

Towards an Intelligent Collaborative Robotic System for Mixed Case Palletizing

Edoardo Lamon^{1,2}, Mattia Leonori¹, Wansoo Kim¹, and Arash Ajoudani¹

Abstract—In this paper, a novel human-robot collaborative framework for mixed case palletizing is presented. The framework addresses several challenges associated with the detection and localisation of boxes and pallets through visual perception algorithms, high-level optimisation of the collaborative effort through effective role-allocation principles, and maximisation of packing density. A graphical user interface (GUI) is additionally developed to ensure an intuitive allocation of roles and the optimal placement of the boxes on target pallets. The framework is evaluated in two conditions where humans operate with and without the support of a Mobile Collaborative robotic Assistant (MOCA). The results show that the optimised placement can improve up to the 20% with respect to a manual execution of the same task, and reveal the high potential of MOCA in increasing the performance of collaborative palletizing tasks.

I. INTRODUCTION

The e-commerce revolution, whose progress in Europe is estimated at 10% every year [1], has placed significant demands on the supply chain performances. In response to the increasing demand, several industrial sectors have automated the processes of goods storage and shipment. In particular, recent robotic palletizing solutions have permitted to increase considerably the processes' performances in terms of throughput (products per hour) and time (run products for longer periods) and contributed to perceiving the e-commerce benefits. In food and beverage industry, for instance, palletizing similar package types, using heavy robotic arms that can lift up to one thousand kilos, has become a benchmark for quality and high performance [2].

In high-mix environments such as warehouses, however, not all the processes are automated and human presence is still a requirement. Heavy manipulators cannot populate these environments, since they are not safe for humans and, to avoid potential human injuries, need to be enclosed in proper fences and to be equipped with reliable safety monitoring systems. Besides, in distribution centres, a wide variety of articles must be placed in boxes and then palletized, which demand for a certain level of manipulation flexibility. In such situations, redundant lightweight robotic arms, while dealing with different case sizes and weight, can ensure human safety due to the ability to regulate interaction forces both with humans [3] and with the environment [4]. More details on the application of cobots in industrial environments can be found in [5], while a review with recent developments

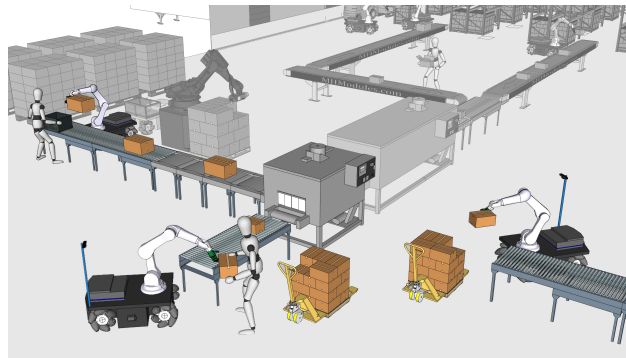


Fig. 1. Conceptual illustration of a warehouse environment where MOBILE Collaborative Assistants (MOCA) contribute to mixed case palletizing autonomously or in collaboration with human partners.

of robotized and autonomous warehouses is presented by *Azadeh et al.* [6]. Thanks to their flexibility, cobots can be exploited for different tasks and easily taught with different learning techniques. Nevertheless, they lack mobility, which is a fundamental requirement for warehouse operations.

A possible solution to address the mobility requirements is to attach the robot base to a mechanism that can roll/slide along rails. This solution restricts the motion of the base of the robot to be along the rails, which makes the construction of such warehouses space inefficient and costly. Other flexible solutions such as humanoid robots, whose kinematic structure resembles humans, offer a high potential in dealing with the variability of tasks in warehouses. However, their performances are still not satisfactory enough especially when it comes to execute logistics tasks that demand simultaneous locomotion and manipulation.

A very promising solution towards addressing the problem of mobility has been the introduction of *Autonomous Ground Vehicles (AGVs)*, i.e., wheeled mobile robots. AGVs can achieve repetitive and monotonous transportation of goods and increase process efficiency, while, at the same time, sharing their workspace with humans [7]. More advanced and recent forms of the AGVs have been developed to fulfil the manipulation requirements. Such platforms, due to robust loco-manipulation capabilities, have the potential to navigate in flat indoor environments using planning algorithms [8], avoid obstacles [9], pick goods from shelves, conveyors, etc., and place them in pallets and pushcarts, autonomously or in collaboration with humans [10].

In this direction, the goal of this work is to evaluate the potential of a MOBILE Collaborative robot Assistant (MOCA) in autonomous palletizing of mixed-size and -weight box containers (see Fig. 1). We propose a novel

This work was supported by the European Unions Horizon 2020 research and innovation programme under Grant Agreement No. 871237 (SOPHIA).

¹ HRI² Lab, Dept. of Advanced Robotics, Istituto Italiano di Tecnologia, Via Morego 30, 16163, Genova, Italy. ² Robotics and Automation group, Dept. of Information Engineering, Università degli Studi di Pisa, Pisa, Italy. edoardo.lamon@iit.it

loco-manipulation framework for logistic applications, which enables MOCA to navigate in free space and, using an embedded RGB-D camera, to scan all the boxes on a conveyor that must be palletized. After recording the sizes and the weights (through markers) of all boxes, an optimisation algorithm is activated to maximise the number of boxes that can be sorted on the detected pallet space. Subsequently, a role allocation algorithm is activated to designate the carrying and sorting actions to MOCA, in autonomous mode, or in collaboration with a human partner. The latter mode is activated when dealing with large-sized boxes, where single-arm manipulation poses a limit.

The control of MOCA loco-manipulation in different phases of the task is achieved by a weighted whole-body impedance controller, which is also capable of assigning larger movements to the arm or to the mobile base based on the task requirements/constraints. The algorithm gives higher mobility to arm movements in close-proximity reaching/sorting actions, and enables larger base mobility when navigating in free spaces.

Two different experiments were performed to evaluate the proposed framework. The first aims to prove that the guidance of the GUI, that displays the solution of the optimised mixed case palletizing, improves the performances, in terms of time and throughput. The second one, instead, is a proof-of-concept collaborative mixed case palletizing, executed by a human subject and MOCA as a coworker.

II. MODULAR FRAMEWORK FOR THE COLLABORATIVE MIXED CASE PALLETIZING

We aim to create a modular framework that enables and optimises the collaborative mixed case palletizing. These aspects range from the allocation of roles between the agents (humans and robots), up to exploring the environment, detecting relevant features from the perception sensors, and planning (collision-free) motions in space.

The framework consists of three main modules (see Fig 2). The modules will be listed by their operating frequency, from the slowest to the fastest. First, a task allocation algorithm, depending on the throughput of the system, computes, for each box, the destination of the box (pose in the pallet), and the agent in charge of the placement. The possible placement modes, considered in this work, are: (MODE A) MOCA places the box in the pallet; (MODE B) the human operator *can* place the box in the pallet, e.g., for light-weight boxes, up to his/her choice, to speed up the process, otherwise MODE A will take over; (MODE C) MOCA and the human operator place the box in the pallet, sharing collaboratively the load (e.g., for large boxes where dual-arm manipulation is a requirement. This cannot be achieved by MOCA, and might pose health risks to the human workers if handled individually). Another module consists in the visual perception system that, using the RGB-D data from the camera, estimates the poses of the boxes in the conveyor and the pallet in the area, defined in the same frame of MOCA. Finally, the motion handler is in charge of generating feasible paths and sending spatial references for the

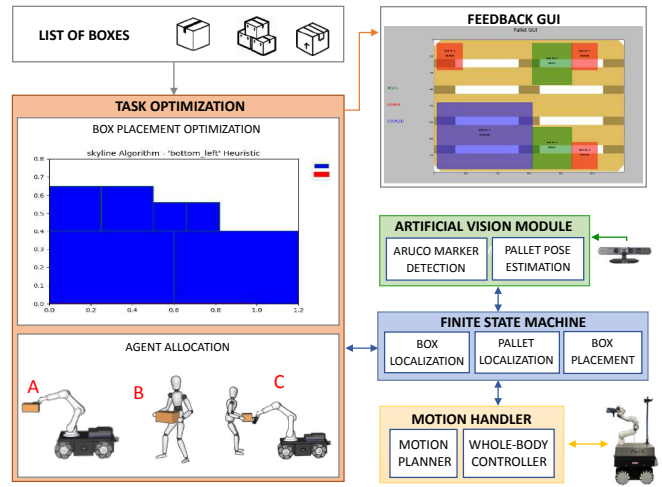


Fig. 2. Scheme of the proposed framework. The mixed case palletization optimisation consists of a bin packing algorithm, in charge of establishing the position of each box in the pallet, and a role allocation algorithm, that select the placement mode. A FSM, coordinates the robot actions, from the boxes and pallet detection, to the commands to pick and place the boxes, according to the results of the optimisation algorithm displayed on a GUI, that informs the human coworker of the current work status.

controller. The controller is a weighted whole-body Cartesian impedance controller, designed to regulate the behaviour of the mobile platform with respect to the end-effector task. The joint space weights of the controller are selected to obtain, alternately, higher flexibility of the arm in close-proximity reaching/sorting operations and larger mobile base mobility when navigating in free spaces. The controller settings, such as impedance parameters and weights for the priorities can be changed online. During each operation, the modules are always active but their services and actions are triggered by different Finite State Machine (FSM) requests.

A. MOCA: Platform and Control

In this section, the hardware and the prioritised whole-body impedance control algorithm of the MOBILE Collaborative Assistant is presented. MOCA [11] is a research platform, designed for human-robot physical collaboration (HRC), with loco-manipulation capabilities that makes it potentially suitable for logistic and flexible manufacturing. It is composed by a lightweight torque-controlled 7-DoFs Franka Emika Panda robotic arm, equipped with the under-actuated Pisa/IIT SoftHand, which is mounted on top of a velocity-controlled 3-DoFs Robotnik SUMMIT-XL STEEL mobile platform. An ASUS Xtion Pro Live RGB-D camera supported by a pole is also attached to the mobile base.

In Wu *et al.* in [11], we presented the first attempt to deal with the control framework of MOCA. The loco-manipulation capabilities were addressed using two different control modes, i.e., the whole-body manipulation mode, that features a whole-body Cartesian impedance controller for the manipulation tasks and the locomotion mode, that consists of a Cartesian impedance controller on the arm uncoupled with an admittance controller on the base. The latter mode was implemented to activate uniquely the base mobility while navigating between two distant points in the environment,

which otherwise would have generated unnecessary arm movements if the first whole-body mode was exploited.

In this paper, we exploit a similar whole-body control strategy for MOCA, since the targeted logistics tasks involve multiple contacts with humans and the external environment. However, the decoupling of the arm and the base movements in locomotion phases is not practical anymore, since the collaborative palletizing actions require simultaneous arm and mobile base movements and interactions. In addition, different phases of such logistics tasks require higher arm or mobile base mobility while maintaining contact at hand, to ensure their successful execution.

In order to fully exploit the redundancy provided by MOCA, we designed a two-level priority impedance controller by solving a weighted inverse dynamics problem. The whole-body decoupled dynamics of MOCA can be written as the parallel of the admittance controller of an m-DoFs mobile base and the dynamics of an n-DoFs torque-controlled arm [11]:

$$\begin{aligned} & \underbrace{\begin{pmatrix} M_{adm} & \mathbf{0} \\ \mathbf{0} & M_r \end{pmatrix}}_M \underbrace{\begin{pmatrix} \ddot{\mathbf{q}}_v \\ \ddot{\mathbf{q}}_r \end{pmatrix}}_{\ddot{\mathbf{q}}} + \underbrace{\begin{pmatrix} D_{adm} & \mathbf{0} \\ \mathbf{0} & C_r \end{pmatrix}}_C \underbrace{\begin{pmatrix} \dot{\mathbf{q}}_v \\ \dot{\mathbf{q}}_r \end{pmatrix}}_{\dot{\mathbf{q}}} + \underbrace{\begin{pmatrix} \mathbf{0} \\ \mathbf{g}_r \end{pmatrix}}_g \\ & = \underbrace{\begin{pmatrix} \tau_v^{vir} \\ \tau_r \end{pmatrix}}_{\boldsymbol{\tau}} + \underbrace{\begin{pmatrix} \tau_v^{ext} \\ \tau_r^{ext} \end{pmatrix}}_{\boldsymbol{\tau}^{ext}}. \end{aligned} \quad (1)$$

where $M_{adm} \in \mathbb{R}^{m \times m}$ and $D_{adm} \in \mathbb{R}^{m \times m}$ are the virtual inertial and virtual damping, $\dot{\mathbf{q}}_v \in \mathbb{R}^m$ is the input velocity sent to the mobile platform, $\tau_v^{ext} \in \mathbb{R}^m$ and $\tau_v^{vir} \in \mathbb{R}^m$ are the external and the virtual torque. Concerning the manipulator, $\mathbf{q}_r \in \mathbb{R}^n$ is the joint angles vector, $M_r \in \mathbb{R}^{n \times n}$ is the symmetric and positive definite inertial matrix of the arm, $C_r \in \mathbb{R}^n$ is the Coriolis and centrifugal force, $\mathbf{g}_r \in \mathbb{R}^n$ is the gravity vector, $\tau_r \in \mathbb{R}^n$, and $\tau_r^{ext} \in \mathbb{R}^n$ are the commanded torque vector and external torque vector, respectively.

The two-level priority impedance torque controller is inspired by the work by *Mingo Hoffman et al.* [12]. For the sake of simplicity, the dependencies on \mathbf{q} and \mathbf{x} will be dropped from now on. Let's consider the problem of finding the closest input torques $\boldsymbol{\tau}$ to some desired $\boldsymbol{\tau}_0$ that realises the operational forces \mathbf{F} :

$$\min_{\boldsymbol{\tau}} \frac{1}{2} \|\boldsymbol{\tau} - \boldsymbol{\tau}_0\|_{\mathbf{W}}^2 \quad \text{s.t. } \bar{\mathbf{J}}\boldsymbol{\tau} = \mathbf{F} \quad (2)$$

where $\mathbf{W} \in \mathbb{R}^{(n+m) \times (n+m)}$ is a positive definite weighting matrix, $\bar{\mathbf{J}} = (\mathbf{J}\mathbf{M}^{-1}\mathbf{J}^T)^{-1}\mathbf{J}\mathbf{M}^{-1}$ is the *dynamically consistent pseudo-inverse of $\mathbf{J}(\mathbf{q})$* , and the constraint $\bar{\mathbf{J}}\boldsymbol{\tau} = \mathbf{F}$, is the general relationship between the generalised joint torques and the operational forces [13]. The set of solutions of problem (2) can be found using the Lagrangian multipliers method. Differentiating the following Lagrangian function

$$\mathcal{L} = \frac{1}{2}(\boldsymbol{\tau} - \boldsymbol{\tau}_0)^T \mathbf{W}(\boldsymbol{\tau} - \boldsymbol{\tau}_0) + \lambda^T (\mathbf{J}\mathbf{M}^{-1}\mathbf{J}^T \mathbf{F} - \mathbf{J}\mathbf{M}^{-1}\boldsymbol{\tau}), \quad (3)$$

leads to

$$\boldsymbol{\tau} = \underbrace{\mathbf{W}^{-1}\mathbf{M}^{-1}\mathbf{J}^T\boldsymbol{\Lambda}_W\boldsymbol{\Lambda}^{-1}\mathbf{F}}_{\boldsymbol{\tau}_{task_1}} + \underbrace{(\mathbf{I} - \mathbf{W}^{-1}\mathbf{M}^{-1}\mathbf{J}^T\boldsymbol{\Lambda}_W\mathbf{J}\mathbf{M}^{-1})\boldsymbol{\tau}_0}_{\boldsymbol{\tau}_{task_2}}, \quad (4)$$

where

$$\boldsymbol{\Lambda} = \mathbf{J}^{-T}\mathbf{M}\mathbf{J}^{-1}, \quad (5)$$

and

$$\boldsymbol{\Lambda}_W = \mathbf{J}^{-T}\mathbf{M}\mathbf{W}\mathbf{M}\mathbf{J}^{-1} \quad (6)$$

is the *weighted Cartesian inertia*, analogous to the Cartesian inertia in Eq. (5). In Eq. (4) it is possible to recognise the two tasks in the controller, the Cartesian force task \mathbf{F} with higher priority and the joint-space torque $\boldsymbol{\tau}_0$ projected in the null space of the first task through a *dynamically consistent null-space projector*. The formulation in (4) contains the prioritised tasks, but the input torques of the controller need also to compensate for the other terms that are not present in (2), like gravity and Coriolis/centrifugal terms:

$$\boldsymbol{\tau} = \boldsymbol{\tau}_{task_1} + \boldsymbol{\tau}_{task_2} + \mathbf{g}(\mathbf{q}) + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} \quad (7)$$

In order to obtain the desired impedance behaviour, we need to define properly \mathbf{F}^{ext} and $\boldsymbol{\tau}_0$. \mathbf{F}^{ext} is computed according to the following relationship

$$\mathbf{F}^{ext} = \boldsymbol{\Lambda}_d \ddot{\tilde{\mathbf{x}}} + \mathbf{D}_d \dot{\tilde{\mathbf{x}}} + \mathbf{K}_d \tilde{\mathbf{x}} \quad (8)$$

where $\tilde{\mathbf{x}} = \mathbf{x}_d - \mathbf{x} \in \mathbb{R}^6$ is the Cartesian error computed with respect to the desired Cartesian pose \mathbf{x}_d , and $\mathbf{K}_d \in \mathbb{R}^{6 \times 6}$ and $\mathbf{D}_d \in \mathbb{R}^{6 \times 6}$ are the desired Cartesian inertia, damping and stiffness matrices, respectively. The desired joint-space impedance behaviour $\boldsymbol{\tau}_0 \in \mathbb{R}^n$ is computed according to [14],

$$\boldsymbol{\tau}_0 = -\mathbf{D}_0 \dot{\tilde{\mathbf{q}}} - \mathbf{K}_0(\tilde{\mathbf{q}} - \mathbf{q}_0), \quad (9)$$

where $\tilde{\mathbf{q}} = \mathbf{q}_0 - \mathbf{q} \in \mathbb{R}^n$ is the joint position error computed with respect to the reference joint position \mathbf{q}_0 , $\mathbf{K}_0 \in \mathbb{R}^{n \times n}$ and $\mathbf{D}_0 \in \mathbb{R}^{n \times n}$ are the desired joint-space stiffness and damping. We used a weight matrix \mathbf{W} of the form

$$\mathbf{W}(\mathbf{q}) = \mathbf{H}^T \mathbf{M}^{-1}(\mathbf{q})\mathbf{H} \quad (10)$$

where $\mathbf{H} \in \mathbb{R}^{n \times n}$ is the tunable positive semidefinite weight matrix of the controller. In particular, in this paper, we select \mathbf{H} as a positive definite diagonal matrix, dynamically selected depending on the task, of the form:

$$\mathbf{H} = \text{diag}\{\eta_{v_1}, \dots, \eta_{v_m}, \eta_{r_1}, \dots, \eta_{r_n}\}, \quad \eta_i > 0. \quad (11)$$

The diagonal elements are tuned to fit each particular task. For example, in manipulation tasks $\eta_{v_i} > \eta_{r_j}$, while, during locomotion, $\eta_{v_i} < \eta_{r_j}$.

Noteworthy, if $\mathbf{H} = \mathbf{I}$, $\mathbf{W} = \mathbf{M}^{-1}$, and the equation (4) can be simplified in the notable result [15]:

$$\boldsymbol{\tau} = \mathbf{J}^T \mathbf{F} + (\mathbf{I} - \mathbf{J}^T \boldsymbol{\Lambda} \mathbf{J} \mathbf{M}^{-1}) \boldsymbol{\tau}_0. \quad (12)$$

In other words, the control algorithm proposed by *Wu et al.* in [11] is the solution of problem (2) weighted by the inverse of the joint-space inertia matrix.

B. Vision Module

Despite warehouses can be considered structured environments, human presence increases the variability of the scenario. Moreover, collaborative robots, for safety reasons, must integrate robust perception systems. For instance, random box positions on conveyors and imprecise pallet location are problems that have to be considered in the framework. For this reason, a strong and reliable visual feedback is designed, to increase robot state awareness and accuracy in the pick and place task. In this section, the vision module is presented, in charge of the detection of boxes and pallets.

Box Detection: For the depicted task, it was necessary to detect the pose of several boxes on a conveyor with a high level of accuracy. These boxes have to be grasped precisely from the top, and placed on the pallet each one close to the other; the high precision in the detection of position and orientation affects the computation of the grasping point, allowing consistent placement of boxes on the pallet. Current literature proposes several techniques for general objects recognition in the scene: for example, in [16] a learning-based method has been implemented for pose estimation of various and heterogeneous objects, while in [17] the author takes advantage of a template matching approach, used for the same purpose. In this work *ArUco* markers detection were used, relying on the robustness and speed in the marker pose computation, as reported in [18] and [19]. Moreover, since the role allocation algorithm requires information about box weights and dimensions, such information can be easily integrated into a list, where each box is identified by its marker ID. This approach presents promising results since, in logistic scenarios, boxes are already equipped with markers or bar codes that specify weight, dimensions, address, etc. The *ArUco* markers detection has been performed through the integration in the module of the *aruco_detect* ROS package, using the images provided by the *Asus Xtion Pro Live* camera mounted on MOCA. Once the *ArUco* marker pose is estimated (Fig. 3(a)), the grasping location is computed through a rigid transformation based on prior information about boxes dimensions.

Pallet Detection: The second perception requirement consists of the detection and the pose estimation of pallets. In theory, the exploitation of the *ArUco* markers on the pallet could solve also this problem. In general, pallets are not equipped with such artificial tags and are subjected to wear over time. In any case, different features can be exploited for the detection. We implemented a detection algorithm based on the pallet geometrical dimensions. This algorithm processes the pointcloud acquired from *Orbbec Astra* depth camera integrated on the Robotnik platform. As reported in [20], in order to perform the pallet recognition, first, the pointcloud is pre-processed using a pass-through filter implemented in the used Point Cloud Library (PCL). It is possible to remove all the points which lie outside a certain region of interest, defined as $d \leq d_{threshold}$, where $d_{threshold}$ is a parameter which defines the dimension of the region from the ground. The detection is performed by looking at

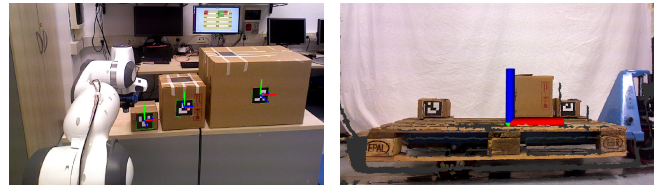


Fig. 3. (left) Box frames detected by the *ArUco* marker detector using the RGB images from the *ASUS* camera. (right) Coloured point-cloud acquired by the *Orbbec* camera. The frame has been placed in the centre of the top surface of the pallet, after detecting the three wooden blocks.

the long side of the pallet, based on its particular structure: as a matter of fact, the pallet is characterised by the presence of three equidistant wooden blocks on its side which are very specific. The algorithm uses a region growing segmentation, which merges the points that belong to the same smooth surface. This approach returns several clusters; to extract the ones corresponding to the wooden blocks, the procedure searches for three clusters whose centroids lie on the same line, as well as at a certain distance one from the other. Once these clusters are obtained, the pallet is detected. Then, a frame is placed on the centre of the top surface: the pose can be obtained through a rigid transformation from the central wooden block, based on pallet dimensions (Fig. 3(b)).

C. Task Optimisation

In this section, we will discuss how the framework intends to improve the execution of mixed case palletizing tasks in terms of time and density of packing and to decrease heavy lifting requirements for human workers. The mixed case palletizing problem is NP-complete and the stacking rules, based on heuristics, must be adapted to each customer accordingly, taking into account all the system constraints. Optimising in this way the packing of goods has several benefits: the packing density is maximised (90% with respect to the average 70%-80% of the manually packed pallets), while the number of pallets needed for the same volume of units is reduced. Usually, the solution is pre-computed offline and consists of the packing unit pick sequence and the location in the pallet, where large, heavy products are placed on the bottom of the pallet while small, lightweight products on the top. During the package placement design process, a small gap between the packages is added, as safety tolerance, to avoid potential collisions. In literature, the mixed case palletizing problem is formulated as the well-known *Bin Packing problem* [21]–[23], which goal is to orthogonally pack a set of rectangular-shaped items into the minimum number of identical containers (bins). Each item can be rotated by 90 degrees on each axis. The theoretical formulation of the 3D bin packing problem and the application on the pallet loading problem, including also technological constraints and stability of the load on the pallet, is presented by *Terno et al.* in [22]. Since most 3D packing heuristics are built upon 2D ones, in this paper, we focus on solving the 2D problem. To evaluate the packing in different scenarios, we enabled our framework to run three different Bin Packing algorithms: the *Guillotine algorithm*, the *Maximal Rectangles algorithm* and the *Skyline algorithm*.

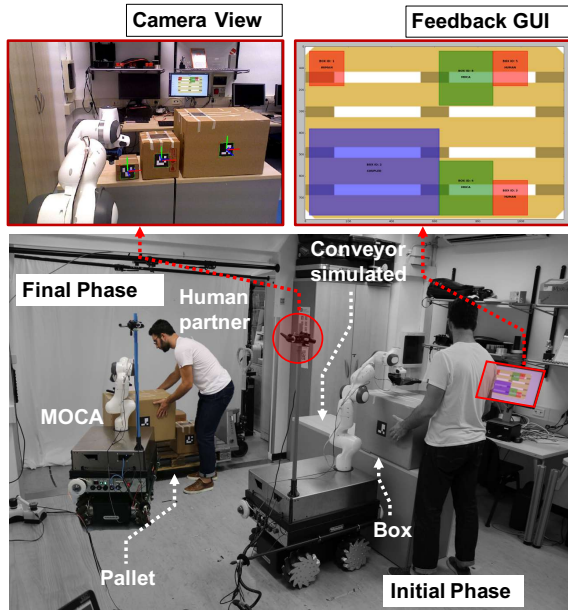


Fig. 4. Illustration of the experimental setup for the MOCA-human partner collaborative palletizing task. Camera view (top left) obtained the orientation of the boxes on the conveyor, and the feedback GUI (top right) provided the visual feedback to the human partner about the optimised palletizing configuration.

These algorithms differ mainly on the left space splitting strategy. A complete description of the implementation and the computational complexity of the algorithms can be found in [24]. Once the Bin Packing algorithm returns the desired box placement, a suitable policy for the role allocation problem should be designed. Introducing the human-robot collaboration in such mixed case palletizing tasks, as a matter of fact, opens to the problem of designating systematically a suitable agent to perform each box placement in the pallet. To design the role allocation algorithm, we assign each task according to each agent capability. This concept was introduced by the authors in [25]: exploitation of robots for repetitive and hard tasks and human employment for monitoring and solving complex tasks. Among the three metrics proposed, namely task complexity, agent dexterity and agent effort, to allocate the pick and place tasks depicted in this work, just the last one will be addressed. The limitations for the role allocation are imposed by the size and weight of the boxes. We assume that the boxes that we are dealing with weight less than 25 kg (the value is imposed by law as the maximum weight that can be lifted by a male worker). On the other hand, the maximum load that MOCA can carry depends on the payload of the mounted robotic arm. Thus, the algorithm will negotiate three different modes (MODE A, B, and C, as explained above) for achieving the task. The pseudocode can be summarised as: $\forall \text{box} \in \text{list_of_boxes}$, if $\text{box.weight} > \text{ROBOT_PAYLOAD}$ or $\text{box.volume} > \text{HIGH_VOLUME}$, then $\text{box.agent} \leftarrow \text{MODE_C}$, else if $\text{box.weight} < \text{LOW_WEIGHT}$ or $\text{box.volume} < \text{LOW_VOLUME}$, then $\text{box.agent} \leftarrow \text{MODE_B}$, else $\text{box.agent} \leftarrow \text{MODE_A}$. In order to achieve the human-robot collaboration, the human worker is expected to monitor

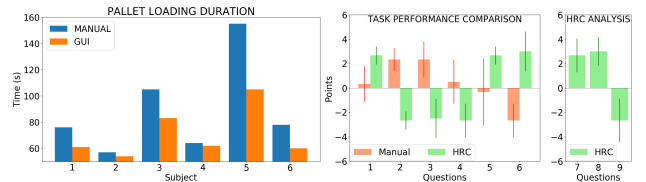


Fig. 5. (left) Duration of the manual (blue) and instructed (orange) palletizing in the first experiment. (right) Likert scale-based questionnaire scores for the manual (red) and collaborative instructed (green) palletizing related to both experiments.

periodically the GUI and to acknowledge the execution of the task. To do that and also achieve a greater level of cooperation between the agents, minimising waiting times, mobile devices or AR headsets [25], [26] could be exploited.

III. EXPERIMENTAL RESULTS

The proposed framework was validated with two different experiments, executed by 6 different naive subjects, not aware of the scope of the experiments. In the first experiment, we asked the subjects to place 11 numbered boxes, of different size and weight, on the surface of a pallet, in such a way all the boxes fits in the lower layer of the pallet. Later, with the same settings, we computed the box placement through the Skyline algorithm and we asked the subjects to place them according to the algorithm result, displayed in the GUI. We compared the results in terms of time spent in completing the palletizing. The results (Fig. 5 (left)) show that, using such algorithm to help human workers improved their performance, saving up to the 20% of the execution time.

The second experiment consists of a proof-of-concept collaborative mixed case palletizing. In this experiment, the subject and MOCA have to collaborate to achieve the pallet loading. Three boxes, marked with an ArUco marker, were placed on top of a simulated conveyor. The pallet was placed close to the conveyor (Fig. 4). The placement is computed, as in the previous experiment, by the Skyline algorithm, using as input the total number of boxes that are supposed to be placed in that pallet. Moreover, the agent in charge of the placement is computed by the role allocation algorithm, that will assign one of the three modes (MODE A, MODE B, MODE C). The robot has just a rough idea of the pallet and conveyor location, and, because of considerable errors in the odometry, has to re-detect the pose of the pallet before each new placement. Also, the boxes have to be detected every new placement, since, in general, the number of boxes in the conveyor might change, due to newly delivered boxes. Just the detected boxes will appear in the GUI, with a different colour depending on the mode (red for MODE A, green for MODE B, and blue for MODE C). To enable MOCA to grasp boxes we planned to use a vacuum gripper, but due to delay in the shipment, we had to opt for a different momentary solution. We placed an electromagnet in the wrist of MOCA and a sheet made of ferromagnetic material on the upper surface of each box. In this way, we could emulate the functioning of the vacuum gripper, that was activated in the box picking phase and

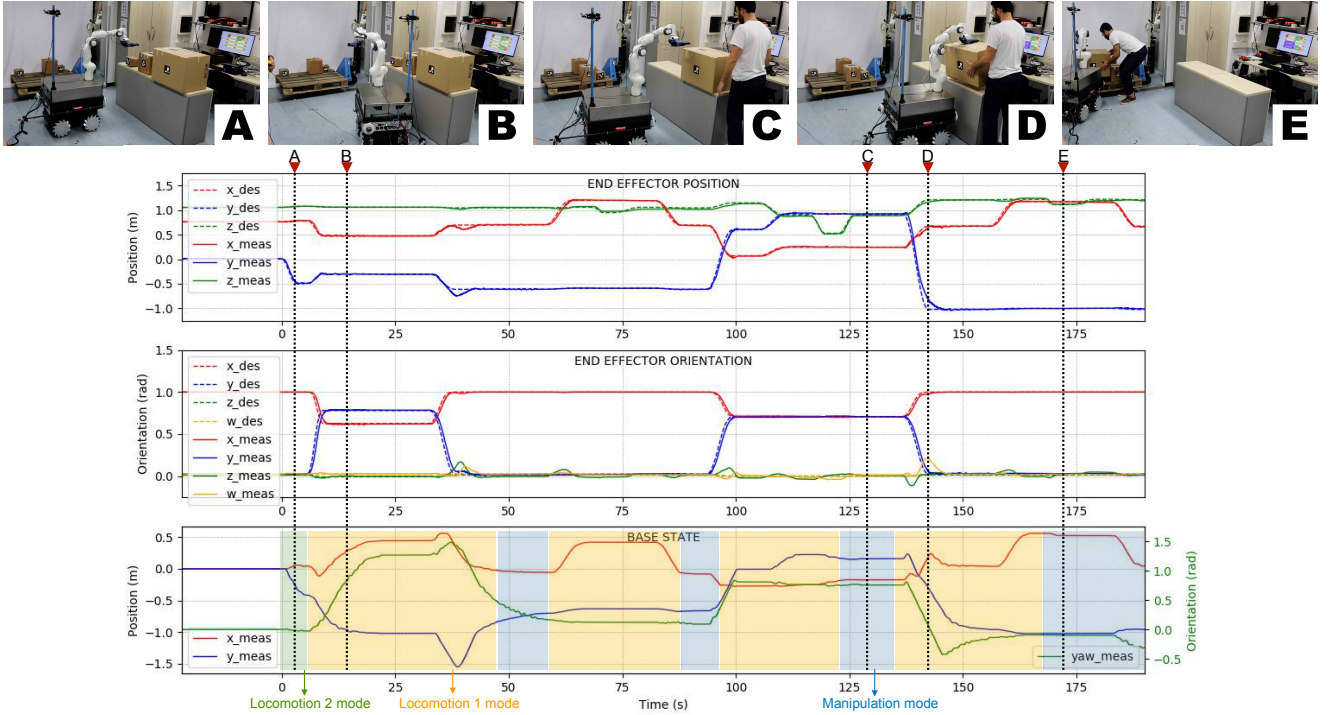


Fig. 6. Snapshots of the second experiment: to detect the boxes in the conveyor, the robot first scans the ArUco markers with locomotion 2 mode (A), then rotates with locomotion 1 mode to estimate the pallet pose (B). Next, it approaches and grasps each box in manipulation mode (C), then moves towards the pallet in locomotion 1 mode (D), and finally places the box in manipulation mode (E). Lower plots of the full experiment demonstrate the prioritised mobility of the arm and the mobile base in these modes.

deactivated in the box placement phase. Finally, the control parameters were designed. We selected a compliant behaviour, keeping a low valued diagonal of Cartesian stiffness ($K_x = K_y = K_z = 300$, $K_{roll} = K_{pitch} = K_{yaw} = 30$), except in the picking and in the placing phase, in which, to ensure higher precision, we increased the stiffness in the x - y axis up to 1000. The Cartesian damping was computed using *Double Diagonalisation design* [14]. The weight matrix \mathbf{H} of the controller was tailored to define three main control modes: the *manipulation mode*, where $\mathbf{H} = \text{diag}\{\sqrt{5}, \sqrt{5}, \sqrt{5}, 1, 1, 1, 1, 1, 1, 1\}$, the *locomotion 1 mode*, a locomotion that exploits all the DoFs provided by the base, where $\mathbf{H} = \text{diag}\{1, 1, 1, \sqrt{5}, \sqrt{5}, \sqrt{5}, \sqrt{5}, \sqrt{5}, \sqrt{5}, \sqrt{5}\}$, and the *locomotion 2 mode*, a locomotion that tries to avoid unnecessary rotations of the base about z axis, where $\mathbf{H} = \text{diag}\{1, 1, \sqrt{20}, \sqrt{5}, \sqrt{5}, \sqrt{5}, \sqrt{5}, \sqrt{5}, \sqrt{5}, \sqrt{5}\}$. The snapshots from the footage of the experiment (Fig. 6) depict the main phases of the experiment. Finally, we asked the subjects to compile a Likert scale-based questionnaire, to have a subjective evaluation of the experiments, approved by the ethics committee Azienda Sanitaria Locale Genovese (ASL) N.3 (Protocollo IIT_HRII_001 (rif. interno:108/2018)). The questionnaire included 9 sentences. Q.1 The palletizing task was easy to perform; Q.2 It was physically tiresome to accomplish the palletizing; Q.3 It was psychologically tiresome to accomplish the palletizing; Q.4 The cognitive load to achieve the task was high; Q.5 Overall, I felt satisfied with the current task performance; Q.6 It was easy to understand where to place each box in the pallet; Q.7

I felt safe in performing the palletizing with the robot; Q.8 Overall, I think that using the current collaborative framework I could perform the same task for a longer duration and better quality time; Q.9 Given the current task performance, I think that collaborative robots do not help to improve such logistic tasks. The participants stated that, even though the complexity of the task results still acceptable (Q.1), the manual performance requires greater physical and psychological effort (Q.2-3), with high cognitive load (Q.4) with respect to the smart collaborative approach. On the other hand, the proposed framework, presents promising results in terms of performance and work quality (Q.8-9), leading to broader satisfaction with the proposed collaborative system (Q.5) (Fig. 5 (right)).

IV. CONCLUSION

In this work, we presented a novel human-robot collaborative approach to the mixed case palletizing problem. Several problems were addressed, from the box and pallet detection and localisation achieved by the visual perception module, to the task and role allocation algorithm, that computes the optimal location in the pallet and agent in charge of the placement, passing through the design of a Cartesian impedance controller capable of changing the joint space behaviour of the robot. The results demonstrated that such a framework, together with MOCA, have a high potential in improving human workers productivity and ergonomics, and represent a promising first step towards an intelligent collaborative robotic system for mixed case palletizing.

REFERENCES

- [1] S. Lone, A. Khelladi, and S. Packiarajah, "European ecommerce report: 2018 edition," *Brussels: Ecommerce Europe Publications*, 2018.
- [2] C. Wurrll, "Mixed case palletizing with industrial robots," in *Proceedings of ISR 2016: 47th International Symposium on Robotics*, June 2016, pp. 1–6.
- [3] S. Haddadin, S. Haddadin, A. Khoury, T. Rokahr, S. Parusel, R. Burgkart, A. Bicchi, and A. Albu-Schffer, "On making robots understand safety: Embedding injury knowledge into control," *The International Journal of Robotics Research*, vol. 31, no. 13, pp. 1578–1602, 2012.
- [4] P. Balatti, D. Kanoulas, N. G. Tsagarakis, and A. Ajoudani, "Towards robot interaction autonomy: Explore, identify, and interact," in *2019 International Conference on Robotics and Automation (ICRA)*, May 2019, pp. 9523–9529.
- [5] A. Ajoudani, A. M. Zanchettin, S. Ivaldi, A. Albu-Schäffer, K. Kotsuge, and O. Khatib, "Progress and prospects of the human-robot collaboration," *Autonomous Robots*, vol. 42, pp. 957–975, 2018.
- [6] K. Azadeh, R. De Koster, and D. Roy, "Robotized and automated warehouse systems: Review and recent developments," *Transportation Science*, vol. 53, no. 4, pp. 917–945, 2019.
- [7] P. R. Wurman, R. D'Andrea, and M. Mountz, "Coordinating hundreds of cooperative, autonomous vehicles in warehouses," in *Proceedings of the 19th National Conference on Innovative Applications of Artificial Intelligence - Volume 2*, ser. IAAI'07. AAAI Press, 2007, pp. 1752–1759.
- [8] D. Berenson, J. Kuffner, and H. Choset, "An optimization approach to planning for mobile manipulation," in *2008 IEEE International Conference on Robotics and Automation*, May 2008, pp. 1187–1192.
- [9] O. Brock and O. Khatib, "Elastic strips: A framework for motion generation in human environments," *The International Journal of Robotics Research*, vol. 21, no. 12, pp. 1031–1052, 2002.
- [10] O. Khatib, K. Yokoi, K. Chang, D. Ruspini, R. Holmberg, and A. Casal, "Vehicle/arm coordination and multiple mobile manipulator decentralized cooperation," in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS '96*, vol. 2, Nov 1996, pp. 546–553 vol.2.
- [11] Y. Wu, P. Balatti, M. Lorenzini, F. Zhao, W. Kim, and A. Ajoudani, "A teleoperation interface for loco-manipulation control of mobile collaborative robotic assistant," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3593–3600, Oct 2019.
- [12] E. M. Hoffman, A. Laurenzi, L. Muratore, N. G. Tsagarakis, and D. G. Caldwell, "Multi-priority cartesian impedance control based on quadratic programming optimization," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, May 2018, pp. 309–315.
- [13] O. Khatib, "Inertial properties in robotic manipulation: An object-level framework," *The International Journal of Robotics Research*, vol. 14, no. 1, pp. 19–36, 1995.
- [14] C. Ott, *Cartesian Impedance Control of Redundant and Flexible-Joint Robots*. Springer, 01 2008, vol. 49.
- [15] O. Khatib, "A unified approach for motion and force control of robot manipulators: The operational space formulation," *IEEE Journal on Robotics and Automation*, vol. 3, no. 1, pp. 43–53, February 1987.
- [16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *NIPS 2012*, 2012.
- [17] S. Hinterstoisser, S. Holzer, C. Cagniart, S. Ilic, K. Konolige, N. Navab, and V. Lepetit, "Multimodal templates for real-time detection of texture-less objects in heavily cluttered scenes," in *2011 International Conference on Computer Vision*, Nov 2011, pp. 858–865.
- [18] S. Garrido-Jurado, R. Muñoz-Salinas, F. Madrid-Cuevas, and R. Medina-Carnicer, "Generation of fiducial marker dictionaries using mixed integer linear programming," *Pattern Recognition*, vol. 51, pp. 10 2015.
- [19] F. Romero Ramirez, R. Muñoz-Salinas, and R. Medina-Carnicer, "Speeded up detection of squared fiducial markers," *Image and Vision Computing*, vol. 76, pp. 06 2018.
- [20] B. Molter and J. Fottner, "Real-time pallet localization with 3d camera technology for forklifts in logistic environments," in *2018 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI)*, July 2018, pp. 297–302.
- [21] S. Martello, D. Pisinger, and D. Vigo, "The three-dimensional bin packing problem," *Operations Research*, vol. 48, no. 2, pp. 256–267, 2000.
- [22] J. Terno, G. Scheithauer, U. Sommerwei, and J. Riehme, "An efficient approach for the multi-pallet loading problem," *European Journal of Operational Research*, vol. 123, no. 2, pp. 372–381, 2000.
- [23] G. S. Taylor, Y. Chan, and G. Rasool, "A three-dimensional bin-packing model: exact multicriteria solution and computational complexity," *Annals of Operations Research*, vol. 251, no. 1, pp. 397–427, Apr 2017.
- [24] J. Jylänki, "A thousand ways to pack the bin – a practical approach to two-dimensional rectangle bin packing," 2010.
- [25] E. Lamon, A. De Franco, L. Peternel, and A. Ajoudani, "A capability-aware role allocation approach to industrial assembly tasks," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3378–3385, Oct 2019.
- [26] A. De Franco, E. Lamon, P. Balatti, E. De Momi, and A. Ajoudani, "An intuitive augmented reality interface for task scheduling, monitoring, and work performance improvement in human-robot collaboration," in *2019 IEEE International Work Conference on Bioinspired Intelligence (IWOBI)*, 2019.