichali, Anna P, Ioannis P Georgilas, and Vassilios D Tourassis (2009). "Design of an advanced prototype robot for white asparagus harvesting." In: *Advanced Intelligent Mechatronics, 2009. AIM 2009. IEEE/ASME International Conference on*. IEEE, pp. 887–892 (cit. on p. 92).
- Chaumette, François and Seth Hutchinson (2006). "Visual servo control. I. Basic approaches." In: *IEEE Robotics & Automation Magazine* 13.4, pp. 82–90 (cit. on p. 68).
- Colledanchise, Michele and Petter Ögren (2017). "Behavior Trees in Robotics and AI, an Introduction." In: *arXiv preprint arXiv:1709.00084* (cit. on p. 47).
- Cucci, Davide A and Matteo Matteucci (2014). "Position tracking and sensors self-calibration in autonomous mobile robots by gauss-newton optimization." In: *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, pp. 1269–1275 (cit. on p. 18).
- Cucci, Davide Antonio and Matteo Matteucci (2013). "A Flexible Framework for Mobile Robot Pose Estimation and Multi-Sensor Self-Calibration." In: *ICINCO* (2), pp. 361–368 (cit. on p. 18).
- Diago, Maria P, Javier Tardaguila, et al. (2015). "A new robot for vineyard monitoring." In: *Wine & Viticulture Journal* 30.3, p. 38 (cit. on p. 32).
- Dudek, Wojciech, Wojciech Szynkiewicz, and Tomasz Winiarski (2016). "Nao robot navigation system structure development in an agent-based architecture of the RAPP platform." In: *Challenges in Automation, Robotics and Measurement Techniques*. Springer, pp. 623–633 (cit. on p. 18).
- Elmenreich, Wilfried (2002). "Sensor fusion in time-triggered systems." In: (cit. on p. 16).

- Foote, Tully (2013). "tf: The transform library." In: *Technologies for Practical Robot Applications (TePRA), 2013 IEEE International Conference on*. Open-Source Software workshop, pp. 1–6 (cit. on p. 9).
- Fox, Dieter, Wolfram Burgard, and Sebastian Thrun (1997). "The dynamic window approach to collision avoidance." In: *IEEE Robotics & Automation Magazine* 4.1, pp. 23–33 (cit. on p. 59).
- Fox, Dieter, Wolfram Burgard, Frank Dellaert, and Sebastian Thrun (1999). "Monte carlo localization: Efficient position estimation for mobile robots." In: *AAAI/IAAI 1999*.343-349, pp. 2–2 (cit. on p. 20).
- Gage, Douglas W (1995). *UGV history 101: A brief history of Unmanned Ground Vehicle (UGV) development efforts*. Tech. rep. NAVAL COMMAND CONTROL, OCEAN SURVEILLANCE CENTER RDT, and E DIV SAN DIEGO CA (cit. on p. 41).
- Hart, Peter E, Nils J Nilsson, and Bertram Raphael (1968). "A formal basis for the heuristic determination of minimum cost paths." In: *IEEE transactions on Systems Science and Cybernetics* 4.2, pp. 100–107 (cit. on p. 23).
- Herring, David (2001). "Precision farming: Feature articles." In: (cit. on p. 28).
- Hinkel, Georg, Henning Groenda, Lorenzo Vannucci, Oliver Denninger, Nino Cauli, and Stefan Ulbrich (2015). "A domain-specific language (DSL) for integrating neuronal networks in robot control." In: *Proceedings of the 2015 Joint MORSE/VAO Workshop on Model-Driven Robot Software Engineering and View-based Software-Engineering*. ACM, pp. 9–15 (cit. on p. 34).
- Hu, Ming-Kuei (1962). "Visual pattern recognition by moment invariants." In: *IRE transactions on information theory* 8.2, pp. 179–187 (cit. on p. 74).
- Hutchinson, Seth, Gregory D Hager, and Peter I Corke (1996). "A tutorial on visual servo control." In: *IEEE transactions on robotics and automation* 12.5, pp. 651–670 (cit. on p. 39).
- Joshi, Sameer A and Lung-Wen Tsai (2002). "Jacobian analysis of limited-DOF parallel manipulators." In: *ASME 2002 International design engineering technical conferences and computers and information in engineering conference*. American Society of Mechanical Engineers, pp. 341–348 (cit. on p. 67).
- Langley, Richard B (1998). "Rtk gps." In: *GPS World* 9.9, pp. 70–76 (cit. on p. 58).
- Latombe, Jean-Claude (2012). *Robot motion planning*. Vol. 124. Springer Science & Business Media (cit. on p. 22).
- LaValle, Steven M (1998). "Rapidly-exploring random trees: A new tool for path planning." In: (cit. on p. 25).
- Lopes, Carlos M, Albert Torres, Robert Guzman, João Graça, Miguel Reyes, Gonçalo Vitorino, Ricardo Braga, Ana Monteiro, and André Barriguinha (2017). "Using an unmanned ground vehicle to scout vineyards for non-intrusive estimation of canopy features

- and grape yield." In: *GiESCO International Meeting, 20th, Sustainable viticulture and wine making in climate change scenarios, 5-10 November 2017*. GiESCO (cit. on p. 33).
- Lozano-Perez, Tomas (1983). "Spatial planning: A configuration space approach." In: *IEEE transactions on computers* 2, pp. 108–120 (cit. on p. 22).
- Madhavan, Raj, Lino Marques, Edson Prestes, Renan Maffei, Vitor Jorge, Baptiste Gil, Sedat Dogru, Gonçalo Cabrita, Renata Neuhold, and Prithviraj Dasgupta (2015). *2015 humanitarian robotics and automation technology challenge* (cit. on p. 34).
- Marder-Eppstein, Eitan, Eric Berger, Tully Foote, Brian Gerkey, and Kurt Konolige (2010). "The Office Marathon: Robust Navigation in an Indoor Office Environment." In: *International Conference on Robotics and Automation* (cit. on p. 20).
- Mohanty, Vikram, Shubh Agrawal, Shaswat Datta, Arna Ghosh, Vishnu Dutt Sharma, and Debashish Chakravarty (2016). "DeepVO: a deep learning approach for monocular visual odometry." In: *arXiv preprint arXiv:1611.06069* (cit. on p. 18).
- Moore, T. and D. Stouch (2014). "A Generalized Extended Kalman Filter Implementation for the Robot Operating System." In: *Proceedings of the 13th International Conference on Intelligent Autonomous Systems (IAS-13)*. Springer (cit. on p. 18).
- Munoz-Salinas, Rafael (2012). "ARUCO: a minimal library for Augmented Reality applications based on OpenCv." In: *Universidad de Córdoba* (cit. on p. 78).
- Oberti, Roberto, Massimo Marchi, Paolo Tirelli, Aldo Calcante, Marcello Iriti, M Hocevar, and H Ulbrich (2014). "Crops agricultural robot: application to selective spraying of grapevine's diseases." In: *Proceedings of the Second RHEA International Conference on Robotics and associated High-technologies and Equipment for Agriculture* (cit. on p. 33).
- Ostafew, Chris J, Angela P Schoellig, and Timothy D Barfoot (2013). "Visual teach and repeat, repeat, repeat: Iterative learning control to improve mobile robot path tracking in challenging outdoor environments." In: *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*. IEEE, pp. 176–181 (cit. on p. 34).
- Remagnino, Paolo, Dorothy N Monekosso, and Lakhmi C Jain (2011). *Innovations in Defence Support Systems-3: Intelligent Paradigms in Security*. Vol. 336. Springer Science & Business Media (cit. on p. 17).
- Ringdahl, Ola, Polina Kurtser, Ruud Barth, and Yael Edan (2016). "Operational flow of an autonomous sweetpepper harvesting robot." In: *The 5th Israeli Conference on Robotics 2016, 13-14 April 2016. Air Force Conference Center Hertzilya, Israel* (cit. on p. 33).
- Rösmann, Christoph, Wendelin Feiten, Thomas Wösch, Frank Hoffmann, and Torsten Bertram (2012). "Trajectory modification considering dynamic constraints of autonomous robots." In: *Robotics;*

- Proceedings of ROBOTIK 2012; 7th German Conference on.* VDE, pp. 1–6 (cit. on p. 59).
- Santos, Pablo Gonzalez-de, Angela Ribeiro, and C Fernandez-Quitanilla (2012). “The RHEA Project: using a robot fleet for a highly effective crop protection.” In: *Proceedings of the International Conference of Agricultural Engineering (CIGR-Ageng’12), Valencia, Spain* (cit. on p. 33).
- Serrano, Daniel, Pietro Astolfi, Gianluca Bardaro, Alessandro Gabrielli, Luca Bascetta, and Matteo Matteucci (2017). “GRAPE: Ground Robot for vineyard Monitoring and Protection.” In: *ROBOT 2017: Third Iberian Robotics Conference*. Vol. 1. Springer, p. 249 (cit. on pp. 29, 53).
- Shamah, Benjamin (1999). “Experimental Comparison of Skid Steering vs. Explicit Steering for Wheeled Mobile Robot,” M. Sc.” In: (cit. on p. 15).
- Singhal, Amit (1997). “Issues in autonomous mobile robot navigation.” In: *Computer Science Dept, U. of Rochester*, p. 74 (cit. on p. 49).
- Thrun, Sebastian (2002). “Probabilistic robotics.” In: *Communications of the ACM* 45.3, pp. 52–57 (cit. on p. 18).
- Thrun, Sebastian, Dieter Fox, Wolfram Burgard, and Frank Dellaert (2001). “Robust Monte Carlo localization for mobile robots.” In: *Artificial intelligence* 128.1-2, pp. 99–141 (cit. on p. 53).
- Wang, Tianmiao, Yao Wu, Jianhong Liang, Chenhao Han, Jiao Chen, and Qiteng Zhao (2015). “Analysis and experimental kinematics of a skid-steering wheeled robot based on a laser scanner sensor.” In: *Sensors* 15.5, pp. 9681–9702 (cit. on pp. 16, 17).
- Yaguchi, Hiroaki, Kotaro Nagahama, Takaomi Hasegawa, and Masayuki Inaba (2016). “Development of an autonomous tomato harvesting robot with rotational plucking gripper.” In: *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*. IEEE, pp. 652–657 (cit. on p. 92).
- Yi, Jingang, Junjie Zhang, Dezhen Song, and Suhada Jayasuriya (2007). “IMU-based localization and slip estimation for skid-steered mobile robots.” In: *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*. IEEE, pp. 2845–2850 (cit. on p. 15).