Integration of a Multi-Camera Vision System and Admittance Control for Robotic Industrial Depalletizing

Davide Chiaravalli, Gianluca Palli
DEI - Department of Electrical, Electronic and Information Engineering, University of Bologna
Bologna, Italy
{davide.chiaravalli2, gianluca.palli}@unibo.it

Riccardo Monica
CIDEA
University of Parma
Parma, Italy
riccardo.monica@unipr.it

Jacopo Aleotti, Dario Lodi Rizzini
RIMLab - Department of Engineering and Architecture, University of Parma
Parma, Italy
{jacopo.aleotti, dario.lodirizzini}@unipr.it

Abstract—This work addresses the task of robot depalletizing by means of a mobile manipulator, taking into account the problem of localizing the boxes to be removed from the pallet and a manipulation strategy that allows to pull the boxes without lifting them with the robot arm. The depalletizing task is of particular interest in the industrial scenario in order to increase efficiency, flexibility and economic affordability of automatic warehouses.

The proposed solution makes use of a multi-sensor vision system and a force-controlled collaborative robot in order to detect the boxes on the pallet and to control the robot interaction with the boxes to be removed. The vision system comprises a fixed 3D Time-of-flight camera and an eye-in-hand 2D camera. Preliminary experimental results performed on a laboratory setup with a fixed-based robotic manipulator are reported to show the effectiveness of the perception and control system.

Index Terms—Robotic Manipulation, Robot Vision, Industrial Depalletizing.

I. INTRODUCTION

Modern industrial automated warehouses are designed to optimize transportation and distribution of loads such as cardboard boxes [1]. The problem addressed in this paper is related to this relevant industrial field, and in particular to automated depalletizing, that is the process of unloading a pallet, namely the origin pallet, containing homogeneous goods, usually in forms of a matrix of cardboard boxes arranged on multiple stacked layers, with the purpose of composing a new pallet, namely the mixed pallet, collecting the items coming from different origin pallets.

Robot depalletizers are usually limited to 4DoFs and feature large end effectors able to move an entire layer of boxes from the origin pallet to a distribution and serialization system composed of multiple conveyor belts. It results that the robot should have a large payload even in case of light boxes just to move the end effector. Moreover, considering that up to ten thousand different articles can be stored in a food and beverage warehouse, it results that the plant footprint should be extremely large due to the need of intermediate buffers to temporarily store the depalletized items before moving them to the mixed pallets.

Vacuum grippers are often used for depalletizing cardboard boxes [2]–[4]. However, cardboard boxes can not always hold the weight of items they contain, therefore this solution is not generally safe. Nakamoto et al. [4] designed a depalletizing system where the package is supported thanks to a second robot arm. In [2] Doliotis et al. developed a vision system for a depalletizing task of sacks using a vacuum gripper. In [3] a robot manipulation system was presented, based on 3D vision and a vacuum gripper, for detection and unloading of cardboard boxes from shipping containers or semi-truck trailers.

Similarly to what is proposed in this paper, Hashimoto et al. [5] presented a genetic algorithm to recognize loads on a pallet. The main differences are that the method in [5] adopted a fixed gray scale camera placed above the pallet, while we adopt a multi-camera system and admittance control to achieve simultaneously an accurate robot positioning and...
a controlled interaction with the boxes. Yunardi et al. [6] investigated a method to determine the size of parcels moving on a conveyor belt using RGB cameras. Prasse et al. [7]–[9] proposed a system to detect the pose of parcels located on a pallet by combining a Time of flight (ToF) sensor and RFIDs, which were used to compute the 3D structure of the layer. Katsoulas et al. [10]–[12] proposed methods for the recognition of arbitrary size boxes in cluttered environments using a planar laser scanner, mounted on a robot arm in eye-in-hand configuration. A drawback of [10]–[12] is that being based on 2.5D edge detectors they could fail to detect aligned boxes in contact with one another. In [13] a robotic depalletizing system was proposed using uncalibrated vision and 3D laser assisted image analysis. All sensors were attached to the ceiling.

In this paper, we show the results of preliminary experiments performed on a simplified laboratory setup, shown in Fig. 1 and described in details in Sec. II. This setup is aimed at testing a new depalletizing system for P-COORSA, in Fig. 1 and described in details in Sec. II. This setup is aimed at developing a robotic platform able to perform the in-situ depalletizing of a single box one at a time, to increase the flexibility, scalability and affordability of automated warehouses. A box is removed from the pallet by inserting a blade into the gap between adjacent parcel boxes, and then by dragging the desired box toward the robot. As in industrial scenarios the boxes are often tightly packed, the gap between the boxes may be very thin and it must be detected with high accuracy for successful blade insertion avoiding any damage to the box itself. For this reason, the proposed system exploits integration of a multi-camera vision system and admittance control. The robot is, indeed, equipped with a second sensor, a 2D RGB in-hand camera, used to refine the detection of the gap between boxes in order to insert the tool. If the tool insertion fails because of incorrect gap detection, the robot is provided with a tool wrench force controller that stops the robot to prevent from damaging the boxes, and the gap detection procedure is repeated again.

The paper is organized as follows: Sec. II reports a description of the experimental setup and how the box detection and manipulation are performed in details, while Sec. III shows the results obtained during experimental sessions. Finally, Sec. IV draws some conclusions and future developments on this work.

II. TASK AND SYSTEM OVERVIEW

The depalletizing task that is considered in this work requires two main problems to be addressed: localization of parcel boxes on the top layer of a pallet as well as the gaps between them, and the execution of the depalletizing operation, which is achieved by inserting a blade into the gap between parcel boxes and by pulling the desired box toward the robot. To solve these problems a multi-camera vision system was adopted as well as a force-controlled collaborative robot. The multi-camera vision system, as shown in Fig. 1, consists of a fixed Time-of-Flight (ToF) sensor mounted above the arm that provides an initial estimated pose of the boxes over the whole scene, in order to locate the boxes and the pallet plane in space, and an RGB camera mounted on the robot end effector used to refine the detection of the gaps among the parcel boxes. To observe the whole pallet at once, the ToF camera (an IFM Electronics O3D303) was placed at a height of about 1.10 m from the pallet plane. Given the nominal focal length of about 305 pixels of the ToF camera, one pixel error leads to an error of approximately 3.6 mm. Hence, narrow gaps cannot be detected with sufficient precision using the ToF camera alone. Conversely, the use of a multi-camera vision system assures a higher precision and accuracy. The information provided by the vision system also enables to efficiently plan the box extraction sequence from the pallet.

The motion of the in-hand camera and the execution of the box extraction operation is handled by a cooperative 6-DoFs robotic manipulator (UR5 from Universal Robots) equipped with a blade mounted as end effector that the robot uses to drag the boxes. The choice of a collaborative robot is motivated by the need of reducing the weight of the manipulator in order to mount it on a mobile base in the final system and by the increasing demand of sharing the workspaces with human workers in modern facilities. Since the end effector must necessarily be thin to be inserted in between the boxes, a suitable cover can be designed to make the end effector safe while the robot is moving in free space. Moreover, the tool wrench provided by the collaborative robot, which can be computed from the measured currents or from torque sensors at joint level, enables to control the interaction between the robot and the objects to be manipulated. This characteristic is exploited in this work to detect if the robot tool is correctly aligned with the gap between the boxes, and if blade insertion is properly executed. In case of misalignment, the robot will detect a sudden increase of the tool wrench in the vertical direction, that in turn triggers the request of a new evaluation of the gap position and orientation by the robot to correct the problem, as detailed in the following.

The algorithm governing the depalletizing task considered in this paper is reported in Alg. 1. In the parcel box detection phase, an initial coarse detection of all the poses is performed by the fixed ToF camera, and the box depalletizing plan exploits this information to define a sequence of picking operations according to the relative position of the robot w.r.t. boxes. Then, in the refine gap estimation phase the 2D camera
Algorithm 1: Robotized Depalletizing Algorithm

Data: ToF camera image, 2D camera image, robot tool wrench

Result: depalletizing of a box layer

perform parcel box detection;
define box depalletizing plan;
while depalletizing plan is not empty do
    consider the first box in the depalletizing plan;
    move the 2D camera to the estimated gap location;
    refine gap estimation using the 2D camera image;
    insert the robot tool in the gap;
    if vertical tool wrench < wrench threshold then
        drag the box toward the robot;
    else
        remove current box from the plan;
    end
end

mounted on the robot tool is used to refine the estimated location of the gap that separates the selected box to be extracted from the opposite box w.r.t. the desired tool insertion point. Once the line that defines the gap is found, the insert the robot tool operation is performed and the tool wrench (via calculation or via direct measurement) is monitored to detect if the tool is correctly entering in the gap, or if it is in contact with the box upper surfaces. If the tool wrench remains under a suitable limit value, the drag the box toward the robot operation is performed and the current box is removed from the plan, otherwise the tool is extracted and gap detection is repeated, using the position of the gap estimated at the previous iteration as a reference to reposition the 2D camera.

In the following, a detailed explanation of the main steps of the algorithm is presented.

A. Parcel Box Detection

The detection of parcel boxes relies on the industrial ToF camera IFM Electronics O3D303, which returns a point cloud and an intensity image, with resolution 352 × 264 pixels and a focal length of about 305 pixels. The point cloud is organized, i.e. the 3D points \( P = \{ p_{ij} \} \) are assigned to pixel coordinates \((i, j)\). Likewise the intensity image \( R \) consists of raw intensity value \( r_{ij} \) for each pixel (see Figure 2(b)). The camera is mounted on a pole and tilted in order to observe the layer of parcel boxes from a high viewpoint at about 1.10 m from the layer (top right in Figure 1). We assume that the size of the boxes is known.

We also assume that the plane \( \Omega \) which includes the top face of the highest layer of boxes is known with respect to the camera reference frame. If the plane parameters are not available, they can be estimated by repeatedly applying RANSAC to the range image and by manually selecting the target plane among all those found in the scene. A pixel \((i, j)\) is a plane inlier if point \( p_{ij} \) is within a tolerance \( t_{\text{inlier}} \) from plane \( \Omega \). A plane inlier image \( H = \{ h_{ij} \} \) is defined, where \( h_{ij} = 1 \) if point \( p_{ij} \) is an inlier, otherwise \( h_{ij} = 0 \) (see Figure 2(c)). To recover potentially missing inliers due to noise, the topological closure operator is applied to \( H \) to obtain the filtered inlier image \( H' = \{ h'_{ij} \} \). The goal of the parcel boxes detection step is the generation of a set \( S \) of parcel boxes, identified by their 3D pose. Set \( S \) must be selected so that the top faces of the boxes, when projected on the image, contain the maximum number of inlier pixels and the minimum number of outlier pixels.

The algorithm is outlined as follows. First, a set of possible box edges is generated, by fitting 2D lines on the pixel with strong discontinuity in the point cloud and/or in the intensity image. These lines define a connectivity graph, where intersection points are the nodes. Box hypotheses are found by traversing this graph. Several conflicting box hypotheses may be generated, which cannot be true simultaneously as the boxes would interpenetrate. Hence, a subset of non-conflicting boxes is selected by maximizing an objective function using a genetic algorithm. These steps are detailed in the following.

The algorithm initially detects the edges of parcel boxes. An edge mask \( E = \{ e_{ij} \} \) is computed where the pixel \( e_{ij} = 1 \) in a box edge, and 0 otherwise (see Figure 2(d)). The detection is performed in both intensity and depth images in order to spot contours, even where boxes are tightly packed and there is no depth discontinuity. Edges in the intensity image \( R \) are detected by applying the standard Canny algorithm. Conversely, edge pixels are detected in the depth image using the plane inliers image, i.e., an edge pixel is an inlier pixel with an outlier in its 4-neighborhood. Edges are detected only in pixels where \( h'_{ij} = 1 \), as only plane inliers can belong to boxes (see Figure 2(e)).

Borders of parcel boxes potentially correspond to lines in the edge mask \( E \) and are detected using standard Hough transform (see Figure 2(f)). A connectivity graph is defined with the nodes corresponding to the intersection pixels and the edges linking the pixels that belong to the same edge line. The cycles of length four in connectivity graph are quadrilaterals in the edge image, and they may represent the top face of candidate box hypotheses \( \mathcal{B} = \{ B_n \} \) (see Figure 2(g)). A quadrilateral is accepted only if the lengths of the four edges match with the known box dimensions up to a tolerance \( t_{\text{box}} \) and the angles between two edges are right-angled with an angular tolerance \( \theta_{\text{box}} \).

The number of box hypotheses \( N_n \) corresponding to the quadrilaterals is usually larger than the true number of boxes due to multiple detection of the same borders or to spurious lines generated by noise, patterns, tags and adhesive tapes on the box. That is, there may be multiple overlapping box hypotheses, which correspond to the same real box. Thus, the inference phase selects a subset \( S = \{ S_n \}, S \subset \mathcal{B}, \) of box hypotheses from the power set \( \mathcal{P}(\mathcal{B}) \) so that conflicts between the hypotheses \( S_n \) are minimized. The inference is formulated as an optimization problem where the function of \( S \) to be
maximized is

\[
F(S) = \sum_{S \in \mathbb{S}} A(S) - \gamma_O \sum_{S \in \mathbb{S}} O(S) - \gamma_I \sum_{S \in \mathbb{S}} I(S) + \\
- \gamma_C \sum_{S \in \mathbb{S}} \sum_{S' \in \mathbb{S}, S' \neq S} C(S', S)
\]  

(1)

The objective function \(F(S)\) is the sum of a first term representing the total area of the hypotheses \(S \in \mathbb{S}\) minus three penalty terms. The first penalty term is the total area of the outlier pixels, i.e. the number \(O(S)\) of pixels with \(h_{ij} = 0\) for each \(S\). The second penalty term is the sum of the differences \(I(S)\) (in pixels) between the quadrilateral area and the expected area of the top face of a box for each hypothesis \(S\). The third penalty term is the total overlapping area (in pixels) between all hypotheses pairs \(S'\) and \(S'\). Coefficients \(\gamma_O, \gamma_I\) and \(\gamma_C\) are parameters weighting each penalty term.

Since online exact solution of the optimization problem is unfeasible, a genetic algorithm approach is adopted. An individual in the population is encoded by a dynamic array of box hypothesis indexes representing each \(\mathbb{S} \subset \mathbb{B}\). Two mutation operators and a crossover operator acting on the population are defined. Operator Mutate1(\(\mathbb{S}\)) removes up to three random elements. Operator Mutate2(\(\mathbb{S}\)) randomly replaces an element. The crossover operator Crossover(\(\mathbb{S}_1, \mathbb{S}_2\)) generates a new individual from two individuals \(\mathbb{S}_1\) and \(\mathbb{S}_2\) by selecting the new individual from hypotheses in \(\mathbb{S}_1 \cup \mathbb{S}_2\). To ensure faster convergence, a Fill operation is applied on the individuals resulting from the above operators. The Fill operation randomly augments the individual \(\mathbb{S}\) with new hypotheses \(B \in \mathbb{B}\) s.t. the augmented set has objective function \(F(\mathbb{S} \cup \{B\}) > F(\mathbb{S})\). An example of the final outcome is displayed in Figure 2(h).

B. Detection of gaps between boxes

The detection of gaps between boxes relies on a 2D camera with 8 Mpixel resolution mounted on the robot arm. It is assumed that during gap detection the end effector is moved in such a way that the image plane of the 2D camera is always parallel to the box plane and that their distance is kept constant. Moreover the image \(x\)-axis is oriented in the expected direction of the box edge as provided by the ToF camera estimation. This allows a further filtering of the results and an additional increase of detection accuracy. All previous choices do not affect the generality of the approach since they are only related to the definition of the specific motion trajectory of the robot end point. The algorithm consists of three steps. First, 2D lines are detected through the Probabilistic Hough Transform. The detected lines are defined by the image coordinates of their extremes \(p_k^s\) and \(p_k^e\). Then, extracted 2D lines are filtered according to their position, orientation and length in the image:

\[
(p_k^s, p_k^e) \in L \text{ if } \begin{cases}
\arctan \left(\frac{p_k^s - p_k^e}{p_k^s - p_k^e}\right) < \theta_{\min}, \\
\{s, e\} \frac{\{s, e\}}{p_{\min_y}^y} < \{s, e\} \frac{\{s, e\}}{p_{\max_y}^y}
\end{cases}
\]

(2)

where \(L\) is the set of accepted lines and \(\theta_{\min}\) and \(p_{\min_y}^y, p_{\max_y}^y\) are the orientation and vertical position threshold respectively. Through a proper choice of those parameters, only the horizontal lines in the center of the image are considered and most of outliers are removed. Therefore, the gap position \(p_{\text{ref}}\) and orientation \(\theta_{\text{ref}}\) used as reference for the picking operation are computed as the mean value center position and orientation of
\[ p_{\text{ref}} = \frac{1}{2n} \sum_{k=1}^{n} p_k^s + p_k^e \] (3)
\[ \theta_{\text{ref}} = \frac{1}{n} \sum_{k=1}^{n} \arctan \left( \frac{p_k^s - p_k^e}{\dot{p}_k^s - \dot{p}_k^e} \right) \] (4)
with \( n \) the number of elements in \( L \).

C. Tool Insertion and Extraction of the boxes

The tool insertion and the extraction of the boxes relies on different robot behaviours. During the insertion a compliance behaviour is required to assess the environment and verify that the gap has successfully being located. On the other hand both the alignment of the tool with the gap and the extraction of the box require a stiff behaviour and precise motions. To address these different control issues an admittance control scheme is proposed. The choice of an admittance control strategy is driven by the characteristics of the majority of industrial robots, that do not allow to directly control motor or joint torques. On the other hand, many collaborative industrial manipulators, as the UR5 used in this work, provide wrench estimation by means of current readings, torque sensors on the joints or a force/torque sensor on the wrist interface. Therefore an admittance control driving the robot position according to the tool wrench estimation results to be very effective. The workspace reference position \( x_{\text{ref}} \) provided to the manipulator controller is defined as the sum of the desired trajectory generated by the motion planner \( x_{\text{des}} \) and a displacement \( \Delta x_{\text{des}} \)

\[ x_{\text{ref}} = x_{\text{des}} + \Lambda \Delta x_{\text{des}} \] (5)

where the selection matrix \( \Lambda \) allows to dynamically set the direction in which the admittance control is active. According to the admittance control scheme, the displacement \( \Delta x_{\text{des}} \) depends on the robot estimated wrench \( F_{\text{ext}} \). Moreover, since the tool wrench is very noisy due to the fact that it is estimated from the motor current in the robot used in the experiments, a wrench threshold and a first-order digital filter are used to attenuate the wrench estimation error and noise. Therefore, the admittance displacement is defined as

\[ \Delta x_{\text{des}} = \frac{K_F z^{-1}}{1 + K_P z^{-1}} dz (\Lambda^TF_{\text{ext}}, F_{\text{th}}) \] (6)

where \( z^{-1} \) represents the sample-time delay according to the Z-transform notation, \( F_{\text{th}} \) is proper positive threshold used to remove the wrench estimation noise and the deadzone function \( dz(\cdot, \cdot) \) is defined as

\[
dz(F_1, F_2) = \begin{cases} 
F_1 - F_2, & F_1 > F_2 \\
0, & -F_2 < F_1 < F_2 \\
F_1 + F_2, & F_1 < -F_2 
\end{cases}
\] (7)

The ratio between the wrench coefficient \( K_F \) and the position coefficient \( K_P \) determines the level of compliance imposed to the system. During the tool insertion, if a collision with a box is detected, the current task is suspended and a new estimation of the gap between the boxes is performed. Therefore, this ratio should be selected to avoid damaging the boxes while performing the tool insertion before the control system decides that the insertion is failed. In case of failure, the tool insertion process is repeated after a new evaluation of the gap position. As an additional controller feature, the absolute value of \( K_F \) and \( K_P \) can be adjusted in order to the change the admittance controller reactivity.

Since the robot needs to be compliant along the \( z \)-axis only, while the motion along the other workspace directions are required to be stiff, see Fig. 1, the selection matrix is defined as a diagonal matrix with the main diagonal characterized by:

\[ \Lambda = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T \] (8)

Once the tool insertion is successful, the box is extracted through the motion of the blade along the \( x-y \) plane, see Fig. 1. Since the admittance behaviour is not active on this plane, see the selection matrix \( \Lambda \) definition in eq. (8), the stiff motion required for the extraction of the box is guaranteed.

III. Experiments

The experimental evaluation of the proposed depalletizing system has been carried out on the laboratory setup described in section II. Both perception and control algorithms have been implemented under the Robot Operating System (ROS) on a desktop PC running Ubuntu 18.04. Data communication between the robot arm, the two cameras and the PC was obtained through a Gigabit Ethernet network. The control loop is set to 125 Hz and the parameters used in the experiments are reported in Table I. All the reported data are expressed with respect to the robot base frame, as shown in Fig. 1. In particular, the \( x \) axis points toward the boxes, and the \( z \) axis points upward. Six boxes which size is 30x22x20 cm are considered, organised into three rows and two columns. All the boxes were placed next to each other on a rigid plane perpendicular to the robot base with no initial gap between them. The goal of the experiments is to collect a single box or a sequence of boxes as discussed in the following sections, regardless of their initial relative configuration.

A. Picking a single box

The experiment illustrated in Fig. 3 shows the robot removing a single box from the pallet in the experimental setup. Once the pose of the boxes has been determined by the fixed camera, a picking task is performed on the lowermost left box with respect to the robot base. In Fig. 4 the complete motion
of the robot is reported including the tool coordinates along the $x$ and $z$ axes (first two rows), and the tool wrench (third row). The robot starts by positioning the end-effector camera over the expected position of the gap between the boxes where the tool will be inserted to perform the refinement of the gap detection. Then the robot moves forward along the $x$ axis to position the tool according to the new estimation of the gap position starting at $t=5\text{ s}$. At time $t=11\text{ s}$ the first tool insertion is performed, but a peak in the tool wrench is detected, see the tool wrench plot at time $t\approx 14\text{ s}$, due to misalignment between the gap and the robot tool and the consequent contact of the tool with the boxes upper surface. This event triggers the backward motion of the tool and the repetition of the complete robot task. Then, starting at $t=20\text{ s}$ the gap detection is repeated, see Fig. 3b, and the tool is aligned with the gap as reported in Fig. 3c, while at $t=23\text{ s}$ the tool insertion starts, see Fig. 3d. The robot successfully completes the tool insertion at $t=33\text{ s}$ as shown in Fig. 3e, and it drags the box toward the robot as reported in Fig. 3f. It is worth noticing that the task execution time can be significantly reduced exploiting the full robot speed. It can also be noted that the robot exhibits a smooth transition of the controller behavior when a contact force is perceived by the manipulator.

The accuracy of the box detection and pose estimation algorithm described in section II-A has been assessed on a dataset of 125 depth images. The ToF camera observed the top layer of the pallet from about $1.10\text{ m}$. Five trials have been performed with a number of parcel boxes ranging from 3 to 9 in different configurations. For each trial 25 depth images of the scene were collected to estimate the boxes pose. The average standard deviation of the box position estimated by the proposed algorithm was $3.05\text{ mm}$, which is comparable with the position error of $3.60\text{ mm}$ due to pixel resolution.

This result confirms the usefulness of the robot admittance controller based on the eye-in-hand camera.

### B. Whole layer depalletizing

Several depalletizing experiments are performed on a full layer of boxes. The goal of each test is to unload all the boxes that could be reached by the robot. Due to the limited workspace of the robot arm, only the first two rows of boxes are accessible. This issue does not represent a limitation since in the future the robot arm will be mounted on a mobile base, therefore, it will be possible to move the arm around the pallet to extract all the boxes. Once the boxes are detected by the ToF camera, the picking order is planned first in ascending order along the $x$ axis, and then in ascending order along the $y$ axis. Fig. 5 shows the results of a successful experiment where the manipulator collects all the four boxes by performing at most two gap detection attempts.
A different experiment is proposed in Fig. 6, where before starting collecting the boxes, the boxes are slightly moved by hand in order to evaluate the robustness of the system. In this case it can be noted that the system is able to quickly adapt to the pose changes of the boxes performed after the first gap detection at time $t=40$ s and successfully completes the task. Also, in this case the manipulator collects all the boxes within a maximum of two gap detection attempts (the additional attempt in the collection of the first box is due to the change of box position). In Fig. 7 the success rate of the box extraction procedure evaluated over 12 extraction attempts is presented. The results show that for 50% of the tool insertions the first gap detection is sufficient, 41.6% of the times a second detection is required and only in a single case a third detection is performed before completing the task. This is probably due to changes in the lighting conditions, that should be less frequent in real industrial warehouses.

Fig. 6. A depalletizing task with real-time changes in box position. The boxes are manually moved at time $40$ s after the first gap detection.

Fig. 7. Overall insertion success rate with respect to gap detection repetitions.

IV. CONCLUSIONS AND FUTURE WORKS

In this work a robot setup is proposed for robotic industrial depalletizing based on a multi camera vision system and a collaborative manipulator controlled by an admittance control scheme. This setup is aimed at testing a new depalletizing system for the P-COORSA prototype described in [14]. The experimental trials reported in this paper show that the robot is able to autonomously detect and extract boxes from the top layer of a pallet. The experiments also demonstrate that the proposed solution based on dragging the boxes by means of a thin tool inserted between adjacent boxes greatly simplifies the box collection process, and significantly reduces the required robot payload. Future work will be focused on installing the robot manipulator on the P-COORSA prototype and to the implementation of more complex depalletizing tasks. Furthermore, the experimental evaluation in a real industrial warehouse are planned.

REFERENCES


