RESEARCH ON CONTROL SCHEMES WITH THE VISION SYSTEM FOR ADVANCED MECHATRONICS DEVICES

XUELING LUO

A DISSERTATION SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN
MECHANICAL ENGINEERING
YORK UNIVERSITY
TORONTO, ONTARIO

AUGUST 2020

ABSTRACT

With the rapid development of machine vision-related technologies, machine vision has drawn great attention from researchers in other engineering areas. In fields like mechanical engineering, a visual system could play a significant role in solving problems for robots' motion control. If the pose of a machine tool manipulator can be obtained easily and accurately, the control strategies for mechanisms would become more efficient. And this would highly promote newly designed machines into the real application. The trend of advanced manufacturing is that mechanical systems are moving towards full automation. However, during the transition, certain human involved processes of manufacturing still cannot be replaced completely. How to create a safe and efficient human-in-the-loop working environment with well-developed machine vision technologies would be one of the keys to the success of advanced manufacturing.

This study focuses on control schemes with the visual system for different types of mechatronics devices. This dissertation has been organized with three steps of studies. In the first step, we study how to promote well-designed machines for application. A vision-based motion control approach is proposed and validated for a recently designed generalized parallel robot. As we understand that it is hard to ignore that human-involved manufacturing is happening before the full automation, the second stage of this study is on how to design a control scheme for a robot to understand information from the surroundings and to be able to accomplish certain tasks by itself. A mobile robot motion control scheme is proposed and implemented in this study with the advanced visual technology to achieve the function of a certain moving target following. In the third part, the study focuses on an advanced Human-Robot Interaction with providing a control scheme for a mobile manipulator by signals from common hand gestures in real-time, which would contribute to the technology access ability in areas that robots has replaced intensive labor work but human experience is highly required. Each of these three parts has been present with a project that includes a control approach design for a specific robot system and the system validation experiment.

ACKNOWLEDGEMENTS

I would like to express my great gratitude to my supervisor Prof. Dan Zhang for being a great supervisor and mentor in the past few years. It was fortunate to me for working with Prof. Dan and his extraordinary research group. I greatly appreciate that he saw my research potential and accepted me as a candidate without any mechanical engineering background. To be encouraged by his dedication to international engineering research and education, not just my research ability in interdisciplinary areas have grown, I have also found my passion for my future career. With his mentorship and guidance, this doctoral program becomes an incredible experience.

I also appreciate my committee members Prof. George Zhu and Prof. Alex Czekanski for their constructive feedbacks in every meeting we had and for their kindness advice whenever I researched out. My great thanks to my exam committee members Prof. Siyuan He and Prof. Jimmy Huang for your kindness suggestions.

Thanks to those faculties and staffs who have shown their great support for my research, teaching, learning experience during those years. And my special thanks to Dr. Garrett Melenka for all his advice and guidance outside of research. I am lucky to work with an incredible research team and all my labmates' kindness support. I would like to thanks to my labmates Hongyan, Guangyu, Xiaodong, Haiqiang, Moritz, Andrea A., Andrea P., Qi, Naijing, Zhibo and Chong for their great ideas in terms of inspiring me to explore new research topics, their help in my study of mechatronics related topics, those hand-on experiences they have shared. And special thanks to Andrea A. for some experimental

assistance and contribution he had made for Chapter 6. Besides, thank you to my friends I have made in this wonderful department and those incredible minds which inspired and supported me throughout the past four years.

Thank you to my parents for your endless love and all your supports in every decision I have made. I also grateful for Haoduo's encouragement, accompany, love and commitment as a life partner. And thank you all for being my rock. Thanks to my supportive families and friends in my home country.

I dedicate this work to my uncle a passionate family chef whose encouragement for me pursuing this degree and his inspiration for my life was invaluable. Although he did not see my completion, his spirit will be with me and always in my cooking. May he rest in peace.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

2D	Two-Dimensional
3D	Three-Dimensional
AR	Augment Reality
DOF	Degrees of Freedom
EE	End-effector
EKF	Extended Kalman Filter
GPM	General Parallel Mechanisms
HCI	Human-Computer Interaction
HITL	Human-In-The-Loop
HRI	Human-Robot Interaction
IBVS	Image Based Visual Servoing
KCF	Kernelized Correlation Filter
KF	Kalman Filter
ML	Machine Learning

OOI	Objects of Interest
PBVS	Position Based Visual Servoing
PM	Parallel Mechanism
SLAM	Simultaneous Localization and Mapping
VR	Virtual Reality

Chapter 1 Introduction

What is a robot? I believe there would be a number of definitions from different individuals. The word “robot” was firstly invented in a science fiction play called “Rossum’s Universal Robots” by the Czech writer Karel Capek. In the past decades, fiction movies and novels already helped us visualized what robots could be. However, the reality is that existing robots are far behind the fantasy of films. There is no doubt that robot-related technology will play a significant role in this century.[1] Luckily, we already see a lot of smart robotic appliances active in our daily life, such as cleaning robots, smart speakers, even robotic waiters in certain restaurants. And we could simply summarize the future of robotics as machines with intelligence.

1.1 Background and Motivation of the Research

1.1.1 Background of the Research

To review the industrial development process, machines have done a lot for us and promoted us to move towards the fourth revolution of industry. Since the innovation of steam power in the late 17th century, industrial processes rapidly developed into the mass production century in the 18th century. Now we are in the information century with a highly automated manufacturing level and renovating towards the next stage. The next level of requirement for manufacturing processes would be higher productivity, better working environment, the flexibility of manufactory, etc. Machines have been doing great so far.

They have been significantly promoted production efficiency, they have significantly been released humans from working hazards, and they have been making our daily life convenient. With the diversity of demands in manufacturing and daily life, we start to expect more from machines such as instead of having massive numbers of the limited variety of products we expect for product diversity, instead of driving cars around we expect cars moving to the destination by themselves. And the smart machine will make those expectations into a reality.

The invention of machines derived from the need that humans would like to have replacements in the roles of heavy labour work. The process of its development could be simply described into three steps. In the first stage, we solved the actuation problem and the machine can move by itself with certain power sources. In the next stage, we solved the sensing problem with the development of sensor technology. So, the machine could get information from the surroundings and be able to accomplish certain tasks by itself. For the third stage, we expected that machines could not only operate as desired but also could ‘think’ as desired. It worth mentioning that the machine vision has played an important role in moving from stage two to stage three. Since the machine vision works as the function of human eyes. After it could see as humans, it is now developing towards ‘seeing and thinking’ as humans. Comparing with traditional sensing devices, the vision system has great advantages such as non-contacting sensing, easy for installing and maintenance and reusability. With the development of machine vision technology, it shows great potential in motion control for machines.

The vision system could understand the position or behaviour of the machine with the camera system mounting on the machine or monitoring the working process of the machine. In Figure 1-1, a typical vision-based robot motion control system is shown. The camera system is able to recognize locations of objects for picking up and the recognition result could contribute to the control strategies of the robot system. After years of study about the system's reliability and stability, this type of highly automated system starts to be applied to today's manufacturing. However, there are still a lot of challenges for this type of robot system such as payload, efficiency, flexibility which can be solved by researches about mechanisms. The robot arm in Figure 1-1 is a typical serial robot system with a large workspace and less complicated control strategies. However, it is also highly limited in load capability, flexibility and efficiency.

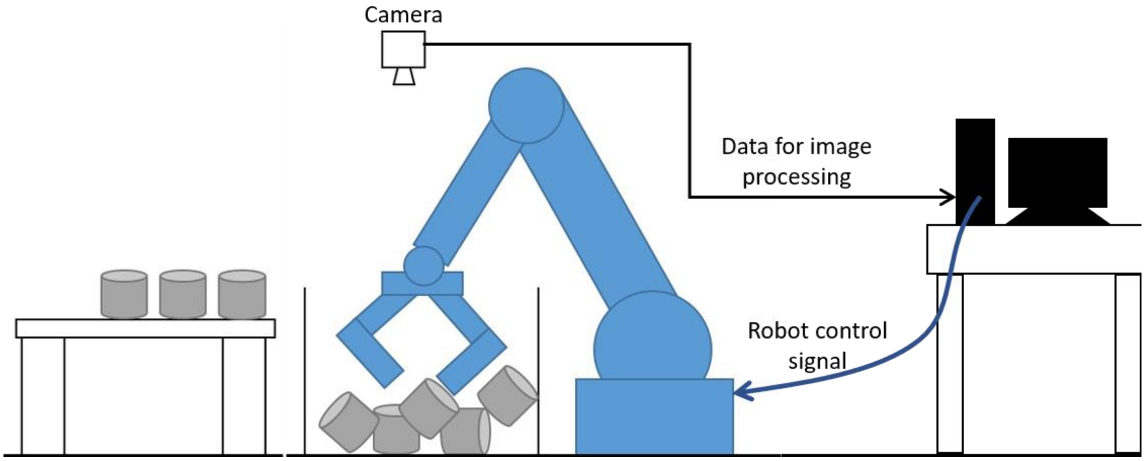


Figure 1-1. A typical robot motion control system with visual feedback

As the development of the industry, the future manufactory processes will be less labour intensive but become more flexible, efficient, reliable for high precision assembly, surface finishing and fast product handling[12]. To meet those demands, researchers on mechanism design have been focused on parallel robots and to solve problems of flexibility, payload, and efficiency of mechanisms. Today, parallel robots like 'Stewart Platform'[7], 'Delta-robot'[13], etc. are widely applied in manufacturing. Besides, parallel robots like [14, 109], etc. have shown great potential in modern manufacturing. Such mechanisms are normally consisted of multiple linkages to support one common platform or EE and formed as closed-loop chains [180].

With the boosting number of studies about parallel mechanisms, there are a lot of excellent designs shown in recent years. One of the challenges for promoting a well-designed parallel mechanism into the real application is that the validation economy and control efficiency. An increasing number of researches start to focus on coupling the visual servo technologies and parallel robots to create accurate, flexible and efficient control schemes. With the revolution of technologies and shifts of market demands, we must study the machine from design to application with new scopes. The visual servo will just become a part of the robot motion control strategies.

Studies show that most of the human jobs with lower-skilled would be replaced or eliminated with the development of technologies in countries like the US, German, Sweden and other Northern European countries [181, 182] and the remaining jobs need human involvement would become more comprehensive and complex [183, 184]. Although a lot of labour works will be replaced by machines in the soon future, there is the fact that hard to be denied that during this transition there will be a period that humans still involved in a lot of manufacturing processes. How to adapt advanced visual systems into human-in-the-loop manufacturing will become another popular research area. Visual devices have already widely been used as surveillance systems in a lot of scenarios, and a few numbers of advanced intelligent technologies have been widely applied commercially such as vehicle number plates recognition [113] and bio-information recognition [185, 186]. Integrating the exiting visual system with intelligence in a manufacturing scenario would provide a safe, reliable and efficient manufacturing system that would also require a high level of reliability and accuracy of the smart system.[187] The machine could work as an assistant, an alarm or a collaborator for a human involved manufacturing process.

Beyond visual servo for motion control of machines and advanced visual assisted system, in some specific industrial tasks which highly rely on experiences, the Human-Robot Interaction (HRI) will become an ideal manufacturing process before the smart machine becomes mature and reliable enough. Keys of this type of technology do not just make sure the system is accurate and dependable which is also about user experience. To optimize human involved manufacturing systems for better interaction, how to make new technology accessible to common users will also become an important research area such as an easy interface process, common communication between humans and machines, etc.

In a word, the research about adapting the visual system into robot motion control and HRI would play an important role in the development the advanced manufacturing. How to make the machine become smart will become one of the focuses of future manufacturing as well. The intelligent machine will become a general engineering study with a wide range of engineering fields involved such as mechanical, electrical and software engineering, etc. And there will be more and more new terms of research invented related to this study such as mechatronics, cyborg and bioroid.

1.1.2 Motivations of the Research

Technologies in machine vision have been highly valued in a lot of areas such as biometrics recognition[187], video surveillance analysis[111], medical imaging diagnosis[112], optical character recognition[113], defect detection in the industry [114], etc. and start to be recognized in areas as autonomous driving[115], augmented reality, etc. With visual applications in commerce and transportation that have been widely studied, have proved with high reliability, high accuracy and great flexibility and such researches become mature application products in the market, researches about visual applications in agriculture and industry would become the focus in the next decades. The expectation of machine vision technology in the area as industry and agriculture which traditionally have highly intensive labour involved workload would be how the technology helps the industry transformation with the shifting of the labour market caused by the new technology.

To adapt the vision system for advanced manufacturing, there are a few challenges that need to overcome. The first one is how the machine vision system could promote the development of the technology in terms of increasing the productivity of the manufacturing processes and how to combine the visual technology into machine design so the next generation of the machine would become not just efficient but also more flexible, more robustness to the working environment. As researches about mechanisms' design have proved those novel parallel mechanisms would suitable for more diverse requirements in the future industry theoretically such as with high flexibility, high load ability and high accuracy, how to evaluate the novel design with a prototype efficiently and accurately would be a big challenge for researchers in the area of the mechanism design[12, 14].

Although we already realized that the labour market for manufacturing would have a massive change with the development of advanced machines in the next decade, advanced machines get into manufacturing is unstoppable and human involved manufacturing process would still exist at the same time. During the transition stage of manufacturing, how to create a safe and friendly co-working environment between humans and robots will become a huge challenge for researchers as well. The machine vision system could bring intelligence into machines so the machine would work by itself or work with humans for specific manufacturing processes. Additionally, if the visual system could understand the manufacturing process and predict environment hazards, the safety of the manufacturing process will get into another level.

With the understanding that labour workers would be replaced by machines eventfully, the job market would start to shift according to novel technologies. Although manufacturing is working towards full automation, before machines become intelligent enough for working all by themselves, there are certain processes still highly rely on humans' experience. Besides, certain human involved processes of manufacturing cannot be replaced such as maintenance. Novel technologies have responsivity to build a smooth path for adapting new manufacturing processes with traditional workers. So how to make new technology accessible to common users such as easy interface processes, common communication between humans and machines, etc. will also become another challenge for researchers.

1.2 Research Objectives

The goal of this research is to develop control schemes for mechatronics devices with different types of visual technologies. According to the flexibility of machine vision technology, this study could be conducted into three levels of studies and the big picture of research objectives are as follows:

1) To design a methodology on how to integrate the visual system into control of well-designed mechanisms. The system should be able to control the motion of parallel mechanisms and provide evaluation information for design improvement. The system should have characters as reusable, easy for setting up, accurate, real-time and economy. Besides, the methodology should be easy to be adopted for other systems. With the result of this research object, the control scheme of a well designed parallel robot would have one more topic that a vision-based control strategy.

2) Although the visual system has been proved that it would improve the level of autonomy of machines and it would promote newly designed machines into applications, the fact that we would still have a large number of Human-In-The-Loop (HITL) scenarios in daily life or manufacturing processes. One of our research objectives would be to design a methodology on how to integrate advanced machine vision into controller design for the human-machine co-working environment.

3) One other research objectives of this study would be to design an easy interaction framework between humans and robots. Beyond vision-based control technologies, to optimize human involved manufacturing systems for a better interactive experience. The focus of this topic would be how to create an accurate, robustness interaction strategy that could easy understand by users with limited knowledge about the interaction knowledge itself.

1.3 The Dissertation Organization

This dissertation is organized as follows:

Chapter 1 has introduced the background of this dissertation which includes how the machine and machine vision get into our daily life and how they affected areas as commercial, manufacturing, etc. As the background have been introduced, the motivation of the study on vision-based motion control for parallel mechanisms, the research on advanced visual technology for robots’ motion control and vision system for HRI are presented. Then a brief about the dissertation organization is presented at the end of this Chapter.

In Chapter 2, a range of research areas that relate to this study are reviewed respectively since this study is interdisciplinary research among machine vision, mechatronics and mechanical engineering. Topics include machine vision, vision system for robots' motion control, parallel robot control strategies, real-time control for a mobile robot with a visual system, human-computer/robot interaction. In this chapter, related works under each topic are reviewed. Besides, discussion and challenges with those works are also presented in this chapter.

In Chapter 3, theories about stereo vision and related work that we have contributed to are presented. As the later chapters are all related to 3D vision, this chapter is the base of our following research. A novel circle detection method is presented for the camera's fast and accurate calibration.

In Chapter 4-6, three different branches of studies about vision-based control schemes are presented, respectively. In each chapter, a research project is presented for demonstration of the thesis' major contributions.

Chapter 4 present a novel vision-based control scheme for a well designed parallel mechanism. The system model is analyzed and verified theoretically. And the system prototype is built and demonstrated the efficiency and accuracy with the proposed model.

This chapter provides a new approach for a well-designed mechanism to create a control strategy that overcame challenges with the parallel robot forward kinematics.

In Chapter 5, an advanced target tracking method is proposed for a mobile robot motion control. And the experiment result has proved the efficiency of the proposed method.

In Chapter 6, a novel two-hand gesture recognition method is proposed for a mobile manipulator. The proposed method is designed according to certain common senses of gesture directional language which would provide a super easy understandably control signals from human gesture to intact with the movement of a mobile manipulator.

Chapter 7 concludes these three studies by discussing their performance and limitations. Besides, we have summarized this interdisciplinary study and its impact on the manufacturing process. Additionally, a few potential research topics are highlighted based on the current achievement of this dissertation.

Chapter 2 Literature Review

2.1 Machine Vision

The establishment of machine vision is derived from that human devoted to designing an information processing system that consists of cameras and computers, and the system could have similar functions as the human vision system. The system would be able to recognize and understand the surrounding environment as the same as humans through image processing technology. The camera system works as the signal receiving device; then, the signal sends to computers for processing. An image processing, recognizing and understanding system is build-in with the computer so the system works as one could understand the real world around it. The purpose of related researches is empowering computers with intelligence that could understand the real scenarios as humans. Machine vision as one of the most important disciplines under computer science is a cross-area subject that involves optical, electrical, image processing, artificial intelligence, etc.[3, 188]. The success of machine vision has to be a good understanding of all characters in its imaging process. According to Figure 2-1, it can be told that machine vision is a complex area that involves a wide range of engineering fields such as optical engineering, electrical engineering, software engineering, etc.

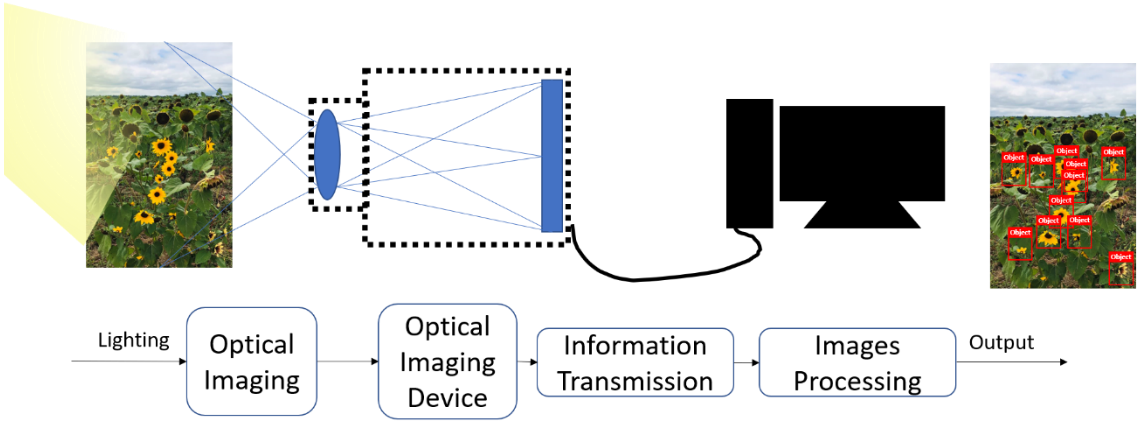


Figure 2-1. A typical imaging and image processing flow

How the imaging process obtain the information from a 3D world to the 2D image with a camera system? The general imaging model was simplified as a linear projection matrix to transfer the world coordinate to the image coordinate. The commonly used model is the pinhole model which shown in Figure 2-2.

The relation between the 3D point $N\ [X_W,\ Y_W, Z_W]$ and the point $n'[u,\ v]$ in the virtual image plane could express with the following equation:

$$\boldsymbol{c}\begin{bmatrix}\boldsymbol{u}\\ \boldsymbol{v}\\ \mathbf{1}\end{bmatrix} = \boldsymbol{M}_{\boldsymbol{wc}}\begin{bmatrix}\boldsymbol{X}_{\boldsymbol{W}}\\ \boldsymbol{Y}_{\boldsymbol{W}}\\ \boldsymbol{Z}_{\boldsymbol{W}}\\ \mathbf{1}\end{bmatrix} \tag{2-1}$$

where c is the scale factor and M_{wc} is the camera parameters which can be written as the combination of the camera extrinsic parameters and intrinsic parameters.

$$\boldsymbol{M}_{\mathbf{wc}} = \begin{bmatrix}\boldsymbol{R}_{\boldsymbol{w}}\\ \boldsymbol{t}_{\boldsymbol{w}}\end{bmatrix}\boldsymbol{K}_{\boldsymbol{w}} \tag{2-2}$$

where $[R_w \quad t_w]^T$ is the extrinsic parameter matrix for the camera to transform the world points to camera coordinates and K_w is the intrinsic parameters to map the camera coordinates into the image plane.

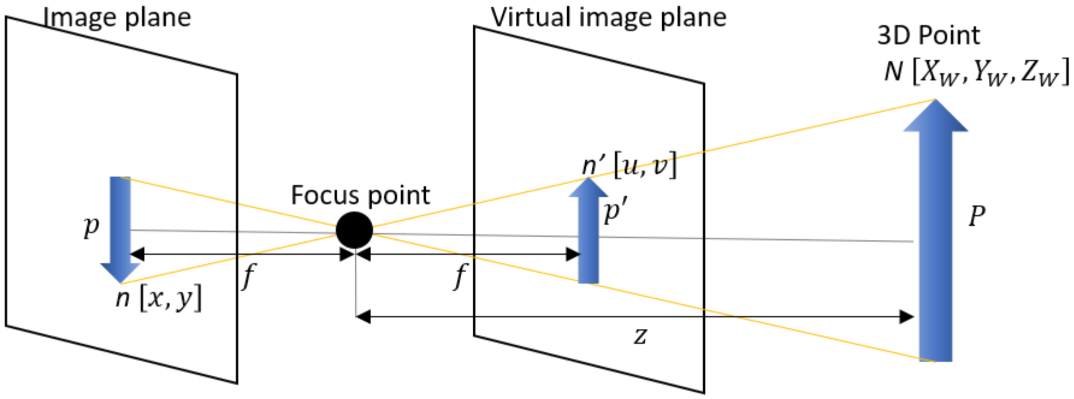


Figure 2-2. The pinhole camera model

During the camera parameter calibration, we could put the above parameters together to obtain a 3 × 4 camera matrix as follow:

$$\boldsymbol{M}_{\mathbf{wc}} = \begin{bmatrix} \mathbf{m}_{\mathbf{11}}^{\boldsymbol{wc}} & \boldsymbol{m}_{\mathbf{12}}^{\boldsymbol{wc}} & \boldsymbol{m}_{\mathbf{13}}^{\boldsymbol{wc}} & \boldsymbol{m}_{\mathbf{14}}^{\boldsymbol{wc}} \\ \mathbf{m}_{\mathbf{21}}^{\boldsymbol{wc}} & \boldsymbol{m}_{\mathbf{22}}^{\boldsymbol{wc}} & \boldsymbol{m}_{\mathbf{23}}^{\boldsymbol{wc}} & \boldsymbol{m}_{\mathbf{24}}^{\boldsymbol{wc}} \\ \mathbf{m}_{\mathbf{31}}^{\boldsymbol{wc}} & \boldsymbol{m}_{\mathbf{32}}^{\boldsymbol{wc}} & \boldsymbol{m}_{\mathbf{33}}^{\boldsymbol{wc}} & \mathbf{1} \end{bmatrix} \tag{2-3}$$

With the understanding of a camera imaging process, researches also notice that the distortion of the camera system caused by the manufacturing process would also affect the efficiency of understanding 3D information from the 2D image. So there are a lot of researchers focused on solving the distortion problems of the camera system. [189, 190]

During the 1950s, researchers start to focus on pattern recognition with two-dimensional image processing. This is the first time in human history that the computer has been used as an equipment to sense the real world. In those decades, the compute vision technology was mainly applied in optical character recognition and aerial image analysis [2]. In the 1970s, the Massachusetts Institute of Technology officially opened the course Machine vision which becomes an important milestone for machine vision application. [191] The 1980s is the decade that the number of researches about machine vision is rapidly growing. Additionally, in the middle of the 1980s, the diversity of researches about machine vision shown up. There were a lot of new concepts and frames proposed such as the framework of perceptual feature-based objects recognition theory, the theory of the active vision, the framework of vision integration etc. [192] With those new concepts and frameworks, new challenges showed up which promoted researches about related areas. After we are moving into the 20 centenaries, researchers started to devote to artificial intelligence which drives researches about machine vision towards general applications.[188] In the early 90s, there were a few commercial products in the market and those products had been applied to specific academic and industrial tasks. However, the system limitation did not promote the wide application. In the recent few years, researches about machine vision started to consider characteristics of biological vision such as high resolution, high robustness and real-time. Besides, with the rapid development of computer hardware, machine vision technology is moving towards the next generation. The machine vision shows great potential in areas as entertainment, commerce, the medical, industry even aerospace. It is not difficult to realize that machine vision technology has begun to enter people's daily life, bringing a lot of convenience and entertainment such as biometrics recognition on the cell phone, vehicle number plates recognition for a parking garage, intraoral digital scanners for dental treatment, etc.

The early vision was focusing on understanding environmental information and extracting objects of interest from images. Besides, features of the object are limited within a certain range like color, texture or shape to ensure the computing efficiency. With the development of artificial intelligence, computer vision starts to move toward a higher level in terms of duplicating the abilities of human vision [3]. Methods such as neural network [278], support vector machine [279], principal components regression [275], etc. highly support the development of machine vision. With those supervised learning methods, machine vision technology would be able to deal with more complex recognition problems with high accuracy like disease diagnosis assistance [278], facial recognition [276, 277], etc. Supervised learning is a typical machine learning method that could make decisions and predictions with the training dataset. With training datasets, a learning model could be created for predicting the response values for new datasets. In order to increase the accuracy of the model, a lot of efforts have to be put in gathering efficient datasets. Since image training methods also highly rely on labels of the object of interest, and effective labelling method would also be the key to supervised training. To increase the flexibility of the learning algorithm model for multiple target detection, there are a certain number of efforts putting on active learning like learning with multiple labels [284], multi-class learning [285], learning with Gaussian processes [286], multi-level prediction [287], etc.

With the potential of reinforcement learning and unsupervised learning shown in data training, researchers in machine vision also start to consider their possibility in vision recognition like facial recognition [288] and other image feature recognition [289, 290].

Although machine vision is rapidly growing with great potential for application in almost everywhere, to apply machine vision technology for more general applications, there are still many problems that need to be solved such as accuracy, real-time, etc. In the past a few years, the number of new methods has promoted the visual system becomes more intelligent [11]. Besides, there are a huge number of researchers still committed to this field and its related applications field such as virtual reality, and argument reality, etc. Applications with machine vision would be a boost at an unpredictable rate in the next decade.

2.2 Vision System for Robots' Motion Control

There are more than 800,000 robots have been widely applied in industrial environments and this population is continuedly growing [18]. However, the application for such robotic systems is excluded from those working or manufacturing scenarios where are required for high precision control and high reliability. Now the modern robots still highly rely on the robots' model to find the position of the end-effector (EE) the manipulator of the robots, especially highly reliant on forwarding kinematic models. So there are a few challenges for promoting application for such systems that need to be overcome. Firstly, sensor accuracy is a challenge for robot motion control. Although the number of sensing technologies is still growing, the commercial machines are still limited by the cost of high precision sensor systems, nonefficient sustainability, and requirements of maintenance of such systems. Secondly, there are some other factors that would cause errors of robots' final position such as manufacturing error in the machine parts. The other challenge for robot motion control is the stiffness of the robot. In order to avoid the error of the rigid body caused by material elastic deformation, the common solution to that is to build robots with high stiffness material, especially heavy metal in a lot of cases. As a consequence of that, manufacturing robots normally need to be driven by powerful motors but have limited capability in loading. As the machine vision technology has been highly valued as an alternative sensing technology which has high accuracy, non-contacting sensing feature and great flexibility as well as the rapid development of imaging devices commercial products, vision systems start to draw attention in the area of robot motion control study.

If the visual system can be efficiently used for robots' motion control, the control strategies for mechanisms would become more economical with high flexibility and high accuracy. Vision-based motion control for robots must be related to the research of the humanoid robot. There are a certain number of researches shown that vision plays a significant role in human motion control [72-74] that inspire researchers in robot motion control area to consider vision-based control schemes [75, 84]. As the expansion of robot application, various vision-based motion control approaches for different types of robots have shown such as eye-in-hand manipulator[61, 89] which has been considered in a lot of industrial scenarios, vision-based localization approaches for mobile robot control [90, 91] which have promoted a vital research area under artificial intelligence known as Simultaneous Localization and Mapping (SLAM) [92]. In the meantime, researchers also located challenges for this area, such as hardware qualities[79], image processing efficiency[85], image stabilization [84], image feature selection [87], etc. With the feature of flexibility, vision-based control has been applied for not just rigid mechanisms but also novel designs as cable-driven robots[55,62], continuum robots[94, 96] and flexible manipulator [95]. Continuedly, as the improvement of accuracy and reliability of image processing technology, vision-based control has shown great potential in medical applications as surgical robots[99, 100], medical assisting mechanical systems[98].

In the early stage of the development of vision-based vision control for industrial applications, researches are mainly focused on transfer the information from the real 3D world to 2D image accurately or rely on the information from the 2D image to estimate the position of the robot in the 3D scene [138]. The visual servoing is the most typical method which uses for transfer the position information to certain algorithm as the feedback of a robot autonomous control loop without human interaction. There are two main kinds of set up for implementing the vision-based motion control. The first one is an eye-in-hand system that equips the camera system on the robot gripper or manipulator to measure the location of the robot's end-effect to the target or object of interest [120, 123]. One of the most significant advantages of such systems is that with an ideal algorithm the effect of system default error would be minimized on the motion control accuracy. However, this type of approach needs typically to continuously update the target location until the manipulator reaching the object which would be a computationally expensive process. Besides, such systems could only have the location of the object but the coordinate of the gripper which means that the feedback control does not include the location of the end effect of the system. Additionally, if there is any interference from the environment which causes the error between the end effect to the destination, the system could not be able to make the adjustment for that. The other type of vision-based motion control setup is eye-to-hand which fixes the camera system in a static position and has the robot manipulator on the camera scene.[122] The advantage of such an approach is that the camera system could work as not just the sensor between the manipulator and the object of interest but also could work as the surveillance system for the whole robot system as well as the working environment. Under this case, the camera system could assist the robot system in finding a path or planning for an efficient grasping without conflicting any environment uncertainty. Challenges for such approaches are how to understand the location of the target in a complex 3D environment and how to plan an ideal trajectory of the robot motion which are out of the scope of machine vision or robotic motion control itself [123]. There are also some researches to combine the above two types of approaches together for reliable and robust control schemes[124 -126]. However, those combinations highly increased the cost of the implementation.

As advantages of the vision-based control shown, researchers also notice that one of the keys to increasing the system accuracy is to get the 3D information directly from the vision system. To compare the eye-in-hand system with a monocular camera and a stereo camera system, the system with a stereo camera would provide more flexibility in a complex working environment such as multiple nearby objects picking up, small objects picking up, etc.[121]

Visual servoing methods which are utilized as feedback from image features to control the motion of robot systems have been proposed to handle many stability and reliability issues in vision-based control systems. Beyond visual servoing for robots, a similar idea is proposed for human motion capture [154]. Such a system could detect the motion timely and accurately. Besides, those ideas could contribute to a wide range of areas like the medical and entertainment industry. With the department of artificial intelligence in data science, machine learning (ML), Virtual Reality (VR)/ Augment Reality (AR), there will be more and more novel technologies and control schemes that appeared in robot motion control-related research.

2.3 Parallel Robots Control Strategies

Parallel robots have drawn great attention with their higher stiffness, higher force-to-weight ratio and efficiency compare with serial robots [155]. In terms of the parallel robots control scheme, it could be generally designed from two frameworks: designing from leg space/ actuator space or designing from the workspace. The former approach focuses on designing a control scheme to compute the actuation signal from a desired moving platform position by inverse kinematic. If all actuators are well controlled in the system, the moving platform or manipulator of the robot could move desirably. One of the challenges for such control schemes is the dynamic model for actuators. To take hydraulic actuator as a typical example, the dynamic model of hydraulic systems is complicated as they resemble velocity sources, they are forces or torques dynamic range limitation and highly nonlinear [156, 157, 158]. So there are a certain number of researches devoted to the actuator modelling for creating high-performance controllers for parallel robots by minimizing fluid compressibility or flow pressure characteristics [159, 160, 161, 162, 163]. On the other hand, there are also researchers focusing on taking advantage of hydraulic systems for creating control schemes with high responding speed under large forces and high stiffness [164, 165, 166, 167]. Some researches are focusing on improving robots' dynamic models by considering removing the mass and inertia of the parallel robots' linkages [169, 170] which could provide simplified dynamic models for controllers with a certain range of error. The second approach is to describe the dynamic system model by the workspace coordinates which means the position of the end-effect of the robot needs to be provided. Normally, this type of measurement is done by orientation sensors. Without those direct measurements, the workspace coordinates would be calculated by the robot forward kinematic model which is well known as a challenging problem for parallel robots [168, 171, 172, 173]. Although we could use a sensor to make the measurement happening, the challenge is the cost of such a 6 degree-of-freedom (DOF) sensor and its maintenance [174, 175]. As this approach could directly evaluate the system performance, there are a lot of researchers devoted to solving problems in finding the desired way to obtain the parallel robot forward kinematic [176, 177, 178]. Although above mentioned parallel robot control strategies are devoting on making an effective control scheme possible, the parallel robot kinematic could still be highly nonlinear and the performance of the robot could still be unpredictable with modelling uncertainties. [179]

A vision-based sensing system as the most promising novel sensing technology has drawn great attention from researchers in the arear of robot motion control. The visual servo system for parallel robot control is first introduced by Koreichi and Kallio[5, 6] and becomes an important research aspect in parallel robots' motion control. With essential features of parallel robots as high load capacity， flexibility high stiffness shows excellent potential in modern industry as well as the advantages of visual sensing technology, visual servo becomes increasingly important research subject in motion control. For serial robots, the vision-based control kinematic model could provide a Cartesian desired velocity which could be easily and relatively accurately converted into robots' joint velocity with the inverse kinematic model. In the case of parallel robots, the inverse kinematic model highly depends on the Cartesian pose which needs to be estimated. [101-103] Traditionally, parallel mechanisms are designed with particular structures to have analytical forward kinematic models and the inverse kinematic models would be formed with estimation poses of the end effect according to the forward kinematic models correspondingly. [7, 69, 104] Consequently, there are two main challenges in terms of motion control to promote those well-designed parallel mechanisms into a real application. The first challenge is to kinematic modelling accuracy. Researchers not just in mechanical engineering but also in mathematic have been devoted to such studies.[7, 105] However, solving the forward kinematic model for parallel mechanisms are always one of the great challenges for modelling. The vision-based control could provide feedback from direct measuring results which efficiently replaced the role of the forward kinematic model and simplified the control schemes for parallel mechanisms.[1] To measure the end effect of mechanisms in 3D, the visual servo technology is widely studied for creating control schemes with 3D poses measuring[23, 106]. The six degrees of freedom (DOF) platform Gough-Stewart [7] is one of the typical examples. Once the visual system directly measures the pose of the end-effect, the inverse kinematic model is obtained with the visual feedback[65] which would minimize errors from modelling and create a robust control scheme. There are two main types of visual servoing: Image-Based Visual Servoing (IBVS) and Position-Based Visual Servoing (PBVS) with different transfer approaches of frames in the system.[23, 29, 106] Both of them are widely used in robot motion control. These two approaches would have a lot of differences in detail. But we could summarize and introduce them as follows. The control block of IBVS and PBVS could be found in Figure 2-3 and Figure 2-4, respectively.

The IBVS is an image processing focused approach which the main idea is to minimize the position error of the object in the image plane. This method does not need a pose model of the object of interest. Features of the object in the image are important. Besides, features of the object desired position in the camera image must know, so this approach could keep compare the current position with the desired one and drive the system to the desired pose. In this case, the control law designed with the image plane is critical for the system efficiency which means a reasonable choice about image features would be significant. The features could be points coordinates, lines, or geometric parameters in the image plane.[1, 22, 25] Since this method is processing in the image plane, one of the most significant advantages of this approach is the system would be less sensitive to the camera's intrinsic or extrinsic parameters. However, also because of the 2D control process, this approach is relatively difficult for a 6 DOFs motion control [193, 194].

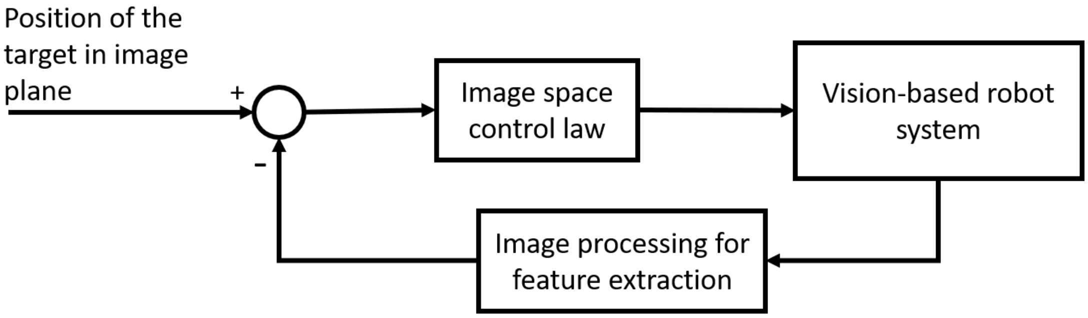


Figure 2-3. The control block of image-based visual servoing

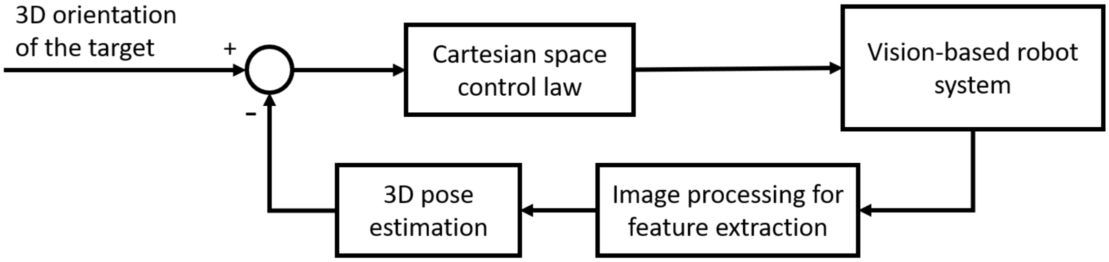


Figure 2-4. The control block of position-based visual servoing

The PBVS is based on a geometric model of the object of interest and highly relied on the known camera parameters. This approach is to estimate the 3D position of the object and to find the relation from the camera to the object, then the comparison result of the current and desired 3D pose would drive the robot to the desired position. It is straightforward to compare the current directly and desired position and easier to plan the motion for the robot comparing with IBVS.[195, 196] However, this method is challenged by the accuracy of calibrated camera parameters and the method which would profoundly affect the accuracy of the 3D pose estimation process.[197]

To integrate the advantages and discard the drawbacks of these two approaches. Some hybrid methods have been proposed by integrating features of the object of interest in both 2D and 3D space [198, 199, 200].

With new ideas about parallel manipulators shown, the research about their control strategies would still be a popular area. Focuses in this area would not only just on the improvement of dynamic modelling for parallel robots but also on adapting highly growing technologies such as machine learning for minimizing computing cost or for predicting control disturbances. With the development of novel sensor technology, there will be more and more researchers join this area with their unique perspectives.

2.4 Real-time Control for A Mobile Robot with A Visual System

Mobile robots are one of the most commonly used robotic systems in an industrial or commercial application. As its relatively simple dynamic model, a mobile robot is one of the first types of robots that have been widely used for studying novel control schemes such as vision-based control or HRI. Traditionally, mobile robots are controlled based on human experience similar to the driving process. The human visual system is feedback or sensing process during the motion control to evaluate if the robot is moving desirably or not. Then the control signal of actuators on the robot would be adjusted by human operation for making changes for the motion. With the development of sensing technology, mobile robots start to move towards full automation. Distance sensors have been equipped on a mobile robot for avoiding obstacles so the robot could move in a designed scene without collision. [201] The more advanced system could able to plan a moving path with certain prior knowledge about the environment. [202] According to the experience of manual control of a mobile robot, researchers are expected for a powerful scheme that is comparable to manual control or more advanced. The visual system has changed researchers' vision of mobile robot motion control. When the robot equipped with 'eyes', focus on its motion control becomes how to make it 'think' like a human. In the last decade, there are tones of research about vision-based mobile robot motion control. The mobile robot starts to be able to understand the static environment around it and would be able to plan a path for itself to move towards a given destination [203]. The more advanced system would be able to find multiple paths for its movement and to choose the best one for later action [204]. There are also some derivative applications about such systems like applying a mobile robot with a visual system for autonomous 3D environment information reconstruction [205]. However, most of those applications are not time-sensitive. Researchers believe real-time application would still be a great challenge for mobile robot motion control and visual technologies could make mobile robot systems even more flexible for different task requirements. One of the biggest challenges for real-time applications is how to process complicated surrounding information timely for mobile robot motion control. Some researchers work simplifying the surrounding information by reducing image features so there would be less kind of information for the visual system to process and more would be more efficient in generated control signals [206]. In a dynamic environment, existing visual technologies already been about to understand or capture the information of interest such as pedestrian recognition and tracking, human behaviour identification and environment hazard learning [207]. How to apply such technologies for real-time mobile robot control would be a challenge. There are some researches devoted to developing novel learning or tracking technologies so the mobile robot could be able to work in a real-time application such as patrol without collision or avoiding dynamic pedestrians [208, 209]. Besides, after the robot is able to understand the working environment timely, how to respond to the environment is the next stage of the mobile robot motion control. Some contributions have been made in path planning for a mobile robot in a dynamic environment [210, 211]. As a result, there is an important research area about this topic is the autonomous car which has been a popular area in the last decade and would still be an important research area in the next decade. As there are some limitations on visual sensing technologies such as sensing range and lighting sensitivity, the future of the motion control of mobile robots would also move towards senor fusing [212, 213, 214].

2.5 Human-Computer/ Robot Interaction

With computers become one of the most important kinds of electronic devices in our daily life, users are looking for novel interaction ways with computers for applying the computer system in more complex scenarios such as commercial, education, entertainment, etc. So the interaction between users and computers can no longer be limited with interaction devices like keyboards and mice. With the development of wireless technologies, interaction devices move toward interacting-in-distance like wireless keyboards, mice and laser pointers. The great potential of interacting-in-distance starts to show in areas as commercial and entertainment. Researches are inspired and devoted to more user-friendly interaction formats such as Human-Computer Interaction(HCI) by tracking laser point on projection aera[142], sensor-based somatosensory gaming with Nintendo Wii, or the more advanced HCI with Kinect by Microsoft based on tracking human skeleton. In recent years, as the rapid development of computer hardware, computers in a small or even micro size become affordable and have been equipped in a wide range of devices or machines such as robots. With inspiration from HCI, research starts to work on the communication between humans and robots. HRI is the collaboration between humans and robots which is a popular area in robotics research in the past few years. The main ideas of such types of systems are to understanding orders send by humans and operate to accomplish certain tasks with humans' experiences, to build the communication between robots and humans, to promote applications of robots in daily life and industrial environment. As a result, the main evaluations about such systems are focus on the capabilities of the design of the robot, how well robots are trained, etc. In nature, this research draws increasing attention in disciplines as artificial intelligence, mechatronics, computer vision, etc. and becomes an interdisciplinary research area.[49] As the goal of the interaction is to work together to accomplish a goal, robots could be introduced as co-workers, helpers or assistants to humans[37,38,45,47]. There are a few certain factors that could be used for evaluating an

HRI system: autonomy level, the natural level of information exchange, how well the collaboration in terms of accomplishing a given task and adaptability, reliability of the system.

The most common and efficient existing interaction method with computers/robots is using body language for communication specially hand gestures. There are certain numbers of researches devoted to increasing the efficiency of hand gesture recognition such as creating a model for hand pose estimation [143], mapping the shape of certain gestures for recognition [144], using motion history image for recognition [145], etc. There are also other methods for HCI/HRI such as wearable devices [146] or recognizing other human body parts [147]. The reason that body language is widely used for interaction is that certain features from those gestures are relatively easy or convenient to be extracted from the background and understood by computers. Now such methods become mature for real-life applications in medical [148, 151], robot motion control [149] and entrainment [150]. With the development of machine learning technologies that computers would be able to understand more complicated features in real-time, there are more alternative ways shown for HCI/HRI such as dialogue recognition [152] and eye movement [153].

Chapter 3 Vision-based Three-Dimensional Measurement

3.1 Stereo vision

As the machine vision is to try to help computers understanding the 3D world from 2D images, there are a lot of efforts have been put on 3D reconstruction from 2D images. The most common method that has been used is the stereo vision with two cameras following the binocular vision theory. The binocular stereoscopic calculates the corresponding position of the two images of the deviation relationship to obtain the measured object of three-dimensional information which has numerous advantages such as efficient, accurate, simple system, low cost, etc. Such systems are also called passive stereo vision as they consisted only of image receiving devices. The working principle of passive stereo vision could be described as a geometry theory which is called epipolar geometry which is shown in Figure 3-1. In one on the two-dimensional view, each pixel is a line in the real 3D space. In the other view, this line in the pixels becomes a line in the new view which provides the corresponding information to calculate the 3D coordinate of each point in 3D space. [17]

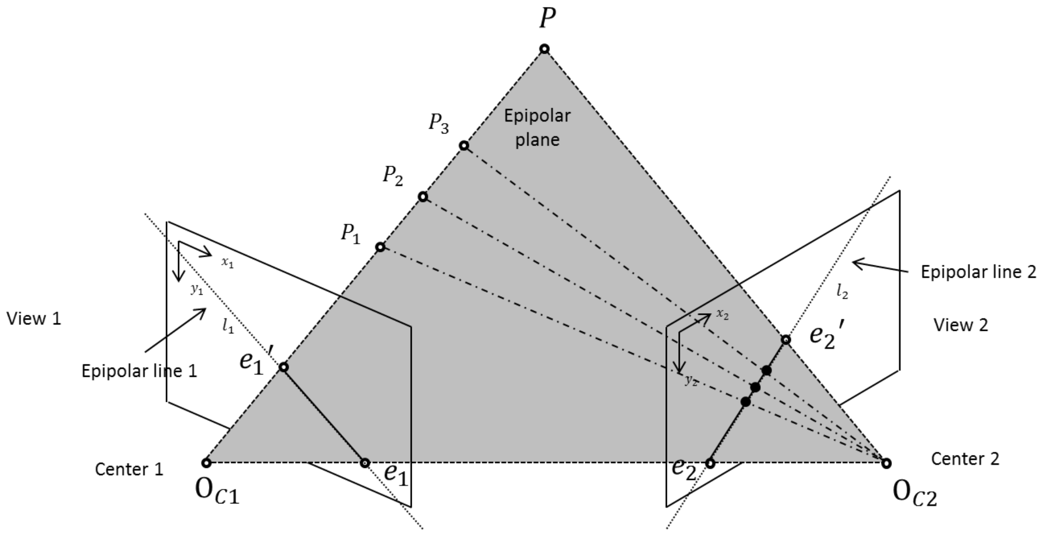


Figure 3-1. Epipolar geometry

To obtain 3D information by binocular vision could be divided into the following six steps: camera calibration, image acquisition, feature extraction, stereo matching, 3D information recovery and processing of recovery information.

3.2 Camera calibration

Camera calibration is a process that must be included in 3D vision for building the relationship between 2D images and 3D real-world information. The technique of camera calibration can be generally classified into two categories: 1) the traditional camera calibration method. Under this type of approach, it is necessary to put a reference object in front of the visual system, so the internal and external parameters of the camera could be calculated by a 3D reconstruction process with a set of known 3D points' coordinates. This method could provide high precision, but this process also limits the camera system in a fixed location. Camera parameters need to be updated once the standing location of the camera is changed. 2) the camera self-calibration. Directly use constraints relationship in image views to calculate the internal camera parameters. This method could be convenient for a special working environment such as a moving visual system.

The camera calibration process is to build the relationship between the real-world and cameras. So we put a reference object in front of cameras and to establish the relationship between 3D coordinates of the reference object which are known and its image in cameras to calculate the parameter of cameras. We called the reference object standard target which should put in front of the cameras also has certain features to be detected. Figure 3-2 shows the configuration of the cameras' calibration.

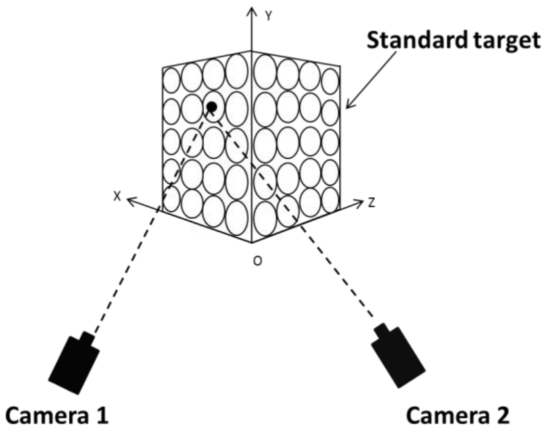


Figure 3-2. Binocular cameras calibration configuration

In Figure 3-2, the standard target we applied has multiple circles as the target feature. So one of the challenges of this calibration process is to detect centers of the circles on the target in image views. This detection process is called feature extraction.

3.3 Feature Extraction

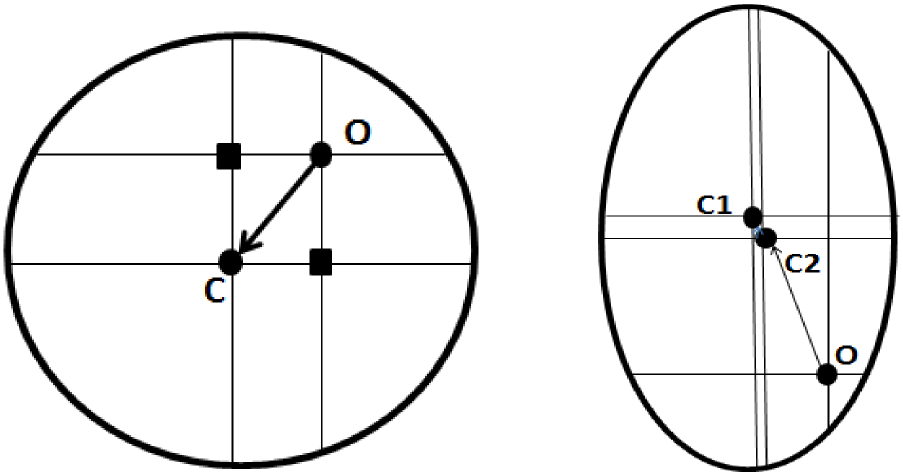
In this part. a feature extraction method we used in the calibration is introduced. In the calibration, we use a circular target as a three-dimensional target detection target instead of the traditional calibration of the checkerboard as a detection target [270]. In the detection of the checkerboard, the Harris corner detection method [133] is usually applied which would impact by the error from the black and white edge of the checkerboard. To avoid this problem, centers of circles as feature points are created for detection.

The feature point extraction method of the center detection method is a feature point detection method based on geometric principles. At first, we detected the edge of each circle or ellipse. Then, a recursive detection starts at any point inside of the circle or ellipse. The detection process is presented as follow:

- To start the detection from a random point O, a horizontal line, as well as a vertical line, are drawn from the point;
- According to the above two lines, the intersection point C is found by drawing the bisector line for the two lines, which could be found in Figure 3-3 (a). Then the distance between the point O and the point C need to be calculated;

- Repeating the above processes, the position of the next point C_i could be found according to the location of point C. The distance between the point C and the point C_i need to be calculated;

- When distance between C_i and $C_{(i+1)}$ are less than 1 which means the distance less than one pixels, we believe that $C_{(i+1)}$ is the center of a circle



(a)Center detection in circle (b)Center detection in ellipse

Figure 3-3. Center detection configuration [270]

The same process could be used for processing the center detection of an ellipse. In Figure 3-3 (b), the distance between C_1 and C_2 is less than 1, so we find that C_2 is the center of the ellipse.

The above present idea is a method that could be widely used for feature detection with simple features such as labels or markers' detection. For complex target detection such as a moving human target or a dynamic hand gesture, there are more advanced feature extraction approached applied for detection. Instead of find edges or points like this approach, the feature of the target is normally considered as an image pitch from a whole image view. Mathematically, the feature would be considered as a feature matrix that contains features of the target like texture, color or optical flow. In chapter 5, an advanced feature extraction method is presented for moving target detection. There are also some hybrid approaches for target detection as the approach for hand gesture recognition we present in chapter 6. The 3D geometry and the morpho of the target are combined as the feature of the target for detection.

3.4 Three-dimensional Information Reconstruction

Once the camera system parameter is obtained and the feature of the target is detected from the image view, the next step is to understand the location of the target in the 3D space or respect to the camera system in the space. To take the classic stereo vision setup as an example, stereo matching is a necessary process for 3D information reconstruction.

When two paired images from two cameras in a stereo vision system are obtained, any feature in two image views could be matched with the camera calibration result. According to the epipolar geometry theorem, in a binocular vision system, a feature point in one of the cameras has a unique correspondence in another camera. According to the definition of the camera definition in equation (2-1) to (2-3), the camera parameter matrix could be obtained with the calibration process. After calibration, we have two matrixes $\mathrm{m}^{\mathrm{mc1}}$ and

$\mathrm{m}^{\mathrm{mc2}}$for the two cameras in the system. According to the epipolar geometry shown in Figure 3-1, the Normalized Cross-Correlation [271, 272] is applied for effectively matching the corresponding point in a paired image. An experiment result is shown in Figure 3-4, there are five pairs of corresponding points shown in a calibrated camera system.

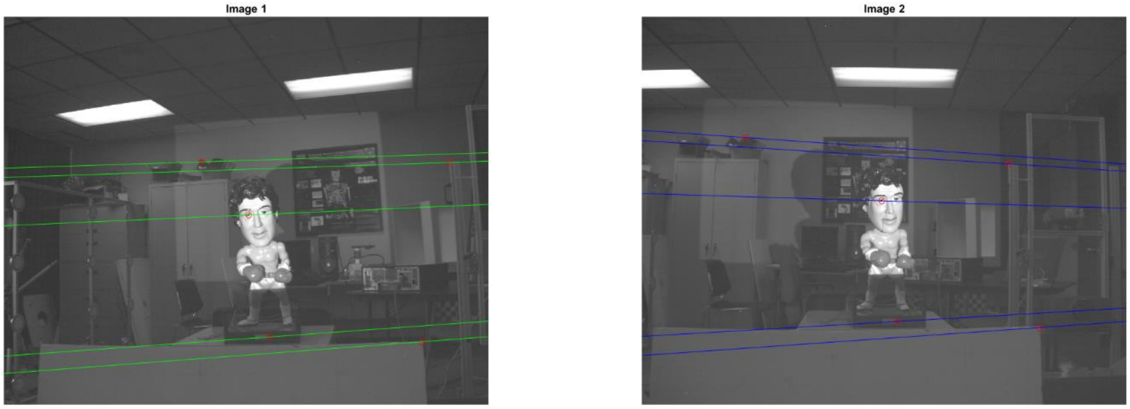


Figure 3-4. Stereo vision matching result from two camera views

After paired features are detected in image views, the 3D information of the object could be calculated. With the camera parameters and detected features in image views, the 3D coordinates $\mathbf{W} = [\mathrm{X}^{\mathrm{W}} \quad \mathrm{Y}^{\mathrm{W}} \quad \mathrm{Z}^{\mathrm{W}}]^{\mathrm{T}}$ of a pair of corresponding points $[\mathrm{x}^{\mathrm{c1}} \quad \mathrm{y}^{\mathrm{c1}}]$ and $[\mathrm{x}^{\mathrm{c2}} \quad \mathrm{y}^{\mathrm{c2}}]$could be calculated by the equation

$$\mathbf{W} = (\mathbf{A}^{\mathbf{T}}\mathbf{A})^{-\mathbf{1}}\mathbf{A}^{\mathbf{T}}\mathbf{b} \qquad \textbf{(3-1)}$$

where

$$\mathbf{A} = \begin{bmatrix} \mathbf{m}_{11}^{\mathbf{wc1}} - \mathbf{m}_{31}^{\mathbf{wc1}}\mathbf{x}^{\mathbf{c1}} & \mathbf{m}_{22}^{\mathbf{wc1}} - \mathbf{m}_{32}^{\mathbf{wc1}}\mathbf{x}^{\mathbf{c1}} & \mathbf{m}_{33}^{\mathbf{wc1}} - \mathbf{m}_{33}^{\mathbf{wc1}}\mathbf{x}^{\mathbf{c1}} \\ \mathbf{m}_{21}^{\mathbf{wc1}} - \mathbf{m}_{31}^{\mathbf{wc1}}\mathbf{y}^{\mathbf{c1}} & \mathbf{m}_{22}^{\mathbf{wc1}} - \mathbf{m}_{32}^{\mathbf{wc1}}\mathbf{y}^{\mathbf{c1}} & \mathbf{m}_{23}^{\mathbf{wc1}} - \mathbf{m}_{33}^{\mathbf{wc1}}\mathbf{y}^{\mathbf{c1}} \\ \mathbf{m}_{11}^{\mathbf{wc2}} - \mathbf{m}_{31}^{\mathbf{wc2}}\mathbf{x}^{\mathbf{c2}} & \mathbf{m}_{22}^{\mathbf{wc2}} - \mathbf{m}_{32}^{\mathbf{wc2}}\mathbf{x}^{\mathbf{c2}} & \mathbf{m}_{33}^{\mathbf{wc2}} - \mathbf{m}_{33}^{\mathbf{wc2}}\mathbf{x}^{\mathbf{c2}} \\ \mathbf{m}_{21}^{\mathbf{wc2}} - \mathbf{m}_{31}^{\mathbf{wc2}}\mathbf{y}^{\mathbf{c2}} & \mathbf{m}_{22}^{\mathbf{wc2}} - \mathbf{m}_{32}^{\mathbf{wc2}}\mathbf{y}^{\mathbf{c2}} & \mathbf{m}_{23}^{\mathbf{wc2}} - \mathbf{m}_{33}^{\mathbf{wc2}}\mathbf{y}^{\mathbf{c2}} \end{bmatrix} \quad \textbf{(3-2)}$$

$$\mathbf{b} = \begin{bmatrix} \mathbf{m}_{34}^{\mathbf{wc1}}\mathbf{x}^{\mathbf{c1}} - \mathbf{m}_{14}^{\mathbf{wc1}} \\ \mathbf{m}_{34}^{\mathbf{wc1}}\mathbf{y}^{\mathbf{c1}} - \mathbf{m}_{24}^{\mathbf{wc1}} \\ \mathbf{m}_{34}^{\mathbf{wc2}}\mathbf{x}^{\mathbf{c2}} - \mathbf{m}_{14}^{\mathbf{wc2}} \\ \mathbf{m}_{34}^{\mathbf{wc2}}\mathbf{y}^{\mathbf{c2}} - \mathbf{m}_{24}^{\mathbf{wc2}} \end{bmatrix} \quad \textbf{(3-3)}$$

According to the stereo vision theory, other types of 3D vision systems could apply a similar approach to reconstruct 3D information from images or combined image-sensor information.

During the validation, the proposed method is achieved with the following pseudocode in Figure 3-5. In the later chapters of the thesis, the main visual system applied for vision-based control is an RGB-D camera which consists of one RGB view and one depth map view which could be found in Figure 3-6. Instead of providing traditional feature information, the depth map is providing the depth length from the scene to the camera.

```
Input  : captured image of the standard target
Output: calibration result of the camera system $M_{\mathrm{wc}}$ and $z_D$

edge detection of circles
for  i = 1: 40
    j = 1;
    for  j = 1:10
        to detect from a random point O in the circle i;
        to draw a horizontal line $h_j$ and a vertical ling $v_j$ through point O
        finding the intersection point $C_j = [x_{c_j}, y_{c_j}]$ of the above to line;
        to draw the horizontal line $h_{j+1}$ and a vertical ling $v_{j+1}$ through the middle
        points of $h_j$ and $v_j$, and find the intersection point $C_{j+1} = [x_{c_{j+1}}, y_{c_{j+1}}]$;
        if  $|C_{j+1} - C_j| < 1$
            center of the circle i located at $D_i = [x_{c_{j+1}}, y_{c_{j+1}}]$;
            break;
        end
    end
end
calculating the $M_{\mathrm{wc}}$ through equation (2-1) with above result $[D_1, D_2, D_3, \ldots, D_i]$
edge detection of the standard target in the depth map
calculating the $z_D$ through equation (3-4)
```

Figure 3-5. The calibration algorithm

Figure 3-6. Images obtained from the Kinect camera: RGB image and depth image

The calibration process would be similar to the binocular cameras system calibration. For the RGB camera calibration, we applied the above-mentioned process to detect the features on the standard target. The detection result is shown in Figure 3-7. According to equations (2-1) to (2-3). The camera parameters could be calculated. The different part of this process compare to the binocular system is the parameter calibration with the depth map. The parameter for the depth sensor could be expressed as

$$\boldsymbol{z_D = \frac{1}{c_{D2}d_u + c_{D1}}} \tag{3-4}$$

where c_{D1}

and c_{D2} are the depth camera parameters, and the d_u is the undistorted disparity [273, 274].

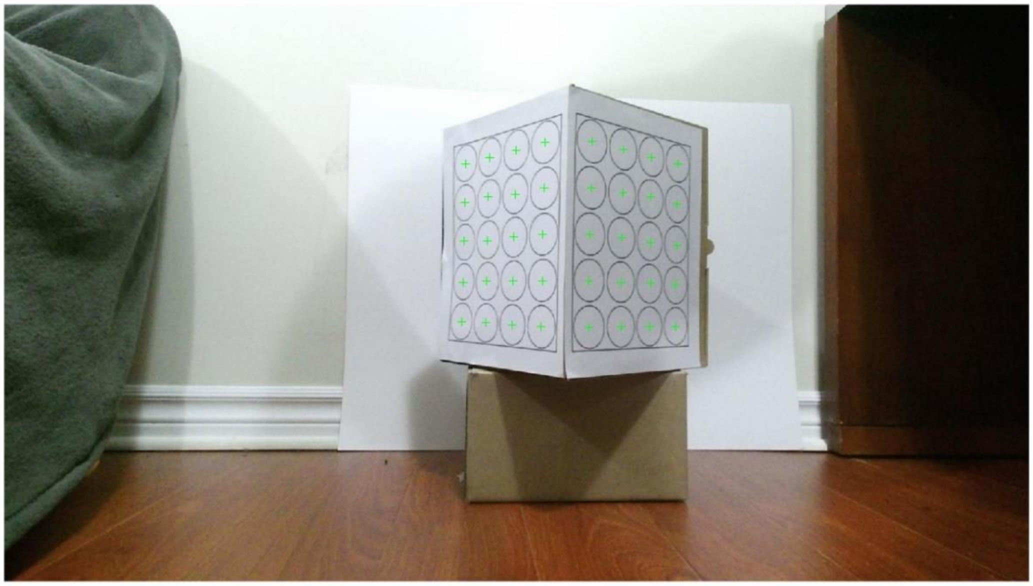


Figure 3-7. Color camera calibration with the feature detected result

As the depth map is not able to provide texture or color information, the calibration needs to be processed with certain known depth information. While building the standard target, the 3D dimension of the target is also measured as pre-knowledge. So during the calibration, the depth map detection would focus on finding edges of the standard target. The detection result could be found in Figure 3-8. With the equation (3-4), the depth image view could be calibrated with the known depth information from the target.

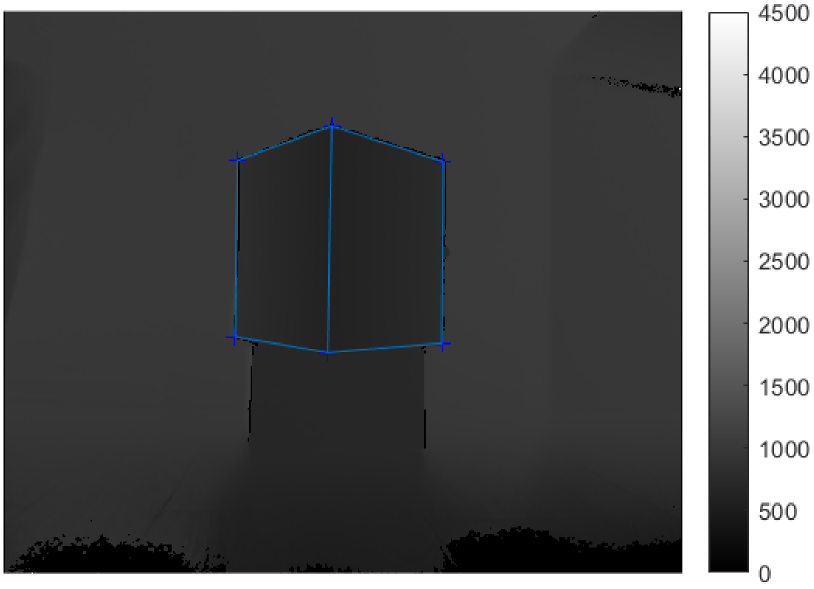


Figure 3-8. Depth camera calibration with the feature detected result

It would always be recommended to calibrate a visual system before applying for measurement or applications since the manufacturing process would make almost every individual device has its own parameter, especially for accurate measurement. However, in certain application scenarios, it is not very necessary to calibrate the camera system. For example, in chapter 5 a moving target following method is presented which not really need the camera to be calibrated since the dynamic system would need a relatively larger tolerance on error.

3.5 Discussion

In this chapter, the foundation of 3D measurement via the vision system is presented through the analysis of a stereo vision 3D reconstruction process. With the understanding of the theory of the stereo vision, a 3D reconstruction process is derived with three steps. First, the geometry relationship between the camera and the real-world has been built with the camera calibration process. In the feature extraction process, we have introduced a novel feature detection strategy for fast and accurate camera calibration. The key to the accuracy of calibration is to detect features from the standard target. This method has avoided the detection error in traditional checkboard detection causing by the corner detection error. Additionally, since the standard target proposes is designed as a 3D target, the camera calibration could be finished with one picture taking from the system instead of multiple images. This process also benefits the calibration for the depth camera, which needs a depth standard for the calibration process. Once the coordinates on the standard target have accurately detected, the 3D measurement for the location of the object of interest could be achieved with the camera parameters.

Chapter 4 Vision-based Control: An Approach for Generalized Parallel Robots

4.1 Introduction of Generalized Parallel Robot and Its Control Strategy

Parallel mechanisms (PM) have shown great potential in industrial or manufacturing applications such as pick-and-place and high-speed operation which requires high repeatability and high load [50, 51, 52]. As the traditional PMs, in general, are built by connecting rigid platforms and limbs which are parallel arranged [53]. This type of typical design is also limited to such systems in terms of workspace, the capability of platforms' rotation. Researchers in mechanisms design are devoted to solving such problems in the past few years. There are groups of general parallel mechanisms (GPM) proposed to overcome the platform rotational limitation and workspace limitation such as a family of singularity-free redundant planar parallel mechanisms [51] and a class of generalized parallel mechanisms with large rotational angles [52]. Such systems are validated theoretically with their new features which would benefit the future manufacturing in terms of productivity, efficiency and flexibility, to verify those features with prototypes become urgent and necessary. However, with a large number of links and a certain amount of passive joints, precision and accuracy are challenges for PM and GPM which are also the reason that the number of different types of PMs for such applications is still limited.

With the development of additive manufacturing, to build a prototype for a well-designed mechanism is less challenged than before for the research on well-designed mechanism system performance analysis with prototypes. Now the most challenging part of this type of validation is how to build the control schemes in an economical, accurate and efficient way. The cartesian control [215] is the most common strategy for PM's motion control with a massive number of researches about building the relationship between the joint velocity and the Cartesian velocity through the inverse kinematic model or the robot Jacobian. Unlike serial robot inverse kinematic, the kinematic model of the PM not just build with joint configurations but of the pose of the robot's EE. Building a control model with the inverse kinematic brings a new challenge for PM's motion control that how to estimate or measure the pose of the EE. There are a lot of efforts in creating approaches for EE pose estimation with the forward kinematic model and measurements from joint space. However, researchers also proved that how challenging it is to form an analytic formulation for the PM's forward kinematic model [215]. There is a promising alternation for PM's motion control called metrological redundancy [216] by involving more sensors into mechanisms in order to simplify the kinematic model of the system [217]. There is another efficient way for estimating the pose of EE – vision-based control which could be applied as an alternative control for Cartesian control of PM. Unlike traditional control strategies for PM which are designed from the joint space or moving platform space [54] to estimate the position of the platform. With the vision-based sensing system, the pose of EE could be measured accurately and timely as a feedback signal for the control scheme. With those features, a vision-based control scheme could significantly promote the efficiency of PM's motion control and could remove barriers for mechanisms' research in terms of system validation. There two types of typical methods for such systems: vision as a sensor for motion control [218] and visual servo control [56]. With electronic sensors shows drawbacks as expensive, hard for maintenance, and unsustainable, visual servo control has drawn great attention in the past decades as features of non-contact sensing, flexibility, sustainability and high robustness to application environment uncertainties[57- 60].

According to the different ways of setting up a visual system, visual servo systems could be generally classified into two categories: the dynamic visual system which mounted on and moved with the robotic system and static visual system which keep the robotic system on the scene and monitor its movement. The former setup is well-known as an eye-in-hand system [61] and widely applied in the motion control for serial manipulators, unmanned aerial vehicles (UAV) or cable-driven robots [62 - 64]. Another configuration of a visual system can be that cameras are located on fixed positions in which the movement of the robotic system can be monitored, such as systems are introduced in [65] or [66]. In general, visual servo control can have two basic approaches: IBVS and PBVS [67]. In IBVS, the visual system is considered as a 2D sensor most likely. Such a system can be robust and straightforward. However, the IBVS approach focus on building the relationship between 2D images and the 2D information of the robotic system, the accuracy of the system may not be stable enough if there is massive position displacement occurred in the robotic system. Besides, it is also challenging to convert 3D information into 2D accurately. Thanks to the robustness design of the IBVS scheme, the method could perform very well in such working scenarios that object of interest could be stably located in a certain place and wait for the camera system to match with its orientation. For PBSV, the visual system is considered as a 3D sensor in Cartesian space [67]. If there is enough knowledge of the system such as the system's model, a particular relationship from a fixed coordinate to the robot moving coordinates can be determined by using image information [68]. The visual servo control scheme is able to minimize the difference between the current pose and the desired pose of the robotic system.

This chapter focused on vision-based control modelling and performance analysis for a generalized parallel mechanism with a fixed camera system. As a novel designed method, instead of dig into its application potential, the focus is to verify how efficient the system and if any potential problems need to be solved or if the system needs to be improved in a certain perspective. For the intent of controlling the system with observation, the camera-to-hand setup is an ideal choice for this research. We took a planar mechanism with a parallelogram EE from [52] as an illustrative example to verify the analysis process. The proposed method can be used for spatial mechanisms or other types of parallel mechanisms with a similar modelling process. Besides, the model of the system can be used for advanced control schemes design in future research as well as applying advancing visual technologies for systems observation.

4.2 Design of A Generalized Parallel Robot Vision-based Control Model

4.2.1 The kinematic of the mechanism

The configuration of a planar mechanism with a parallelogram EE [52] is shown in Figure 4-1, and our approach of the control scheme is based on this mechanism which has 3-DOF for the manipulator and 1 DOF for the end-effect. In the configuration, parameters $a, b, c, d, e,$ and f are constants as the length of the linkage of the parallelogram EE, the linkage length of AB, BC, CD and parameters on the based that distances between actuators, respectively.

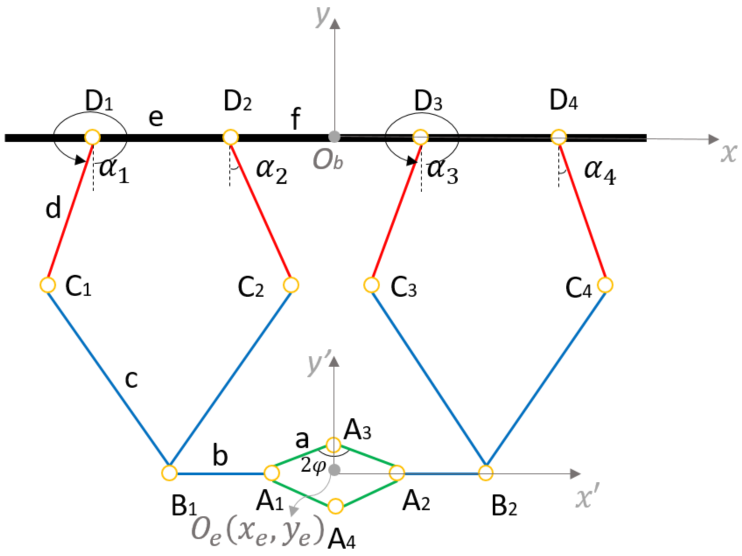


Figure 4-1. The planar general parallel mechanism configuration

In the Cartesian coordinates $O_e x' y'$ of the end effect, the position vector $\boldsymbol{p^e} = [\boldsymbol{x_e}\ \ \boldsymbol{y_e}]$ of the reference point O_e on the moving platform A_i is defined and can be written as,

$$\boldsymbol{A_i^e} = (\boldsymbol{-1})^{\boldsymbol{i}} \boldsymbol{a} * [\mathbf{sin\varphi} \quad \mathbf{0}]^{\mathrm{T}}, (\boldsymbol{i} = \mathbf{1}, \mathbf{2}) \tag{4-1}$$

where a is the distance between each two adjacent A_j $(j = 1,2,3,4)$ and φ is the operational angel of the end effect which would change for different tasks such as pick-and-place. With the change of the φ, the EE of parallelogram would be able to 'open' and 'close' to a certain degree with a different dimension of the object of interest.

The distance $|A_iB_i|$ $(i = 1,2)$ is a constant b which is connection limb between the EE and one side of the five-bar mechanism. So, the coordinates of points B_i respect to the EE frame could be written as,

$$\boldsymbol{B_i^e} = [(-\mathbf{1})^{\boldsymbol{i}}\boldsymbol{b} + \boldsymbol{A_i^e} \quad \mathbf{0}]^{\boldsymbol{T}}, (\boldsymbol{i} = \mathbf{1}, \mathbf{2}) \tag{4-2}$$

Let rewritten B_i^e to a vector $\boldsymbol{v_i}$ a simplified expression

$$\boldsymbol{v_i} = [\boldsymbol{t_i} \quad \mathbf{0}]^{\boldsymbol{T}} \tag{4-3}$$

where $\boldsymbol{t_i} = (-\mathbf{1})^{\boldsymbol{i}}(\mathbf{a} * \mathbf{sin\varphi} + \boldsymbol{b}), (i = 1,2)$ and the derivative of this position would be written as $\dot{\boldsymbol{t_\iota}} = (-\mathbf{1})^{\boldsymbol{i}}\mathbf{cos\varphi}\dot{\boldsymbol{\varphi}} = \boldsymbol{n_i}\dot{\boldsymbol{\varphi}}$.

The rotation of the moving platform $O_ex'y'$ to the reference frame O_bxy on the base could be represented by the rotational matrix R which was written as,

$$\mathbf{R} = \begin{bmatrix} \mathbf{cos\theta} & -\mathbf{sin\theta} \\ \mathbf{sin\theta} & \mathbf{cos\theta} \end{bmatrix} \tag{4-4}$$

where θ is the angle between the fixed frame O_bxy and the moving frame $O_ex'y'$. Then, points B_i^e could be written from the EE frame to the base frame as

$$\boldsymbol{B_i^b} = \boldsymbol{p} + \boldsymbol{Rv_i}, (\boldsymbol{i} = \mathbf{1}, \mathbf{2}) \tag{4-5}$$

where $p = (x,\ y)$ is the coordinates of the vector p^e in the $O_b xy$ frame. Differentiating (3-5) with respect to time would obtain the velocity vector of points B_i in the frame $O_b xy$

$$\dot{\boldsymbol{B}}_i^b = \dot{\boldsymbol{p}} + \dot{\boldsymbol{\theta}}\boldsymbol{QRv}_i + \boldsymbol{R}\dot{\boldsymbol{v}}_i, (\boldsymbol{i} = \mathbf{1}, \mathbf{2}) \tag{4-6}$$

where

$$\mathbf{Q} = \begin{bmatrix} \mathbf{0} & -\mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix} \tag{4-7}$$

Now, let's move to the expression of points C_i. Position vectors of joint C_i in $O_b xy$ frame could be written as the expression with actuators' parameters

$$\mathbf{C}_j^{\mathbf{b}} = [\ \boldsymbol{d}_i + \boldsymbol{dsin\alpha}_j \quad -\boldsymbol{dcos\alpha}_j]^{\boldsymbol{T}} \quad (\boldsymbol{j} = \mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}) \tag{4-8}$$

where $[d_1, d_2, d_3, d_4] = [-(e+f),\ -f\ ,\ f,\ (e+f)]$, respectively.

Considering legs $\mathrm{B_iC_j}$ of the mechanism, length of legs is constant c and which provide a constraint by the distance between points $\mathrm{B_iC_j}$ of the system as

$$\boldsymbol{c}^{\mathbf{2}} = \left(\boldsymbol{B}_i - \boldsymbol{C}_j\right)^{\boldsymbol{T}}\left(\boldsymbol{B}_i - \boldsymbol{C}_j\right) \ \ (\boldsymbol{for\ i} = \mathbf{1}, \boldsymbol{j} = \mathbf{1}, \mathbf{2}\ \boldsymbol{and\ for\ i} = \mathbf{2}, \boldsymbol{j} = \mathbf{3}, \mathbf{4}) \tag{4-9}$$

Differentiating the equation (3-9) respect to time in $O_b xy$ frame would obtain

$$\left(\boldsymbol{B}_i^b - \boldsymbol{C}_j^b\right)^{\boldsymbol{T}} \dot{\boldsymbol{B}}_i^b = \left(\boldsymbol{B}_i^b - \boldsymbol{C}_j^b\right)^{\boldsymbol{T}} \dot{\boldsymbol{C}}_j^b \tag{4-10}$$

where $\dot{\boldsymbol{C}}_j^b = \boldsymbol{d} * \dot{\boldsymbol{\alpha}}_j[\boldsymbol{cos\alpha}_j \quad \boldsymbol{sin\alpha}_j]^{\boldsymbol{T}}$**.** Let write $\left(\boldsymbol{B}_i^b - \boldsymbol{C}_j^b\right)^{\boldsymbol{T}} = \boldsymbol{u}_j$

To rearrange (3-6) and (3-9), the equation would be expressed as follow

$$\boldsymbol{u_j}\left(\dot{\boldsymbol{p}} + \dot{\boldsymbol{\theta}}\boldsymbol{QRv_i} + \boldsymbol{R}\dot{\boldsymbol{v}}_{\boldsymbol{i}}\right) = \boldsymbol{u_j}\left(\dot{\boldsymbol{\alpha}}_{\boldsymbol{j}}\boldsymbol{m_j}\right) \tag{4-11}$$

Where $m_j = d * [cos\alpha_j \quad sin\alpha_j]^T$.

In the equation (3-11), $\dot{p}$ is the Cartesian velocity of the end-effect in frame $O_b xy$, $\dot{\alpha}_J$ (j=1,2,3,4) is the velocity of the joint on the base which locates on point D_j.

According to the Jacobian analysis of parallel manipulators [70], we can write the relation between the velocity of the end-effect and joints as

$$\boldsymbol{J_x}\dot{\boldsymbol{s}} = \boldsymbol{J_q}\dot{\boldsymbol{\alpha}}_{\boldsymbol{J}} \tag{4-12}$$

where $\dot{s} = [\dot{p}^T \quad \dot{\varphi} \quad \dot{\theta}]^{\mathrm{T}}$. $\dot{p}^T$ is the velocity of translation and $\dot{\varphi}$ is the velocity of the manipulator which only occurs when a task is happening such as picking up or placing. And $\dot{\theta}$ is the velocity of rotation, respectively. To rearrange the equation (3-11), the matrix J_x can be express as

$$\boldsymbol{J_x} = \begin{bmatrix} \boldsymbol{u_1} & \boldsymbol{u_1Tn_1} & \boldsymbol{u_1QRv_1} \\ \boldsymbol{u_2} & \boldsymbol{u_2Tn_1} & \boldsymbol{u_2QRv_1} \\ \boldsymbol{u_3} & \boldsymbol{u_3Tn_2} & \boldsymbol{u_3QRv_2} \\ \boldsymbol{u_4} & \boldsymbol{u_4Tn_2} & \boldsymbol{u_4QRv_2} \end{bmatrix} \tag{4-13}$$

where $T = R * [1\ 0]^T$and n_i here is $(-1)^i\cos\varphi$ according to equation (3-3). And

$$\boldsymbol{J_q} = (\boldsymbol{u_j m_j})\boldsymbol{I_{4\times 4}} \tag{4-14}$$

Then the relation of joints velocities and velocity of the manipulator in Cartesian coordinate can be written as

$$\dot{\boldsymbol{\alpha}}_{\boldsymbol{J}} = \boldsymbol{A}^{\boldsymbol{b}} \dot{\boldsymbol{s}} \tag{4-15}$$

where $\boldsymbol{A}^{\boldsymbol{b}} = \boldsymbol{J}_{\boldsymbol{q}}^{-1} \boldsymbol{J}_{\boldsymbol{x}}$.

This is a typical process to create the relationship between the EE and actuators of a parallel mechanism system. The key to building a control scheme with this kinemics model is the pose of the EE. As advantages we have discussed the visual system, a vision-based control scheme would be presented.

4.2.2 Vision-based control scheme

Vision-based control is to control the motion of a robot by vision data which obtain from camera systems. In IBVS, most of the schemes are based on points feature, or few advance schemes are based on lines or image moments [234]. Since IBVS commonly have no requirement of on poses of the manipulator only changes in depth of the feature need to be considered, IBVS is not robust in certain scenarios such as large rotational motion happened in the robotic system. In PBVS, the camera system can be a camera fixed in the workspace so that the robot motion can be observed [69], or a set of cameras mounted on pan-tilt heads for observation [60] which is so-called eye-to-hand. The eye-in-hand setup could also apply the PBVS. According to the system kinematic model (3-15) which depends on the 3D pose of the manipulator, both eye-in-hand and eye-to-hand visual systems with PBVS can provide an adequate scheme for this mechanism. As the proposed system consists of a novel mechanism that has not been widely studied, the focus of this work is to create a convenient control system to verify the system efficiency. The eye-to-hand system for observing the EE of the system seems like an ideal setup for the research purpose. So, we are modelling the proposed scheme. The eye-to-hand system with PBSV could be considered for future applications of the mechanism.

For the mechanism in Figure 4-1, a mobile reference frame $\mathcal{F}^m$ is defined which is attached to the frame of the manipulator and $\mathcal{F}^{\mathrm{e}}$. And $\mathcal{F}^d$which are considered as the fixed frame of the base. According to the PBVS control scheme in Figure 4-2., the pose vector is defined as P and P^* which are located in the current frame $\mathcal{F}^m$ and the desired frame $\mathcal{F}^{m^*}$, respectively. The error e can be defined as $\boldsymbol{e} = (\boldsymbol{e_t}\ \ \boldsymbol{e_\vartheta})$, in which $\boldsymbol{e_t} = \left[\boldsymbol{e_x}, \boldsymbol{e_y}, \boldsymbol{e_\varphi}\right]^{\boldsymbol{T}}$is the translation error and $\boldsymbol{e_\theta} = \boldsymbol{u\vartheta}$ in which u is the axis and ϑ is the angle of the rotation matrix ${}^{m}R_{m^*}$. In the mechanism, φ is contributed to the translation in the z direction of the manipulator.

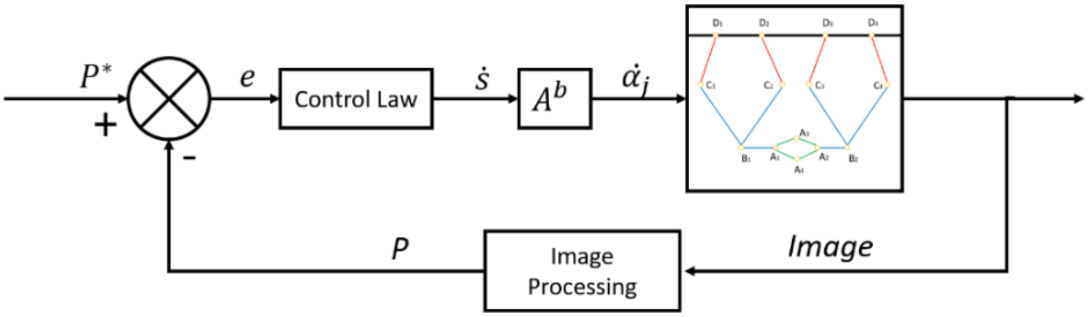


Figure 4-2. Control sachem of a generalized planar parallel robot

The relationship between $\dot{e}$ and the velocity of the platform in its own Cartesian coordinate can be written as:

$$\dot{\boldsymbol{e}} = \boldsymbol{L_s} \dot{\boldsymbol{s}}_{\boldsymbol{m}} \tag{4-16}$$

where an exponential decoupled form is selected as $\dot{\boldsymbol{e}} = -\lambda \boldsymbol{e}$.

According to the definition of the interaction matrix, the matrix related to error e can be written as:

$$\boldsymbol{L_s} = \begin{bmatrix} -\boldsymbol{I_3} & [\boldsymbol{e_t}]_\times \\ \boldsymbol{0} & \boldsymbol{L_{u\vartheta}} \end{bmatrix} \tag{4-17}$$

where I_3 is the identity matrix, $[e_t]_\times$ is the skew symmetric matrix:

$$[\boldsymbol{e_t}]_\times = \begin{bmatrix} \boldsymbol{0} & -\boldsymbol{e_\varphi} & \boldsymbol{e_y} \\ \boldsymbol{e_\varphi} & \boldsymbol{0} & -\boldsymbol{e_x} \\ -\boldsymbol{e_y} & \boldsymbol{e_x} & \boldsymbol{0} \end{bmatrix} \tag{4-18}$$

and $\mathrm{L}_{\mathrm{u}\theta}$ is given by [18]:

$$\mathbf{L_{u\vartheta}} = \boldsymbol{I_3} - \frac{\boldsymbol{\vartheta}}{\mathbf{2}} [\boldsymbol{u}]_\times + \left(\mathbf{1} - \frac{\mathbf{sinc}(\boldsymbol{\vartheta})}{\mathbf{sinc}^{\mathbf{2}}\left(\frac{\boldsymbol{\vartheta}}{\mathbf{2}}\right)} \right) [\boldsymbol{u}]_\times^{\mathbf{2}} \tag{4-19}$$

where $\mathrm{sinc} = \frac{\sin \vartheta}{\vartheta}$.

So the instantaneous velocity of the manipulator can be presented as a function of the pose error:

$$\dot{\boldsymbol{s}}_{\boldsymbol{m}} = -\lambda \widehat{\boldsymbol{L}_{\boldsymbol{s}}^{-1}} \boldsymbol{e} \tag{4-20}$$

where $\widehat{\boldsymbol{L}_{\boldsymbol{s}}^{-1}}$ is the inverse of an estimate of L_s and it is a full rank matrix[18].

The relation between the velocity of the platform in the $\mathcal{F}^m$ and the $\mathcal{F}^d$ can be expressed as:

$$\dot{\boldsymbol{s}} = \boldsymbol{A}_{\boldsymbol{d}} \dot{\boldsymbol{s}}_{\boldsymbol{m}} \tag{4-21}$$

where A_d is the adjoint matrix which can be expressed as:

$$\boldsymbol{A}_{\boldsymbol{d}} = \begin{bmatrix} {}^{d}\boldsymbol{R}_{\boldsymbol{m}} & \left[{}^{d}\boldsymbol{t}_{\boldsymbol{m}}\right]_{\times} {}^{d}\boldsymbol{R}_{\boldsymbol{m}} \\ \boldsymbol{0}_{\boldsymbol{3}} & {}^{d}\boldsymbol{R}_{\boldsymbol{m}} \end{bmatrix} \tag{4-22}$$

where ${}^{d}t_m$ is the vector from O_m to O_d.

Under our control scheme, since we chose an eye-to-hand setup the $\mathcal{F}^m$ is the same frame as the $\mathcal{F}^e$. If the model is applied in an eye-in-hand setup, the $\mathcal{F}^m$ is the same frame as the camera frame.

4.3 Validation of the Models

To verify the proposed method of in previous sections, a simulated generalized parallel mechanism with a parallelogram EE is shown in Figure 4-3. The mechanism is a 3DOF (2-translational motion, 1 rotational motion) + 1DOF (1 motion in the end-effect) parallel robot. In the simulation, all lengths of the robot are designed into a realistic dimension.

Corresponding to Figure 4-1, {a, b, c, d, e, f} = {50, 30, 100, 80, 60, 35} where all lengths are in mm.

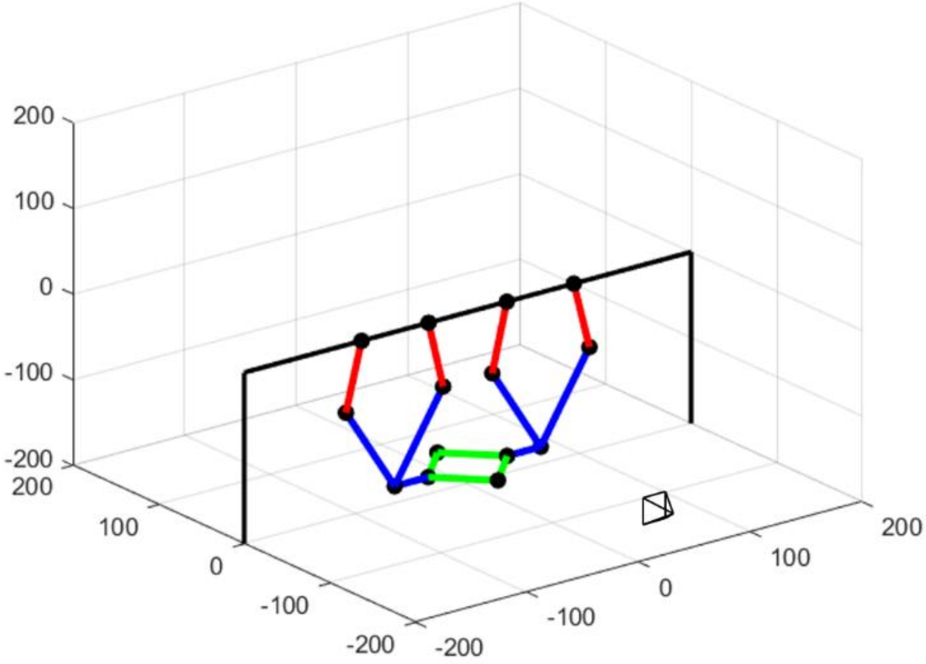
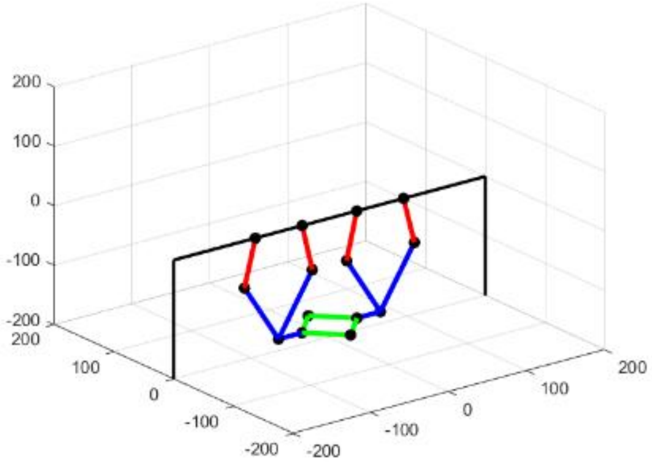


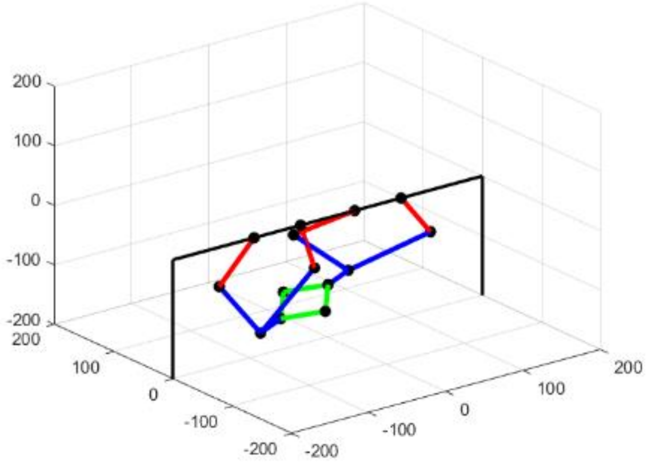
Figure 4-3. The simulation of the parallel platform

In the first simulation, we choose a case that no end-effect motion happened in the system. It is a motion from the initial position $\boldsymbol{s_1} = \left[0, -130, 0, \frac{pi}{4}\right]^T$ to the desired position $\boldsymbol{s_2} = \left[-30, -90, \frac{pi}{6}, \frac{pi}{4}\right]^{\mathrm{T}}$ which are shown in Figure 4-4.

.



(a) The initial pose $\boldsymbol{s_1}$



(b) The desired pose $\boldsymbol{s_2}$

Figure 4-4. Poses of the mechanism

To verify the robustness of the proposed vision-based control, random noise is added to the robot pose in the Cartesian coordinate. In Figure 4-5, the simulation result is shown.

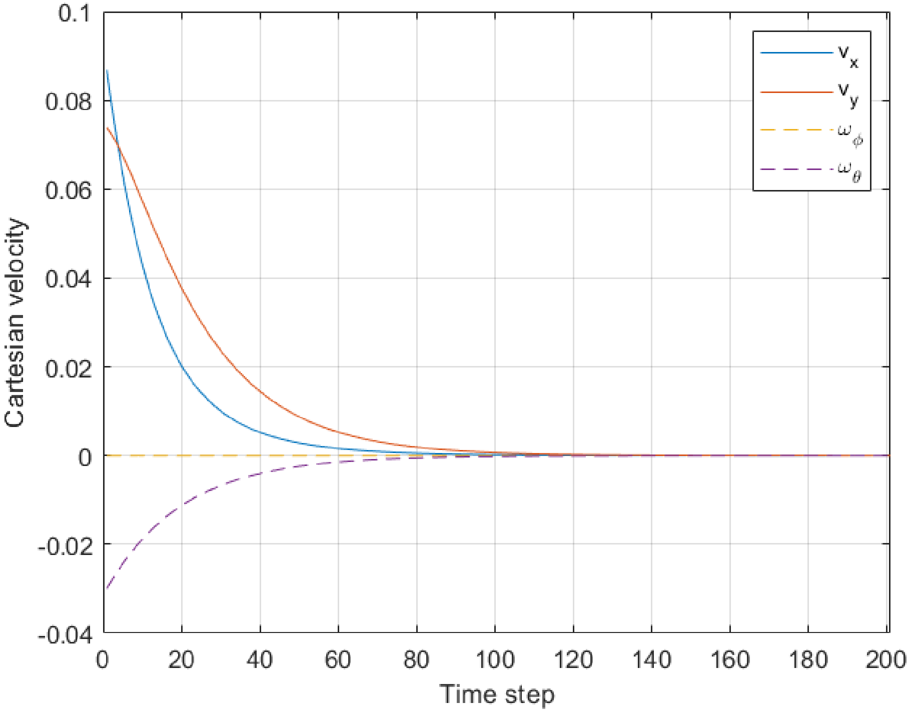


Figure 4-5. Simulation result

4.4 Experiment on Proposed Method

During the validation, the proposed method is achieved with the following pseudocode in Figure 4-6.

```
Input  : mechanism parameters $[a, b, c, d, e, f]$
Output: actuation signals: $[\alpha_1, \alpha_2, \alpha_3, \alpha_4]$
Initial calibration of the system to get the camera model $L_s$ and the depth error of the system $e_z$
should be a constant or zero
mechanism kinematic modelling: solve the inverse kinematic with [a, b, c, d, e, f]
set the mechanism to initial position with know $[\alpha_{o1}, \alpha_{o2}, \alpha_{o3}, \alpha_{o4}]$
image detection to get the EE initial position $p_o = (x_o, y_o)$ and find the movement
  plane of the mechanism
for i = 1;
    image detection to get the EE position $p_i = (x_i, y_i)$ and find the movement
    plane of the mechanism
    calculating the error between the detection position and desired position
    $e_t = [e_x, e_y, e_\varphi]^T$
    calculating the system status $\dot{s}$ by equation (4-16) to (4-22) to obtain the actuation
    signals $[\alpha_{i1}, \alpha_{i2}, \alpha_{i3}, \alpha_{i4}]$;
    if  the $[\alpha_{i1}, \alpha_{i2}, \alpha_{i3}, \alpha_{i4}]$< minimum step
        break;
    end
end
```

Figure 4-6. The vision-based control scheme

The system is built with the same parameters used in the simulation part. All mechanism linkages are made through 3D printing. The printer is the LulzBot TAZ 5 with printing material PLA. The printing error is about ± 2-3 mm. The camera system that is applied here is an RGB-D camera that could use a stereo visual set up as an alternative by Microsoft. The main processor here is an ASUA laptop with an i7-6700 processor. Before the experiment starts, the camera system should be calibrated to get accurate detection later. This calibration process of the camera system is achieved according to the process mentioned in chapter 3. During this experiment, a spatial standard target was used for the camera's calibration, as shown in figure 4-7.

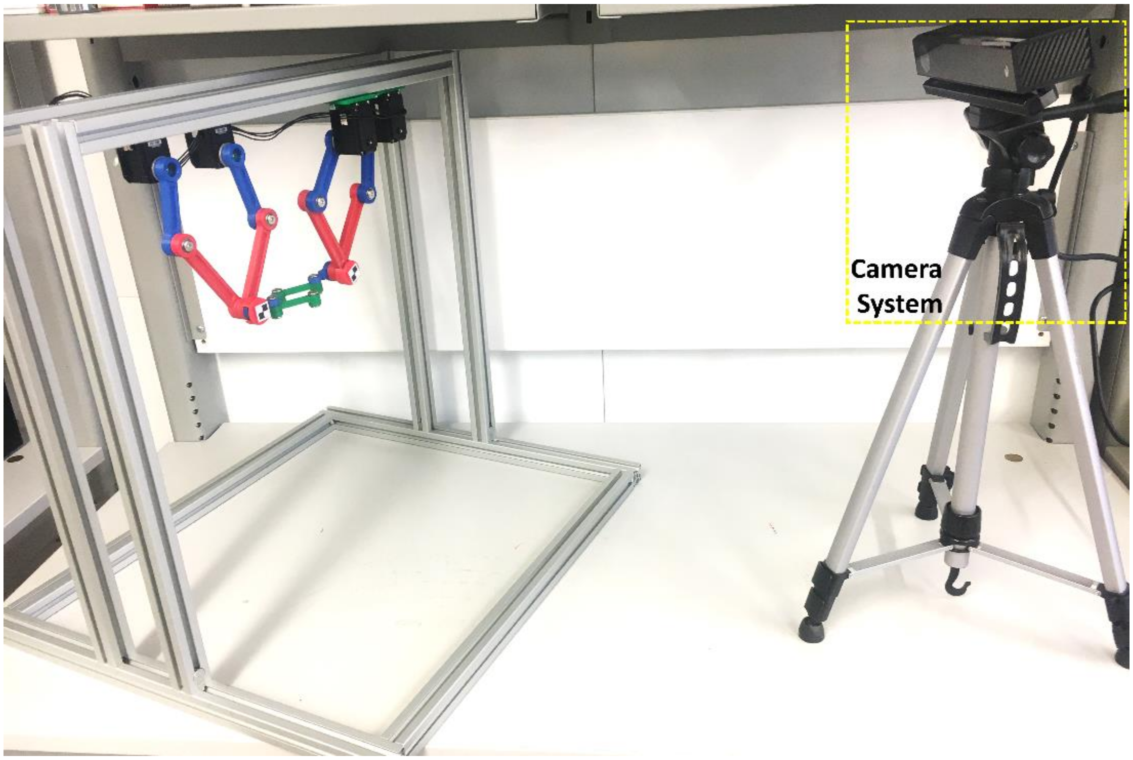


Figure 4-7. Experiment setup

The result of the experiment is shown in Figure 4-8 with the same movement in simulation. The performance of the experiment has the same trend as the simulation result. However, due to system error caused by linkage errors and assembling errors, the system noise is hard to be eliminated completely. The experiment result has validated the efficiency of the proposed model. Based on this result, the proposed approach could be extended for dynamic control of the system. Besides, with the position control result, the system could be applied for applications like picking up and placing.

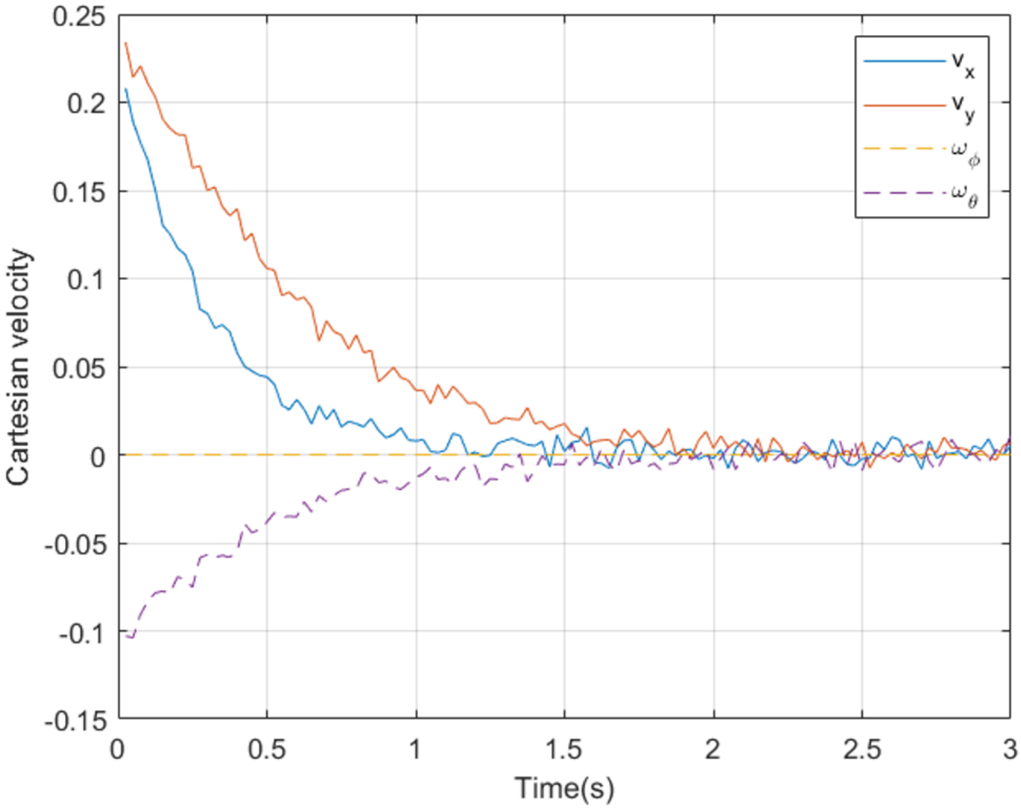
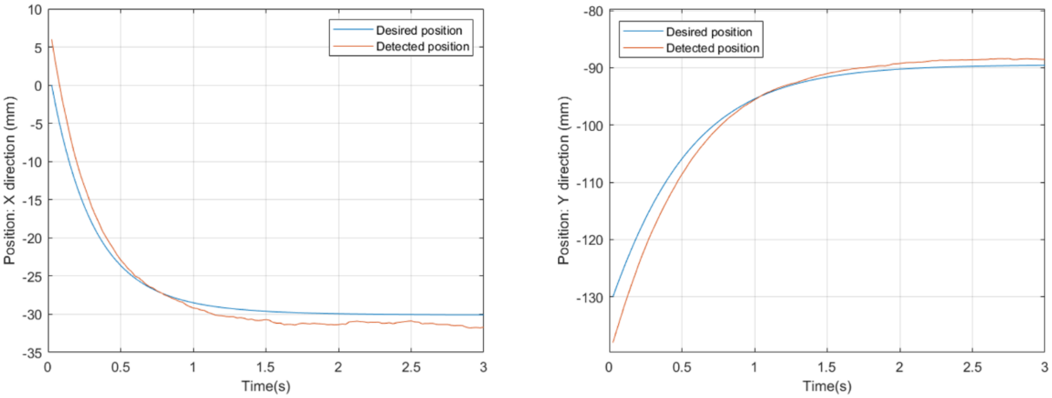


Figure 4-8. Experiment result

In the experiment result, we notice that the system would still be able to reach the desired position in a short period, although the system error is greater than the ideal system.This result has validated the efficiency of the proposed control scheme.

4.5 Discussion

In this chapter, a novel vision-based control scheme is proposed for a generalized parallel mechanism. The methodology building process starts with a traditional kinematic analysis of a parallel mechanism. Then, instead of predicting the position of EE, we create a control logic with the know position of EE. To minimize the error of the EE from the visual feedback, a visual servo control scheme is proposed. In the simulation, the proposed control strategy is verified in an ideal system. The simulation result has shown a great response to the system that would minimize the system error into zero. Additionally, a prototype is built for the validation of the control scheme. The prototype is not a perfect system that would have errors from the 3D printing of linkages, assembling. Additionally, in order to avoid interference between linkages, linkages are intentionally arranged with displacement in the depth direction. In the experiment result, the system error starts with a relatively high level, the control scheme would be able to minimize the system error to a low level within 2 seconds and the system error could be controlled under 3mm. The comparison of the EE position in translational directions between the experiment result and the simulation result is shown below.



(a) Position of the EE in the x direction (b) Position of the EE in the y direction

Figure 4-9. Comparison of the EE position in translational directions between experiment result and the simulation result

In Figure 4-9, the desired position of the EE and the measured result of the EE in both x and y directions are shown below. The comparison of the EE position in the rotational direction between the experiment result and the simulation result is shown in Figure 4-10.

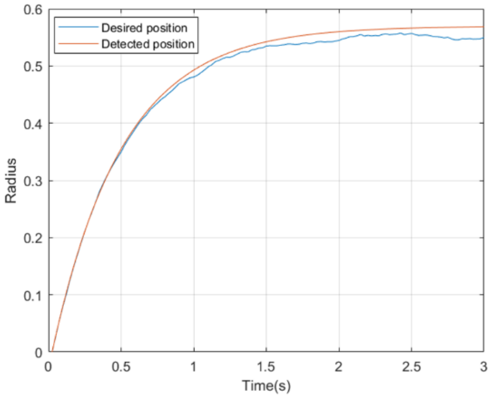


Figure 4-10. Comparison of the EE in the rotational direction between the experiment result and the simulation result

Apparently, the mechanism starts with a position that has a certain distance to the desired initial position. This could be the result of assembling error and linkage manufacturing errors. In order to see the difference between the experiment and the desired movement, the comparison of the EE's moving trajectory is shown in Figure 4-11.

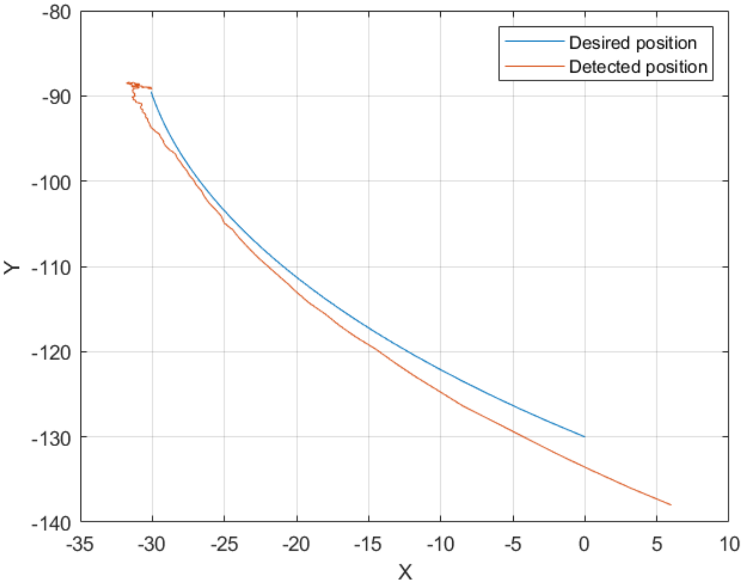


Figure 4-11. Comparison of the EE's moving trajectory

During the experiment, the mechanism starts at the position [6.73, -136.76, 0] and ends at the position [-30.23, -89.12, 31.62°]. The system would be able to move to a position near the desired position in a short time. Then, the system would adjust its position towards the desired position quickly as well. The actuation signal range is the position of the motor that from 0 – 4096 correspond to degrees 0 - 360°. When the step of the actuation signal is smaller than 10 which is less than 1 degree, the motor will not move. During the experiment, after 1.7s, the mechanism has arrived at the final position. Since the detection stopped after 3s, there are still detection results after 1.7s. The detection errors from there are all less than 1mm.

As the simulation and experiment result has implemented the desired performance of the system, this method could be applied for other parallel mechanisms. Although the proposed method is implemented on a planar mechanism that only has movement in a planar area, the system model is built according to a 3D pose measurement. In other words, if the pose of the EE on any mechanism could be kept in the view of the visual system, the control scheme analysis process that we proposed here would be able to adapt to any parallel robot.

Since the method provides a promising sensing approach to measure the pose of the EE which has highly simplified the modelling of the mechanism kinematics, the proposed approach is valuable for an easy, economic and accurate for the mechanism's position control. The next step of this study could extend this work for the system's torque control. As we only created a signal feedback control with the visual information, the extended work about this study could also focus on building control strategy with sensor fusion technology like include additional feedbacks from joint space. Additionally, since the visual system is supervising the motion of the mechanism, there is another approach that could be studied in the future that to observing linkages of the system for motion control. If all the locations of joints in the system could be measured, the visual feedback signal could not just only be applied for motion control but also as the performance feedback for the mechanism kinematic model calibration. In this case, the mechanical system could also have calibration guidance from the visual information. Furthermore, with the accurate measurement of the position of EE as the output of a system, this approach could also be considered to extend into a system recognition study.

Chapter 5 Advanced Visual System for Robot Control

5.1 Introduction Real-time Moving Target Following Systems

When we start to see a delivery robot on the street, it is easy to feel that the robot is really about our daily life now. However, the truth about the delivery robot is that a pilot is monitoring the robot behind the screen to make sure it could be functional [265]. But it is not hard to see the potential of applications of a mobile robot in our daily life. Before the technology could send such mobile robots to achieve full automation, human tracking is one of the most promising applications for a robotic system in the service industry and manufacturing. A human tracking robot could play an important role in daily life such as cargo transportation assisting robots at the workshop or individual luggage carrying robots at any transportation terminal [220]. Therefore, a human following robotic system should able to detect the target continuously and follow the movement of the target while could acknowledge the surrounding environment to find its own path. Unlike surveillance systems with static cameras, a human following robotic system is equipped with a typical eye-in-hand visual system and needs to work in a dynamic environment with timely changing positions of the system itself and the target. In this case, the idea from the visual servo is not suitable here for a mobile robot system since the pose of the objects of interest (OOI) cannot be easily obtained. Before controlling the motion of the robot, the OOI should be located accurately and timely. There are a lot of challenges in terms of locating a specified moving target such as illumination variation, the target's shape variation and target occlusion. [221] Researchers have seen the potential of HRI application with a human following robot system. So there is a large number of efforts have been put on moving target detection and tracking [222]. To achieve the human following for a robotic system, there are approaches presented with multiple sensor sets like using a laser scanner to discriminate human legs for a complex background [223] or applying integrated range sensors for moving targets tracking [224]. Besides, there are several studies present on path planning and localization for moving target following [225]. Among those researches, to apply a visual system for moving target tracking is one of the most popular approaches.

There is a vast number of existing studies devoted to it in terms of achieving the moving target recognition and tracking. For discriminating the OOI from a complex background, texture features are most commonly applied on targets localization with RGB images, and those features could be generally classified as global features and local features. A common challenge of such methods is how to locate and track the OOI in real-time. There are a few remarkable works devoted to improving image processing efficiency by using the local feature in visual tracking filed [226, 227, 228]. Discrete Fourier Transform (DFT) is another typical method that has been widely studied to build correlation filters and to achieve high calculating efficiency [229]. However, such methods are applied in two-dimensional image processing in most cases. In a real-world application scenario, the OOI is in a 3D environment and its tracking should be a three-dimensional problem as well. Besides, it is different from general pedestrians tracking that all humans in the scene are targets for the system and the key is to discriminate humans from other objects in the picture.

We have proposed an efficient and accurate method for dynamic target detection and tracking. And the tracking result is used for a mobile robot motion control in real-time with a mounted RGB-D camera system on the robot. The visual tracker is achieved by applying a modified Kernelized Correlation Filter (KCF) which would detect and track the target timely and accurately. Besides, the camera system could provide surrounding information in three-dimensional for robot motion control and obstacles avoiding. The validation experiment of the proposed system shows that the system could achieve the detection, tracking and following of the determined target smoothly, timely and accurately.

5.2 A Real-time Moving Target Following System with Depth Camera

5.2.1 A Modified Kernelized Correlation Filter Tracker

This part will start with the introduction of the traditional Kernelized Correlation Filter (KCF) which is the foundation of the proposed tracking algorithm. There are a few numbers of trackers designed according to KCF. Such approaches are designed for searching target in a window without drift and tracking without failure when the target is moving with changing speed. As we are looking for a real-time application, to achieve the above approaches, more calculations are needed which highly reduce the computing speed of the algorithm [230] is not applicable in our system requirement. In our proposed approach, the variation velocity of the target is considered to modify the KCF tracker, and the same vector of the movement for the mobile robot is applied for predicting the target's trajectory. In the meantime, we also include the depth camera information in the algorithm for the mobile robot to avoid obstacles in its path.

The focus of KCF tracker [226] is to solve the ridge regression problem which is an approach for minimizing the error between input image patches $t_{w,h}$ $(\mathrm{w} \times \mathrm{h})$ and its regression labels $y_{w,h}$. The model of this problem could be written as,

$$\mathbf{min} \sum_{\boldsymbol{w},\boldsymbol{h}} \left| \langle \boldsymbol{f}(\boldsymbol{t}_{\boldsymbol{w},\boldsymbol{h}}), \boldsymbol{W} \rangle - \boldsymbol{y}_{\boldsymbol{w},\boldsymbol{h}} \right|^{\mathbf{2}} + \boldsymbol{\gamma} \|\boldsymbol{W}\|^{\mathbf{2}} \tag{5-1}$$

where f represent the mapping to the nonlinear space that introduced by the kernel, $t_{w,h}$ is the image patch of the OOI, the γ is a regulation parameter to control overfitting W. So, the target has to be selected manually at the beginning of the tracking. The shifts $t_{w,h}$ $((\mathrm{w}, \mathrm{h}) \in \{0, \dots, W-1\} \times \{0, \dots, H-1\})$ around the target would be considered as the training sample. W is the solution of $W = \sum_{w,h} \alpha_{w,h} f(t_{w,h})$. The variable is presented as calculating discrete Fourier transformation and its inverse of the following expression:

$$\boldsymbol{\alpha} = \boldsymbol{F}^{-\mathbf{1}} \left[\frac{\boldsymbol{F}(\boldsymbol{y})}{\boldsymbol{F}\left(\boldsymbol{K}(\boldsymbol{t}_{\boldsymbol{w},\boldsymbol{h}}, \boldsymbol{t})\right) + \boldsymbol{\gamma}} \right] \tag{5-2}$$

where $K(t_{w,h}, t)$ is the kernel correlation of the target model t. Then the target position within a pitch z could be found by the maximum response of

$$\hat{\boldsymbol{f}}_{\boldsymbol{z}} = \boldsymbol{F}^{-\mathbf{1}} \left(\boldsymbol{F}\left(\boldsymbol{K}(\boldsymbol{t}_{\boldsymbol{w},\boldsymbol{h}}, \boldsymbol{t}) \right) \, \boldsymbol{F}(\boldsymbol{\alpha}) \right) \tag{5-3}$$

With the above training model, real-time tracking could be built. To track a target without constant moving speed and to predict its motion, the OOI variable vector could be defined as $\boldsymbol{x} = \left[p_x, p_y, v_x, v_y\right]^{\mathrm{T}}$, where (p_x, p_y) is the centre location of the target and (v_x, v_y) is the velocity of the target.

As it is mentioned above, the initial position of the target is selected manually and initial velocities vectors are zero. A second-order autoregressive model is applied here for target status estimating [230]:

$$\boldsymbol{x^{k+1}} = \boldsymbol{Ax^k} + \boldsymbol{N(\tau)} \tag{5-4}$$

where A is the state matrix and $N(\tau)$ is the noise matrix. The noise matrix satisfies the Gaussian distributed.

Additionally, to build a more efficient tracking model and a real-time controller for the mobile robot, the OOI is described as a three-dimensional vector for position prediction as $\boldsymbol{x_p} = \left[p_x, p_y, p_z, v_x, v_y, v_z\right]^T$, where $\left(p_y\right)^2 = (d_x)^2 - (p_x)^2 + (p_z)^2$ is the depth position of the target respect to the mobile robot which is obtained by the depth camera and d_x is the depth of position x. The position offset between the x direction and the z direction is $sin\,\theta_r = \frac{d_y}{p_x}$, where θ_r is the rotation vector which would be applied as the signal for mobile robot motion control. The moving velocity of the OOI in the depth direction is the differential of the depth length respect to time as $v_y = d\,\left(p_y\right)$.

In order to obtain a smooth tracking process, A Extended Kalman filter (EKF) is applied for the OOI's position predicting. The state of it could be written as

$$\boldsymbol{\hat{X}_{k+1}} = \boldsymbol{\hat{X}_{k}} \tag{5-5}$$

and the measurement equation is

$$\boldsymbol{P_{k+1}} = \boldsymbol{\phi_{k+1} P_k \phi_{k+1}^T + Q} \tag{5-6}$$

where P_{k+1} is the covariance matrix and ϕ_{k+1} is the observation matrix. Q is the noise of the system. The state of the system would be updated by

$$\boldsymbol{K_{k+1}} = \boldsymbol{P_{k+1} H_{k+1}^T \left(H_{k+1} P_{k+1} H_{k+1}^T + R\right)^{-1}} \tag{5-7}$$

$$\boldsymbol{\hat{X}_{k+1}} = \boldsymbol{\hat{X}_{k+1} + K_{k+1}\left(Z_{k+1} - \hat{h}(X_{k+1})\right)} \tag{5-8}$$

$$\boldsymbol{P_{k+1}} = \boldsymbol{[I - K_{k+1} H_{k+1}] P_{k+1}} \tag{5-9}$$

Where $\hat{h}(X_{k+1})$ is the predicted measurement model of the system and Z_{k+1} is the measurement result of the system.

5.2.2 The Mobile Robot Control Scheme

As the proposed method would be applied on a system that equipped with an RGB-D camera that provides RGB and depth images synchronously, the control scheme of the robot motion could include a process to avoid obstacles in the camera view with the depth information. The information that needs to be determined here is the relative position of obstacles to the camera view.

When there is no obstacle in the scene, the directional vector $[\mathrm{X_r},\ \mathrm{Y_r},\ \theta_\mathrm{r},\ V_r]^\mathrm{T}$ with $[\mathrm{X_r},\ \mathrm{Y_r},\ \theta_\mathrm{r}]$ as moving direction would be the control signal for the robot motion which is obtained from the prediction in 5.2.1. And V_r is the moving velocity of the robot which is obtained from the $v_x\ and\ v_y$ since the robot is not able to move in the vertical direction. In Figure 5-1, the module of the tracking system is shown when $p_x = 0$.

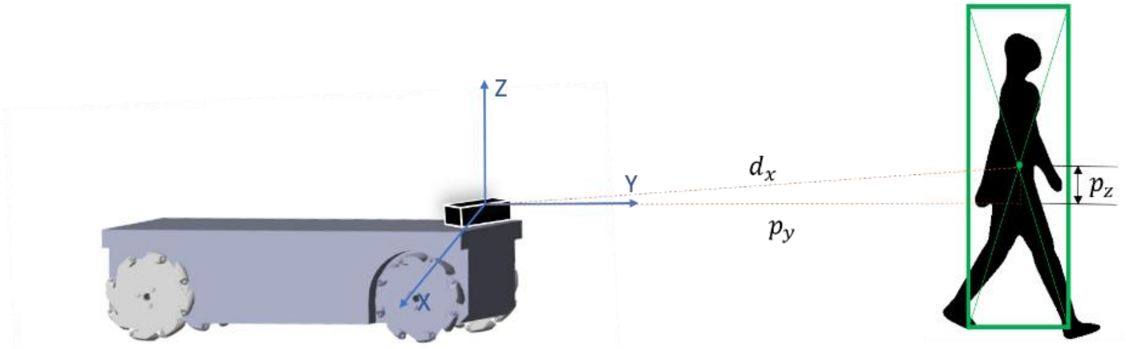


Figure 5-1. The system coordinate during tracking

When obstacles occurred in the path of the robot, the information from the depth camera would provide additional navigation signals. To reduce the processing time and calculation in terms of sensing abnormal objects in the path, only part of the information from the depth camera is considered. As the robot would only operating within a limited height range, the maximum height $\mathrm{H_{r,max}}$ would be the range that considered from the depth camera. In Figure 5-2, a detection result is presented with a depth camera in robot height.

Marked points in the figure are the edges of obstacles in the scene which would be used for accessibility judgement with the understanding of the perspective. The robot is assumed that could pass through an alley with a minimum width $\mathrm{W_{min}}$ or it will stop. To adjust the control vector for the mobile robot's motion

So the robot will have an adjusting vector $[x_r,\ y_r,\ \rho_r,\ v_r]^T$. where x_r is the location in the x direction of the obstacle, y_r will be the same as $\mathrm{Y_r}$, $\sin \mathrm{\rho_r} = f_1(X_r - x_r, Y_r)$ and $v_r = f_2(X_r - x_r, Y_r)$.

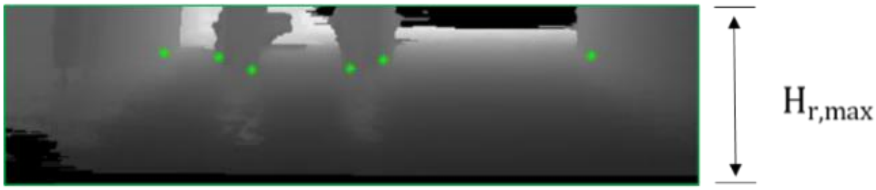


Figure 5-2. The detection result from a depth camera with robot height

The Omnidirectional robot that proposed to use in our approach is able to move in more direction. Unlike a traditional mobile robot, different motions of this robot are decoupled like translation or rotation which could provide improved performance in terms of flexibility and fast response for the system. [231] The mecanum wheeled robot is a typical omnidirectional system which is shown in Figure 5-3,

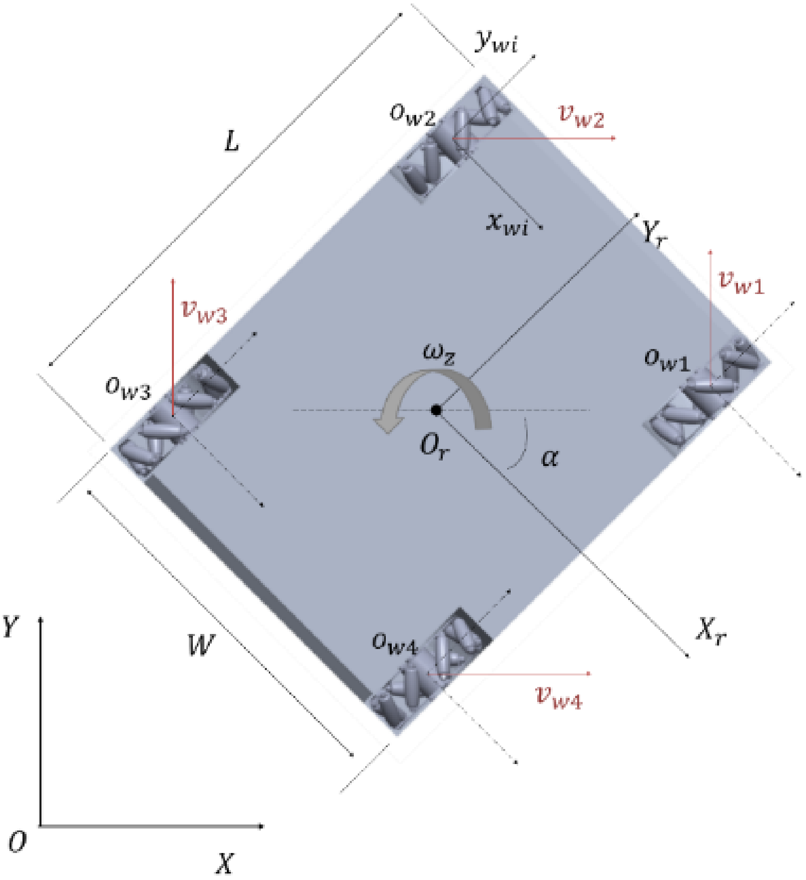


Figure 5-3. The configuration of the robot

where the coordinates OXY is the fixed coordinate and the $\mathrm{O_r}X_rY_r$ is the moving coordinates for the robot. Besides, the coordinates o_{wi} (i=1, 2, 3, 4) are the coordinates frame for each wheel.

The position of each wheel is described as $\mathrm{S_{wi}} = [x_{wi},\ y_{wi},\ \alpha_i]^T$ (i=1, 2, 3, 4). The angular velocity around the hub and the roller are $\dot{\theta}_{xi}$ and $\dot{\theta}_{ri}$, respectively. And the angular velocity about the contact point is $\dot{\theta}_{zi}$. The radius of the wheel and the rollers are R and r, respectively. As the wheel is designed with certain angles, in the proposed approach, the slope angler for the roller and in the robot is φ_i in which $\varphi_1 = \varphi_3$ =45° and $\varphi_2 = \varphi_4 = -45°$. So the velocity of the robot could be written as

$$\dot{\mathbf{S}}_{\mathrm{wi}} = \begin{bmatrix} \dot{\boldsymbol{x}}_{\boldsymbol{w\imath}} \\ \dot{\boldsymbol{y}}_{\boldsymbol{w\imath}} \\ \dot{\boldsymbol{\alpha}}_{\boldsymbol{\imath}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \boldsymbol{r}\,\mathbf{sin}\,\boldsymbol{\varphi_i} & \mathbf{0} \\ \boldsymbol{R} & -\boldsymbol{r}\,\mathbf{cos}\,\boldsymbol{\varphi_i} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{\theta}}_{\boldsymbol{x\imath}} \\ \dot{\boldsymbol{\theta}}_{\boldsymbol{r\imath}} \\ \dot{\boldsymbol{\theta}}_{\boldsymbol{z\imath}} \end{bmatrix} \tag{5-10}$$

To build the relationship between the wheel rotational angel with the robot coordinates T_{wi}, the translation matrix could be formed as a homogenous matrix:

$$\boldsymbol{T}_{\boldsymbol{w2ri}} = \begin{bmatrix} \boldsymbol{cos\alpha_{w2ri}} & \boldsymbol{-sin\alpha_{w2ri}} & \boldsymbol{d_{xw2ri}} \\ \boldsymbol{sin\alpha_{w2ri}} & \boldsymbol{cos\alpha_{w2ri}} & \boldsymbol{d_{yw2ri}} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix} \tag{5-11}$$

where α_{w2ri} (i=1,2,3,4) is the rotation angle of wheel respect to the $0_{\mathrm{r}}X_rY_r$ coordinates, d_{xw2ri} and d_{yw2ri} are the translational displacement.

So the position vector of the robot could be written as $\mathrm{S_r} = [x_r,\ y_r,\ \alpha_r]^T$=$T_{w2ri}S_{wi}$.

Then, the kinematics of the robot is obtained as

$$\begin{bmatrix} \dot{\boldsymbol{x}}_{\boldsymbol{r}} \\ \dot{\boldsymbol{y}}_{\boldsymbol{r}} \\ \dot{\boldsymbol{\alpha}}_{\boldsymbol{r}} \end{bmatrix} = \frac{\boldsymbol{R}}{\mathbf{4}} \begin{bmatrix} \mathbf{-1} & \mathbf{1} & \mathbf{-1} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \frac{\boldsymbol{W+L}}{\mathbf{2}} & -\frac{\boldsymbol{W+L}}{\mathbf{2}} & -\frac{\boldsymbol{W+L}}{\mathbf{2}} & \frac{\boldsymbol{W+L}}{\mathbf{2}} \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{\theta}}_{\mathbf{1}} \\ \dot{\boldsymbol{\theta}}_{\mathbf{2}} \\ \dot{\boldsymbol{\theta}}_{\mathbf{3}} \\ \dot{\boldsymbol{\theta}}_{\mathbf{4}} \end{bmatrix} \tag{5-12}$$

Where $[\dot{\theta}_1, \dot{\theta}_2, \dot{\theta}_3, \dot{\theta}_4]$T is the angular velocity of each wheel.

In the control of the robot, the signal could be sent in as the coordinate$[\dot{x}_r, \dot{y}_r, \dot{\alpha}_r]$.

5.3 Validation and Experiment

During the validation, the proposed method is achieved by the following pseudocode in Figure 5-4.

```
Manual selected target at i-1 -th frame $z^{i-1}$
Input  : i-th frame $z^i$, target state $x^{i-1}$, model template $t^{i-1}$ and $\alpha^{i-1}$
Output: target state x , model template $t^i$ and $\alpha^i$
'searching for target' : the center of the searching window comes from the prediction
  result by equation (5-1) and (5-2);
calculating confidence measurement $f_z$;
the detected target motion vector x and depth motion vector $[p_z, v_z]$ with extended
  Kalman filter to find the final position of the target;
updated target motion vector $x^i = [p_x, p_y, v_x, v_y]$ and $\alpha^i$;
$p_{target} = f(p_x, p_y, p_z)$ ; %target position relative to the camera
$X_r = f(p_x, p_{target})$;
$Y_r = f(p_y, p_{target})$;
$v_x = f(p_x(i), p_x(i+1))$; $v_y = f(p_y(i), p_y(i+1))$;
$V_r = f(p_{target}(i), p_{target}(i+1))$;
$\theta_r = f(p_x, p_{target})$
If  'obstacles' detected  %movement adjustment
    $p_{obs}(x_r) = f(p_o, p_o, p_o)$ ;
    $Y_r = Y_r$;
    $\rho_r = f_1(X_r - x_r, Y_r)$;
    $v_r = f_2(X_r - x_r, Y_r)$;
    $X_r = X_r + x_r$;
    $\theta_r = \theta_r + \rho_r$;
    $V_r = V_r + v_r$;
end;
```

Figure 5-4. The proposed algorithm

5.3.1 The Validation of the Modified KCF Tracker

Traditionally, the efficiency of a tracker would be tasted with available online benchmark videos like trackers [226, 230, 291, 292] are all tested with certain benchmarks. To verify the efficiency and accuracy of the proposed tracker, the proposed approach has been tested on videos of a benchmark [266]. The result could be found in Figure 5-5. During the tracking, the algorithm is able to track the target when the target is occluded and lost for a short time. When the target is partially occluded and fully occluded, the target box is still located on the target path. With a relatively long tracking period with complex interference, the target could still be tracked.

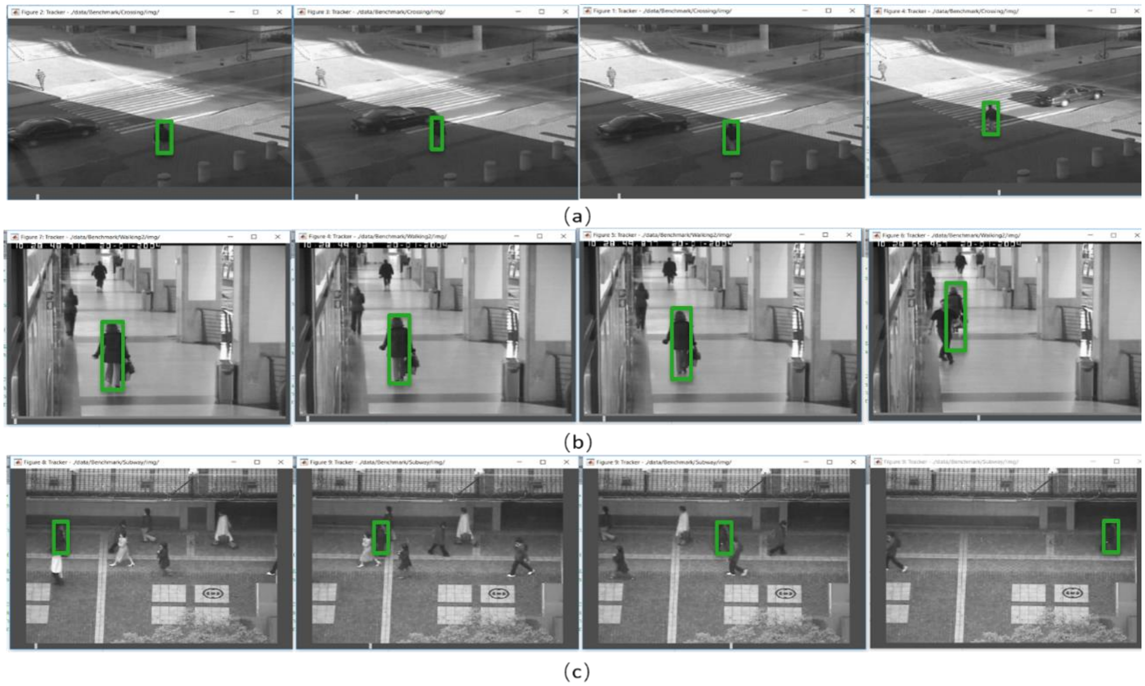


Figure 5-5. Test result on videos of a benchmark [266]

This test is done with videos from static camera systems which make the detection and tracking less complicated. There are a few more tests done on a dynamic benchmark. In Figure 5-6, the test result on a dynamic video is shown which is a benchmark taken by a dynamic camera located on a mobile robot. The proposed method has shown the expected result during detecting and tracking. And the result has been demonstrated that the system would be able to be tracking in a complex environment. Additionally, with the prediction, the target would be able to be located even there is a short period of full-occlusion.

Figure 5-6. Test result in Dataset BAHNHOF [232]

To implement the proposed approach for a real-time application, the mobile robot system used in the experiment is shown in Figure 5-7. The system consists of an omnidirectional mobile robot, a Kinect camera system by Microsoft and an ASUS laptop with an i7-6700 processor as the system main processor. As the system required for a wireless application, a wireless power supply is designed for the camera.

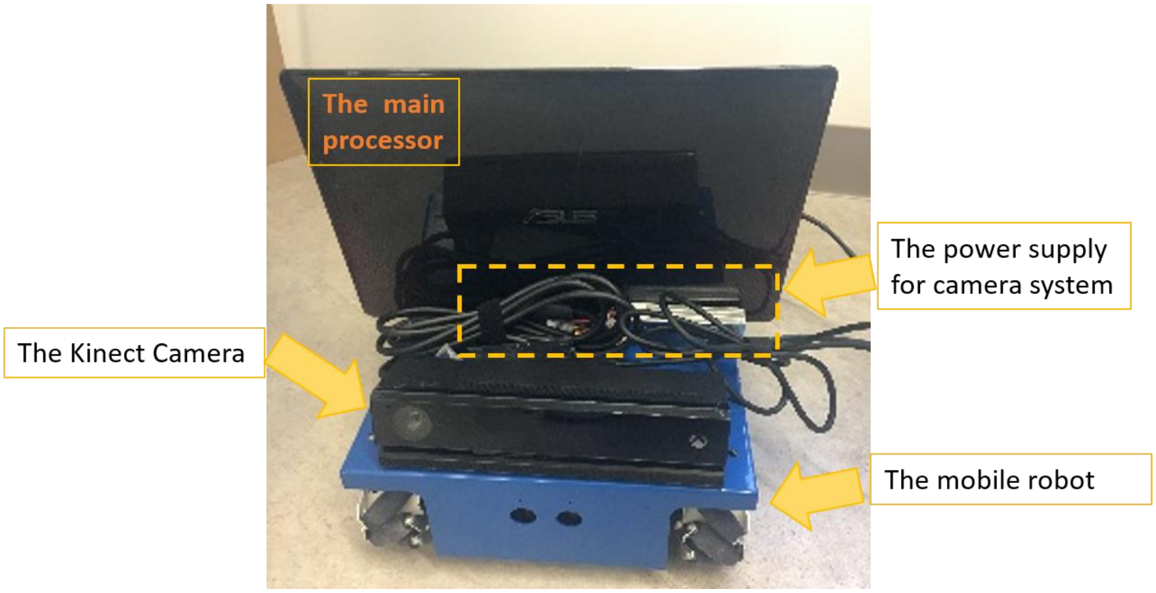


Figure 5-7. The mobile robot system with an RGB-D camera

The system is tested in an experimental environment. In Figure 5-8, a few scenes from a series of experiments are shown. The first row is shown the tracking result in the colour image space, and the second row shows the accessibility judgment result from the proposed method. The maximum moving velocity for the robot is 10m/s. And the accessibility detection range $\mathrm{H}_{\mathrm{r,max}}$ is set to 40 cm. The distance from the target to the camera system is around 2.7 – 3.5 m during the tracking process.

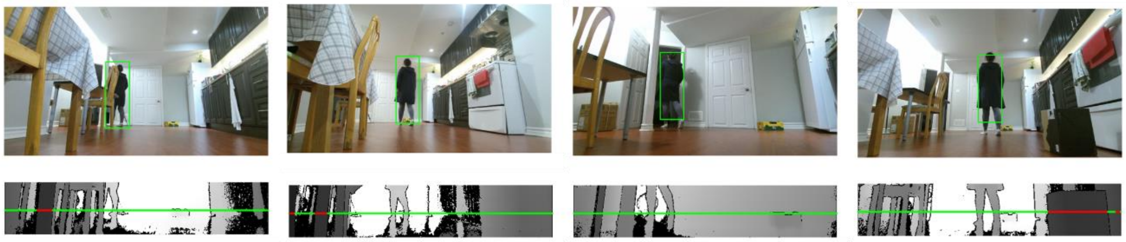


Figure 5-8. Test result in the experimental scene

In the figure of the accessibility result (the second row in Figure 5-8), a line is drawn to indicate if there are any obstacles in the path of the robot within the height range.

5.4 Discussion

In this chapter, we have proposed an effective tracker for dynamic target detection and tracking which is achieved by applying a modified KCF for OOI’s learning and tracking. The tracking performance of the modified tracker has been tested on multiple benchmarks and has been demonstrated its accuracy and efficiency. To verify the real-time tracking feature of the tracker, the proposed method has also been adapted for a mobile robot system which has equipped with an RGB-D camera. And an EKF has been applied for enhancing the tracking performance. Besides, the prototype system would be able to provide information on the robot working environment to guide the robot avoiding obstacles on its path with the three-dimensional information providing by the camera system. The validation experiment of the proposed system shows that the system could achieve the detection, tracking and following of the determined target smoothly, timely and accurately. With the feature of the designed system that sensing information with an RGB-D camera, this novel approach provides a promising application that could be widely applied in a human involved working environment. Such a system could be applied to carrying robots in the workshop as a robot assistant, the following robot for transporting luggage for tourists, or it could be applied in hospitals for patient's transportation.

The tracker we have applied in this chapter is a modification from the KCF tracker with extended application in 3D tracking. In Figure 5-9, we have shown the tracking result comparison between the proposed tracker and the KCF tracker. For the most challenge tracking scenario that targets occlusion, the KCF tracking result shows an obvious drifting. And the modified KCF tracker has a better performance on following the target moving trajectory.

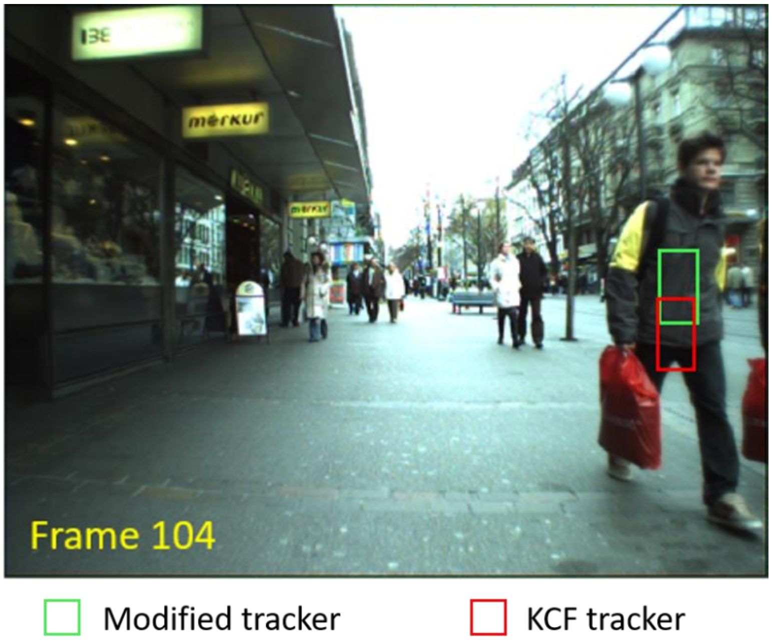


Figure 5-9. Tracking result comparison

Although the experiment result has proved the efficiency of the proposed method, there are still a few improvements that could be made for the system. In terms of obstacles avoiding, we considered a relatively simple strategy to achieve a real-time application. So these obstacles preventing approach could be less efficient in a dynamic working environment that obstacles may not always be in static statues. Besides, the mobile robot moving speed is set with a relatively low moving speed range to ensure the system efficiency which performed very well with a normal walking speed target. If the moving target raising the moving speed, the mobile robot may not be able to avoid obstacles with a high operational speed effectively. There are some effort could be put on the path planning for this system since the depth map could provide information within five-meter. With certain planning strategies, the system would provide more flexibility while tracking a complex dynamic target. With the improvement mentioned above, this human-following robot would be able to work in a general environment that consists of more unpredictable noises or a more critical environment like criminal following and chasing as an AI police.

Chapter 6 Beyond Vision-based Control: The Vision System for Human-Robot Interaction

6.1 Introduction of Hand Gesture Recognition

With the development of HCI, among all kind of interaction approaches, hand gesture recognition based HCI becomes one of the most popular implementations as this is the most convenient method for sending the interacting signal in the natural distance and it is one of the closest approaches to daily operation activities between humans [241]. Comparing with other human natural communication ways like vocal signal or eye contacting signal, hand gestures have the most common and notable features between different individuals which make the recognition of hand gestures become realistic and reliable with current technologies and could be widely used and could be easily adapted into different applications.

There are several researchers devoted to the study of effective hand gesture recognition in the past decades. Those approaches could be generally classified into two categories with the way of sense. The first one is hand gesture recognition with contact sensors such as 'Cyberglove' [242] which is a glove-liked system would detecting gestures with changing signal from electronic components[233], and the more advanced system like body-worn inertial sensors [243] which sensing bio-signal for gesture recognition. Most of those approaches are focusing on converting gestures into electrical signals, then creating a mapping relation between different electrical signal features to their corresponding gestures. With this kind of framework, the detection result is accurate and reliable as well as less sensitive to lighting conditions, and this type of technology has already been applied for HCI and used in military, industry and commercial fields. However, challenges for those wearable devices are system manufacturing and maintenance since such systems normally consist of multiple electronics parts needed to be wired together and systems normally require complex calibrations [245]. Besides, some of those systems may not be very user friendly since the wearing process and the wearing period could be uncomfortable. As an alternative approach, there is a comprehensive gesture sensing system proposed which combined inertial measurement unit (IMU) and Electromyography (EMG) sensors for detecting the motion of hand [256]. But such a system would need a certain number of pre-learned patterns to be created for effective recognition.

The second classification of the hand gesture recognition method is achieved with non-contact sensing technology which mainly achieved with camera systems. Although the former approach for gesture recognition has already been widely studied and already has mature products on the market with high accuracy and reliability, non-contact sensing start to get more attention as drawbacks of contact sensing systems have shown with wide application. At the same time, features such as natural interaction, measuring without physical contact of non-contact sensing seen to be promising for better interaction and better user experience. In certain application scenarios, non-contact sensing shows great advantages comping with contact sensing systems like the object of interest is fragile, concerned for contact or preferred to be contactless. With the booming of machine vision technologies, a vision-based hand gesture recognition system has been widely studied and products with such functions start to take the market in areas as smart surveillance, teleconferencing, HCI, etc. [248]. Besides, such systems with comprehensive sensing and recognition ability have shown great potential in filed like medical and military. For example, the vision-based sensing system in a visual environment for very large-scale biomolecular modelling [249] and the integrated controller for the virtual environment Battlefield [250].

Although hand gestures own relatively notable features, there are a lot of different perspectives in terms of understanding features from gestures. One of the approaches is based on the colour space to determine gestures from the background. Such a method could be which could effectively reach the desired result and could also solve problems like partially occluded [251]. By applying a pattern recognition approach like mean filtering in HSV color space and through morphological image processing, the system could be able to recognize static gestures effectively [252]. Applications for such color space based methods would be troublesome in a cluttered environment due to their sensitivity to color models [246, 247]. The above-mentioned approaches are mainly studied in the image plane which is hard to solve the recognition problems in 3D. There is a lot of effort on 3D recognition from image planes with the understanding of projection perspective and the inspiration from human eyes, and that valuable information provided by 3D data would be able to solve problems like self-occlusion[262, 263]. However, the efficiency of 3D reconstruction is still a challenge for applications that require high accuracy, real-time and high robustness [264, 265]. There is another promising alternative technology for 3D contactless measurement – 3D sensors, especially depth sensors (e.g., Kinect and Intel RealSense 3D camera). Among numerous applications, one of the most interesting applications is using gestures for mobile robot control. With the inspiration from the traffic police officers' hand signals, there are some different approaches proposed for mobile robot motion control with understanding a series of hand gestures. There is an electronic sensor-based system proposed for an omnidirectional wheel-chair motion control [253] with a hand gestures classifier algorithm. The system consists of a wearable IMU sensor and two EMG sensors for obtaining wrist movement and two forearm muscle activities. To encode electrical signals that corresponding to different gestures as well as to train the data received, there are seven different gestures that could be accurately detected for controlling maneuverability of the wheelchair. As the system efficiency has highly relied on sensors and a designed training process, this approach may need to be modified or calibrated for different users. And sensors have to be handled carefully during the application to make sure the system would keep functioning. Besides, time-consuming would also be a challenge for an effective application of such a method. Another approach proposes in [252] presented a gesture-based control scheme for a mobile robot with five different gestures that corresponding to certain control signals. In this method, a library with multiple pre-learned gestures is required, then the later recognition would be processed as a looking-for-table format to find the best contour correspondence. Although a gesture-based state-feedback control strategy is applied for robot motion control to guarantee the real-time operation, the limited number of gestures that could understand by the system would narrow the application of such a system for complex operational requirements. Besides, since the library is built in the color space, as we mentioned above, the system performance could be concerned in a non-experimental environment with complex background information. To take advantage of the depth sensor, a robotic wheelchair motion control scheme which based on hand gesture recognition by a Kinect sensor is proposed [253]. This novel approach applies a Fuzzy controller to establish a corresponding relation from gestures to speed and direction for the wheelchair which could efficiently provide the understanding of a gesture movement trend as a command of the robot’s moving trend. The system is well functional and could accurately work in real-time. However, the system is less robust to vibrations that occurred during the operation which is not avoidable for this operation process that users’ hands must hold in the air. Besides, the system is also not working perfectly during a high-speed operational model [253].

So, the mobile robot motion control with hand gesture recognition needs to overcome challenges like accurate and time-efficient recognition in a complex environment, variable gesture recognition, efficient processing for real-time motion control and less sensitivity to users’ diversity. In this chapter, a comprehensive approach is presented. The experiment of the proposed method has validated features like real-time, accurate and flexible of the system.

6.2 Real-time Two-hand Gesture Recognition for Mobile Robot Control

We developed a gesture recognition approach with a depth camera for motion control of a mecanum wheeled mobile robot manipulator. Control signals of the robot are corresponding to gestures from users' both hands. The system proposed system consists of two parts of processes: the process of real-time gesture recognition and the interaction process between the detected gesture and robot motion controller. This approach proposes a 3D geometry-based method for gesture recognition and the system working flow is shown in Figure 6-1. To start with the system initialization, a few seconds pre-setting for the individual user is required. Once it is finished, the system starts to process the detection and recognition in real-time. As the proposed recognition scheme could understand gestures from both hands, there is a primary set of gestures and a secondary set of gestures for control logic. When an operational gesture is detected, the system will determine if the mobile robot is moving or not. The operational signal for the designed gripper would only send when the robot is not moving.

There are seven different gestures defined for mobile robot moving operation and three gestures defined for the gripper's operation. Once the system started, the gesture signal would be detected and recognized timely, then a corresponding operational signal for the motion control of the robot or the gripper would be sent. The system could be able to move desirably and able to achieve certain tasks with users' orders.

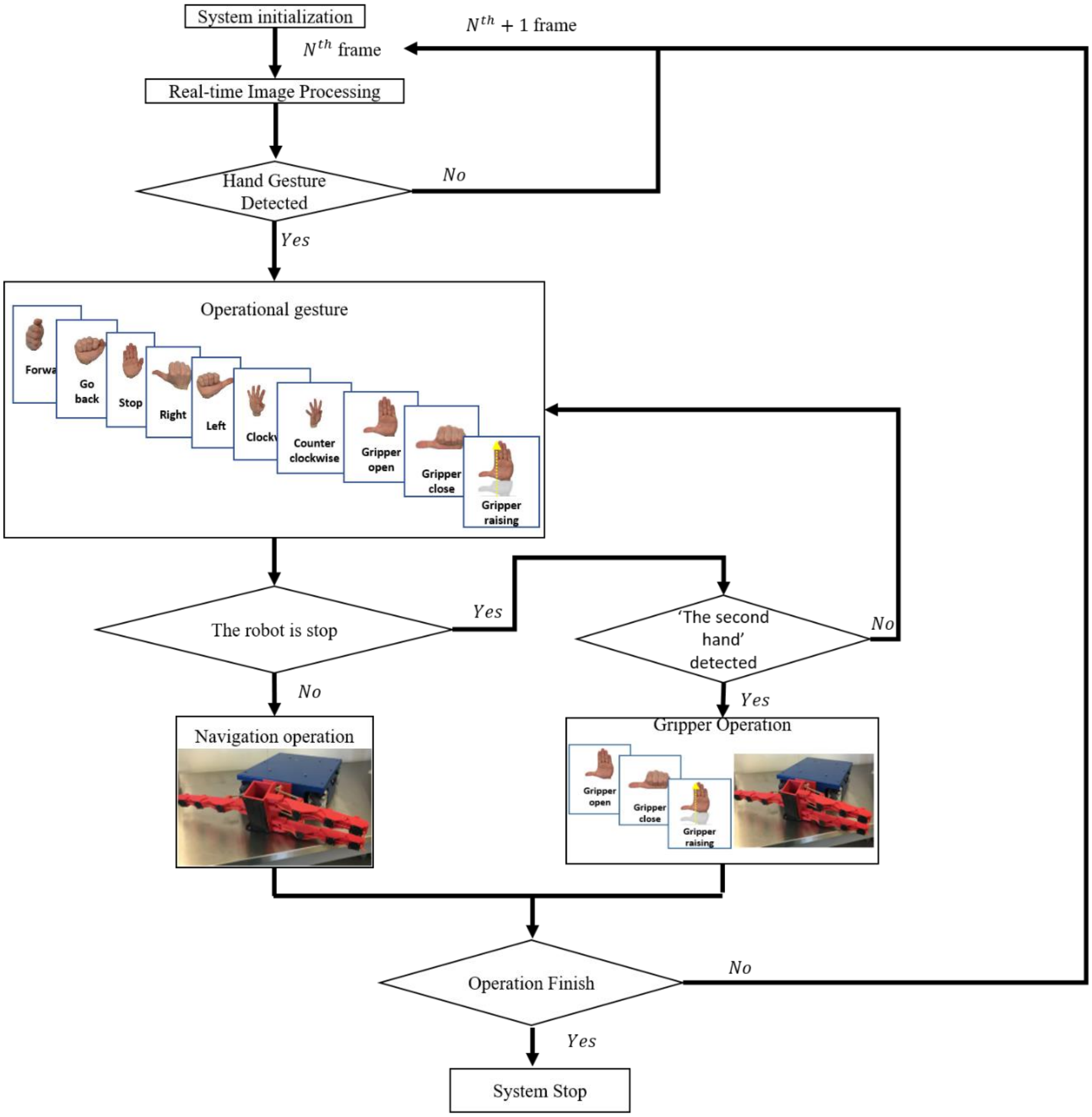


Figure 6-1. The system working flowchart

6.2.1 Two-hand Gesture Recognition

As we understand that one of the biggest challenges for gesture recognition is to distinguish the gesture of interest from a complex background, the proposed method determines a depth range for detection to eliminate the background noise affection for the detection. Besides, gestures' features are defined in 3D space for creating a series of flexible operational gestures. To discriminate 7 different gestures with geometric features, a coordinate system of a hand is defined and shown in Figure 6-2. In the coordinates, there are three length parameters used for gestures definition h, w and d which are the length of the mass centre to the tallest point of the finger in the vertical direction, the mass centre to the farthest point in the thumb horizontal direction and the mass centre to the thumb in-depth direction, respectively.

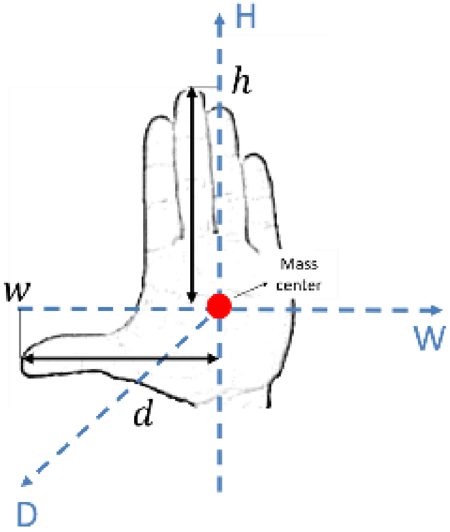


Figure 6-2. Hand gesture coordinate

Before the detection started, the system needs to be initialized for getting initial hand parameters according to the coordinates system. There are two gestures required for this process which are the gesture "Left" and the gesture "Stop". The former is to set the user's hand with thumb open and the other four fingers close in order to get the initial number for

w. The latter gesture is to set the hand with the thumb close and the other four fingers upright to acquire the initial number for h. This process could be finished within 2 seconds. And with this process the proposed method could have a unique parameter set for each individual. Additionally, this quick initialization process could also provide an operational range for users automatically. This process is achieved with a pre-arranged detection range in depth space which would only have detection results. In another word, the user would only see an effective initialization process when they put their hands in the designed range from the camera to the hand.

After a successful initialization of the system, the logic for discriminating different gestures could be described with the following expression (6-1). When the geometric parameters are under certain conditions, the directional signal could be obtained.

$$G(w,h,d) = \begin{cases} w(T_{WN}) \quad \&\& \quad h(T_{HN}) & \ldots Backward \\ w(T_{WX}) \quad \&\& \quad h(T_{HN}) & \ldots Right \\ w(-T_{WX}) \quad \&\& \quad h(T_{HN}) & \ldots Left \\ w(T_{WN}) \quad \&\& \quad h(T_{HX}) & \ldots Stop \end{cases} \tag{6-1}$$

where, $G(w,h,d)$ is the current gesture geometric parameters as $T_W = w, T_H = h,$ and $T_D = d$. The gestures for orientational operation recognition could be divided into two situations. When there is no rotation trend detected, the gesture detected would be discriminated by the logic equation (6-1). With the initial hand parameters, thresholds are set for geometric recognition as – T_{WN} for the thumb closed, - T_{HN} for the four-finger closed status, - T_{WX} for the thumb opened and – T_{HX} for the opened four-finger, respectively. There are five detection scenarios with the above logic expression:

- While both $T_W < T_{WN}$ and $T_H < T_{HN}$ are detected, the state of the detection is "Forward";
- When $T_W \in T_{WN}$ and $T_H \in T_{HN}$ are detected, the state of the detection is "backward";
- With the $T_W \in T_{WX}$ and $T_H \in T_{HN}$ detected, the state of the detection is "right";
- For the detected result as $T_W \in -T_{WX}$ and $T_H \in T_{HN}$, the state of the detection is "left";
- As the $T_W \in T_{WN}$ and $T_H \in T_{HX}$ detected, the state of the detection is "stop";

When there is a rotational trend appearing in the system, the rotation direction of the system can determinate by the changing trend of the T_D which means checking the movement of thumb in the depth direction. The method "convexity defects" [254, 255] is adapted for the rotational detection which is the status when fingers are open and a few contours shown between fingers. Then there are two detection scenarios with this condition:

- When T_{WN} and changing T_D are detected, the state of the detection is clockwise rotation;
- When $-T_{WN}$ changing T_D are detected, the state of the detection is counter-clockwise rotation.

The above 7 gesture that defined for mobile robot navigation could be present as Figure 6-3. Images of gestures in the figure are detected results from the system.

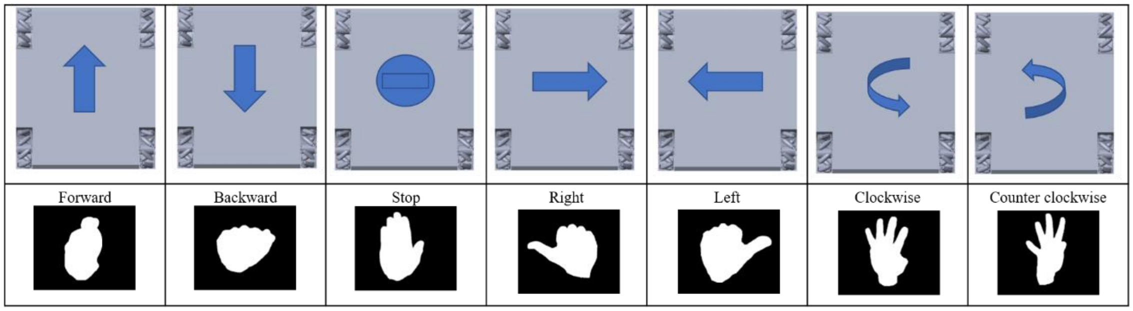


Figure 6-3. The 7 different operational gestures for mobile robot navigation

A similar logic is used for defining gestures for the gripper's operation but less complex than the navigation gestures. There are two static gestures and one dynamic gesture for the gripper's operation which is shown in Figure 6-4. One more condition for this detection is that when there is a 'stop' detected and the second hand detected as well then the recognition would be processed. As the understanding of two hands may locate in a different depth location, second-hand recognition also starts with a pre-set value. The simple gesture used here for initialization is thumb open with four fingers folding to obtain the $-T_{WX}$ and T_{HN}.

- When the $T_W \in -T_{WX}$ and $T_H >> T_{HN}$ detected, the state of the detection is "to open the gripper";
- For the detected result as $T_W \in -T_{WX}$ and $T_H \in T_{HN}$, the state of the detection is "to close the gripper";
- With the "to open the gripper" or the "to close the gripper" detected, if there is an additional movement of the gesture detected in vertical direction of on the center

mass of the area of interest, the state of this operation would be “gripper move” as the signal of lifting or putting down;

Figure 6-4. Hand gesture for the gripper’s operation

During the operation, if the depth of the mass centre is detected with the change in the same position from the initial value, then there is an update of all threshold ranges with the same ratio. But this change of the depth length can not happen outside of the designed detection range.

6.3 The Experiment of Proposed Method

For the proposed system validation, an HRI experiment is designed with certain tasks like operating the mobile robot in a designed map. There are some objects on the map for the robot system to do some picking up and placing operation. Additionally, the system is operated in real-time. As hand gesture recognition is the key to the proposed method, before we applied it for the HRI implementation, the efficiency of the detection must be tested. The testing process is created according to methods in [141, 240, 253]. In the test, all hand gestures are repeated for detection and recognition multiple times in a certain order with nature changing speed which means the user would present all gestures one-by-one in a certain order and repeat the order several times. Besides, during the testing, the user is required to present gestures as normal as she/he can.

Table 6-1 Accuracy of the hand recognition algorithm

	Detected (Number)	Correctly (Number)	detectedAccuracy (%)
Forward	217	206	94.93
Backward	194	185	95.36
Right	226	213	94.25
Left	207	196	94.69
Stop	191	183	95.81
Clockwise rotation	213	197	92.49
CCW rotation	192	179	93.23
Gripper close	230	221	96.09
Gripper open	217	209	96.31
Gripper lifting	208	201	96.63

To set up the HRI implementation with this hand gesture recognition method, a laptop with windows 10 system, Inter I7-6700 CPU is arranged as the main processor for gesture detection and recognition. The mobile robot with mecanum wheels is equipped with a designed 2-DOF gripper for operation. And the robot has a micro-controller inside as the second processor for sending control signals to actuators on wheels and on the grippers.

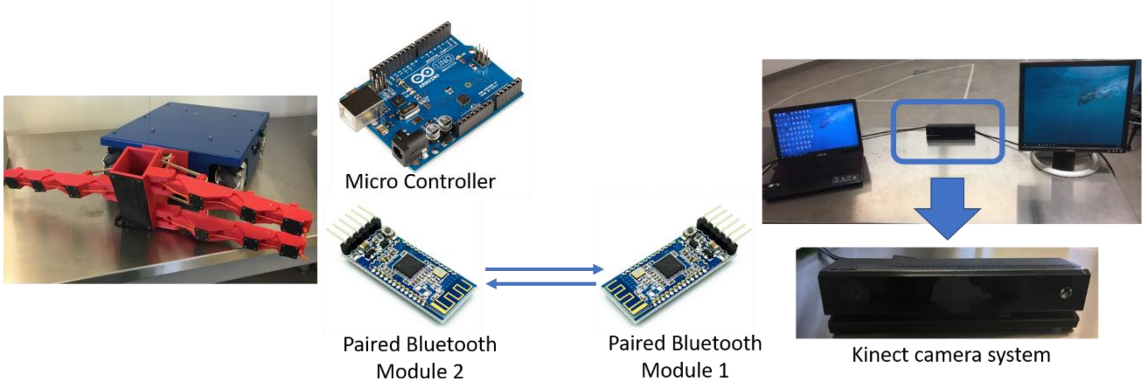


Figure 6-5. The system configuration

The communication between the main processor (an ASUS laptop with i7-6700 processor) and the second processor (an Arduino Mega) is achieved with a pair of Bluetooth chips. The system configuration could be found in Figure 6-5, where the Kinect camera by Microsoft is applied in the system as the visual sensor for gesture recognition. The operational depth range is set as 50 cm – 70 cm from the camera to the hand. This is the key to the proposed method for eliminating background noises for the detection. The mentioned range is for this specific experimental setup. For other working scenarios, this range could be adjusted to meters away from the camera. For the Kinect sensor that we applied here, depth range within 3 meters would be recommended and the 20cm operational range would also recommend ensuring recognition accuracy.

Figure 6-6. The experimental field

As the implementation is an HRI process that relies on human experience or judgment about the working environment, a labelled experimental field is arranged for verifying if the mobile robot could follow the navigation signal correctly. The labelled map is shown in Figure 6-6 with straight paths as well as curved paths. The latter would need the robot to be more flexible and the system must have a high level of real-time that could respond to the change. During the real-time experiment, there are several obstacles and objects placed on the map for the robot to remove and pick up.

During the validation, the proposed method is achieved with the following pseudocode in Figure 6-7 :

```
Input  : Detected hand gestures, $G(w,h,d)$
Output: Operational signals, s
Initial setup for obtaining values: $T_{WX}$, $T_{HX}$, $T_{WN}$ and $T_{HN}$
Detected values $G(w,h,d) = T_W, T_H, T_D$
If  'fingers' = 'open' && $T_{WN}$ && 'changing $T_D$'
    s = 'clockwise rotation';
  elseif  'fingers' = 'open' && $-T_{WN}$ && 'changing $T_D$'
        s = 'counterclockwise rotation';
end
if  $T_W < T_{WN}$ && $T_H < T_{HN}$
    s = 'forward';
  elseif  $T_W \in T_{WN}$ && $T_H \in T_{HN}$
      s = 'backward';
  elseif  $T_W \in T_{WX}$ && $T_H \in T_{HN}$
      s = 'right';
  elseif  $T_W \in T_{WX}$ && $T_H \in T_{HN}$
      s = 'left';
  elseif $T_W \in T_{WN}$ and $T_H \in T_{HX}$
      s = 'stop';
      if  $T_W \in -T_{WX}$ && $T_H >> T_{HN}$   %'second-hand' detected
          s = 'to open the gripper';
          if  'vertical direction movement' detected
              s = 'the opened gripper move up/down';
          end
        elseif  $T_W \in -T_{WX}$ && $T_H \in T_{HN}$
            s = 'to close the gripper';
          if  'vertical direction movement' detected
              s = 'the closed gripper move up/down';
          end
      end
end
```

Figure 6-7. The proposed hand recognition algorithm

In this experiment, for the recognition initial value setup, the ‘stop’ and the ‘left’ gesture are detected from the user’s right hand and the ‘gripper close’ gesture is detected from the user’s left hand. To ensure later detection has a reliable result, this process is designed to be finished manually. So the user has to make sure those three gesture is detected clearly during the initialization. In Figure 6-8, a series of pictures are shown which are taken during the real-time experiment. In the figure, gestures are detected as forward, backward, right and counter-clockwise rotation in the order of left to right. In the meantime, the robot would be able to follow the gesture order correctly and timely.

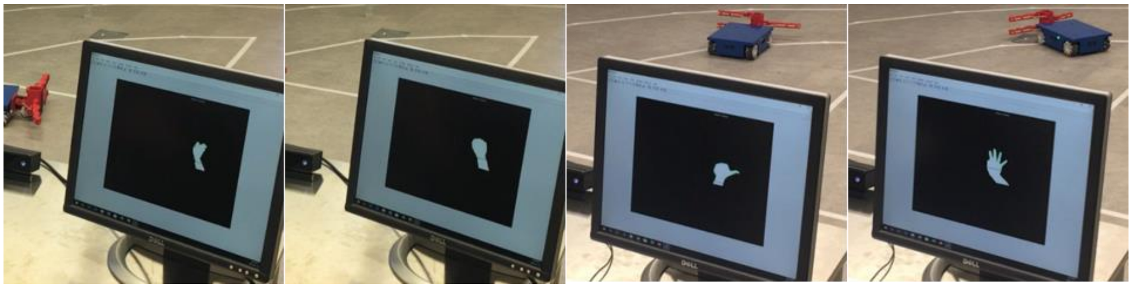


Figure 6-8. The result of mobile robot navigation with the hand gesture

To implement the ability of the proposed system as working for certain tasks, a few objects are arranged on the map. The user should operate the robot for moving through the path and clearing objects located on its way with the two-hand operation.

In Figure 6-9, a set of pictures are shown which are taken during a path following while finishing certain tasks HRI operation. In Figure 6-9 (a) and 6-9 (c), the robot is responding to the navigation signal “forward” and “counter-clockwise rotation”, respectively. In Figure 6-9 (b), the robot is executing the “counter-clockwise rotation” order while holding the same lifting operation on the gripper. This is the process that the robot tries to move the object out of its path. Similarly, Figure 6-9 (d) is shown the robot stopping for picking up the final object on the path.

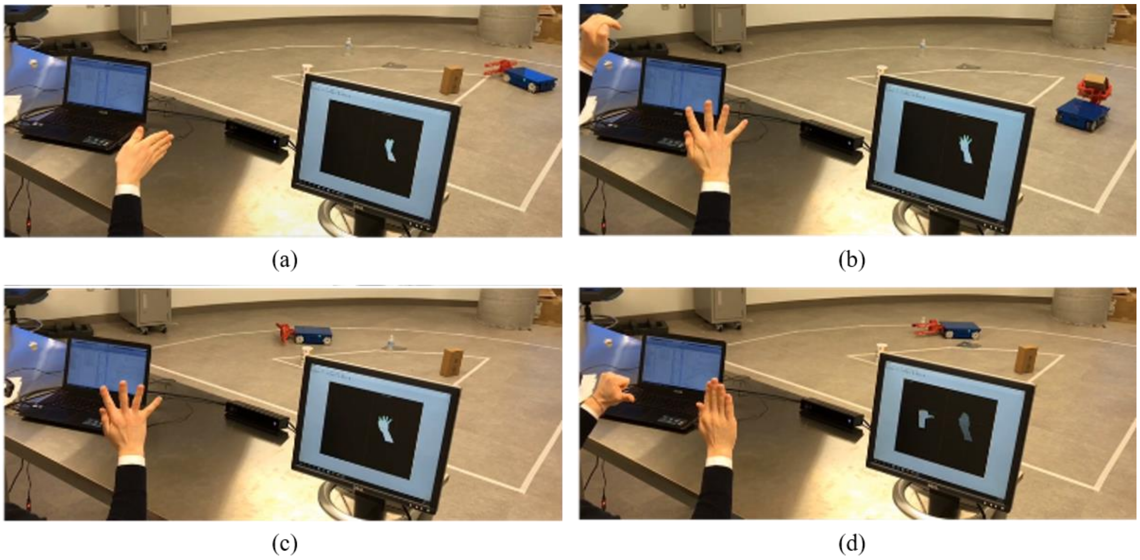

(a) (b)
(c) (d)

Figure 6-9. The result of mobile robot navigation and operation with the hand gesture

During this experiment, the HRI process performed very well. The robot could follow the signal from the hand gesture accurately and timely.

6.4 Discussion

As the key to this research is on the accuracy of the hand recognition algorithm, to further evaluate this work, we take two existing research for comparison. These two methods are one of the most successful methods under their approach category. The first one is an online detection method which is also using the Kinect camera as the sensor [244]. This method could recognize five different gestures with 93.9% accuracy and 4s initial setup time. Although it provides robustness and real-time recognition method, the capability of detecting more gestures would improve its flexibility. The second method is a learning-based recognition method that could recognize 7 different gestures with a 97.71% accuracy [141]. This approach provides a competitive result in terms of accuracy and the capability of detecting different hand gestures. However, data training is a long process that can not be done within seconds. When the system has been applied with more and more users, the accuracy of the detection may decrease since the training library only contains gestures from 20 users. Although the library could always be updated, each updating would require another training process. A comparison result is present in table 6-2 among the proposed method in this chapter and the above-mentioned methods.

Table 6-2 Comparison of the proposed method and existing methods

	Proposed method	Ren's method [244]	Li's method [141]
Accuracy	94.73%	93.9%	97.71%
Initial setup time	1-2s	4s	20 people and 7 gestures for training (a long training period)
Recognition goal	7 gestures	5 gestures	7 gestures

From the numeric data of the table, the proposed method could be a competitive approach with existing methods. And in the experiment, the proposed method has an excellent performance in terms of accuracy and real-time. Additionally, the designed gestures in the proposed method are easy to be understood by any user with the common directional sense, which could be one of the greatest advantages of it. This user-friendly design would allow those who are new to the system to learn and operate the system with a brief period. For users who have been practiced with different interaction systems, this design would have less chance to make them feel confused about the different operational gestures.

Although the proposed method seems reliable and effective for a real-time HRI application, there are still some improvement could be done with this approach. During the initialization, a bad process would lead to a high failure rate in the later detection. In order to have high accuracy, the user should try to keep each initialization gesture as stable as possible with less dynamic noises such as vibration or quick changing in the depth direction. Although the system has certain tolerance on initialization errors that small rotation angle of each gesture could be adjusted by the threshold ranges and the slow change in depth direction could also be accepted by the algorithm, the process for setting initial system parameters still need some attention from users. So if the initialization system could be improved by becoming more robustness with users' unconscious behaviours which could affect the quality of system parameters, this method would become more user-friendly and more efficient.

Chapter 7 Conclusion and Future Work

7.1 Conclusions

This dissertation has present studies of vision-based control schemes for different mechatronics devices. In the first part, a novel vision-based control approach is proposed for a well designed parallel mechanism. A prototype is built for the validation of the control scheme. As the simulation and experiment result has implemented the desired performance of the system, this method could be applied for other parallel mechanisms. Since the method provides a promising sensing approach to measure the pose of the EE which has highly simplified the modelling of the mechanism kinematics, the approach could be further studied for the system dynamic control. Additionally, since the visual system is supervising the motion of the mechanism, there is another approach that could be studied in the future that to observing linkages of the system for motion control. If all the locations of joints in the system could be measured, the visual feedback signal could not just only be applied for motion control but also as the performance feedback for the mechanism kinematic model calibration.

With the discussion made in the research background, we believe it is valuable to study a human involved robot motion control scheme. The second part of this dissertation has presented a human-following robot application approach. An advanced target tracking method is proposed for a mobile robot motion control. With the feature of the designed system that sensing information with an RGB-D camera, this novel approach provides a promising application that could be widely applied in a human involved working environment. Although the experiment result has proved the efficiency of the proposed method, there are still a few improvements that could be made for the system. In terms of obstacles avoiding, we considered a relatively simple strategy to achieve a real-time application. So the proposed obstacles preventing approach could be less efficient in a dynamic working environment that obstacles may not always be in static statues.

Another important aspect that has been studied in this dissertation is in a popular research area – HRI. An advanced motion control approach is proposed for a mobile robot manipulator. The proposed method has an excellent performance in terms of accuracy and real-time with a series of designed hand gestures that could be easily understood by any user with their common sense. Although the proposed system seems reliable and efficient, there are still some improvement could be done with this approach. For obtaining the high accuracy of the system, the user should be careful with the initialization gesture setting up process, although the system has tolerance on small errors So if the initialization system could be improved by becoming more robustness with users unconscious behaviours which could affect the quality of system parameters, this method would become more user-friendly and more efficient.

With our wide range of studies on vision-based control schemes, it is not hard to list the advantages of such visual systems, and it is also not hard to tell how promising it is for such application with a visual system. However, there is still a lot of work that must be done to promote these technologies in solving general engineering problems. Besides, there are some limitations of visual systems that cannot be ignored, like its sensitivity to lighting condition and computing intensity which is always a concern for image processing technology. For a reliable, real-time, and efficient application in areas like advanced manufacturing, real-life interaction, combining different types of sensors together for designing a comprehensive system could be one of the promising solutions.

7.2 Contribution of the Dissertation

7.2.1 Vision-based Control: An Approach for Generalized Parallel Robots

The vision-based control scheme for a well designed parallel mechanism has provided a simplified kinematic model for the mechanism that helps to create the feedback control. The methodology of the proposed method is based on 3D visual information that could be easily adapted to other types of parallel mechanisms.

And the system prototype is built and demonstrated the efficiency and accuracy with the proposed model. The system would respond to the change and stabilize in 2s and the system error would be minimized in 3mm. The potential application of the proposed method could be for system model calibration as the system performance would be directly measured, and this research could be extended into a study for system recognition for systems like a complicated soft parallel mechanism.

7.2.2 Advanced Visual System for a Mobile Robot Motion Control

The advanced target tracking method for a mobile robot’s motion control has proposed a real-time and accurate tracker that could be applied in a relatively challenging environment. Additionally, the 3D tracking strategy brings more flexibility to this method.

In the implementation and experiment, the system has shown great performance in tracking the selected target and control the movement of the mobile robot in real-time. The potential application for such a system could work as an assistant robot for carrying the weight and following the user in a warehouse or work as a service robot.

7.2.3 Beyond Vision-based Control: The Vision System for HRI

The two-hand gesture recognition method for a mobile manipulator has provided an efficient mobile manipulator interaction strategy. The proposed method has designed 7+3 of user-friendly gestures that could be detected timely for real-time applications. Besides, the 3D detection strategy makes the system highly robust to environmental noise.

In the system implementation and validation, the proposed method has presented high accuracy with a 94% correct detection rate. The application of such an interaction method has presented its flexibility and real-time during communication with the mobile manipulator. The potential application for such a system could carry chemical or other types of hazard products in the manufacturing process or interact with other types of robotic systems.

7.3 Future Work

As the development of the industry, the future manufactory processes will be less labour intensive but become more flexible and reliable. Let's take a manufacturing process as an example to understand the current status of manufacturing. The traditional mechanical manufactory system such as vibration feeders [267] are longer suitable for industrial requirements as flexibility and labour economy. Randomized bin-picking was raised attention since the 1980s. This problem is related to an industrial application of automatic picking an object from a randomly stacked group of items by a robotic system. In the recent decade, With the development of vision systems, a visual assisted randomized bin-picking system became a popular research topic. And part of those research results already been used for industrial applications. Machines involve more and more manufacturing process and our daily life is unstoppable. There are more works could be done in the future of this study.

In terms of using a vision system for machine tools motion control, there are a few future work that could be done with the study as a foundation.

How the machine vision system could promote the development of the technology in terms of increasing the productivity of the manufacturing processes and how to combine the visual technology into machine design so the next generation of the machine would become not just efficient but also more flexible, more robustness to the working environment. There are a few aspects in details that we would like to devote to:

- To further apply a visual system for parallel mechanisms motion control, the vision system as a position sensor could also be applied for linkage observation. If the mechanism could be observed not just from the EE but also from all linkages, the system would have less uncertainty during the control.

- With the mechanism system been controlled with visual signal feedback, this approach could be further studied for the system's dynamic control. As force or momentum is not a visible vector for a system, the future dynamic control strategy could be conducted with a sensor fusion method.

- As the visual system is applied for system observation, another possible future research could be the combination of MV and machine learning for system recognition. With a certain amount of learning process, the advanced approach of system recognition could provide a system model for the future application.

- Since the visual system is working as a surveillance system to observe the system performance, it could also be applied for observing the working environment for machine tools path planning or obstacles avoiding.

How to create a safe and friendly co-working environment between humans and robots will become a huge challenge for researchers as well. The machine vision system could bring intelligence into machines so the machine would work with humans for specific manufacturing processes. There are a few future directions that we would like to contribute to in the future.

- As the application potential of the human-following robot technology, to continue working on the current system and to minimize its current limitation would be interesting work. And we believe that this technology would be widely applied in the soon future.

- If the visual system could understand the manufacturing process and predict environment hazards, the safety of the manufacturing process will get into another stage. To study on machine learning for creating a reliable environment library is one of the directions we would like to work on. One of the challenges of this study would be how to extract features of these working scenarios of interests as they are formed with dynamic information.

- The real-time application would always be one of the most significant challenges for applying those well-designed visual technologies into the application. We would like to work on optimization for an advanced visual system for providing less computing intensity approaches for robot motion control.

So how to make new technology accessible to common users such as natural interface processes, common communication between humans and machines would also be another future area of this study. The HRI system from this work already has the ability for a non-experimental application. But there is more to be done since we believe that novel technologies have responsivity to build a smooth path for adapting new manufacturing processes with traditional workers.

- The presented work in this dissertation focuses on interaction with a relativity simple robot system by hand gestures. Further study could be working on an interaction scheme with an advanced robot system like a robot arm by not just gestures but also other body languages.

- As we all know that if you like to train yourself for programming in a certain language, you will able to take certified training online. But if you like to learn to operate a machine, there is no such way yet. Assisting by VR/AR technology and HRI approaches, this could become a future.

- To teach the robot with human experience would be another promising area. In the working environment or manufacturing process which highly relies on human experience, if the machine would be trained through the HRI process, this would make a huge step for machines walking towards full automation.

Bibliography

[1]. P. Corke, “Robotics, vision and control: Fundamental algorithms in MATLAB”, Springer, 2011

[2]. R. C. Gonzalez and R. E. woods, “Digital image processing, Third Edition” Pearson International Edition, 2008

[3]. D. Forsyth and J. Ponce, “Computer vision: A modern approach, 2nd edition,” Pearson, 2011

[4]. R. Jain, R. Kasturi, and B. G. Schunck, “Machine vision,” McGraw-Hill, 1995

[5]. M. Koreichi, S. Babaci, and F. Chaumette, et. al., “Visual servo control of a parallel manipulator for assembly tasks,” in Proceedings of the 6th International Symposium on Intelligent Robotic Systems, pp. 109–116, 1998

[6]. P. Kallio, Q. Zhou and H. N. Koivo, “Three-dimensional position control of a parallel micromanipulator using visual servoing,” in Proceedings of the SPIE Microrobotics and Microassembly II, 4194, pp. 103–11, 2000

[7]. M. Husty, “An algorithm for solving the direct kinematics of general Gough-Stewart platforms, Mechanism and Machine Theory, Vol. 31(4), pp. 365-380, 1996

[8]. W. N. Klarquist and A. C. Bovik. “A foveated vergent active stereo vision system for dynamite three-dimensional scene recovery”. IEEE Transactions on Robotics and Automation, Vol. 14(5): 755-770, 1998

[9]. P. M. Sharkey, D. W. Murray, and P. F. MeLauehlan, et al., “Hardware development of the yorick series of active vision systems”, Microproeessors and Mierosystems, 21: 363-375, 1998

[10]. O. Sutherland, H. Truong, and S. Rougeaux, et al., “Advancing active vision systems by improved design and control,” In: Proceedings of International Symposium on Experimental Robotics (ISER2000), 2000, Honolulu, USA

[11]. R. Szeliski, “Computer vision: Algorithms and applications,” Springer, 2010

[12]. D. Zhang, “Parallel robotic machine tools,” Springer, 2010

[13]. R. Clavel, “Delta, a fast robot with parallel geometry”, in Proceedings 18th International Symposium on Industrial Robots, pp. 91-100, 1988, Lausanne, Switzerland

[14]. D. Zhang and Z. Gao, “Optimal kinematic calibration of parallel manipulators with pseudoerror theory and cooperative coevolutionary network,” IEEE Transactions on Industrial Electronics, Vol. 59(8), pp. 3221-3231, 2012

[15]. F. Pierrot and O. Company, “H4: a new family of 4-dof parallel robots”, in AIM'99: IEEE/ASME International Conference on Advanced Intelligent Mechatronics, pp. 508-513, 1999, Atlanta, USA

[16]. K. Sehlus, W. Fellenz, and A. Koschanetal, et al., “A modular 10-DOF vision system for high-resolution active stereo,” The University of Auckland, 1997

[17]. R. Hartley and A. Zisserman, “Multiple view geometry in computer vision,” Cambridge University Press. 2004

[18]. S. Hutchinson, G. Hager, and P. Corke, “A tutorial on visual servo control,” IEEE Transactions on Robotic Automation, vol. 12, pp. 651–670, 1996

[19]. E. Malis, “Improving vision-based control using efficient second-order minimization techniques,” IEEE International Conference on Robotics and Automation (ICRA), pp. 1843–1848, 2004, New Orleans, USA

[20]. L. Weiss, A. Sanderson, and C. Neuman, “Dynamic sensor-based control of robots with visual feedback,” IEEE Journal on Robotics and Automation, vol. 3(5), pp. 404–417, 1987

[21]. J. T. Feddema and O. R. Mitchell, "Vision-guided servoing with feature-based trajectory generation (for robots)," IEEE Transactions on Robotics and Automation, vol. 5(5), pp. 691-700, 1989

[22]. F. Chaumette, “Potential problems of stability and convergence in image-based and position-based visual servoing,” in the Confluence of Vision and Control, vol. 237, Lecture Notes in Control and Information Sciences, D. Kriegman, G. Hager, and S. Morse, Eds. New York: SpringerVerlag, 1998, pp. 66–78

[23]. B. Thuilot, P. Martinet, L. Cordesses, and J. Gallice, “Position based visual servoing: Keeping the object in the field of vision,” in Proceedings 2002 IEEE International Conference on Robotics and Automation, pp. 1624–1629, 2002, Washington, DC, USA

[24]. E. Malis, “Visual servoing invariant to changes in camera intrinsic parameters,” IEEE Transactions on Robotics and Automation, vol. 20 (1), pp. 72–81, 2004

[25]. E. Malis, F. Chaumette, and S. Boudet, “2-1/2D visual servoing,” IEEE Transactions on Robotics and Automation, vol. 15(2), pp. 238–250, 1999

[26]. D. Lowe, “Three-dimensional object recognition from single two dimensional images,” Artificial Intelligence, vol. 31(3), pp. 355–395, 1987

[27]. P. Corke and S. Hutchinson, “A new partitioned approach to imagebased visual servo control,” IEEE Transactions on Robotics and Automation, vol. 17(4), pp. 507–515, 2001

[28]. D. Dementhon and L. Davis, “Model-based object pose in 25 lines of code,” International Journal of Computer Vision, vol. 15, pp. 123–141, 1995

[29]. W. Wilson, C. Hulls, and G. Bell, “Relative end-effector control using cartesian position based visual servoing,” IEEE Transactions on Robotics and Automation, vol. 12(5), pp. 684–696, 1996

[30]. T. Hoeniger, "Dynamically shared control in human-robot teams through physical interactions," in Proceedings. 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems. Innovations in Theory, Practice and Applications (Cat. No.98CH36190), 1998, Victoria, Canada

[31]. V. V. Unhelkar, J. Perez, J. C. Boerkoel, J. Bix, S. Bartscher and J. A. Shah, "Towards control and sensing for an autonomous mobile robotic assistant navigating assembly lines," 2014 IEEE International Conference on Robotics and Automation (ICRA), pp. 4161-4167, 2014, Hong Kong

[32]. V. V. Unhelkar, C. Pérez-D'Arpino, L. Stirling and J. A. Shah, "Human-robot co-navigation using anticipatory indicators of human walking motion," 2015 IEEE International Conference on Robotics and Automation (ICRA), pp. 6183-6190, 2015 Seattle, USA

[33]. B. Vanderborght, A. Albu-Schäeffer, and A. Bicchi, E. Burdet, et. al., “Variable impedance actuators: A review,” Robotics and Autonomous Systems, 61(12), pp. 1601–1614, 2013

[34]. C. Vogel, C. Walter, and N. Elkmann, “A projection-based sensor system for safe physical human-robot collaboration,” 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 5359-5364, 2013, Tokyo, Japan

[35]. J. Wolfe, B. Marthi, and S. Russell. Combined task and motion planning for mobile manipulation. In Proceedings of International Conference on Automated Planning and Scheduling (ICAPS), pp. 254-257, 2010, Toronto, Canada

[36]. A. M. Zanchettin, N. M. Ceriani, and P. Rocco, et al., "Safety in human-robot collaborative manufacturing environments: Metrics and control," IEEE Transactions on Automation Science and Engineering, vol. 13(2), pp. 882-893, April 2016

[37]. C. Pérez-D'Arpino and J. A. Shah, "Fast target prediction of human reaching motion for cooperative human-robot manipulation tasks using time series classification," 2015 IEEE International Conference on Robotics and Automation (ICRA), pp. 6175-6182, 2015, Seattle, USA

[38]. J. Mainprice and D. Berenson, "Human-robot collaborative manipulation planning using early prediction of human motion," 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 299-306, 2013, Tokyo, Japan

[39]. J. H. Hong, Y. S. Song, and S. B. Cho, “A hierarchical bayesian network for mixed-initiative human-robot interaction,” Proceedings of the 2005 IEEE International Conference on Robotics and Automation, pp. 3808-3813, 2005, Barcelona, Spain

[40]. T. Kanda, T. Hirano, and D. Eaton, et al., “Interactive robots as social partners and peer tutors for children: A field trial,” Human-Computer Interaction, vol. 19(1–2), pp. 61–84, 2004

[41]. T. Kanda, R. Sato, and N. Saiwaki, et al., “Friendly social robot that understands human’s friendly relationships,” 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 2215-2222, 2004, Sendai, Japan

[42]. W. Burgard, A. B. Cremers, and D. Fox, et. al, “The interactive museum tourguide robot,” Proceedings of the fifteenth national/tenth conference on Artificial intelligence/Innovative applications of artificial intelligence, pp.11–18, 1998, Monona Terrace, USA

[43]. H. G. Okuno, K. Nakadai, and H. Kitano, “Realizing audio-visually triggered ELIZA-like non-verbal behaviors,” PRICAI2002, LNAI 2417, Springer-Verlag, pp. 552-562, 2002

[44]. T. Kanda, H. Ishiguro, and M. Imai, et al., “Body Movement Analysis of Human-Robot Interaction,” International Joint Conference on Artificial Intelligence (IJCAI 2003), pp. 177-182, 2003, Acapulco, Mexico

[45]. H. Ishiguro, T. Ono, and M. Imai, et al., “Development of an interactive humanoid robot "Robovie" -An interdisciplinary approach, R. A. Jarvis and A. Zelinsky (eds.),” Robotics Research, Springer, pp. 179-191, 2003

[46]. T. Kanda, H. Ishiguro, and M. Imai, et al., “A constructive approach for developing interactive humanoid robots,” IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1265-1270, vol.2, 2002, Lausanne, Switzerland

[47]. T. Kanda, T. Hirano, and D. Eaton, et al., “A practical experiment with interactive humanoid robots in a human society,” in Proceeding IEEE International Conference of Humanoid Robots, 2003, Munich, Germany

[48]. T. Kanda and H. Ishiguro, “Reading human relationships from their interaction with an interactive humanoid robot,” International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems (IEA/AIE 2004: Innovations in Applied Artificial Intelligence), pp. 402-412, 2004, Ottawa, Canada

[49]. M. A. Goodrich, and A. C. Schultz, “Human-Robot Interaction: A Survey,” Foundations and Trends in Human-Computer Interaction, pp. 203-275. 2007

[50]. F. Bourbonnais, P. Bigras, and I. A. Bonev, “Minimum-time trajectory planning and control of a pick-and-place five-bar parallel robot,” IEEE/ASME Transactions on Mechatronics, vol. 20(2), pp. 740-749, 2015

[51]. C. Gosselin, T. Laliberte and A. Veillette, “Singularity-free kinematically redundant planar parallel mechanisms with unlimited rotational Capability,” IEEE Transactions on Robotics, vol. 21(2), pp. 457- 467, 2015

[52]. X. Jin, Y. Fang and D. Zhang, “Design of a class of generalized parallel mechanisms with large rotational angles and integrated end-effectors,” Mechanism and Machine Theory, vol. 134, pp. 117-134, 2019

[53]. X. Jin, Y. Fang, and H. Qu, et al., “A class of novel 4-DOF and 5-DOF generalized parallel mechanisms with high performance,” Mechanism and Machine Theory, vol. 120, pp. 57-72, 2018

[54]. V. Stojanovic and N. Nedic, “A Nature Inspired Parameter Tuning Approach to Cascade Control for Hydraulically Driven Parallel Robot Platform,” Journal of Optimization Theory and Applications, vol. 168, pp. 332-347, 2016

[55]. G. Meunier, B. Boulet and M. Nahon. “Control of an overactuated cable-driven parallel mechanism for a radio telescope application,” IEEE Transaction on Control Systems Technology, vol. 17(5), pp. 1043-1054, 2009

[56]. T. Dallej and P. Martinet, “Image-based visual servoing of a Gough-Stewart parallel manipulator using leg observations,” International Journal of Robotics Research, vol. 26(7), pp. 677-687, 2007

[57]. L. Weiss, A. Sanderson and C. Neuman, “Dynamic sensor-based control of robots with visual feedback,” IEEE Transaction on Robotics, vol. 3, pp. 404-417,1987

[58]. B. Zi, J. Lin and S. Qian, “Localization, obstacle avoidance planning and control of a cooperative cable parallel robot for multiple mobile cranes,” Robotics and Computer-Integrated Manufacturing, vol. 34, pp. 105-123, 2015

[59]. M. Gupta, S. Kumar, and L. Beher, et al., “A novel vision-based tracking algorithm for a human-following mobile robot,” IEEE Transaction on Systems, Man, Cybernetics: Systems, vol. 47(7), pp. 1415-1427, 2017

[60]. T. Dallej, M. Gouttefard, N. andreff and et al., “Vision-based modeling and control of large-dimension cable-driven parallel robots,” 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1581-1586, 2012, Vilamoura, Portugal

[61]. G. Flandin, F. Chaumette, E. Marchand, “Eye-in-hand/eye-to-hand cooperation for visual servoing,” Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings, vol. 3, pp. 2741-2746, 2000, San Francisco, CA, USA

[62]. T. Dallej, M. Gouttefarde, and N. Andreff, et al., “Towards vision-based control of cable-driven parallel robots,” 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 2855-2860, 2011, San Francisco, USA

[63]. P. K. Allen, A. Timcenko, B. Yoshimi and P. Michelman, “Automated tracking and grasping of a moving object with a robotic hand-eye system,” IEEE Transactions on Robotics and Automation, vol, 9(2), 1993

[64]. N. Kummer, A. Beresowskaja, H. Firouzi and H. Najjaran, “Autonomous UAV controlled collision landing via eye-in-hand visual servoing,” Canadian Aeronautics and Space Journal, Vol. 62(13), pp. 122, 2016

[65]. T. Dallej, N. Andreff, Y. Mezouar and P. Martinet, “3D pose visual servoing relieves parallel robot control from joint sensing,” 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 4291-4296, 2006, Beijing, China

[66]. F. Paccot, P. Lemoine, N. Andreff and et al., “A vision-based computed torque control for parallel kinematic machines,” 2008 IEEE International Conference on Robotics and Automation, pp. 1556-1561, 2008, Pasadena, USA

[67]. F. Chaumette and S. Hutchinson, “Visual servo control part I: Basic approaches,” IEEE Transactions on Robotics and Automation Magazine, vol. 13(4), pp. 82-90, 2006

[68]. Z. Zake, F. Chaumette, N. Pedemonte and S. Caro, “Vision-based control and stability analysis of a cable-driven parallel robot,” IEEE Robotics and Automation Letters., vol. 4(2), pp. 1029-1036, 2019

[69]. N. Andreff and P. Martinet, “Unifying kinematic modeling, identification, and control of a Gough-Stewart parallel robot into a vision-based framework,” IEEE Transactions on Robotics, vol. 22(6), pp. 1077-1086, 2006

[70]. L. W. Tsai, “Robot Analysis: The mechanics of serial and parallel manipulators,” John Wiley & Sons, New York, NY, 1999

[71]. F. Chaumette and S. Hutchinson, “Visual servo control part II: Advanced approaches,” IEEE Transactions on Robotics and Automation Magazine, vol. 14(1), pp. 109-118, 2007

[72]. Y. Fukuoka, K. Tanaka, A.Ishida and H. Minamitani: “Characteristics of visually feedback in postural control during standing,” IEEE Transactions on Rehabilitation Engineering, Vol. 7(4), pp. 427-434, 1999

[73]. S. Ushida, J. Terashita and H. Kimura, “Switching structural biomechanical model of multisensory integration during human quiet standing,” in Proceedings of 43rd International Conference on Decision and Control (IEEE CDC2004), pp. 959– 965, 2004, Atlantis, Bahamas

[74]. J. Seara and G. Schmidt: “Intelligent gaze control for vision-guided humanoid walking: Methodological aspects,” Robotics and Autonomous Systems, Vol. 48, pp. 231-248, 2004

[75]. S. Ushia, K. Yoshimi, T. Okatani and et. al., "The Importance of gaze control mechanism on vision-based motion vontrol of a biped robot," in Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, 2006, Beijing, China

[76]. N. Afzulpurkar, S. Rougeaux, L. Berthouze, and Y. Kuniyoshi, "A gaze based approach for visual servoing," in Proceedings of International Conference on Computational Engineering in System Application, pp. 570-575, 1996, Lille, France

[77]. J. Demiris, S. Rougeaux, and G. M. Hayes, et al., "Deferred imitation of human head movements by an active stereo vision head," in Proceedings of 6th IEEE International Workshop on Robot and Human Communication, pp. 88-93, 1997, Roman, Italy

[78]. Y. Nakabo, N. Fujikawa, and T. Mukai, et al., "High speed and bio-mimetic control of a stereo head system," in Proceedings of SICE Annual Conference, Vol. 3, pp. 2371-2376, 2004, Sapporo, Japan

[79]. Y. Kuniyoshi, N. Kita, S. Rougeaux, and T. Suehiro, "Active stereo vision system with foveated wide angle lenses," in Proceedings of Second Asian Conference on Computer Vision (ACCV95), pp. 359-363, 1995, Singapore

[80]. L. Berthouze, S. Rougeaux, F. Chavand, and Y. Kuniyoshi, "A learning stereo-head control system," in Proceedings of World Automation Congress, International

Symposium on Robotics and Manufacturing, In Robotics and Manufacturing: Recent trends in Research and Applications, Vol. 6, pp. 43- 47, 1996

[81]. M. Ishikawa, T. Komuro, Y. Nakabo, and A. Namiki, “The 1ms-vision system and its application examples, workshop: Innovative SensoryMotor fusion opens a new robotic world (Organizers: Masatoshi Ishikawa, Makoto Kaneko),” IEEE International Conference on Robotics and Automation (ICRA), 2002, Washington, D. C., USA

[82]. S. Takizawa, S. Ushida, T. Okatani and K. Deguchi, “2DOF Motion stabilization of biped robot by gaze control strategy,” in Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 3809-3814, 2005, Edmonton, Canada

[83]. J. Gutman, M. Fukuchi and M. Fujita, “Stair climbing for humanoid robots using stereo vision,” in Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1407-1413, 2004, Sendai, Japan

[84]. R. Kurazume and S. Hirose, “Development of image stabilization system for a remote operation of walking robots,” in Proceedings of IEEE International Conference on Robot and Automation, pp. 1856-1861, 2000, San Francisco, USA

[85]. T. Shibata and S. Schaal, “Biomimetic gaze stabilization based on feedback-error-learning with nonparametric regression networks”, Neural Networks, Vol. 14, pp. 201-216, 2001

[86]. N. Rescher and P. Oppenheim, “Logical analysis of gestalt concepts,” British Journal for the Philosophy of Science, Vol. 6, pp. 89-106, 1955

[87]. D. Navon, “Forest before tree: the precedence of global features in visual perception,” Cognitive Psychology, Vol. 9, pp. 353-383, 1977

[88]. R. A. Kinchla and J. M. Wolfe, “The order of visual processing: “Top-Down,” “Bottom-Up,” or “Middle-Out”,” Perception and Psychophysics, Vol. 25, pp. 225-231, 1979

[89]. M. Ichikawa, S. Motoki, and M. Shibata, “Exploring motion control for three-dimensional object recognition based on eye-in-hand visual servoing,” IEEJ Transactions on Industry Applications, 126- D(6), pp. 726–731, 2006

[90]. A. J. Davison, I. D. Reid, N. D. Molton and O. Stasse, "MonoSLAM: Real-Time single camera SLAM," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29(6), pp. 1052-1067, 2007

[91]. A. Nüchter, K. Lingemann, J. Hertzberg, and H. Surmann, “6D SLAM—3D mapping outdoor environments,” Journal of Field Robotics, vol. 24, pp. 699–722, 2007

[92]. G. Zhang, J. H. Lee, J. Lim and I. H. Suh, "Building a 3-D line-based map using stereo SLAM," IEEE Transactions on Robotics, vol. 31(6), pp. 1364-1377, 2015

[93]. K. Zhang, L. Zhang, and M. H. Yang, “Real-time compressive tracking,” in European Conference on Computer Vision, Springer, pp. 864–877, 2012, Florence, Italy

[94]. J. M. Croom, D. C. Rucker, J. M. Romano, et al., “Visual sensing of continuum robot shape using self-organizing maps,” in IEEE International Conference on Robotics and Automation, pp. 4591–4596, 2010, Anchorage, Alaska

[95]. A. Vandini, S. Giannarou, S.-L. Lee, and G.-Z. Yang, “3D robotic catheter shape reconstruction and localisation using appearance priors and adaptive c-arm positioning,” in MIAR 2013, ser. LNCS, vol. 8090. Berlin, Heidelberg: Springer-Verlag, 2013.

[96]. A. Reiter, R. E. Goldman, and A. Bajo, et al., “A learning algorithm for visual pose estimation of continuum robots,” 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 2390–2396, 2011, San Francisco, USA

[97]. V. Chitrakaran, A. Behal, D. Dawson, and I. Walker, “Setpoint regulation of continuum robots using a fixed camera,” Robotica, vol. 25(5), pp. 581–586, 2007

[98]. D. B. Camarillo, K. E. Loewke, C. R. Carlson, and J. K. Salisbury, “Vision based 3-d shape sensing of flexible manipulators,” in IEEE International Conference on Robotics and Automation, pp. 2940–2947, 2008, Pasadena, USA

[99]. B. Bardou, F. Nageotte, P. Zanne, and M. de Mathelin, “Design of a robotized flexible endoscope for natural orifice transluminal endoscopic surgery,” in Computational Surgery and Dual Training, pp. 155–170, 2010

[100]. A. Vandini, A. Salerno, C. J. Payne, “Vision-based motion control of a flexible robot for surgical applications,” 2014 IEEE International Conference on Robotics & Automation (ICRA, pp. 6205-6211), 2014, Hong Kong

[101]. T. Dallej, N. Andreff and P. Martinet, “Contribution to generic modeling and vision-based control of a broad class of fully parallel robots,” Robotica, 36, pp. 1874-1896, 2018

[102]. J. Angeles, “Fundamentals of robotic mechanical systems: Theory, methods and algorithms,” Springer, 1995

[103]. J. Merlet, “Parallel robots,’ Kluwer Academic Publishers, 2000

[104]. J. P. Merlet, “An algorithm for the forward kinematics of general 6 D.O.F parallel manipulators,” Research Report 1331, 1990

[105]. W. J. Wilson, C. C. Williams Hulls and G. S. Bell, "Relative end-effector control using Cartesian position based visual servoing," IEEE Transactions on Robotics and Automation, vol. 12(5), pp. 684-696, 1996

[106]. Z. Qi and J. E. Mcinroy, “Improved image based visual servoing with parallel robot,” Journal of Intelligent and Robotic Systems, vol. 53, 359–379, 2008

[107]. G. Palmieri, M. Palpacelli, M. Battistelli and M. Callegari, “A comparison between position-based and image-based dynamic visual servoings in the control of a translating parallel manipulator,” Journal of Robotics, Article ID: 103954, 2012

[108]. I. Bonev, ‘‘Delta parallel robot—The story of success,’’ 2001, on line available: http://www.parallelmic.org

[109]. H. B. Choia1, A. Konnoa, and M. Uchiyama, “Design, implementation, and performance evaluation of a 4-DOF parallel robot,” Robotica, vol. 28, pp. 107–18, 2010

[110]. N. Hayes, “Ethics and biometric identity: Biometrics represent a huge advancement from the centuries-old use of paper documentation but with a price,” 2019, [online]. Available: https://www.securityinfowatch.com/access-identity/biometrics/article/21072152/ethics-and-biometric-identity

[111]. W. Lao, J. Han and P. H. n. De With, "Automatic video-based human motion analyzer for consumer surveillance system," IEEE Transactions on Consumer Electronics, vol. 55(2), pp. 591-598, 2009

[112]. K. Doi, “Computer-aided diagnosis in medical imaging: Historical review, current status and future,” Computerized Medical Imaging and Graphics 31, pp. 198–211, 2007

[113]. H. Li and C. Shen, “Reading car license plates using deep convolutional neural networks and LSTMs,” CoRR, 2016, [Online]. Available: https://arxiv.org/abs/1601.05610

[114]. Q. Essid, H. laga and C. Samir, “Automatic detection and classification of manufacturing defects in metal boxes using deep neural networks,”, 2018, [online]. Available: https://doi.org/10.1371/journal.pone.0203192

[115]. K. Jo, J. Kim, and D. Kim, et al., "Development of autonomous car—Part II: A case study on the implementation of an autonomous driving system based on distributed architecture," IEEE Transactions on Industrial Electronics, vol. 62(8), pp. 5119-5132, 2015

[116]. M. Sonka, V. Hlavac, and R. Boyle, “Image processing, analysis, and machine vision,” Cengage Learning, 2014

[117]. A. Visioli, “Tuning of PID controllers with fuzzy logic,” IET Proceeding Control Theory and Application, vol. 148(1), pp: 1–8, 2001

[118]. C. Juang, G. Chen, C. Liang and D. Lee, “Stereo-camera-based object detection using fuzzy color histograms and a fuzzy classifier with depth and shape estimations,” Applied Soft Computing, vol. 46, pp. 753-755, 2016

[119]. C. Kao, C. Chuang, and R. Fung, “The self-tuning PID control in a slider–crank mechanism system by applying particle swarm optimization approach,” Mechatronics, vol. 16, pp: 513-522, 2006

[120]. H. Tian, Y. Kou, and R. Funakubo, et al., “Tracking performances of Eye-vergence visual servoing system under different light condition with evolutionary pose tracking,” in Proceedings of the 2016 IEEE/SICE International Symposium on System Integration, 2016, Sapporo, Japan

[121]. D. Kim, R. Lovelett, and A. Behal, “Eye-in-hand stereo visual servoing of an assistive robot arm in unstructured environments,” IEEE International Conference on Robotics and Automation, 2009, Kobe, Japan

[122]. H. Ka, C. Chung, and D. Ding, et al, “Performance evaluation of 3D vision-based semi-autonomous control method for assistive robotic manipulator,” Disability and Rehabilitation: Assistive Technology, Vol. 13(2), pp. 140-145, 2018

[123]. D. Kim, Z. Wang, and N. Paperno, et al., “System design and implementation of UCF-MANUS—An intelligent assistive robotic manipulator,” IEEE/ASME Transactions on Mechatronics, Vol. 19(1), pp: 225-237, 2014

[124]. N. Bourbakis, I. Ktistakis, and L. Tsoukalas, et al., “An autonomous intelligent wheelchair for assisting people at need in smart homes: A case study,” 2015 6th

International Conference on Information, Intelligence, Systems and Applications (IISA), pp. 1-7, 2015, Corfu, Greece

[125]. T. Lampe and M. Riedmiller, "Acquiring visual servoing reaching and grasping skills using neural reinforcement learning," The 2013 International Joint Conference on Neural Networks (IJCNN), pp. 1-8, 2013, Dallas, USA

[126]. D. Kim, R. Hazlett-Knudsen, and H. Culver-Godfrey, et al, “How autonomy impacts performance and satisfaction: Results from a study with spinal cord injured subjects using an assistive robot,” IEEE Transactions on System, Man, and Cybernetics-Part A: Systems and humans, vol. 42(1), pp: 2-14, 2012

[127]. H. Ka, D. Ding, and R. Cooper, “Three dimentional computer vision-based alternative control method for assistive robotic manipulator,” Symbiosis, 1(1), 2016

[128]. O. Sutherland, H. Truong, and S. Rougeaux, et al., “Advancing active vision systems by improved design and control,” In Proceedings of International Symposium on Experimental Robotics (ISER2000), 2000, Honolulu, USA

[129]. A. J. Davison, “Mobile robot navigation using active vision,” Ph.D. dissertation University of Oxford, 1999

[130]. K. Schiluns, W. Fellenz, A. Koschan and M. Teschner, “A modular 10-DOF vision system for high-resolution active stereo,” The University of Auckland, Tamaki

Campus, Computing and Information Technology Research Computer Vision Unit, no. CITR-TR-2/CS-TR141, 1997

[131]. S. J. D. Prince, “Computer vision: Models, learning, and inference”. Cambridge University Press, 2012

[132]. R. Hartley and A. Zisserman, “Multiple view geometry in computer vision,” Cambridge University Press. 2004

[133]. Chris Harris, Mike Stephens. “A combined corner and edge detector”. In Processing of 4th Alvey Vision Conference, pp.147-151, 1988, Manchester, England

[134]. P. Grice and C. Kemp, “Assistive mobile manipulation: Designing for operators with motor impairments,” RSS 2016 Workshop on Socially and Physically Assistive Robotics for Humanity, 2016, Ann Arbor, USA

[135]. G. K. M. Cheung, S. Baker and T. Kanade, "Visual hull alignment and refinement across time: a 3D reconstruction algorithm combining shape-from-silhouette with stereo," in Proceedings of 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. II-375, 2003, Madison, USA

[136]. E. Cervera, F. Berry, P. Martinet, “Is 3D useful in stereo visual control?” in Proceedings of the 2002 IEEE International Conference on Robotics & Automation, pp. 1630-1635, 2002, Washington, USA

[137]. P. Martinet and E. Cervera, "Stacking Jacobians properly in stereo visual servoing," in Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation, pp. 717-722, 2001, Seoul, South Korea.

[138]. W. G. Pence, F. Farelo, R. Alqasemi, Y. Sun and R. Dubey, "Visual servoing control of a 9-DoF WMRA to perform ADL tasks," 2012 IEEE International Conference on Robotics and Automation, pp. 916-922, 2012, Saint Paul, USA

[139]. H. Cheng, H.Chen, and Y. Liu, "Object handling using Autonomous Industrial Mobile Manipulator", 2013 IEEE 3rd Annual International Conference on Cyber Technology in Automation Control and Intelligent Systems (CYBER), pp. 36-41, 2013, Nanjing, China

[140]. C. Stanton, A. Bogdanovych, and E. Ratanasena, "Teleoperation of a humanoid robot using full-body motion capture, example movements, and machine learning," In Proceedings of Australasian Conference on Robotics and Automation, 2012, New Zealand

[141]. W. Li, C. H and et. al. "Hand gesture recognition for post-stroke rehabilitation using leap motion", 2017 IEEE International Conference on Applied System Innovation, pp. 386-388, 2017, Sapporo, Japan

[142]. K. Liu, X. Luo, and W. Wu, et al., “A fast and accurate calibration algorithm for laser pointer tracking via structured light illumination,” Electronics Letters, vol. 50(25), pp. 1946 – 1948, 2014

[143]. M. de La Gorce, D. J. Fleet and N. Paragios, "Model-based 3D hand pose estimation from monocular video," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33(9), pp. 1793-1805, 2011

[144]. L. Y. Deng, Dong-Liang Lee, Huan-Chao Keh and Yi-Jen Liu, "Shape context based matching for hand gesture recognition," IET International Conference on Frontier Computing. Theory, Technologies and Applications, pp. 436-444, 2010, Taiwan

[145]. C. Hsieh, D. Liou and D. Lee, "A real time hand gesture recognition system using motion history image," *2010* 2nd International Conference on Signal Processing Systems, pp. V2-394-V2-398, 2010, Dalian, China

[146]. T. Starner, J. Auxier, D. Ashbrook and M. Gandy, "The gesture pendant: a self-illuminating, wearable, infrared computer vision system for home automation control and medical monitoring," Digest of Papers. Fourth International Symposium on Wearable Computers, pp. 87-94, 2000, Atlanta, USA

[147]. J. W. Davis, and S. Vaks, “A perceptual user interface for recognizing head gesture acknowledgements,” In Proceedings of the 2001 workshop on perceptive user interfaces, ACM Press, pp. 1–7, 2001, Orlando, USA

[148]. M. Schultz, J. Gill, and S. Zubairi, et al., “Bacterial contamination of computer keyboards in a teaching hospital,” Infection Control and Hospital Epidemiology, Vol.4(24), pp.302–303, 2003

[149]. S. M. Goza, R. O. Ambrose, M. A. Diftler, I. M. Spain, “Telepresence control of the nasa/darpa robonaut on a mobility platform,” In Conference on human factors in computing systems. ACM Press, pp 623–629, 2004, Vienna, Austria

[150]. S. S. Rautaray and A. Agrawal, "Interaction with virtual game through hand gesture recognition," 2011 International Conference on Multimedia, Signal Processing and Communication Technologies, pp. 244-247, 2011, Aligarh, India

[151]. L. E. Sucar, F. Orihuela-Espina, R. L. Velazquez, D. J. Reinkensmeyer, R. Leder and J. Hernández-Franco, "Gesture therapy: An upper limb virtual reality-based motor rehabilitation platform," IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 22(3), pp. 634-643, 2014

[152]. T. Boros and S. Daniel Dumitrescu, "A “small-data”-driven approach to dialogue systems for natural language human computer interaction," 2017 International

Conference on Speech Technology and Human-Computer Dialogue (SpeD, pp. 1-6), 2017, Bucharest, Romania

[153]. P. Kasprowski, K. Harezlak and M. Niezabitowski, "Eye movement tracking as a new promising modality for human computer interaction," 2016 17th International Carpathian Control Conference (ICCC), pp. 314-318, 2016, Tatranska Lomnica, Slovak

[154]. L. Sigal, A. O. Balan, and M. J. Black, “HUMANEVA: Synchronized video and motion capture dataset and baseline algorithm for evaluation of articulated human motion”, International journal of computer vision, Vol. 87, pp. 4-27, 2010

[155]. I. Davliakos and E. Papadopoulos, “Model-based control of a 6-dof electrohydraulic Stewart–Gough platform,” Mechanism and Machine Theory, Vol. 43 (11), pp. 1385-1400, 2008

[156]. M.R. Sirouspour and S.E. Salcudean, “Nonlinear control of hydraulic robots,” IEEE Transactions on Robotics and Automation, Vol. 17 (2), pp. 173-182, 2001

[157]. D.H. Kim, J.Y. Kang and K.I. Lee, “Robust tracking control design for a 6 DOF parallel manipulator,” Journal of Robotic Systems, Vol. 17 (10), pp. 527-547, 2000

[158]. X. S. Gao, D. L. Lei, Q. Z. Liao, and G. F. Zhang, “Generalized Stewart–Gough platforms and their direct kinematics. IEEE Transactions on Robotics, Vol. 21(2), pp. 141–151, 2005

[159]. A. Ghobakhloo, M. Eghtesad, and M. Azadi, “Position control of a Stewart– Gough platform using inverse dynamics method with full dynamics,” In 9th IEEE International Workshop on Advanced Motion Control, pp. 50–55, 2006, Istanbul, Turkey

[160]. C. Guan, and S. X. Pan, “Adaptive sliding mode control of electro-hydraulic system with nonlinear unknown parameters,” Control Engineering Practice, Vol. 16(11), pp. 1275–1284, 2008

[161]. H.B. Guo, and H.R. Li, “Dynamic analysis and simulation of a six degree of freedom Stewart platform manipulator,” in Proceedings of the Institution of Mechanical Engineers Part C - Journal of Mechanical Engineering Science, Vol. 220(1), pp. 61–72, 2006

[162]. H.B. Guo, Y.G. Liu, G. R. Liu and H. R. Li, “Cascade control of a hydraulically driven 6-DOF parallel robot manipulator based on a sliding mode,” Control Engineering Practice, Vol. 16(9), pp. 1055–1068, 2008

[163]. D. H. Kim, J. Y. Kang and K. I. Lee, “Robust tracking control design for a 6 DOF parallel manipulator,” Journal of Robotic Systems, Vol. 17(10), pp. 527–547, 2001

[164]. G. Lebret, K. Liu, and F. L. Lewis, “Dynamic analysis and control of a Stewart platform manipulator,” Journal of Robotic Systems, Vol. 10(5), pp. 629–655, 1993

[165]. S. H. Lee, J. B. Song, W. C. Choi, and D. H. Hong, “Position control of a Stewart platform using inverse dynamics control with approximate dynamics,” Mechatronics, 13(6), pp. 605–619, 2003

[166]. J. P. Merlet, “Solving the forward kinematics of a Gough-type parallel manipulator with interval analysis,” International Journal of Robotics Research, 23(3), pp. 221–235, 2004

[167]. M. R. Sirouspour, and S. E. Salcudean, “Nonlinear control of hydraulic robots,” IEEE Transactions on Robotics and Automation, 17(2), pp. 173–182, 2000

[168]. G. A. Sohl, and J. E. Bobrow, “Experiments and simulations on the nonlinear control of a hydraulic servo system,” IEEE Transactions on Control Systems Technology, vol. 7(2), pp. 238–247, 1999

[169]. T. Burg, D. Dawson, and P. Vedagarbha, “A redesigned DCAL controller without velocity measurements: Theory and demonstration,” Robotica, vol. 15, pp. 337–346, 1997

[170]. C. G. Kang, “Closed-form force sensing of a 6-axis force transducer based on the Stewart platform,” Sensors and Actuators A: Physical, vol. 90(1–2), pp. 31–37, 2001

[171]. N. I. Kim and C. W. Lee, “High speed tracking control of Stewart platform manipulator via enhanced sliding mode control,” in Proceedings of IEEE International Conference on Robotics and Automation, pp. 2716–2721, 1998, Leuven, Belgium

[172]. K. Kosuge, M. Okuda, and T. Fukuda, et al., “Input/output force analysis of Stewart platform type of manipulators,” in Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1666–1673, 1993, Yokohama, Japan.

[173]. F. Pierrot, F. Marquet, O. Company, and T. Gil, “H4 parallel robot: modeling, design and preliminary experiments,” in Proceedings of the 2001 IEEE International Conference on Robotics & Automation, pp. 3256–3261, 2001, Seoul, Korea

[174]. F. H. Ghorbel, O. Chetelat, R. Gunawardana and R. Longchamp, "Modeling and set point control of closed-chain mechanisms: theory and experiment," IEEE Transactions on Control Systems Technology, vol. 8,(5), pp. 801-815, 2000

[175]. F. X. Wu, W. J. Zhang, Q. Li, and P. R. Ouyang, “Integrated design and PD control of high speed closed-loop mechanism,” ASME Journal of Dynamic Systems, Measurement, and Control, Vol. 124(4), pp. 522–528, 2002

[176]. I. A. Bonev, and J. A. Ryu, “New method for solving the direct kinematics of general 6-6 Stewart platforms using three linear extra sensors,” Mechanism and Machine Theory, Vol. 35, pp. 423–436, 2000

[177]. G. Innocenti and V. Parenti-Castelli, “Forward kinematics of the general 6-6 fully parallel mechanism: An exhaustive numerical approach via a mono dimensional search algorithm,” Journal of Mechanical Design, Vol. 12, pp.932–937, 1993

[178]. J. Y. Kang, D. H. Kim, and K. I. Lee, “Robust estimator design for forward kinematics solution of a Stewart platform,” Journal of Robotic Systems, Vol. 15, pp. 30–42.

[179]. Q. Li and FX. Wu, “Control performance improvement of a parallel robot via the design for control approach,” Mechatronics, Vol.14, pp: 947-964, 2004

[180]. Y. Jin, H. Chanal, and F. Paccot, “Parallel Robots,” In: Nee A. (eds) Handbook of Manufacturing Engineering and Technology. Springer, 2015

[181]. L. Bonekamp and M. Sure, “Consequences of industry 4.0 on human labour and work organisation,” Journal of Business and Media Psychology, Vol. 6(1), pp: 33-40, 2015

[182]. T. Stock and G. Seliger, “Opportunities of sustainable manufacturing in industry 4.0,” Procedia, Vol. n40, pp. 536-541, 2016

[183]. J. Bowles, “The computerisation of European jobs - Who will win and who will lose from the impact of new technology onto old areas of employment? ” 2014, Online only at: http://www.bruegel.org/nc/blog/detail/article/1394-the-computerisation-of-european-jobs/

[184]. U. Schmitt, “Personal knowledge management for development (PKM4D) framework and its application for people empowerment,’ Elsevier Procedia Computer

Science, International Conference on Knowledge Management (ICKM), Vol.99, pp.64-78, 2016, Vienna

[185]. S. Z. Li and A. K. Jain, “Handbook of face recognition,” Springer, 2005

[186]. T. Ahonen, A. Hadid and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28(12), pp. 2037-2041, 2006

[187]. J. Wang, Y. Ma, L. Zhang and et. al., “Deep learning for smart manufacturing: Methods and applications,” Journal of Manufacturing Systems, Vol. 48, pp.144-156, 2018

[188]. A. Hornberg, “Handbook of machine vision,” Wiley–VCH Verlag GmbH & Co, KGaA, Weinheim, 2006

[189]. C. D. Herrera, J. Kannala and J. Heikkilä, "Joint depth and color camera Calibration with distortion correction," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34(10), pp. 2058-2064, 2012

[190]. J. Weng, P. Cohen and M. Herniou, "Camera calibration with distortion models and accuracy evaluation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 14(10), pp. 965-980, 1992

[191]. B. Horn, “Robot vision,” Cambridge, MA: MIT Press, 1986

[192]. Y. Zeng, X. Xu, D. Shen, Y. Fang and Z. Xiao, "Traffic sign recognition using kernel extreme learning machines with deep perceptual features," IEEE Transactions on Intelligent Transportation Systems, vol. 18(6), pp. 1647-1653, 2017

[193]. A. H. A. Hafez, E. Cervera and C. V. Jawahar, "Hybrid visual servoing by boosting IBVS and PBVS," 2008 3rd International Conference on Information and Communication Technologies: From Theory to Applications, pp. 1-6, 2008, Damascus, Syria

[194]. G. Chesi, K. Hashimoto, D. Prattichizzo, and A. Vicino, “Keeping features in the field of view in eye-in-hand visual servoing: A switching approach,” IEEE Transactions on Robotics, vol. 20(5), pp. 534–549, 2004

[195]. Y. Mezouar and F. Chaumette, “Path planning for robust image-based control,” IEEE Transactions on Robotics and Automation, vol. 18(4), pp. 534–549, 2002

[196]. V. Kyrki, D. Kragic, and H. I. Christensen, “New shortest-path approches to visual servoing,” in IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS’04, pp. 349– 354, 2004, Sendai, Japan,

[197]. E. Malis, F. Chaumette, and S. Boudet, “2 1/2 D visual servoing,” IEEE Transactions on Robotics and Automation, vol. 15(2), pp. 238–250,1999

[198]. E. Cervera, A. P. D. Pobil, F. Berry, and P. Martinet, “Improving image - based visual servoing with three-dimensional features.” International Journal of Robotics Research, vol. 22(10-11), pp. 821–840, 2003

[199]. A. H. Abdul Hafez and C. V. Jawahar, "Integration framework for improved visual servoing in image and cartesian spaces," *2006* IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 2320-2325, 2006, Bejing, China

[200]. N. Gans and S. A. Hutchinson, “An asymptotically stable switched system visual controller for eye in hand robots,” in IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 3061–3068, 2003, Las Vegas, USA

[201]. B. H. Lee and C. S. G. Lee, "Collision-free motion planning of two robots," IEEE Transactions on Systems, Man, and Cybernetics, vol. 17(1), pp. 21-32, 1987

[202]. S. B. Marapane, M. M. Trivedi, N. Lassiter and M. B. Holder, "Motion control of cooperative robotic teams through visual observation and fuzzy logic control," in Proceedings of IEEE International Conference on Robotics and Automation, pp. 1738-1743, 1996, Minneapolis, USA

[203]. K. Sabe, M. Fukuchi, J. -. Gutmann, T. Ohashi, K. Kawamoto and T. Yoshigahara, "Obstacle avoidance and path planning for humanoid robots using stereo vision," IEEE International Conference on Robotics and Automation, pp. 592-597, 2004, New Orleans, USA

[204]. C. Juang and C. Hsu, "Reinforcement ant optimized fuzzy controller for mobile-robot wall-following control," IEEE Transactions on Industrial Electronics, vol. 56(10), pp. 3931-3940, 2009

[205]. M. Seder and I. Petrovic, "Dynamic window based approach to mobile robot motion control in the presence of moving obstacles," in Proceedings 2007 IEEE International Conference on Robotics and Automation, pp. 1986-1991, 2007, Roma, Italy

[206]. J. Tani, "Model-based learning for mobile robot navigation from the dynamical systems perspective," IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 26(3), pp. 421-436, 1996

[207]. S. Dogramadzi, M. Giannaccini, and C. Harper, et. al., “Environmental hazard analysis – a variant of preliminary hazard analysis for autonomous mobile robots,” Journal of intelligent & robotic systems, Vol. 76 (1), pp. 73-117, 2014

[208]. S. Y. Chung and H. P. Huang, "A mobile robot that understands pedestrian spatial behaviors," 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 5861-5866, 2010, Taipei

[209]. P. T. Szemes, H. Hashimoto and P. Korondi, "Pedestrian-behavior-based mobile agent control in intelligent space," IEEE Transactions on Instrumentation and Measurement, vol. 54(6), pp. 2250-2257, 2005

[210]. W. G. Han, S. M. Baek and T. Y. Kuc, "Genetic algorithm based path planning and dynamic obstacle avoidance of mobile robots," 1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation, pp. 2747-2751, 1997, Orlando, USA

[211]. C. Qixin, H. Yanwen and Z. Jingliang, "An evolutionary artificial potential field algorithm for dynamic path planning of mobile robot," 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 3331-3336, 2006, Beijing, China

[212]. M. Kam, Xiaoxun Zhu and P. Kalata, "Sensor fusion for mobile robot navigation," in Proceedings of the IEEE, vol. 85(1), pp. 108-119, 1997

[213]. J. Kim, Y. Kim and S. Kim, "An accurate localization for mobile robot using extended Kalman filter and sensor fusion," 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), pp. 2928-2933, 2008, Hong Kong

[214]. F. Lizarralde, E. V. L. Nunes, Liu Hsu and J. T. Wen, "Mobile robot navigation using sensor fusion," 2003 IEEE International Conference on Robotics and Automation, pp. 458-463, 2003, Taipei, Taiwan

[215]. N. Andreff and P. Martinet, "Unifying kinematic modeling, identification, and control of a Gough–Stewart parallel robot into a vision-based framework," IEEE Transactions on Robotics, vol. 22(6), pp. 1077-1086, 2006

[216]. L. Baron and J. Angeles, “The on-line direct kinematics of parallel manipulators under joint-sensor redundancy,” in Advances in Robot Kinematics. Norwell, MA: Kluwer, pp. 126–137, 1998

[217]. P. Renaud, N. Andreff, P. Martinet and G. Gogu, "Kinematic calibration of parallel mechanisms: a novel approach using legs observation," IEEE Transactions on Robotics, vol. 21(4), pp. 529-538, 2005

[218]. B. Dasgupta and T. Mruthyunjaya, “Force redundancy in parallel manipulators: Theoretical and practical issues,” Mechanism and Machine Theory, vol. 33(6), pp. 727–742, 1998

[219]. B. Espiau, F. Chaumette, and P. Rives, “A new approach to visual servoing in robotics,” IEEE Transactions on Robotics Automation, vol. 8(3), pp. 313–326,1992

[220]. N. Hirose, R. Tajima, and K. Sukigara, “Personal robot assisting transportation to support active human life—Human-following method based on model predictive control for adjacency without collision,” in Proceeding IEEE International Conference on Mechatronics (ICM), 2015, pp. 76–81, 2015, Nagoya, Japan,

[221]. M. Gupta, S. Kumar, L. Behera, and V. K. Subramanian, “A novel vision-based tracking algorithm for a human-following mobile robot,” IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 47(7), pp. 1415-1427, 2017

[222]. N. Bellotto and H. Hu, “Multisensor-based human detection and tracking for mobile service robots,” IEEE Transactions Systems, Man, and Cybernetics, B, Cybernetics., vol. 39(1), pp. 167–181, 2009

[223]. Q. Xiao, F. Sun, R. Ge and et al., “Human tracking and following of mobile robot with a laser scanner,” 2017 2nd International Conference on Advanced Robotics and Mechatronics (ICARM), pp. 675-680, 2017, Anhui, China

[224]. Hoshino F and Morioka K, “Human following robot based on control of particle distribution with integrated range sensors,” IEEE/SICE International Symposium on System Integration, pp. 212-217, 2011, Kyoto, Japan

[225]. T. Yoshimi, M. Nishiyama, and T. Sonoura, et al., “Development of a person following robot with vision based target detection,” IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China, 2006, pp. 5286-5291.

[226]. J. F. Henriques, R. Caseiro, P. Martins, and et al., “High-speed tracking with kernelized correlation filters,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37(3), pp. 583–596, 2015

[227]. S. Liu, T. Zhang, X. Cao, and C. Xu, “Structural correlation filter for robust visual tracking,” in Proceeding IEEE Conference Computer Vision Pattern Recognition (CVPR), pp. 4312–4320, 2016, Les vegas, USA

[228]. T. Liu, G. Wang, and Q. Yang, “Real-time part-based visual tracking via adaptive correlation filters,” in Proceeding IEEE Conference Computer Vision Pattern Recognition (CVPR), pp. 4902–4912, 2015, Boston, USA

[229]. M. Wang, Y. Liu, and D. Su and et al., “Accurate and real-time 3-D tracking for the following robots by fusing vision and ultrasonar information,” IEEE/ASME Transactions On Mechatronics, vol. 23(3), pp. 997-1006, 2018

[230]. X. Wang, M. O'Brien, C. Xiang, B. Xu and H. Najjaran, "Real-time visual tracking via robust kernelized correlation filter," 2017 IEEE International Conference on Robotics and Automation (ICRA), pp. 4443-4448, 2017, Singapore

[231]. H. Taheri, B. Qiao and N. Ghaeminezhad, “Kinematic model of a four mecanum Wheeled Mobile,” International Journal of Computer Applications, vol. 113(3), pp. 6-9, 2015

[232]. A. Ess, B. Leibe and L. Van Gool, “Depth and appearance for mobile scene analysis,” IEEE 11th International Conference on Computer Vision, pp. 1-8, 2007, Rio de Janeiro, Brazil

[233]. G. Dewaele, F. Devernay, and R. Horaud, “Hand motion from 3D point trajectories and a smooth surface model,” in Proceeding Europe Conference Computer Vision, pp. 495–507, 2004, Prague, Czech Republic

[234]. F. Chaumette, "Image moments: a general and useful set of features for visual servoing," IEEE Transactions on Robotics, vol. 20(4), pp. 713-723, 2004

[235]. X. Zhao, and S. Peng, “Direct displacement analysis of parallel manipulators,” Journal of Robotics Systems, Vol. 17(6), pp. 341—345, 2000

[236]. P. Renaud, N. Andreff, G. Gogu, and P. Martinet, “On vision-based kinematic calibration of a Stewart—Gough platform,” 11th World Congress in Mechanism and Machine Science (IFToMM2003), pp. 1906— 1911, 2003, Tianjin, China,

[237]. E. Marchand, F. Spindler and F. Chaumette, "ViSP for visual servoing: A generic software platform with a wide class of robot control skills," IEEE Robotics & Automation Magazine, vol. 12(4), pp. 40-52, 2005

[238]. R. Mahony and T. Hamel, "Visual servoing using linear features for under-actuated rigid body dynamics," in Proceedings 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems. Expanding the Societal Role of Robotics in the Next Millennium, vol.2, pp. 1153-1158, Maui, USA, 2001

[239]. H. Kino, T. Yahiro, F. Takemura and T. Morizono, "Robust PD control using adaptive compensation for completely restrained parallel-wire driven robots: Translational systems using the minimum number of wires under Zero-gravity condition," IEEE Transactions on Robotics, vol. 23(4), pp. 803-812, 2007

[240]. Z. Ren and J. Yuan, “Depth camera based hand gesture recognition and its applications in Human-computer-interaction,” IEEE 8th International Conference on Information, Communications and Signal Processing, pp. 1-5, 2011, Singapore

[241]. C. Lee and Y. Xu, “Online, interactive learning of gestures for human/robot interfaces”, IEEE International Conference on Robotics and Automation, vol.4, pp. 2982-2987, 1996, Minneapolis, USA

[242]. H. Junker, O. Amft, and P. Lukowicz, et. al., “Gesture spotting with body-worn inertial sensors to detect user activities”, Pattern Recognition, Vol. 41, pp. 2010-2024, 2008

[243]. D. L. Miranda, M. J. Rainbow, and J. J. Crisco, et. al., “Kinematic differences between optical motion capture and biplanar videoradiography during a jump–cut maneuver”, Journal of Biomechanics, Vol. 43, pp. 567-573, 2013

[244]. Z. Ren, J. Yuan, and J. Meng, et. al., “Robust part-based hand gesture recognition using Kinect sensor”, IEEE Transactions on Multimedia, Vol. 15(5): pp. 1110-1120, 2013

[245]. N. Shimada, Y. Shirai, Y. Kuno and J. Miura, "Hand gesture estimation and model refinement using monocular camera-ambiguity limitation by inequality constraints," in Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition, pp. 268-273, 1998, Nara, Japan

[246]. C. Chua, H. Guan, and Y. Ho, “Model-based 3d hand posture estimation from a single 2d image,” Image and Vision Computing, Vol. 20, pp.191 – 202, 2002

[247]. C. Hsieh, D. Liou and D. Lee. “A real time hand gesture recognition system using motion history image”, The 2nd International Conference on Signal Processing Systems, pp.394-439, 2010, Dalian, China

[248]. M. Zeller, J. C. Phillips, and A. Dalke, et. al, “A visual computing environment for very large scale biomolecular modeling,” IEEE 24th International Conference on Application-Specific Systems, Architectures and Processors, pp. 3-12, 1997, Zurich

[249]. C. Wu, W. Chen and C. Lin, “Depth-based hand gesture recognition,” Multimedia Tools and Applications, Vol. 75(12), pp. 7065 – 7086, 2016

[250]. S. Soleimanizadeh, D. Mohamad, T. Saba, and Rehman A, “Recognition of partially occluded objects based on the three different color spaces (RGB, YCbCr, HSV),” 3D Research, Vol. 6(3): pp.1–10, 2015

[251]. H. Zhao, J. Hu, Y. Zhang and H. Cheng, "Hand gesture based control strategy for mobile robots," 2017 29th Chinese Control And Decision Conference (CCDC), pp. 5868-5872, 2017, Chongqing, China

[252]. S. Kundu, O. Mazumder, P. K. Lenka and S. Bhaumik, “Hand gesture recognition based omnidirectional wheelchair control using IMU and EMG sensors,” Journal of Intelligent & Robotic Systems, vol. 91, pp. 1-13, 2017

[253]. C. Manresa, J. Varona, R. Mas and F. j. Perales, “Perales hand tracking and gesture recognition for human-computer interaction”, Electron Letters Computer Vision and Image Analysis, 5(3), pp. 96–104, 2005

[254]. H. Yeo, B. Lee and H. Lim. “Hand tracking and gesture recognition system for Human-computer interaction using Low-cost hardware,” Multimedia Tools and Application, vol. 74, pp. 2687-2715, 2015

[255]. M. Georgi, C. Amma, and T. Schultz, “Recognizing hand and finger gestures with IMU based motion and EMG based muscle activity sensing,” In: Proceedings of the international conference on bio-inspired systems and signal processing. Science and Technology Publications, 2015

[256]. R. Xu, S. Zhou and W. J. Li, "MEMS accelerometer based nonspecific-user hand gesture recognition," IEEE Sensors Journal, vol. 12(5), pp. 1166-1173, 2012

[257]. L. Colasanto, R. Suarez and J. Rosell, "Hybrid mapping for the assistance of teleoperated grasping tasks," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 43(2), pp. 390-401, 2013

[258]. G. Heumer, H. B. Amor and B. Jung, "Grasp recognition for uncalibrated data gloves: A machine learning approach," in Presence, vol. 17, no. 2, pp. 121-142, 2008

[259]. A. H. Al-Timemy, R. N. Khushaba, G. Bugmann and J. Escudero, "Improving the performance against force variation of EMG controlled multifunctional upper-limb

prostheses for transradial Amputees," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 24(6), pp. 650-661, 2016

[260]. F. Riillo, L. R. Quitadamo, and F. Cavrini, et al., “Optimization of EMG-based hand gesture recognition: supervised vs. unsupervised data preprocessing on healthy subjects and transradial amputees,” Biomed Signal Process Control, vol. 14, pp. 117–125, 2014

[261]. M. Bray, E. Koller-Meier and L. Van Gool, "Smart particle filtering for 3D hand tracking," in Proceeding Sixth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 675-680, 2004, Seoul, South Korea

[262]. J. Y. Lin, Ying Wu and T. S. Huang, "3D model-based hand tracking using stochastic direct search method," in Proceeding Sixth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 693-698, 2004, Seoul, South Korea

[263]. J. Segen and S. Kumar, “Gesture vr: Vision-based 3d hand interface for spatial interaction,” in Proceeding ACM International Conference on Multimedia, pp. 455–464, 1998

[264]. M. J. Reale, S. Canavan, L. Yin, K. Hu and T. Hung, "A multi-gesture interaction system using a 3-D iris disk model for gaze estimation and an active appearance model for 3-D hand pointing," IEEE Transactions on Multimedia, vol. 13(3), pp. 474-486, 2011

[265]. A. Ohnsman, “Postmates’ delivery robots have a phantom babysitter,” 2019, online, Available: https://www.forbes.com/sites/alanohnsman/2019/09/23/postmates-delivery-robots-have-a-phantom-babysitter/#2f7db0d719a7

[266]. Y. Wu, J. Lim, and M. H. Yang, “Online object tracking: A Banchmark,” The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013, pp. 2411-2418

[267]. W. Abbeloos, and T. Goedeme, “Point pair feature based object detection for random bin picking,” 2016 13th Conference on Computer and Robot Vision (CRV), 2016, Victoria, Canada

[268]. R. Bloss, “Smart robot that picks parts from bins,” Assembly Automation, Vol. 26(4), pp. 279-282, 2006

[269]. Brown, E., Rodenberg, N., Amend, J. and et al,; “Universal robotic gripper based on the jamming of granular material”, in Proceedings of the National Academy of Sciences of the United States of America, 107(44), pp. 18809–18814, 2010

[270]. X. Luo, “Study on application of machine vision in sources-based human computer interaction,” Master’s thesis, Sichuan University, 2013

[271]. S. Wei and S. Lai, "Fast template matching based on normalized cross correlation with adaptive multilevel winner update," in IEEE Transactions on Image Processing, vol. 17(11), pp. 2227-2235, 2008

[272]. Feng Zhao, Qingming Huang and Wen Gao, "Image matching by normalized cross-correlation," *2006* IEEE International Conference on Acoustics Speech and Signal Processing Proceedings, 2006, Toulouse, France

[273]. C. Raposo, J. P. Barreto and U. Nunes, "Fast and accurate calibration of a kinect sensor," 2013 International Conference on 3D Vision - 3DV 2013, pp. 342-349, 2013, Seattle, USA

[274]. D. Herrera C., J. Kannala and J. Heikkilä, "Joint depth and color camera Calibration with distortion correction," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34(10), pp. 2058-2064, Oct. 2012

[275]. W. Flores-Fuentes, J. C. Rodriguez-Quinonez, and D. Hernandez-Balbuena, et al., "Machine vision supported by artificial intelligence," 2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE), pp. 1949-1954, 2014, Istanbul, Turkey

[276]. C. Shan, S. Gong, and P. W. McOwan, “Facial expression recognition based on local binary patterns: A comprehensive study,” Image and Vision Computing, vol. 27, pp. 803-816, 2009

[277]. G. Zhao, X. Huang, and M. Taini, et. al., “Facial expression recognition from near-infrared videos,” Image and Vision Computing, vol. 29, pp. 607-619, 2011

[278]. R. Pydipati, T. F. Burks, and W. S. Lee, “Statistical and neural network classifiers for citrus disease detection using machine vision,” Transactions of the ASAE, Vol. 48(5): 2007−2014, 2005

[279]. C. L. Huang and C. J. Wang, “A GA-based feature selection and parameters optimization for support vector machines,” Expert Systems with Applications, vol. 31, pp. 231–240, 2006

[280]. J. Krumm and S. A. Shafer, "Shape from periodic texture using the spectrogram," in Proceedings 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 284-289, 1992, Champaign, USA

[281]. M. P. Kumar, P. H. S. Torr and A. Zisserman, "OBJCUT: Efficient segmentation using top-down and bottom-up cues," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32(3), pp. 530-545, 2010

[282]. J. Krumm and S. A. Shafer, "Local spatial frequency analysis of image texture," in Proceedings Third International Conference on Computer Vision, pp. 354-358, 1990, Osaka, Japan

[283]. C. Thurau and V. Hlavac, "Pose primitive based human action recognition in videos or still images," 2008 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-8, 2008, Anchorage, Alaska

[284]. S. Vijayanarasimhan and K. Grauman, "What's it going to cost you?: Predicting effort vs. informativeness for multi-label image annotations," 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 2262-2269, 2009, Miami, USA

[285]. Yan, Jie Yang and Hauptmann, "Automatically labeling video data using multi-class active learning," in Proceedings Ninth IEEE International Conference on Computer Vision, pp. 516-523, 2003, Nice, France

[286]. A. Kapoor, K. Grauman, R. Urtasun and T. Darrell, "Active learning with Gaussian processes for object categorization," 2007 IEEE 11th International Conference on Computer Vision, pp. 1-8, 2007, Rio de Janeiro, Brazil

[287]. S. Vijayanarasimhan and K. Grauman, “Multi-level active prediction of useful image annotations for recognition,” In Advances in Neural Information Processing Systems 21, Vancouver, Canada, 2008

[288]. A. Hadid, O. Kouropteva and M. Pietikainen, "Unsupervised learning using locally linear embedding: experiments with face pose analysis," Object recognition supported by user interaction for service robots, 2002, Quebec City, Canada,

[289]. Y. Li, M. Paluri, J. M. Rehg, and P. Dollar, “Unsupervised learning of edges,” The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1619-1627, 2016, Las Vegas, USA

[290]. J. Janai, F. Guney, and A. Ranjan, et al., “Unsupervised Learning of Multi-Frame Optical Flow with Occlusions,” The European Conference on Computer Vision (ECCV), pp. 690-706, 2018, Munich Germany

[291]. S. Hare, S. Golodetz, and A. Saffari, et al., "Struck: Structured Output Tracking with Kernels," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38(10), pp. 2096-2109, 2016

[292]. W. Zhong, H. Lu and M. Yang, "Robust object tracking via sparsity-based collaborative model," 2012 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1838-1845, 2012, Providence

Appendices

Appendix A: List of Publication

This list includes papers published by the author that related to this dissertation.

[1]. Xueling Luo, Xiaodong Jin, Dan Zhang, and Qi Zou. Vision-based Control Approach for Generalized Planar Parallel Robots. The IEEE International Conference on Automation, Control and Robotics Engineering (IEEE-CACRE 2020). (Accepted)

[2]. Xueling Luo, Dan Zhang and Xiaodong Jin. A Real-time Moving Target Following Mobile Robot System with Depth Camera. 2018 4th International Conference on Mechanical and Aeronautical Engineering, Bangkok, 2018

[3]. Xueling Luo, Andrea Amighetti and Dan Zhang. A Human-Robot Interaction for a Mecanum Wheeled Mobile Robot with Real-Time 3D Two-Hand Gesture Recognition, 2019 3rd International Conference on Artificial Intelligence, Automation and Control Technologies (AIACT 2019), Xi'an, 2019 (awarded "Excellent Oral Presentation")