Fully Autonomous Picking with a Dual-Arm Platform for Intralogistics

Alessandro Palleschi[+], Marco Gugliotta[+], Chiara Gabellieri[+], Dinh-Cuong Hoang[#], Todor Stoyanov[#], Manolo Garabini[+], Lucia Pallottino[+]

***Abstract*—Despite the strong demand for solutions and the intense research effort, fully autonomous object manipulation is not yet solved. One of the main challenges is the large variety of conditions in which the objects should be handled. There is a need for solutions that integrate both advanced perception systems for object recognition and pose estimation and robust planning algorithms. In this work, we present a fully autonomous dual-arm picking platform for intralogistics. The system is the result of the integration of a dual-arm manipulation system with a perception system for object registration, 3D instance-aware mapping, and 6D pose estimation of the objects. Eventually, we present the results of experiments on objects and conditions relevant to the intralogistics domain to assess the performance of the system.**

***Index Terms*—motion planning, dual-arm manipulation, intralogistics**

I. Introduction

A growing need from industry has been one of the primary driving factors of research in autonomous manipulation. Although several production processes have reached a high level of automation, intralogistics still requires manual operations, especially in picking tasks [1]. With companies worldwide investing in automation driven by economic and efficiency considerations, the recent spread of the COVID-19 pandemic is also expected to accelerate and facilitate the deployment of autonomous robotics solutions to reduce human contact and favor social distancing. However, the commercially available autonomous solutions for material handling work mostly for a limited variety of objects and specific configurations [2], and often based on custom grippers purposely designed to match the features of a specific object. These ad hoc solutions are not economically sustainable for small and medium-sized enterprises (SMEs) [3] that have to face different challenges: i) the large variety of objects [4], due to shorter product life cycles and dynamic production changeovers; ii) unstructured environments [2]; iii) the diversity at operating conditions.

This paper presents a manipulation system for logistics capable of performing picking operations in a completely autonomous way. The system is the result of the integration of the WRAPP-up robotic platform, a dual-arm solution for intralogistics presented in [5], with a cutting-edge perception system [6]. Experiments on different objects and conditions are presented to assess the performance of the proposed solution.

This work has received funding from the European Union's Horizon 2020 research and innovation program under agreement no. 732737 (ILIAD) and by the Italian Ministry of Education, and Research (MIUR) in the framework of the CrossLab project (Departments of Excellence)

[+]Centro di Ricerca E. Piaggio e Dipartimento di Ingegneria dell'Informazione, Università di Pisa, Pisa, Italia

[#]Centre for Applied Autonomous Sensor Systems (AASS), Orebro University, Sweden

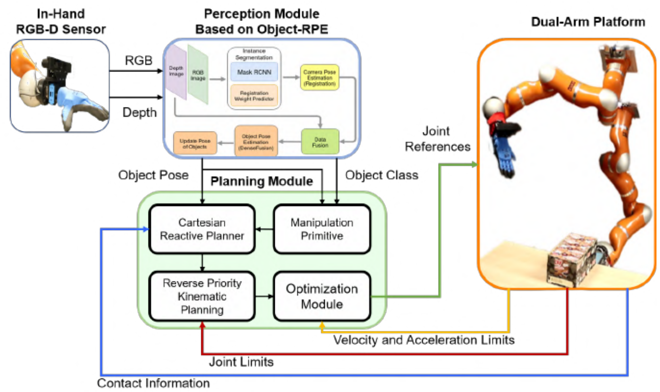


Fig. 1. Block scheme of the integrated system. The reactive planner algorithm, that take as input the information given by the perception system, is used to adapt and re-plan online the picking strategy based on contact information from force torque sensors at the end-effectors.

II. Manipulation for Intralogistics

Analyzing the different manipulation tasks that characterize the intralogistics scenario, we could classify them into three main categories, listed in order of increasing difficulty: A) Manipulation of a single object; B) Manipulation of multiple loosely-packed objects; C) Manipulation of multiple tightly-packed objects. Cases A) and B) are typical of picking objects from a conveyor belt, while Case C) is typical of picking operations from pallets or shelves.

In this work, we focus on addressing cases A)-B). The main challenge to solve case C) is posed by the inherent difficulty of the underlying perception problem. It requires to reliably distinguish between very close instances of the same object and accurately estimate their 6D pose from only a partial view. For a general object perception system, these conditions still pose an open problem.

III. The Dual-Arm Manipulation Platform

The solution here proposed to solve cases A)-B) is a dual-arm robotic platform, WRAPP-up, equipped with an advanced planning module and a perception system. WRAPP-up is an autonomous dual-arm manipulation system made of two KUKA LWR provided with an adaptive anthropomorphic hand, the Pisa/IIT SoftHand, and an end-effector based on an actuated belt, the Velvet Tray [5]. Force torque sensors ATI mini45 and ATI mini75 are fixed at the wrist of the hand and of the velvet, respectively. An Asus Xtion PRO[1] camera is fixed on top of the hand.

[1]https://www.asus.com/3D-Sensor/Xtion_PRO/

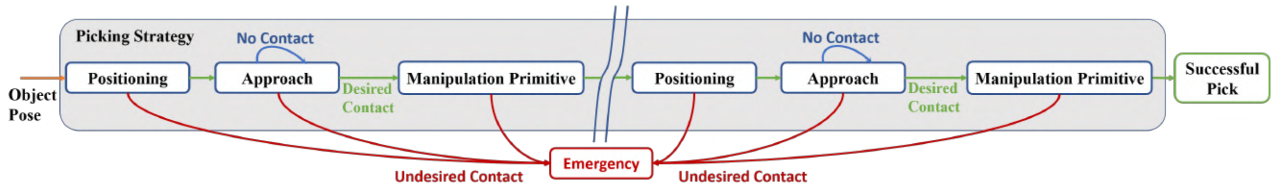


Fig. 2. General Scheme of the Reactive Planning Framework

Fig. 3. Objects used in the experiments.

The main components that constitute the intelligence of the integrated system are i) an object detection algorithm and ii) a reactive motion planner. A scheme explaining how the system works, and the connections between its building-blocks, is displayed in Fig. 1.

A. Object Detection Algorithm

We employ the Object-RPE framework [6] which couples Convolutional Neural Networks (CNNs) and a state-of-the-art dense Simultaneous Localization and Mapping (SLAM) system, ElasticFusion [7], to achieve high-quality semantic reconstruction as well as robust 6D pose estimation. The input RGB-D data is processed by a segmentation module and the reported object instance detections are filtered and matched to the existing 3D reconstruction. The predicted pose is used as a measurement update in a Kalman filter to estimate an optimal 6D pose of the object. Object-RPE exploits the availability of multiple observations of the scene acquired from different viewpoints. For more details on the algorithm, the interested reader can refer to [6].

B. Reactive Motion Planner

The second main block that constitutes the intelligence of the system is the reactive motion planner, the green block in Fig. 1. It generates a manipulation strategy to correctly pick the object. We identified a set of atomic actions, named manipulation primitives, based on considerations about the physical characteristics and the possible configurations of the objects to manipulate in an intralogistics context. The primitives, named **Sliding**, **Horizontal Rotation**, and **Vertical Rotation**, and presented in more details in [5], are expressed as a sequence of Cartesian waypoints for the end-effectors, relative to the pose of the object. As highlighted in Fig. 1, the planning module takes as inputs the classes of the (possibly multiple) detected objects, their poses, and dimensions. Then, it autonomously selects the picking strategy and generates a reference path for the end-effectors. Given the path, a minimum-time trajectory that takes into account constraints up to the jerk level is generated using the optimization algorithm presented in [8], adapted to run online.

To increase the robustness of the platform, we integrated force information at the planning stage. More in detail, the measured forces are used to detect a possible contact with an object whenever they exceed an user-defined threshold. The contact information is used as a trigger to plan online the motion of the robot in a reactive fashion. This reactive approach exploits the fact that we can decompose each picking strategy in a finite number of steps. The planner is then based on finite state machine, in which each state represents a step of the strategy. Indeed, we are not planning the complete trajectory all at once, but we plan it step-by-step and the decision on which step to plan is triggered online based on contact information and the current positions of the end-effectors.

Given the pose of the object, a picking strategy, as depicted in Fig. 2, is represented as a finite state machine. The states are defined by the *Manipulation Primitives* and two other classes of actions to be executed before actually perform a primitive. These two actions have been defined as *Positioning* (P), in which one, or both, the end-effectors are placed near the target object, and *Approach* (A), that represents the state in which the end-effectors approach the object and establish a contact before starting the manipulation primitive (M).

The pose of the target object is used to start the *Positioning* phase and plan a trajectory for the robot. Then, one, or both, the end-effectors start approaching, and contact information is used to trigger the transition. If no contact is detected, the robot keeps moving toward the object, whereas, if an unexpected contact is detected, it could trigger the transition to an *Emergency* state. Once the expected contact is detected, the planner starts the selected *Manipulation Primitive*.

The transitions from the *Manipulation Primitive* state to the other states, *Positioning* or *Emergency*, are triggered by contact information and/or by the end-effectors reaching a specified position, depending on the strategy.

IV. Experimental Validation

We tested the platform on a representative subset of the set of items described in [5] (see Fig. 3). Three of the selected items are cuboids, and two are cylinders with differing size, weight, and texture. Examples of the poses retrieved by the perception system for these objects are reported in 4.

We performed several tests to address manipulation cases A)-B). Given the shapes of the objects used for the tests, and considering the possible configuration that they can assume

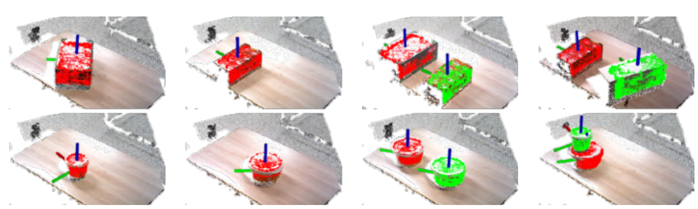
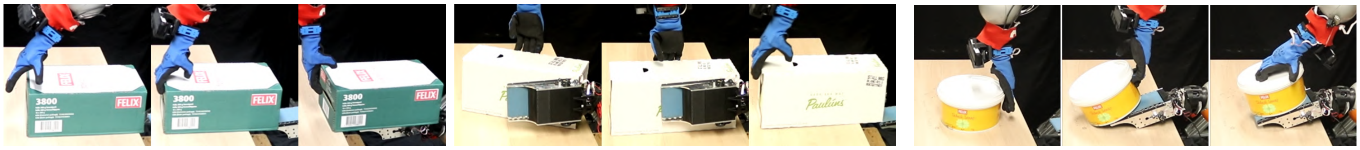


Fig. 4. The output of the perception algorithm is shown overlaid on the point cloud of the scene.



(a) Sliding - Average Execution Time: 12 s (b) Vertical Rotation + Sliding - Average Execution Time: 19 s (c) Horizontal Rotation + Sliding - Average Execution Time: 40 s

Fig. 5. The three picking strategies used to pick box-shaped and cylinder-shaped objects in various configurations.

TABLE I
PERFORMANCE OF THE MANIPULATION SYSTEM - SINGLE OBJECT

Object	# of picking actions	Picking rate
yellow bucket	10	100%
green bucket	10	80%
brown box	10	100%
green box	10	90%
white box	10	100%

Picking strategy	# of picking actions	Picking rate
Sliding	13	92.3%
Vertical Rotation + Sliding	17	100%
Horizontal Rotation + Sliding	20	90%

TABLE II
PERFORMANCE RESULTS FOR MULTIPLE OBJECTS

Object	# of picking actions	Picking rate
Loosely-packed boxes	20	95%
Loosely-packed cylinders	20	90%

in the workspace, three different picking strategies have been used. They are shown in Fig. 5, together with their average execution times. These strategies represent a particular case of those described in Section III-B. Two strategies are dedicated to the picking of box-shaped objects, namely a **Sliding** action for non rotated boxes, i.e., boxes aligned with the nominal picking direction, Fig. 5(a), and a combination of **Vertical Rotation** and **Sliding** to pick rotated boxes, Fig. 5(b). For the two cylinders, we defined a strategy based on the **Horizontal Rotation** primitive, depicted in Fig. 5(c). The hand is used to lift the cylinder and the Velvet Tray is placed below it. Then, a collaborative **Sliding** is performed to complete the picking.

Case A) We performed 10 picking actions per object. A picking action is considered successful if the object is carried to the unloading position without falling during the task. The success rate associated to each object and to each picking strategy is reported in Tab. I.

It is worth noting that for cuboid objects the combination of Vertical Rotation and Sliding corresponds to a 100% picking rate, while a failure is present when the simple Sliding strategy is used. More specifically, the failure is on the picking of the green box.

On the other hand, for the two cylinders we implemented only one picking strategy. We can see that this strategy is very effective for the bigger one, the yellow bucket, while for a smaller object turned out to be less performing and robust. Indeed, we noted how the reduced footprint and weight of the green bucket makes it more likely for the object to tip over while lifted by the hand, or during the sliding action.

Case B) To address manipulation case B), we performed 10 tests with two boxes placed in a loosely-packed configuration, for a total of 20 picking actions. More specifically, we picked 7 times the white box, 7 times the green box, and 6 times the brown box. Of the 20 picking actions, 11 where performed using a Sliding strategy, while 9 used the combination of Vertical Rotation and Sliding. Furthermore, we performed 10 tests with two buckets in a loosely packed configuration, picking 10 times each the yellow and the green bucket. The results of the tests are reported in Table II. As can be seen, the picking performance for the cylinders is not strongly affected by their relative configuration on the pallet, sticking around 90% in both configurations A) and B). The performance of the boxes, instead, slightly degrades when multiple objects are considered rather than single ones.

V. CONCLUSION

In this work, we presented a fully autonomous integrated system for picking operations. The potential of the platform emerged, due to its adaptability in handling objects with a wide range of sizes. The current integrated system is potentially able to grasp an even more diverse set of objects, exploiting recent results in the combination of perception systems and data-driven algorithms for grasping with robotic hands [9]. An important part of our future works will be dedicated to solving manipulation case C), focusing on the recognition and pose estimation of very close instances of the same goods.

REFERENCES

[1] G. H. Lim *et al.*, "Mobile manipulation for autonomous packaging in realistic environments: Euroc challenge 2, stage ii, showcase," in *2018 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, pp. 231–236, April 2018.

[2] W. Echelmeyer *et al.*, "Performance indicators for robotics systems in logistics applications," in *IROS Workshop on Metrics and Methodologies for Autonomous Robot Teams in Logistics (MMARTLOG)*, p. 55, 2011.

[3] G. Michalos, S. Makris, J. Spiliotopoulos, I. Misios, P. Tsarouchi, and G. Chryssolouris, "Robo-partner: Seamless human-robot cooperation for intelligent, flexible and safe operations in the assembly factories of the future," *Procedia CIRP*, vol. 23, pp. 71 – 76, 2014. 5th CATS 2014 - CIRP Conference on Assembly Technologies and Systems.

[4] G. Q. Huang, M. Z. Q. Chen, and J. Pan, "Robotics in ecommerce logistics," *HKIE Transactions*, vol. 22, no. 2, pp. 68–77, 2015.

[5] M. Garabini *et al.*, "Wrapp-up: A dual-arm robot for intralogistics," *IEEE Robotics & Automation Magazine*, 2020.

[6] D. Hoang, A. J. Lilienthal, and T. Stoyanov, "Panoptic 3d mapping and object pose estimation using adaptively weighted semantic information," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1962–1969, 2020.

[7] T. Whelan *et al.*, "Elasticfusion: Real-time dense slam and light source estimation," *The International Journal of Robotics Research*, vol. 35, no. 14, pp. 1697–1716, 2016.

[8] A. Palleschi, M. Garabini, D. Caporale, and L. Pallottino, "Time-optimal path tracking for jerk controlled robots," *IEEE Robotics and Automation Letters*, vol. 4, pp. 3932–3939, Oct 2019.

[9] C. Gabellieri *et al.*, "Grasp it like a pro: Grasp of unknown objects with robotic hands based on skilled human expertise," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2808–2815, 2020.