MODEL-BASED POSE ESTIMATION OF THE TOOL CENTER POINT OF COLLABORATIVE INDUSTRIAL ROBOTS (UR5) USING A PROPOSED FIDUCIAL MARKER SYSTEM

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

HAMID MAJIDI BALANJI

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
MECHANICAL ENGINEERING

SEPTEMBER 2019

Approval of the thesis:

MODEL-BASED POSE ESTIMATION OF THE TOOL CENTER POINT OF COLLABORATIVE INDUSTRIAL ROBOTS (UR5) USING A PROPOSED FIDUCIAL MARKER SYSTEM

submitted by **HAMID MAJIDI BALANJI** in partial fulfillment of the requirements for the degree of **Master of Science in Mechanical Engineering Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. M. A. Sahir Arıkan
Head of Department, **Mechanical Engineering** _____

Assist. Prof. Dr. Ali Emre Turgut
Supervisor, **Mechanical Engineering, METU** _____

Assist. Prof. Dr. Lütfi Taner Tunç
Co-Supervisor, **Manufacturing Engineering Dept., Sabancı University** _____

Examining Committee Members:

Assist. Prof. Dr. A.Buğra Koku
Mechanical Engineering Dept., METU _____

Assist. Prof. Dr. Ali Emre Turgut
Mechanical Engineering, METU _____

Prof. Dr. M. Kemal Özgören
Mechanical Engineering Dept., METU _____

Assist. Prof. Dr. Lütfi Taner Tunç
Manufacturing Engineering Dept., Sabancı University _____

Assist. Prof. Dr. Selçuk Himmetoğlu
Mechanical Engineering Dept., Hacettepe University _____

Date: 05.09.2019

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: HAMID MAJIDI BALANJI

Signature:

ABSTRACT

MODEL-BASED POSE ESTIMATION OF THE TOOL CENTER POINT OF COLLABORATIVE INDUSTRIAL ROBOTS (UR5) USING A PROPOSED FIDUCIAL MARKER SYSTEM

MAJIDI BALANJI, HAMID
Master of Science, Mechanical Engineering
Supervisor: Assist. Prof. Dr. Ali Emre Turgut
Co-Supervisor: Assist. Prof. Dr. Lütfi Taner Tunç

September 2019, 130 pages

Real-time pose estimation of Tool Center Point (TCP) of industrial robots is very important in industrial robotic applications. The TCP pose information is used in many industrial robotic tasks such as calibration and control tasks for on-line TCP pose corrections and improvement of the TCP pose accuracy and repeatability. Nowadays, in industry and robotic research laboratories, laser trackers, optical CMMs, stereo vision and photogrammetry techniques are applied for TCP pose estimation purposes; however, each of these has their own problems that affect their pose estimation performance and efficiency in industrial robots.

This study presents a novel model-based pose estimation method based on computer vision and augmented reality markers. The proposed system is able to estimate the pose of the Tool Center Point (TCP) of the industrial robotic manipulators in point-to-point applications. An innovative Rhombicuboctahedron ArUco Mapper (RAM) is designed for pose estimation of the TCP. The performance analysis of the proposed system proved its repeatability and accuracy. The absolute positional measurement accuracy of the proposed system in the robot base frame is in the range of 0.12-0.48mm and for orientational measurements in the robot base frame are in the range

of 0.003-0.012^{o}. The pose measurement accuracy and repeatability in the camera base frame show better results ±50μm-500μm for the positional accuracy and ±0.009^{o}-0.05^{o} for the orientational accuracy. The proposed method is considered as a cost-effective, high accuracy and repeatable TCP pose estimator for industrial robotic applications.

Keywords: Computer vision, Industrial Robotic Manipulator, TCP pose, Rhombicuboctahedron ArUco Mapper, Absolute pose accuracy, Absolute pose repeatability

ÖZ

Model Tanimli, UR5 Manipulator Ortasinin Posizyonunu Tahim eden, Olcum Usaretli Marker Sistemi

MAJIDI BALANJI, HAMID
Yüksek Lisans, Makina Mühendisliği
Tez Danışmanı: Doç. Dr. Ali Emre Turgut
Ortak Tez Danışmanı: Doç. Dr. Lütfi Taner Tunç

Eylül 2019, 130 sayfa

Endüstriyel robotların Alet Merkez Noktasının (AMN) konumunun ve yöneliminin kestirimi çok önemlidir. AMN poz bilgisi noktadan-noktaya endüstriyel robot uygulamalarında kullanılmaktadır. Günümüzde lazer tarayıcılar, optik CMM'ler, stereo görüntüleme ve fotogrametri teknikleri AMN poz belirlenmesi için kullanılmaktadır, ne varki bu tekniklerin her birinde poz belirlenmesi performanslarını ve verimliliklerini negative yönde etkileyen bazı eksiklikler mevcuttur.

Bu çalışma bilgisayar görü ve artırılmış gerçeklik işaretlerine dayalı yeni bir model tabanlı poz belirleme sistemi sunmaktadır. Bu çalışmada önerilen sistem nokta-nokta uygulamalarında kullanılan endüstriyel robotların (AMN) ait pozu (konum ve yönelim) belirlemek için geliştirilmiştir. Yenilikçi bir Rhombicubocathedron ArUCO Mapper AMN poz belirlenmesi amacı ile oluşturulmuş ve uygulanmıştır. Önerilen sistemin performans analizi yüksek hassasiyetli poz belirlenmesinin ve tekrarlanabilirliğin mümkün olduğunu göstermiştir. Robot taban koordinat sistemine göre mutlak konum ölçüm hassasiyeti 0.12-0.48 mm, yönelim ölçüm hassasiyeti ise 0.003-0.012^o aralıklarındadır. Kamera koordinat sistemine göre mutlak konum ölçüm hassasiyeti ±50μm-500μm, yönelim ölçüm hassasiyeti ise ±0.009^o- 0.05^o olmak üzere daha iyi sonuçlar göstermiştir. Önerilen sisteme endüstriyel robot uygulamaları için

uygun maliyetli ve yüksek hassasiyetli tekrarlanabilir ölçüm metodu olarak değerlendirilebilir.

Anahtar Kelimeler: Bilgisayar görüntüsü, Endüstrıyel robot kolu, TCP poz, Rhombicuboctahedron ArUco Mapper, Mutlak poz doğruluğu, Mutlak poz tekrarlanabilirlik

To My Parents For their endless love and support

ACKNOWLEDGEMENTS

Foremost, I offer my deepest gratitude to my supervisor, Dr. Ali Emre Turgut, for his great technical supports, precious caring, and managing the project. I hope the results of this project can convey my appreciations for his supports and guidance. It was a great honor, and scientific experience for me to work with him. I would also express my gratitude to Dr. Lütfi Taner Tunc, thesis co-advisor, for his excellent comments and points during our collaborations in this work. I learned more from him.

I also acknowledge Prof. Dr. Erol Şahin and Prof. Dr. Sinan Kalkan for their precious supports of KOVAN robotic lab researchers with providing some modern robots such as the UR5. Their valuable efforts in providing a friendly and scientific atmosphere in KOVAN Research lab are admirable. I acknowledge the contribution of Prof.Dr. M. Kemal Özgören, Dr. Buğra koku, and Dr. Selçuk Himmetoğlu for serving as the jury members.

I would also acknowledge Cem Bilaloğlu, my lab colleagues in KOVAN for his help in 3D printing tasks. Special thanks go to Tugay Alperen Karagüzel and Negin Bagherzadi, KOVAN Lab researchers, for their great ideas and thoughts that really flourished my thesis contents.

In the end, I would like to express my special gratitude to Prof. Dr. Rafael Muñoz Salinas, the leading creator and developer of the ArUco markers (AVA research lab-Cordoba university) for providing me with the latest library of the ArUco and also his priceless advises and comments concerning ArUco library installation, using and debugging.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

ABBREVIATIONS

TCP: Tool Center Point

D-H: Denavit-Hartenberg

CMMs: Coordinate Measurement Machines

RAM: Rhombicuboctahedron ArUco Mapper

FWK: Forward Kinematic

CHAPTER 1

INTRODUCTION

1.1. Background

The industrial revolution term was coined for the first time by Arnold Toynbee to portray Britain's economic evolution from 1760 to 1840 [1].

The term industrial revolution points to a period from the 1770s to 1870s [2], when technological transition happened in manufacturing processes. In this period, mechanical and electrical forces were brought under control, enabling the manufacturing of novel products and tools. Consequently, social and economic advances occurred as a result of this progress in product manufacturing techniques from agrarian-based methods to industrial-based techniques, known as the industrial revolution.

In the existing literature, the industrial revolution classifies into four eras:

1. **Industry 1.0**: this era started with the introduction of steam and water power from 1760 to 1830. Two significant inventions of this period are forging and shaping operations.
2. **Industry 2.0**: In this period, factories achieved mass production capabilities using the introduction of electrical energy into the industry. From a historical viewpoint, this era can be traced back to the years between 1870 to 1914. The telephone, the light bulb, phonograph, and internal combustion engines are some outstanding inventions of those times.
3. **Industry 3.0**: Information technology and electronics have contributed to the third industrial revolution. Some industrial historians have called this era as the "Digital Revolution." For instance, the personal computer, the internet, and

the telecommunication technology brought into reality as a result of this industrial revolution.

4. **Industry 4.0**: This era includes the application of the Cyber-physical systems, the internet of things (IoT), data exchange, cloud computing, and cognitive computing in manufacturing technologies. In a short expression, digitalization of the manufacturing procedures and fusion of the production using artificial intelligence, information, and communication technologies are some main elements of industry 4.0. In fact, the idea of "smart factories" was promoted by industry 4.0. Three-dimensional printing, robotic applications in the manufacturing process such as product assembly, welding, and machining are some examples of this industrial era. Figure 1.1 outlines the development stages of the industrial revolution.

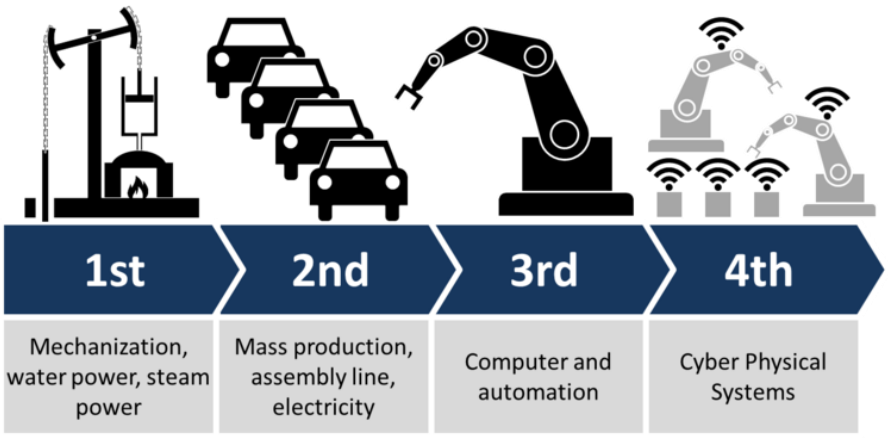

Figure 1.1. Industrial revolution progression [1]

Currently, living in the Industry 4.0 period, applications of the sensitive, intelligent and perceptive robotic agents remain at the forefront of manufacturing sectors.

1.2. Motivation

The idea of replacing human workforce with machines is not a new concept. For thousands of years, humankind has been trying to minimize the burden of manual work by inventing labor-saving devices, such as the plow or potter's wheel [3]. Among invented machines, for most of the people, automatons were appealing candidates to replace human beings because of their life-like behaviors and actions. In the eighteenth

century, people of Europe experienced the first form of the automata, known as Vaucanson's duck, which used a cam mechanism to sequence its movements to mimic human-like body movement [4].

The term "robot" coined in a science fiction play called "Rossum's Universal Robots" in Czech in 1921 [5]. The root of the "robot" word traces to the Czech word "Robotnik" means slave [6].

Although the term "robot" is familiar to most people and its physical appearance distinguishes them from other human-made objects such as cars and toys, there is not a unified or all-in-agreement definition for robots. Finding out a unique definition for the "robot" through literature is a challenging task. The concept of a robot has used for a broad range of machine types such as industrial robot manipulators, Unmanned-Aerial Vehicles (UAVs), humanoid robots, etc. However, the most straightforward definition has been offered by Bekey [7], as a machine that senses, thinks, acts, and is programmable. The sensing ability of robots enables them to collect information from their surrounding environment. For example, proprioceptive sensors measure the robot state such as robot pose in a workspace [5].

Robots need to be programmed and controlled in order to be able to carry out their tasks. For this purpose, robot pose information considers as essential and efficient factors. In other words, robot pose data enables them to communicate with the surrounding environment. The robot localization problem divides into two categories:

1. **Position tracking** or **incremental localization**: This type of localization refers to the case wherein the initial pose of the robot is known in advance, and the current pose of the machine is estimated based on updating the previous pose.
2. **Landmark-based pose estimation:** In this category, some natural or artificial landmarks are used for pose estimation in the environment. In other words, in this way, the robot computes its relative pose based on the landmarks and in this way the robot localizes itself in the environment.

Most robot pose estimation and localization algorithms use the information obtained from sensors such as laser rangefinders, sonars, GPS, wheel encoders and cameras. Although the laser rangefinders are simple to use and have excellent measurement precision, they are slow especially when the data need to be exchanged between the robot and the laser rangefinder at a high rate. Sonars are cheap and faster, but their pose data contains a high degree of uncertainty or error. Pose estimation based on odometry methods relies on robot dynamics such as wheel speeds, axes length, and wheel diameter. Odometry methods suffer from some drift-induced problems. In other words, an anomaly in the robot physics and dynamics will contribute to the erroneous results. When compared to the above sensors, cameras are more flexible, cheaper and have higher resolution measurements which can be used for object recognition, 3D reconstruction, and localization problems. Additionally, advancement in the computing infrastructures and tools decreases the computation time of the vision-based algorithms and make them accessible.

1.3. Tool Center Point

Tool Center Point (TCP) is defined as the tool-related reference coordinate that lies along the last wrist axis at a user-specified distance from the wrist [8]. Thus, Tool Center Point (TCP) locations lie on line passing through the center of the last link to the end of end-effector. The exact position of the Tool Center Point (TCP) is a function of the end-effector type. For example, for a gripper, it could be the center of the fingers. For a paint gun, it could be a specified distance from the painting gun [9].

1.4. Objective

The main objective of this study is to propose a vision-based method for estimating the 6 DOF pose (position and orientation) of the Tool Center Point (TCP) for the industrial robots. High-precision measurements, economic efficiency, and flexibility were some of the main goals of this study it can be applied in some manufacturing to improve the of the pose accuracy and repeatability of the industrial robots.

1.5. Thesis Outline

The thesis is organized as follows; first, the problem definition, motivation, and objective of this study are proposed in Chapter 1. Chapter 2, is focused on the literature review and past works concerning the TCP pose estimation of the industrial robots proposed by other researchers. Chapter 3 discusses the theoretical principles and implemented pose estimation algorithms in this study. Chapter 4 describes the experimental setup and procedures used for the conducting of the experiments. The results are interpreted and some discussions made over the proposed method in Chapter 5. Finally, Chapter 6 concludes the study and suggested some points for future works.

CHAPTER 2

LITERATURE SURVEY

2.1. Introduction

A statistical survey conducted on 7000 employees in manufacturing sectors from seven countries showed that nearly 70 percent of the interviewees believed that robotics and automation had improved their productivity considerably. According to the statistical analysis of the survey, countries involving industrial automation were experiencing high Gross Domestic Products (GDP) [10].

Rapid advances in industrial robotics and automation have attracted the attention of manufacturing and industry in mass production. According to the results of a survey [11] conducted by the International Federation of Robotics (IFR), more than three million industrial robots will be applied in industrial plants all around the world by 2020, Figure 2.1. The industrial robots selling market increased with a 15% growth in 2015 over 2014 (Figure 2.1). The IFR has an estimated average annual growth of 12% of the application of robots in industry and other sections from 2016 to 2019, which will be amount to the more than 2.5 million industrial robots.

Nowadays, industrial robots are used widely in manufacturing, especially in the automotive and aerospace industries [12]. Some advantages of the industrial robots' applications are reduction in waste materials, and production of high-quality goods with high precision.

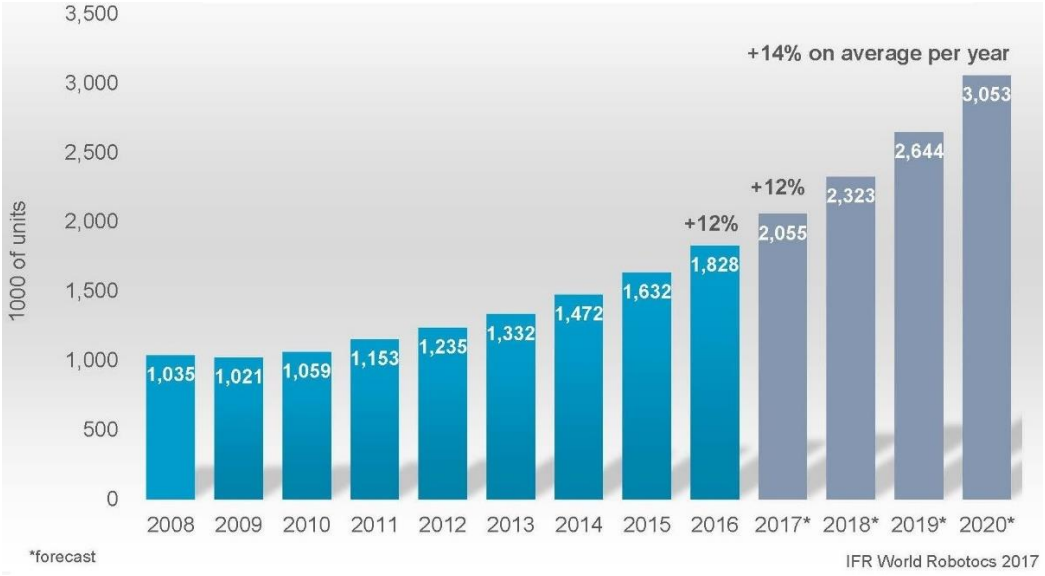

Figure 2.1. Estimated worldwide stock of industrial robots marketing [13]

The standard industrial robot classifications are done based on the type of the Tool Center Point (TCP) that installs at the last links which determine the application of the robot in manufacturing and industry. Based on this classification method, industrial robots are classified into three categories [14]:

1. Welding Robot: This category includes in Arc welding, electron beam welding, flux-cored welding, laser welding, MAG welding, MIG welding, orbital welding, oxyacetylene welding, plasma cutting, plasma welding, resistance welding, shielded metal Arc welding, spot welding, submerged Arc welding, TIG welding, TIP/TIG welding, and welding automation.

2. Material handling robot: In this category, industrial robots play the role of an assistant called collaborative robots, dispensing robots, injection molding robots, machine loading robots, machine tending robots, material handling robots, order picking robots, packaging robots, palletizing robots, part transfer robots, pick and place robots, press tending robots, and vision robots.

3. Third-party applications: In this category, industrial robot applications are restricted to the 3D laser vision robots, appliance automation robots, assembly robots,

coating robots, cutting robots, deburring robots, drilling robots, grinding robots, fiberglass cutting robots, laser cutting robots, material removal robots, milling robots, paint robots, polishing robots, sanding robots, spindling robots, and waterjet robots.

The demand for precise industrial robotic manipulators is increasing year-over-year in manufacturing sectors, especially in the aerospace industry. Thus, the problem of improving industrial robot precision has been one of the main concerns of the robot manufacturers and users.

2.2. The Significance of the Pose Knowledge of Industrial Robotic Manipulators

An industrial robot pose accuracy or pose precision is defined as the robot's ability to precisely position its TCP at a specified pose commanded by the robot controller. For this purpose, it is required to estimate the TCP pose with respect to a reference point [15].

The controllers of the industrial robotic manipulators contain an ideal mathematical model of the manipulators which relates its encoder readings to the Tool Center Point (TCP) pose. However, due to some unmodeled phenomena that is discussed in the following sections, the actual pose of the industrial robotic manipulators deviates from the ideal pose estimated by the controller. The ideal mathematical model implemented in the industrial robotic manipulator controller is called the nominal model, and the actual industrial robotic manipulator is called the actual model [16]. Accordingly, to increase the pose accuracy, all efforts should be put to decrease the error between the ideal and actual models which leads to higher positional accuracy. Industrial Robotic manipulators' positional accuracy is an important factor in applying them in the manufacturing sections mainly in the aerospace industry which requires a high degree of accuracy [17]. Nowadays, most of the industrial robotic systems are optimized for repetitive automation tasks by using teaching methods. In this way, the robot learns manually the desired pose of the TCP to accomplish its task in a repetitive manner. But, teaching approach is inefficient in some non-repetitive tasks especially in digital building constructions which require the on-line pose estimation of the TCP [18]. The

positional inaccuracy of a drilling end-effector mounted on an industrial robotic manipulator, Airbus UK [19], showed that the drilled-holes produced by a robotic manipulator did not meet the specific requirements by the Airbus standards

Thus, the 3D on-line pose estimation of the Tool Center Point (TCP) of robotic manipulators is an important factor in high-precision manufacturing and medical sections which demands more accurate robotic manipulators.

2.3. Error Sources of the TCP pose inaccuracy

This section discusses the primary sources of the positional errors of the industrial robotic manipulators that must be detected by the positioning systems. The positional errors of the industrial and non-industrial robotic manipulators is categorized into the following classes:

1. Geometric errors: Geometric errors are one of the principal error sources in the TCP positioning [20]. It is related to the manufacturing inaccuracies of the robots' parts. This kind of error is as a result of the non-ideal geometry of the links and joints of manipulators, such as errors due to machining tolerances.

2. Environmental errors: The common form of environmental errors occurs in the form of thermal errors. The robots' links and other structures are fabricated from materials that can be expand. Therefore, when temperature increase in the workspace of the robot; the temperature of motors, bearings, and structural elements increase causing thermal deformations causing dimensional inaccuracies.

3. Dynamic errors: This kind of error occurs as a result of the dynamic loads associated with motion. Dynamic errors are due to inertial loading and structural resonance created by the motion. Therefore, in the pose accuracy experiments conducted in static points, it can be negligible. However, in the path-based controls, they become significant [20].

4. System errors: These errors are due to improper calibration, sensor inaccuracies, drive train backlash, and installation errors [20][21].

All of these errors affect the TCP pose accuracy and impede it to reach to the desired pose calculated by the nominal controller, thus, decreasing of the robot's pose accuracy. Based on the stated facts in this part of the study, actual pose estimation of the TCPs of the robotic manipulators such as industrial and non-industrial types is a critical operation for pose correction in the manufacturing and medical applications. In the following section, the standard and popular methods applied for the actual pose estimation of the robotic manipulators have been discussed in more detail.

2.4. TCP Pose Estimation Techniques

Although most of the industrial robots have high positional repeatability around 30-50μm [21], however, they suffer from accuracy problems. The poor accuracy of the industrial robotic manipulators has limited their application, particularly in the aerospace, automobile manufacturing, and carbon-fiber-reinforced polymers industries which requires a high-degree of robot accuracies [22]. For instance, in many aerospace manufacturing tasks, the required accuracy constraints may be several orders of magnitude higher than the robot's accuracy reported by the factory. Thus, there is a great need for the improvement of the pose accuracy of the TCPs. Doubtlessly, the robotic manipulator calibration operation is one of the robust solutions to this problem. The central part of the calibration operations is the measurements of the relationship between the robot's base position and the robot's TCP accurately [22]. Because of the non-linearity of the industrial robotic manipulators due to the error sources (Section 2.3) most TCP errors cannot be detectable at robot's joints. For example, in many cases, joint encoders of the robotic manipulators are not able to measure the deflections of the links of robotic manipulators. Therefore, position and orientation errors accumulate at the TCP pose and make the robot as a less-precise device [23].

Most of the mathematical models developed of the TCP pose correction operations have not incorporated the effects of the error sources mentioned the Section 2.3. As a result, some researchers have proposed and developed external metrology systems to

obviate the problems of the positional inaccuracy due to the error sources mentioned in ection 2.3 such as robot's joint encoders. Identification of robot kinematic errors to improve pose accuracy by calibration or other pose correction methods requires the knowledge of accurate full pose measurements (position and orientation) of robot Tool Center Point (TCP) in Cartesian space [24].

In the following sections, the applied methods for the measurement of the actual pose of the industrial robot's TCP using external metrologies have discussed in more details.

2.4.1. TCP Pose Estimation Using Theodolites

Theodolite is an optical instrument that is used for angle measurements in the horizontal and vertical planes. Theodolites are used in land surveying, navigation, metrology, construction sector, and rocket launching [25].

The theodolites consist of three parts, including the base, the alidade, and the telescope [26], Figure 2.1. The base is installed on a tripod, the alidade rotates about the vertical axis, and the telescope rotates about the horizontal axis. The measurements principle of a theodolite is based on a line of sight of the telescope, which is usually defined by the two angles. There are circular vertical and horizontal circular scales mounted on the same axis of the rotation on the horizontal and vertical axes, Figure 2.2 [27]. In the old versions of the theodolites, measurement reading did by a Vernier scale which was a challenging task to read and extract precise value. However, modern theodolites have digital displays that reduce the chances of error readings and enables the device to connect with a computer to process or store the data acquired by the theodolite [28].

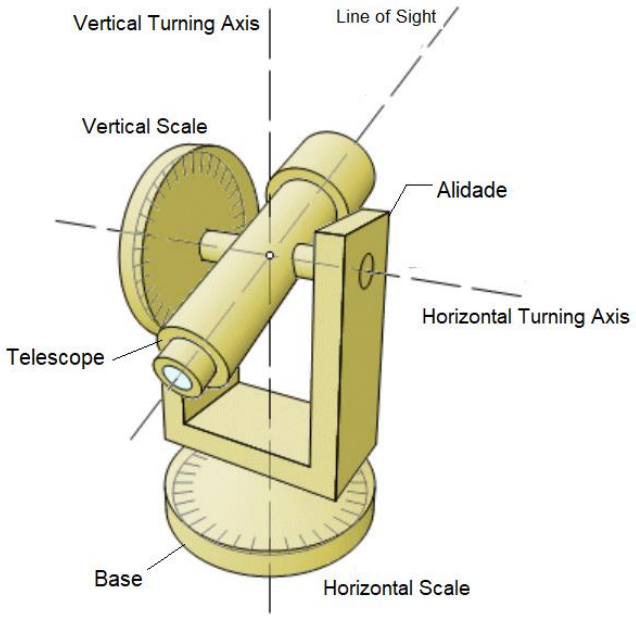

Figure 2.2. A schematic view of a theodolite structure [27]

It is surprising to say that theodolites do not compute distance information. They are only able to measure the target's angle information. Theodolites provide two independent angle quantities in the horizontal and vertical directions per pose. Directly speaking, for each pose, it gives a vertical and a horizontal angle. For example, if theodolite has focused at a target point on a robot TCP, it is deduced that target point lies somewhere along a line passing through the center of the telescope and target point. The measurement accuracy of the theodolites is around ± 0.5 arc second [27]. Based on the principle of the pose estimation using theodolites, it has been applied by some researchers for the industrial robotic manipulators for calibration operation. The [29] utilized a single theodolite measurement system for the calibration of a PUMA 560 manipulator to improve the accuracy of its forward model. The experimental setup was configured such that the target on the TCP viewed as a large subset of the workspace. The world reference frame is defined as the intersection of the device axes and represented by $X_wY_wZ_w$, and the target point reference frame is shown by $X_tY_tZ_t$ according to Figure 2.3.

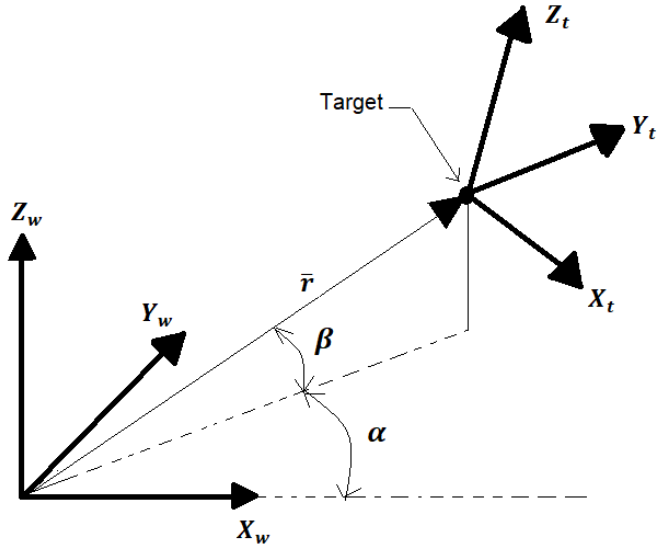

Figure 2.3. The theodolite poses estimation approach

According to Figure 2.3, the horizontal and vertical angles and distances between the theodolite's reference frame and the target are **α**, **β** and $\bar{r}$, and the transformation from the target reference frame into the world reference frame is defined by [29]:

$$\begin{bmatrix} ? & ? & ? & \bar{r}\cos(\beta)\cos(\alpha) \\ ? & ? & ? & \bar{r}\cos(\beta)\sin(\alpha) \\ ? & ? & ? & \bar{r}sin(\beta) \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.1)$$

The question mark elements in the (2.1) represent the elements of the transformation that are not specified by the theodolite device. The distance variable $\bar{r}$ cannot be determined by the theodolite. Since only two variables of the (2.1) are known, the absolute value of the pose is unknown. If the robotic manipulator changes its TCP into another pose, the observer looking the target point on the TCP would have no idea whether he is watching a small robot at close range or a giant robot at long-range. In order to obviate this problem, it is required to make an additional measurement such as length scale tools [30].

In order to increase the performance of a single theodolite, some researchers have combined it with other measuring tools like vision. A vision-based automatic

theodolite (VBAT) [30] was proposed and implemented to obtain an automatic theodolite for robot calibration purposes. The system consisted of a digital camera mounted on a servo stage rotating in the horizontal (azimuth) and vertical (elevation) planes. The whole system controlled using a computer system. In this proposed arrangement, the camera replaced in place of the usual theodolite telescope. Despite being inexpensive, the developed system could be applied for estimating of the positioning inaccuracies industrial robots in calibration processes.

2.4.2. TCP Pose Estimation Using Coordinate Measuring Machines (CMMs)

Doubtlessly, one straightforward approach to determine three independent coordinates of a point in space is the coordinate measuring machines (CMMs). They are applied as one of the most powerful metrological instruments in most manufacturing sites. The primary task of a CMM machine is to measure the actual shape of a workpiece and compare it against the desired shape. CMMs can measure the shape, form, location, and orientation of any object in 3D. The three-dimensional metrological information of the parts can be collected using different sensors [31]. Figure 2.4 shows a typical coordinate measuring machine.

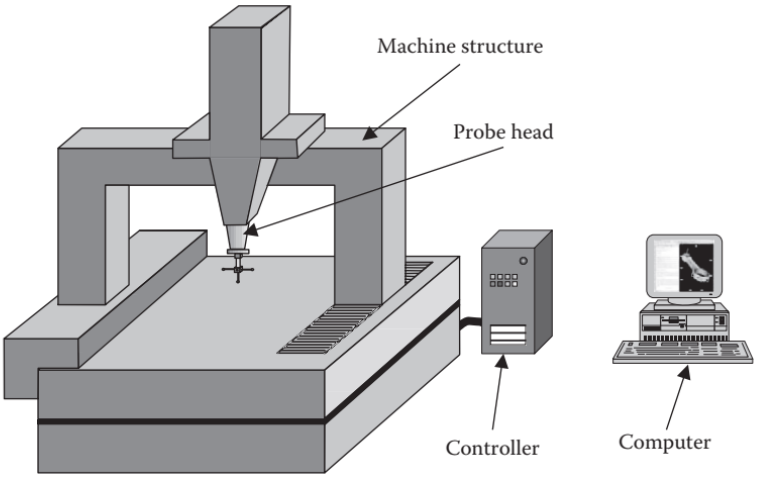

Figure 2.4. The components of a CMM machine [32]

The main components of a CMM machine [32] are composed of:

1. A three Axes mechanical frame equipped with displacement transducers at each axis.

2. A probe head or sensitive touch probe which contains the sensor to measure the part.

3. Control Unit.

4. Computer with a peripheral device such as printer, plotter, etc. and software to calculate and display the measurement results.

CMMs are also used for calibration purposes in industrial robots. For the industrial robotic manipulator calibration, a uniform sphere is attached to the Tool Center Point (TCP) of the robot. Then, the touch probe is moved so that it touches several points on the surface of the target (sphere). In this way, it can calculate the center of a sphere from the coordinates of the points on the sphere surface [29].

Veitschegger et al. have used CMMs for robot calibration using a fixture for PUMA 560 calibration. They reported the positional accuracy of their measurements system are on the same order of the positioning accuracy of the calibrated robot [33]. Mooring et al. [34] have applied a Mitutoyo CX-D2 coordinate measuring machine to estimate the pose of the PUMA 560 for calibration. They reported the workspace limitation. The CMMs does not cover the entire workspace of the PUMA.

2.4.3. TCP Pose Estimation Using Laser Trackers

The automotive and aerospace industries need some high-precision metrological equipment to measure the three-dimensional characteristics of the parts and large objects. Laser trackers are one of the precise metrological devices that can meet their requirements[35].

Laser trackers apply the principle of the interferometry principle to create point-to-point measurements [36]. Interferometry [37] is a measurement technique that is based on the phenomenon of interference of waves like light, radio or sound waves. The simplest form of a laser interferometer is shown in Figure 2.5. A laser resource creates

a beam of coherent, monochromatic light that is passed through a beam splitter. A portion of the laser beam is transmitted through the beam splitter and directed toward the moving mirror and other portion of light reflected toward the fixed mirror. The re-reflected lights from both the fixed and moving mirrors recombine in the beam splitter. As the moving mirror is displaced, the recombined beam interferes causing the photodetector to sense an alternating intensity with different wavelength.

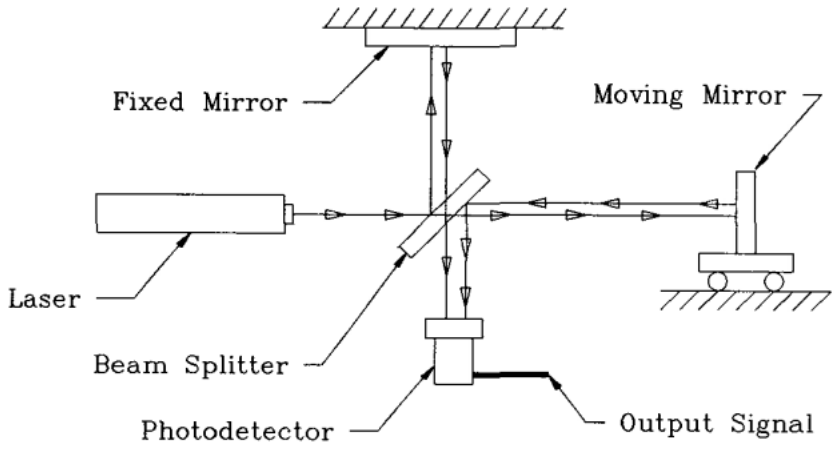

Figure 2.5. Laser interferometry [29]

By counting the number of interferences of the reflected light as a result of moving mirror displacement, it is possible to measure the displacement of the moving mirror. The accuracy of the interferometer performance is a function of the wavelength [29]. A laser tracking system is composed of a fixed base and a rotating head which emits a laser beam toward a moving retro-reflective target held against an object to be measured. The light re-reflected from the retro-reflective target re-enters the tracker in the same direction it has left. This reflected beam used as a means of the measurements of the target's three-dimensional property like position and dimension [37]. The spherically mounted retro-reflectors (SMR) are the standard type of targets offers by the manufacturers [37]. A laser tracker has two angular encoders and a distance meter. The angular encoders measure the azimuth and elevation of the laser tracker's head, and the distance is measured from the tracker's header to the center of

the spherical SMR target [37]. Figure 2.6 shows the schematic of a laser tracker and its working principle.

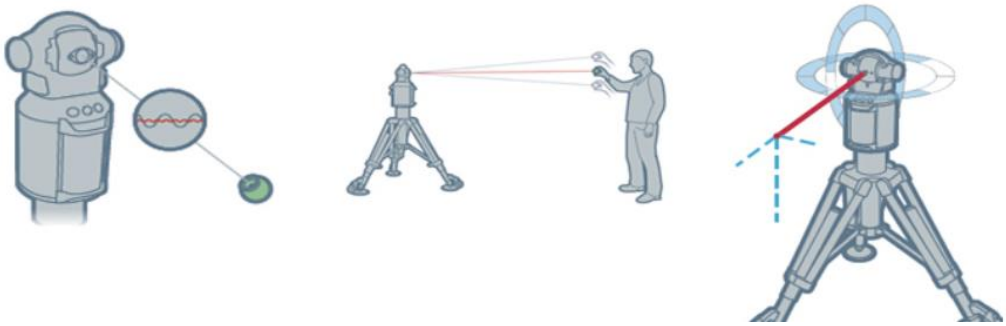

Figure 2.6. A schematic of a laser tracker [37]

In practice, laser trackers (only three producers in the world: Leica, FARO, or API) and optical CMMs (Nikon Metrology, Northern Digital, Metronor, Geodetic Systems, AICON, GOM) are two available pose measurements devices that are used for calibration of industrial robots to improve their positional accuracy [38]. Laser trackers are one of the reliable and available metrology tools used by most of the industrial robot manufacturers for robot calibration for many years. One of the earliest forms of the position measurements using laser trackers was due to Lau, and Haight [39]. Their system consisted only of a steerable beam laser interferometer, Figure 2.7.

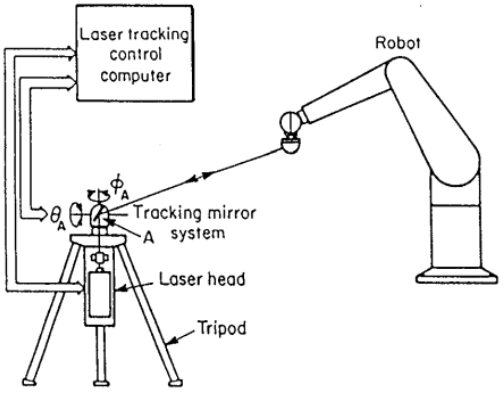

Figure 2.7. TCP measurement system developed by [40]

According to Figure 2.7, laser head sends out a laser beam in the vertical direction. As the beam strikes a mirror which can be precisely controlled by rotating the mirror

about the horizontal and vertical axes, the beam directed toward a reflector mounted on the TCP of the industrial robotic manipulator. The returning beam is directed into the laser interferometer. If the TCP starts to move, the reflecting beam alignment is changed and this is sensed by the laser interferometer detectors and, causes the mirror to rotate so that the returning beam realigns with the previous one. Since the mirror angles and distance to the target are known precisely it is possible to calculate the location of the target. The first advantage of this system is that it is fully automatic and can track the mounted reflector on the TCP. The second advantage is the high accuracy of the measurements. However, this system has a significant disadvantage of beam interruption which causes the destruction of the interference counting causing the measurements to restart. Decker et al. [41] have proposed a laser tracking system (LTS) for the real-time pose estimation (position and orientation) of a moving robot's TCP along an arbitrary trajectory. Figure 2.8 shows a schematic of the proposed system in this study.

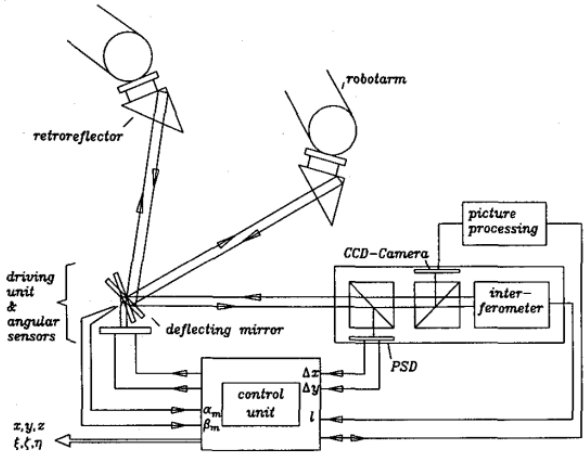

Figure 2.8. The graphical representation of the laser tracking system developed by [41]

A retroreflector consisted of three perpendicular mirrors mounted on the TCP of an industrial robotic manipulator. The mirrors configurations of the retroreflector caused the reflected beam from the retroreflector to be in parallel with the sent beam to the retroreflector, Figure 2.8. The returning beam is directed by a deflecting mirror into the interferometer, a position sensing diode (PSD), and a CCD camera. Since the laser

beam hits the center point of the retroreflector, accordingly there is not any displacement between the emitted and reflected beam. When the robot moved, the laser beam did not hit the center of the retroreflector, thus causing a displacement of the reflected laser beam. The PSD measured this displacement and considered it as the tracking error. To minimize the tracking error, the deflecting mirror's joints, which is called Cardon, rotated and allowed the emitted laser beam to follow the retroreflector mounted on top of the TCP. By this mechanism, the proposed system was able to follow arbitrary movements of the retroreflector. The TCP position is estimated using the data from the interferometer, the angular encoders, and the PSD signals. For orientation estimation of the retroreflector, the CCD camera analyzed the intensity profile of the reflected laser beam. The darks lines in the intensity profile were representative of the blackened edges of the retroreflector. By analyzing these lines, it was possible to calculate the orientation of the retroreflector with respect to the robot's TCP. The main advantages of the developed LTS system in [41] reported as being flexible and fast. It could measure a big set of points in a low time. The other advantage of the system was its ability for on-line measurements which caused to improve the robot dynamic properties leading to a higher path accuracy and precise point approaching. However, the system had issues in orientation measurements inaccuracy that needed to be improved. The orientation measurements of the developed system was improved in [41] . They added some wires at the front of the retroreflector, Figure 2.9. The supplementary shadows of the added wires created high-resolution images of the edges of the retroreflector which facilitated the orientation measurements analysis. Despite most efforts put on the [41] and [42], the resulting system could only detect the pitch and yaw angles in the range of ± 30°.

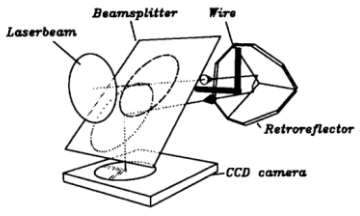

Figure 2.9. Modified retroreflector for pose estimation [42]

A systematic approach for the estimation of the positioning errors of a Motoman SK120 robotics manipulator was proposed in [43]. The main aim of their study was to identify the kinematic parameter and its corresponding errors corrections in order to improve the ability of the robot in reaching a specified pose more accurately. They applied a laser tracker (Leica LT500) for the pose estimation of the robot TCP. Since the utilized laser tracker was not able to measure the orientation, therefore their works suffered from the orientation estimation problem. Some appealing features of laser trackers such as giving high-resolution measurements, covering large workspace and contactless measurements have conceived most users to apply them in industrial robot calibration. The [42] have applied a Laser Tracking System (LTS) for the calibration of the PUMA 560 at the Robotics Center of the Florida Atlantic University for estimation of the calibration parameters. They obtained a calibration accuracy about 0.1mm which was four times precise than the results obtained using a CMM which estimated a 0.4mm for the PUMA robot [44].

2.4.4. TCP Pose Estimation Pose Estimation Using Computer Vision

Nowadays, with the rapid developments of the computational power of computers along with advanced progress in the visual techniques and algorithms, computer vision has placed itself at the top of the science, medical, industry, and high precision manufacturing. Some practical applications of the vision systems are [45]:

- Quality control to the inspection of the defected parts such as in manufacturing industries,
- Pose estimation in Robotic based manufacturing,
- Food and beverage industry,
- Medical diagnosis using X-rays or ultrasound images.
- Navigation for autonomous or mobile robots,
- Detecting events such as visual surveillance or people counting,
- Military applications.

A lens and a light-sensitive array are the two of the main components of every vision system. The light-sensitive array composed of a large number of discrete cells, called pixels that senses the frequency of the striking light on them. The frequency information from each pixel is digitized, and a discrete number between 0 to 255 is assigned to each sensed frequency. The values from 0 to 255 assigned to each pixel are called the gray-level values. Figure 2.10 (a), shows a schematically representation of the light-sensitive array, the working principle of a cell (pixel) of a light-sensitive array, Figure 2.10 (b), and the concept of the gray-scale images and gray-scale values, Figure 2.10 (c)-(d) [46][47] .

One of the primary concerns of a vision system is the resolution of the light-sensitive array, or in other words, the resolution of the image [29]. Nowadays, vision systems play important roles in the robotic and industry. In general, robots use vision systems for tasks such as obstacle avoidance, human-interaction, identifying and locating parts, quality control inspections, and to improve positioning accuracy [48]. One of the core requirements of industry 4.0 is the applications of flexible robots in the manufacturing processes. The flexibility of robots can be achieved by incorporating vision systems and other sensor technologies [49]. Industrial robots require high accuracy and repeatability in order to accomplish the assigned tasks. For example, in the automotive industry the required accuracy for spot welding is about 1mm, or in the aerospace industry it is at least ten- to twenty-fold higher accuracy [50].

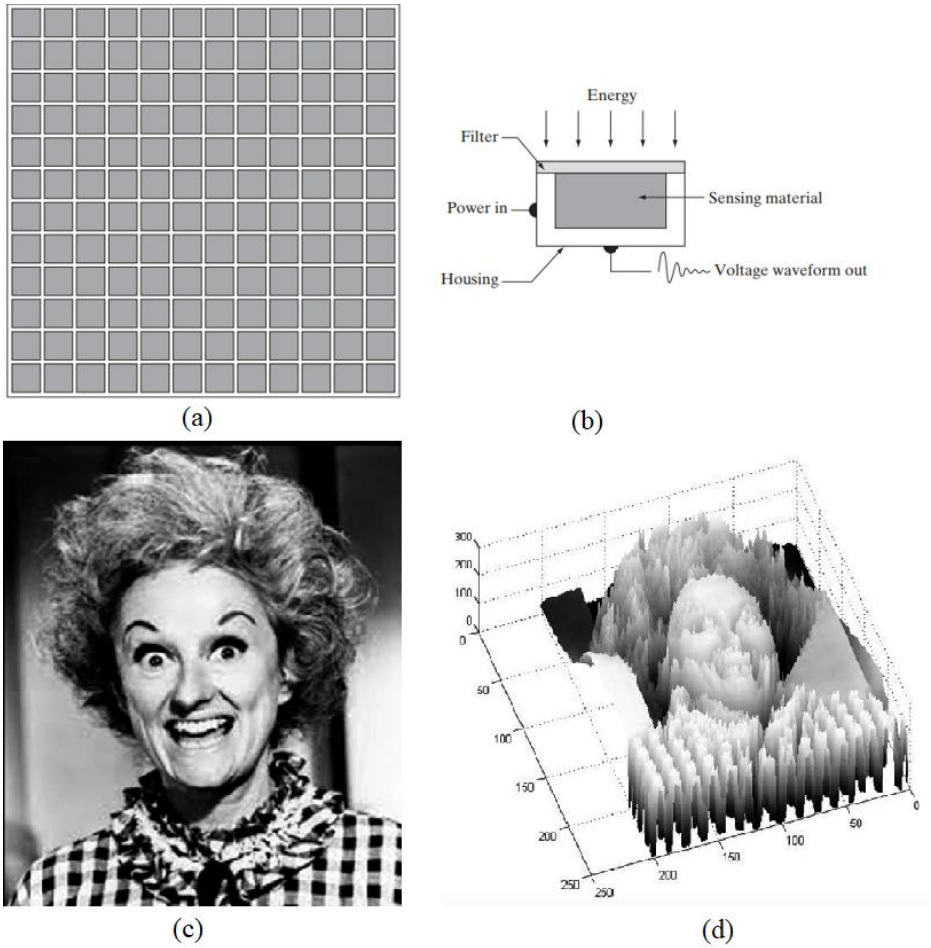

Figure 2.10. A light-sensitive array(a), the working principle of a cell array (b) [46], gray-scale image (c), and the concept of the gray-level values (d) [47]

One solution to improve the industrial robots' accuracy is to equip them with optical-based sensors and equipment such as laser trackers systems, photogrammetry or vision systems with high-resolution cameras for applying them in pose estimation operations of the robots TCP [50].

Vision techniques in industrial robotic fields for the TCP pose estimations tasks can be classified according to Table 2.1 [50]. In the following sections, each of the vision techniques for the pose estimation of the Tool Center Point (TCP) of industrial robots is discussed in more detail.

Table 2.1. *Vision-based classification for robotic applications* [51]

	Single Camera	Multiple Cameras
Passive vision	*2D*	Stereo vision Photogrammetry
Active vision	Time of flight Structured light Light coding Laser triangulations	Structured light Projected texture stereovision

2.4.4.1. TCP Pose Estimation Using Stereo Vision and Photogrammetry

Photogrammetry is a visual technique for the measurements of the dimensions of an object from their digital images using 3D reconstruction techniques and algorithms. In general, there are two ways to obtain three-dimensional information of an object from their 2D dimensional images [52]:

● **Model-based pose estimation:** A single calibrated camera and a model with a known geometry is used to determine the pose of the model, Figure 2.11.

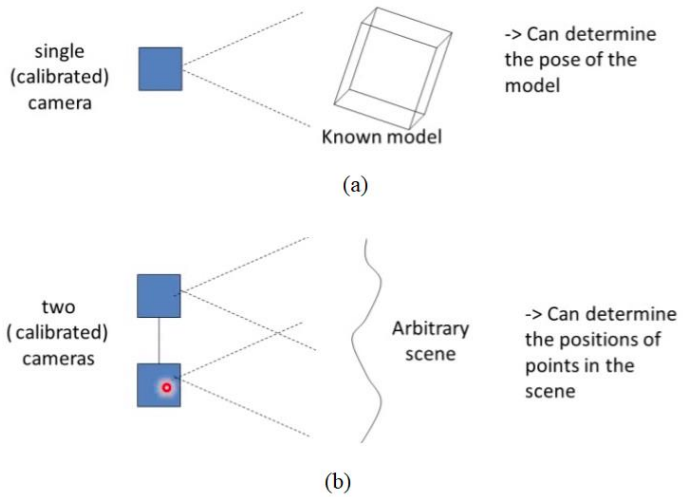

Figure 2.11. Model-based pose estimation (a), and stereo vision-based pose estimation (b) [52]

● **Stereo vision:** In this method, two calibrated cameras with a known pose between them is used for obtaining the 3D information of an unknown scene, Figure 2.11 (b).

In stereo vision and photogrammetric techniques the same point in a real-world scene by different cameras and then the intersection of the projection lines is computed to obtain its 3D position. For this purpose, some physical markers like stickers or laser points are necessary to stick over and around the object for better detections, Figure 2.12. (a)-(b) [50].

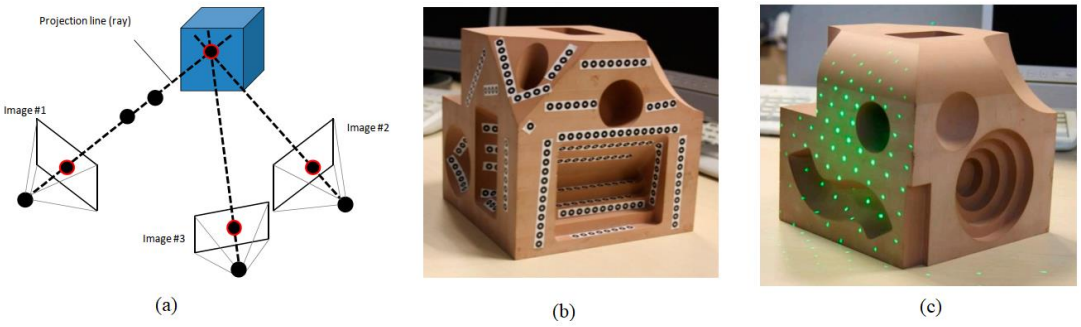

Figure 2.12. Concept of the projection lines in stereo-vision (a), physical markers-stickers (b), physical markers-laser points (c) [51]

Stereo vision-based pose estimations have been widely used in the industrial robotic manipulators such as in assembly. In [52], it was proposed an automatic smartphone assembly system in order to assemble the phone back-shell to the phone itself, Figure 2.13. Their proposed and developed system consisted of two-color CCD cameras mounted on the TCP of a six-degree-of-freedom Mitsubishi industrial robot, a vacuum absorption tool, and a personal computer to control the whole system. The calibrated CCD cameras was used to estimate the pose (position and orientation) of the back-shells and smartphones itself using the extraction of the corner features of the smartphone and its back-shell from the images acquired by the two cameras. In their study, the visual pose estimation was used for the robot control system to perform of the robotic assembly tasks. Although, the experiment results were satisfactory for the laboratory environment, in a practical factory environment, the visual field of view of the two cameras was not able to the entire workspace, since in the factory, the distance of the smartphone and its back-shell may be far away from each other to be seen inside in single frame.

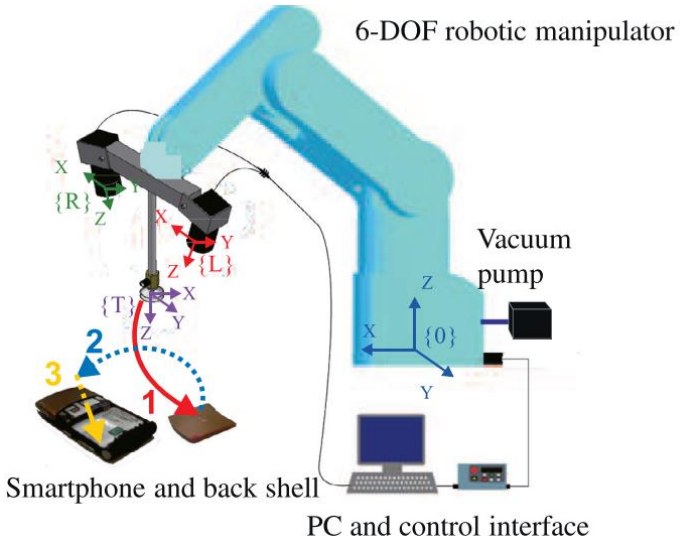

Figure 2.13. The automatic smartphone back-shell assembly system [52]

Thus, when the camera got closer to smartphones, it caused the loss of the camera fields of view.

Švaco et al. [53] has proposed a non-contact stereo vision-based system attached to a robotic manipulator, KUKA KR 6 R 900, which was used to measure the pose of the calibration points in space. The main aim of this study was to improve the absolute accuracy of the industrial KUKA robot which is a common problem in 6 DOF serial link manipulators [54]. Figure 2.14 shows the experimental setup configuration used in [53]. The system consisted of two perpendicular CCD cameras, with macro lenses, that constituted a virtual robot Tool Center Point (TCP), Figure 2.14 (a). A black polymer sphere was used as a calibration point with known pose, and its center position coordinates estimated by the cameras. Then, the virtual TCP (two perpendicular cameras) manually moved to the desired position of the black sphere (target). At this point, the vision system automatically adjusted the robot position with respect to the calibration sphere in the camera coordinate frame. These cycles were repeated over and over and at each stage, the joint variables were recorded for the calibration purposes. The final results of the calibration have proven the reliability and

accuracy of the vision system applied in this study for the improvement of the absolute accuracy of the industrial robots.

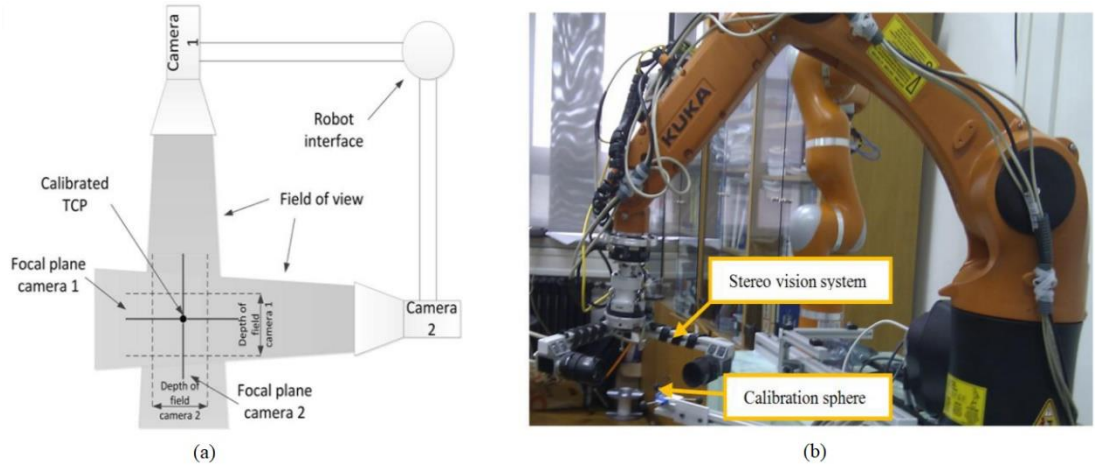

Figure 2.14. TCP pose estimation using a stereo-vision system developed by [53]

Möller et al. [55] have proposed a method for increasing the absolute positional accuracy of an industrial milling robot using stereo vision camera system. They estimated the pose (position and orientation) of the tool center point (TCP) of the robot with a milling spindle. Their vision system was able to measure the pose of the retro-reflective target attached on the milling spindle, Figure 2.15, and during the milling process, the target was observable by the camera systems. A laser tracker system was used for the calculations of pose relations between the workpiece, the robot, and the camera system as well as the pose relationship between the TCP and the milling tool.

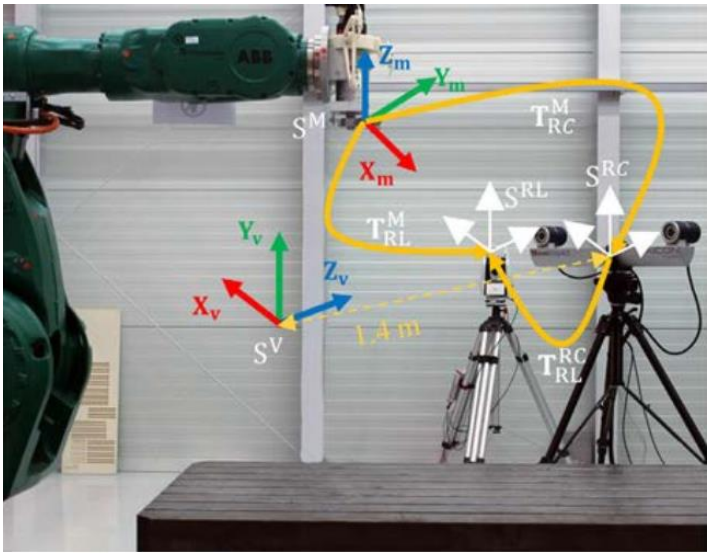

Figure 2.15. Experimental setup of [56]

The applied optical systems in this study consisted of a Leica laser tracker AT 401 and an AICON multi-camera system which enabled the high-accuracy pose estimation of the robot spindle in big workspaces. However, their proposed vision system accuracy decreased in longer distances. The proposed system was able to measure the target's pose at the frequencies up to 10Hz. The final results of their studies showed that the stereo vision systems can be considered as a reliable external metrological device for improving the absolute accuracy of the industrial robots for machining operations in big workspaces. Stereo vision and photogrammetry techniques have applied in applications other than the TCP pose estimations. One of the applications of the vision systems in modern manufacturing is in advanced robotic welding, especially for the path corrections of off-lined programmed welding robots and for the localization of the weld seam pose of robotic arc welding [56][57].

2.4.4.2. TCP Pose Estimation Using Structured-light

Active vision technique is composed of a light source (light projector) and one or two data receptor such cameras. One of the most common active vision techniques is structured light vision systems. Unlike passive vision systems like stereo vision and photogrammetry techniques which requires the visible features of an object from

different viewpoints for pose estimation, in the structured light techniques, artificial visual patterns such as a white light is projected on the scene in order to create visual patterns. One of the main obstacles of applying of this technique is related to the size and weights of sensors which comprised of a projector and a camera that makes it challenging to mount them on top of the industrial robot [50]. A solution to this problem was purposed by [58], Figure 2.16.

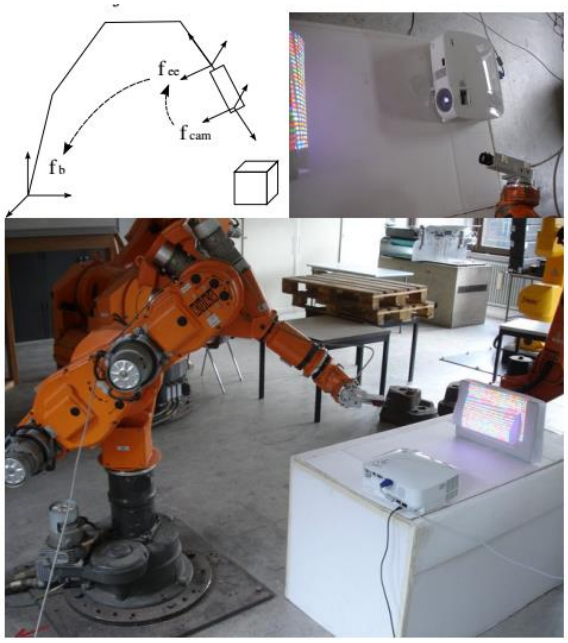

Figure 2.16. Structured-light in Robot Positioning [59]

In this configuration, Figure 2.16, the camera was mounted on the robot TCP in an eye-in-hand configuration and the projector installed on a static and fixed frame illuminating the workspace. They reported a positional accuracy of 3mm, which met their project requirements. They found the eye-in-hand configuration of the proposed system as a positive point which solved the occlusion problems as well. The [59] have applied the structured light techniques for robot positioning in a visual servoing project. They installed the projector aside the robotic manipulators, and a structured light created a coded light pattern in order to create visual features for robot positioning. Their proposed system was able to position the robot with respect to

planar objects with reasonable accuracy, but the camera motion became noisier when using non-planar objects for robot positioning.

2.4.4.3. TCP Pose Estimation Using Optical Coordinate Measuring Machines

Optical CMMs apply the image processing techniques and algorithms to perform the automatic measurement which is fast and highly accurate [60]. One of the popular commercial brands of the optical sensors is the Nikon Metrology optical CMM (KCMM) which consisted of three linear CCD camera. Optical CMMs apply the triangulation techniques for pose estimation procedure. In order to use this measurement device for metrology purposes, it is required to attach active LEDs on the surface of a target. If the optical CMM detect only one LED, the three-camera of the system estimates the pose of the LED in space using triangulation techniques, if two LEDs are detected by the system, they can calculate the distance of the LEDs as well. Using nine LEDs, the optical CMMs is able to track the position of the object in a volume up to 6 meters away from the device [61]. Recently, optical CMMs are becoming popular in the industrial robot sections because of their high-accurate measurements. Optical CMMs can measure the pose of the robot TCP dynamically at frequencies of 30Hz or more with respect to the robot base. Some CMMs use active targets like LEDs or infrared diodes. The main advantage of active targets is that they are insensitive to environmental conditions like light variations. Other types such as C-Track series, use passive targets which basically consisted of small circular stickers covered with retroreflective materials [39]. Figure 2.17, shows two different brands of the optical CMMs commonly used by researchers in industrial robotic applications.

Figure 2.17. AICON optical CMM (a) [64], and Nikon optical CMM (b) [62]

A few researchers have applied Optical CMM for the improvement of the positional performance of industrial robots. The [63] proposed an online dynamic pose correction algorithm to improve the pose accuracy of industrial robots at a single pose for applications such as drilling or riveting. This algorithm applied the dynamic pose measurements as feedback to the robot control system to guide the robot TCP to the desired location. They used an optical coordinate measure machine, C-Track 780, for the online pose estimation, and a 6 DOF industrial robot (FANUC LR Mate 200iC)). Figure 2.18, shows the experimental setup applied by [63].

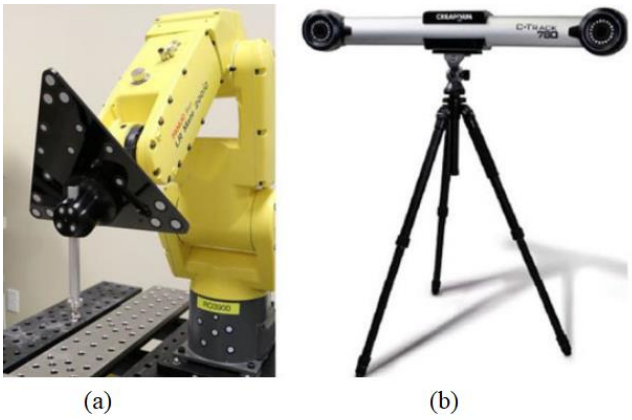

Figure 2.18. Experimental setup: FANUC LR Mate 200iC (a), C-Track (b) [63]

The C-Track system equipped with a binocular vision sensor that can measure the position of the retroreflective targets. The measurement rate of the vision system was 29 Hz. In their study, C-Track measured the pose of the robot TCP with respect to a

workpiece. They applied four retro-reflective targets to evaluate the position and orientation of the robot. The 100 static points were measured using the C-Track (Optical CMMs) for the evaluation of the pose estimation. Figure 2.19 shows the schematic view of the experiment and reference frames in this study.

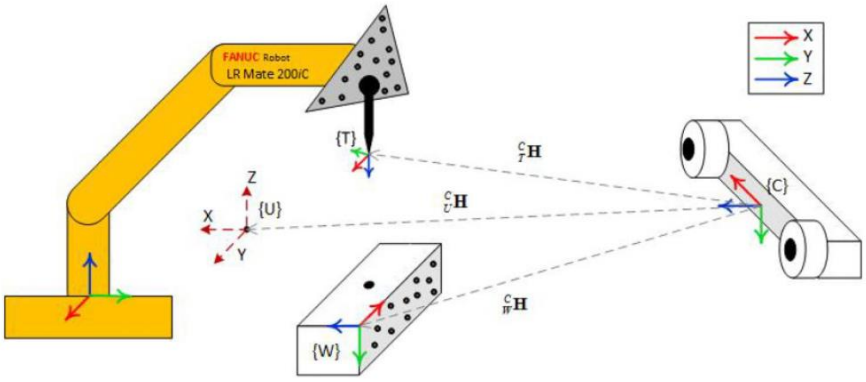

Figure 2.19. The schematic of experimental setup [63]

The experimental results demonstrate that using the optical CMM as the online pose measurement tool can improve the positional accuracy of the robots in the orders of 0.05mm for the position and 0.050° for the orientation.

2.4.5. Sensor-Fusion based TCP Pose Estimation

In precise applications, applying a single sensor for pose estimation purposes (position and orientation) of a robot such a TCP is not so precise. All sensors have some degree of uncertainty which causes loss of precision [64]. To solve this problem, multi-sensor fusion techniques have been proposed by some robotic researchers. The main objective of the sensor fusion algorithms is to obtain a high performance acquired by integration of the individual sensors. In [65], data from inertial measurement sensors (IMU) and vision systems were fused in order to improve the accuracy of the proposed pose estimation system. Applying IMU sensors alone accompanies with uncertainties such as sensor drift errors and using vision by itself cannot handle the problem of occlusion moreover, vision cannot cover the whole area of the robot workspace. Kalman filter and extended Kalman filter are two of the most common sensor fusion algorithms have that been used in robot mapping and localization projects for many

years. Siciliano et al. [65] implemented an algorithm for the pose improvement of a robot manipulator by implementing a multi-sensor combination measuring system (MCMS). In this study, two data fusion algorithms have been applied to fuse the position and orientation data such as the Kalman filter (KF) and multi-sensor optimal information fusion algorithm (MOIFA). The implemented system setup was composed of an industrial robotic manipulator, an industrial photogrammetry system, a digital inclinometer, and a PC and software, Figure 2.20.

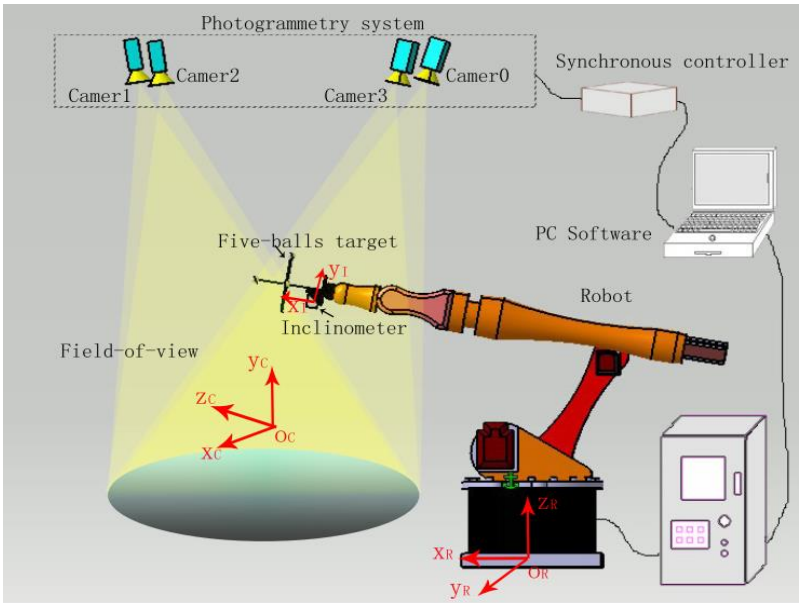

Figure 2.20. Multi-sensor combination measuring system (MCMS) [65]

A high precision industrial three-dimensional photogrammetry system and a high accuracy digital inclinometer (in the accuracy of 0.01°) have been selected for the position and orientation measurements. The experiment results showed that a significant decrease in the positional error of the robot manipulator less than 1 mm after pose calibration of the robot using these proposed methods. Moreover, the best accuracy of the photogrammetry system was obtained in the center of the camera field of view. They reported a 38%-78% decrease in pose error by the applying of the fusion techniques.

2.5. Review of Robotic Manipulator TCP Pose Estimation Techniques

In this section, a brief review is carried out over the TCP pose estimation techniques used in literature. Theodolite is one of the earliest tools used for the evaluation of the industrial robot manipulator's pose accuracy. Despite being cheap, or being rented at a reasonable cost, and using a simple target, its major drawback is the measurement recordings process which requires long times to read the measurements. Mooring et al. [30] has applied a single theodolite for the pose estimation of a PUMA 560 robot end-effector. They reported that the reading procedure as time-consuming and subject to the errors. Podoloff et al. [16] has mentioned theodolite as "slow and fatigue". Moreover, they are only able to acquire the angle information of the targets. Add to above; theodolites are not able to estimate targets at distances closer than 3.048 to 4.572 meters because the target would be too large. Therefore, theodolites can be used in big areas [29].

Perhaps one of the appealing devices as a solution to the pose estimation problems of theodolites is the coordinate measurement machines (CMMs). Being easy-to-use, requiring only a simple target, and giving easily interpretable data are some of the advantages of this system [29]. Most CMMs have smaller workspace and this hinders their usage in industrial robotic applications and CMMs with large workspace are very expensive. Even with the availability of large workspace CMMs, in most cases, there is a need for fixtures [29,33] to accommodate the CMM for robot pose estimation purposes. Nguyen et al. [24] has considered the theodolites and the CMM based methods relying on jigs and fixtures as traditional tooling measurements techniques that do not meet the precision requirements imposed by the automatic and aerospace industries. Doubtlessly, laser tracking technologies have the greatest impact on metrology instrumentation developments and techniques. Recently, laser trackers have kept the attentions of the industry and researchers pose measurement of robot manipulators in calibration and control operations. In simple versions, laser trackers

can measure a point in 3D coordinate, however, they are able to measure the full pose of any object by attaching at least three spherically mounted reflectors (SMRS) [39]. Being precise and able to measure the pose data in a large working volume are some advantages of applying the laser trackers in industrial robotic TCP pose estimations. However, similar to any physical systems, laser trackers suffer from some measurements problems. Firstly, it is not so easy to measure the TCP pose in any arbitrary orientation especially when the attached target on the TCP is not in the field of view of the laser. Erdem et al. [67] has proposed a configuration of the active LEDs homogenously distributed in a spherical shape, Figure 2.21, for the pose estimation of the Tool Center Point (TCP) in any orientation.

Figure 2.21. An innovative LEDs target for the TCP pose estimation purposes *[66]*

The other problem of pose estimation of industrial robots using laser tracker is associated with their precision degradation in manufacturing environment. The laser trackers measure the pose of the robot TCP with respect to the robot base which is secured to a fixed point on the factory floor. Any vibration due to the working of other manufacturing tools such as pressing machines decreases the accuracy of the pose estimations [39]. The other disadvantage of the laser trackers is related to their sensitivity to the ambient conditions of the factories or workshops such as air currents [29]. Moreover, laser trackers are extremely expensive starting at the minimum pricing of 100000$ to approximately 300000$ or more [39].

Unlike laser trackers, optical CMMs measures dynamically the robot TCP with respect to the robot's base at frequencies of 30Hz. In contrast to laser trackers, optical CMMs are cost-effective, however, the workspace of the optical CMMs are smaller than laser trackers [64]. Optical CMMs are less accurate than laser trackers [65][66]. In literature, there are a few researchers that have applied optical CMMs in TCP pose estimations operations [67]. Some optical CMMs, such as Nikon Metrology and Northern digital [68], use infrared light-emitting diodes as their active targets. The main advantage of optical CMMs with active targets in TCP pose estimation operations can be associated with being easily detectable and not affected by the environmental light conditions, however, they are so expensive. Most researchers preferred to use optical CMMs with passive reflectors over laser trackers due to their high prices [64][69].

Vision-based techniques for the TCP pose estimation purposes can be evaluated based on several criteria such as accuracy, range, weight, safety, processing time, and environmental effects. In robotic TCP pose estimation applications, stereo-vision, and photogrammetry techniques have been used by most researchers [53] [55][58][67][70]. Stereo vision and photogrammetry techniques give accurate position results in the order of 50μm [71]. However, photogrammetry and stereovision system uses several images of a single point and compares its disparities in several images, it has the problem [53] of a massive amount of information which causes an unacceptable increase in processing time [50], laser trackers are also effected from the environmental influences mainly ambient light.

Time-of-flight cameras are not used commonly for the industrial robots pose estimations since they are not so accurate despite their insensitivity to environmental disturbances. Konolige [72] reported the pose accuracy of the Time-of-flight cameras around 10mm. The application of the structured light equipment for the industrial robotic pose estimation purposes does seem to be appealing because of their heavy-weights, long processing time, and being sensitive to environmental lightning [59].

Despite, high-performance pose estimation capability of the sensor fusion-based techniques, and being fully automatic measurement processes, there are few numbers of researches in this field and it seems, this technique is not so popular by industry and scientific researches for industrial robot pose estimation. However, applying IMU sensors and vision in fusion-based for the pose estimation purpose in the industrial sector is promising.

In conclusion, based on the literature reviews, currently, the most common, reliable and popular pose estimation devices, and tools for the pose estimation of the industrial robots are the laser trackers, optical CMMs, and stereo-vision and photogrammetry techniques.

2.6. Contribution

The main contribution of this study is using a single and cost-effective web-camera, unlike stereovision and photogrammetrical methods, which use several expensive cameras, while keeping the proposed system's accuracy close to or even better than these systems. The most outstanding feature of this system is its cost-effectiveness, it is composed of a 3D geometrical solid with several pieces of print-out papers that are not comparable with most expensive laser trackers, optical CMMs, and stereo vision systems which costs even up to 300000$ for some versions able to measure the orientation. Moreover, the proposed system has a high pose measurement accuracy and repeatability. It is easy-to-use and does not require any complex setup or extra tools to perform the pose estimation.

CHAPTER 3

METHODOLGY

3.1. Introduction

The proposed method for the 6 DOF pose estimation of TCP is discussed in this section. The principle of the purposed method is based on using the square fiducial markers in the form of an innovative marker map called *Rhombicuboctahedron ArUco Mapper* (RAM). ArUco markers are used as the fiducial markers in this study. Firstly, the principle of the ArUco markers, as the elements of RAM, is discussed briefly. Following that, the idea behind the geometry of the map is addressed in more detail. Then, the idea behind the geometry of the marker map is discussed in more detail.

3.2. Absolute TCP Pose Estimation Method

In this study, the 6 DOF absolute pose of the TCP is estimated in the robot's base coordinate frame. The whole method is composed of four sub-methods, including:

- Marker detection and identification
- Marker mapping and localization
- Model pose estimation with respect to the camera
- TCP pose estimation with respect to the robot base

A general overview of the implemented method is shown in Figure 3.1. In the following sections, the details of each sub-methods are discussed in more detail.

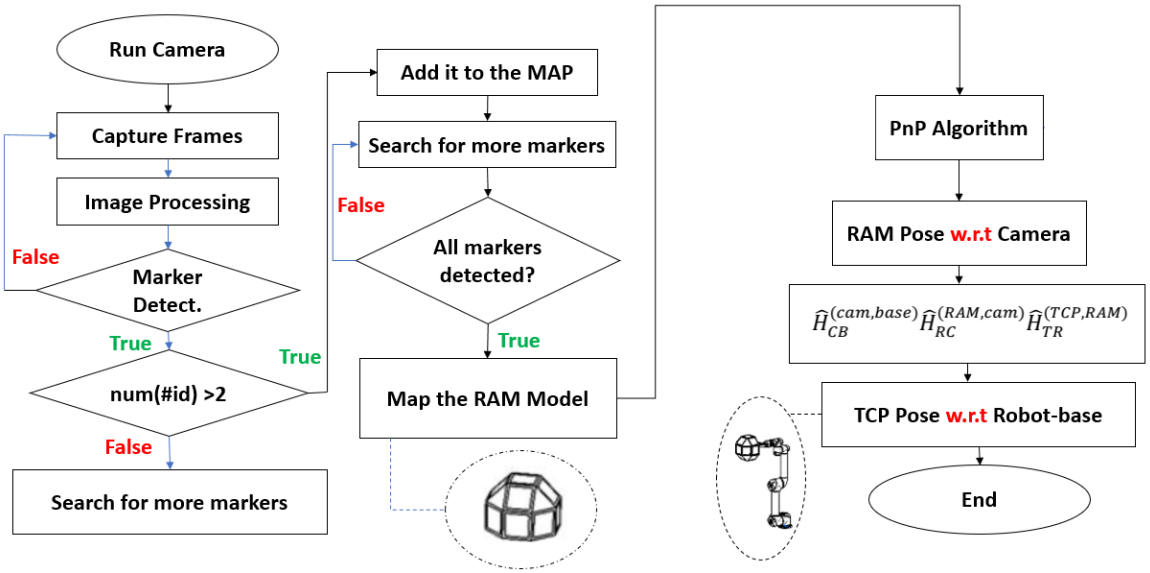

*Figure 3.1.*Absoulute 6 DOF TCP pose estimation method developed in this study

3.2.1. The principle of the fiducial markers (ArUco.3)

The structural elements of the developed TCP pose estimation method are fiducial markers. The fiducial markers are used as 3D landmarks for pose estimation applications in robotics and augmented reality for many years. Nowadays, the more common types of fiducial markers are in square shapes and have bitonal patterns in white and black. Being in square shape helps the fiducial marker system to recognize the markers and distinguish them from other objects in the environment. Moreover, the square shape provides four corner points for the pose estimation algorithm. The bitonality of the markers enables the fiducial marker system to codify each marker. Thus, each marker is identified by a unique identity as an id number. Additionally, black and white patterns eliminate the need for lighting and camera sensitivity problems. There are various types of fiducial markers to be applied to pose estimation problems such as ARToolKit, ARToolKit Plus, ARTag, and ArUco.

In this study, an optimized version of the ArUco markers called ArUco3 is applied. Compared to other markers, its processing time has decreased considerably, and memory-related problems have been solved.

Each fiducial marker has its own coordinate frame. The ArUco marker coordinate frames are located in the center of the marker. If it is assumes that the side length of the marker is *s*, the coordinates of the corner points of the ArUco are according to the convention shown in Figure 3.2.

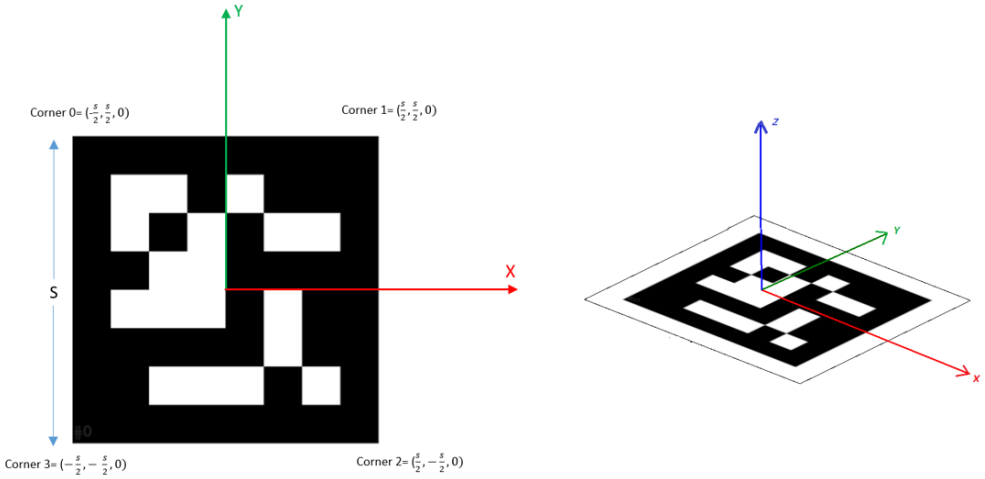

Figure 3.2. ArUco marker coordinate system

The coordinates of the corner points of an ArUco marker expressed in the marker's coordinate frame are:

$$c_1 = (\frac{s}{2}, -\frac{s}{2}, 0) \quad (3.1)$$

$$c_2 = (\frac{s}{2}, \frac{s}{2}, 0) \quad (3.2)$$

$$c_3 = (-\frac{s}{2}, \frac{s}{2}, 0) \quad (3.3)$$

$$c_4 = (-\frac{s}{2}, -\frac{s}{2}, 0) \quad (3.4)$$

3.2.2. ArUco Marker Detection

The first part of the proposed method is called marker detection. In this sub-method, some image processing and computer vision algorithms are applied to input frames. Figure 3.3 shows the marker detection procedure of the developed method.

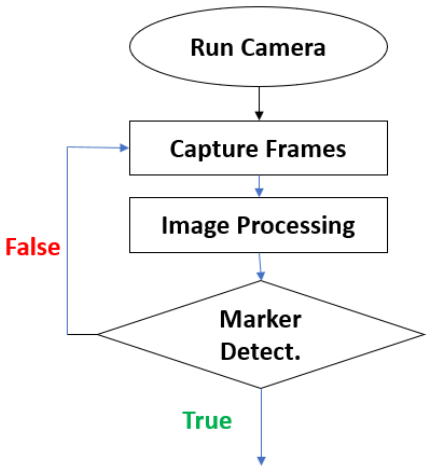

Figure 3.3. Marker detection procedure of the proposed method

The ArUco marker detection procedure is as follows:

1. **Frame acquisition:** In this stage, the camera shutter opens and frames acquired as input to the proposed method.
2. **Pre-processing:** In this stage, some low-level processing like noise suppression and the color frames are converted into gray-scale frames.
3. **Applying the adaptive thresholding**: Adaptive thresholding is applied to extract all the borders of markers in frames.
4. **Contour extraction**: In this step, the closed curve of points or lines illustrating the objects' boundaries are extracted. However, it is worth to mention that not all the extracted boundaries are markers.
5. **Filtering**: All the extracted borders with a small number of points removed.
6. **Polygonal approximation of contours:** In this step, only the concave contours with precisely four corners points are kept.
7. **Sorting**: Counter-clockwise corner sorting is carried out.

8. **Filtering-out too small rectangles**: In this step, only rectangular shapes with most external borders are kept.
9. **Marker identification**: Marker detection step comprises the following steps:
   a) Homography algorithm is applied to remove the projective perspective of the markers to obtain a frontal view of the rectangles.
   b) Internal code reading: The extracted rectangles are divided into 6×6 grids, of which 5×5 grids contain the marker id information.
   c) If the marker ID is valid, then it is confirmed that the detected region is an ArUco marker, then the corners are refined using subpixel algorithm.

In a nutshell, the output of the marker detection sub-method is the detected markers that includes marker ID numbers and four extracted corner points of each marker.

3.2.3. Model-based ArUco Mapping and Localization Algorithm

In this section, the marker mapping techniques of the implemented method are discussed in more detail. Based on the project specifications, the camera-robot configuration is defined to be eye-to-hand because of the risk of the camera lens occlusion by the spraying oils or water particles ejected from the spindle tool such as water-jets and drills, etc. Based on this fact, for pose estimation purposes, it is necessary to apply a map of markers instead of a single marker. The map of markers should be mountable on top of a robotic manipulator. In pose estimation problems based on the marker mapping techniques as the number of the map's markers increases, the pose estimation accuracy of the developed algorithm improves. Muñoz et al. [73] recommended having at least two detected-markers at each frame are necessary to have a precise pose estimation. To increase the number of markers in each frame of the camera field of view, the marker map geometrical model should represent more markers to the camera to increase the estimated pose accuracy of the developed system. Moreover, increasing of the number of markers also solves the ambiguity problem. The ambiguity problems commonly arise in the marker-based pose estimation problems. In this case, the z-axis of the camera and markers are

colinear, and two pose solutions are calculated. The other specification of the map of markers is that all markers in the map should be in the same sizes. Therefore, the map model is required to have regular faces (faces with equal lengths).

The evaluations of the 3D geometrical shapes and solids showed that the solution might exist among convex polyhedral shapes which are composed of regular polygons. The assessments of some the of polyhedron solids demonstrated that the truncated rhombicuboctahedron was an ideal candidate for this study based on the availability of fabrication facilities. Figure 3.4 (a), shows a representation of the rhombicuboctahedron designed using CAD software (CATIA), and Figure 3.4 (b) shows the manufacturing process using a 3D printer in the KOVAN lab.

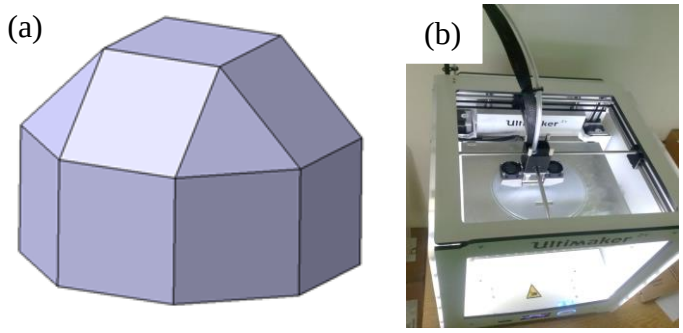

Figure 3.4. Designed Rhombicuboctahedron (a), and its 3D printing process (b)

The truncated rhombicuboctahedron has thirteen square-shaped and four isosceles-shaped triangles as its facets. According to Figure 3.5, at least four ArUco markers are detected from vertical perspective, and at most seven markers are detectable from side-view perspectives. The detection of several markers increases the robustness of the system in case of marker-occlusion and light-varying conditions. The dimensions of the Rhombicuboctahedron ArUco Mapper (RAM) are shown in Figure 3.6. The length of each square side is 11 cm, and the internal angles between facets are 135°.

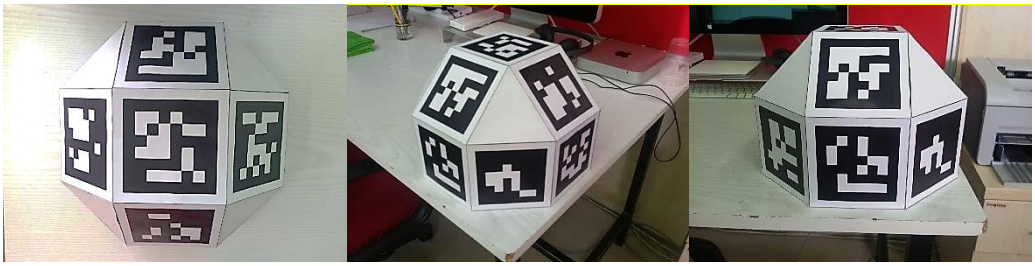

Figure 3.5. Different perspective of the Rhombicuboctahedron

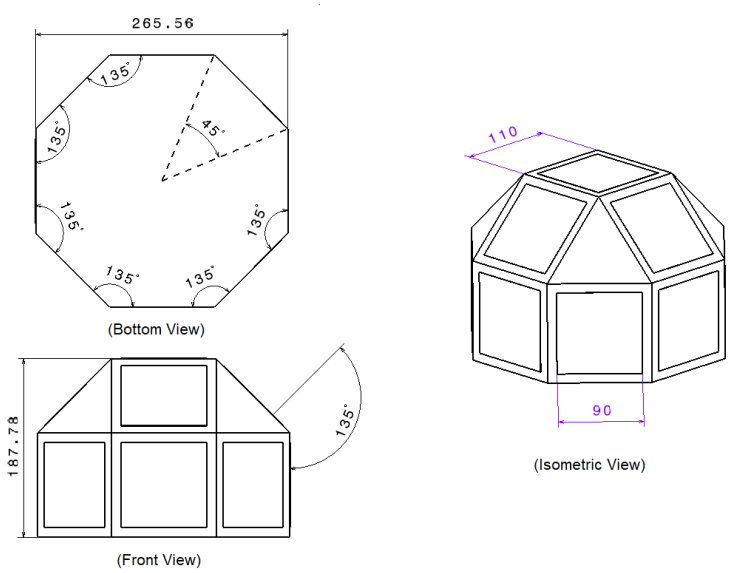

Figure 3.6. Geometrical specifications of the Rhombicuboctahedron ArUco Mapper

After fabricating the map of markers in rhombicuboctahedron shape, and sticking ArUco3 markers on its faces, it is required to prepare the map of the rhombicuboctahedron model, Figure 3.5. The aim of mapping of the model is to find the markers' pose (position and orientations) on the rhombicuboctahedron. The localized and mapped rhombicuboctahedron model are used as a supplementary pose estimation tool.

The used method for the mapping and localization of the proposed model is due to [73]. The Rhombicuboctahedron ArUco Mapping (RAM) principle is as follows:

Each marker on the Rhombicuboctahedron faces is depicted with four corners, Figure 3.2, and their poses are represented by T_m transformation. The T_m transforms a marker's corner points from the marker reference frame to the robot base frame. The set of the 13 markers attached on the Rhombicuboctahedron solid is depicted:

$$M = \{m_i\} \qquad for\ i = 0,1,2,\dots,12 \qquad (3.5)$$

In (3.5), m_i is the markers in the set. If a video sequence from the rhombicuboctahedron is recorded and the marker mapping algorithm is applied to each frame of the video sequence, then, the set of the all detected markers in frame *t* is:

$$f^t = \{i | i \in M\} \qquad (3.6)$$

Moreover, the pixel locations of the corner points of marker *i* is depicted by:

$$\omega_i^t = \left\{ \boldsymbol{u}_{i,k}^t \;\middle|\; \boldsymbol{u} \epsilon \mathbb{R}^2, k = 1 \dots 4 \right\} \qquad (3.7)$$

The essential condition for marker mapping using the method in [73] is that only frames containing at least two markers have to be processed for the marker mapping purposes. In mathematical terms:

$$|f^t| > 1 \qquad (3.8)$$

In the marker mapping technique developed in [73], the main aim is to minimize the squared of the re-projection error of a detected marker. The estimated re-projection error of the detected marker *i* in frame f^t is defined as:

$$e_i^t = \sum_j [\Psi(\delta, T^t, T_i . \boldsymbol{c}_j) - \boldsymbol{u}_{i,j}^t]^2 \qquad (3.9)$$

In (3.9), δ is the camera intrinsic values, and Ψ is the projection function. Thus, the total reprojection error in the whole set of frames is the function of the marker poses, frame poses, and camera intrinsic parameters:

$$e(T_1, \dots, T_M, T^1, \dots, T^N, \delta) = \sum_t \sum_{i \epsilon f^t} e_i^t \qquad (3.10)$$

Where M is the total number of markers (M=13), and N is the total numbers of frames.

The marker mapping algorithm runs offline on a video of Rhombicuboctahedron ArUco Mapper and creates a pose quiver with relative poses of the observed markers. Then, an initial pose graph is created, and corner refinement using Levenberg-Marquardt optimization is applied to reduce the reprojection errors (3.9) of the marker corners in all observed frames. The output of this sub-method or marker mapping is the Rhombicuboctahedron ArUco Mapper (RAM) that is used by the proposed pose estimation method, Figure 3.7.

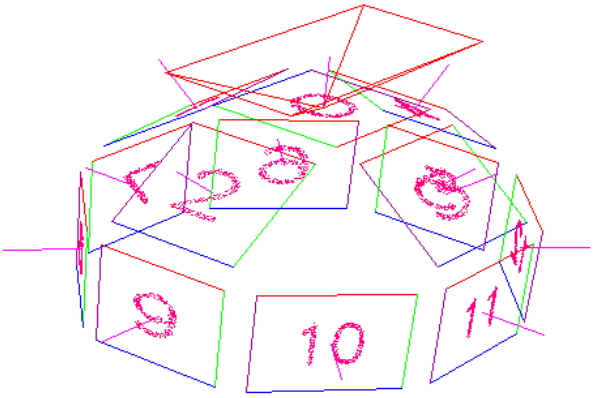

Figure 3.7. Mapped and localized Rhombicuboctahedron

After the Rhombicuboctahedron Marker Mapping, the map of the model is stored in a YML format and feed into the proposed pose estimation method developed in this study. Since, the model and pose information of the rhombicuboctahedron is known in advance, the proposed pose estimation technique in this study is categorized in model-based class. The flow-chart of marker mapping sub-method is shown in Figure 3. 8.

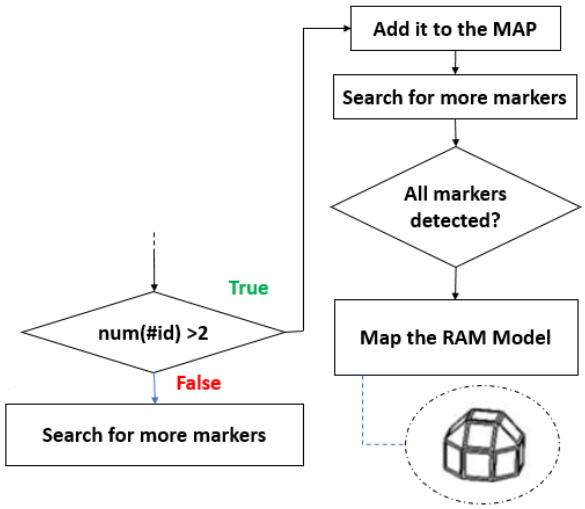

Figure 3.8. Marker mapping sub-method

3.2.4. RAM Pose Estimation With Respect to Camera (PnP Theory)

Fiducial markers are used to estimate the camera pose (position and orientation) in their coordinate frames. The camera pose estimation includes of estimation of the camera rotation (roll, pitch, and yaw) and translation (x, y, z) with respect to a marker reference coordinate such a marker coordinate. Camera pose estimation using ArUco markers is based on the *Perspective-n-Point* theory or PnP algorithm. The perspective-n-point algorithm estimates an object position and orientation with respect to a known reference frame from a given set of *n* 3D-points in the world and their corresponding 2D projections in the image. Pose estimation based on the points correspondence or point-matching has many applications in AR and 3D robot pose estimation problems. The direct linear transformation algorithm [45] is one of the earliest forms of the pose estimation problem based on the 3D-to-2D points matching theory.

In order to improve the accuracy of the pose estimation problem based on the 3D-to-2D point matching, several simplifications have been made to the DLT problem. The *Perspective-n-Point* (PnP) problem, is one of the optimized and robust pose algorithms proposed for the solving of the pose estimation problems. Camera matrix should be known when using the *Perspective-n-Point* algorithm.

The *Perspective-n-Point* approach have many different forms [74], [75], [76], [77], [78], and [79]. The straight forward form of the PnP algorithm requires three sets of point correspondences, called P3P, which produces four solutions to the pose estimation problem. However, by adding the fourth points, called P4P, the problem becomes more promising, which called *P4P* problem.

Riba et al. [80] have formulated and implemented the PnP problem in a way suitable for software implementations. According to the [81], Figure 3.9, given a set of 3D points P_i expressed in the 3D real-world and their 2D projections u_i onto the image plane, the pose of the camera including Rotation and Translation with respect to the world and the focal length *f* can be extracted from (3.11).

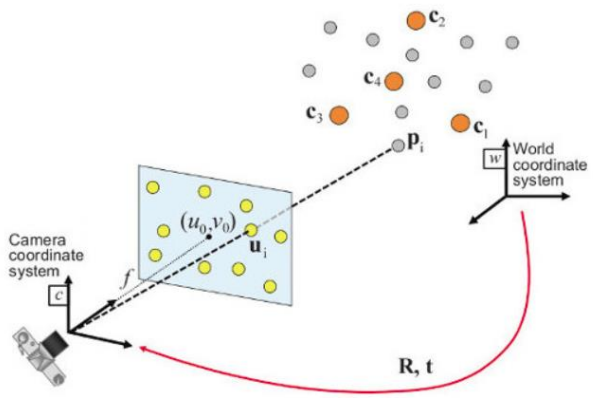

Figure 3.9. Perspective-n-Point theory [80]

$$sm' = A[R \quad | \quad t]M' \qquad (3.11)$$

$$s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \qquad (3.12)$$

In (3.11) and (3.12):

- X, Y, and Z are the coordinates of a 3D point in the world coordinate system,
- u and v are its projections on the image coordinate system,

- A is called the camera matrix, or camera intrinsic parameters (Appendix.B),
- c_x and c_y are the coordinates of camera principle point which is located at the image center and f_x and f_y are the focal lengths in the pixel units.

Thus, by solving the (3.12) the rotation and translation of each correspondence 3D-to-2D points can be calculated. The pose of the square-shape fiducial markers such as ArUco marker is estimated using the theory of the Perspective-n-Points algorithm. Each marker provides four points to the perspective-n-point algorithm for pose estimation operation.

After, Rhombicuboctahedron ArUco Mapping, coordinates of markers corner points determined in the RAM coordinate frame, Figure 3.10.

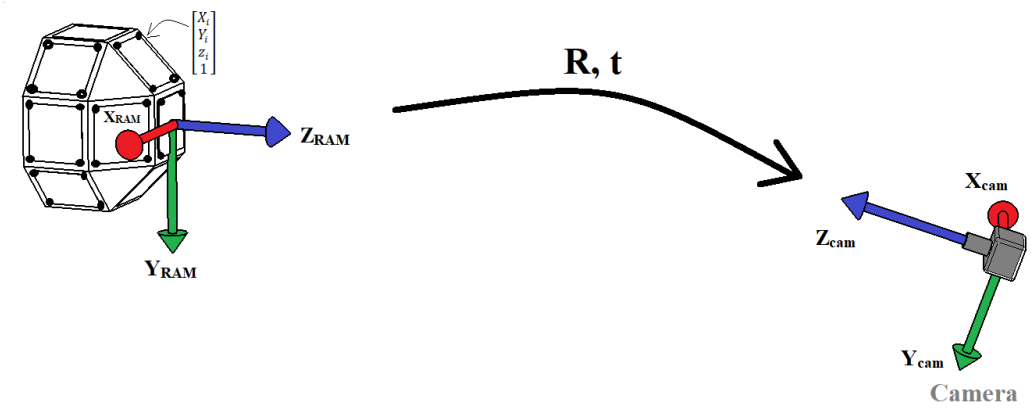

Figure 3.10. RAM pose estimation with respect to the camera using PnP theory

Based on the Perspective-n-point theory, to estimate an object pose with respect to the camera, at least three points on its surface should be known in a reference frame. Since each marker provides precisely four points in space, the thirteen markers provide fifty-two 3D-points in space. The three-dimensional coordinates of these points are known in the RAM coordinate frame. As mentioned earlier, at least four markers and at most seven markers can be detected by the camera during the pose estimation operations. Accordingly, at least sixteen points and at most twenty-eight 3D-points provided to the PnP algorithm for the RAM pose estimation with respect to the camera coordinate frame. (3.13) estimated the pose of the RAM with respect to the camera.

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} = \frac{s}{X^2+Y^2+Z^2+1} \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} [X \quad Y \quad Z \quad 1] \quad (3.13)$$

Where in (3.13) :

$$[\boldsymbol{R} \quad | \quad \boldsymbol{t}] = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \qquad (3.14)$$

By substitution of (3.14) into (3.13), the pose of the RAM with respect to the camera will be estimated:

$$[\boldsymbol{R} \quad | \quad \boldsymbol{t}] = \frac{s}{X^2+Y^2+Z^2+1} \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} [X \quad Y \quad Z \quad 1] \qquad (3.15)$$

3.2.5. TCP Pose Estimation with Respect to the Robot Base

Till this point, the pose of the RAM is estimated in the camera coordinate frames. However, the absolute accuracy and repeatability of the proposed pose estimation system are considered in this study, the proposed system should be able to calculate the TCP pose in the robot base frame. For this purpose, the homogenous transformations are carried out. Figure 3.8, shows the schematic view of the homogenous transformations from the TCP to the robot base frames.

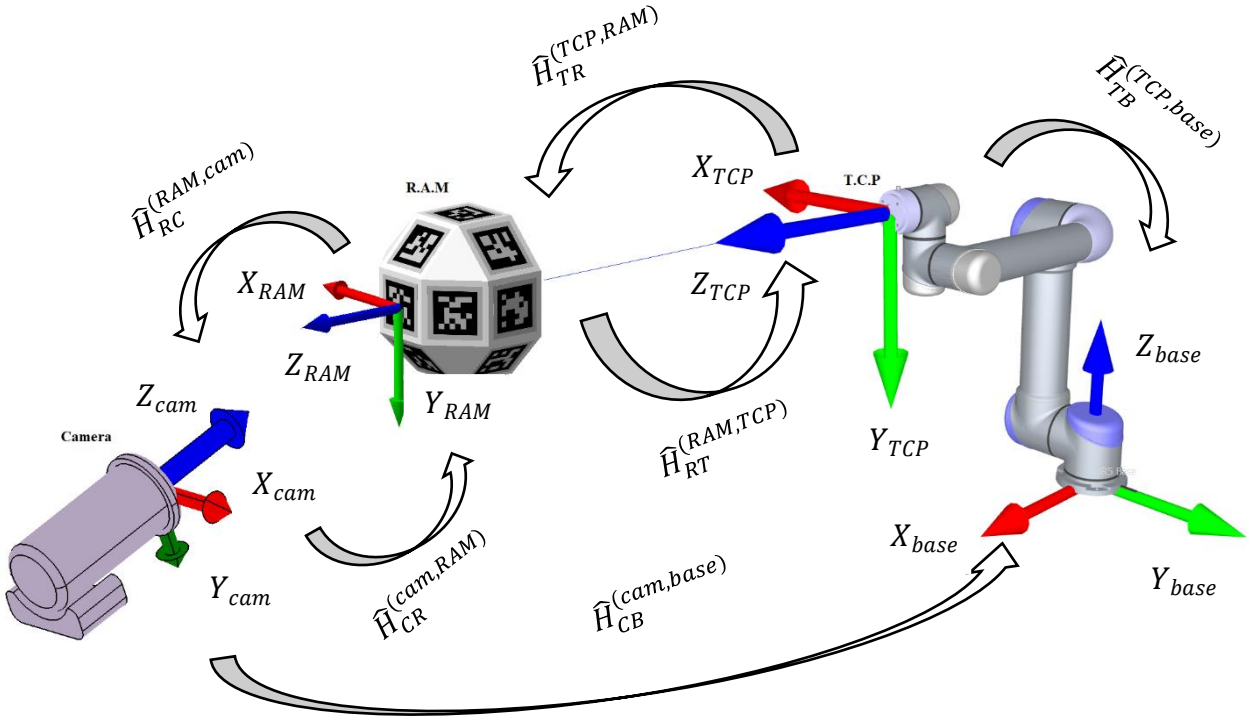

Figure 3.11. Homogenous transformations from TCP to Robot base

Where in Figure 3.11:

$\widehat{H}_{TR}^{(TCP,RAM)}$ is the homogenous transformation matrix from the TCP frame to RAM frame

$\widehat{H}_{RC}^{(RAM,cam)}$ is the homogenous transformation matrix from the RAM frame to the camera frame

$\widehat{H}_{CR}^{(cam,RAM)}$ is the homogenous transformation matrix from the camera frame to RAM frame

$\widehat{H}_{RT}^{(RAM,TCP)}$ is the homogenous transformation matrix from the RAM frame to TCP frame

$\widehat{H}_{TR}^{(TCP,RAM)}$ is the homogenous transformation matrix from the TCP frame to RAM frame

$\widehat{H}_{CB}^{(cam,base)}$ is the homogenous transformation matrix from the camera frame to the base frame.

$\widehat{H}_{TB}^{(TCP,base)}$ is the homogenous transformation matrix from the TCP frame to the base frame.

To estimates the TCP pose in the robot base frame, the following transformation carried to take the TCP point to the robot base frame.

$$\widehat{H}_{CB}^{(cam,base)}\widehat{H}_{RC}^{(RAM,cam)}\widehat{H}_{TR}^{(TCP,RAM)} \qquad (3.16)$$

In (3.16), $\widehat{H}_{TR}^{(TCP,RAM)}$ is calculated from the CAD models of the UR5, gripper and the Rhombicuboctahedron ArUco Mapper (RAM).

CHAPTER 4

EXPERIMENTS

In this chapter, experiment that are carried out to evaluate the performance of the developed method is explained. At first, the UR5 robotic arm, the web camera, camera calibration, and verification methods are discussed. Then, details of the standard experiments conducted to judge the performance of the implemented pose estimation method are presented.

4.1. The UR5 Robotic Manipulator

The robotic manipulator used in this study is one of the most popular, lightweight, highly flexible, and collaborative (COBOT) robots called UR5. Collaborative robots (COBOTs) provide a safely human-robot collaboration in industrial environments without any risks of injuries or health damages. The UR5 stands for the Universal Robotics, and postfix five refers to the robot payload which is 5kg [82]. The UR5 robotic manipulators are equipped with a force sensor that stops the robot when it hits an object or a human being.

4.1.1. The UR5 Anatomy

UR5 is a six-degrees-of-freedom robotic manipulator. It consists of a chain of the six rigid links connected by revolute joints as base, shoulder, elbow, wrist1, wrist2, and wrist3, Figure 4.1.

Figure 4.1. UR5 Robotic Manipulator [74]

One of the different features of the UR5 Robots compared to the other 6 DOF industrial robots is that their shoulder and elbow joints rotation axes are perpendicular to the rotation axis of the base joint. The main task of the wrist joint is to move the Tool Center Point (TCP) to a specified orientation and position. The technical drawing of the UR5 depicted in Figure 4.2.

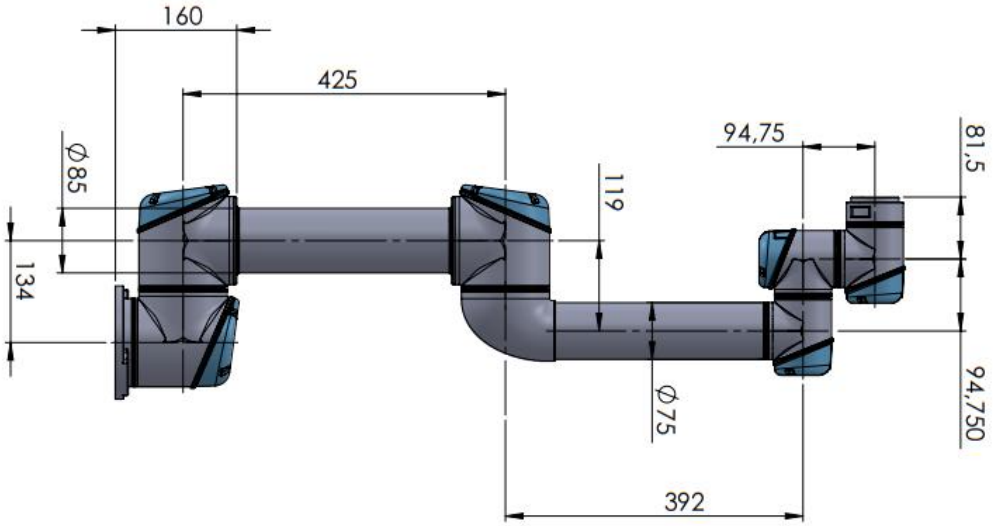

Figure 4.2. Technical drawing of the UR5 robot arm (all dimensions are in mm) [83]

4.1.2. The UR5 Specifications

The manipulating payload of the UR5 is 5kg/11lbs. UR5 itself weights 18.4kg. Figure 4.3 shows the UR5 workspace in top and side view. The reach distance of the arm is 850 mm or 33.5 inches. In general, its workspace is in spherical with the center at base and diameter as 1700 mm [83].

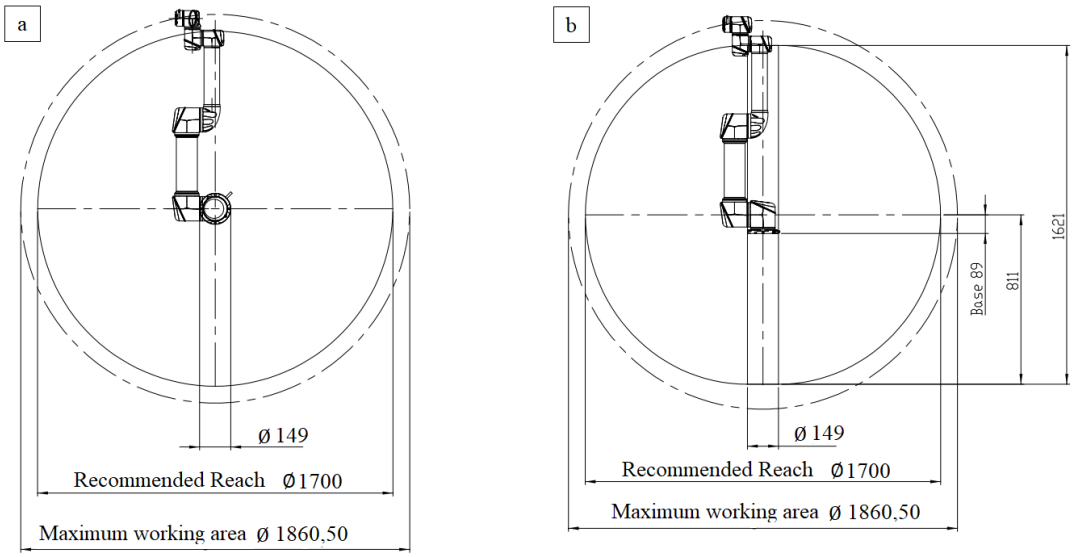

Figure 4.3. UR5 workspace [83]

The UR5 repeatability is given as 0.1 mm. Table 4.1 depicts some specifications of the UR5 robot manipulator [83].

Table 4.1. *The UR5 Specification*

Specification	
Payload	5 kg/ 11 lbs
Reach	850 mm / 33.5 in
Degrees of freedom	6
Programming	Polyscope GUI
Repeatability	±0.1 mm / ± 0.0039 (in)
Joint Working Range	± 360°

4.1.3. Communication with the UR5

The simplest way to communicate with the UR5 is through the PolyScope GUI controller called teach-pendant. Figure 4.4 shows the UR5 control panel. Low-level controlling can be done using this controller. Robot Operating System (ROS) is the other means of UR5 controlling [84]. Currently, ROS can be installed and used in Linux. It has wealthy libraries and forums to solve robotics problems. In some cases, when it is required to work in Windows systems, a MATLAB driver that is developed by Delft Center for Systems and Control (DCSC) is used [85]. The only shortcoming of this driver compared to ROS is that it does not have many libraries and programs to be used.

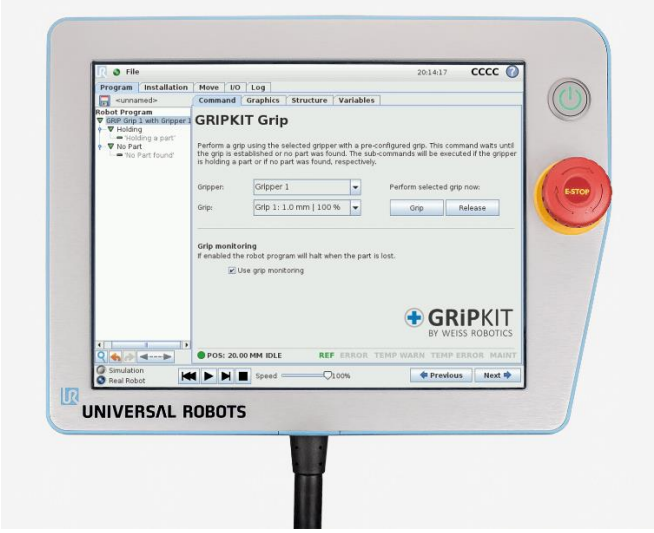

Figure 4.4. PolyScope of the UR5 robot *[86]*

4.1.4. Camera and Robot Configurations

There are three different robot-camera configurations, such as Eye-in-hand, Eye-to-hand, and active camera head. Based on the requirements of the thesis project definition, Eye-to-hand configuration is selected.

The camera is fixed at a point in the UR5 workspace. Figure 4.5 shows the camera and robot configuration used in this study.

Figure 4.5. Experimental test setup

4.2. The Lighting System

The illumination source is one of the main factors of a successful computer vision application. The high-quality lighting system produces a high contrast between the objects to be detected or measured and the background. The fluorescent lamps constitute the lighting system of the KOVAN laboratory where the experiment is conducted. Fluorescent, quartz halogen, and LED are the mostly used lighting systems for small and medium-scale machine vision applications [87]. The regular installation of the lamps in the array form with equal spaces between them created uniform optics in the experiment environment.

4.3. Experimental Workspace

Figure 4.6 shows a schematic view of the experiments' cubic workspace. The UR5 workspace dimension is in a cubic volume of 1860.50×1860.50×1860.50mm. The camera field of view covered the UR5 workspace volume.

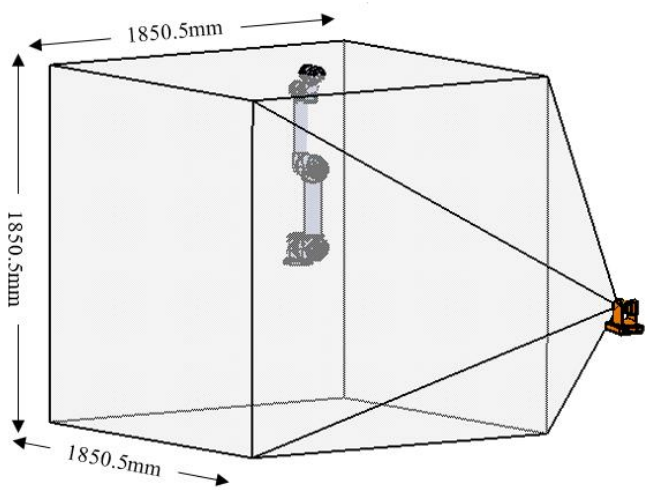

Figure 4.6. Camera field of view (FOV)

4.4. The UR5 Kinematic Modeling

In this section the forward kinematics (FWK) of the UR5 is discussed. The applied principles and notations for the forward kinematic (FWK) analysis are implemented using [88]. The UR5 zero structure condition (all joint angles are zero) is shown in Figure 4.7 (a) and its corresponding schematic view is depicted in Figure 4.7 (b). Joints and links are labelled according to Figure 4.7 (b).

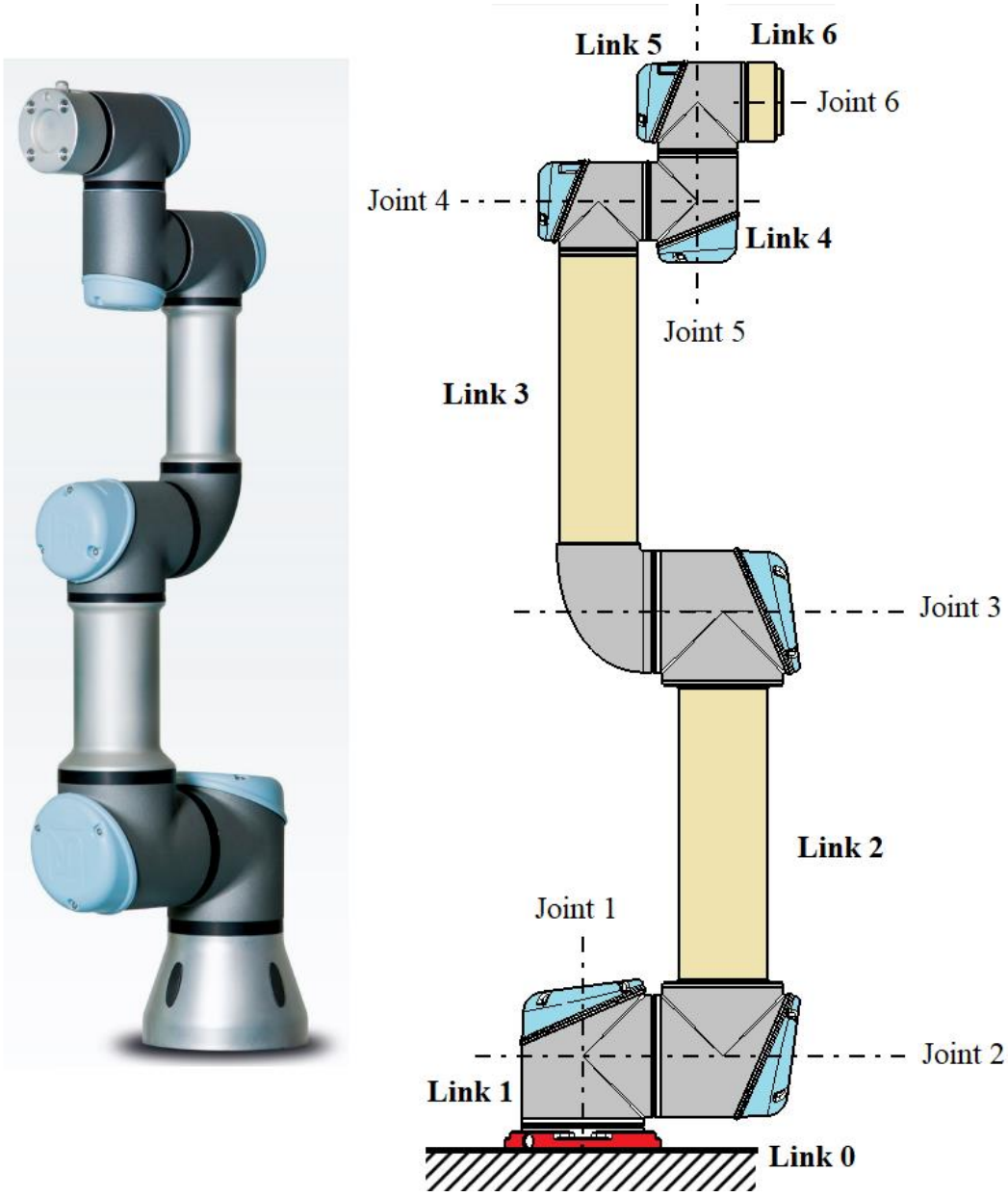

Figure 4.7. UR5 zero condition structure (a) and schematic view of robot (b)

Figure 4.8 shows the conventions and notations used in assigning reference frames to links.

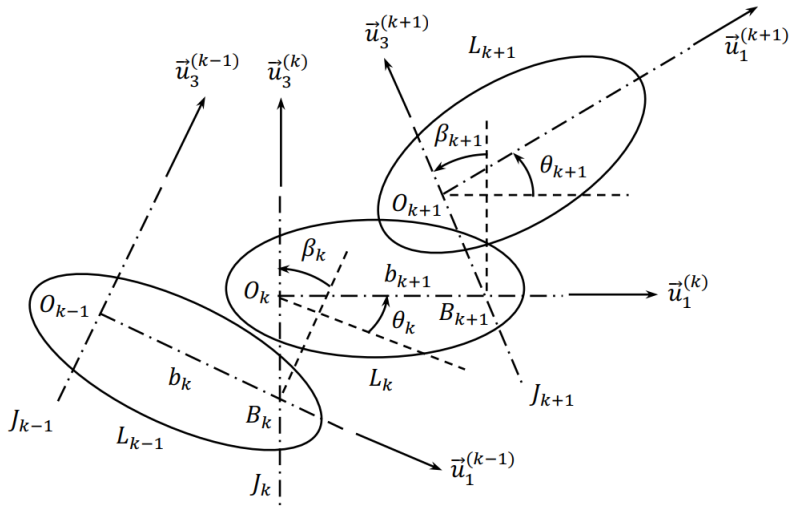

Figure 4.8. Convention used for forward kinematic analysis of UR5 arm [88]

The Denavit-Hartenberg (D-H) parameters principle and convention are described in the Appendix. D.

The Denavit-Hartenberg parameters of the UR5 (Figure 4.9) are:

$$\beta_1: \left[\vec{u}_3^{(0)} \rightarrow \vec{u}_3^{(1)}\right] @ \vec{u}_1^{(0)} = \frac{\pi}{2}$$

$$\beta_2: \left[\vec{u}_3^{(1)} \rightarrow \vec{u}_3^{(2)}\right] @ \vec{u}_1^{(1)} = 0$$

$$\beta_3: \left[\vec{u}_3^{(2)} \rightarrow \vec{u}_3^{(3)}\right] @ \vec{u}_1^{(2)} = 0$$

$$\beta_4: \left[\vec{u}_3^{(3)} \rightarrow \vec{u}_3^{(4)}\right] @ \vec{u}_1^{(3)} = \frac{\pi}{2}$$

$$\beta_5: \left[\vec{u}_3^{(4)} \rightarrow \vec{u}_3^{(5)}\right] @ \vec{u}_1^{(4)} = -\frac{\pi}{2}$$

$$\beta_6: \left[\vec{u}_3^{(5)} \rightarrow \vec{u}_3^{(6)}\right] @ \vec{u}_1^{(5)} = 0$$

$$d_1: \left[\vec{u}_1^{(0)} \rightarrow \vec{u}_1^{(1)}\right] along\ \vec{u}_3^{(1)} = 0.89159$$

$$d_2: \left[\vec{u}_1^{(1)} \rightarrow \vec{u}_1^{(2)}\right] along\ \vec{u}_3^{(2)} = 0$$

$$d_3: \left[\vec{u}_1^{(2)} \rightarrow \vec{u}_1^{(3)}\right] along\ \vec{u}_3^{(3)} = 0$$

$$d_4: \left[\vec{u}_1^{(3)} \rightarrow \vec{u}_1^{(4)}\right] along\ \vec{u}_3^{(4)} = 0.10915$$

$$d_5: \left[\vec{u}_1^{(4)} \rightarrow \vec{u}_1^{(5)}\right] along\ \vec{u}_3^{(5)} = 0.09465$$

$$d_6: \left[\vec{u}_1^{(5)} \rightarrow \vec{u}_1^{(6)}\right] along\ \vec{u}_3^{(6)} = 0.0823$$

$$b_1: \left[\vec{u}_3^{(0)} \rightarrow \vec{u}_3^{(1)}\right] along\ \vec{u}_1^{(0)} = 0$$

$$b_2: \left[\vec{u}_3^{(1)} \rightarrow \vec{u}_3^{(2)}\right] along\ \vec{u}_1^{(1)} = -0.425$$

$$b_3: \left[\vec{u}_3^{(2)} \rightarrow \vec{u}_3^{(3)}\right] along\ \vec{u}_1^{(2)} = -0.39225$$

$$b_4: \left[\vec{u}_3^{(3)} \rightarrow \vec{u}_3^{(4)}\right] along\ \vec{u}_1^{(3)} = 0$$

$$b_5: \left[\vec{u}_3^{(4)} \rightarrow \vec{u}_3^{(5)}\right] along\ \vec{u}_1^{(4)} = 0$$

$$b_6: \left[\vec{u}_3^{(5)} \rightarrow \vec{u}_3^{(6)}\right] along\ \vec{u}_1^{(5)} = 0$$

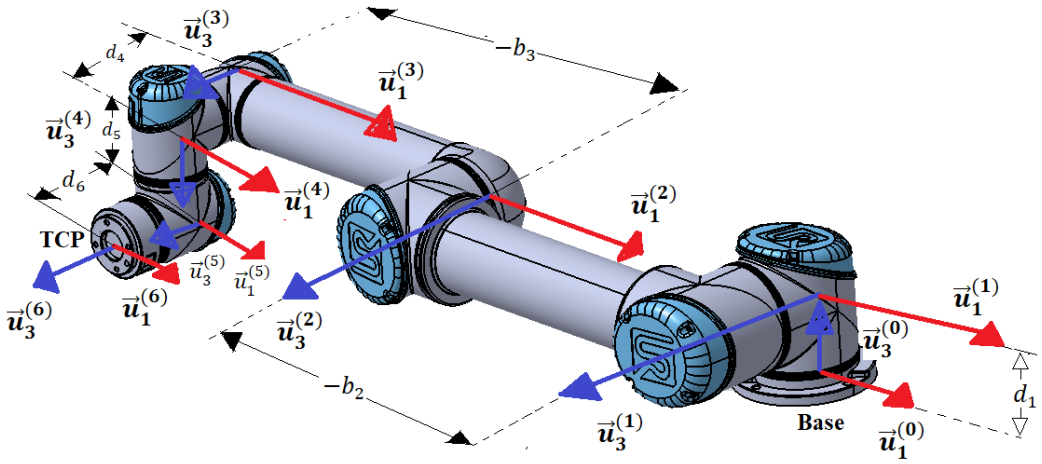

Figure 4.9. Frame assignment of the UR5 links based on the notation convention

Table 4.2 shows the D-H parameters of the UR5.

Table 4.2. *Denavit-Hartenberg Parameters of the UR5*

D-H Table	*Link 1*	*Link 2*	*Link 3*	*Link 4*	*Link 5*	*Link 6*
β_k	$\frac{\pi}{2}$	0	0	$\frac{\pi}{2}$	$-\frac{\pi}{2}$	0
b_k	0	-0.425	-0.39225	0	0	0
θ_k *or* δ_k	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6
s_k *or* d_k	0.089159	0	0	0.10915	0.09465	0.0823
O_k	O	S	E	W1	W2	P

4.4.1. The UR5 Forward Kinematic

Using extracted D-H parameters, the homogenous transformation matrix for pose estimation of the TCP in the robot base reference frame is:

$$\hat{H}_{06}^{(0,6)} = \hat{H}_{01}^{(0,1)}\hat{H}_{12}^{(1,2)}\hat{H}_{23}^{(2,3)}\hat{H}_{34}^{(3,4)}\hat{H}_{45}^{(4,5)}\hat{H}_{56}^{(5,6)} \qquad (4.1)$$

Figure 4.10. shows the schematic representation of the transformation operations (4.1) from the TCP to the robot base reference frame.

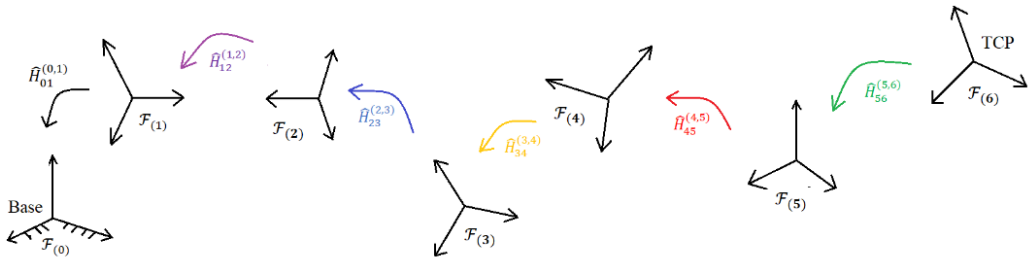

Figure 4.10. Schematic representation of transformation operation from TCP to base frame

The principle of the applied forward kinematic in this study is explained in the Appendix. E.

In (4.1), each transformation matrix H is defined as:

$$\widehat{H}_{k-1k}^{(k-1,k)} = \begin{bmatrix} \hat{C}^{(k-1,k)} & | & \bar{r}_{k-1k}^{(k-1)} \\ ----- & | & --- \\ \bar{0}^t & | & 1 \end{bmatrix} \qquad (4.2)$$

In (4.1), $\hat{C}^{(k-1,k)}$ is the 3×3 orthonormal orientation matrix, which specifies the orientation of the link frame (k) into same orientation of link frame (k-1), and $\bar{r}_{k-1k}^{(k-1)}$ is the position vector from the origin of the frame (k-1) into frame (k) written in terms of the unit vectors of the reference frame (k-1).

The orientation matrix for the two successive links is [88]:

$$\hat{C}^{(k-1,k)} = e^{\tilde{u}_1\beta_k}e^{\tilde{u}_3\theta_k} \qquad (4.3)$$

The transformation matrices between two successive links are computed :

$$\hat{C}^{(0,1)} = e^{\tilde{u}_1\beta_1}e^{\tilde{u}_3\theta_1} = e^{\tilde{u}_1\frac{\pi}{2}}e^{\tilde{u}_3\theta_1} = e^{\tilde{u}_1\frac{\pi}{2}}e^{\tilde{u}_3\theta_1}\, e^{-\tilde{u}_1\frac{\pi}{2}}e^{\tilde{u}_1\frac{\pi}{2}} = e^{\widetilde{-u}_2\theta_1}e^{\tilde{u}_1\frac{\pi}{2}} \qquad (4.4)$$

$$\hat{C}^{(1,2)} = e^{\tilde{u}_1\beta_2}e^{\tilde{u}_3\theta_2} = e^{\tilde{u}_3\theta_2} \qquad (4.5)$$

$$\hat{C}^{(2,3)} = e^{\tilde{u}_1\beta_3}e^{\tilde{u}_3\theta_3} = e^{\tilde{u}_3\theta_3} \qquad (4.6)$$

$$\hat{C}^{(3,4)} = e^{\tilde{u}_1\beta_4}e^{\tilde{u}_3\theta_4} = e^{\tilde{u}_1\frac{\pi}{2}}e^{\tilde{u}_3\theta_4} = e^{\tilde{u}_1\frac{\pi}{2}}e^{\tilde{u}_3\theta_4}e^{-\tilde{u}_1\frac{\pi}{2}}e^{\tilde{u}_1\frac{\pi}{2}} = e^{\widetilde{-u}_2\theta_4}e^{\tilde{u}_1\frac{\pi}{2}} \qquad (4.7)$$

$$\hat{C}^{(4,5)} = e^{\tilde{u}_1\beta_5}e^{\tilde{u}_3\theta_5} = e^{-\tilde{u}_1\frac{\pi}{2}}e^{\tilde{u}_3\theta_5} = e^{-\tilde{u}_1\frac{\pi}{2}}e^{\tilde{u}_3\theta_5}e^{\tilde{u}_1\frac{\pi}{2}}e^{\widetilde{-u}_1\frac{\pi}{2}} = e^{\tilde{u}_2\theta_5}e^{\widetilde{-u}_1\frac{\pi}{2}} \qquad (4.8)$$

$$\hat{C}^{(5,6)} = e^{\tilde{u}_1\beta_6}e^{\tilde{u}_3\theta_6} = e^{\tilde{u}_3\theta_6} \qquad (4.9)$$

Relative positions of the successive links are computed:

$$\vec{r}_{01} = d_1\vec{u}_3^{(0)} \qquad (4.10)$$

The vector representation of (4.10) in the base link frame is computed as:

$$\bar{r}_{01}^{(0)} = 0.089159\bar{u}_3^{(0/0)} = 0.089159\bar{u}_3 \qquad (4.11)$$

The relative position of the link 2 through link 6 with respect to their previous links' frame are:

$$\vec{r}_{12} = -b_2\vec{u}_1^{(1)} \qquad (4.12)$$

$$\bar{r}_{12}^{(1)} = -b_2\bar{u}_1^{(1/1)} = -b_2\bar{u}_1 \qquad (4.13)$$

$$\vec{r}_{23} = -b_3\vec{u}_1^{(2)} \qquad (4.14)$$

$$\bar{r}_{\mathbf{23}}^{(\mathbf{2})} = -b_3\bar{u}_1^{(2/2)} = -b_3\bar{u}_1 \qquad (4.15)$$

$$\vec{r}_{34} = d_4\vec{u}_3^{(3)} \qquad (4.16)$$

$$\bar{r}_{\mathbf{34}}^{(\mathbf{3})} = d_4\bar{u}_{\mathbf{3}}^{(\mathbf{3}/\mathbf{3})} = d_4\bar{u}_{\mathbf{3}} \qquad (4.17)$$

$$\vec{r}_{45} = d_5\vec{u}_3^{(4)} \qquad (4.18)$$

$$\bar{r}_{\mathbf{45}}^{(\mathbf{4})} = d_5\bar{u}_3^{(4/4)} = d_5\bar{u}_3 \qquad (4.19)$$

$$\vec{r}_{56} = d_6\vec{u}_3^{(5)} \qquad (4.20)$$

$$\bar{r}_{\mathbf{56}}^{(\mathbf{5})} = d_6\bar{u}_{\mathbf{3}}^{(\mathbf{5}/\mathbf{5})} = d_{\mathbf{6}}\bar{u}_{\mathbf{3}}^{(\mathbf{5}/\mathbf{5})} = d_{\mathbf{6}}\ \bar{u}_3 = d_{\mathbf{6}}\bar{u}_3 \qquad (4.21)$$

According to (4.2), the homogenous transformation matrix between successive links starting from base link is:

$$\widehat{H}_{01}^{(0,1)} = \left[\begin{array}{c|c} \hat{C}^{(0,1)} & \bar{r}_{01}^{(0)} \\ \hline \bar{0}^t & 1 \end{array}\right] = \left[\begin{array}{c|c} e^{\widetilde{u}_2\theta_1}e^{\widetilde{u}_1\frac{\pi}{2}} & \bar{r}_{01}^{(0)} \\ \hline \bar{0}^t & 1 \end{array}\right]$$

$$\widehat{H}_{01}^{(0,1)} = \begin{bmatrix} cos\theta_1 & -sin\theta_1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ sin\theta_1 & cos\theta_1 & 0 & 0.089159 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad (4.22)$$

$$\widehat{H}_{12}^{(1,2)} = \left[\begin{array}{c|c} \hat{C}^{(1,2)} & \bar{r}_{12}^{(1)} \\ \hline \bar{0}^t & 1 \end{array}\right] = \left[\begin{array}{c|c} e^{\widetilde{u}_3\theta_2} & \bar{r}_{12}^{(1)} \\ \hline \bar{0}^t & 1 \end{array}\right]$$

$$\widehat{H}_{12}^{(1,2)} = \begin{bmatrix} cos\theta_2 & -sin\theta_2 & 0 & b_2 \\ sin\theta_2 & cos\theta_2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.23)$$

$$\widehat{H}_{23}^{(2,3)} = \left[\begin{array}{ccc|c} & \hat{C}^{(2,3)} & & \bar{r}_{23}^{(2)} \\ & - - - - - & & - - - \\ & \bar{0}^t & & 1 \end{array}\right] = \left[\begin{array}{ccc|c} & e^{\tilde{u}_3\theta_3} & & \bar{r}_{23}^{(2)} \\ & - - - - - & & - - - \\ & \bar{0}^t & & 1 \end{array}\right]$$

$$\widehat{H}_{23}^{(2,3)} = \begin{bmatrix} cos\theta_3 & -sin\theta_3 & 0 & b_3 \\ sin\theta_3 & cos\theta_3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.24)$$

$$\widehat{H}_{34}^{(3,4)} = \left[\begin{array}{ccc|c} & \hat{C}^{(3,4)} & & \bar{r}_{34}^{(3)} \\ & - - - - - & & - - - \\ & \bar{0}^t & & 1 \end{array}\right] = \left[\begin{array}{ccc|c} & e^{\tilde{u}_2\theta_4}e^{\tilde{u}_1\frac{\pi}{2}} & & \bar{r}_{34}^{(3)} \\ & - - - - - & & - - - \\ & \bar{0}^t & & 1 \end{array}\right]$$

$$\widehat{H}_{34}^{(3,4)} = \begin{bmatrix} cos\theta_4 & -sin\theta_4 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ sin\theta_4 & cos\theta_4 & 0 & d_4 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.25)$$

$$\widehat{H}_{45}^{(4,5)} = \left[\begin{array}{ccc|c} & \hat{C}^{(4,5)} & & \bar{r}_{45}^{(4)} \\ & - - - - - & & - - - \\ & \bar{0}^t & & 1 \end{array}\right] = \left[\begin{array}{ccc|c} & e^{\tilde{u}_2\theta_5}e^{\tilde{u}_1\frac{\pi}{2}} & & \bar{r}_{45}^{(4)} \\ & - - - - - & & - - - \\ & \bar{0}^t & & 1 \end{array}\right]$$

$$\widehat{H}_{45}^{(4,5)} = \begin{bmatrix} cos\theta_5 & -sin\theta_5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -sin\theta_5 & -cos\theta_5 & 0 & d_5 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.26)$$

$$\widehat{H}_{56}^{(5,6)} = \left[\begin{array}{ccc|c} & \hat{C}^{(5,6)} & & \bar{r}_{56}^{(5)} \\ & - - - - - & & - - - \\ & \bar{0}^t & & 1 \end{array}\right] = \left[\begin{array}{ccc|c} & e^{\tilde{u}_3\theta_6} & & \bar{r}_{56}^{(5)} \\ & - - - - - & & - - - \\ & \bar{0}^t & & 1 \end{array}\right]$$

$$\widehat{H}_{56}^{(5,6)} = \begin{bmatrix} cos\theta_6 & -sin\theta_6 & 0 & 0 \\ sin\theta_6 & cos\theta_6 & 0 & 0 \\ 0 & 0 & 1 & d_{\mathbf{6}} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.27)$$

The homogenous transformation from the TCP to the robot base frame is extracted by the substitution of (4.22) -(4.27) in (4.1).

4.5. Geometric Camera Modeling and Calibration

Table 4.2 shows the camera specifications used in this study. The camera models is described in Appendix. A. The camera's specifications are shown in Figure 4.11.

Table 4.3. *Camera specifications*

Camera.	
Model	Pinhole
Image Sensing Capacity	16 Megapixel
Video Sensing Capacity	2Megapixel
Resolutions	480, 720, 1080
FPS	30

Figure 4.11. Used camera in this study [89]

Geometric camera calibration calculates the parameters of the lens and image sensor of a camera. The camera calibration procedure is a necessary procedure to correct the lens distortion effects in image metrology applications. Camera calibration procedure ends up with the calculation of the intrinsic, extrinsic, and distortion coefficients.

In order to estimate the camera parameters, it is required to have 3D world points and their corresponding 2D image points. In Appendix B, camera calibration is discussed

in detail. Some common patterns are used for camera calibration. Figure 4.12, shows one of these patterns called the checkerboard. ChArUco patterns was developed for the camera calibration purposes. In this study, ChArUco pattern, which is a combination of the Chessboard and ArUco patterns, is used for camera calibration because of its ability to compute the camera parameters including intrinsic, extrinsic, and distortion parameters, precisely and reliably.

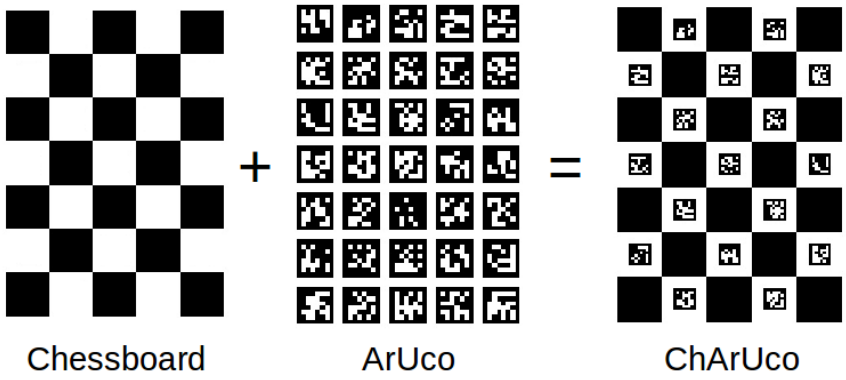

Figure 4.12. Chessboard (a), ArUco (b), and ChArUco calibration patterns (c)

ChArUco board is made up of multiple markers. This property enables them to calculate the camera matrix even in the case of partially occluded calibration board. In other words, in the cases that only some points can be observable or partially observable, the calibration code does not fail. About 250 images of the ChArUco board are acquired at different poses and locations, Figure 4.12 (c). The camera resolution is set up to 1280 × 720. Tables 4.4 and 4.5 show the results of the calibration parameters including the focal lengths, principal point, and distortion coefficients calculated during camera calibration.

Table 4.4. *Estimated camera distortion coefficients from ChArUco camera calibration*

Resolution	k_1	k_2	k_3	p_1	p_2
1280 × 720	0.1056	-0.2703	0.4166	0.0072	-0.0038

Table 4.5. *Estimated camera matrix elements from ChArUco camera calibration (mm)*

Resolution	f_x	f_y	p_x	p_y
1280 × 720	1149.87	1147.74	649.2089	372.3689

Reprojection error is considered as an algebraic criterion to judge the accuracy of camera calibration procedure. Reprojection error is estimated for every image. Hence, in a camera calibration operation which uses several images, the mean of the reprojection error used as a criterion for judging the accuracy of the calibration process. The calculated reprojection error was calculated as 0.211 that proved the success of the calibration procedure.

4.6. Pose Estimation Performance Evaluation

The implemented pose estimation system can be considered as a visual-based pose measurement sensor. Therefore, it is necessary to evaluate the pose estimation performance of the system based on some performance evaluation metrics and criteria. The industrial robots pose performance metrics suggested by ISO 9283::1998 standard (industrial robot performance) is absolute pose accuracy and repeatability.

In pose performance evaluations, pose accuracy is defined as the maximum translational and rotational errors between any two points in the robot workspace. On the other hand, repeatability or precision is defined as the error between several successive attempts to reach an actual pose. Repeatability is one of the most critical factors in evaluation of the robotic machines and sensors' performances. Figure 4.13, shows the graphical representation of the repeatability and accuracy which helps the intuition of these concepts.

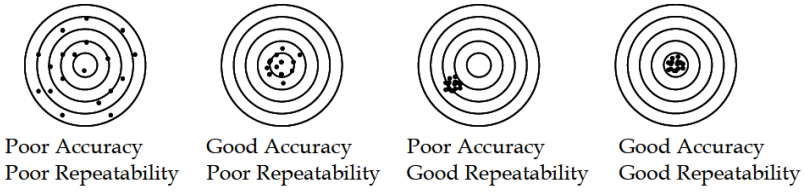

Figure 4.13. Repeatability and accuracy graphical concepts [12]

It is surprising to tell that repeatability and accuracy are different terms. According to Figure 4.13 a system can be repeatable but not accurate. In terms of pose measurements, repeatability is a measure of how close are the successive pose measurements by the purposed pose estimation system when the target (TCP) is placed in a specified point in robot workspace. In order to evaluate the pose measurement performance of the proposed system in this study, some techniques and methods are taken from ISO 9283:1998 standards. However, it is worth to mention that ISO 9283 is defined for the industrial robot's pose performance evaluation, not for evaluation of the pose measurement system performance. In other words, ISO 9283:1998 is a series of international standards dealing with industrial robots. It is worth to mention all the measurements are absolute or estimated in the robot base coordinate frame.

4.6.1. Pose Measurement Terms and Definitions

Before discussing the test procedures, it makes sense to define some technical terms used in this study [90]:

Estimated Pose (**Cluster)**: A set of the TCP absolute pose measurements estimated by the proposed system.

Barycenter: For a cluster of n-points ((x_j, y_j, z_j) for j = 1, 2,….,n,) the barycenter of the cluster of n-points is a point whose coordinate is the mean of the cluster.

The ground-truth pose (actual pose): The ground-truth pose is a set of points in the robot workspace whose pose knowledge including position and orientation are known in advance. The ground-truth poses are used for the performance assessment of the proposed pose estimation system.

4.6.1.1. Pose Estimation Accuracy

Pose estimation accuracy of the proposed system is defined as the measurement ability of the developed system in minimizing the deviations between a ground-truth pose (actual TCP pose) and the barycenter of the estimated poses (cluster) calculated by the proposed system.

In this study, two kinds of pose accuracies are evaluated:

a) Position accuracy: the difference between the ground truth pose (TCP actual position) and the barycenter of the measured positions of the TCP in the robot base.

b) Orientation accuracy: the difference between the orientation of the ground-truth pose (TCP actual orientation) and the average of the estimated orientations of the TCP in the robot base.

4.6.1.2. Pose Estimation Repeatability

In the visual-based pose estimation systems repeatability or test-retest reliability is a function of the time. In other words, it is a time-lapsing process. Repeatability is defined as the consistency or agreement between the successive measurements of the same quantity performed under the same measurement conditions [91]. In this study, the pose measurement repeatability of the proposed system is defined as the closeness of estimated poses of the ground-truth poses (TCP actual pose) over a specified period. The following conditions are fulfilled in the assessment of the pose's measurement repeatability of the proposed system:

- The same camera as the measurement instrument
- The same computer as the processor
- The same location for the camera
- The same location for the UR5 robot
- Repetition over a short and continuous period (camera not interrupted)

4.7. Absolute Pose Estimation Performance Evaluation Experiments

In this section, the test procedure used for the evaluation of the pose estimation performance of the proposed system is discussed. As mentioned earlier, all the measurements are carried out with respect to the robot base reference frame.

4.7.1. Selection of the Ground-truth Poses (Calibrated Points)

In order to judge the performance of the measurement system, it is required to have some ground-truth points in which their pose information (position and orientation) are specified in advance. For this purpose, five points on a plane are selected as the ground-truth poses (TCP actual poses). The selection criteria for the ground truth pose is inspired by a test cube offered by the ISO 9283 standard. ISO 9283:1998 have suggested four test cubes for the industrial robotic manipulators pose performances evaluations, Figure 4.14. According to this standard, the cube should be located in the robot workspace; moreover, it should occupy the maximum volume of the robot workspace with the edges parallel to the base coordinate system.

According to Figure 4.14, test cubes are distinguished from each other based on the configuration of the measuring plane (cross-hatching plane). Based on this criterion, test cube naming conventions are:

a) C_1-C_2-C_7-C_8
b) C_2-C_3-C_8-C_5
c) C_3-C_4-C_5-C_6
d) C_4-C_1-C_6-C_7

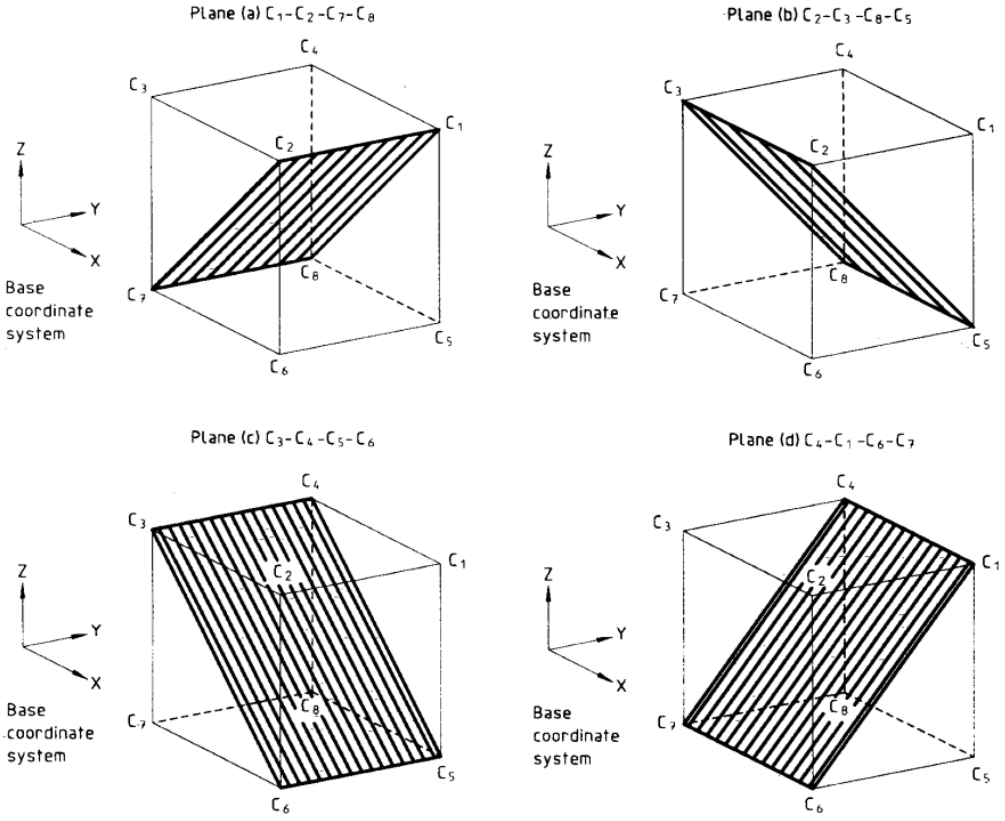

Figure 4.14. Test cubes offered by the ISO 9283

The test cube C3-C4-C5-C6 is selected for the pose measurement evaluation of the proposed system in this study. The main criterion for this selection was based on the UR5 workspace limitation in the KOVAN research lab. Five points are placed on the diagonals of the measuring plane called P_1 to P_5. The points P_1 to P_5 are called the ground-truth points (TCP actual pose). It assumed that the pose knowledge of the ground-truth points with respect to the UR5 base coordinate are known in advance. The ground-truth poses are estimated from the UR5 forward kinematic. In some industrial robotic manipulators, ground-truth points pose have been provided to customers in terms of the joint variables to be used in calibration applications. According to the ISO 9283 specifications, points P_2 to P_5 are located at a distance of (10±2)L% from the ends of the diagonals and P_1 is at the intersection point of the diagonals of the measuring plane

[86]. Figure 4.15, shows a graphical representation of the test cube and calibrated points used for the pose accuracy and repeatability evaluation of the proposed system.

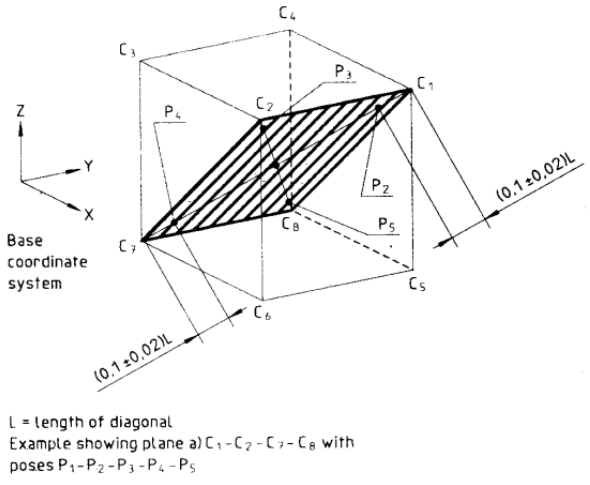

Figure 4.15. Used test cube for pose performance evaluations

Based on the UR5 workspace limitations in the laboratory environment, the maximum side length of the test cube is computed as the 1240 mm, Figure 4.16.

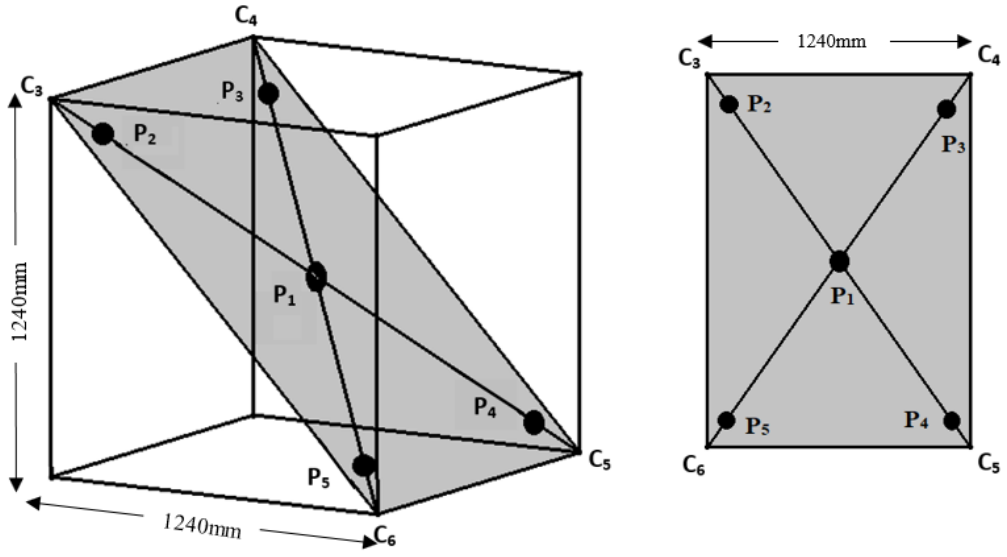

Figure 4.16. The applied test cube specification in this study

The ground-truth point's coordinates (TCP actual poses) is expressed in the UR5 base frame as:

$$P_i(x_i, y_i, z_i, a_i, b_i, c_i) \qquad (4.28)$$

Where in (4.28), x_i, y_i, and z_i are the position coordinates of the ground-truth points and a_i, b_i, and c_i are the orientations of the ground-truth points in the UR5 base frame. The position coordinates are expressed in the form of the distance (millimeter) and orientation expressed in the form of the angles in radians or degrees unit.

$$P_1(-79.32, -367.29, 241.58, 0.5987, -2.2848, 2.3725) \qquad (4.29)$$

$$P_2(-329.1, -117.34, 491.41, 0.6131, -1.6377, 1.1613) \qquad (4.30)$$

$$P_3(170.88, -117.34, 491.41, 0.8382, -2.0172, 2.8499) \qquad (4.31)$$

$$P_4(170.88, -616.75, -8.32, 0.0133, -2.2143, 2.2143) \qquad (4.32)$$

$$P_5(-329.1, -616.75, -8.32, 0.404, 2.0036, -2.0041) \qquad (4.33)$$

4.7.2. TCP Absolute Pose Estimation Accuracy Experiment

The experimental setup is depicted in Figure 4.17. First, the Rhombicuboctahedron ArUco Mapper (RAM) is mounted on the UR5 gripper, and then the UR5 moves its TCP to the point P_1.

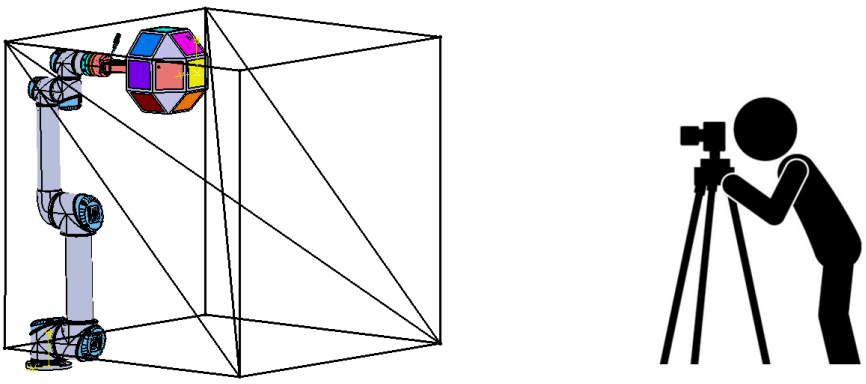

Figure 4.17. The graphical representation of the test setup

At point P_1, 30 pose measurements are recorded by the proposed pose estimation system. Then, the UR5 moves through points P_2 to P_5 and at each point, 30 pose estimation are recorded by the proposed system. In total, about 150 pose estimation of RAM is recorded by the system in the camera reference frame. Each pose estimation includes six pose parameters, three translational, x, y, and z, and three orientations a, b, and c. In order to evaluate the pose estimation performance in the robot base frame, it is required to transform the estimated poses into the robot base frame. For this purpose, the coordinate frame transformations are carried out as shown in Figure 3. 8 (Section 3.2.5).

Table 4.6 shows the coordinates of the ground-truth poses (TCP actual poses) of P_1 to P_5 in joint space, and Table 4.7 shows their corresponding pose knowledge of the ground-truth points in the robot base frame expressed in cartesian coordinate representation.

Pose knowledge of the TCP is calculated using forward kinematic implemented in MATLAB. Table 4.8 shows the pose of the ground truth points in the camera coordinate frame.

Table 4.6. *Pose knowledge of the ground-truth poses (TCP actual pose) in joint space*

Point No.	θ_1 (°)	θ_2 (°)	θ_3 (°)	θ_4 (°)	θ_5 (°)	θ_6 (°)
P_1	-90.5	-76.99	-97.13	-4.62	113.6	39.55
P_2	-149.95	-90.28	-43.25	-49.21	125.04	-98.85
P_3	30.42	-80.31	-46.36	9.74	18.97	-256.45
P_4	-61.09	-117.4	-88.52	25.53	60.77	0.06
P_5	-111.58	-126.08	-74.47	20.85	111.345	27.41

Table 4.7. *Pose knowledge of the ground-truth poses (TCP actual pose) in robot base frame*

Point No.	x (mm)	y (mm)	z (mm)	a (°)	b (°)	c (°)
P_1	-79.41	-367.1	640	-14.2781	-101.963	89.9830
P_2	-328.9	-118.7	890.2	-56.3847	77.0456	-44.3927
P_3	172.7	-115.4	890.2	-65.6265	-33.4149	-21.1249
P_4	169.4	-615.7	389.4	-47.8248	-96.0391	95.7527
P_5	-328.9	-616	390.2	-27.3988	-111.7841	112.4429

Table 4.8. *The ground truth points pose in the camera coordinate frame*

Point No.	Distance (mm)	a (°)	b (°)	c (°)
P_1	1442.9	50.6486	141.5528	9.4372
P_2	1803.6	123.43	-111.52	-20.972
P_3	1914.5	112.42	73.644	30.124
P_4	1151.3	77.921	153.95	1.1959
P_5	1168.8	41.134	167.86	-0.99665

The UR5 robot configurations at each ground-truth points P_1 to P_5 represented in Figure 4.18. The simulations have performed in Robodk (Free edition).

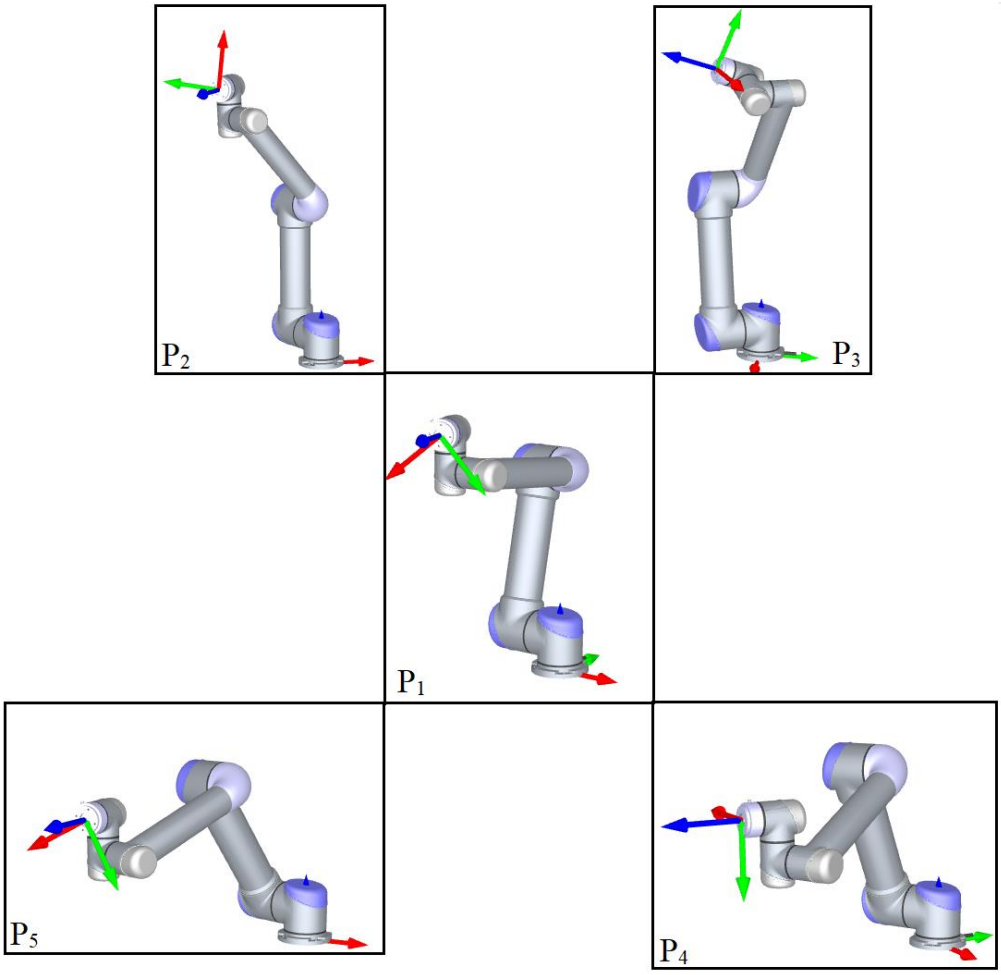

Figure 4.18. The UR5's TCP configuration at the ground-truth points

4.7.3. Calculation of the Absolute Pose Accuracy of the TCP

TCP estimated pose (position and orientation) at the ground-truth points recorded by the proposed system. At each of the points P_1 to P_5, the vision algorithm calculates the RAM pose in camera coordinate frame and then using (4.1), the absolute pose estimation of the TCP at each ground truth points in the robot base frame is calculated. The absolute positioning accuracy of the TCP in the ground truth points calculated as:

$$PA_{P_i} = \sqrt{PA_X^2 + PA_y^2 + PA_Z^2} \qquad (4.34)$$

Where in (4.34):

Positioning accuracy in the x-axis:

$$PA_x = (\bar{x} - x_c) \qquad (4.35)$$

Positioning accuracy in the y-axis:

$$PA_y = (\bar{y} - y_c) \qquad (4.36)$$

And the positioning accuracy in the z-axis

$$PA_z = (\bar{z} - z_c) \qquad (4.37)$$

In (4.35) - (4.37):

$\bar{x}$, $\bar{y}$, and $\bar{z}$ are the barycenter (mean) of the estimated positions by the proposed vision system:

$$\bar{x} = \frac{1}{n}\sum_{k=1}^{n} x_k \qquad (4.38)$$

$$\bar{y} = \frac{1}{n}\sum_{k=1}^{n} y_k \qquad (4.39)$$

$$\bar{z} = \frac{1}{n}\sum_{k=1}^{n} z_k \qquad (4.40)$$

and x_c, y_c, and z_c are the ground truth point (TCP actual position) of the UR5, which is calculated using the forward kinematic formulas.

Orientation accuracy about x, y, and z axes of the proposed system is estimated using:

$$OA_a = (\bar{a} - a_c) \qquad (4.41)$$

$$OA_b = \left(\bar{b} - b_c\right) \qquad (4.42)$$

$$OA_c = (\bar{c} - c_c) \qquad (4.43)$$

With:

$$\bar{a} = \frac{1}{n}\sum_{k=1}^{n} a_k \qquad (4.44)$$

$$\bar{b} = \frac{1}{n}\sum_{k=1}^{n} b_k \qquad (4.45)$$

$$\bar{c} = \frac{1}{n}\sum_{k=1}^{n} c_k \qquad (4.46)$$

Where in (4.41) to (4.43), a_c, b_c, and c_c are the actual angles of the UR5 TCP and $\bar{a}$, $\bar{b}$, and $\bar{c}$ are the averages of the estimated orientations of the TCP by the proposed system.

Table 4.9 shows the barycenter (mean) of the estimated positions and the ground-truth positions (TCP actual pose) of TCP at the points P_1 to P_5.

Table 4.9. *Positions knowledge of the ground-truth poses and barycenter position in robot base frame*

Point No.	$\bar{x}$(mm)	$\bar{y}$(mm)	$\bar{z}$(mm)	x_c (mm)	y_c (mm)	z_c (mm)
P_1	-79.4481	-366.6834	639.84	-79.41	-367.1	640
P_2	-328.9754	-119.0754	890.4381	-328.8	-118.7	890.2
P_3	172.8175	-115.5044	889.5824	172.7	-115.4	890.2
P_4	169.4676	-615.7949	389.4127	169.4	-615.7	389.4
P_5	-328.9930	-616.0199	390.0517	-328.9	-616	390.2

Table 4.10 shows the mean of the estimated orientation and the ground-truth orientations (TCP actual orientation) at the points P_1 to P_5.

Table 4.10. *Orientations knowledge of the ground-truth and mean points in robot base frame*

Point No.	$\bar{a}$ (°)	$\bar{b}$ (°)	$\bar{c}$ (°)	a_c (°)	b_c (°)	c_c (°)
P_1	-14.3840	-101.9593	89.9834	-14.2765	-101.9654	89.9813
P_2	-56.3960	77.0483	-44.3927	-56.3848	77.0496	-44.3975
P_3	-65.6256	-33.4142	-21.1247	-65.6275	-33.4166	-21.1245
P_4	-47.8260	-96.0523	95.7490	-47.8243	-96.0505	95.7467
P_5	-27.3980	-1.950985	-111.7832	-27.3997	-111.7894	112.4427

The statistical analysis of the pose (position and orientation) estimation accuracy performance of the proposed system for the 150 measurements (5points × 30 measurements) are shown in Table 4.11 and Table 4.12.

Table 4.11. *Positioning accuracy of the proposed system*

Point No.	$PA_X = \bar{x} - x_C$	$PA_y = \bar{y} - y_C$	$PA_z = \bar{z} - z_C$	PA_{P_i}
P_1	0.0628	0.3800	0.0770	0.3928
P_2	-0.1753	-0.3753	0.2380	0.4778
P_3	0.1175	-0.1044	-0.1176	0.1963
P_4	0.0676	-0.0948	0.0127	0.1172
P_5	-0.0930	-0.0199	-0.1482	0.1761

Units are in millimeters

Table 4.12. *Orientation accuracy of the proposed system*

Point No.	$OA_a = \bar{a} - a_C(°)$	$OA_b = \bar{b} - b_C(°)$	$OA_c = \bar{c} - c_C(°)$
P_1	-0.009	-0.002	0.0005
P_2	-0.01	-0.001	0.005
P_3	0.002	0.003	-0.002
P_4	-0.002	-0.002	0.002
P_5	0.001	0.006	0.002

Figure 4.19-4.23, shows the 3D representations of the ground-truth poses and TCP estimated poses at the points P_1 to P_5.

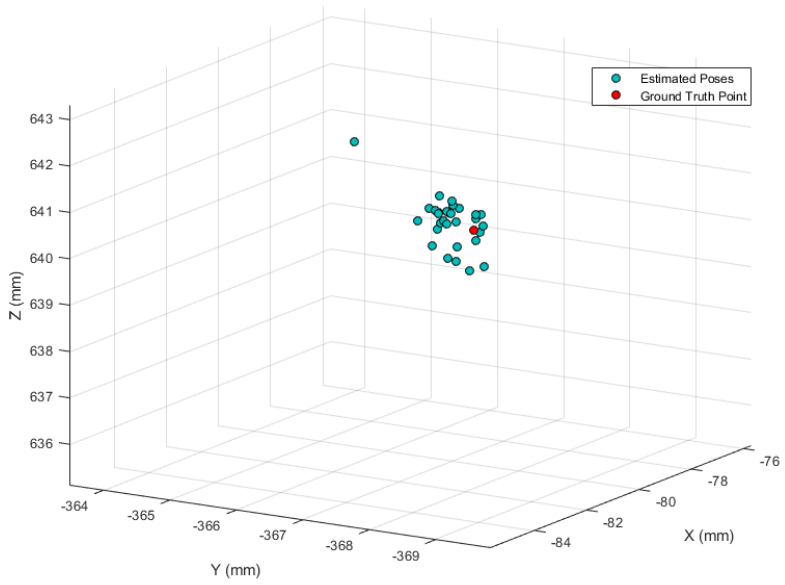

Figure 4.19. TCP actual position at the ground-truth point P_1 (red) and TCP estimated pose (greens)

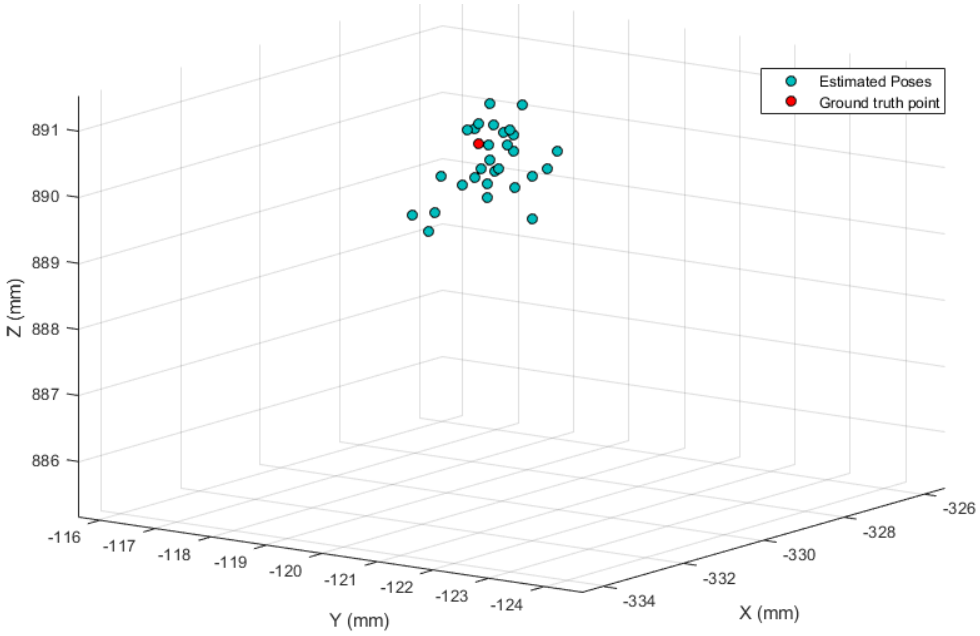

Figure 4.20. TCP actual position at the ground-truth point P_2 (red) and TCP estimated pose (greens)

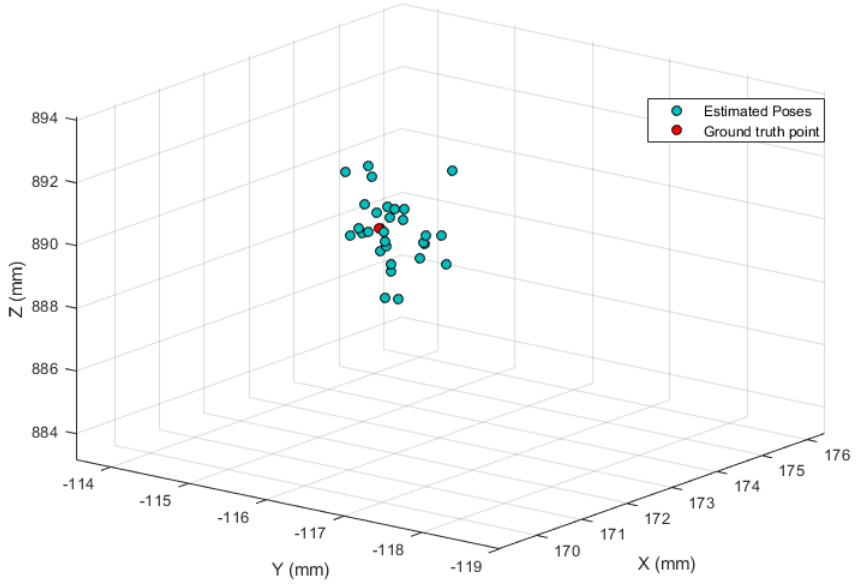

Figure 4.21. TCP actual position at the ground-truth point P_3 (red) and TCP estimated pose (greens)

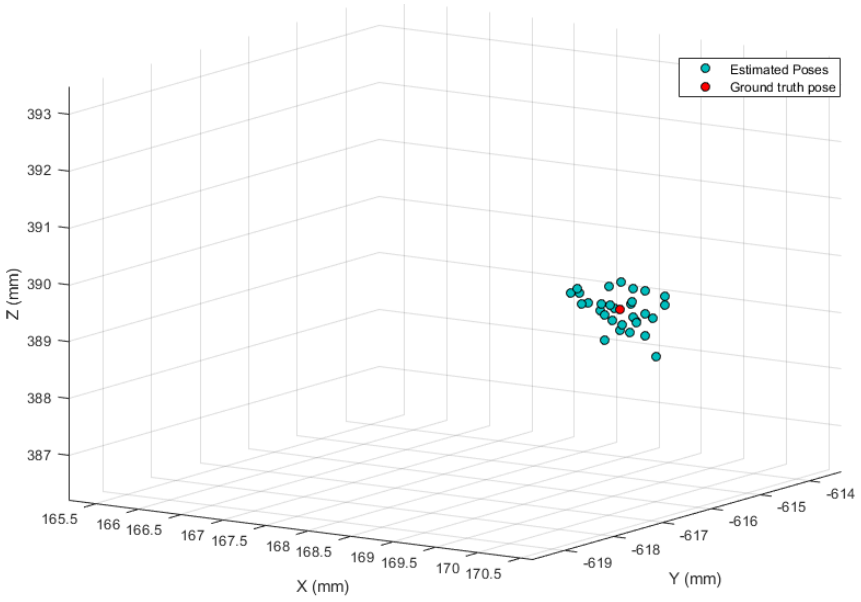

Figure 4.22. TCP actual position at the ground-truth point P_4 (red) and TCP estimated pose (greens)

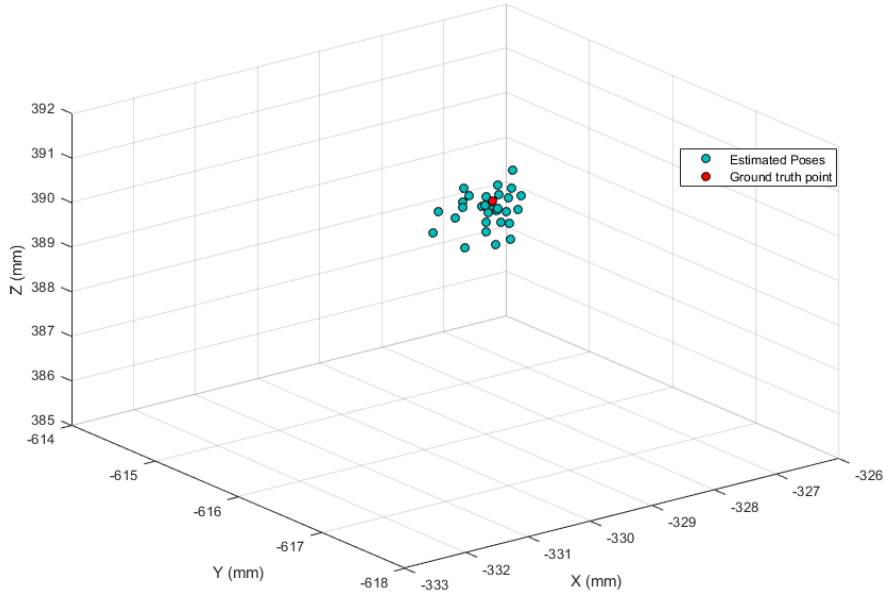

Figure 4.23. TCP actual position at the ground-truth point P_5 (red) and TCP estimated pose (greens)

Figs.4.24-4.28 shows the distributions of the positions and orientations of the TCP calculated using the proposed system.

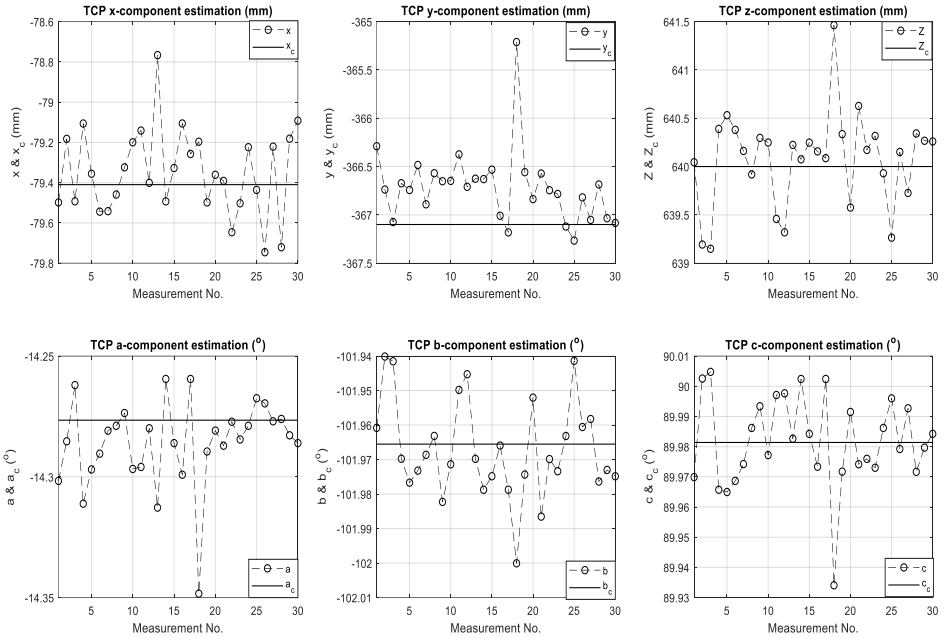

Figure 4.24. TCP estimated and actual pose at ground-truth point P_1

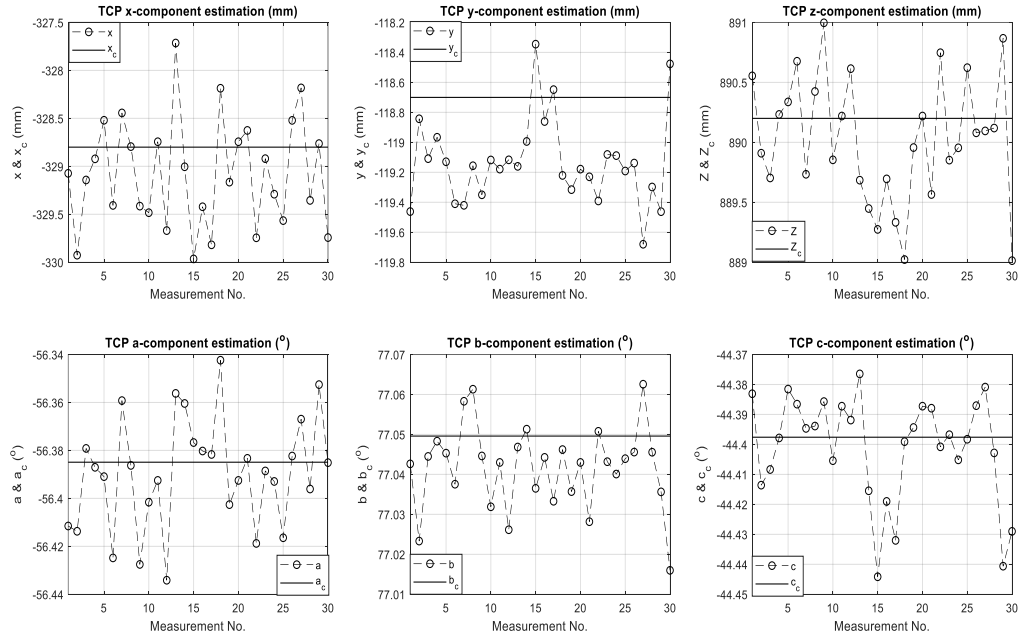

Figure 4.25. TCP estimated and actual pose at ground-truth point P_2

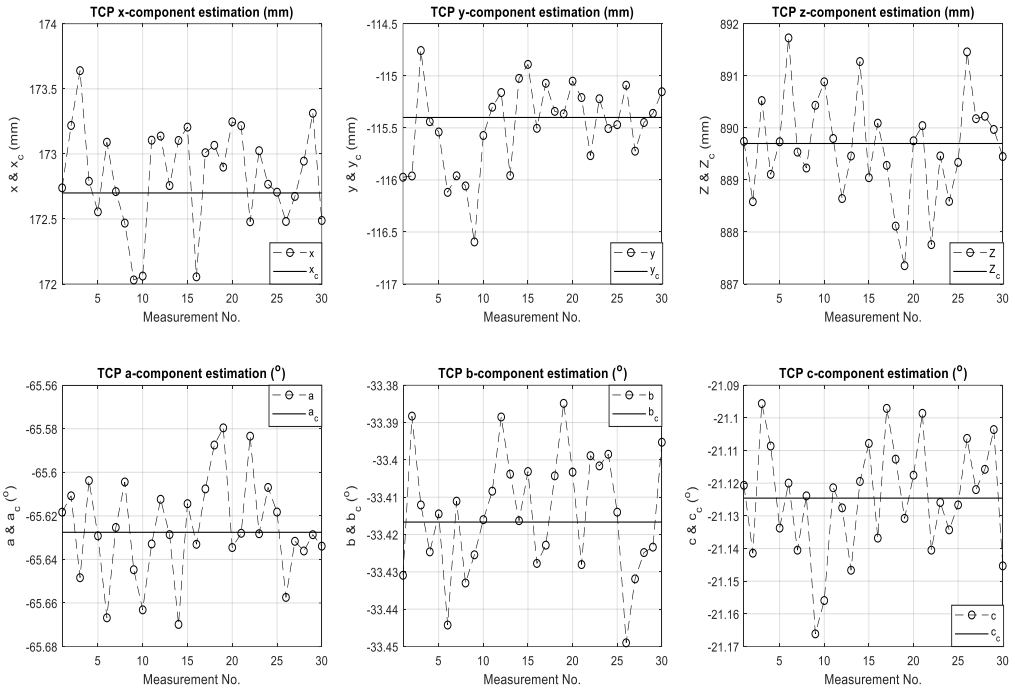

Figure 4.26. TCP estimated and actual pose at ground-truth point P_3

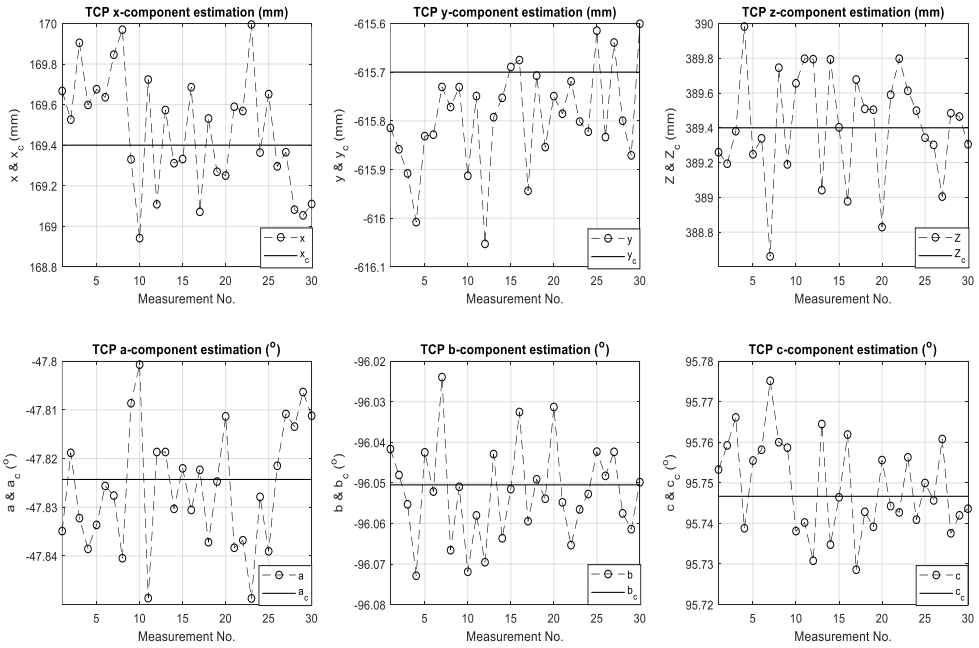

Figure 4.27. TCP estimated and actual pose at ground-truth point P_4

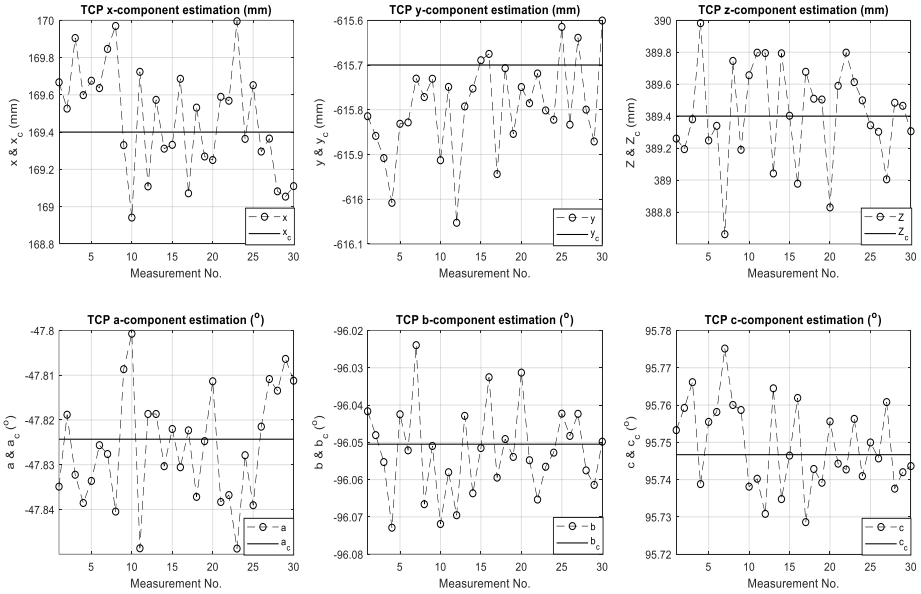

Figure 4.28. TCP estimated and actual pose at ground-truth point P_5

4.7.4. Calculation of the Pose Repeatability of the Proposed System

At each ground-truth poses P_1 to P_5, the developed pose estimation algorithm is executed for 10 seconds. During this period of experiment conduction, about 315-368 poses information (position and orientation) recorded for the TCP pose locations.

Repeatability of the measured pose data at each ground-truth poses is calculated using:

$$PR_{P_i} = \sqrt{\frac{\sum_{i=1}^{n}(P_i - \bar{P})^2}{n-1}} \qquad (4.48)$$

Where in (4.48):

$$P_i = \sqrt{(x_{p_i} - \bar{x})^2 + (y_{p_i} - \bar{y})^2 + (z_{p_i} - \bar{z})^2} \qquad (4.49)$$

Table 4.13 shows the calculated pose repeatability using (4.48).

Table 4.13. *Repeatability of TCP position measurements by the proposed system*

Point No.	PR_{P_i}(mm)
P_1	±0.0002
P_2	±0.0005
P_3	±0.0004
P_4	±0.00006
P_5	±0.00004

The repeatability of the orientation data measurement is calculated for each orientation angle about the axes x, y, and z in values of a, b, and c radians:

$$OR_a = \pm\sqrt{\frac{\sum_{i=1}^{n}(a_i-\bar{a})^2}{n-1}} \qquad (4.50)$$

$$OR_b = \pm\sqrt{\frac{\sum_{i=1}^{n}(b_i-\bar{b})^2}{n-1}} \qquad (4.51)$$

$$OR_c = \pm\sqrt{\frac{\sum_{i=1}^{n}(c_i-\bar{c})^2}{n-1}} \qquad (4.52)$$

Table 4.14 depicts the repeatability of the measured orientations of the TCP at each ground truth points.

Table 4.14. *Repeatability of TCP orientation measurements by the proposed system*

Point No.	OR_a(°)	OR_b(°)	OR_c(°)
P_1	$\pm$0.03	$\pm$0.02	$\pm$0.03
P_2	$\pm$0.03	$\pm$0.03	$\pm$0.05
P_3	$\pm$0.03	$\pm$0.04	$\pm$0.03
P_4	$\pm$0.01	$\pm$0.02	$\pm$0.03
P_5	$\pm$0.009	$\pm$0.02	$\pm$ 0.02

4.8. Pose Accuracy of the Proposed System in the camera frame

In this section, the pose accuracy and repeatability of the proposed system are evaluated in the camera coordinate camera frame. The RAM's pose in the camera coordinate frame depicted by (x_{RAM}, y_{RAM}, z_{RAM}, a_{RAM}, b_{RAM}, c_{RAM}). This evaluation method, is comparable to the pose estimation methods used in laser trackers, and optical CMMs which performs in the tool's reference frame. The position measurement accuracy of RAM in the x_{camera}, y_{camera}, and z_{camera} directions calculated from [92]:

$$CPA_x = \frac{|X_i - \bar{X}|}{N} \qquad (4.53)$$

$$CPA_y = \frac{|Y_i - \bar{Y}|}{N} \qquad (4.54)$$

$$CPA_Z = \frac{|Z_i - \bar{Z}|}{N} \qquad (4.55)$$

Where in (4.53)-(4.55), N was the number of measurements, X_i, Y_i , and Z_i are the position measurements at the i-th step, $\bar{X}$, $\bar{Y}$, and $\bar{Z}$ are the barycenter of the measurements in the x, y, and z-directions. The orientation measurement accuracy of the RAM in the camera coordinate frame calculated:

$$COA_a = \frac{|a_i - \bar{a}|}{N} \qquad (4.56)$$

$$COA_b = \frac{|b_i - \bar{b}|}{N} \qquad (4.57)$$

$$COA_c = \frac{|c_i - \bar{c}|}{N} \qquad (4.58)$$

In (4.53) -(4.55), CPA stands for the Coordinate Position Accuracy, and COA stands for Coordinate orientation accuracy. Table 4.15 and 4.16 shows the RAM's position and orientation measurements in the camera reference frame.

Table 4.15. *Accuracy of the coordinate position measurement*

Point No.	N	CPA_x(mm)	CPA_y(mm)	CPA_z (mm)
P_1	368	0.0326	0.0189	0.1908
P_2	356	0.0002	0.1575	0.6350
P_3	315	0.00004	0.0632	0.2526
P_4	320	0.00002	0.0124	0.0822
P_5	318	0.00004	0.0144	0.0630

Table 4.16. *Accuracy of the coordinate orientation measurement*

Point No.	COA_a(°)	COA_b (°)	COA_c(°)
P_1	0.02	0.02	0.02
P_2	0.03	0.02	0.04
P_3	0.02	0.02	0.02
P_4	0.01	0.01	0.02
P_5	0.007	0.01	0.02

4.9. Pose Precision Evaluation in the case of Marker-Occlusion

One of the main advantages of the application of the map of markers instead of a single marker is its robustness to marker occlusion. Occlusion can occur by other objects or by the proposed system failure in detecting of some RAM's markers. For this purpose, the robot configured in a specified pose such that all of the RAM's markers can be detected by the system. Then, marker is incrementally on from one to seven, and the effects of the marker numbers in the pose estimation are studied. Figure 4.29 shows the effects of the marker's occlusion on the precision (repeatability) of the proposed system.

In Figure 4.29, at first only one marker (marker #0) is available to the system, and other markers occluded (turned down in the RAM map, Figure 4.29(a)). For this case, RAM's pose precision (repeatability) is measured in the camera coordinate frame.

This case was the worst situation that the proposed system might encounter during its action. Then, the markers with ID numbers 2, 3, 4, 10, 11, 12 are turned on one by one, Figure 4.29(b) - 4.29(g).

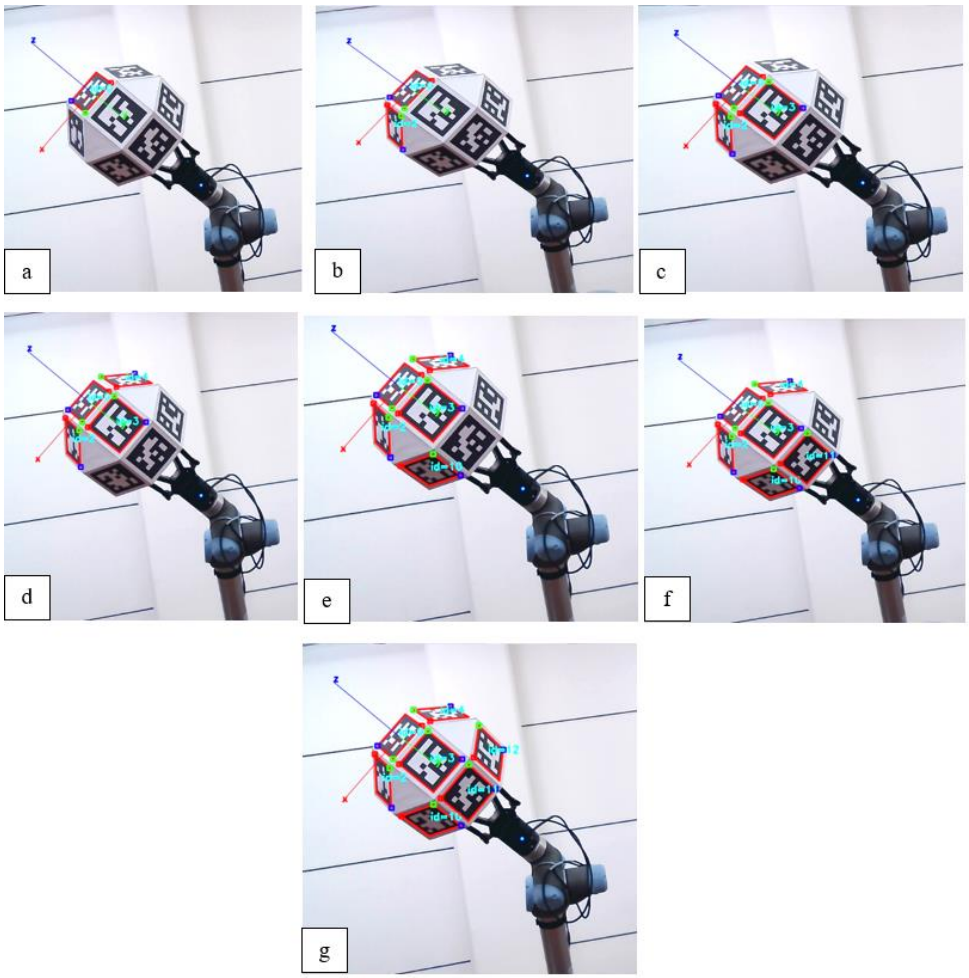

Figure 4.29. Effects of the marker's occlusion in the proposed pose estimation system

Table 4.17 and 4.18 shows the position and orientation precision values estimated by the proposed system.

Table 4.17. *Marker occlusion effect on the position precision (repeatability)*

Number of Markers	Position Repeatability (Precision)-mm
1 (Figure4.19(a))	± 0.4000
2 (Figure4.19(b))	± 0.2344
3 (Figure4.19(c))	± 0.1300
4 (Figure4.19(d))	± 0.1225
5 (Figure4.19(e))	± 0.0881
6 (Figure4.19(f))	± 0.0642
7 (Figure4.19(g))	± 0.0704

Table 4.18. *Marker occlusion effect on the orientation precision (repeatability)*

Number of Markers	OR_a (°)	OR_b (°)	OR_c (°)
1 (Figure4.19(a))	± 0.05	± 0.03	± 0.04
2 (Figure4.19(b))	± 0.04	± 0.02	± 0.04
3 (Figure4.19(c))	± 0.02	± 0.02	± 0.03
4 (Figure4.19(d))	± 0.02	± 0.02	±0.02
5 (Figure4.19(e))	± 0.02	± 0.02	±0.02
6 (Figure4.19(f))	±0.02	± 0.02	±0.02
7 (Figure4.19(g))	± 0.02	±0.02	±0.02

CHAPTER 5

RESULTS AND DISCUSSIONS

5.1. Vision System

In this section, the effects of the vision system on the performance of the proposed system in pose estimating are discussed in more detail.

5.1.1. Illumination (Lighting)

Illumination techniques is an essential factor in computer vision applications. In order to have robust and reliable results, researchers try to create uniform illumination during computer vision applications. Environmental lighting intensity is a function of time, and this matter makes the computer vision applications a time-dependent process. Therefore, in the computer vision applications, researchers try to control the light intensities by using particular kinds of lighting systems and by pulling black curtains in front of windows to have a uniform illumination in the experiment environment.

During the experiments, there was a combined illumination from the fluorescent and outdoor lights.

that caused non-uniform illumination and, some of the markers are detected as if they are blinking. Blinking markers are those markers that their detection was interrupted by the illumination of the environment and the camera resolution. As a result, the proposed vision system detected some markers in on-off mode. This negative phenomenon of blinking markers was one of the primary sources of the presence of the outlier pose data in the pose estimation process using the Rhombicuboctahedron ArUco Mapper (RAM). Figure 5.1 shows the presence of a blinking marker as a result of uncontrolled illumination in one of the experiments. This phenomenon degraded the pose estimation accuracy and repeatability in some degrees. In Figure 5.1 (a), marker #9 gets on-off in a fraction of milliseconds, and consequence of this

phenomena is the presence of outlier poses that are far from the TCP actual pose or ground-truth poses.

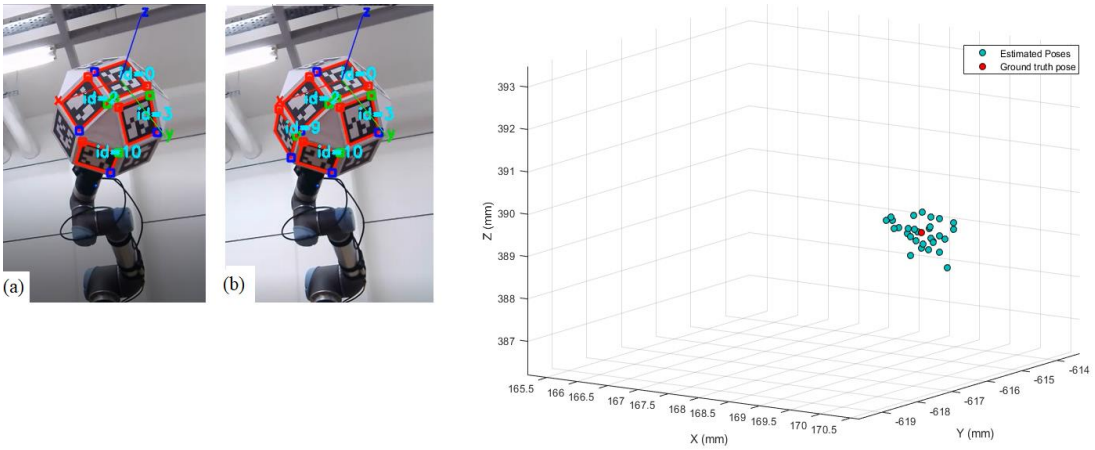

Figure 5.1. The blinking marker phenomena in the marker detection operation (a) off, (b) on, and (c) arising of outlier poses

5.1.2. Camera Resolution

The camera resolution is an essential factor in image processing applications. In the early experiments, to decrease the processing time, low-resolutions such as 640×480 was used for processing. Although low processing time was obtained using 640×480 resolution, errors in pose estimation procedures increased significantly. It can be described based on the perspective-n-point pose estimation procedure applied in fiducial markers. In low-resolution images, detection of corner of markers was not accurate. Based on this fact, to increase the pose estimation accuracy, the resolution is set to 1280×720. In this case, high pose estimation accuracy and high-repeatability of the data measurements proved the reliability of the applied resolution.

5.2. Rhombicuboctahedron ArUco Mapper (RAM)

In this section, some considerations about the Rhombicuboctahedron ArUco Mapping system is discussed.

5.2.1. The 3D-Geometrical Consideration of the RAM

The geometrical characteristics of the mapper model enable the vision system to detect at least four markers and at most seven markers during the pose estimation operations. Figure 5.2, shows the minimum and maximum detected ArUco markers during one of the experiments. Results showed that as the number of detected ArUco markers increased, the pose accuracy of the proposed system improved considerably.

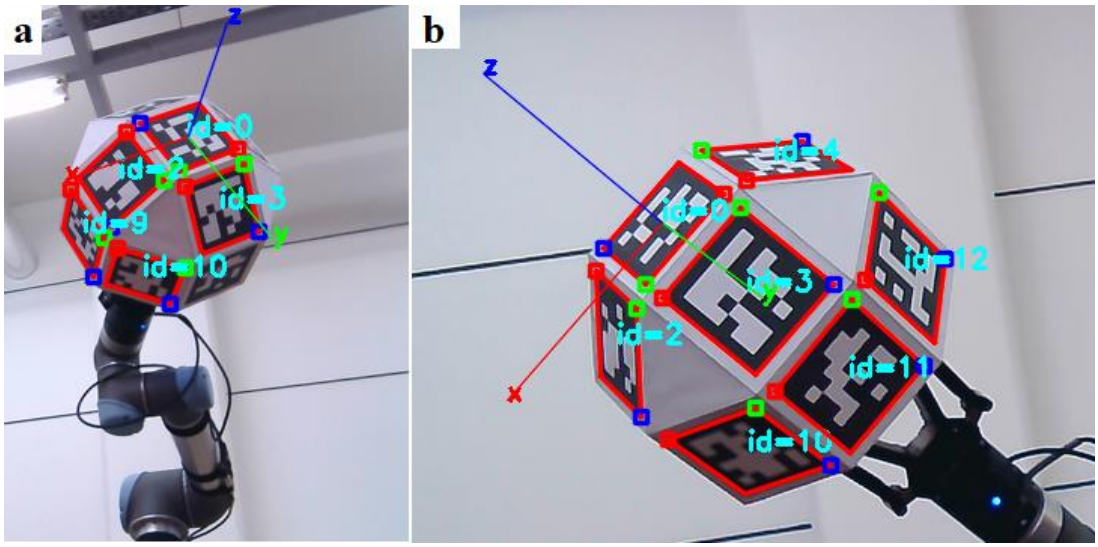

Figure 5.2. The minimum detected marker (a), and maximum detected marker (b)

5.2.2. ArUco.3 as the applied fiducial marker in RAM structure

The ArUco.3 markers are used as the fiducial markers in the structure of the RAM. ArUco.3 is an optimized version of the conventional ArUco marker. ArUco.3 is about 17 to 40 times faster than the conventional ArUco markers. Moreover, in ArUco.3 the memory problem has been solved using mixed-integer linear programming algorithms (MILP) [93]. In the previous ArUco versions, as the size of markers increased, memory requirements of the detection algorithm increased exponentially. Accordingly, as experienced in this study, there is not any limitation in using markers with bigger sizes for the RAM pose estimations.

5.2.3. Rhombicuboctahedron ArUco Mapper (RAM) Performance Evaluation

Fiala [94] has defined some metrics to evaluate the performance of the fiducial markers. In this study, three essential metrics considered.

- **The false-positive rate**: It is defined as the false detection of a non-existing marker. In this study, false positive detection has never been observed in Rhombicuboctahedron ArUco Mapper (RAM).
- **The inter-marker confusion rate**: It is a metric for the detection of the wrong marker ID. This also has never been observed during the experiments.
- **The false-negative rate**: In this case, despite the presence of the fiducial marker in the scene, its presence is not reported by the fiducial marker system. This has never been observed during the experiments.

The Vertex-jitter: Vertex-jitter is a significant problem in fiducial markers systems. This problem shows itself in the form of the vibration of marker vertices during marker detections step. This vibration causes errors during pose estimation. Despite efforts put into eliminating the vertex-jitter, in some orientations and positions, the proposed system experienced this undesirable effect which degraded the pose accuracy of the proposed system. A demo video [95] shows the occurrence of the vertex-jitter.

5.3. The Proposed Absolute Pose Estimation System

In this section, some aspects of the pose accuracy and repeatability are discussed.

5.3.1. The Absolute Pose Accuracy

Five ground-truth poses are selected whose pose knowledge with respect to the robot base are known. These five poses are representative of a sample of poses with different positions and orientations that TCP might experience during a typical run. The distance of these points from the camera are listed in Table 4.8. The distance of points P_1, P_2, P_3, P_4, and P_5 from the camera were 1442.9mm, 1803.6mm, 1914.5mm, 1151.3mm, and 1168.8mm, respectively. The closest and farthest points to the camera were P_5, and P_3, respectively. The lowest absolute positional accuracy estimated by

the proposed system is obtained as 0.4778 mm for point P_2, Table 4.11. By considering the RAM configuration at P_2, it observed that the system detected was due to four markers; therefore, the position estimation of the RAM had resulted from the participation of the four markers or sixteen 3D points. However, what degraded the positional accuracy was the presence of blinking markers, which disturbed the position estimation of the TCP at point P_2.

According to Table 4.11, and Figure 4.20, it is concluded that the most significant errors happened in the y-component of the TCP position at point P_2. Although P_1 is relatively close to the camera, it suffered from the marker blinking problem. Its absolute positional accuracy is about 0.3928mm. The farthest point from the camera is P_3, but since six markers were detected in pose estimation procedure and also there was not any blinking marker, its absolute positional accuracy was 0.1964mm which was a significant improvement over the absolute positional accuracy of the points P_1 and P_2. Figure 4.26, shows the symmetrical distribution of the estimated positions x, y, and z of the TCP about the P_3. The best absolute positional accuracy of the TCP happened at the points P_4 and P_5, that are about 0.1172mm and 0.1762mm, respectively. The main reason for the high positional accuracies at these points are associated with their closeness to the camera, and the absence of the blinking markers. Considering the x, y, and z components of the estimated positions of TCP in Figures.4.27 and 4.28, showed small positional errors in the x, y, and z directions, moreover symmetrical distribution of the estimated position components about the ground truth values proved this fact.

Test results for the orientation accuracy of the proposed TCP pose estimation system at the ground truth points showed that the most accurate results obtained at the points P_4, is 0.003°. The high orientation accuracy at this point can be associated with the closeness of P_4 to the camera, and also the absence of the marker blinking problem. The orientation accuracies of points P_5 and P_1 are in the desirable limits, 0.007°, and 0.0044° respectively. The points P_5 and P_1 are close to the camera (1168.8mm, and 1442.9mm) and did not experience the marker blinking problem.

The lowest orientation accuracy belongs to the point P_2, 0.01°. It is surprising to mention that at point P_2, there is the absolute position accuracy problem as well. Similar to position accuracy problem, at point P_2, the low accuracies in orientation estimation by the proposed system can be attributed to the presence of the blinking markers. The blinking markers and vertex-jitter phenomena are the two main factors in obtaining high accuracy for orientation and position by the proposed TCP pose estimation system in this study. Closeness to the camera is also an important factor in pose estimation accuracy.

5.3.2. The Pose Estimation Repeatability

The repeatability of the TCP pose measurement system is analyzed in the camera coordinate system. The tests are carried out with the same hardware (camera and computer), at the same location (UR5 and camera locations), and in relatively short periods (10 seconds) without interrupting the camera. The results of the pose repeatability analysis are depicted in Table 4.13 and Table 4.14. The statistical results of repeatability tests show that there is a high level of repeatability of the developed pose estimation algorithm. The highest position measurements repeatability belong to the points P_4 and P_5, about 0.00006 and 0.00005 mm, respectively. P_2 has the lowest repeatability of the position measurement (0.0005184 mm). Results show that the pose measurement repeatability of the system was reliable. Similar to the absolute position accuracy analysis, with the occurrence of the marker blinking phenomena, the positional measurement accuracy of the proposed system degraded significantly. There are agreements between the orientation measurements by the proposed system for all the point P_1-P_5, in the range of ±0.009° to ±0.05°. The best orientational repeatability, (±0.03°) occurred at the points P_4 and P_5 where the distance between the camera and TCP were short, and there were not any blinking markers. Similarly, the lowest orientation repeatability belongs to P_2, which suffers from position, orientation and pose repeatability problem as the result of blinking markers phenomena.

5.4. Affecting Factors of the Pose Estimation System Performance

Hardware and software problems affect the pose performance capability of the proposed pose estimation system.

5.4.1. Robot

Similar to the other industrial robotic manipulators, the UR5 robot can be subjected to pose error problems mentioned in chapter 2. UR5 is only used for educational purposes; accordingly, it has not experienced any overloading problems. Thus, it was assumed that the UR5 forward kinematic is correct and it is not subject to the high pose errors. It was also assumed that the UR5 robot is calibrated.

5.4.2. Camera

Some parts of the TCP pose measurement error are due to the camera specifications. The camera used during the experiments is an A4TECH 1080 HD web camera, equipped with a CMOS sensor. Although CMOS sensors are high-speed sensors and are suited for the detection of the fast-moving objects, they suffer from distortion problems. This problem is associated with CMOS sensors and affected the detection of the corners of the markers significantly. Consequently, any disturbances in corner point detection affected the Perspective-n-point pose estimation algorithm which would lead to the erroneous pose estimation. Figure 5.3 shows the schematic representation of the distortion effect by the CMOS sensors in marker corner detection process which affects the pose estimation operation using fiducial marker systems.

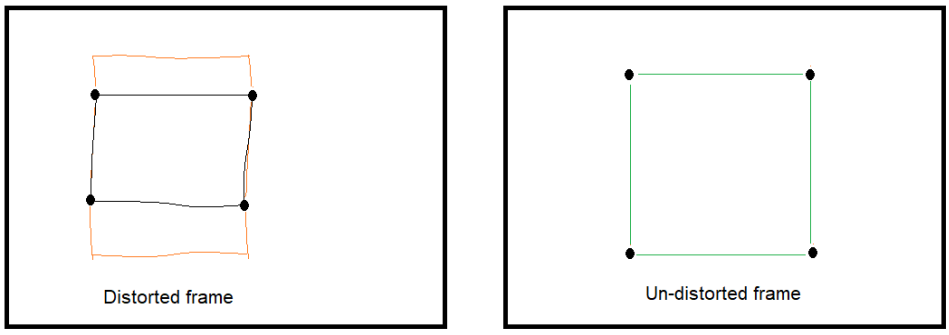

Figure 5.3. CMOS sensors distortion effects on the marker's corner detection

5.4.3. Reference Frame Transformations

The sources of the TCP pose estimation error came from the homogenous transformation from the TCP to the robot base frame. After pose estimation of the Rhombicuboctahedron ArUco Mapper (RAM) with respect to the camera, using (4.1), the transformation from the TCP to the robot base was performed. Since the homogenous transformation is a numerical operation, and moreover the distance between the TCP frame and the RAM frame measured manually and also checked from CAD information, therefore, some numerical errors are introduced during pose estimation.

In conclusion, by minimizing the errors mentioned in Sections 5.4.1 to 5.4.3, the pose accuracy of the developed system would improve significantly.

5.5. Effects of Occluded-markers on the Position Precision (Repeatability)

An experiment is carried out to study the effects of the occluded markers in the TCP pose measurement precision (repeatability) in Section 4.9. It is concluded that with the increasing of the occluded markers of the RAM, the pose precision decreases considerably as shown in Figure 5.4.

According to Figure 5.4 (and Table 4.12), with increasing of the markers from one marker to seven markers, the positioning precision (repeatability) increases six times. Thus, by decreasing the effects of the RAM's marker-occlusion, the pose estimation precision would improve considerably.

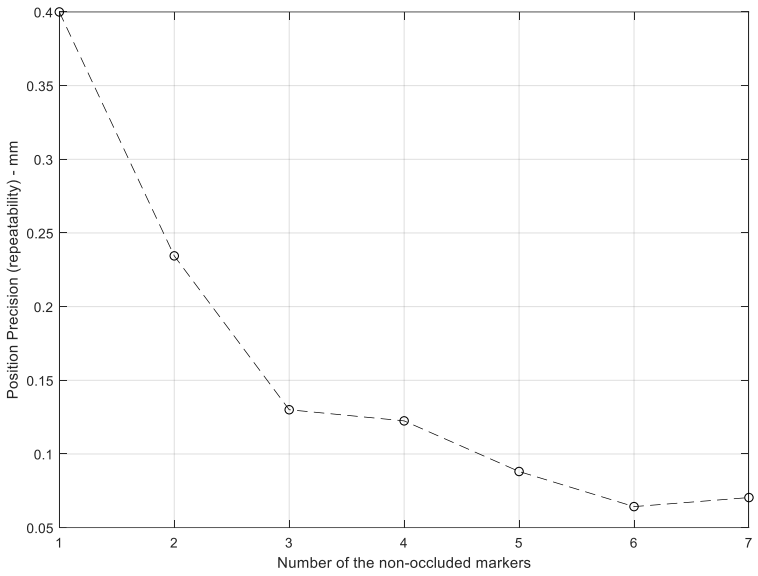

Figure 5.4. Effects of the non-occluded markers on the positioning precision (Repeatability)

5.6. Applications of the Proposed TCP Pose Estimation System

Improving the positional accuracy and repeatability of the industrial robotic manipulators become a critical task that robot producers compete over it to take the market control. The industrial robot calibration process is one of the practical techniques for improvement of the industrial robot's pose accuracy.

Based on the estimated position accuracy and repeatability of the proposed system, the developed method is suitable for some industrial applications, especially for the improvement of the TCP pose accuracy and repeatability. Currently, based on the complexity level of the proposed method, it is applicable in point-to-point industrial robotic applications. In the following sections, some point-to-point industrial tasks suitable for the developed method are explained.

5.6.1. Point-to-point Industrial Robots

A point-to-point robot performs its tasks at discrete points defined in the robot workspace. In point-to-point tasks, the robot drives its TCP to a specified discrete point in the robot workspace, then stops there and performs its task. After finishing

the task at that point, the robot drives its TCP to the next static point, and it repeats the working cycle.

5.6.1.1. Spot Welding

Welding robots attract the primary attention of the customers in the industry and manufacturing sectors. Welding operation is a laborious and unhealthy operation because of generation of noise, fume, and intense light. The automation of the welding tasks in industrial level requires repeatable robots. According to some industrial robots' catalogues, the required pose repeatability is in the ±0.04 –0.2 mm [96][97]. In this study, the pose measurement repeatability of the proposed method is in the range of 0.00005 – 0.0005 mm. Thus, it is concluded that the proposed method has the required repeatability to be used in welding operations.

5.6.1.2. Handling, loading, and Unloading

Machines and tools loading, unloading and handling tasks requires robots with high accuracy and repeatability. In these tasks, the robot picks a tool or part from a tool holder or a conveyor belt and then position and orient it into a particular configuration required by the receiver machine or unit. Some suggested repeatability for the handling-robots to accomplish these kinds of tasks is ± 0.1 mm, and absolute pose accuracy is ± 0.1mm [92]. Based on the estimated pose accuracy and repeatability, the developed method can be applied for pose improvement of the handling-robots' TCP.

5.6.1.3. Assembly

Assembly robots are mainly used in the automotive and aerospace industries. The robots serve as an assistant or as a master robot to accomplish the assembly tasks. The recommended repeatability for assembly task is ± 0.5mm, and for positional accuracy is ± 0.5mm [96]. The obtained pose repeatability and accuracy of the proposed system in this study proved that it could be applied for the assembly task process.

5.6.1.4. Inspection

Customers' growing interest in purchasing high-quality products, made the producers manufacture the zero defects products. Since the human inspection in quality control tasks is subject to failures, inspection robots are replacing with human elements in manufacturing sites. The most critical item for the inspection robots is repeatability. Based on the high-repeatability of the TCP position measurements obtained in this work, the proposed system can be applied in industrial inspection robots.

5.6.1.5. Hole drilling

Many aerospace manufacturing tasks require robots with high positional and orientational accuracy and repeatability. One of the point-to-point robotic applications in the aerospace industry is hole drilling operations of carbon composites. In [23], the required positional accuracy of 6 DOF serial robots to be applied in hole drilling tasks of the aerospace parts is reported up to 0.2mm. The proposed pose estimation system developed in this study can be applied in this sector.

5.6.1.6. Component insertion

Component insertion task is a flexible industrial task. The required pose accuracy and repeatability to automate this operation by robots is dependent on the design and manufacturing tolerances specified by a particular application. These tolerances varies between 0.05mm to more than several millimeters depending on the manufacturing operations. The estimated pose measurements accuracy and repeatability of the proposed system in this study proved that it could use for the positional and orientational accuracy and repeatability improvement of the industrial robots in component insertion operations.

5.6.2. Continuous path applications

The continuous path applications require robots to accomplish their tasks while the robot TCP is in dynamic mode. Continuous path industrial applications include spraying, finishing, gluing, Arc welding, cleaning of metal particles, and some

complex assembly processes. To evaluate the robots' positional and orientational accuracies and repeatability for continuous path applications requires measurement devices capable of evaluating the path accuracy and repeatability of the robots. Currently, because of the frame-rate limitations of the applied camera in this study (30fps), path accuracy and repeatability evaluation of the proposed system for the continuous path applications are not possible.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1. Conclusion

The demands for industrial robots with high accuracy and repeatability, especially in the automotive and aerospace industries dictate the robot manufacturers and users to make their robots precise by improving robots' pose accuracy, and repeatability. There is a need for external metrology tools such as laser trackers, optical CMMs and stereo vision and photogrammetry techniques to fill this gap. Although laser-trackers and optical CMMs are precise pose estimation tools, they are not cost-effective and, in most cases, there is a need for more interference of the operator to conduct the measurements.

In this study, a novel 6 DOF pose estimation algorithm for the real-time absolute position and orientation measurements of the TCP was proposed and implemented based on the computer vision and augmented reality markers principles. For this purpose, an innovative marker mapping method is proposed. An Archimedean solid called the truncated Rhombicuboctahedron which has thirteen regular faces is used as the mapping model. In this study, ArUco.3 markers are used in the map model structure.

The developed pose estimation system can compete with the currently used pose estimation devices and tools. The proposed method in this study is considered as a cost-effective device. It just uses a regular CMOS-sensor web camera and some print-out markers. Unlike laser trackers and optical CMMs which costs hundreds thousands of dollars and uses some other expensive peripheral devices such as retroreflectors, or in stereo vision systems which needs several expensive CCD cameras for the absolute pose accuracy in the orders of 1/10(mm) and orientation estimation in the order of

1/1000 (°). Despite being cost-effective, its pose estimation accuracy and repeatability calculated on the orders of 1/10(mm) for position and 1/1000 (^{o}) for the orientation measurements. Thus, it can be applied for the calibration operation of TCP of industrial robots in some point-to-point applications.

Characteristics of the proposed TCP pose estimation techniques can be summarized as follows:

- **Non-contact process**: It is not necessary for physical contact of the TCP for pose measurement.
- **High accuracy in pose measurement**: The statistical analysis of the pose accuracy showed a satisfactory and more accurate position accuracy of 0.12mm-0.48mm and for orientation in the range from 0.003^o – 0.01^o.
- **High repeatability in pose measurements**: The proposed pose estimation system showed a high level of agreement between successive pose measurements. The measurement repeatability was in the range of ±0.00005mm-0.0005mm and for the orientation in the range of $\pm 0.009^o$-0.05^o.
- **Cost-effective**: Only some printed black-and-white papers and a cheap camera webcam are used for the TCP pose estimation.
- **Easy-to-use and flexibility**: There is no need for special skills or setup configuration in order to use the proposed system. The proposed TCP pose estimation system can be used in most places with enough illumination to detect markers.
- **Appropriate for real-times applications**: The proposed pose estimation system in this study is fast enough to be applied in the real-time tasks. Depending on the number of detected markers of RAM, its processing time is in between the 19 to 25 milliseconds.
- **Comparable with commercial pose estimation systems**: Currently, the most reliable calibration devices for the pose estimation purposes of the industrial robots are laser trackers, optical CMMs and stereo vision. By applying these tools, the absolute positional accuracy of the manipulators has increased in

ranges of less than 1 mm. The positional and orientational accuracy and repeatability of the proposed pose estimation system in this study showed that it is comparable in performance with the mentioned devices for pose estimation.

In conclusion, the developed pose estimation system is an accurate and repeatable TCP pose estimation method that can be applied for the improvement of the industrial robots' pose accuracy and repeatability for point-to-point applications.

6.2. Future Works

The proposed pose estimation system can be considered as a prototype that needs some performance improvements to be more practical and efficient in manufacturing operations. In the future works, high-speed CCD machine vision cameras will be used as the visual input for the proposed pose estimation system. CCD cameras obviate the problem of distortions; therefore, their application will solve the problem of vertex-jitter. Using high frame-rates cameras will increase the application of the system for more applications such as path continuous tasks like machining, deburring, spraying, polishing and cutting.

REFERENCES

[1] H. A. G. Lopez and M. A. P. Cisneros, “Industry 4.0 & Internet of Things in Supply Chain,” in *Proceedings of the 8th Latin American Conference on Human-Computer Interaction - CLIHC ’17*, 2017, pp. 1–4.

[2] M. Skilton, *The 4th Industrial Revolution*. Basingstoke: Palgrave Macmillan, 2018.

[3] C. Gifford, F. Picini, and Australian Broadcasting Corporation., *Robots*. ABC Books for the Australian Broadcasting Corporation, 2008.

[4] P. Corke, *Robotics, Vision and Control*, 2nd ed., vol. 73. Berlin: Springer, 2011.

[5] Matt Simon, “The Complete History And Future of Robots | WIRED.” [Online]. Available: https://www.wired.com/story/wired-guide-to-robots/.

[6] “robot | Origin and meaning of robot by Online Etymology Dictionary.” [Online]. Available: https://www.etymonline.com/word/robot. [Accessed: 15-Aug-2019].

[7] George A. Bekey, *Autonomous Robots: From Biological Inspiration to Implementation and Control - George A. Bekey - Google Books*. Cambridge, Massachusetts, 2005.

[8] S. Y. Nof, *Handbook of Industrial Robots*, 1st ed. New York, NY: John Wiley & Sons Inc, 1999.

[9] “The 5 Minute Guide to Use Any End Effector with RoboDK - RoboDK blog.” [Online]. Available: https://robodk.com/blog/robot-end-effector-guide/#targetText=The Tool Center Point (TCP,the center of the fingers).

[10] I. International Federation of Robotics, “Robots and the Workplace of the Future,” no. March, p. 35, 2018.

[11] “Robots double worldwide by 2020 - International Federation of Robotics.” [Online]. Available: https://ifr.org/ifr-press-releases/news/robots-double-worldwide-by-2020.

[12] M. Abderrahim, A. Khamis, S. Garrido, and L. Moreno, “Accuracy and Calibration Issues of Industrial Manipulators,” *Ind. Robot. Program. Simul. Appl.*, no. December, 2006.

[13] “Global industrial robot sales doubled over the past five years - International Federation of Robotics.” [Online]. Available: https://ifr.org/ifr-press-

releases/news/global-industrial-robot-sales-doubled-over-the-past-five-years.

[14] “RobotWorx - Industrial Robot Applications.” [Online]. Available: https://www.robots.com/applications.

[15] R. M. Podoloff and W. P. Seering, “an Accuracy Test Procedure for Robotic Manipulators,” pp. 19–22.

[16] J. S. Shamma and D. E. Whitney, “A method for inverse robot calibration,” *J. Dyn. Syst. Meas. Control. Trans. ASME*, vol. 109, no. 1, pp. 36–43, 1987.

[17] M. Y. Amirat, J. Pontnau, and F. Artigue, “A three-dimensional measurement system for robot applications,” *J. Intell. Robot. Syst.*, vol. 9, no. 3, pp. 291–299, Mar. 1994.

[18] L. Stadelmann, T. Sandy, A. Thoma, and J. Buchli, “End-Effector Pose Correction for Versatile Large-Scale Multi-Robotic Systems,” *IEEE Robot. Autom. Lett.*, vol. 4, no. 2, pp. 546–553, 2019.

[19] M. Summers, “Robot Capability Test and Development of Industrial Robot Positioning System for the Aerospace Industry,” 2005.

[20] M. A. Meggiolaro, S. Dubowsky, and C. Mavroidis, “Geometric and elastic error calibration of a high accuracy patient positioning system,” *Mech. Mach. Theory*, vol. 40, no. 4, pp. 415–427, 2005.

[21] “Strategies for Improving Robotic Accuracy with Metrology – Metrology and Quality News - Online Magazine.” [Online]. Available: https://metrology.news/strategies-for-improving-robotic-accuracy-with-metrology/.

[22] J. M. Corum, R. L. Battiste, K. C. Liu, and M. B. Ruggles, “Basic Properties of Reference Crossply Carbon-Fiber Composite,” vol. 27, p. 53, 2000.

[23] S. van Duin, P. J. Crothers, P. Steele, and J. Newberry, “A Comparison between Indoor GPS versus Laser Tracking Metrology for Robotic Drilling,” *SAE Tech. Pap. Ser.*, vol. 1, no. 2006, pp. 1016–1022, 2010.

[24] H.-N. Nguyen, J. Zhou, and H.-J. Kang, “A new full pose measurement method for robot calibration.,” *Sensors (Basel).*, vol. 13, no. 7, pp. 9132–47, Jul. 2013.

[25] N. Thyer, “Double Theodolite Pibal Evaluation by Computer,” *J. Appl. Meteorol.*, vol. 1, no. 1, pp. 66–68, Mar. 1962.

[26] M. A. R. (Michael A. R. Cooper, *Modern theodolites and levels*. Granada, 1982.

[27] “Theodolite Error.” [Online]. Available: http://whistleralley.com/surveying/theoerror/.

[28] M. R. D. Benjamin W. Mooring, ZVI S. Roth, *Fundamentals of manipulator*

calibration, vol. 32, no. 1–2. New York: John Wiley & Sons Inc, 1991.

[29] D. E. Whitney, C. A. Lozinski, and J. M. Rourke, "Industrial robot forward calibration method and results," *J. Dyn. Syst. Meas. Control. Trans. ASME*, vol. 108, no. 1, pp. 1–8, 1986.

[30] M. R. Driels and U. S. Pathre, "Vision-based automatic theodolite for robot calibration," *IEEE Trans. Robot. Autom.*, vol. 7, no. 3, pp. 351–360, Jun. 1991.

[31] R. J. Hocken and P. H. Pereira, *Coordinate Measuring Machines and Systems, Second Edition*. 2012.

[32] K. Harding, *Handbook of Optical Dimensional Metrology*. Boca Raton: CRC, 2016.

[33] W. K. Veitschegger and C. H. Wu, "Robot Calibration and Compensation," *IEEE J. Robot. Autom.*, vol. 4, no. 6, pp. 643–656, 1988.

[34] B. W. Mooring and S. S. Padavala, "The effect of kinematic model complexity on manipulator accuracy," in *Proceedings, 1989 International Conference on Robotics and Automation*, pp. 593–598.

[35] "How Laser Trackers Work | Quality Digest." [Online]. Available: https://www.qualitydigest.com/inside/metrology-article/how-laser-trackers-work-062509.html.

[36] T. Yoshizawa, *Handbook of Optical Metrology*. Boca Raton: CRC, 2009.

[37] "Interferometry explained." [Online]. Available: https://www.renishaw.com/en/interferometry-explained--7854.

[38] A. Nubiola, M. Slamani, A. Joubair, and I. A. Bonev, "Comparison of two calibration methods for a small industrial robot based on an optical CMM and a laser tracker," vol. 32, no. 2014, pp. 447–466, 2019.

[39] K. Lau, R. J. Hocken, and W. C. Haight, "Automatic laser tracking interferometer system for robot metrology," *Precis. Eng.*, vol. 8, no. 1, pp. 3–8, Jan. 1986.

[40] J. P. Prenninger, "Contactless position and orientation measurement of robot end-effectors," in *[1993] Proceedings IEEE International Conference on Robotics and Automation*, pp. 180–185.

[41] S. Decker, H. Gander, M. Vincze, and J. P. Prenninger, "Dynamic measurement of position and orientation of robots," *IEEE Trans. Instrum. Meas.*, vol. 41, no. 6, pp. 897–901, 1992.

[42] Y. Bai, H. Zhuang, and Z. S. Roth, "Experiment study of PUMA robot calibration using a laser tracking system," *SMCia 2003 - Proc. 2003 IEEE Int. Work. Soft Comput. Ind. Appl.*, pp. 139–144, 2003.

[43] G. Alici and B. Shirinzadeh, “A systematic technique to estimate positioning errors for robot accuracy improvement using laser interferometry based sensing,” *Mech. Mach. Theory*, vol. 40, no. 8, pp. 879–906, 2005.

[44] H. Zhuang, “Modeling identification and compensation of robot manipulators,” Florida Atlantic University, 1989.

[45] R. J. Radke, *Computer vision for visual effects*. Cambridge University Press, 2013.

[46] R. E. Rafael C. Gonzalez, Woods, *Digital Image Processing*. New Jersey: Pearson Education, 2008.

[47] A. Bobick, A., Essa, I., and Chakraborty, *Introduction to computer vision*. Udacity Online Course Series on Computer Vision, 2015.

[48] G. Alenyà, S. Foix, and C. Torras, “ToF cameras for active vision in robotics,” *Sensors Actuators A Phys.*, vol. 218, pp. 10–22, Oct. 2014.

[49] L. Pérez, Í. Rodríguez, N. Rodríguez, and R. Usamentiaga, “Robot Guidance Using Machine Vision Techniques in Industrial Environments : A Comparative Review,” 2016.

[50] and X. W. T. Clarke, “The Control of a robot End-effector using Photogrammetry,” vol. XXXIII, pp. 137–142, 2000.

[51] L. Pérez, Í. Rodríguez, N. Rodríguez, R. Usamentiaga, and D. F. García, “Robot guidance using machine vision techniques in industrial environments: A comparative review,” *Sensors (Switzerland)*, vol. 16, no. 3, 2016.

[52] W. C. Chang, “Robotic assembly of smartphone back shells with eye-in-hand visual servoing,” *Robot. Comput. Integr. Manuf.*, vol. 50, no. August 2017, pp. 102–113, 2018.

[53] M. Švaco, B. Šekoranja, F. Šuligoj, and B. Jerbić, “Calibration of an industrial robot using a stereo vision system,” *Procedia Eng.*, vol. 69, pp. 459–463, 2014.

[54] G. Siciliano, B., Sciavicco, L., Villani, G., and Oriolo, *Robotics: Modeling, Planning, and Control*. London: Springer, 2010.

[55] C. Möller, H. C. Schmidt, N. H. Shah, and J. Wollnack, “Enhanced Absolute Accuracy of an Industrial Milling Robot Using Stereo Camera System,” *Procedia Technol.*, vol. 26, pp. 389–398, 2016.

[56] A. Ryberg, M. Ericsson, A. K. Christiansson, K. Eriksson, J. Nilsson, and M. Larsson, “Stereo vision for path correction in off-line programmed robot welding,” *Proc. IEEE Int. Conf. Ind. Technol.*, pp. 1700–1705, 2010.

[57] M. Dinham and G. Fang, “Autonomous weld seam identification and localisation using eye-in-hand stereo vision for robotic arc welding,” *Robot.*

Comput. Integr. Manuf., vol. 29, no. 5, pp. 288–301, Oct. 2013.

[58] K. Claes and H. Bruyninckx, "Robot positioning using structured light patterns suitable for self calibration and 3D tracking," *Proc. 2007 Int. Conf. Adv. Robot.*, 2007.

[59] J. Pagès, F. Chaumette, C. Collewet, and J. Salvi, "A camera-projector system for robot positioning by visual servoing," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2006, 2006.

[60] "Optical CMMs | Measurement System Types and Characteristics | Measurement Fundamentals | KEYENCE America." [Online]. Available: https://www.keyence.com/ss/products/measure-sys/measurement-selection/type/image.jsp.

[61] "Optical CMMs: The future of CMMs? | Quality Digest." [Online]. Available: https://www.qualitydigest.com/inside/metrology-article/optical-cmms-future-cmms-072810.html.

[62] "3D optical measurement device for vehicle safety testing - RITM IndustryRITM Industry." [Online]. Available: http://ritmindustry.com/catalog/3d-measuring-arms-laser-trackers/3d-optical-measurement-device-for-vehicle-safety-testing/.

[63] S. Gharaaty, T. Shu, A. Joubair, W. F. Xie, and I. A. Bonev, "Online pose correction of an industrial robot using an optical coordinate measure machine system," *Int. J. Adv. Robot. Syst.*, vol. 15, no. 4, p. 172988141878791, Jul. 2018.

[64] A. T. Erdem and A. O. Ercan, "Fusing inertial sensor data in an extended kalman filter for 3D camera tracking," *IEEE Trans. Image Process.*, vol. 24, no. 2, pp. 538–548, 2015.

[65] M. B. Alatise and G. P. Hancke, "Pose Estimation of a Mobile Robot Based on Fusion of IMU Data and Vision Data Using an Extended Kalman Filter.," *Sensors (Basel).*, vol. 17, no. 10, Sep. 2017.

[66] F. Boochs, R. Schutze, C. Simon, F. Marzani, H. Wirth, and J. Meier, "Increasing the accuracy of untaught robot positions by means of a multi-camera system," in *2010 International Conference on Indoor Positioning and Indoor Navigation*, 2010, pp. 1–9.

[67] T. Hsiao, "Iterative Learning Control for Trajectory Tracking of Robot Manipulators," *Int. J. Autom. Smart Technol.*, vol. 7, no. 3, pp. 133–139, 2017.

[68] "Measurement Sciences Products - Measurement Sciences." [Online]. Available: https://www.ndigital.com/msci/products/.

[69] "K-CMM | Optical CMM | Portable measuring | Nikon Metrology." [Online].

Available: https://www.nikonmetrology.com/en-gb/product/k-cmm.

[70] A. Filion, A. Joubair, A. S. Tahan, and I. A. Bonev, "Robot calibration using a portable photogrammetry system," *Robot. Comput. Integr. Manuf.*, vol. 49, no. June 2017, pp. 1339–1351, 2018.

[71] D. Rieke-Zapp, W. Tecklenburg, J. Peipe, H. Hastedt, and C. Haig, "Evaluation of the geometric stability and the accuracy potential of digital cameras — Comparing mechanical stabilisation versus parameterisation," *ISPRS J. Photogramm. Remote Sens.*, vol. 64, no. 3, pp. 248–258, May 2009.

[72] K. Konolige, "Projected texture stereo," *Proc. - IEEE Int. Conf. Robot. Autom.*, pp. 148–155, 2010.

[73] R. Muñoz-Salinas, M. J. Marín-Jimenez, E. Yeguas-Bolivar, and R. Medina-Carnicer, "Mapping and localization from planar markers," *Pattern Recognit.*, vol. 73, no. June 2016, pp. 158–171, 2018.

[74] M. a Fischler and R. C. Bolles, "Paradigm for Model," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[75] L. Quan and Z. Lan, "Linear N-point camera pose determination," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 8, pp. 774–780, 1999.

[76] B. Triggs, "Camera pose and calibration from 4 or 5 known 3D points," *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 1, pp. 278–284, 1999.

[77] M. Dhome, M. Richetin, J. T. Lapresteé, and G. Rives, "Determination of the Attitude of 3-D Objects from a Single Perspective View," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 12, pp. 1265–1278, 1989.

[78] X. S. Gao, X. R. Hou, J. Tang, and H. F. Cheng, "Complete solution classification for the perspective-three-point problem," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 8, pp. 930–943, 2003.

[79] R. H. ; B. C. ; O. L. ; B. Lacolle, "An analytic solution for the perspective 4-point problem," in *Computer Vision and Pattern Recognition (CVPR)*, 1989.

[80] E. R. Pi, "Implementation of a 3D pose estimation algorithm," no. June, pp. 12–17, 2015.

[81] "Real Time pose estimation of a textured object — OpenCV 3.0.0-dev documentation." [Online]. Available: https://docs.opencv.org/3.0-beta/doc/tutorials/calib3d/real_time_pose/real_time_pose.html.

[82] "UR5 collaborative robot arm | flexible and lightweight robot arm." [Online]. Available: https://www.universal-robots.com/products/ur5-robot/.

[83] "(User Manual)," *UR5/CB3 User Man.*, vol. 1, no. 1, 2016.

[84] “ROS.org | Powering the world’s robots.” [Online]. Available: https://www.ros.org/.

[85] N. M. López, F. Sciascio, C. M. Soria, and M. E. Valentinuzzi, “Control of a Robotic Arm,” *Biomed. Eng. Online*, vol. 13, pp. 1–13, 2009.

[86] “GRIPKIT by Weiss Robotics | Gripping Modules for Universal Robots.” [Online]. Available: https://www.weiss-robotics.com/gripkit/en/.

[87] “Machine Vision Lighting Demystified.” [Online]. Available: https://www.qualitymag.com/articles/89399-machine-vision-lighting-demystified.

[88] M. K. Özgören, *Robot Manipülatörlerin Dinamiği ve Kontrolu*. Ankara: Makina Teorisi Derneği, 2016.

[89] “1080p Full-HD WebCam(PK-910H) |.” [Online]. Available: https://www.a4tech.com/product.aspx?id=36.

[90] P. A. Systems, “ISO 9283:1998 Manipulating Industrial Robots-Performance Criteria and Related Test Methods,” 2005.

[91] “Repeatability - an overview | ScienceDirect Topics.” [Online]. Available: https://www.sciencedirect.com/topics/engineering/repeatability.

[92] “How to Calculate the Accuracy of Measurements.” [Online]. Available: https://sciencing.com/calculate-accuracy-measurements-6391160.html.

[93] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and R. Medina-Carnicer, “Generation of fiducial marker dictionaries using Mixed Integer Linear Programming,” *Pattern Recognit.*, vol. 51, pp. 481–491, 2016.

[94] M. Fiala, “Designing highly reliable fiducial markers,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 7, pp. 1317–1324, 2010.

[95] *Vertex-Jitter effects in the Rhombicuboctahedron ArUco Mapper (RAM) pose estimation - YouTube*.

[96] and K. P. Andre Cugy, *Industrial Robot Specifications*. Paris: Kogan Page, 1984.

[97] “Top 10 Advantages of Robotic Welding » Scott.” [Online]. Available: https://www.scottautomation.com/news/articles/top-10-advantages-of-robotic-welding/.

[98] S. Savarese, “Computer Vision, From 3D Reconstruction to Recognition, Lecture Notes,” 2018.

[99] S. Maybank, *Theory of Reconstruction from Image Motion*. Berlin: Springer Verlag, 1993.

[100] “Affine and Projective Transformations - Graphics Mill 5.5 for .NET.” [Online]. Available: https://www.graphicsmill.com/docs/gm5/Transformations.htm.

[101] R. T. Collins and Y. Tsin, “Calibration of an outdoor active camera system,” in *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, pp. 528–534.

[102] J. Weng, P. Cohen, and M. Herniou, “Camera calibration with distortion models and accuracy evaluation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 10, pp. 965–980, 1992.

[103] “Calculate robot coordinates from measured chessbaord corners (hand-eye calibration) - OpenCV Q&A Forum.” [Online]. Available: https://answers.opencv.org/question/204910/calculate-robot-coordinates-from-measured-chessbaord-corners-hand-eye-calibration/.

[104] C. Loughlin, “Eye-in-hand robot vision scores over fixed camera,” *Sens. Rev.*, vol. 3, 1983.

APPENDICES

A. Camera Models

A camera is a device which maps the 3D space into 2D space. This mapping is called the perspective projection.

$$Camera: \mathbb{R}^3 \Longrightarrow \mathbb{R}^2$$

Three camera models, namely the central perspective model (Pinhole Camera Model), Orthogonal projective model and affine projective model.

Central Perspective Model (Pin-hole Camera Model):

The pin-hole camera is the simple model for the perspective projection modelling. Image formation in pin-hole camera is depicted in Figure A.1.

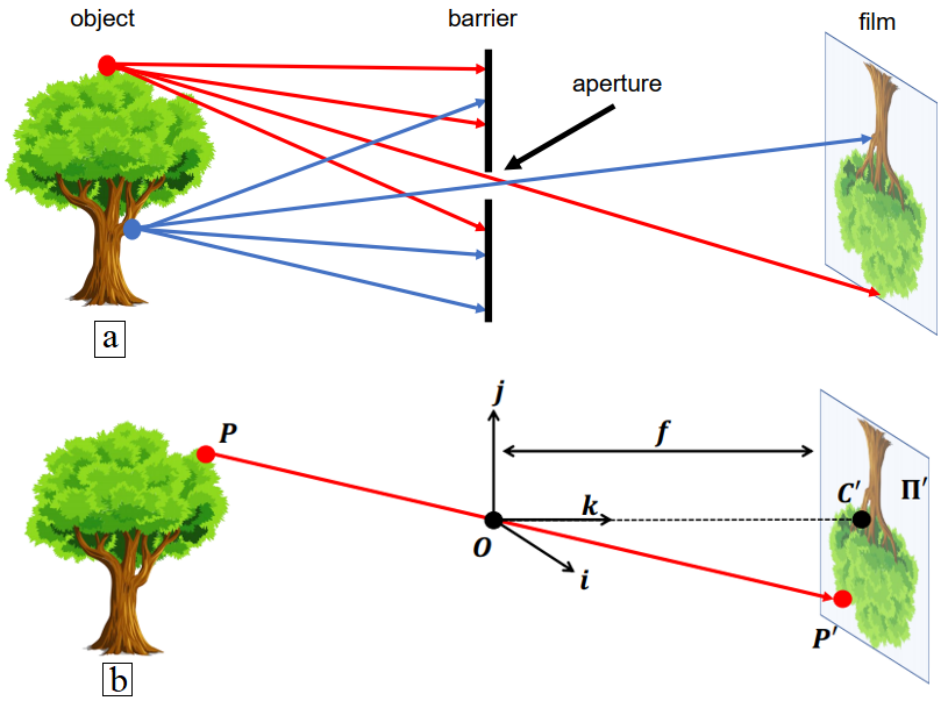

Figure A.1. Image formation in a pinhole camera model

In this architecture shown in Figure A.1 (a), a barrier called aperture is placed between the 3D object and a photographic film or sensor. A ray emanating from points on the

3D object surface passes through the aperture and strikes the photographic film. In vision processing terminologies, the film is called the image or retinal plane, the aperture is referred to as the pinhole or camera center *O*, and the distance between the pinhole and image plane is called focal length *f* [98]. The pinhole camera model is considered as a basic camera model. Figure A. 2 shows a schematically representation of the pinhole camera model.

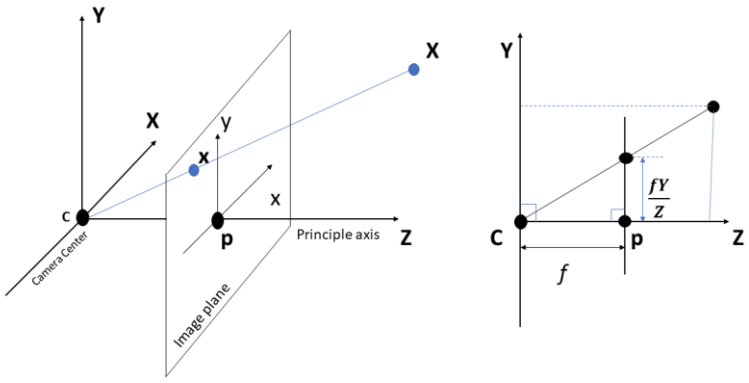

Figure A.2. Pin-hole camera model [98]

The central projection mapping from the 3D world (X) into the image plane is:

$$(X, Y, Z)^T \quad \rightarrow \quad (\frac{fX}{Z}, \frac{fY}{Z})^T \qquad \text{(A.1)}$$

For the pinhole-camera model, representation of the projection equation from the 3D world into the image plane using in matrix representation is:

$$\begin{bmatrix} fX + Zp_x \\ fY + Zp_y \\ Z \end{bmatrix} = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \qquad \text{(A.2)}$$

Where, in (A.2):

$$K = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad \text{(A.3)}$$

is called the camera matrix.

Orthogonal Projective Model:

In this model, images were formed on the photosensitive plane of the camera from the parallel beams coming from the object surface [99]. The orthogonal projective model is an adequate model when there is a restriction in the range of depth and in the cases when the image is small.

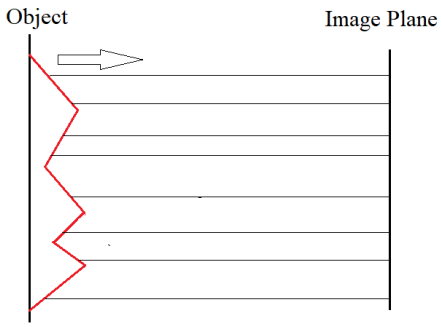

Figure A.3. The orthogonal image projection model[99]

This is method cannot measure the accurate results when the distance between the camera and the object is too small.

Affine Projective Model:

This model is a special case of the projective transformation. The affine projective models preserve points, straight lines and planes. For example, parallel lines remain parallel after affine transformation [100]. The application of this kind of cameras are so popular among researchers; however, the linearity of this model allows the camera calibration to be done with ease.

B. Camera Calibration

The pinhole cameras introduce significant distortions to the images. Radial and tangential distortions are two kinds of major distortions caused by camera sensors. Radial distortions cause the straight lines to appear curved. This distortion intensity becomes larger as they go farther from the image center. When the camera lens is not parallel with the image plane due to some reasons such as gravitational problems, it causes tangential distortions. In computer vision and image processing applications, it is necessary to obviate any distortion before any computer vision applications in order to obtain accurate results. Some researchers [101] have argued that the main source of the image distortion has come from lens distortion due to handling, temperature, and humidity conditions. Geometric camera calibration or camera calibration computes the parameters of a lens and image sensor of an image or camera in order to remove this adverse effect. Simply speaking, the camera calibration is a process to find the distortion coefficients.

$$(k_1, k_2, k_3, p_1, p_2) \qquad \text{(B.1)}$$

Camera calibration has applications to correct lens distortion, measure the size of an object in world coordinate systems, or determine the location of the camera in the scene in applications such as machine vision, robotics, 3D scene reconstruction, pose estimation and navigation systems. The camera calibration process includes the calculation of the intrinsic, extrinsic, and distortion coefficients. Camera calibration is an especially important action in dimensional quantitative measurements, depth from stereoscopy, or motion from images [102].

Intrinsic Camera Parameters:

Intrinsic camera parameters include some camera parameters such as focal length, the scale factor, and the optical center point coordinates. Intrinsic camera parameters are obtained through a process called camera calibration, which resulted in the intrinsic camera matrix or calibration matrix:

$$K = \begin{bmatrix} \alpha_x f & sf & p_x \\ 0 & \alpha_y f & p_y \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{(B.2)}$$

In (B.2), *f* is called the focal length, and p_x and p_y are called the principle focal point in the x and y directional.

C. Camera-Robot Configuration

In robotic applications, based on the camera pose with respect to the robot, three cameras-robot configurations are defined. It is worth to mention that in robotic applications, the term “eye” refers to the camera, and “hand” refers to the robot TCP.

Eye-to-hand System:

In this configuration, the camera is fixed in a place in the robot workspace and is calibrated with respect to the robot base frame. Figure C.1 shows this kind of configuration schematically.

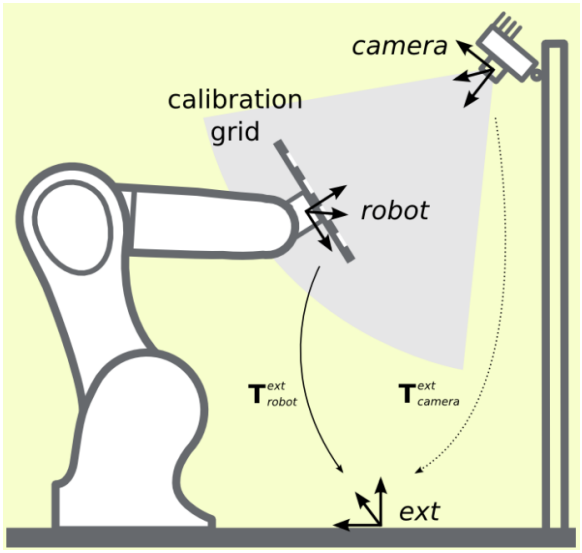

Figure C.4. Eye-to-hand configuration [103]

Although, the eye-to-hand system can see the robot and its workspace, however, it is possible to lose the target because of the robot links motion can occlude the robot TCP view.

Eye-in-hand System:

In this configuration, the camera is mounted on the robot structure and looking at the end-effector, Figure C.2. The main advantage of this method is that the robot cannot occlude the camera field of view. But, in this configuration, the robot arm should be

stopped in some cases, so that vision system can do its pose calculation of a part and sent it to the robot control system [104].

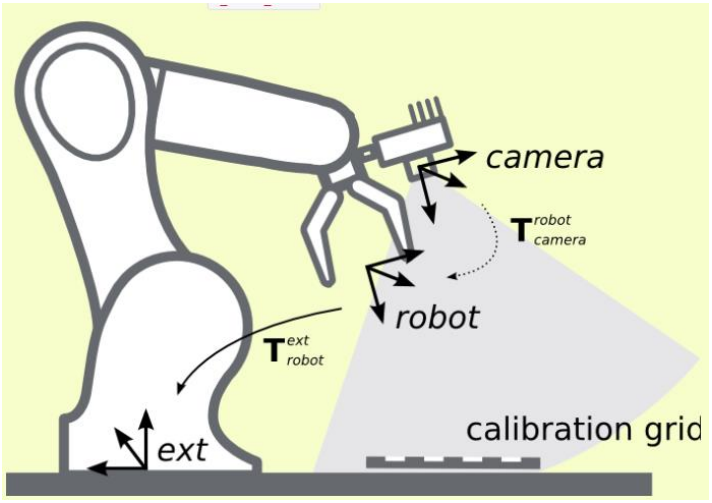

Figure C.5. Eye-in-hand configuration [105]

Active Camera Head:

In this configuration, the camera or eye is neither mounted on the robot manipulator nor overhead of the robot. Instead, it is mounted on a pan-and-tilt device. The pan-and-tilt device can change the pose of the camera automatically in order to track the object in motion.

D. Denavit-Hartenberg Parameters

Six-Degree-Of-Freedom industrial robot manipulators consist of a series of links connected to each other by means of joints. It is necessary to attach an arbitrary reference frame to each link to be able to extract the kinematical relations between the successive links. A kinematic relation between two successive links can be explained in terms of the homogenous transformations between the frames of links in the chains.

In order to be systematic in the choice of the attached frames to the links, a commonly used reference frame called the Denavit-Hartenberg analysis method or D-H convention is used in robotic applications. Figure D.1 shows the Denavit-Hartenberg representation of a kinematic chain of successive links with single-axis joints [88].

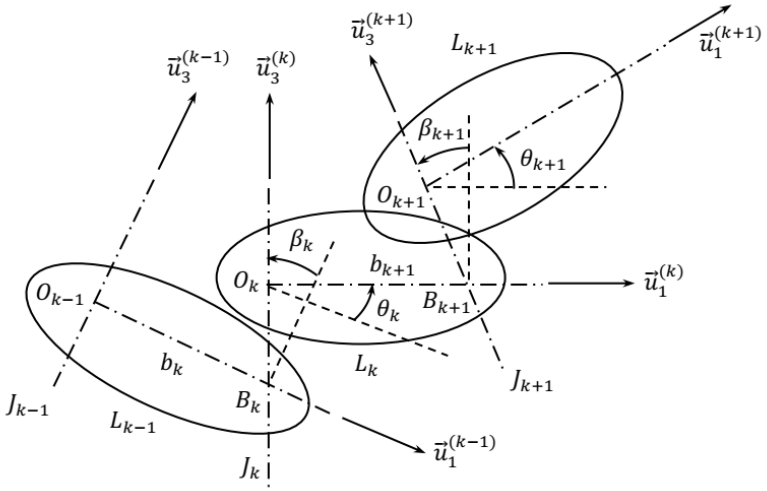

Figure D.6. The Denavit-Hartenberg convention for some links in the kinematic chain[88]

In the Denavit-Hartenberg convention, any homogenous transformation $\widehat{H}_{AB}^{(a,b)}$ between the reference frame of link, a and b can be represented as a product of four basic transformations.

$$\widehat{H}_{AB}^{(a,b)} = \boldsymbol{R}_{\vec{u}_3^{(k)},\theta_k}\boldsymbol{Trans}_{\vec{u}_3^{(k)},\theta_k}\boldsymbol{Trans}_{\vec{u}_1^{(k)},b_k}\boldsymbol{R}_{\vec{u}_1^{(k)},\beta_k} \qquad \text{(D.1)}$$

$$\widehat{H}_{AB}^{(a,b)} = \begin{bmatrix} cos\theta_k & -sin\theta_k & 0 & 0 \\ sin\theta_k & cos\theta_k & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & d_k \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & b_k \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & cos\beta_k & -sin\beta_k & 0 \\ 0 & sin\beta_k & cos\beta_k & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(D.2)

The four quantities θ_k, d_k, b_k, and β_k are parameters associated with the relation of the links *k-1* and *k* in the kinematic chain. These four kinematic parameters θ_k, d_k, b_k, and β_k are defined as the joint angle, link offset, link length, and twist angle, respectively. They are also called Denavit-Hartenberg or D-H parameters.

The D-H parameters are defined:

a) θ_k is the rotation angle that rotated the link L_{k-1} to the link L_k about joint axis $\vec{u}_3^{(k)}$, or mathematically:

$$\theta_k = \sphericalangle\left[\vec{u}_1^{(k-1)} \rightarrow \vec{u}_1^{(k)}\right] about\ \vec{u}_1^{(k)} \qquad \text{(D.3)}$$

If the two successive links have a revolute between them, then θ_k will be joint variable, however, if the connecting joint be prismatic joint then $\theta_k = \delta_k$ will be constant. δ_k is called the deflection angle.

b) $\boldsymbol{s_k}$ is the link offset of the two successive links L_{k-1} and L_k.

$$s_k = B_k O_k \text{ along } \vec{u}_3^{(k)} \qquad \text{(D.4)}$$

This distance shows the offset between the link L_{k-1} and L_k along $\vec{u}_3^{(k)}$. If the connecting joint between link L_{k-1} and L_k is prismatic then $\boldsymbol{s_k}$ will be joint variable. Otherwise, if it is rotation angle then $s_k = d_k$ is called offset.

c) $\boldsymbol{b_k}$ is the distance between the joint axes of the successive links.

$$b_k = O_{k-1}B_k\ along\ \vec{u}_1^{(k-1)} \qquad \text{(D. 5)}$$

d) $\boldsymbol{\beta_k}$ is called the twist angle between two successive links which is defined as:

$$\beta_k = \measuredangle\left[\vec{u}_3^{(k-1)} \rightarrow \vec{u}_3^{(k)}\right] about\ \vec{u}_1^{(k-1)} \qquad \text{(D. 6)}$$

Thus, in a nutshell, $(\theta_k, s_k, \beta_k, b_k)$ is called the D-H parameter .

E. Forward Kinematic of Robotic Manipulators

Kinematics is a branch of dynamic which studies the motion of the rigid bodies in the absence of the external forces. In other words, kinematics concentrates on the study of the geometrical description of motion using position, orientation, and their derivatives. In terms of the robotic manipulators, forward kinematics is the pose estimation of the Tool Center Point (TCP) or end-effectors with respect to a specified coordinate reference frame using kinematic equations. In order to achieve this goal, a coordinate frame is embedded in each link of the manipulator. Two successive links relative position and orientation can be calculated using homogenous transformations. Any homogenous transformation matrix can be represented by a matrix **H** called Denavit matrix. It is worth to mention that, a homogenous transformation matrix describes the relative translation and rotation between link coordinate systems. Mathematically, if $\mathbf{A_a}$ describes the position and orientation of the first link and $\mathbf{A_b}$ describe the position and orientation of the second link, then the relative position and orientation of the second link with respect to the first link would be:

$$\widehat{H}_{AB}^{(a,b)} = \boldsymbol{A}_a \boldsymbol{A}_b \qquad \text{(E.1)}$$

$\widehat{H}_{AB}^{(a,b)}$ is defined as a transformation matrix describing the relative pose of the link with the coordinate system {**B**} with respect to the link with the coordinate system {**A**}. Accordingly, for a six-degree-of-freedom serial manipulator, a transformation matrix which describe TCP pose with respect to the robot base will be:

$$\widehat{H}_{06}^{(0,6)} = \widehat{H}_{01}^{(0,1)} \widehat{H}_{12}^{(1,2)} \widehat{H}_{23}^{(2,3)} \widehat{H}_{34}^{(3,4)} \widehat{H}_{45}^{(4,5)} \widehat{H}_{56}^{(5,6)} \qquad \text{(E.2)}$$

In (E.2):

$$\widehat{H}_{k-1k}^{(k-1\,,k)} = \begin{bmatrix} \widehat{\boldsymbol{C}}^{(k-1\,,k)} & \bar{\boldsymbol{r}}_{k-1,k}^{(k-1\,)} \\ \bar{\mathbf{0}}^t & 1 \end{bmatrix} \qquad \text{(E.3)}$$

Where, $\widehat{\boldsymbol{C}}^{(k-1\,,k)}$ is called the orientation of the link L_k with respect to L_{k-1}.

$$\widehat{\boldsymbol{C}}^{(k-1,\ k)} = \hat{R}\left(\bar{u}_1^{(k-1/k-1)}, \beta_k\right)\hat{R}\left(\bar{u}_3^{(k/k)}, \theta_k\right) = \hat{R}(\bar{u}_1, \beta_k)\hat{R}(\bar{u}_3, \theta_k) \quad \text{(E.4)}$$

$$\widehat{\boldsymbol{C}}^{(k-1,\ k)} = e^{\tilde{u}_1\beta_k}e^{\tilde{u}_3\theta_k} \quad \text{(E.5)}$$

$$\widehat{\boldsymbol{C}}^{(k-1,\ k)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos\beta_k & -sin\beta_k \\ 0 & sin\beta_k & cos\beta_k \end{bmatrix}\begin{bmatrix} cos\theta_k & -sin\theta_k & 0 \\ sin\theta_k & cos\theta_k & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\widehat{\boldsymbol{C}}^{(k-1,\ k)} = \begin{bmatrix} cos\theta_k & -sin\theta_k & 0 \\ cos\beta_k sin\theta_k & cos\beta_k cos\theta_k & -sin\beta_k \\ sin\beta_k sin\theta_k & sin\beta_k cos\theta_k & cos\beta_k \end{bmatrix} \quad \text{(E.6)}$$

Position equation of the link k-1 with respect to link k can be written as:

$$\vec{r}_{k-1,k} = \vec{r}_{O_{k-1}O_k} = \vec{r}_{O_{k-1}B_k} + \vec{r}_{B_kO_k} \quad \text{(E.7)}$$

$$\vec{r}_{k-1,k} = \vec{u}_1^{(k-1)}b_k + \vec{u}_3^{(k)}s_k \quad \text{(E.8)}$$

The orientation and position of the link k in reference frame k-1 is written as:

$$\bar{r}_{k-1,k}^{(k-1)} = \bar{u}_1^{(k-1/k-1)}b_k + \bar{u}_3^{(k/k-1)}s_k$$

$$\bar{r}_{k-1,k}^{(k-1)} = \bar{u}_1 b_k + e^{\tilde{u}_1\beta_k}e^{\tilde{u}_3\theta_k}\,\bar{u}_3 s_k = \bar{r}_{k-1,k}^{(k-1)} = \bar{u}_1 b_k + e^{\tilde{u}_1\beta_k}\,\bar{u}_3 s_k$$

$$\bar{r}_{k-1,k}^{(k-1)} = \bar{u}_1 b_k + (cos\beta_k + \tilde{u}_1 sin\beta_k)\,\bar{u}_3 s_k$$

$$\bar{r}_{k-1,k}^{(k-1)} = \bar{u}_1 b_k - \bar{u}_2 s_k sin\beta_k + \bar{u}_3 s_k\, cos\beta_k \quad \text{(E.9)}$$

The matrix representation of (E.9) is:

$$\bar{r}_{k-1,k}^{(k-1)} = \begin{bmatrix} b_k \\ -s_k sin\beta_k \\ s_k\, cos\beta_k \end{bmatrix} \quad \text{(E.10)}$$

After substitution of (E.6) and (E.9) into (E.3), the transformation matrix of link k-1 with respect to link k becomes:

$$\widehat{H}^{(k-1,k)} = \begin{bmatrix} cos\theta_k & -sin\theta_k & 0 & b_k \\ cos\beta_k sin\theta_k & cos\beta_k cos\theta_k & -sin\beta_k & -s_k sin\beta_k \\ sin\beta_k sin\theta_k & sin\beta_k cos\theta_k & cos\beta_k & s_k\, cos\beta_k \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \text{(E.11)}$$