

Received November 3, 2020, accepted November 24, 2020, date of publication December 7, 2020, date of current version December 23, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3043069

Precise Robotic Manipulation of Bulky Components

PAOLO FRANCESCHI¹, (Associate Member, IEEE), NICOLA CASTAMAN^{2,3}, (Member, IEEE), STEFANO GHIDONI², (Member, IEEE), AND NICOLA PEDROCCHI¹, (Member, IEEE)

¹Institute of Intelligent Industrial technologies and Automation, National Council of Research of Italy, 20133 Milan, Italy

²Department of Information Engineering, University of Padova, 35131 Padova, Italy

³IT + Robotics srl, 36100 Vicenza, Italy

Corresponding author: Paolo Franceschi (paolo.franceschi@stiima.cnr.it)

This work was supported in part by the European Union's Horizon 2020 Research and Innovation Programme under Grant 738039, and specifically under the Clean Sky 2 Programme, and in part by the MIUR (Italian Ministry for Education) under the initiative "Departments of Excellence" (Law 232/2016).

ABSTRACT Increasing the flexibility of robots needs systems more capable in perceiving and interacting with the environment. A challenge is still to easy design the robotic system around the application, especially when the objects to be manipulated are bulky, and the relative positions between the robot and the objects to be manipulated are uncertain and high precision is required to successfully complete a task. In this paper a possible guideline to design a system capable to localize itself, identify a target, bulky, object and manipulate it, is presented. A method for tuning the impedance control parameters is shown, to keep interaction forces below dangerous values. The autonomous localization, grasping and assembly of a sidewall panel of an airplane is used as test. Experiments show that the success rate of completing a task increases, combining vision perception and force control, with respect to the single use of visual localization and position control.

INDEX TERMS Computer integrated manufacturing, autonomous systems, intelligent robots, robot control, robot motion, robot vision systems.

I. INTRODUCTION

The robotic manipulation and assembly tasks are widely studied and implemented in industry [1]. Typically the goal in industrial automated assembly applications is to reduce the cycle time, reducing costs and increasing productivity [2], [3]. In many applications, it is already possible to achieve high speed motion with autonomous free-collision trajectory generation [4], and, in the same time, high positioning performance [5], [6]. Many examples can be found in automotive applications [7]–[9] and electronic manufacturing where high precision in positioning is required, as well as flexibility to manipulate different small components [10].

In a wide field of applications, however, the relative position of the robot and the objects to be picked and handled is known with limited accuracy. In such a condition, when the inaccuracies to be compensated are few millimeters, or tens of degrees, the impedance control is widely used to supervise the interaction [12]–[14]. These algorithms have been proved in applications that need high-accurate force tracking [15]–[21],

but cannot compensate large deviations, since they rely only on the information given from measurements of forces and torques.

Large positions deviations represent an issue not only when the position of the objects to be manipulated is partially unknown, but also when the manipulated objects are large, and a small error in the picking orientation introduces large errors at the object extremities. In such a condition, the use of visual feedback in robotic control has been widely investigated from the begin of the '90s [23], realized mounting the camera on the robot wrist (eye-in-hand) or on a fixed position on the robot base (hand-to-eye). On the one hand, eye-in-hand is better in term of achievable accuracy since the calibration of the position of the camera with respect to the robot hand is often linearizable, and less sensitive to modelling errors. However such configuration suffers from occlusions, and when the handled object is texture-less, as close the camera is, as less references it can see. Therefore, when large objects have to be handled, the eye-in-hand configuration cannot be implemented. On the other hand, the hand-to-eye configuration allows greater flexibility in the positioning of the camera (or of the cameras), granting the possibility to avoid

The associate editor coordinating the review of this manuscript and approving it for publication was Okyay Kaynak¹.

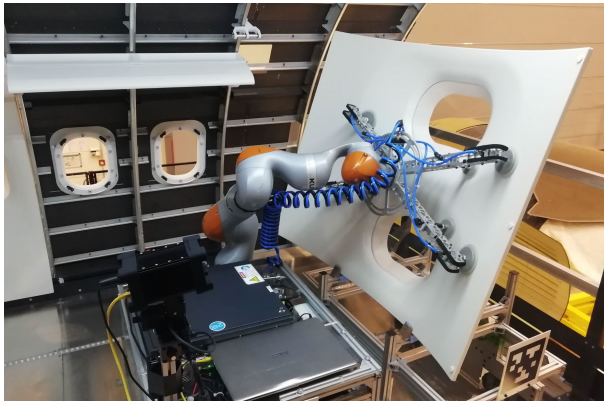


FIGURE 1. Example of manipulation of bulky objects in unstructured environment from EURECA project [11].

occlusion. However, such configuration is less accurate, since the calibration model is strongly non-linear [24], and if the vision system is mounted far from the object, the accuracy of the estimated position decreases.

The coupling of impedance controllers and visual servoing have been investigated from the very beginning of the 90's, to overcome these limitations [25], [26]. As matter of example, in [27] a vision system is used to online generate the set point for a robotic manipulator controlled by impedance control during peg-in-hole task. In [28] the authors use impedance control to improve the performances in the estimate, coming from a camera, of the position of a known and planar object. [29] propose an hybrid force and vision based impedance control in which vision is used to generate fictitious forces capable to modify the robot motion according to impedance law. Similarly [30] propose to use vision system to detect obstacles and to generate fake forces to perform an insertion. In [31] the authors used a vision system along with impedance control to implement an hybrid position/impedance control for the assembly of a small component. With the proposed approach it is necessary to teach the robot and the vision system by a demonstration assembly to calibrate and close the vision-robot-object chain, moreover impedance control is active in the direction of assembly to not exceed assembly forces, while in the other directions, position control is used, making impossible any error compensation due to the vision system.

An important step ahead in the state of the art is [22], where a position-based visual impedance control is proposed, shown in figure 2. Specifically, the methodology encapsulates a motion planner between the low-level impedance controller and the pose estimation stage, resulting in a framework that can be adapted in many industrial applications. The drawback of this approach is that the vision-system should have a continuous, clear, vision of the scene to feed impedance controller. This assumption is far to be realistic when bulky objects are manipulated, due to occlusions and to the fact that even a small error in the position estimate can become huge proportionally to the distance of the edges (e.g., an error of 1° becomes ~ 17 mm at the distance of 1 m).

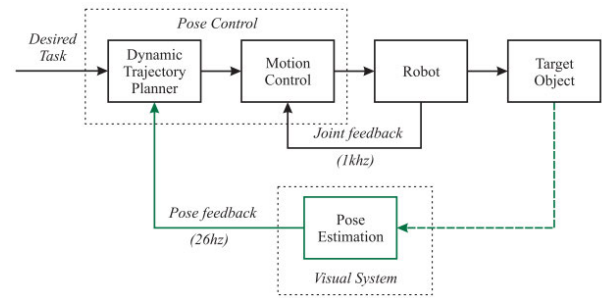


FIGURE 2. Position-based visual servoing, Schema from [22].

With the aim of deploying a methodology that increases the performance of a robotized system when picks and handles bulky objects, the here proposed approach extends the approach in [22], and the target position is computed asynchronously, triggering a new behavior for the robot. Making asynchronous the generation of the set-point, makes challenging the tuning of the impedance controller, since the vision system accuracy cannot grant the compensation of minor errors. Indeed, when large objects are manipulated, precision of 10 mm $\sim 0.5^\circ$ in the pose identification can be considered as a technical challenge. Impedance control is typically already implemented by the robot producer, and in any case is of easy implementation for a typical user, hence a fine tuning of its parameters is introduced to fulfill tolerances and force constraints, considering large objects. The work, therefore, will describe the criteria for the selection of the best methodology for each step, and how to tune the different modules, in an easy way. Specifically, the aim of this paper is to show how precision position of a robotic manipulator as well as low interaction forces can cohabit, exploiting vision system and impedance control, compensating for each other errors, involving the manipulation and assembly of a huge and bulky component, under restrictive tolerances. Such approach is tested in a real industrial case such as the autonomous assembly of an airplane sidewall panel, showing the power of the cooperation.

II. METHOD

In this section, easy and ready-to-use rules for the choice and tuning of various modules, for the precise robotic manipulation of a bulky object in a cluttered environment, is presented. The description of a method for choosing the hardware/software components according to the requirements, is introduced, than the driving choices for vision system selection are presented, along with a fine tuning of the classical impedance control parameters.

A. APPROACH

The flowchart in figure 3 presents the driving choices to select the proposed method, when commonly used approaches are not feasible.

Considering the pick and place tasks, usually, in industrial environments, fixed base robots are used to pick an object and

place, or assemble, it. The positions are predefined, in assembly lines, and precomputed, collision-free, trajectories are easily executed by the robot. In case the object and assembly poses are not known a priori, a vision sensor is typically introduced to localize them.

Many previous studies suggest 2D vision as a reliable and good approach, nevertheless, in the absence of strong features as textures, colors or high contrasts, 2D approach lacks in performance. Moreover, in the considered case, where distance and orientation of the object is not fixed, different viewpoints of the same object can create a different perception of the object, making 2D vision approach not feasible. When non-structured environments are considered, light source can not be fully controllable, and the object has different location and orientation with respect to the camera, 3D vision overcomes many challenges related to low contrast and poor lighting, as well as distance and orientation uncertainties, typical of 2D systems. Popular techniques for object recognition and pose estimation, based on 3D data, exploit 3D CAD models for object recognition and shape matching [32]. Many of these techniques exploits local descriptors [33], [34]. Local descriptors rely on strong geometrical local features. They do not have the notion of the whole object, but describe the local geometry around key points. As alternative, global descriptors are used. They are calculated for a set of points (*i.e.* a cluster) that represents an object. In the presented use case, an object with poor geometrical features is considered, hence a pipeline based on object segmentation and principal component analysis can be used as global descriptor. A segmentation step is required before extracting the cluster that represents the object.

Given the repeatability and accuracy of the vision system/algorithm, if tolerances required by the process are respected, an easy way to perform the manipulation is to compute online the trajectories according to the visual information, and execute them with a pure position trajectory following control. On the contrary, if corrections are required, visual servoing is a widely used approach to online compensate the motion with respect to errors in poses. Visual servoing might not be feasible, in particular considering the manipulation of large objects, that occlude the field of view of the camera. In such cases providing the robot with additional sensing can solve the problem. Interaction is involved, hence with a properly tuned impedance control, the robot can be provided with the compliance necessary to successfully perform the manipulation task.

The remainder of this section addresses the description of the vision system and algorithm, and a tuning method for the impedance control parameters, subject to geometrical and force constraint.

B. VISION SYSTEM AND POSE ESTIMATION

Considering a bulky object, camera specifications and its position with respect to it are crucial. In order to frame the entire part, the vision system must be in Eye-to-Hand

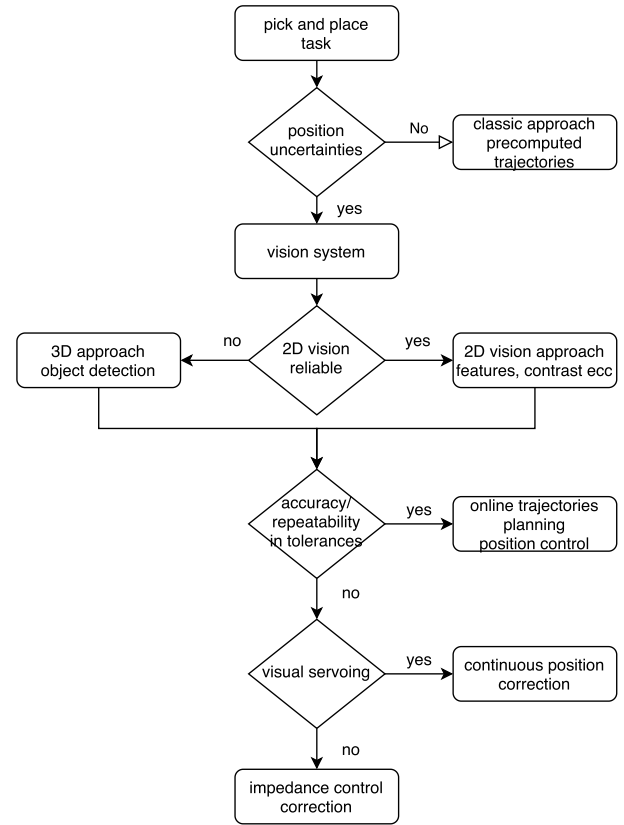


FIGURE 3. The flowchart of decision process. The flowchart highlights the most important decision to take into account in the system design phase.

configuration, meaning that the camera is rigidly placed with respect to the robot arm base.

The field of view (FOV) of the camera represents a limit with respect to the minimum distance between the camera and the object, that can be easily computed as

$$d_{camera}^{min} = \frac{s}{2 \tan(\frac{FOV}{2})} \tag{1}$$

where s is the maximum size of the object in one direction.

Hence the object must be placed with respect to the camera at a certain distance greater than d_{camera}^{min} . Moving the camera (e.g. with a PTU) can partially relax the constraint on the minimum camera distance at the cost of increased computational time and a more complex hand-eye calibration, introducing calibration errors. Moreover every 3D sensor has its specific optimal working range, hence, as additional constraint for choosing the position of the camera, this must be taken into account:

$$d_{range}^{min} < d_{camera}^{min} < d_{range}^{max} \tag{2}$$

Once the position of the camera is defined, pose estimation algorithms aiming at finding the coordinates of the object (position $P = (x, y, z)$ and orientation $O = (rx, ry, rz)$) with respect to the camera frame, can be designed. The developed algorithm is divided in 3 main steps: Segmentation, Matching, and Registration. Segmentation exploits the Region Growing [35] algorithm. This algorithm combines

the points that are close enough in terms of smoothness constraint, so that each extracted cluster is a set of points considered part of the same smooth surface. Region Growing also eliminates clusters that have a number of points below a defined threshold. Such threshold can be computed as

$$\#points_{min} = \frac{(size_x * size_y) * \#points}{4d^2 * \tan(FOV_x/2) * \tan(FOV_y/2)} \quad (3)$$

where d is the distance of the object from the camera, $size_x$ and $size_y$ are the sizes of the object respectively on x and y axes, and $\#points$ is the number of points in the cluster of the point cloud. In the matching phase, the first step consists in calculate the Oriented Bounding Box (OBB) for each cluster. The OBB is the smallest bounding box containing the set of points that compose the cluster. The OBB dimensions are compared with the ones of the part. If the dimensions are not comparable within a tolerance, the corresponding cluster is rejected. Instead, clusters whose OBB fulfill the geometric constraints are selected as possible match. In the end, Iterative Closest Point (ICP) [36] is used to refine the alignment and improve the pose estimation accuracy with each possible match. The distance between the reference object point cloud and the cluster of the possible match is also computed. The cluster with lower distance from the CAD model is selected and the relative pose is considered as the reference for the pose estimation.

To evaluate the accuracy of the registration the RMSE between the reference point cloud and the extracted point cloud is computed. However, pose estimation accuracy is also affected by different error sources, such as the camera calibration and the hand-to-eye calibration.

In the end, the goal of the vision system presented above is to provide the robot with good pose information, i.e., within the required tolerances imposed by the external features.

C. CARTESIAN IMPEDANCE CONTROLLER

The Cartesian impedance loop aims at modifying the tool set point, to make the robot end effector behavior as an equivalent mass-spring-damper system, described by the following:

$$\mathbf{M}_i \ddot{\mathbf{x}} + \mathbf{C}_i \dot{\mathbf{x}} + \mathbf{K}_i \mathbf{x} = \mathbf{F}_{ext} \quad (4)$$

where matrices \mathbf{M}_i , \mathbf{C}_i , \mathbf{K}_i represent impedance control inertia, damping and stiffness, respectively, and can be defined according to the desired compliance of the robot, and \mathbf{F}_{ext} is the vector of external forces.

From (4) it is possible to obtain the desired set of cartesian accelerations for the end effector [37]:

$$\ddot{\mathbf{x}}_d = -\mathbf{M}_i^{-1} [\mathbf{C}_i \Delta \dot{\mathbf{x}} + \mathbf{K}_i \Delta \mathbf{x} + \mathbf{F}_{ext}] \quad (5)$$

Figure 4 shows the block diagram of the control schema. The behavior of the end-effector of the robot it is now equivalent to a mass-spring-damper system, reacting and adapting to external forces, making the contact with the external environment compliant instead of being rigid.

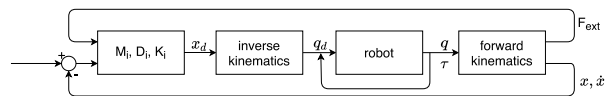


FIGURE 4. The control scheme of the impedance controller deployed in Cartesian space.

1) IMPEDANCE PARAMETER TUNING

To ensure low interaction forces with a desired dynamical behavior and a good position compensation, a method for tuning M_i , C_i and K_i is presented in this section. Without any loss of generality, for the rest of this section, only the translational direction along one axis is considered for the impedance parameters tuning. The system equations, while Cartesian impedance control is applied, correspond to the motion of a mass-spring-damper system. The reference set-point is chosen to be zero for the tuning of the parameters, with the environment applying a force. Considering the environment as a pure stiff wall, with stiffness defined as K_{env} , the force applied is produced by the relative motion between the environment and the system, computed as

$$F_{ext} = K_{env}(x_{env} - x) \quad (6)$$

where x_{env} and x are the environment and system positions, respectively.

Considering a funnel-shaped reference as external environment, the relative motion is described by a ramp limited in time

$$x_{env} = \alpha t, \quad 0 \leq t \leq \bar{t}$$

where α correspond to the slope, t is the time and \bar{t} the time where the slope flattens.

If the environmental stiffness K_{env} is not known, an estimate can be provided by an observer as in [38].

The system behavior (i.e. end-effector position) in a single direction of motion, is described by (7)

$$M_i \ddot{x} + C_i \dot{x} + K_i x = F_{ext} \quad (7)$$

Substituting (6) in (7) and rearranging leads to:

$$M_i \ddot{x} + C_i \dot{x} + K_{aug} x = K_{env} x_{env} \quad (8)$$

where $K_{aug} = K_i + K_{env}$ is the equivalent stiffness of the augmented system.

Considering an overdamped augmented system, the solution of the second order differential equation (8) is

$$x(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t} + \frac{K_{env} \alpha}{K_{aug}} t - \frac{C_i K_{env} \alpha}{K_{aug}^2} \quad (9)$$

For sufficiently large times, the two exponential terms will go to zero, living with the equation of a straight line with coefficients $m = \frac{K_{env} \alpha}{K_{aug}}$ and $q = -\frac{C_i K_{env} \alpha}{K_{aug}^2}$.

The relative position between the system and the environment is $x_{env} - x = \alpha t - mt + q$. Considering the maximum allowed force $F_{max} = K_{env}(x_{env} - x_{max})$ and evaluating the

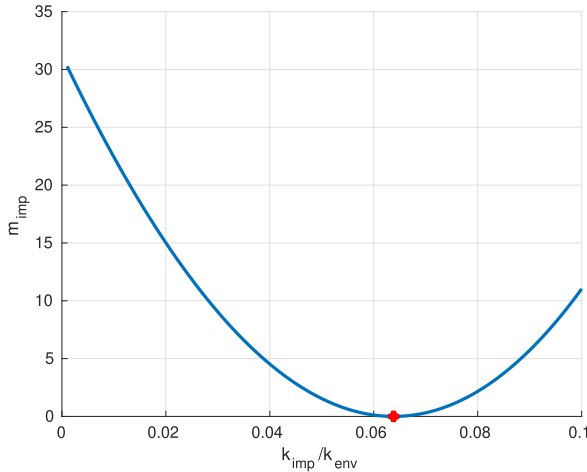


FIGURE 5. Values of M_i varying stiffness ratio. The value k_{lim} computed as in (13), is marked in red.

distance at $t = \bar{t}$ (where $x = x_{max}$), the following inequality must holds to avoid force overshoot.

$$\frac{F_{max}}{K_{env}} \geq \alpha \bar{t} - m \bar{t} + q \tag{10}$$

The solution (9) is computed for overdamped systems, hence in the following ζ_{aug} is taken to be larger than one, and the corresponding damping can be computed

$$C_i = 2\zeta_{aug}\sqrt{K_{aug}M_i} \tag{11}$$

From (10), substituting (11) in q and solving for M_i , the critical mass for which the maximum force is reached can be obtained.

$$M_{cr} = \left\{ \left[\frac{F_{max}}{K_{aug}} - \alpha \bar{t} \left(1 - \frac{K_{env}}{K_{aug}} \right) \right] \frac{K_{aug}^2}{2\zeta_{aug}\alpha\sqrt{K_{aug}}} \right\}^2 \tag{12}$$

Equation (12) satisfies (10) only for positive values of the term $\frac{F_{max}}{K_{aug}} - \alpha \bar{t} \left(1 - \frac{K_{env}}{K_{aug}} \right)$, after that point for the same value of mass there will be an higher value of stiffness, the transient will be the same but the system will “penetrate” more the environment, leading to higher forces that exceed the maximum allowed.

A limit stiffness ratio k_{lim} can be computed, where the mass is equal to zero

$$\frac{K_i}{K_{env}} < k_{lim} = \frac{1}{1 - \frac{F_{max}}{K_{env}\alpha\bar{t}}} - 1 \tag{13}$$

Given F_{max} , δx , \bar{t} , $\alpha = \frac{\delta x}{\bar{t}}$, it is possible to tune the values of M_i , C_i , K_i of the desired impedance behavior that guarantees no force overshoot.

$$\begin{cases} K_i = k_s k_{lim} K_{env} \\ M_i = m_s M_{cr} \\ C_i = 2\zeta_{aug}\sqrt{K_{aug}M_i} \end{cases} \tag{14}$$

where $k_s < 1$, $m_s < 1$ are safety coefficients used to avoid (i) too high natural frequency w the first, (ii) force overshoot due to imperfect modeling the second.

In figure 6 the time response of the system and the force exchanged are shown for different values of $k_s = [1.1, 0.9, 0.5, 0.1]$, for $m_s = 1$ and $\zeta_{aug} = 1.2$, $F_{max} = 30N$, $\delta x = 0.01m$, $\bar{t} = 2.2s$. It can be seen the wrong behavior of the system for $k_s > 1$, the value of the masses is very similar for the first two cases, but the difference in stiffness produces different behavior hence different forces.

The steady state position error, when the ramp flattens can be computed imposing to the system a constant force $F_{ext} = K_{env}\delta x$, solving equation (8) leads to the final

$$x_{ss} = \frac{K_{env}}{K_{aug}} \delta x \tag{15}$$

The steady state error will be

$$e_{ss} = \delta x - x_{ss} = \delta x \left(1 - \frac{K_{env}}{K_{aug}} \right) \tag{16}$$

D. CONSIDERATIONS

The 3D vision system is able to estimate the poses within a certain tolerance, hence some considerations regarding the maximum produced errors must be taken. Errors can be introduced by (i) a wrong estimate of the pose and (ii) a wrong impedance compensation both for grasping and place position. The errors in pose estimation by the camera are due to its resolution, hence the maximum δx to be compensated by the impedance control depends on it. During the grasping phase, in the worst case scenario of maximum error in position estimate, the error after impedance compensation will be e_{ss}^g as in (16), the same applies to place position and we will have e_{ss}^p , where superscripts g and p stands for *grasping* and *place* respectively. Knowing the repeatability range rr of the camera for both the poses and assuming normal distribution, the maximum displacements to be compensated, due to linear position estimate, can be computed as $\delta x_{lin}^g = \frac{rr_{lin}^g}{2}$ and $\delta x_{lin}^p = \frac{rr_{lin}^p}{2}$, moreover δx^p has to compensate for e_{ss}^g too, hence $\delta x^p = \delta x_{lin}^p + e_{ss}^g$. Considering large objects, the references may be positioned at the edges, far from the nominal pose estimate, hence a wrong estimate in the orientation of the pose can produce a large error in the relative position between references on the environment and references on the object/robot. Taking this into account, having the reference feature at a length l from the nominal estimated pose, having an orientation error $\delta\theta = \frac{rr_{\theta}}{2}$, the position error of the feature due to $\delta\theta$ will be $\delta x_{rot} = l \sin(\delta\theta)$. In this case $\delta x^g = \delta x_{lin}^g + \delta x_{rot}^g$ and $\delta x^p = \delta x_{lin}^p + \delta x_{rot}^p + e_{ss}^g$.

III. CASE STUDY

The proposed methodology is implemented in a real industrial scenario, presented in this section. The results here shown are intended to be an anecdotal description of a use case where the suggested approach generates a great improvement with respect to other two typical approaches in robotic manipulation.

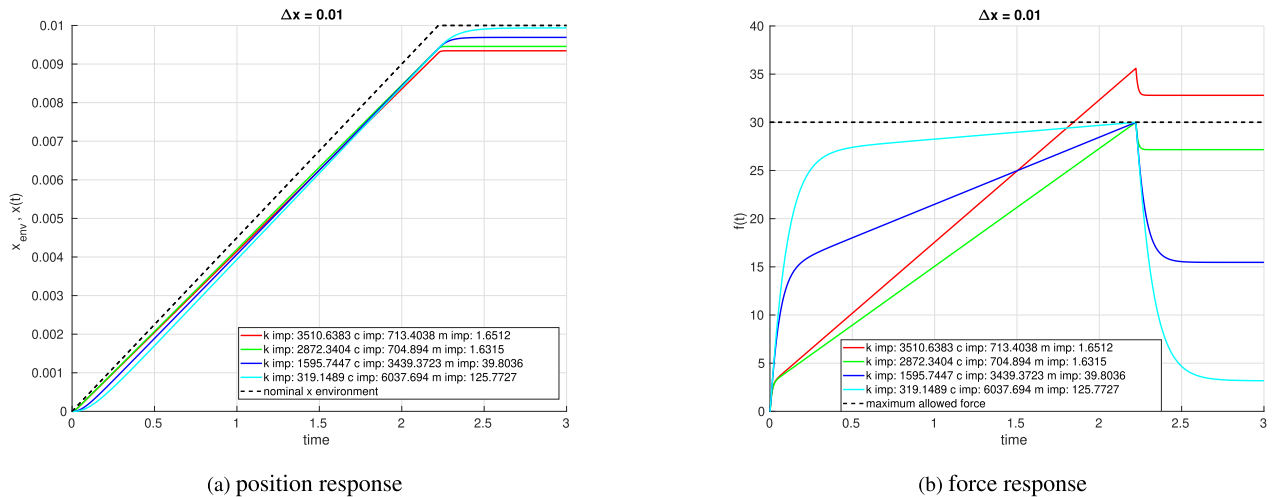


FIGURE 6. Position and force response of the system with different values of k_s .

A. SYSTEM DESCRIPTION

The chosen robot manipulator is a KUKA iiwa 14 R820, which is suitable, thanks to its redundancy, to perform motions in a small and cluttered environment as the interior of an airplane fuselage. Moreover it has high payload (14 kg), required for heavy objects manipulation. The robot is equipped with a tool provided with suction cups and a vacuum system to grasp, manipulate, and assemble the panel. Two independent circuits can be activated, via electrovalve, to manipulate curve objects (sidewall panels) as well as planar objects (cargo panels). The vision system is composed by a pan-tilt unit (PTU) in order to cover all the working area and a depth camera to acquire 3D data. The selected pan-tilt unit is a FLIR PTU-46 that guarantee a pan range of $\pm 159^\circ$ and a tilt range from -47° to 31° with a resolution of 0.003° . 3D data are acquired with a ODOS StarForm Swift Time-of-Flight (ToF) depth camera. The camera has a resolution of 640×480 px at 40 fps and works in the range of 0.5-6 meters and a precision around 1 cm. The robot and the camera, as well as the sidewall panel, are mounted on passive wheeled mobile carts, necessary to move them inside the aircraft fuselage easily. This element introduces uncertainties in the relative positions between robot and panel and robot and assembly point.

B. TASK DESCRIPTION

The task the robot has to perform is the autonomous assembly of an aircraft sidewall panel, depicted in fig 7.

Manipulating bulky components inside narrow spaces such as the interior of aircraft fuselage, limits the robot's movements which must be precise to avoid collisions. The correct localization of the robot and the object is fundamental to online create a correct environment for trajectory planning. The vision system initially identifies the fuselage windows to provide the proper position of the robot in the environment, than it scans the panel and localize it with respect to the robot pose. A 3D map of the environment is than created according to the information coming from the camera, and the trajectories are planned online.

Impedance control, conveniently tuned, is adopted for the tasks that requires contacts and interaction between the robot gripper and the panel or the environment, in this application for grasping and assembly. The grasping task begins with the robot positioned in front of the panel, according to the position received from the camera system. A forward motion is performed along the outgoing direction of the tool, while impedance control is active. Since it is required that all the four suction cups are properly aligned and fully in contact with the surface to guarantee a successful grasping, small reorientations are allowed by the compliant controller. Indeed even a small error in the position and/or orientation estimate of the panel, even due to the camera resolution, may lead to a misalignment of the tool with respect to the curved surface of the panel. The assembly procedure of the sidewall panel is the most complex of the whole process, indeed, since the tolerances of the assembly features are very tight ($1 \div 10$ mm depending on the directions), the allowed error is very small. Relying on pose estimation only, it is very difficult, usually impossible, to properly fit the features. As for the grasping task, the success of the assembly strongly relies on interaction control to ensure the correct position.

C. EXPERIMENTAL RESULTS

The experimental results are shown in this section. First of all a measurement of the accuracy and repeatability of the camera in estimating the poses is presented. Than a comparison between three different approaches (vision system without impedance control, impedance control only, vision system and interaction control working together) is done keeping the poses fixed. Finally the robustness of the combined approach is shown in terms of success rate, randomly changing the poses.

1) VISION SYSTEM

The most important values that the vision system has to estimate, for both the assembly and the panel positions, are x and y distances and rotation θ around z axis, with respect

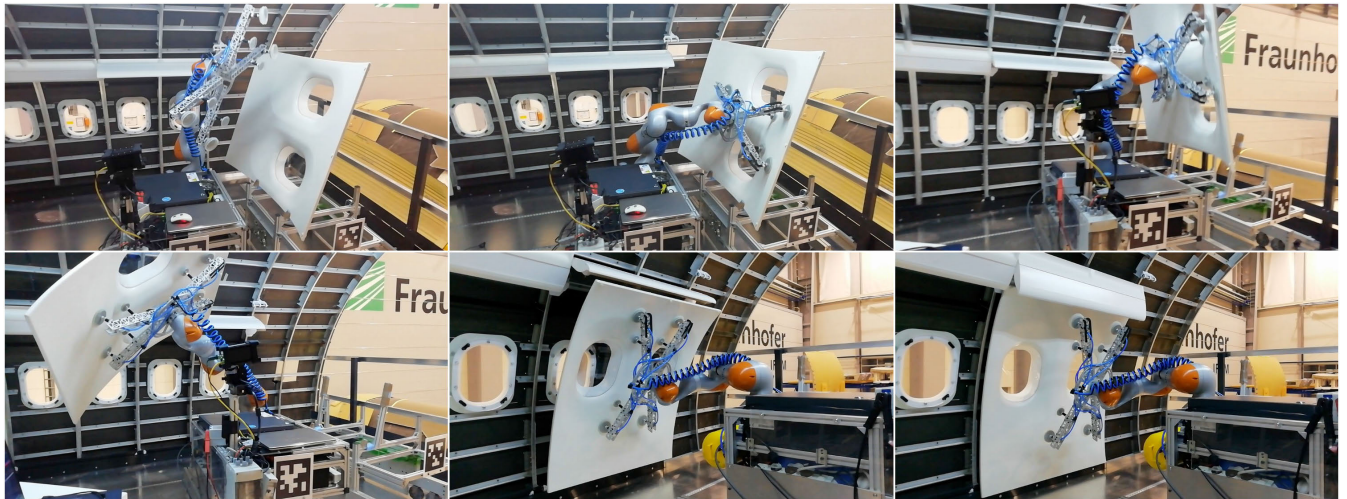


FIGURE 7. The autonomous assembly task of the sidewall panel inside the aircraft cabin.

TABLE 1. x , y and θ estimated for assembly and panel poses.

	mean	diff
x_w [m]	1.0660	0.0136
y_w [m]	0.1319	0.0034
θ_w [°]	0.6037	0.8756
x_p [m]	-0.4273	0.0129
y_p [m]	-0.7399	0.0346
θ_p [°]	89.108	1.6363

to the origin of the robot, since the other values (height and other rotations) are fixed and known. In figure 8 the estimates of x and y from the vision system are shown, with respect to the robot base. The estimate of the x , y , and θ values for both the windows and panel positions, as well as the mean and maximum deviations, are presented in table 1.

2) EXPERIMENTAL TESTS WITH QUASI-FIXED POSITIONS

In this section three different approaches are presented and compared in order to show the robustness of our approach, considering quasi-fixed positions (small errors in positioning can be introduced since the panel position was manually restored after each trial). Table 2 summarizes the success for each approach in localizing the robot and the panel trough vision system (when used), as well as the success in performing the grasping and assembly tasks.

- a. **Vision/position control** The first 10 trials were made relying on vision information only: the robot target positions were exactly the positions given by the camera. Despite the success rate of identifying the panel and the windows is very high, the success of the complete task is null. This because, even if the grasping of the panel sometimes is successful, it is not done in the correct position: because of the complex shape of the sidewall panel, even a small error in the position estimate makes the grasping not feasible, or the coupling between the tool and the panel is not done according to the nominal

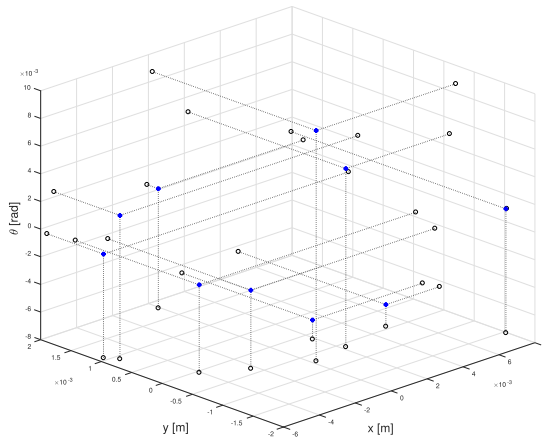
one. It has to be underlined that the 0/10 successes of the assembly is with respect to the 10 trial, but, since it directly depends on the 4/10 grasping successes, it is actually measured on 4 trial only.

- b. **Impedance control** If only impedance control is used, without any precise information coming from the vision system, the success rate still remains very low. In this case, the grasping is made more robust due to the capability to adapt to small errors trough impedance control, but the assembly is very complex since the tolerances don't allow any big error. Moreover if the position between the panel and the robot, or the robot and the fuselage change, even a little, there is no possibility to compensate for any error, making this approach not flexible at all.
- c. **Vision/impedance control** Combining the capabilities of both the approaches, the vision system is able to compensate for huge error and estimate a good initial position, on the other hand, impedance control is very good in compensating for small errors in the estimate of the camera. In this case, for the grasping task there is no problem. In the assembly task, the success rate is very high with respect to the two other cases. Despite this some failure may still happen because of the very restrictive tolerances on the assembly features. The time responses of the position and force of the robot end-effector, in the direction of the environmental correction, are shown in figure 9.

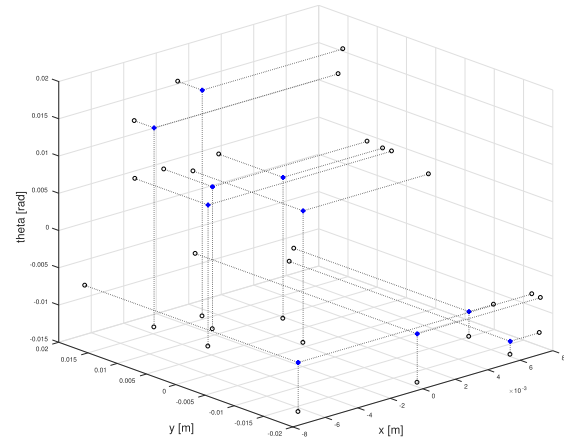
3) ERRORS ANALYSIS

In this section the errors causing failure to the process are discussed with respect to the three different approaches illustrated above.

- a. **Error in poses estimation:** Considering the vision/position and vision/impedance control it is not correct to talk about errors in pose estimate, rather it is a resolution



(a) estimate of windows poses wrt robot base



(b) estimate of panel poses wrt robot base

FIGURE 8. x , y and θ values estimated by the vision system with respect to the robot base, relative to 10 experiments repetitions.

TABLE 2. Comparison between success of the three approaches: vision based, impedance control based and combination between vision and impedance. Results are acquired in 10 repetitions.

	Robot Localization	Panel Localization	Panel Grasping	Panel Assembly
Vision and position control	100%	100%	40%	0%
Impedance control without vision	n/a	n/a	70%	10%
Vision and impedance control	100%	100%	100%	90%

limit that arises. The repeatability is presented above, in table 1. From the experiments, it was observed that errors may occur only in cases in which the field of view of the camera is partially occluded and the object can be only partially seen. In the cases in which vision system is not present the poses are not estimated, but taken for granted even if small errors may occur, due to manual/autonomous (with an AGV) positioning.

- b. **Error in panel grasping:** Errors in panel grasping occur when the coupling between the tool and the panel is not properly done, and the panel is not grasped. They can be the result of an improper estimate/knowledge of the pose: it occurs mainly in the case of vision/position control since there is no possibility to compensate for a small error in the pose estimate. The panel surface has a complex, non-planar shape: since all the four suction cups must be in contact with the surface to grasp the panel, a small error in the estimate can produce a misalignment, preventing the possibility to properly touch the surface. Impedance control performs better in rejecting this error, since the compliance of the robot allows small reorientation, having the correct alignment and contact between the tool and the panel surface. Errors can still occur if the pose of the panel is not the nominal one. Vision/impedance control allows to have a good estimate of the pose and a compensation for small errors (in particular in orientation), increasing the success rate.
- c. **Error in assembly:** Errors in the assembly phase occur when the panel is not properly inserted in the features

nor properly fixed. They can be caused either by an improper estimate of the assembly position and/or a grasping not in the nominal position, that is an incorrect coupling between the manipulator tool and the panel. In these cases even if the assembly pose is properly estimated, during the insertion phase the panel is not properly positioned. In the case of vision/position control a compensation can not be done, while considering a compliant behavior of the robot (impedance control, vision/impedance control), if the error is into the allowed tolerances, the insertion can be properly done, adjusting small errors thanks to the robot behavior.

In conclusion vision/impedance control combines the potentiality of the vision system to estimate a correct enough initial position, compensating for small errors with the robot compliant behavior, allowing to have a very high success rate.

4) ROBUSTNESS WITH DIFFERENT POSES

Proven the reliability of the combined approach composed by vision information and impedance control, the robustness of this solution is tested, changing for 10 repetitions the position of the robot with respect to the panel and the assembly position. The positions are randomly selected, with respect to the nominal, in the range ± 100 mm both in x and y direction and $\pm 5^\circ$ for the orientation along z axis, for both panel and assembly positions. The success rate in this test is the same as in the previous case, with 9/10 successfully assembly performed. The only condition that must be satisfied is the

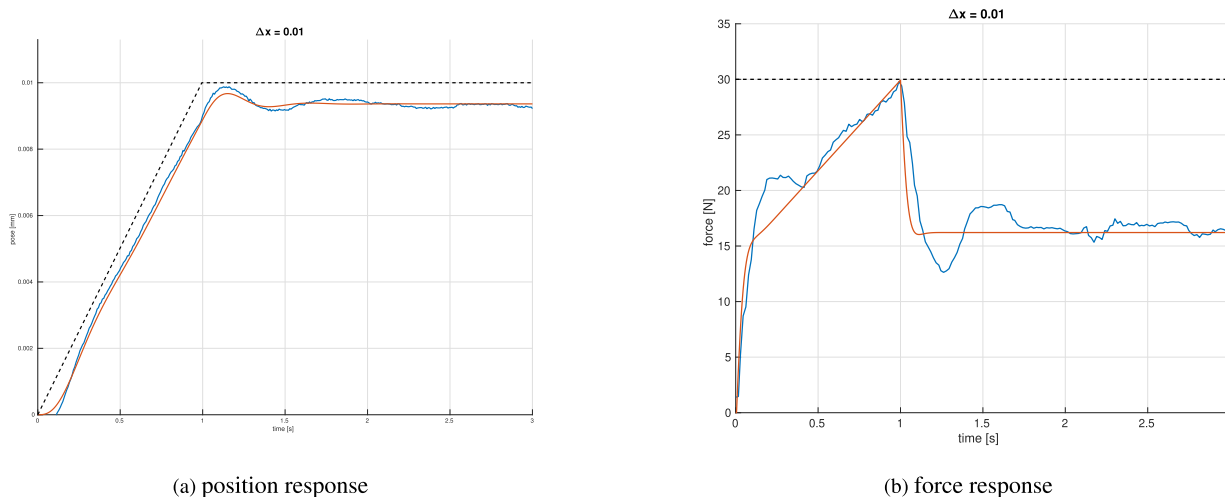


FIGURE 9. Position and force response of the system. The dotted lines are the reference displacement over the time, the blue lines are the measured displacement and force, the red lines are the simulated displacement and force.

reachability of the robot of both the panel and the assembly pose. A second important condition is to keep the robot far enough from singularity positions while force control is used, this because forces are estimated through the Jacobian matrix, which becomes singular. With the usage of a force/torque sensor at the end-effector this constraint can be relaxed.

IV. CONCLUSIONS

Combining the potentiality of vision and force control it is shown that a complex task can be successfully completed, due to their ability to compensate for each other errors.

This paper aims to show how precise position of a robotic manipulator as well as low interaction forces can cohabit, exploiting vision system and impedance control, compensating for each other errors, involving the manipulation and assembly of a huge and bulky component under restrictive tolerances. Such approach is tested in a real industrial case such as the autonomous assembly of an airplane sidewall panel, showing the powerful of the cooperation.

In conclusion in this paper is shown how vision assisted impedance control can be useful to solve a real industrial application. For the selected application the object to be manipulated is huge and bulky and the environment is very reduced, having very low allowed errors. The possibility to perform visual servoing was precluded due to the huge dimensions of the panel, which occluded the possibility to have reference. In the end it is demonstrated the robustness with respect to error in positioning of this approach: a priori knowledge of the positions of the robot and the panel is not required.

This work considers the manipulation of a huge and bulky, rigid object and its precise assembly under strict tolerances. Future research work will address different objects manipulation, involving non-rigid objects manipulation, where some degree of deflection must be taken into account. Moreover a more complex model for the environment with a damping component will be investigated, for a better comprehension

of the interaction forces that may lead to an unsuccessful assembly or damage to the object.

ACKNOWLEDGMENT

The work has been developed within the project EURECA.

REFERENCES

- [1] H. S. Cho, H. J. Warnecke, and D. G. Gweon, "Robotic assembly: A synthesizing overview," *Robotica*, vol. 5, no. 2, pp. 153–165, Apr. 1987.
- [2] F. Chen, F. Cannella, C. Canali, T. Hauptman, G. Sofia, and D. Caldwell, "In-hand precise twisting and positioning by a novel dexterous robotic gripper for industrial high-speed assembly," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2014, pp. 270–275.
- [3] Y. Koseki, T. Arai, K. Sugimoto, T. Takatiji, and M. Goto, "Design and accuracy evaluation of high-speed and high precision parallel mechanism," in *Proc. IEEE Int. Conf. Robot. Automat.*, vol. 2, May 1998, pp. 1340–1345.
- [4] C.-L. Chen, T.-C. Wu, and C.-C. Peng, "Robust trajectories following control of a 2-link robot manipulator via coordinate transformation for manufacturing applications," *Robot. Comput.-Integr. Manuf.*, vol. 27, no. 3, pp. 569–580, Jun. 2011.
- [5] S. N. Gottschlich and A. C. Kak, "A dynamic approach to high-precision parts mating," in *Proc. IEEE Int. Conf. Robot. Automat.*, vol. 2, Apr. 1988, pp. 1246–1253.
- [6] W. M. Spong, *Robot Dynamics and Control*, 1st ed. Hoboken, NJ, USA: Wiley, 1989.
- [7] S. Peters, G. Lanza, J. Ni, J. Xiaoning, Y. Pei-Yun, and M. Colledani, "Automotive manufacturing technologies—An international viewpoint," *Manuf. Rev.*, vol. 1, p. 10, 2014, doi: [10.1051/mfreview/2014010](https://doi.org/10.1051/mfreview/2014010).
- [8] G. Michalos, S. Makris, N. Papakostas, D. Mourtzis, and G. Chryssolouris, "Automotive assembly technologies review: Challenges and outlook for a flexible and adaptive approach," *CIRP J. Manuf. Sci. Technol.*, vol. 2, no. 2, pp. 81–91, Jan. 2010.
- [9] S. Pellegrinelli, N. Pedrocchi, L. M. Tosatti, A. Fischer, and T. Tolio, "Multi-robot spot-welding cells for car-body assembly: Design and motion planning," *Robot. Comput.-Integr. Manuf.*, vol. 44, pp. 97–116, Apr. 2017.
- [10] A. C. Sanderson and G. Perry, "Sensor-based robotic assembly systems: Research and applications in electronic manufacturing," *Proc. IEEE*, vol. 71, no. 7, pp. 856–871, Jul. 1983.
- [11] L. Roveda, S. Ghidoni, S. Cotecchia, E. Pagello, and N. Pedrocchi, "Eureca h2020 cleansky 2: A multi-robot framework to enhance the fourth industrial revolution in the aerospace industry," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA), Workshop Ind. Future, Collaborative, Connected, Cogn. Novel Approaches Stemming from Factory Future Ind. 4.0 Initiatives*, 2017.
- [12] N. Hogan, "Impedance control: An approach to manipulation," in *Proc. Amer. Control Conf.*, Jul. 1984, pp. 304–313.

- [13] X. Wang, Y. Wang, and Y. Xue, "Adaptive control of robotic deburring process based on impedance control," in *Proc. IEEE Int. Conf. Ind. Informat.*, Aug. 2006, pp. 921–925.
- [14] S. Jung and T. C. Hsia, "Adaptive force tracking impedance control of robot for cutting nonhomogeneous workpiece," in *Proc. IEEE Int. Conf. Robot. Automat.*, vol. 3, May 1999, pp. 1800–1805.
- [15] L. Roveda, N. Iannacci, F. Vicentini, N. Pedrocchi, F. Braghin, and L. M. Tosatti, "Optimal impedance force-tracking control design with impact formulation for interaction tasks," *IEEE Robot. Autom. Lett.*, vol. 1, no. 1, pp. 130–136, Jan. 2016.
- [16] L. Roveda, N. Iannacci, and L. M. Tosatti, "Discrete-time formulation for optimal impact control in interaction tasks," *J. Intell. Robot. Syst.*, vol. 90, nos. 3–4, pp. 407–417, Jun. 2018.
- [17] L. Roveda, "Adaptive interaction controller for compliant robot base applications," *IEEE Access*, vol. 7, pp. 6553–6561, 2019.
- [18] L. Roveda, N. Pedrocchi, F. Vicentini, and L. M. Tosatti, "An interaction controller formulation to systematically avoid force overshoots through impedance shaping method with compliant robot base," *Mechatronics*, vol. 39, pp. 42–53, Nov. 2016.
- [19] L. Roveda, N. Pedrocchi, M. Beschi, and L. M. Tosatti, "High-accuracy robotized industrial assembly task control schema with force overshoots avoidance," *Control Eng. Pract.*, vol. 71, pp. 142–153, Feb. 2018.
- [20] X. Sheng and X. Zhang, "Fuzzy adaptive hybrid impedance control for mirror milling system," *Mechatronics*, vol. 53, pp. 20–27, Aug. 2018.
- [21] L. Roveda, N. Pedrocchi, and L. M. Tosatti, "Exploiting impedance shaping approaches to overcome force overshoots in delicate interaction tasks," *Int. J. Adv. Robot. Syst.*, vol. 13, no. 5, pp. 1–11, 2016.
- [22] V. Lippiello, B. Siciliano, and L. Villani, "A position-based visual impedance control for robot manipulators," in *Proc. IEEE Int. Conf. Robot. Automat.*, Apr. 2007, pp. 2068–2073.
- [23] B. Espiau, F. Chaumette, and P. Rives, "A new approach to visual servoing in robotics," *IEEE Trans. Robot. Autom.*, vol. 8, no. 3, pp. 313–326, Jun. 1992.
- [24] F. Vicentini, N. Pedrocchi, M. Malosio, and L. Molinari Tosatti, "High-accuracy hand-eye calibration from motion on manifolds," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2011, pp. 3327–3334.
- [25] B. Nelson, N. Papanikolopoulos, and P. Khosla, "Visual servoing for robotic assembly," in *Visual Servoing: Real-Time Control of Robot Manipulators Based on Visual Sensory Feedback*. Singapore: World Scientific, 1993, pp. 139–164.
- [26] B. J. Nelson, N. P. Papanikolopoulos, and P. K. Khosla, "Robotic visual servoing and robotic assembly tasks," *IEEE Robot. Autom. Mag.*, vol. 3, no. 2, pp. 23–31, Jun. 1996.
- [27] G. Morel, E. Malis, and S. Boudet, "Impedance based combination of visual and force control," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 1998, pp. 1743–1748.
- [28] V. Lippiello, B. Siciliano, and L. Villani, "A position-based visual impedance control for robot manipulators," in *Proc. IEEE Int. Conf. Robot. Automat.*, Apr. 2007, pp. 2068–2073.
- [29] R. Carelli, E. Oliva, C. Soria, and O. Nasisi, "Combined force and visual control of an industrial robot," *Robotica*, vol. 22, no. 2, pp. 163–171, Mar. 2004.
- [30] Y. Nakabo and M. Ishikawa, "Visual impedance using 1 ms visual feedback system," in *Proc. IEEE Int. Conf. Robot. Automat.*, vol. 3, May 1998, pp. 2333–2338.
- [31] Y. Lei, J. Xu, W. Zhou, W. Sun, Y. Wu, J. Luo, and H.-T. Zhang, "Vision-based position/impedance control for robotic assembly task," in *Proc. Chin. Control Conf. (CCC)*, Jul. 2019, pp. 4620–4625.
- [32] A. Aldoma, M. Vincze, N. Blodow, D. Gossow, S. Gedikli, R. B. Rusu, and G. Bradski, "CAD-model recognition and 6DOF pose estimation using 3D cues," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCV Workshops)*, Nov. 2011, pp. 585–592.
- [33] R. B. Rusu, N. Blodow, and M. Beetz, "Fast point feature histograms (FPFH) for 3D registration," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2009, pp. 3212–3217.
- [34] F. Tombari, S. Salti, and L. D. Stefano, "Unique signatures of histograms for local surface description," in *Proc. Eur. Conf. Comput. Vis.* Berlin, Germany: Springer, 2010, pp. 356–369.
- [35] T. Rabbani, F. Van Den Heuvel, and G. Vosselmann, "Segmentation of point clouds using surface smoothness constraint," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. 36, no. 5, pp. 248–253, 2006.
- [36] P. J. Besl and N. D. McKay, "Method for registration of 3-D shapes," *Proc. SPIE*, vol. 1611, pp. 586–607, Apr. 1992.

[37] B. Siciliano and L. Villani, *Robot Force Control*, 1st ed. Norwell, MA, USA: Kluwer, 2000.

[38] L. Roveda, F. Vicentini, and L. M. Tosatti, "Deformation-tracking impedance control in interaction with uncertain environments," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Nov. 2013, pp. 1992–1997.



PAOLO FRANCESCHI (Associate Member, IEEE) received the B.S. degree in aerospace engineering and the M.S. degree in mechanical engineering from the Politecnico of Milano, in 2015 and 2017, respectively.

In 2016, he joined as a Visiting Scholar at the IBM Almaden Research Center, San José, CA, USA, to complete his master's thesis dealing with modeling and control of a bipedal robot. Since 2018, he joined STIIMA-CNR, Milan, Italy, as a Research Fellow. His research interests include interaction control for robotic manipulators and optimization of complex robotics manipulation working area and tasks.



NICOLA CASTAMAN (Member, IEEE) received the M.Sc. degree in computer engineering from the University of Padova, Italy, in 2016, where he is currently pursuing the Ph.D. degree in autonomous robotics with the Department of Information Engineering. From 2016 to 2017, he was a Research Fellow with the Intelligent Autonomous Systems Laboratory (IAS-Lab), Department of Information Engineering, University of Padova. In 2017, he completed in the MBZIRC international robotics challenge with the Desert Lion Team of the IAS-Lab, where they obtain 3rd place in the Grand Challenge. He is also an R&D Developer at IT + Robotics srl. His research interests are in task and motion planning and mobile manipulators for industrial robotics applications.



STEFANO GHIDONI (Member, IEEE) is currently an Assistant Professor with the University of Padova, Italy. His research interests are in semantic segmentation, people re-identification systems, and deep learning for scene understanding, mainly targeting robotics applications. His research interests also include deep learning for human-robot cooperation, focused on robot task and motion planning when a human is operating inside the robot working range. He was involved in a number

of industrial projects for bringing intelligence into the industrial processes, including airplane and boat production.



NICOLA PEDROCCHI (Member, IEEE) received the M.S. degree in mechanical engineering from the University of Brescia, Brescia, Italy, in 2004, and the Ph.D. degree in applied mechanics and robotics, in 2008.

From 2008 to 2011, he was a Research Assistant with the Institute of Industrial Technologies and Automation, National Research Council of Italy, where he has been a Full Researcher, since 2011. His research field is on control techniques for industrial manipulators in advanced application requiring the interaction robot-environment (e.g. technological tasks) or robot-human operator (e.g. workspace sharing, teach-by-demonstration). He is involved in researches for accurate elastic modeling and dynamic calibration of industrial robots. Since 2015, he has been coordinating the activity of the Robot Motion Control and Robotized Processes Laboratory at CNR-STIIMA.

...