

Peg-in-Hole Using 3D Workpiece Reconstruction and CNN-based Hole Detection

Michelangelo Nigro[†], Monica Sileo[†], Francesco Pierri[†],
 Katia Genovese[†], Domenico D. Bloisi^{*}, Fabrizio Caccavale[†]

I. INTRODUCTION

The evolution of manufacturing in last decade demands for more autonomous, safe and effective robotic systems, capable of fast adaption to rapidly changing production requirements. Such a flexibility can be achieved by endowing robotic systems with advanced capabilities of executing complex tasks in partially structured environments. Thus, new approaches, integrating different methodologies and technologies, are to be investigated. To this aim, in this paper, a novel approach to semi-autonomous assembly tasks execution is developed and tested by considering a Peg-in-Hole operation, which can be considered a good benchmark case for assembly tasks.

The common approach to Peg-in-Hole includes two steps: the search, aimed at localizing the hole and aligning the peg to its axis, and the insertion. Regarding the insertion phase, it can be performed via pure positional control only in the case of perfect localization of the hole. Otherwise, active or passive force control methods are to be adopted. In practical application scenarios, the accumulated errors due to imperfect knowledge of the hole position and robot positioning errors, make the pure position control approach possible only in the presence of generous clearances, while the approaches based on impedance control [1] are the most popular. In this work the Peg-in-Hole task has been considered for an industrial scenario where the target object is positioned manually in the workcell. Thus, it is possible to assume that the workpiece is located within the robot workspace, but uncertainties on both the holes' position and their tilt can be far larger than the task tolerance.

A Convolutional Neural Network (CNN) based approach for object detection is combined with a Stereo-DIC surface reconstruction method to accurately localize the holes. Finally, the peg insertion is performed by means of an admittance control, that confers a mechanical impedance behavior to the peg, without resorting to a wrist-mounted force/torque sensor.

II. PROPOSED STRATEGY

The following strategy is proposed:

- The robot moves an eye-in-hand stereo camera system over a semi-sphere spanning its workspace, in such a way to scan the workpiece surface. In each position, a pair of

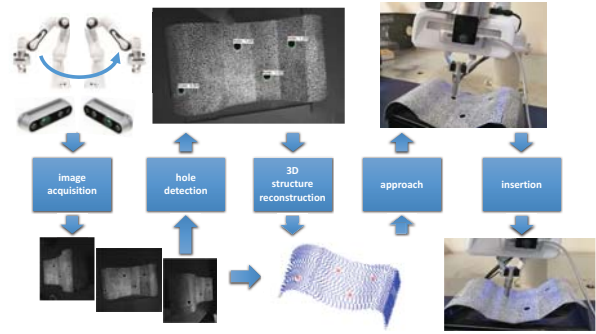


Fig. 1. The proposed peg in hole pipeline.

images is acquired with the previously calibrated hand-eye stereo-cameras system.

- The acquired images are the input to a CNN that detects the holes on the workpiece surface.
- Each pair of stereo-images from the acquired series is processed with a 3D Digital Image Correlation (DIC) approach [2] to retrieve the 3D position of a highly dense set of points on a portion of the workpiece surface; a subsequent merging operation allows the transformation of each reconstructed point cloud from the local coordinate system of the master camera to a common global reference system. Stereo-triangulation is used to determine the 3D position of the centers of the holes. Finally, the distribution of the local normal vector is computed over the tessellated surface reconstructed with stereo-DIC.
- The robot moves the peg close to the hole and aligns the approach unit vector with the hole axis (*approach* phase). Then, the peg is inserted in the hole by moving the robot under an admittance control strategy, so as to confer the suitable compliance to the peg (*insertion* phase).

Fig. 1 shows the above described pipeline.

A. Holes detection

The YOLOv3 [3] supervised approach has been chosen to detect the holes on the surface of the workpiece, since it is faster than classifier-based systems but characterized by a similar accuracy. Given an image, YOLOv3 exploits a 53-layers CNN (called Darknet-53) to generate a vector of bounding boxes and class predictions in the output. To create the "hole detector", 175 images of size 640×480 captured at different distances from the holes, have been annotated using the LabelImg tool. The training set was composed by 108

[†]School of Engineering, ^{*}Department of Mathematics, Computer Science, and Economics, University of Basilicata, 85100 Potenza, Italy.

Corresponding author: monica.sileo@unibas.it

This research has been supported by the project ICOSAF (Integrated collaborative systems for Smart Factory - ARS01_00861), funded by MIUR under PON R&I 2014-2020.

images while the remaining 67 were adopted as test set. Holes are detected into each image pair, creating a binary mask to facilitate the subsequent greyscale thresholding and centroids computation.

B. 3D surface reconstruction

The workpiece surface is provided with a stochastic black speckle pattern on a white background, thus allowing the implementation of an area-based image registration with a subset-based DIC approach. The sensor coordinates of the pairs of corresponding image points are hence used to reconstruct the position of the 3D workpiece point via triangulation. For each position of the scanning sequence, a point cloud of a portion of the workpiece surface is reconstructed in the reference system of the master camera. DIC is then used to match overlapping portions of the regions of interest (ROIs) in image pairs from contiguous hand positions. Finally, the rigid transformation that overlaps with the minimum distance the corresponding points data in two contiguous point clouds is found through non-linear optimization. These transformations are used to move and merge the point clouds into a unique reference system. The merging error, defined as the average Euclidean distance between corresponding points from four contiguous views, is 0.26 ± 0.18 mm. From the set of points measured with DIC, a triangular mesh was automatically built via the Delaunay tessellation algorithm. A plane was calculated for each triplet of points of the mesh, thus allowing to easily retrieve the distribution of the local normal vector over the whole reconstructed surface with the same spatial resolution of the DIC points grid reconstruction (about 2 mm spacing). This procedure allows to compute the positions of the hole centers and their normal unit vector (see Fig.1), in the reference frame of the master camera in the first position of the sequence.

C. Peg insertion

In order to approach the hole, the robot motion is planned to move the peg tip in the neighborhood of the workpiece surface aligning the peg and hole axes. To this aim, a closed-loop inverse kinematics algorithm with task priority has been implemented, where two tasks are assigned: *peg alignment* and *position tracking*.

Once the approach phase has been completed, the insertion is performed by keeping the orientation constant. In order to limit the mechanical stresses and ensure suitable compliance to the peg, an admittance control has been implemented at the peg tip level. It is assumed that the robot manipulator is not equipped with a wrist-mounted force/torque sensor, but, as usual in collaborative robots, with torque sensors at the joints. Thus, the wrench acting on the peg tip is estimated via an observer based on the generalized momenta [4].

III. EXPERIMENTAL RESULTS

The above strategy has been experimentally validated by adopting the robot manipulator Franka Emika Panda equipped with an Intel Realsense D435 camera in an eye-in-hand configuration. The camera holds two stereo infrared imagers,

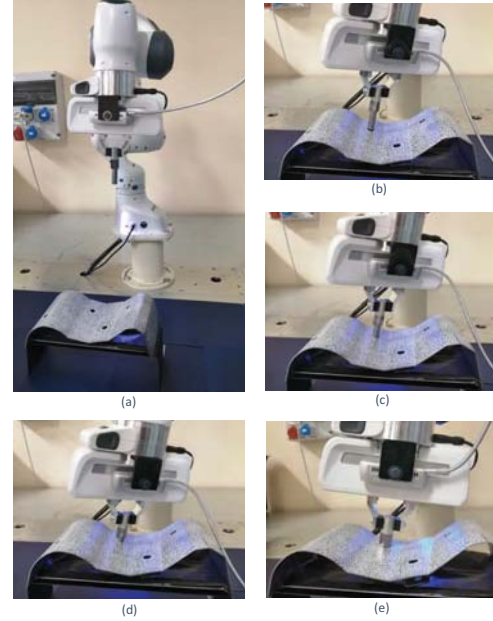


Fig. 2. Snapshots of the Peg-in-Hole assembly task: (a) robot initial pose; (b) approach phase; (c) beginning of the insertion phase; (d) insertion phase at the maximum contact wrench; (e) end of the insertion.

an RGB module and an infrared projector. The assigned task requires the mechanical coupling between a peg with a diameter of 12.2 mm and a hole of diameter of 12.4 mm and unknown tilt. The peg is symmetric and rigidly grasped by the robot gripper.

In order to perform the 3D surface reconstruction, 4 pairs of images have been taken by the stereo infrared imagers of the Realsense camera. Then, hole detection is performed on each image.

Fig. 2 reports some snapshots taken during an experimental test. Fig. 2a shows the random selected robot initial pose; in Fig. 2b, the robot reaches the target point at the end of the approach phase; in Figs. 2c and 2d, the poses corresponding, respectively, to the beginning of the insertion phase and the maximum contact wrench are reported; finally, Fig. 2e shows the end of the insertion phase. During the peg insertion the admittance control allows to keep bounded the wrench acting on the tool tip. Moreover, the experienced error between the actual hole tilt and the estimated one was negligible.

REFERENCES

- [1] F. Caccavale, C. Natale, B. Siciliano, and L. Villani, "Control of two industrial robots for parts mating," in *Proceedings of the 1998 IEEE International Conference on Control Applications (Cat. No. 98CH36104)*, vol. 1. IEEE, 1998, pp. 562–566.
- [2] M. A. Sutton, J. J. Orteu, and H. Schreier, *Image correlation for shape, motion and deformation measurements: basic concepts, theory and applications*. Springer Science & Business Media, 2009.
- [3] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv*, 2018.
- [4] F. Ficuciello, A. Romano, L. Villani, and B. Siciliano, "Cartesian impedance control of redundant manipulators for human-robot co-manipulation," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2014, pp. 2120–2125.