Bin picking pneumatic-mechanical gripper for industrial manipulators

Vladislav Ivanov
Dept. for Machine Elements and Non-Metallic Constructions
Technical University of Sofia
Sofia, Bulgaria
vvi@tu-sofia.bg

Angel Aleksandrov
Dept. for Machine Elements and Non-Metallic Constructions
Technical University of Sofia
Sofia, Bulgaria
alexandrov.angel1@gmail.com

Mohamad Bdiwi
Robot System Department
line 3: *Fraunhofer Institute for Machine Tools and Forming Technology*
Chemnitz, Germany
Mohamad.Bdiwi@iwu.fraunhofer.de

Aleksander Popov
Technical University of Sofia
Sofia, Bulgaria
dextrite@protonmail.com

Aquib Rashid
Robot System Department
line 3: *Fraunhofer Institute for Machine Tools and Forming Technology*
Chemnitz, Germany
aquib.rashid@iwu.fraunhofer.de

Zhanna Pershina
Falulty of Automation, Novosibirsk State Technical Universirty
Russia
pershina@tiger.cs.nstu.ru

Aleksey Kolker
Falulty of Automation, Novosibirsk State Technical Universirty
Russia
a.kolker@corp.nstu.ru

Lubomir Dimitrov
Dept. for Machine Elements and Non-Metallic Constructions
Technical University of Sofia
Sofia, Bulgaria
lubomir_dimitrov@tu-sofia.bg

***Abstract* — The human body is an advanced and complex machine. A simple task like grasping an object, which is done intuitively, has puzzled researchers for decades. For each object, humans have a suitable grip to grab it optimally and safely. Even for objects which we have never seen before, we can choose the best possible grip and can adapt as the environment changes to get a better grip based on cognition and experience. One of the main capabilities of humans is their ability to learn about unfamiliar things and processes. This capability helps us adapt to different situations and still be able to solve a problem.**

This works focuses on developing a self-learning robotic system which can replicate the human learning capabilities in a handing-over task. The proposed system consists of two sub-modules:

1) Vision analysis and environment monitoring, which provides accurate global and local information about the area in which the robot has to hand over the specific object;

2) Safe and flexible bin-picking gripper, which handles various objects with complex geometries.

The work is a consortium of 3 partners: IWU Chemnitz (Germany), Novosibirsk State Technical University (Russia) and Technical University of Sofia (Bulgaria) under European program ERA.Net RUS PLUS 2017

Keywords — gripper, bin picking, vision analysis

I. Introduction

There are many gripper producers and many types of grippers for different applications [1]. The goal of this proposal is to design a gripper which can detect an object, identify it, approach it even in a relatively difficult to grasp place (e.g. in a box corner), take it without destroying it, and carry it to another place.

Robots are supposed to perform a task, which incorporates picking random or semi-ordered stacked objects in a predefined task relevant position and pose, which requires higher mechanical integrity. The process of bin picking includes picking randomly placed, jumbled parts from a bin and placing them into a specific, predefined position for further manipulation or introducing them directly into machines. The worst-case scenario is when the objects to be grasped are placed near the walls or in the corners of the bin. This situation requires a big grasping flexibility to avoid collision between gripper and bin on the one hand and to allow precise and quick extraction/pull of a single object out of a cluster of the same or different objects on the other hand. Although this is an easy task for humans it appears as one of the hardest tasks in the world of robotics since it sets very high requirements towards almost every aspect of robotic manipulation, including machine vision, deep learning, flexible grasping, grasp synthesis.

Currently, bin picking is able to be fully automated only with a huge system integration project that requires multiple advanced technologies to work together. These include:

- A 3D model of the part, the bin, the robot end effector, the placement target, and any environmental obstacles
- A model of one or more ways to pick up the part with the end effector and deposit it at the placement target
- A 3D sensor to map the bin
- Image analysis software to locate each part and potential obstacles in the bin
- Path planning software to find a collision-free route from the part’s pick-up point to the placement target
- Robot control software to maneuver the robot, end effector, and part along the route

There are commercial bin picking systems that include some of these components and address a subset of bin picking challenges. Usually these systems combine a 3D sensor with image analysis software that runs on a separate computer. A robotics expert is expected to integrate the sensor, computer, software, and robot controller, and then write a program to retrieve the location of each part and figure out how to get it to the placement target [2]. Some authors suggest simpler solutions [3,4,5].

The essence of such a solution [3] is instead of trying to detect objects when they are piled up on top of each other, move them to a place where they can be more easily detected by a normal, 2D vision sensor. The proposed solution works in the following way:

1. Use the robot gripper to grab a "handful" of the objects that you want to pick. There's no need to detect the objects for this. Simply move the gripper into the box and grasp.

2. Drop the objects onto a flat surface.

3. Use the robot vision sensor to detect individual objects on the surface.

4. Pick up each object one by one.

Another simple solution is to shuffle the bin until the objects are in better position. In [4] the authors propose a combination of suction tool and a two fingers parallel gripper. A limitation of that design is that the gripper relies mainly on the suction cup and the parallel gripper is only a complimentary system with a limited grasping capability with respect to the dimensions, weight and geometry of the objects to be grasped. Another disadvantage is that both systems can’t work together.

In [5] authors have designed a dual arm – a combination between a suction tool with one finger. Again here the finger is only a complimentary system and the gripper relies mostly on the vacuum system which is not 100% reliable in industrial environment where the objects could be contaminated with oil etc.

Both of the above mentioned solutions, which were developed as a part of the Amazon Picking Challenge, do not solve the problem with extracting or pulling an object out of a cluster of the same objects (like several cups put into each other). The limited load capacity of the finger system in both cases is a major disadvantage. The overall load capacity is low in both designs.

In general, the bin picking problem stays still not fully solved due to its high complexity, despite of the efforts of multiple research teams.

II. Gripper description

The bin picking pneumatic-mechanical gripper is shown in fig.1. Generally, it consists of the following systems: mechanical part, vacuum system, electrical part, and controlling block.

The gripper meets the following requirements:

- Load capacity of at least 5 kg;
- Ability to grasp and safely move objects made of different materials and having complex geometrical shape, including convex and concave surfaces, holes, and dynamic center of gravity;
- Ability to extract a single object out of a group of randomly placed or precisely ordered objects of the same type or different ones;
- Ability to grasp an object that is situated in the corner of the bin or near its walls;
- Ability to ensure human safety by inherent design and global perception system;

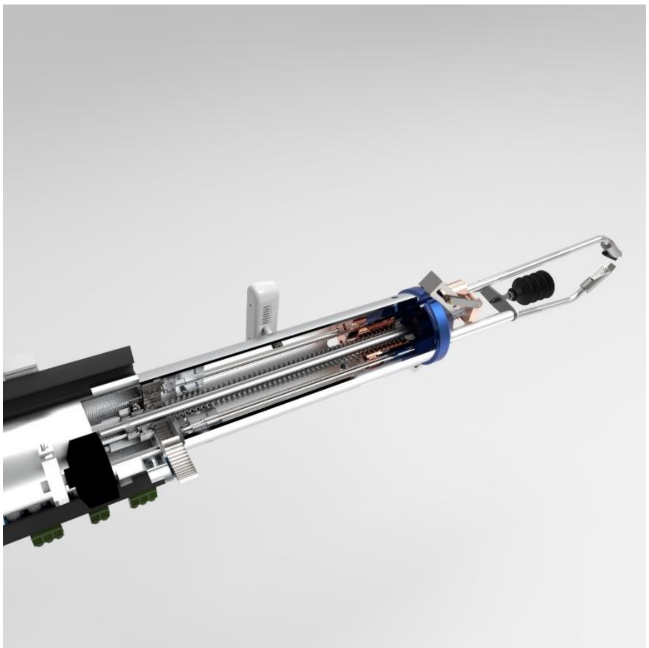

Fig. 1. Bin picking pneumatic-mechanical gripper.

- Compatibility with industrial robots (2900 mm of range);
- Compatibility with the machine vision system which is to develop.

III. Functionality

The gripper has two basic systems – vacuum and mechanical fingers that can work together or separately, depending on the grasping strategy chosen with respect to the geometry of the object that to be grasped.

An example scenario is pulling out of a single part from the edge of a bin or from a tight stack of parts and moving the object grasped at points to vacuum for a stable three point grasp.

The vacuum system is equipped with a smart vacuum ejector and a suction cup that can grasp objects with convex and concave surfaces. Its load capacity is between 0,5 and 2,0 kg depending on the diameter of the cup.

The mechanical system consists of two fingers that, at the same time, execute translational and rotational motion. That design solution allows them to go behind the suction cup when it is necessary or to be extended around 50 mm in front of the suction cup. The motion is realized through power screws which are connected to servo motors with high precision encoders.

The fingertips replicate the human fingertips, thus ensuring a big grasping flexibility and manipulation ability (fig.2, fig.3.). They allow for the extraction of a single object out of a stack of objects of the same type or different like if we are using tweezers. Once pulled out of the cluster, the gripper can use both systems to assure a firm and reliable grip. The fingertips design also allows for manipulation, before robust grasp with finger surfaces and vacuum) grasping metal sheets or objects with convex or concave shape made of metal sheet without using the suction cup.

Fig. 2. Gropper grasping different objects: stone, rolling element bearing, bushing, pen.

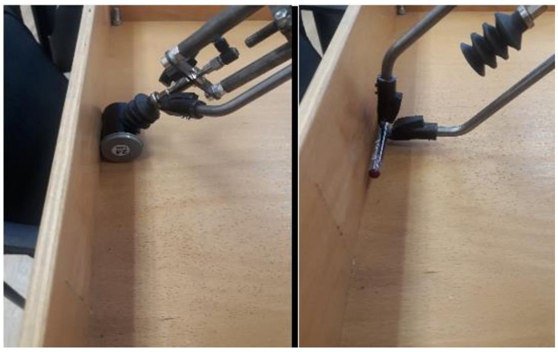

Fig. 3. Gropper grasping different objects from the box corner.

The proposed system (named ICU) has three main subsystems:

- Dual servo controller;
- Vacuum ejector (fig.4). The vacuum subsystem is based around a fully integrated onboard smart ejector. It is controlled by multiple binary inputs and outputs from the main controller;
- Main control board. The main controller is an arm based single board computer running Ubuntu 18.04 LTS. The main code is written in python with all of the computationally expensive tasks, such as the kinematic model, running on optimised C++ based custom libraries. The connection to the master computer is achieved using ROS.

The only required connections to make it run, is a 120A 12V power input, regulated air supply and an ethernet connection for communication.

The movement of the fingers is achieved using two identical force air cooled 5065 BLDC motors each with a 8192 CPR capacitive multiturn encoder directly coupled to them. They are electrically connected to an onboard high performance motor controller, which essentially turns them into servo motors. The choice of motors is governed by the requirement for extreme compactness while maintaining high performance. This whole subsystem runs on the main 12v input bus at which the maximal no-load speed is 3200 RPM. With the current configuration both motors can run at peak current of 65 Amps safely for up to a minute at which they produce around 2 N.m. of torque. This current is achieved only when running at maximal acceleration and grip strength. In a real world environment, the time spent at max load is no more than a second each grasping and releasing cycle by utilising the self-locking nature of the drive screws, thus overheating is unlikely even with high speed processes. The communication with the main controller is achieved using a usb interface which allows for low latency and high frequency operation.

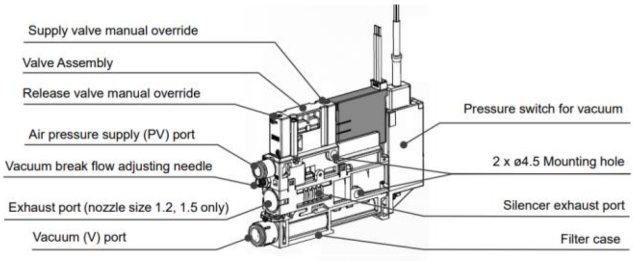

Fig. 4. SMC ZK2A10K5RW-06 Ejector System with Valve – (EJ).

IV. Vision Analysis and Environment Monitoring

A new technology for evaluating 6DoF objects in the workspace of a robot manipulator has been developed to decide on the optimal grasping of the desired object and moving it. The proposed technology is based on the analysis of the observed scene data obtained from the RGBD sensor, which is mounted on the capture device of the robot manipulator.

The solution to the problem of determining the orientation of objects in three-dimensional space is largely based on the approach proposed in the work of Bertram Drost and Slobodan Ilic. [8], and which still remains effective in this area. This approach is based on the comparison of the existing CAD-model of the object with a 3D cloud of points on the features of the singular points. However, the effective application of this approach is possible only on scenes with a dense cloud of points, which requires a large computational resources.

In order to optimize computing resources, we propose an initial preprocessing step, before standard method from Bertram Drost and Slobodan Ilic [8], perform 3D semantic segmentation of objects in the scene, thereby performing remaining computations on lesser data, which is of interest. This reduces the size of the 3D point cloud, which in turn leads to an increase in the performance of the algorithm for evaluating 6DoF objects. Thus, the processing method technology produced is based on the developed algorithm for the semantic segmentation of three-dimensional scenes.

Semantic segmentation of objects in images is implemented on the basis of a convolutional neural network. MaskRCNN [9] is used as a convolutional neural network designed for semantic segmentation in this research. The software environment for training the development model consists the TensorFlow + Keras neural network library. This requires a description of the architecture of the convolutional neural network in terms of a sequence of connected layers and their parameters, as well as a description of the parameters of the learning algorithm.

The semantic segmentation of objects in the image is performed by selecting the class with the highest probability for each pixel of the processed image, examples of the results of semantic segmentation are shown in fig. 5.

Fig 5. Semantic segmentation results example

Fig. 6 demonstrates the overlapping model points (red) and image points (gray) obtained from RGBD sensor.

Fig.6. 6dof estimation result

The key challenge in the deep learning based semantic segmentation model training is the requirement of a large amount of annotated dataset. It requires preliminary annotation of objects in the image in the form of segments (polygons or masks), which should be assigned to a class and ID of the object in the scene. For solving this problem, we developed a simulation environment, which generates automatic labeled data in different poses, lighting conditions and scenarios using raytracing.

The raytracing is a process of modeling of a real physical process of light reflection and consumption. The approach allows the generation of realistic images. These images represent a high quality training database. Moreover, it is created only on artificially obtained images.

For the creation of three-dimensional graphics, we use the Persistence of Vision Raytracer [10], which is high-quality, free software. On the other site, a Python based tool realizes a special set of light sources, structure of material and physically based (physical engine) occlusion of objects. This synthetic system is trained and is able to perform semantic segmentation with metric Interception over a Union (IoU) [11] IoU 0.55, which is comparable to an actual trained data model (IoU= 0.60).

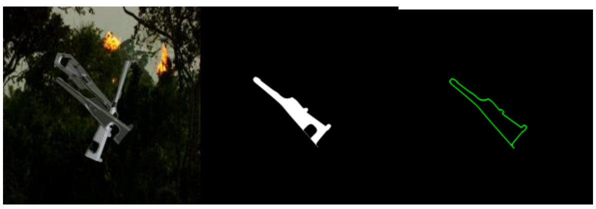

Fig 7. Example training dataset item

The accuracy of the 6DoF estimation (with an accuracy of +/- 5 degrees) during the preliminary experiment is 0.75.

V. Conclusions

As a result of the work done, a bin picking pneumatic-mechanical gripper has been developed. The gripper configuration allows to detect the object, identify it, grasp it even from a relatively difficult to grasp place, take it without destroying it, and carry it to another place. The gripper can be used for industrial robots for grasping different objects with complicated shape. The gripper allows for picking the object by vacuum or by fingers or by both.

Acknowledgment

This work has been accomplished with financial support by National Ministry of Education and Science of Bulgaria and the European program ERA.Net RUS PLUS 2017, under contract Grant No DO 02/3 “Self-Learning Robot Assistant System (ICU)“.

References

[1] Ggrippers market report leading companies. Forecast2022, July 12, 2018 [Online]. Available http://www.gmoutlook.com/grippers-market-report-leading-companies-like-company-name-competitive-landscape-research-forecast-2022-53286.html

[2] Truebenbach E., “Is fully automated bin picking finally here”, February 18, 2019. [Online]. Available: https://www.therobotreport.com/fully-automated-bin-picking-finally-here/.

[3] Owen-Hill A., “Bin picking the easy way vs the hard way with robot vision”, March 21, 2019. [Online]. Available: https://blog.robotiq.com/bin-picking-the-easy-way-vs-the-hard-way-with-robot-vision.

[4] Schwarz M. et al., “Fast object learning and dual-arm coordination for cluttered stowing, picking, and packing”, IEEE International Conference on Robotics and Automation (ICRA), Brisbane, Australia, May 2018.

[5] Wade-McCue S. et al., “Design of a multi-modal end-effector and grasping system”, Amazon Robotics Challenge, 2017, [Online]. Available: https://arxiv.org/pdf/1710.01439.pdf.

[6] Tomov P., “Increasing the Efficiency of Automation of Production Processes by Reporting the Parameters of the Parts’ Flow”, TEM Journal, 2017, 6(3), ISSN: 2217-8309 DOI: 10.18421/TEM63-08 (484-487).

[7] Zhmud V., et al., “Application of ultrasonic sensor for measuring distances in robotics”, Journal of Physics: Conference Series (Open access), 2018, 1015(3),032189

[8] Drost, Bertram, and Slobodan Ilic. "3d object detection and localization using multimodal point pair features." *2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission*. IEEE, 2012.

[9] He, Kaiming, et al. "Mask r-cnn." Proceedings of the IEEE international conference on computer vision. 2017.

[10] Persistence of Vision Raytracer, http://www.povray.org , May 21, 2020, [online]

[11] Garcia-Garcia A. et al. A review on deep learning techniques applied to semantic segmentation //arXiv preprint arXiv:1704.06857. – 2017.