Some Placeholder Title

A Master Thesis

written by

Victor Melbye Staven
vista17@student.sdu.dk

The code for this project is available at
`https://github.com/vmstavens/in_hand_pose_estimation`

University of Southern Denmark
The Technical Faculty

Word Count : 962
October 28, 2022

Abstract

Some abstract text explaining the goal, methods and conclusion of the project.

Contents

Acknowledgements

My acknowledgements

Acronyms

cobots collaborative robots.
CP correspondence problem.
CV computer vision.
DL deep learning.
DOF degrees of freedom.
EE end effector.
PE pose estimation.
PNP pick-and-place.
SOTA state of the art.

Terms

collaborative robots (cobots) [some description].

computer vision (CV) is a scientific field which deals with how computers can gain a high-level understanding from digital inputs in the form of images or videos.

correspondence problem (CP) the problem within computer vision (CV) to localize the projection of the same object in 3D space in 2 or more images.

deep learning (DL) is a technique for data driven pattern recognition.

end effector (EE) [some description].

manipulator , [a manipulator description].

pick-and-place (PNP) [some description].

pose estimation (PE) is a scientific picipline which deals with how computers can gain a high-level understanding from digital inputs in the form of images or videos.

Chapter 1

Introduction

1.1 Context

As of 2022 most of the industrialized world has developed tools for unprecedented growth in wealth and technology on a global scale [1, Chapter 4]. In such times a great deal of consumerism and interconnection is present with people needing products produced faster and more consistently than ever before [1, Chapter 4]. As one would expect, this creates a high demand for manufacturers to reliably and consistently provide products, while also remaining flexible as the demand for different product change rapidly. In order to accommodate the need for ever greater volumes of products, consistent, reliable and flexible labor is essential in assembly, transport and manipulation processes in the production pipeline. Due to these types of manual labor being largely done by unskilled workers, automation alternatives are being adopted which provides benefits [1, Chapter 4]. This different approach to manufacturing has been labeled the fourth industrial revolution or i4.0 for short. The beneficiaries being the employer and employee, with the employer having the benefits: Avoid paying monthly salaries to unskilled laborers doing manual tasks, here the automation alternative only requires electrical energy and potential supervision by few qualified individuals. Potential risks are also involved when hiring humans as the workforce can be inconsistent due to human error [2] or left out due to illness etc. Considerations with regards to workers rights such as working conditions and wage also needs not to be considered. Workers furthermore cause production limitations in the form of stand still hours, such as bathroom and lunch breaks along with after work hours and holidays. This replacement of manual labor also potentially benefits the employee, as boring and physically wearing work is automated, enabling the employees to take on different and less wearing and potentially dangerous roles. While the issue of labor unemployment becomes apparent solutions which provide support to already hired workers have been developed, such as collaborative robots (cobots)[3] which would negate this problem.

When implementing automation of production lines using robotics, certain categories of problems are revealed. These include: Assembly, alteration and pick-and-place (PNP), the last being the one of interest in this project.

1.2 Problem Description

Pick-and-place manipulators are used in a wide variety of different fields such as sorting of waste [4] handling of food [5][6] and factory bin picking [7] [8] [9]. The solutions in these industries are examples of subcategories under the PNP problem, namely sorting and bin picking. Since both of these are subcategories of the PNP problem, they fundamentally follow the same sequential four phases from start to end. These phases being pre-grasping, grasping, transport, and placement [8] for traditional implementations of the PNP pipeline. The pre-grasp phase involves localizing the object(s), potentially estimating their pose and executing the trajectory to move the end effector (EE) grasp, collision free to said object(s). Here different potential grasps can be considered in order to determine the best pose for the EE. In the grasping phase the EE gasps the object in such a manner that the object's entire weight is supported by the EE, and ends when the object no longer is in contact with the environment, which often is the container holding the object. The transportation phase involves the motion of the manipulator to move from the pose achieved after the grasping phase, to a pose ready for placement of the object in the desired placing area or fixture. Here considerations may be needed with regards to how much force and torque the EE's grasp can tolerate while moving without losing the object. Finally the goal of the placing phase is to place the object within the placing area or fixture in a desired end pose. Here the constraints on the end pose might differ significantly based on the application, as the pose of greens in a crate might need less precision than if the manipulator hands a bolt to the another robotics system in the pipeline.

While these phases make up a traditional PNP systems, certain assumptions are made regarding the objects of interest in order for this pipeline to function. Specifically the localization and pose estimation (PE) of the pre-grasp phase are assumed possible due to either ensured object poses or estimated poses through computer vision (CV) sensory system. Due to CV being a mature research field a wide range of solution proposals to these problems have been generated [10]. These include classic vision [11] [12], deep learning based [13] and combinations of these [14]. However, while these may be sufficient for certain tasks they fundamentally suffer from the weaknesses introduced by vision techniques. These being a great number of outliers caused by: occlusions, reflecting, transparent or homogeneous surfaces, and repetitive structures when solving the correspondence problem (CP). Within factory settings the common ones are transparent and reflective objects, due to metallic, plastic and glass product often being the materials used. While deep learning (DL) solutions have been developed for both for reflective [15] and transparent [16] objects, these are use case specific and show limited results in a wider range of applications.

This project suggests a different PNP pipeline for cases where the object's starting pose is unknown. In this PNP pipeline the PE is moved from the pre-grasping phase to a new phase between the grasping and transportation phase, called the PE phase. The specific goal of this project is to develope a solution to this phase using tactile sensors in the EE to determine the object's pose. By using the tactile sensing rather than visual the problems presented above will be eliminated. This will be done using a humanoid gripper as the EE with tactile sensors in each finger, more specifically a Shadow Dexterous Hand [17] with 20 degrees of freedom (DOF).

In depth explanation of traditional vs. suggested manipulator pipeline

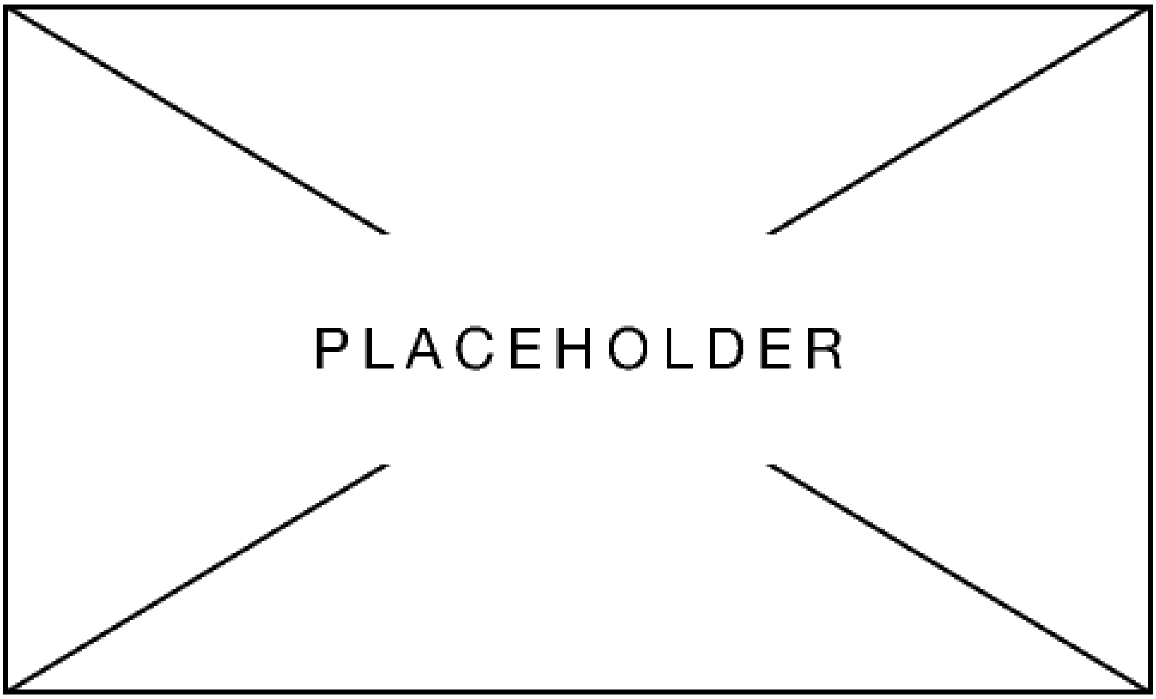

Fig. 1: Cartoon of suggested pnp vs. traditional pnp.

In order to solve this PE problem, three sub problems are identified and labeled problem 1, 2 and 3. Problem 1 involves modeling the contact between the gripper's tactile sensors and the object, also referred to as tactile perception. Problem 2 is to convert the collected data from problem 1 to estimated pose candidates. Finally problem 3 involves in-hand manipulation. Since the initial grasp of the object might not be oriented in a manner where the recognizable features make context with the tactile sensors, manipulating the object within the EE's grasp will enable further information gathering. Thus the final problem is to control the EE in such a manner that the tactile sensors make context with the object in intelligently decided areas for a better pose estimate.

To test if the developed system successfully solves the PE problem, it is hypothesized that the intelligent probing method provides a statistically significant faster average PE convergence, along with a statistically significant greater success rate when determining the correct pose. A correct pose is here defined as the pose being greater than or equal to 95% of the ground truth pose, and statistically significant is defined by an α-level of 95%. This hypothesis will be referred to as H_1, while the null hypothesis H_0 being that there is no statistically significant difference between intelligent and random probing's PE performance as described above.

1.3 Thesis Overview

To present the work done in this project, the system modeling is done in Chapter 2 and state of the art (SOTA) is presented in Chapter 3 for each of the problems presented above. Here the solutions best suited for this project's gripper are chosen. Each solution is described in detail, how they are applied, their performance tested and finally evaluated and conclude upon in their respective chapters i.e. chapter Chapter 4, Chapter 5 and Chapter 6. In Chapter 7 the three methods are combined in the final integration and finally the project is discussed and concluded upon in Chapter 8 and Chapter 9 respectively.

Chapter 2

Modeling

In order to choose the methods best suited for solving the problems presented in Chapter 1 a representative model of the system is needed. This model must thus contain all the physical phenomena which is needed to facilitate methods for solving the presented problems. These phenomena are contact surface pressure responses from the object's texture and topology to the manipulator and frictions between the object and manipulator. Modeling these is done by applying contact and friction models respectively.

To illustrate a single point contact model between the EE's finger and the object, a normal force $\mathbf{f}_N$ at contact point $\mathbf{c}_1$ can be seen in Fig. 1. Once a greater force is applied by the EE's finger the point contact expands to a contact surface which due to the finger's surface being compliant and said finger's convex shape a pressure distribution $p(r)$ is created. The pressure distribution is here a function of the contact surface radius r assuming a non-conforming contact case. This can be seen illustrated in Fig. 2

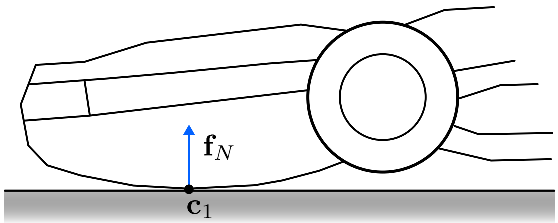

Fig. 1: Single point contact model between compliant manipulator surface and object surface.

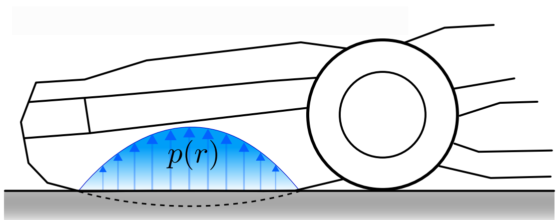

Fig. 2: Contact model of the pressure distribution $p(r)$ caused by the increased force applied from the EE's finger to the object.

An illustration of a point contact friction model can be seen in figure Fig. 3, where $\mathbf{c}_1$ is the point of contact, $\mathbf{f}_N$ is the normal force acting on the EE's finger in $\mathbf{c}_1$, $\mathbf{f}_f$ is the friction force reacting to the external force acting on the object $\mathbf{f}_{\text{ext}}$. The friction force will here be proportional to the normal force applied by the EE's finger.

When the EE applies a greater force on the object the normal force will likewise increase and thus the friction force will increase. This model considers the finger to be compliant and thus a soft finger friction model can be utilized. The soft finger friction model can be seen in Fig. 4

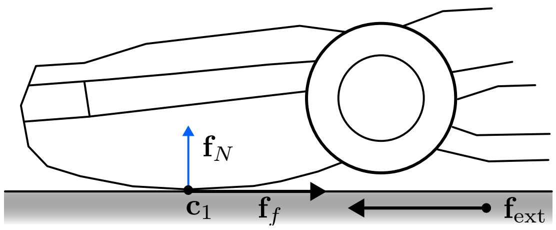

Fig. 3: Single point contact model between compliant manipulator surface and object surface.

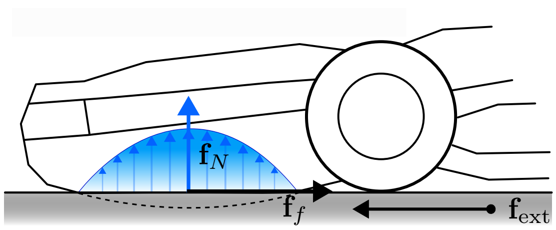

Fig. 4: Friction model of the pressure distribution $p(r)$ caused by the increased force applied from the EE's finger to the object.

During in-hand manipulation the presence of external forces is a common problem and often comes in the form of gravity acting on the object, which can cause the problem to slip. By combining the contact and friction models as described above grasps and in-hand manipulations can be performed by compensating for the gravitational pull with applied force. This can be seen illustrated in Fig. 5.

By describing the kinematic tree of the EE, the

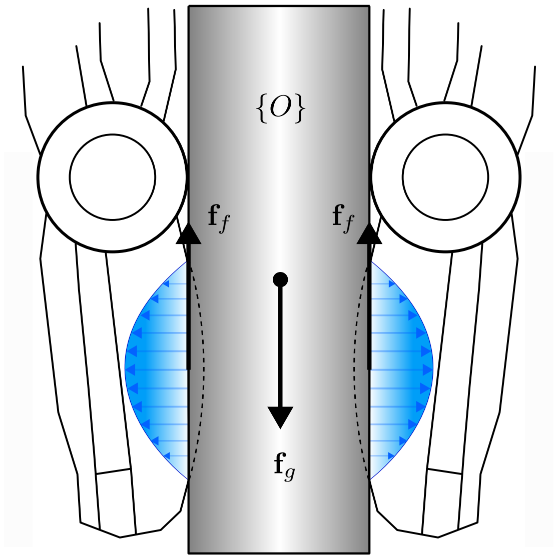

Fig. 5: The friction forces $\mathbf{f}_f$ and contact model keeps the object $\{O\}$ from slipping between the EE's fingers when the gravitational pull $\mathbf{f}_g$ is acting on it.

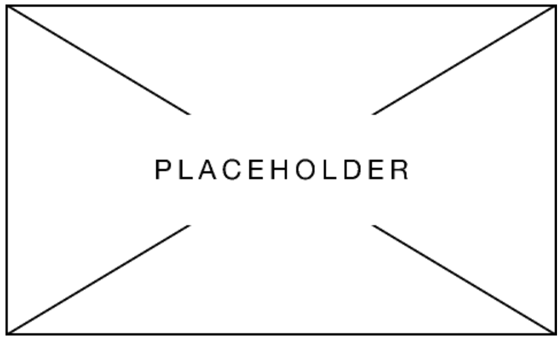

Fig. 6: The model of the world representation for this project.

To determine which methods best describe the models presented above for this project, the SOTA will be presented in Chapter 3.

Chapter 3

State of the Art

3.1 Problem 1 - Tactile Perception

Based on the contact model described in Chapter 2
For tactile perception it is needed to model the contact between the robotic manipulator finger and the object. The model will here
When considering different methods for modeling contact interfaces, each can be categorized depending on which parameters the model describe the relation between. This results in the following groupings: contact-area-force models, stress-strain models, force-displacement models and
Love's formulation
What is tactile perception? Why is it relevant?
How is a tactile sensor constructed [18] what different types exist and which one is present in the model provided.
"Representations of tactile data are commonly either inspired by machine vision feature descriptors"
often used in computer vision context, where each tactile image
Addressing the problem

3.2 Problem 2 - Pose Estimation

3.3 Problem 3 - In-Hand Manipulation

Chapter 4

Tactile Perception

4.1 Introduction

Here we write the introduction for problem 1.

4.2 Related Work

Here we cite the related work by `\cite{source-label}` like this [18]

Chapter 5

Pose Estimation

5.1 Introduction

Here we write the introduction for problem 2.

5.2 Related Work

Here we cite the related work by `\cite{source-label}` like this [18]

Chapter 6

In-Hand Manipulation

6.1 Introduction

Here we write the introduction for problem 3.

6.2 Related Work

Here we cite the related work by `\cite{source-label}` like this [18]

Chapter 7

System Integration

7.1 Introduction

Here we write the introduction for the system integration.

Chapter 8

Discussion

Chapter 9

Conclusion

Bibliography

[1] Publications, United Nations. “Inequality in Asia and the Pacific in the Era of the 2030 Agenda for Sustainable Development”. In: 7May 2018. URL: https://www.unescap.org/publications/inequality-asia-and-pacific-era-2030-agenda-sustainable-development.

[2] Wenwen, Shi et al. “Analysis and Control of Human Error”. In: *Procedia Engineering* 26 (Dec. 2011), pp. 2126–2132. DOI: 10.1016/j.proeng.2011.11.2415.

[3] Galin, Rinat et al. “Cobots and the benefits of their implementation in intelligent manufacturing”. In: *IOP Conference Series: Materials Science and Engineering* 862 (May 2020), p. 032075. DOI: 10.1088/1757-899X/862/3/032075.

[4] Raptopoulos, Fredy, Koskinopoulou, Maria, and Maniadakis, Michail. “Robotic Pick-and-Toss Facilitates Urban Waste Sorting”. In: *2020 IEEE 16th International Conference on Automation Science and Engineering (CASE)*. 2020, pp. 1149–1154. DOI: 10.1109/CASE48305.2020.9216746.

[5] Talpur, Mir Sajjad Hussain and Shaikh, Murtaza Hussain. *Automation of Mobile Pick and Place Robotic System for Small Food Industry*. 2012. DOI: 10.48550/ARXIV.1203.4475. URL: https://arxiv.org/abs/1203.4475.

[6] Yamanaka, Yuta et al. “Development of a Food Handling Soft Robot Hand Considering a High-speed Pick-and-place Task”. In: *2020 IEEE/SICE International Symposium on System Integration (SII)*. 2020, pp. 87–92. DOI: 10.1109/SII46433.2020.9026282.

[7] Lee, Sukhan and Lee, Yeonho. “Real-Time Industrial Bin-Picking with a Hybrid Deep Learning-Engineering Approach”. In: *2020 IEEE International Conference on Big Data and Smart Computing (BigComp)*. 2020, pp. 584–588. DOI: 10.1109/BigComp48618.2020.00015.

[8] Mnyusiwalla, Hussein et al. “A Bin-Picking Benchmark for Systematic Evaluation of Robotic Pick-and-Place Systems”. In: *IEEE Robotics and Automation Letters* 5.2 (2020), pp. 1389–1396. DOI: 10.1109/LRA.2020.2965076.

[9] Wong, Ching-Chang et al. “Generic Development of Bin Pick-and-Place System Based on Robot Operating System”. In: *IEEE Access* 10 (2022), pp. 65257–65270. DOI: 10.1109/ACCESS.2022.3182114.

[10] He, Zaixing et al. “6D Pose Estimation of Objects: Recent Technologies and Challenges”. In: *Applied Sciences* 11.1 (2021). ISSN: 2076-3417. DOI: 10.3390/app11010228. URL: https://www.mdpi.com/2076-3417/11/1/228.

[11] ., Taryudi and Wang, Ming-Shyan. “3D object pose estimation using stereo vision for object manipulation system”. In: May 2017, pp. 1532–1535. DOI: 10.1109/ICASI.2017.7988217.

[12] Oh, Jong-Kyu, Lee, Sukhan, and Lee, Chan-Ho. “Stereo vision based automation for a bin-picking solution”. In: *International Journal of Control, Automation and Systems* 10.2 (Apr. 2012), pp. 362–373. ISSN: 2005-4092. DOI: 10.1007/s12555-012-0216-9. URL: https://doi.org/10.1007/s12555-012-0216-9.

[13] Abdelaal, Mahmoud et al. “Uncalibrated stereo vision with deep learning for 6-DOF pose estimation for a robot arm system”. In: *Robotics and Autonomous Systems* 145 (2021), p. 103847. ISSN: 0921-8890. DOI: https://doi.org/10.1016/j.robot.2021.103847. URL: https://www.sciencedirect.com/science/article/pii/S0921889021001329.

[14] Nakano, Yoshihiro. “Stereo Vision Based Single-Shot 6D Object Pose Estimation for Bin-Picking by a Robot Manipulator”. In: *CoRR* abs/2005.13759 (2020). arXiv: 2005.13759. URL: https://arxiv.org/abs/2005.13759.

[15] Kozák, Viktor et al. “Data-Driven Object Pose Estimation in a Practical Bin-Picking Application”. In: *Sensors* 21.18 (2021). ISSN: 1424-8220. DOI: 10.3390/s21186093. URL: https://www.mdpi.com/1424-8220/21/18/6093.

[16] Xu, Chi et al. “6DoF Pose Estimation of Transparent Object from a Single RGB-D Image”. In: *Sensors* 20.23 (2020). ISSN: 1424-8220. DOI: 10.3390/s20236790. URL: https://www.mdpi.com/1424-8220/20/23/6790.

[17] Straszheim, Troy et al. *Shadow Robot*. URL: https://www.shadowrobot.com/dexterous-hand-series/.

[18] Chi, Cheng et al. “Recent Progress in Technologies for Tactile Sensors”. In: *Sensors* 18.4 (2018). ISSN: 1424-8220. DOI: 10.3390/s18040948. URL: https://www.mdpi.com/1424-8220/18/4/948.

Appendix A

Appendix A Title