Design and Development of Robotic Part Assembly System under Vision Guidance

Bunil Kumar Balabantaray

Department of Industrial Design

National Institute of Technology Rourkela

Design and Development of Robotic Part Assembly System under Vision Guidance

Dissertation submitted in partial fulfillment

of the requirements of the degree of

Doctor of Philosophy

in

Industrial Design

by

Bunil Kumar Balabantaray

(Roll Number: 512ID102)

based on the research carried out

Under supervision of

Prof. Bibhuti Bhusan Biswal

September, 2016

Department of Industrial Design

National Institute of Technology Rourkela

Department of Industrial Design

National Institute of Technology Rourkela

September, 2016

Certificate of Examination

Roll Number: *512ID102*

Name: *Bunil Kumar Balabantaray*

Title of Dissertation: *Design and Development of Robotic Part Assembly System under Vision Guidance*

We the below signed, after checking the dissertation mentioned above and the official record book (s) of the student, hereby state our approval of the dissertation submitted in partial fulfillment of the requirements of the degree of *Doctor of Philosophy* in *Industrial Design* at National Institute of Technology Rourkela. We are satisfied with the volume, quality, correctness, and originality of the work.

Prof. B. B. Biswal
Principal Supervisor

Prof. D. R. Parhi
Chairman, DSC

Prof. M. R. Khan
Member, DSC

Prof. B Subudhi
Member, DSC

External examiner

Prof. Pankaj Sa
Member, DSC

Head of the department

Department of Industrial Design

National Institute of Technology Rourkela

Prof. Bibhuti Bhusan Biswal

Professor and Dean (Faculty Welfare**)**

September, 2016

Supervisor's Certificate

This is to certify that the work presented in this dissertation entitled *"Design and Development of Robotic Part Assembly System under Vision Guidance"* by *Bunil Kumar Balabantaray*, Roll Number 512ID102, is a record of original research carried out by him under my supervision and guidance in partial fulfillment of the requirements of the degree of *Doctor of Philosophy* in *Industrial Design*. Neither this dissertation nor any part of it has been submitted earlier for any degree or diploma to any institute or university in India or abroad.

_____

Bibhuti Bhusan Biswal

Principal Supervisor

Dedicated

To

My Loving

Family

Declaration of Originality

I, *Bunil Kumar Balabantaray*, Roll Number: *512ID102* hereby declare that this dissertation entitled *Design and Development of Robotic Part Assembly System under Vision Guidance* presents my original work carried out as a doctoral student of NIT Rourkela and, to the best of my knowledge, contains no material previously published or written by another person, nor any material presented by me for the award of any degree or diploma of NIT Rourkela or any other institution. Any contribution made to this research by others, with whom I have worked at NIT Rourkela or elsewhere, is explicitly acknowledged in the dissertation. Works of other authors cited in this dissertation have been duly acknowledged under the section "References". I have also submitted my original research records to the scrutiny committee for evaluation of my dissertation.

I am fully aware that in case of any non-compliance detected in future, the Senate of NIT Rourkela may withdraw the degree awarded to me on the basis of the present dissertation.

September, 2016
NIT Rourkela

Bunil Kumar Balabantaray

Acknowledgment

This dissertation is a result of the research work that has been carried out at **National Institute of Technology, Rourkela**. During this period, the author came across with a great number of people whose contributions in various ways helped in the field of research and they deserve special thanks. It is a pleasure to convey the gratitude to all of them.

First of all the author expresses his heartiest gratitude to his supervisor and guide **Dr. B. B. Biswal,** Professor and Dean faculty Welfare, NIT, Rourkela for his valuable guidance, support and encouragement in the course of the present work. The successful and timely completion of the work is due to his constant inspiration and constructive criticisms. The author cannot adequately express his appreciation to him. The author records his gratefulness to Madam **Mrs. Meenati Biswal** for her constant support and inspiration during his work and stay at NIT, Rourkela.

The author take this opportunity to express his deepest gratitude to **Prof. M.R. Khan,** Head of the Department and faculty members of the Department of Industrial Design, NIT Rourkela for constant advice, useful discussions, encouragement and support in pursuing the research work.

The author is grateful to **Prof. A Biswas**, Director and **Prof. B. Majhi**, Dean Academic Affairs, NIT, Rourkela, for their kind support and concern regarding his academic requirements.

The author also expresses his thankfulness to **Ms. Nibedita Mishra**, **Mr. O P Sahu**, **Mr. Raju, Mr. P. Jha**, **Mr. Nagmani**, **Mr. G. Balamurali**, **Mr. Golak Mahanta, Ms. Amruta Rout, Mr. P. Jain and Mr. P. Sahu**, researchers and many research scholar friends at NIT Rourkela for unhesitating cooperation extended during the tenure of the research programme.

The completion of this work came at the expense of author's long hours of absence from home. Words fail to express his indebtedness to his family members father **Mr. Gokulananda Rout**, mother **Ms. Santilata Rout**, mother-in-laws **Ms. Rashimirekha Nayak**, wife **Ms. Rajashree Nayak** (***Shree***), loving brothers **Mr. U. R. Das**, **Mr. S. K. Balabantaray** and sister **Priyadarshine** for their understanding, patience, active cooperation and after all giving their times throughout the course of the doctoral dissertation. The author thanks them for being supportive and caring. His **wife** and **parents relatives** deserve special mention for their inseparable support and prayers.

Last, but not the least, the author thank the one above all, the omnipresent **GOD**, for giving him the strength during the course of this research work.

September, 2016 *Bunil Kumar Balabantaray*

NIT Rourkela Roll Number: 512ID102

Abstract

Robots are widely used for part assembly across manufacturing industries to attain high productivity through automation. The automated mechanical part assembly system contributes a major share in production process. An appropriate vision guided robotic assembly system further minimizes the lead time and improve quality of the end product by suitable object detection methods and robot control strategies. An approach is made for the development of robotic part assembly system with the aid of industrial vision system. This approach is accomplished mainly in three phases. The first phase of research is mainly focused on feature extraction and object detection techniques. A hybrid edge detection method is developed by combining both fuzzy inference rule and wavelet transformation. The performance of this edge detector is quantitatively analysed and compared with widely used edge detectors like Canny, Sobel, Prewitt, mathematical morphology based, Robert, Laplacian of Gaussian and wavelet transformation based. A comparative study is performed for choosing a suitable corner detection method. The corner detection technique used in the study are curvature scale space, Wang-Brady and Harris method. The successful implementation of vision guided robotic system is dependent on the system configuration like eye-in-hand or eye-to-hand. In this configuration, there may be a case that the captured images of the parts is corrupted by geometric transformation such as scaling, rotation, translation and blurring due to camera or robot motion. Considering such issue, an image reconstruction method is proposed by using orthogonal Zernike moment invariants. The suggested method uses a selection process of moment order to reconstruct the affected image. This enables the object detection method efficient. In the second phase, the proposed system is developed by integrating the vision system and robot system. The proposed feature extraction and object detection methods are tested and found efficient for the purpose. In the third stage, robot navigation based on visual feedback are proposed. In the control scheme, general moment invariants, Legendre moment and Zernike moment invariants are used. The selection of best combination of visual features are performed by measuring the hamming distance between all possible combinations of visual features. This results in finding the best combination that makes the image based visual servoing control efficient. An indirect method is employed in determining the moment invariants for Legendre moment and Zernike moment. These moments are used as they are robust to noise. The control laws, based on these three global feature of image, perform efficiently to navigate the robot in the desire environment.

Keywords: object detection; edge detection; corner detection; vision system; moment invariants; visual servoing; Zernike moment; Legendre moment; feature vector; Gaussian filter; fuzzy logic; wavelet transformation.

Contents

List of Figures

List of Tables

Abbreviations

CAD	Computer Aided Design
CCD	Charged Couple Device
CMOS	Complementary Metal Oxide Semiconductor
CSM	Camera-Space Manipulation
CV	Computer Vision
DAQ	Data Acquisition
DoF	Degrees of Freedom
FMS	Flexible Manufacturing System
FOV	Field of View
HVS	Hybrid Visual Servoing
IBVS	Image Based Visual Servoing
IMED	Image Euclidian Distance
IVS	Industrial Vision System
LoG	Laplacian of Gaussian
MM	Mathematical Morphology
MPA	Mechanical Part Assembly
MV	Machine Vision
NN	Nearest Neighbor
PBVS	Position Based Visual Servoing
PCB	Printed Circuit Board
RVS	Robotic Vision System
SE	Structuring Element
SIFT	Scale-Invariant Feature Transform
SURF	Speeded up Robust Features
UAV	Unmanned Aerial Vehicle
VGR	Vision-Guided Robotics
VGRAS	Vision Guided Robotic Assembly System
VS	Visual Servoing
ZM	Zernike Moments

Chapter 1

Introduction

1.1 Industrial Vision System

Machine vision (MV), also called as "industrial vision system" (IVS) or "vision systems" , aims at processing the data acquired by sensors (e.g. camera) which are connected to the hardware and software units for processing data acquired by sensors [Jain et al., 1995; Vernon, 1991; Steger et al., 2008]. IVS is a multi-disciplinary field, covering computer science, optics, mechanical engineering, electronics and industrial automation. Initially, MV systems are focused on manufacturing, which is quickly changing, spreading into medical applications, food processing, research, and even traffic control.

MV is well-defined by The Machine Vision Association of the Society of Manufacturing Engineers and the Automated Imaging Association as the use of components for non-contact sensing to automatically recognize and understand an image of real circumstances in order to acquire information to control a machine or process [Nello, 2000]. It performs the programmed image analysis for the purpose of controlling processes, machines, robot and quality.

IVS employs techniques for extraction of desirable features of images. The methods are used to provide an image-based automatic inspection and investigation for applications like process control, robot guidance and automatic inspection in shop floor [Herakovic, 2010; Malamas et al., 2003]. The scope of IVS is wide, which is related to computer vision (CV). Industrial vision is slightly different from CV, which primarily deals with image processing. IVS are designed for visual assessment and control under challenging industrial applications that require high-speed, high-magnification, and/or repeatable measurements. CV and industrial vision differ in the process of creating and processing images. As such, in industries robots are equipped with IVS, whereas CV is suitable for the robots to work along with human workers. IVS is simpler yet more real-world application, but CV is related to artificial intelligent techniques [http: //www.visiononline. org/vision-resources-details.cfm/vision-resources/Computer-Vision-vs-Machine-ision/content_id /45856-7; Davies, 2012].

Normally, human experts are performing the quality control and inspection by the use of visual features. Human operators perform the task better than machines in several circumstances, but they are slower and get exhausted easily. Further, it is required to develop the skills of the workers to make expert, which is costly and time consuming and also it is difficult to retain them in an industry. There are many cases where, inspection and quality control task are distinct jobs for a best trained expert. In applications such as tracking of desired object and robot guidance, accurate

information must be extracted as fast as possible and used. Inspection is very difficult and dangerous in environments like underwater, industry with nuclear materials, chemical industry etc. IVS can commendably substitute human workers in such challenging environments.

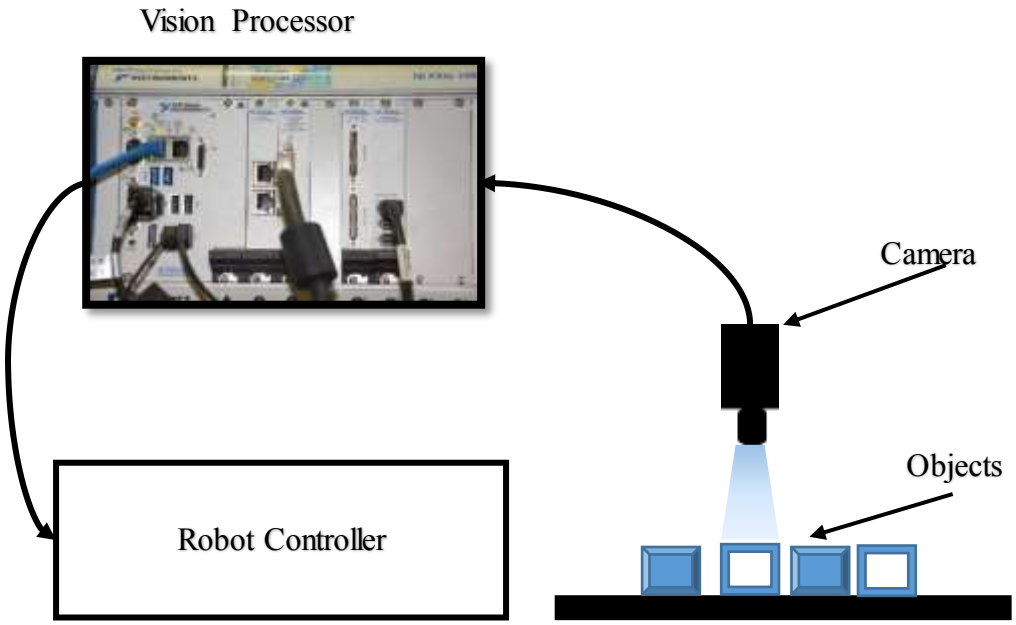


Figure 1.1: A Typical Industrial Vision System [Malamas et al., 2003]

Figure 1.1 represents the system model of an IVS. A typical IVS consists of one or more cameras for capturing images, lighting for illumination of the area of focus, vision hardware such as frame grabber and vision processor, software to process and examine images. First, vision sensors (camera) are used to capture the image of the objects. Cameras are usually positioned at fixed places as in most of the industrial environments. Automation systems are developed to examine only well-known objects available in the work environment. Then the acquired images are processed by a computer system through specially designed image processing software. The outcome of the processing is the desired features of the objects are then fed to manufacturing process control system or robots etc. The entire work environment must be properly illuminated in order to enable the acquiring of the images for processing and analysis. The processing is sometimes take more time or computationally intensive and therefore, exceeds the processing capabilities of the main processor. Hence, application specific hardware like digital signal processors and integrated circuits are used to ease the problem of processing speed. The results of this processing can be used to:

- Control a manufacturing process.
- Provide data/information to other external devices for further processing.
- Detect defective parts.

The functioning of IVS comprises of four important procedures as follows [https://en.wikipedia.org/wiki/Computer_vision].

- Image Acquisition: The vision sensor is used to capture images and converted to digital format and stored in the system. Image acquisition is achieved through careful structuring of the lighting arrangement and proper positioning the cameras to enhance the features of interest.
- Image Processing and Segmentation: The processor uses a set of procedures to increase the quality of element in the image that are significant for inspection. This consists of segregating the desired features from all other parts of the image.
- Feature Extraction Techniques: The feature extraction methods are used to identify and quantify the desired features in the image. Such features are shape, size and invariant features of the object. Feature extraction involves reducing the volume of data that describes the image by using several image processing algorithms like edge detection, corner detection, shape descriptor etc.
- Feature Classification: Classification, or feature recognition, is the process of assigning features that share a common property to a predetermined set of flaws such as sand inclusions or pin holes in a casting.

It is a difficult task to design and develop a general IVS. The reason behind this is application dependent. The elements required for designing a general purpose IVS is as follows:

- Front-end optics for Image Acquisition: The front end consists of the lighting, the lens, and the camera. It is possible to capture the critical features of the object by selecting proper acquisition hardware.
- Frame grabber for Image Conversion: A frame grabber contains an in-built processor that accepts the input as a video stream from the camera, digitizes it, and stores for further analysis. Advanced frame grabbers also speed up the image processing and feature extraction techniques.
- Processor: A processor is just like a computer system that processes the vision tasks.

- Software: Software are used to investigate the image data and perform the tasks and control the overall operation. Normally, the software package includes several image processing and feature detection methods. Some of the techniques are such as:
  - Pixel count: computes the number of light or dark pixels.
  - Thresholding: transforms an image to black and white.
  - Connectivity and Segmentation: used to detect and recognize objects by separating the light and dark connected regions of pixels.
  - Gauging: measures the dimension of the object.
  - Edge Detection: detecting the edge map of the object.
  - Template Matching: finds identical patterns.
  - Pattern Recognition: recognizes the specific object pattern.

A reliable system must have a low "escape rates" (i.e. non-accepted cases reported as accepted) and no 'false alarms' (i.e. accepted cases reported as non-accepted). Therefore, the system adjust to changing scenarios and achieve high performance consistently. It is very difficult to achieve a robust performance. The system can perform is effectively under certain condition like good lighting and low noise environment, appropriate hardware. In summary, an IVS need to be fast in processing the tasks and cost effective.

1.2 Application of Industrial Vision System

The design and development of a good IVS vary with respect to the application environment and are dependent on the associated tasks. In inspection applications, the system must be capable identifying the variations or defects in products within the limit of acceptance. Whereas, guidance, alignment, measurement and verification in assembly operations are performed by users.

It is difficult to have an IVS that can handle each and every tasks associated with the applications. Therefore, the design and development of such system is based on the requirements with respect to the application domain. The main problem in automating an industrial task is to understand the types of information which is needed to be extracted and the processes of interpreting them into measurements or features. In case of identifying the defecting objects, it is necessary to specify, in advance, the meaning of defective in terms of measurements, the set of rules for recognizing the defective objects and algorithms for implementation in software or hardware. A decision is taken carefully on the type of measurements to be performed and on the exact location of the measurements. Vision enabled inspection systems need high speed, high amplification,

uninterrupted operation, and/or repeatability of measurements is often preferred by manufacturers. For example, fabrication in semiconductor manufacturing industries is influenced significantly by vision based inspection, without this the production can be reduced significantly. The vision based inspection technology can inspect the components at high speeds with precision and accuracy. Therefore, the application of IVS is broad.

Application of IVS is in almost all modern industries. MV systems can be used in various range of industries, some of these industries are: [http://www.kuviovision.com /machine-vision-industry; Labudzki and Legutko, 2011; Batchelor, 2012],

- Chip manufacturing
- Pharmaceutical
- Packaging
- General Manufacturing
- Food Processing
- Solar Production
- Defense
- Assembly

MV system is used with robots for identifying the objects. This system is used normally to carry out the inspection task where industrial robots are not used. A MV system is attached to a high speed production line for the acceptance of the parts involved in the work. The parts, either accepted or rejected, are separately moved to different storage places. Applications like assembly, inspection and part handling are currently being used in many industries. Other applications can make use of this system for surveillance, determining the accurate position, tracking, and recognition the desired part. The applications for MV system varies generally dependent on the type of task required to accomplished in industry and production environment, some typical applications include:

- Inspection of assembled product in different phases of assembly
- Inspection of labels or bar codes on products
- Defect detection
- Robot navigation and guidance
- Determining orientation of components
- Traceability of manufactured products

- Packaging Inspection
- Checking laser marks and cuts
- Checking of food pack

1.3 Robot Vision System

No industry has been more transformed by vision-guided robotics (VGR) than manufacturing. The earliest robots were designed for simple pick-and-place operation. Now, with technological advances in sensors, computing power, and imaging hardware, vision-guided robots are much more functional and greatly improve product quality, throughput, and operational efficiency.

Over the years vision based technology has been developed becoming reliable and efficient with a reduced pricing. It helps in enhancing the value of the system by improving the throughput and quality of the product with a higher flexibility. Vision guided robots can localize accurate parts, determine the correct grasping points and inspect the assembled final product. Robotic Vision System (RVS) [Nayar, 1990] makes available a platform with vision guidance to robots for industrial applications like packaging, inspection, manufacturing and defect detection etc. The application of RVS also includes Medical Industries, Agriculture and Food Processing Industries, Pharmaceutical Industries and Manufacturing Industries. The possibilities are endless.

1.3.1 Medical Application

The Medical industry are advancing from the growing use of advanced digital image acquisition and processing with cameras being used widely within medical [Dario et al., 1996], microscopy, and life science applications. Medical technologies based on sensors can assist in diagnosis, premature detection of disease, treatment, operation and training. The range of applications in medical continues to increase due to the creativity in developments in vision technology such as mini-cameras for endoscopy or insignificantly intrusive surgery, industrial cameras to analyze the retina in ophthalmology, analysis in complex orthopaedic related disease. One of the common One of the important application of the MV system is the diagnosis through non-contact optical tracking, which is related to human-machine interface. The use of MV system provides solution to the medical industries with a better speed, improved precision and stability factor.

Finally, it is important to note that medical technology contains an extensive variety of various imaging gadgets like microscopes, endoscopes camera, X-ray units, CRT, MRI units, and cameras to monitor the operation theatre. These gadgets are used with different interfaces, standards and control strategies.

1.3.2 Pharmaceutical Industries

In the pharmaceutical industry, it is critically important to count tablets and gel capsules before filling them into containers. To meet those demands, such industries developed a range of electronic counting and packaging systems that can be deployed either as stand-alone units or as part of an integrated packaging line. The industry moved to replace such old system by IVS to develop a vision-based inspection system for enhancement in the capability of their existing product range, which can detect partly formed and damaged tablets and capsules [Hamilton, 2003; Chauhan et al., 2011].

The pharmaceutical industry was one of the first to use color-based MV system. This was driven by the need to verify that the correct tablets were in the specific location in the package. The colors that were used were challenging for MV systems because they were often subtle intensity variations of the same hue. The following applications in the pharmaceutical industry are addressed properly by IVS.

- Tablet tracking and inspection
- Defect detection
- Cap or Seal Inspection
- Powder or Liquid fill inspection
- Seal surface inspection
- Component inspection (Stoppers / Caps / Glass)
- Final product inspection
- Box / Label inspection and many more

1.3.3 Food Processing Industries

Now-a-days the food processing industries are making automation with the aid of RVS [Gray et al., 2013; Cubero et al., 2011] which has been an essential requirement. It is more than an option to improve the quality of the products, reduce the processing. The type of IVS to perform inspection tasks in food processing industries depends on several factors. IVS based on line-scan camera are most commonly used for inspection and sorting of food materials such as Mango, Cabbage, Carrot and Tomato etc. Another demanding application of MV system in food processing industry is to determine the harvested or baked food materials based on color by using advanced cameras like multi spectral or area-array camera.

Efficient inspection and sorting systems are required to produce higher-grade vegetable.

The "Tegra system" from "Key Technology" sorts the vegetables by its shape and size. It sorts peas only by color because they are generally of consistent size. They are also graded and separated according to specific gravity prior to inspection. The Cayman® BioPrint® sorter [http://www.key.net/products/cayman] is used to detect the distinct biological characteristics of objects. This helps in sorting of the product accurately when the good products are similar to defected or external materials in color, shape and size

"Iris Vision" from "Stemmer Imaging" [http://www.stemmer-imaging.pl/en/applications/at-the-cutting-edge-of-carrots/] is used to inspect Carrots during harvest. Carrots can be in any orientation on the conveyor belt. So, the angle of rotation with respect to the conveyor belt is determined first. It is required to find out the green part from which carrot is to be cut off. The cutting point of the Carrot is determined by assessing the thickness on both the sides. Similarly, IVS is used to guarantee the packaging of correct potion of meat or fish.

1.3.4 Assembly

The use of MV system provides a platform to manufacturing industries dealing with assembly of different parts. Manufacturing and automobile industries like Mahindra, Ford, TATA, and Honda etc. are employing the RVS for assembly of parts to produce final product [Kress, 2004; Tsugawa, 1994].

Now-a-days economy is motivated more towards production of customized products, rather than concentrating on mass manufacturing. The economy is greatly dependent on use of control strategies and operating procedures for assembly in unstructured environments where parts are randomly placed. The robotic assembly can be built to increase the cycle time and throughput by

integrating the MV system with the robot. The visual information helps in guiding the robot to do the task autonomously. Such a setup is called vision guided robotic assembly system (VGRAS) [Sternberg, 1985; Herakovic, 2010]. Considering the competition in market, the robotic system integrated with vision system will be widely used in the industrial assembly environment (as shown in table 1.1). This take care of mass customization in smaller industries with short production time and modular products [Kellett, 2009].

Table 1.1: Robots in Manufacturing Environment: Today versus the Future [Kellett, 2009]

Robots in Manufacturing	Current	Forthcoming
	Mass production	Mass Customization
Production duration	Long	Short
Lot volume	Large	Small
Configuration of manufacturing goods	Non-modular	Modular
Features of goods	Narrow variety	Numerous variety

The need for consistent and flexible operation with a minimum possibility of downtime and expected demand in many industrial assembly environments, where, the VGR systems are becoming more popular. In the past, robots are used in assembly operations to pick up parts from a specific location every time and robot will fail to do so, if the part is even slightly misplaced. This problem can be significantly reduced when robots are integrated with IVS and other necessary sensors. The integrated systems are used in helping the robot to adjust itself by finding out the exact location of the part. The progress in vision technology makes the robot adjust in the changing environment which allows the robot to efficiently execute the tasks that is inconsistent. Presently, the research in area of vision guided robotic application is growing and growth is visible globally. Vision guided robotic assembly operation can be realised in mechanical part assembly, electronics part assembly and micro assembly etc.

A. Mechanical Part Assembly

Mechanical part assembly (MPA) [Wesley et al., 1980] is the most important part of the manufacturing process. It is defined as an ordered sequence of controlling tasks where separate parts are assembled and combined to yield a product. The final product is manufactured by means of adding distinct parts at each station while traveling down the assembly line. In some cases, the conveyor belt stops at each station, whereas in other set-ups, the conveyor belt moves continuously. In MPA, the design engineers provides a logical order of assembly sequence that can produce the final product by ensuring the effectiveness of the assembly.

Mechanical part assembly depends more on automatic part handling and insertion. The automatic part handling is the principal concern of the process and follows certain efficient rules to easily separate out the desired part from the bulk. In order to achieve this, the parts must be transported along the track through feeder with proper orientation suitable for robots.

The benefits of using MPA is that the parts can be disassembled and reassembled without difficulty in remote locations. The cost of production is reduced in this type of assembly as the parts are created off site and final product is produced in the factory. The upgradation and repairing cost is also less as compared to other type of assembly.

B. Electronics Part Assembly

The main issue in producing the electronic product through assembly is that the components must be precisely align and tight tolerance. Traditionally, robotic automation with teach/repeat is high-priced as it utilizes expensive fixtures and time-consuming setup. The issues with VGRAS for electronic part assembly are the grasping error and visually inaccessible assembly features. The vision guidance requires the proper selection of such visual features and the design of control laws for robot control.

Such assembly setup uses advanced technologies with state-of-the-art sensors which produce quick, reliable and economic products with a reduction of error rates. Sensor technology has proven itself in all areas of the electronics industry from chip production to complex assembly. In such industries, identification, inspection monitoring, and grasping is a challenging task. To achieve these tasks, intelligent visual servoing methods and sensors like vision sensors, ultrasonic sensors, force and torque sensors etc. are used. The pick-and-place task involves in processing very small components. A stand-alone solution, the Inspector I40 vision sensor [https://www.sick.com /in/en/product-portfolio/vision/2d-vision/inspector/c/g114860] is more efficient and cost-effective

for doing inspection, positioning and measuring task. During final inspection of circuit boards, smart camera ensures that even the smallest parts are mounted correctly. For example, the presence of seals can be checked and screws can be measured if they are screwed correctly in the housing. Errors are indicated in real time.

Considering design of printed circuit board (PCB), the plain board is fixed with several components to produce a functional board. In through-hole technology [Boggs et al., 2004], component leads are inserted in holes. In surface-mount technology [Page, 2005], the components are glued on pads or lands on the surfaces of the PCB. In both, component leads are then mechanically fixed and electrically connected to the board by soldering. Vision is used to locate specific spot on PCB and do testing and inspection through non-tactile sensing. Robot controllers track electronic devices as they move through the production line. As electronic devices or components become more sophisticated, the need for a more advanced assembly process naturally follows. VGRAS is the best fit for such application as level of sophistication increases day by day. The use of sensory input such as vision has created intelligent robotic solutions.

C. Micro Assembly

The assembly process is the key regulating factor in micro machine technology. Manual assembly is costly and requires a high level of precision which is not practically possible for the workers due to the eye strain and dexterity limitations associated with assembling such miniature parts under a microscope. Such assembly cell requires state-of-the-art motion control equipment including a robotic work cell with nanometer scale resolution and real-time vision system to control servo mechanisms and motors for the alignment and assembly of parts with sub-micron tolerances.

The basic requirement of an automated micro assembly is that it must be able to transport microscale parts and manipulate them so that precise spatial relation with microscale tolerances can be established for die alignment, parts insertion, and for certain packaging processes such as die bonding, device sealing, etc. It is always a challenge task in handling the micro components. It requires high-speed automation capable of capturing such minute parts; in-line metrology; vision systems to ensure recognizing the correct part, required orientation and positional accuracy of these components to one another to sub-micron tolerances. The problem of robotic micro assembly has been explored using high precision actuators and vision feedback. The equipment required to perform micro assembly include the followings [Gauthier and Régnier, 2011]:

- MV system equipped with stereo microscopes and high resolution camera.
- Micro positioner, micro gripper manipulation and position control.
- Advanced clean room.

1.3.5 Benefits of Vision Guided Robotic System

The vision guided robotic system can adjust to the changing scenarios in the assembly environments with a little change in mechanical set up. In most of the cases, the gripper is changed with respect to the parts involved in the assembly. A change in program is the main requirement in handing the different types of components. Such programs make it possible to handle the parts differs in geometry with random orientation. This results in quick changeover times. Some of other benefits of vision based robotics are:

- Automates manual steps and processes
- One camera can replace multiple sensors
- Speeds of the production and increases the throughput
- Greatly reduces programming effort to guide robot
- Increases flexibility of the robot
- Recipe-driven production reduces changeover times
- Allows for less precise positioning of parts
- Use of less expensive and more precise robots
- Eliminates expensive fixtures
- Allows a robot to perform multiple tasks

1.3.6 Limitation of Vision Guided Robotic System

The main drawbacks of vision guided robotic system is the proper integration of vision system with the robot controller. This is due to the lack of knowledge of the developer. The setup is virtually implemented to dummy assembly environments for gathering the knowledge of system parameters that plays important role in making the system. This process of training is time consuming and expensive. Therefore, the setup needs proper training. In industrial environments, the image acquisition and proper lighting of environment are other major issues which makes it difficult to develop such system.

1.4 Mechanical Part Assembly

The use of VGR is guided by the demands for improvement in automation and flexibility in manufacturing industries. Such assembly system requires a clear understanding of the capabilities and methods used for VGR, and a thorough analysis of the parts and application environment [Xiong et al., 2002; Chawla and Deb, 2012]. The development of vision guided robotic system is the significant factor in designing a MPA.

Manufacturing a product through assembly is a complex task. Greater precision is required as components get smaller as well as higher throughput and greater flexibility are essential for reasonable advantage. Manual assembly cannot stand up to the current requirements in manufacturing industries. Many manufacturing tasks would simply be impossible, complex and time consuming without industrial robots.

Integration of suitable MV system with the robot results in higher quality products. In such case, the robot navigation is fast with higher accuracy. Parts with different shape and size can be handled easily at the same assembly line by doing a simple change in programming.

Developments in part recognition methods and support for 3D data enables the user to implement applications like bin picking [Sansoni et al., 2014], 3D pose determination [Zhu et al., 2014; Ulrich et al., 2009], 3D inspection and quality control. For years, vision-guided robots have been a mainstay for manufacturing and assembly tasks such as inspecting and sorting parts. These operations tend to be carried out in highly constrained, structured environments with zero obstacles. However, advances in processing power and sensor technologies are now enabling robots to undertake more unstructured tasks which require better vision systems to identify and avoid obstacles.

The use of time-of-flight sensors and technologies that process 3D images makes the vision guided robotic system an efficient one [http://www.roboticstomorrow.com/article/2015/06/ vision-guided-robotics-advances/6241?]. The inline inspection finds new possibilities due to the advanced imaging technology. The scanning speed of the advanced CMOS sensor can be achieved up to several thousands of high resolutions 3D profiles per second. The main advantage of structured light and stereo 3D cameras over sheet-of-light triangulation camera is that the relative motion between the sensor and the part is not required. This permits fast generation of 3D point clouds with appropriate accuracy for good scene understanding and robot control. The growth in

sensor technology and development of efficient algorithms guides in implementing vision-guided robotics tasks efficiently.

A MPA system with VGR system can be able to performing the following important tasks:

- Part recognition and inspection before and after assembly
- Stable grasping
- Visual servoing and navigation

Part/Object Recognition:

The vision system can be gainfully used for recognition and classification of parts along with inspection. The parts are recognized with respect to the shape and size from the image captured by IVS through image processing techniques. It is also required to determine the exact position of the target object for robot to grasp. A robot takes decision and performs action by following a set of sequential tasks.

Recognizing desired object can be termed as a labelling problem based on known object models. For example, in an automated assembly environment, the conveyor belt may contain one or more similar or different objects. To recognize the object of interest, a set of labels corresponding to known object models are to be set and the system must be capable of assigning the correct labels to the object or region in the image taken by system. In the process of recognizing objects, feature extraction methods plays an important role. There are several feature extraction techniques like edge map, corner point detection, color histograms, mathematical morphology, moment based and connected components labelling etc. discussed in the subsequent chapters. These features are used to compare with the features of the stored object models. The object recognition is fully depended on the efficiency of the feature extraction methods.

Stable Grasping:

Robots are manipulating a variety of objects in an assembly work-cell. A physical connection bet end-effector and objects is developed for manipulation of objects. Effectiveness of the grasping technique depends on grasp planning algorithms and the shape of the object. Stable grasping is to transport an object form one place to another. The grasp planning algorithms have made remarkable advancement in resulting stable robotic grasping.

The correct manipulation does not guarantee the success of stable grasping. Grasping of a "Mug" is possible from side or top-down. The both way helps in carrying it to other place. But for pouring

water out of the "Mug", the suitable grasping position is grasping from side. Such constraint in grasping is called semantic constraint.

Inspection:

MV system is used to support the robots to perform the inspection tasks before and after assembly. During the initial stage of assembly, vision system is used to recognize the part with respect to either shape or size. This checks the finishing of surface, dimension of the product, and labelling errors etc. The vision based inspection requires less time with reduced errors. In manual inspection by the workers requires more time and there is a possibility of erroneous result.

Visual servoing:

Visual servoing (VS) [Chaumette and Hutchinson, 2006; Chaumette and Hutchinson, 2007], also known as vision-based robot control, is a technique which is used to control the motion of a robot by using feedback information extracted from a vision sensor. Tracking of the part from feeder to final product is based on VS techniques. These techniques help in feeding information to robot controller about the part shape, size and its exact position in the work cell. These techniques also help in stable grasping of parts. The servoing method is also used for collision free robot navigation.

Primarily, VS is of two types. The degrees of freedom (DoF) of the robot can be controlled by using visual features of the object, referred to as Image Based Visual Servoing (IBVS) [Corke and Hutchinson, 2001; Mahony and Hamel, 2005]. Other approach of controlling the robot is performed by pose estimation of the desired object. This is called Position Based Visual Servoing (PBVS) [Wilson et al., 1996; Thuilot et al., 2002]. These two types of visual servoing techniques depend on correct feature vectors which is extracted through feature extraction techniques.

1.5 Motivation

Human beings are the highest level of autonomous systems because they can think and they can change plan at any moment due to their high level of intelligence. On the other hand, robots cannot reach the same high level of autonomy as humans as they are programmed to do certain tasks according to certain factors which are completely programmed by human beings. Usually they have no possibilities to change plan like humans or plan new things unless the programmer programs them to change the plan. However, due to technological development in machines, sensors, actuator, and digital electronics and microprocessor technology it is possible to create an autonomous robot.

Precision and quality control are the important issues in manufacturing industries. Robot in an industrial setting executes a wide range of tasks such as pick-and-place, welding, fault detection, and spray painting etc. The commom aspect of these tasks includes identification, measurement and inspection ensuring certain accuracy and repeatability with respect to a certain position or object.

Visual guidance is a rapidly developing approach to the control of robot manipulators, which is based on the visual perception of a robot. The information collected through visual perception is used for part recognition, part grasping, part orientation and visual servoing etc. These actions can be visual guidance which involves the use of one or more cameras and a machine vision system to control the position of the robot's end-effector relative to the work-piece as required by the task.

For years, automatic machines have been efficiently performing labor-saving work and repeating tasks such as churning out millions of one type of part. Unfortunately, automatic machines are not suited to small batch work since physical changes must be made to the machine for it to perform a different task. This inflexibility of automatic machines has led to the term "hard" automation. Robots, on the other hand, are flexible and can be easily switched to perform different jobs by simply changing the robot control program. Robots follow "soft" automation and make the "flexible manufacturing system (FMS)" a reality. Soft automation is economically beneficial for situations with high manufacturing overheads and skilled labor costs.

The challenges in product assembly environment mostly remain in how to precisely align the components for tight tolerance assemblies in quickly reconfigurable production. Traditional automation or teach/repeat robotic automation is often prohibitive because of costly fixtures and time-consuming setup and ramp-up. The conventional methods of VGR in assembly automation solutions cannot perform the assemblies with the grasp error and visually inaccessible assembly features. A correct approach using vision-guided robotic part assembly method in conjunction with vision system can make the system efficient, convenient and economical.

Assembly robots have expanded the possibilities in the world of automation. Robotic assembly technology has found many applications in both industrial and research fields, including semiconductor manufacturing, hybrid MEMS device automation, automated nanoassembly, flexible parts assembly. Much progress has been made towards the automation of robotic assembly tasks over the years. However, enormous numbers of tasks are still performed by human

assemblers. In many cases, human operators performing these assembly tasks run the risk of injury due to accidents or repetitive stress. It is, therefore, desirable to automate these assembly processes. The research community, since early days to till date are attracted towards the design and development of robotic assembly system with the aid of vision system, not only as an exciting technical problem itself but perhaps because of challenges faced by the automation industrial system and the inherent interest in advanced technology. It is observed from the study that several important projects have been launched, and important examples of VGRAS have been developed. Still, the present situation is that reliable and flexible systems are yet not available for real time applications. It is also easy to foresee that in future a consistent research activity is bound to happen in this fascinating field, with design, development and analysis of such system at the technological (sensor, actuator etc.), methodological (control, planning etc.) and application (workspace, assembly environment etc.) levels. Important connections with other scientific fields, such as Human Robot Interaction [Mahony and Hamel, 2005], are also expected.

Summarizing, in present scenario, all industries are opting for robotic based automated or semi-automated systems to meet the production quality with cometitive cost. Therefore, increasing the level of intelligence and thereby increasing the level of autonomy in reprogrammable devices like robots through a fusion of robot technology and sensor technology is a challenging task before the engineers.

1.6 Broad Objective

Having studied the importance of vision guided robotic systems, the objective of the proposed work is to design and develop a robotic assembly system with the aid of vision system and necessary algorithms to carry out part assembly process autonomously.

The developed system should be capable of capturing the information about the availability of parts in the work area. In the work area, there may be several parts present, out of which the system must be able to identify the correct part required for assembly operation. The proposed system should be capable of identifying the parts through extraction of features like shape, size and orientation etc.

Once the system is capable of identifying the correct part, then the part is to be grasped by the robot gripper. Methodology is to be developed for a stable grasping by determining the position and points with a desired manipulation or orientation as per the requirement of the assembly task.

The robotic assembly environment comprises of different obstacles like other parts and accessories etc. In such case, navigating the manipulator is a challenging task. Hence, methodologies are to be developed for navigating the robot manipulator in the environment of multiple obstacles to avoid collisions.

Finally, the end product is to be inspected for completeness. Inspection procedure is to be developed for inspecting the product after assembly.

In order to achieve these objectives the following tasks are suggested.

- Selection of typical application areas or environments
- Selection/determination of appropriate vision system (type and specification)
- Selection of other additional sensors as per the requirement of the tasks and assembly environment.
- Integration of the sensors and their feedback system with the robot control system.
- Development of algorithms to carry out tasks in real time.

1.7 Methodology

The work is planned to be carried out in the following five phases:

Phase –I: Review of literature: Consider various types of multi-fingered robotic hands, resemblance with human anatomic structure. Analyze the reviewed literature based on different aspects like structure, control, application, grasping etc. of the available hand and list out the important outcomes and shortfalls.

Phase –II: Development of Vision System.

- Selection of right types of CMOS, monochrome camera.
- Selection of frame grabber with sufficient memory and compatible software on a personal computer.
- Development of algorithm for part detection, identification and robot navigation.

Phase –III: Development of feature extraction methods for object recognition.

Phase –IV: Design and Development of the system by integrating the robotic system with vision system.

Phase –V: Development of algorithms for the robot navigation based on visual information.

The aforementioned tasks will be designed theoretically, modelled and simulated virtually using appropriate tools so far as the integration and control techniques are concerned, and finally the developed system will be implemented through hardware to realize the system to carry out the desired tasks.

The developed system which can guide to achieve the automated assembly process must ensure the followings.

- Identifying of correct component.
- Finding out the correct position and point of grasping.
- Determining the distance of the gripper from the grasping position.
- Navigating the robot within an unstructured environment with the help of visual feature feedback based control of the robot.
- Inspecting the individual parts and the final product for their correctness.

1.8 Organization of the Thesis

The present chapter 1, is the introduction chapter gives brief idea about the VGRAS and its application in various field. Apart from this introduction chapter, the thesis organized as follows:

Chapter 2 provides a review of literature based on different aspects of designing and development of VGRAS such as vision sensor based robot navigation, feature extraction and identification of parts and the most important is visual servo control of robot. The objective of the research work is also defined and presented based on the analysis of the review of literature.

Chapter 3 presents system requirement specifications for designing of vision guided robotic system for part assembly like hardware and software requirements. The hardware components are vision sensor, vision processor, frame grabber and robotic system etc. The software required for integrating the hardwares are LabVIEW, MATLAB and Digi-Matrix Kawasaki Robot library and AS language.

Chapter 4 discusses about the feature extraction and object identification methods like edge map based, interest point calculation based, moment based are implemented. Comparison study is made between the proposed methods and the popular existing state-of-art methods. The implementation is done in MATLAB and LabVIEW environment. Object detection based on classification using Nearest Neighbor method is implemented.

Chapter 5 describes the experimental setup which includes vision sensors, machine vision and robot controller. In this chapter, the calibration methods for vision sensors are present along with the installation details of the vision processor and robot controller. It also includes the methods for integrating robot controller with all other hardware components.

Chapter 6 presents methods of visual servo control along with the mathematical model. The image based visual servo control (IBVS) is implemented based on moment invariants like Legendre and Zernike moment invariants. The moment invariants are calculated by the proposed indirect method of calculation. The results obtained are presented with analysis.

Chapter 7 concludes the present dissertation with a summary of the contribution and the scope for future work.

1.9 Summary

In this chapter, the general overview of the robotic assembly system is presented. The MV system, elements of MV system and its broad application with respect to the scenario where robots are a part are presented, as per need and technological development. This chapter also contains the description of VGRAS along with limitations and benefits. The applications of VGRAS in manufacturing industries are described. The motivation towards the presentation of thesis is clearly mentioned in this chapter. It also includes the broad objectives and the methodologies to be followed to achieve the objectives.

Chapter 2

REVIEW OF LITERATURE

2.1 Overview

Vision system provides an eye to the robot with which it can recognize, determine the position of the parts and adjust corresponding working sequences in an uncertain environment. The desired intricate steps in designing vision guided robotic system is highly hinged on manipulating or processing the vision information. With growing demand in industrial environment the need is to make it more efficient and faster. However, the object recognition and visual servoing methods involved in designing vision systems are sophisticated, specifically automated part assembly systems.

One of the goals of machine vision researchers is to give humanlike visual capabilities to robot so that it can able to sense the environment in their field of view, understand the sensed data, and proceeds with appropriate actions through programming. Spatial vision with sensors works agreeing to the same principle as that of human eye. Vision includes the physical components of illumination, image formation, reflectivity and geometry as well as the intelligence aspects of understanding and recognition. In short, there are many aspects of machine vision that provide enough opportunities to researchers to work on. A good amount of research work has been carried out by large number of researchers to explore the benefit of machine vision in conjunction with advanced technological tools. Nevertheless, there remain many grey areas that need to be explored and utilized.

In the survey, feature extraction and part recognition techniques are emphasized more. The survey also includes robot navigation with the aid of vision sensor and visual servo control methods. Important attributes of an intelligent vision system are recognition of parts, inspection, and robot control strategies. In order to design and develop a system that preserves these aspects it is important to clearly define its required outputs and the available inputs.

In the following sections, some of the major research literatures are compiled to understand the importance of the subject as well as to figure out the necessity of further research in the area.

2.2 Measurement and Identification of Parts Using Machine Vision

In part assembly application, part identification is an important task. This is fully depended on the efficiency of feature extraction techniques and the algorithms involved in measurement of parts or objects. Accurate measurement is helpful in precise identification of parts and its grasping point determination. The intelligent vision system is considered to be the typical system for identification and position recognition. It provides information about the part, its position and orientation that are placed on a specific plane. In this survey, feature like edge, corner and shape is considered. Several researchers have contributed towards the design of algorithms for edge map detection, corner point detection along with shape. Detection of edges is a set of mathematical methods of finding pixels in an image at which the image brightness changes sharply or has discontinuities. In machine vision and image processing, edge detection is a process which concerns with the measurement of significant variations of the grey level images and the identification of features that originated this. This information is useful for application in image enhancement, recognition and registration etc. The techniques involved in edge detection are differentiation and smoothing where smoothing yields a loss of information and differentiation is an ill-conditioned problem. To design a general edge detection method for all context is a difficult task. The widely used edge detector, proposed by Canny (1986), is a computational approach to edge detection which is using an adaptive thresholding technique with hysteresis to remove streaking of edge contours. The threshold value is set with respect to the noise present. Using feature synthesis the operator finds the edge at different scales.

Various shape features like shape invariants, shape signature, curvature, moments, spectral feature and shape context etc. have been used in object identification process. The performance of shape features are evaluated based on the level of accuracy in retrieving shapes in object image. It is not sufficient to evaluate a representation technique only by the effectiveness of the features used. Some of the important literatures are studied in this chapter.

A knowledge based vision system to identify the overlapping objects was developed by Ming-chien and Peng-fei (1988). They developed a geometric knowledge base comprising of shape information about the objects which is translation and rotation invariant and plays an important role in partially occluded object identification. They also proposed an attributed string matching with merge method for identification of objects, if the object feature degrades.

Yamaguchi and Nakajima (1990) proposed a shape identification system for 3D objects using fiber grating vision sensor. They used space encoding and pattern transformation technique to identify the shape of the 3D object. The pattern transformation technique is used on the spot projection images produced by the sensor. A multi-layered neural network used in the proposed system for producing a rotation and translation invariant output which is trained by transformed pattern. The developed system is capable of recognizing complex irregular objects.

Sardy et al. (1993) developed a vision based identification and defect detection of woven fabrics samples in Textile industry. Artificial neural network is used in the proposed system. The textural features of the woven fabrics is extracted by neighboring grey level dependence matrix and grey level run length matrix. The extracted textural features are used as input to the neural network. The neural network is trained by backpropagation method.

Kim et al. (2001) presented an algorithm for parts measurement in a flexible part assembly system. They also proposed error-corrected algorithm that determines the misalignment between parts and its mating hole. The assembly work is performed by compensating the misalignment based on measure information. The algorithm can be extended by considering more parameters of assembly.

Rousseau et al. (2001) developed an efficient and non-contact measurement technique to identify the kinematic parameters of a 6 degrees of freedom (DoF) industrial robot. A single camera is attached with the robot gripper to measure the position and orientation of a passive target in the robot. The accuracy of the measurement is enhanced by subpixel interpolation method. The critical factor, that affects the measurement of parameters, is performed independently.

Malamas et al. (2003) presented a detailed survey on application of industrial vision system. They identified the features of the product by which it is inspected. Issues and approaches in designing industrial vision systems are pointed out and discussed.

Powell et al. (2003) proposed a detection and identification technique for Sardine eggs at sea by using machine vision system. They used the real-time flow imaging and classification system which is an image acquisition and processing system comprising of line scan camera and other hardware for image processing. In this approach, shape, size and shading features are computed for the eggs. These features are then processed through a classification and regression tree algorithm which yields a decision tree that accurately classify the objects.

An advanced pose estimation and feature extraction technique is proposed by Jin et al. (2003) for identifying the obstacle and robot position information for determining the navigation plan for the robot.

Karagiannis and Astolfi (2005) proposed a method of identification in perspective vision system based on a nonlinear reduced-order observer. The proposed approach identify the position of the moving object in 3D space through a single camera. The identifier simplifies in resulting an asymptotic estimation of the coordinates of the object. The proposed solution is robust to noise.

Jerbic et al. (2005) proposed a machine vision based autonomous robotic assembly system capable of working in a complex environment. They used sobel edge detection along with border tracking algorithm for generating the object contours. Then the proposed recognition algorithm is applied to produce a hypothesis about the object in the acquired image. This object contours is then compared with the contours created from CAD model of data. This technique yields all the instances of the object position and orientation information for the robot to act.

Jusoh (2006) addressed the issues related to localization of targets using single color vision system. Color thresholding technique is used for targets detection. The author selects the threshold value by considering the mean of image samples and standard deviation of the color components. The correctness of the detection is verified by the features like diameter, x-y ratio and area.

Munich et al. (2006) proposed a visual pattern recognition approach for autonomous robotic application. They used visual pattern recognition system developed by Evolution Robotics to recognize the objects which is invariant to rotation, scale and transformation. This system is also capable of handling the issues occurred due to lighting, occlusion.

Pena-Cabrera and Lopez-Juarez (2006) proposed a method of recognition and POSE estimation for the components used in assembly in a distributed manufacturing system which is invariant to scale, rotation and orientation within the work space. They demonstrated the issues related to image processing technique, centroid and perimeter calculation. The grasping and coordinate information of objects are given input to the robot in real time. The POSE estimation is performed based on visual feedback. The author implemented the proposed method in a manufacturing cell. The correct recognition of assembly workpieces is successfully developed by a Fuzzy ARTMAP neural network model.

Object identification and image segmentation based on Hue-Saturation-Value color mode, clustering algorithm is proposed by Peng et al. (2006). The pose estimation of the desired objects is performed by CCD vision sensor.

Part identification method is proposed by Moghadam et al. (2008) by the measurement of parts through a stereo vision camera system. A 2D laser range finder is integrated with vision sensor for obstacle detection and ground plane detection. 2D cost map is generated from both 3D world model and laser range finder. The two 2D cost map are combined for accurate detection of object as the two sensors may detect different parts of an object. Then an occupancy grid map is used for obstacle avoidance.

Gomes et al. (2008) proposed a theoretical real time vision for robotics by using abstract visual data. They presented a mathematical model with a small computational complexity for a moving fovea as multi-resolution representation is used for input images. The mathematical model described is useful for a robot to perform physical motion with a feedback from camera to reach to a desired position.

Gao and Jin (2009) proposed a vision based postal envelop identification system. The system is capable of recognizing the postal address and postcode on envelops moving in a conveyor with a high speed camera. The character recognition rate of this system is 98.72%. The proposed system also store the captured envelop images and the recognition results in real time which help in further tracking.

Identification of pieces in chess system with dual-robot coordination based on vision was developed by Yubo et al. (2010). They used the vision system output to coordinate and navigate two robot for playing the game of chess. The authors used character segmentation method with the help of back propagation neural network to identify objects and targets.

Fujimura et al. (2011) presented a study on development of vision based shack wave detection in a saturated traffic to prevent accidents. The authors proposed a vehicle infrastructure integration system in which the driver is given with a feedback information about the shock wave and the error in propagation is examined by vision sensor network. The proposed future scope of this work is to examine the tolerance with respect to human-factor engineering.

Fernandes et al. (2011) described different aspects in designing autonomous visual inspection system and integration procedure of vision system in industrial process. Several constraints that affects the designing the machine vision aided inspection system is presented by the author by

considering the dynamic behavior of industrial processes. This variability nature can be controlled by adding adaptability behavior to the system.

Mustafah et al. (2012), proposed identify the object by measuring the distance and size of the object with the help of stereo vision system for identification. The system developed is intended to perform object detection and identification on the stereo images by blob extraction and distance and size measurement of the desired object. The system also uses a fast algorithm to realize in a real-time environment.

A vision guidance system for autonomous assembly is proposed by Lee et al. (2012) in 3D. In the developed system, surface patches like planar, cylindrical, conic and spherical, are identified from 3D point clouds. The recognition of patch primitives is performed by geometric surface patch segmentation approach based on Hough transformation. The surface patches are extracted from automotive CAD model in DXF format. The resulting CAD models are decomposed into simple entities like planes and cylinders. These CAD data are used for identification, pose estimation and handling parts by a two arm robot. The accuracy and runtime of the developed system is under standard operational range. They specified the future scope of the work for improvements to an advanced and pose estimation under borderline image quality conditions. The runtime can be reduced as per the requirement in main assembly line throughput. The detection technique can be extended for recognition of objects in number and type.

Murai et al. (2012) proposed a visual inspection system to inspect the drugs. The developed system identify the drugs by hierarchical identification technique. A higher recognition rate is achieved as the result is compared with several kinds of capsules and tablets. In this work, one-dose package drugs with wrapping film and overlapping with other are not considered, which can be studied further.

Hasan and AL Mamun (2012) described the identification and tracking of an object by a robot with the help of a vision sensor. The objective of proposed system is to automatically track a colored object by a robot through a webcam. Further improvement can be done to this work by using sensors like sonar and infrared.

An automatic identification system based on MV system for seedling transplanter presented by Tong et al. (2012). Here, the MV system is used to identify the state of seedlings in the plug tray i.e., detect the empty and unhealthy seedle and guide the manipulator to uproot the seedling. Bolb analysis and morphological operator is used to eliminate the noise present in the captured image.

The identification rate is 96.2% is achieved for tomato seedlings samples and can be used for others.

Jestin et al. (2012) proposed an automatic feature extraction technique for retina images. The developed system is based on automatic diagnosis of Diabetic retinopathy through extraction of static features from retina images. The static features considered in this work are mean, standard deviation, variance, entropy, contrast, correlation, energy and homogeneity etc. The extracted features is used for effective classification of retinal disease. Real-time use of the proposed system can done in future to measure the significance of the technique.

Wang and Yagi. (2012) proposed a location recognition technique based on visual features. Use of one of the features like edge, interest point and color feature is useful for certain environment. But in this work, the author have used three features by integrating the features like edge, interest point and color feature to have a robust location recognition technique. Harris function is used in feature detection and partition of regions. The developed recognition technique is robust to factors like viewpoint changes, partial occlusions, and partial illumination changes. This location recognition can be extended for object recognition in scenes.

Tanoto et al. (2012) presented a flexible and scalable software architecture for a vision based multi robot tracking system. The proposed system produce the position information of robots individually. The system is designed by considering the flexibility and scalability nature, so that, it can be incorporated in other similar systems and can adopt to changes in technology.

Gascón and Barraza (2012) presented a prototype of 6-DoF stereoscopic Eye-in-Hand visual servo control of robot BIBOT. BIBOT is a low cost robot with stereo vision system mounted on the end effector. The robot manipulator can proceed with any required orientation and position within the workspace. The stereoscopic vision is used to calculate the 3D position of a target. It can support further research in image based visual servoing (IBVS), position based visual servoing (PBVS) and nonlinear control of robot. The proposed prototype is modeled and analyzed in computer aided design (CAD).

Owayjan et al. (2012) developed a lie detection system using facial micro-expression based on embedded vision system. The system proposed is using a mathematical model capable of detailed analysis of the facial expressions in frames of a video with an active and dynamic framework. A parametric representation of face is extracted which is compared with a physical model. This analysis yields spatial and temporal patterns of human face while making an attempt to tell a lie.

Zhu et al. (2012) proposed a vision based defect detection system for asphalt pavements in microwave. In the system, a weighted neighborhood average filtering is applied to image of pavements. The defect in the pavements is identified by using cluster based thresholding (OTSU method) with region growing method. Barrel distortion is corrected by cubic polynomial method and inclination distortion is rectified by re-projection technique.

Zakaria et al. (2012) proposed a shape recognition method by applying computer vision technique. The proposed technique is capable of recognizing the circular, triangular and squared shape. In this technique, OTSU method is used for thresholding the captured image, median filter for noise reduction and edge is detected by sobel operator. Thinning method is applied to output of sobel operator for removing the unwanted edges. The shape is identified by the compactness of the region. The compactness is considered 1 to 14 for circle, 15 to 19 for square and for triangle it is ranging from 20 to 40. The detection accuracy of the proposed method is 85%. This method is sensitive to noise and lighting.

Oh et al. (2012) proposed an automatic bin picking system by using stereo vision system in an environment where parts are randomly placed. In this system, the desired work piece is detected in a 2D image by a geometric pattern matching technique. A wide field of view (FOV) is considered here. A stereo camera is used estimating the position with the pattern matching technique. For a consistent result in position calculation, multiple pattern registration and ellipse fitting method is applied. The grasping point in a work piece is computed by using the position and bin information. The author established a practical bin-picking strategy by using the proposed methods. The proposed method can be further improved by considering lens distortion and structured-light technique for texture less parts.

Benjamim et al. (2012) proposed an identification method of medicine boxes through visual features for visually disabled people. They used vision sensor for identification of relevant features like speeded up robust features (SURF) and scale-invariant feature transform (SIFT) in medicine boxes and then processed the information to produce an audio file. This audio file informed the people about the medicine. The proposed method is using image processing technique for identification.

An advance method of manipulation and recognition of object is performed in by Troniak et al. (2013) by Charlie as a kitchen assistant method for navigating a robot among floors.

An efficient target detection technique is proposed by Kim and Hwang (2013) by using Kalman filter with the aid of vision system. The focal plane of landmarks is computed for better detection of desired landmark.

An intelligent binocular vision system for a robot is proposed by Huang and Cheng (2013). The developed vision system is using edge detection method for finding the object contours. A specific object is detected by the use of Hough transform arbitrary lines for random geometry. The 3D coordinates of the target is computed accurately by the binocular vision system by using principle of binocular vision trigonometry and robot localization algorithm. The distance between the robot and desired object is measured by coordinate transformation method. The vision system developed is achieving the accuracy and reliability. This system can be developed further by considering the world coordinates of surrounding environment for more accuracy in grasping and placing of object in desired place.

Kruse et al. (2013) presented a framework for manipulation of a dual-arm telerobot while holding an object through human gestures. A set of human gesture vocabulary is used during object manipulation which is interpreted as the desired object configuration. The developed robot is capable of performing autonomous target identification, desired alignment of object, selecting stable grasping. The developed system is developed by using a software for distributed communication and control system i.e., Robot Raconteur which helps in interfacing all the required components. The author cited that the lack of contact detection in this work during grasping can give rise to a more advanced system in future. Also, the time delay in communication between joint command and action can be considered for further development.

Lin et al. (2013) proposed an obstacle detection method for assisting parking of vehicles by a monocular vision system. The existence of the obstacle is determined by the frame difference method and morphological operator is used to enhance the features of the obstacles existed. The corner of the stationary or object in motion in back side of vehicle is detected by features from accelerated segment test (FAST) method. Then the falsely detected corner features are eliminated by the invers perspective mapping technique. This mapping technique results the exact obstacle region. The proposed system is considered as an alarming system to guide the driver that helps in providing collision free parking. The suggested object detection algorithm is proved to have high detection rate and feasible computational complexity.

Murai and Morimoto (2013) proposed an automated visual inspection system to inspect the packaged drugs. They considered only features which are robust to several factors like light reflection, scale transition and phase difference. Hierarchical identification with weak classifiers like Hue-normalized brightness histogram, Hue-brightness histogram and auto-correlation etc. is used for identification of drugs. In this work, the both side of the packet is not distinguished from each other. Occlusion due to the bands, if any, and split mark is not considered in the suggested system. These issues can be considered in further development.

Rao (2013) proposed a cotton recognition system for picking of cottons by robot. MV system is used for processing of images of cotton plant and color subtraction method for identification of different portions of plant. The accuracy of the recognition method is enhanced by dynamic freeman chain coding technique.

Nguyen et al. (2013) proposed an obstacle detection and mapping technique for an intelligent chair navigation. A combination of two cameras used for detection of objects under uncertainty.

Katyal et al. (2013) proposed control framework for human assistive robot with the help of MV system for recognizing object of interest estimating its position in 3D space. Prosthetic sensor is used for closed-loop robotic control. The optical component in this MV system is Kinect which comprises of optical camera for image acquisition and depth camera for depth measurement in a scene. The 3D points in the acquired image lies outside of the workspace are filtered by pass-through filter. The shape of the objects are identified by a sample consensus algorithm. Only spherical and cylindrical shapes are considered by the author. This algorithm also compute geometric properties like radius and orientation of object which helps in finding stable grasping. This work can be expanded by considering objects with different shapes and size.

Dahiya et al. (2013) proposed a defect identification system by classification of known eye disease based on fuzzy rule base. A data set is used which contains symptoms of some category of eye disease. The new data like dark spot in middle or corners areas, patches, blurring, double vision, etc. is recorded from patients and is compared with stored data set to classify the disease. After acquisition of symptom. The defect is analyzed by fuzzy rule base and image rendering mechanism.

Sinha and Chakravarty (2013) presented an identification system that accurately identify a person through considering walking pattern in some indoor environment by 3D human pose estimation. The pose of a person is modeled by considering skeleton points captured through Kinect sensor.

A spatiotemporal set of key poses and sub poses are used in modeling the gait pattern that occurred in gait cycles periodically. In this work, static and dynamic features related to motion are extracted to model the gait pattern. They proposed that in future this work can be made more robust by employing the automatic rotation of Kinect and using an efficient noise removal technique.

A vision based object identification method is proposed by Resendiz et al. (2013) for inspection of rail road network. Spectral estimation is used for detection of any defect in the track. They proposed an autonomous object detection system.

Kim and Hwang (2013) presented a method of landing an unmanned aerial vehicle (UAV) autonomously based on information acquired by vision sensor. They developed a platform for UAV using several avionic sensors and integrated with a flight control system. The vision algorithm that is color-based recovery-net detection is used to detect the landing information for a reliable landing of UAV. Kalman filter is used to extract the features of object in an image of the environment. The proposed method provides an effective solution to landing of a UAV without GPS information.

Darabkh et al. (2014) proposed an efficient feature extraction method for iris pattern matching. The proposed method is using sliding window control technique with some mathematical procedure on pixels to yield a feature vector of smaller size for an efficient and faster iris recognition system. The effect of varying light intensity is reduced in the proposed method. Several performance metrics like false rejection rate, false acceptance rate and the recognition rate are considered for evaluating the performance of the method.

A vision guided robotic system is proposed for an automated security checking of vehicle by Fareh et al. (2014). The automated vehicle inspection system integrate vision sensor with a robotic system for realizing the suggested system.

2.3 Vision Sensor Based Robot Navigation

Industrial vision system (IVS) for robot navigation provides the location information of the parts which are spatially visible. The location information can be used to configure automated handling, automated assembly and part processing system. This reduces the cycle time in production and cost significantly.

Shimizu et al. (2001) presented a wide angle foveated vision sensor system for navigating a mobile robot autonomously. This system is using two vision system i.e., central vision system and peripheral vision system and their cooperation is contributing towards a high quality navigation of

robot. The central vision system associated with a proposed algorithm is used to determine the obstacle information from 3D image. The peripheral vision system is used to obtain the location and orientation information of the obstacle and boundary of path.

Li and Yang (2003) proposed a vision-based landmark recognition system for navigation of robot in unknown environment. The pattern of landmark is searched within the image captured through a vision sensor by using genetic algorithm. The vision system used in this work is capable of producing the features of corresponding landmarks. To avoid the obstacles ultrasonic sensors are used in combination of a set of fuzzy rules.

Jin et al. (2003) proposed a sensor-fusion technique in which obstacle and robot position information in past are recorded for further use. These data are useful in accurate determination of navigation trajectory. This technique calculate the next position of robot consulting the stored data set with the current data set. These data are computed by the help of vision sensors along with ultrasonic sensors.

Okada et al. (2003) presented a vision based robot navigation system in unknown environment. They used plane segment finder technique which finds the random planner surface areas from the depth of an image. The proposed humanoid is navigating in an unknown environment with the help of stereo vision system.

Peng et al. (2006) proposed an intelligent control and navigation of mobile robot based on visual information acquired from a vision sensor. This information is used for controlling of mobile robot and made it to navigate towards the desired position. They implemented the object identification and image segmentation based on Hue-Saturation-Value color mode, clustering algorithm. They computed the location information of desired objects through CCD vision sensor.

Huwedi et al. (2006) proposed a vision based feature extractor which enables the robot to navigate to an unexplored place based on some important feature of the place. The next target position for the robot navigation is determined based the number of features and their visibility by robot sensor.

Segvic et al. (2007) presented an appearance-based navigation approach by using a single perspective vision sensor. They developed the vision system for outdoor navigation of vehicles in a dynamic environment where the factors like condition of imaging and temporary occlusion are present. The robot is trained in a hierarchical structured environment. The structured work environment contains a graph of important images with their 2D features which enables the robot to navigate in the workspace.

Won et al. (2008) suggested a vision-based navigation of mobile robot. They used a vision sensor and particle filter where vision sensor is used for acquiring images and tracking and a particle filter is used to manage the navigation in a nonlinear observation model.

Moghadam et al. (2008) proposed a navigation plan based on stereo vision system for a mobile robot. They measured the objects by a stereo vision camera. The vision sensor is integrated with a laser range finder which provides a 2D cost map for navigating the robot in messy and complex environments. The occupancy grid map of objects is used for obstacle avoidance.

Neto et al. (2008) proposed a nondeterministic approach for eliminating the redundant data present in the real time autonomous navigation system. They used a machine vision segmentation algorithm i.e., Threshold and Horizon Finder algorithm which efficiently identifies the navigation area from an image. This approach reduces the processing time and further it can be optimized for better performance.

Hagiwara et al. (2009) proposed an improved view-based navigation of robot which is invariant to changes in scenery by adjusting the position resolution. The improved navigation method is able to control the position resolution, and then flexibly respond to the change of surroundings and navigate the robot accurately. In this method illumination issue is handled by applying edge detector with smoothing technique. The smoothing technique is robust to changes in an image and environment.

Li et al. (2010) proposed a 3D positioning system by using omnidirectional vision sensor and landmarks for mobile robot navigation in a known environment. They developed a positioning algorithms to generate the position information of objects. The omnidirectional vision image can provide 360 degree information around the sensor. This helps in navigating the robot in circular or straight line path.

Chang and Chu (2010) presented a robot localization and navigation system with the aid of vision system in an unknown environment. The proposed method is using vision sensors and on-line calibration procedure. They used a recursive calibration of vision sensor with respect to robot kinematics to determine the position information and orientation of stationary cameras. These information are used to build the sensor network in real time. The field of view of the adjacent camera pair is likely to be overlapped, but their position and orientation in Cartesian space are unknown. The mobile robot is controlled by effectively calibrate it with any stationary camera which is capable of observing five preselected color-coded features. This results in establishing

the coordinate transformations among the camera. The proposed approach is an efficient approach as it seems to be cost effective, flexible and efficient.

Zhao and Jiang (2010) presented a guidance model for detection and localization of crops with the help of vision system. The quasi navigation baseline is determined by pattern recognition technique. The dynamic navigation of robot is performed based on the navigation information extracted from quasi navigation baseline with Hough transform. The proposed model is a simple and robust technique.

Yanjun et al. (2010) proposed a multi-sensor vision system for robot navigation. The data from each sensor is unified by world coordinate system. Each sensor is calibrated locally then the least-square processing of this data is done by genetic programming. For robot navigation, a plan is inferred from this global data by delta fuzzy rule which creates accurate and complete map of the environment containing obstacle information.

Zhang et al. (2011) proposed a robotic assembly approach with the aid of vision. They used the camera space manipulation method to achieve the vision based robotic assembly. Local calibrations is performed to achieve high accuracy in aligning the parts for final assembly. The visual feature is extracted by a vision sensor and assembly feature is determined by CAD model which helps in implementing of the proposed technique. The visual information is given as input to robot as visual servoing for navigation of robot to the desired position.

Troniak et al. (2013) presented a semi-humanoid robotic system for navigation in a multi-stored building autonomously. The proposed robot is design and implemented to operate an elevator. This robot is efficient in locating points of interest, manipulating objects and navigating objects among floors. This work is focused on Charlie as a kitchen assistant method for object recognition. The future scope of this work is to design and realize a robot can communicate and find alternate solution to a task by using advanced multimodal semantic scene understanding method.

Nguyen et al. (2013) presented an experimental study of smart wheel chair system i.e., thought-controlled intelligent machine. This intelligent machine is using stereoscopic cameras for 3D depth calculation and for mapping a 360° spherical monocular vision. This combination of camera helps in obstacle detection and mapping for autonomous navigation of chair in unknown environment. The study uses a hand-free control method that consists of head-movement controller and brain computer interface.

Kim and Hwang (2013) proposed a technique for navigation system in a poor lighting environment with less number of visible landmark. The navigation system is developed by integrating the inertial and landmark based vision. The proposed system is using an indirect Kalmann filter with a feedback to the system. They measured the focal plane of landmarks for effective computation of landmarks. They compared the proposed system with a system integrated with a vision navigation solution. Further improvements to this system can be done with a consideration of error characteristic of vision navigation solution.

Resendiz et al. (2013) presented a tele operated navigation system that employs vision based obstacle avoidance and object identification technique for a rail road network. Spectral estimation and signal-processing methods are used for detection and tracking along the track. In future, the track inspection technology can be made more robust by using automatic detection and segmentation of objects. This can lead to an autonomous system for inspecting miles of tracks without human supervision by automatic adjustment of spatial detection filters.

Lai and Lin (2013) proposed a cross-floor navigation system for stair-climbing mobile robot autonomously using vision sensor. The image based navigation system is developed by using upward and forward looking camera setups. The guiding lanes for robot is determined by using image features like stair lines, skirting lines and ceiling landmarks. The long distance navigation of robot is achieved by installing the wireless sensors in the image dead zones. A visual servo control strategy is developed based on image feature error for accurate navigation.

Diskin et al. (2013) presented autonomous navigation method for a mobile system in an unknown environment with fixed obstacles and moving targets. They used several computer vision algorithms. A stereoscopic camera is used for estimating the depths of the environment. The originality of this approach is efficiency of the algorithm and scene reconstruction to take decision in real time. The detection and recognition of human body is performed based on efficient local binary pattern (LBP) descriptors. The correctness of the detection is measured by modular component analysis along with the Voila-Jones face detection and recognition algorithm which enables a pose invariant face detection.

Steiner and Brady (2014) proposed a vision based algorithm for hazard detection and navigation for flights. This system includes two downward-pointing monochromatic cameras, a horizon-pointing monochromatic camera and modular software framework. They used two vision based

hazard detection algorithms, one is for searching of camera field of view for suitable lading position and other one generates a plan for navigation and safety landing in real time.

Fareh et al. (2014) presented an automated vehicle inspection system by the integration of vision guided robotic system. They proposed a seamless and efficient procedure of integration of various sensors and robotic system for navigation of robot manipulator over the region of interest on the surface of vehicle. This helps in achieving an automated security screening of vehicles.

Zheng et al. (2015) proposed a multi-sensor fusion based monocular visual navigation system for a quadrotor. A local feature detector and descriptor is used for feature detection and matching. They suggested a motion estimation algorithm based on the flight characteristics of quadrotor. A pose estimation algorithm is proposed by the authors for precise pose estimation. The proposed method transforms the 6-DoF pose estimation problem into a 4-DoF problem. This results in more accurate result in less time.

Li et al. (2016) proposed an image based navigation approach. The navigation with higher precision is achieved by using a geo-referenced image database in environments with poor information of global positioning system. An efficient matching algorithm is used to search and match a real-time images with image in the database. 3D navigation parameter is calculated from the matched image.

2.4 Visual Servoing and Robot Control

A global overview of research and development in visual servoing is presented here. Individual sections cover traditional and novel topics in visual control that are relevant for this thesis. A division is made between historical, traditional and modern work in visual servoing, as well as on developments in path and trajectory planning. Finally, recent work on visual control is discussed that most resembles the work perceived in this thesis. A few similar studies are highlighted and a comparison is made to emphasize the differences between the both. To limit this review to the field of visual control, a general review for the topics of modelling and control of robotics as well as visual processing is not considered. Furthermore, this review discusses a global overview of existing methods.

Visual servoing, vision-based robot control, is a robot control strategy that uses visual information for motion control of robot. The earliest development in visual servoing was performed in the SRI International Labs in 1979. Such control system is proposed by Agin (1979). It is a method for guiding the robot motion by camera as vision sensor. This robot control method is mainly of two

category. The first one is IBVS which involves in developing the control strategy by using extracted feature vector of object to directly control the DoF of the robot. The second type of visual servoing is PBVS that involves the geometric interpretation of the information extracted from the camera, such as estimating the pose of the target and parameters of the camera with an assumption that some basic model of the target is known.

Malis et al. (1999) focused on the problems that generally appear in image-based visual servoing when the initial camera position is far away from its desired position. They presented that local minima or a singularity of the image Jacobian can be reached during the servoing by suitable example. One of the important issue in visual servoing is the singularity of the Jacobian matrix. It is well known that the image Jacobian is singular if the object is collinear, or belong to a cylinder containing the camera optical center i.e., image with three points. Using more than three points, singularity problem can be avoided. The singularity problem is addressed by combining visual features obtained directly from the image, and position-based features. This method of servoing is called 2 ½ D visual servoing. This approach combines the image features and 3D information. The 3D information is obtained by applying pose estimation algorithm through a projective reconstruction which is resulted from current and desired images. Similarly the image features are retrieved by using image based visual servoing technique. The limitation of this approach is that the corresponding control laws are more sensitive to image noise than classical image-based visual servoing.

Malis and Chaumette (2000) proposed a new method of vision based robot control. In this approach, the author eliminates the drawbacks of both IBVS and PBVS. The control strategy for visual servoing is derived by estimating the partial displacement of camera from the current to the desired position at each iteration of the control law. The control law for controlling the 6-DoF of camera is developed based on the visual features extracted from the partial displacement of camera.

Sim et al. (2002) proposed a position based visual servoing for an eye-in-hand configuration. They modified the Smith Predictor DeMenthon Horaud visual servoing system for 3D visual tracking. The inherent vision delay is eliminated by the use of Smith Predictor which is a key factor in the performance of the servoing scheme.

The method of camera-space manipulation (CSM) is used for a robust and precise technique for vision guided robot manipulation. This method does not require proper calibration of cameras and

manipulator kinematics which is required in calibration-based methods (Li and Yang, 2003). Additionally, the fast and real time image processing technique is not needed in CSM method.

Chaumette (2004) proposed an image based visual servoing based on moments. The interaction matrix is derived from segmented images by using Green's theorem that can be applied to any moment. In the suggested technique, a combination of six moments are used to control the 6-DoF of the system. This technique is yielding better for a symmetrical object or a planner object with complex or unknown shape.

Hrach et al. (2005) developed a design approach for vision based robot control system. The proposed design approach includes 1 MPixel CMOS sensor and digital signal processor. The bandwidth requirement for the integration of the sensor with the host computer system is reduced as the sensor is combining the both image acquisition and processing. The proposed design approach results in low power consumption and high frame rate. The suggested sensor can be used further as a building block for more complex measurement systems.

Chaumette and Hutchinson (2006) presented basic techniques of the formulation of visual servo control problem. They described two visual servo control schemes called IBVS and PBVS. They also discussed the performance and stability issues in these two schemes.

Chaumette and Hutchinson (2007) discussed the advanced techniques that involves in dealing the stability and performance issue in IBVS and PBVS.

Collewet and Marchand (2011) suggested a new approach to control the robot through 2D visual servoing. They used the luminance of all pixels in the image for servoing. The main advantage of this approach is that the visual features are not required to be matched or tracked. The interaction matrix is calculated analytically. This computation is based either on a temporal luminance constancy hypothesis or on a reflection model so that complex illumination changes can be considered. They implemented this approach by mounting the IEEE 1394 camera on a 6-DoF Gantry robot. The control law is calculated in real-time. The implementation in this work indicated that it is possible to use directly the luminance of all the pixels in an image as visual features in visual servoing. This enables in eliminating the complex image processing tasks. The image spatial gradient is used for computing the interaction matrix. In this control scheme, the matching between the initial and desired features is not required. They applied this approach for positioning and tracking tasks. This is robust to approximated depths, low textured objects, partial occlusions and specular scenes. They also showed that luminance leads to lower positioning errors than a classical

visual servoing based on 2D geometric visual features. This approach is not sensitive to partial occlusions and to coarse approximations of the depths required to compute the interaction matrix. Nacereddine et al. (2011) proposed a control plan for a turrent Pan-Tilt system by utilizing the visual features extracted by correlation method. The visual features are optimized by a multiscale pyramidal algorithm. They employ the image based visual servoing scheme to determine the kinematics screw of the camera.

Shi et al. (2012) focused on a critical problem for autonomous capture in space industry. They proposed a hybrid 2D/3D visual servoing technique for space robot to capture the target. The on-orbit servicing autonomously is an important task in space industry. For such task, visual servo of eye- in-hand type configuration is suitable. They developed a 3D simulation platform to verify the proposed method. They used image based visual servoing scheme as it is simple and less sensitive to camera calibration error. When the target is very close to the end effector, then position based visual servoing is used. They developed a 3D simulation platform by considering the different aspects of space environment like lighting and power of computation, safety and real-time scenario, cooperative visual marker, processing time delay, and capturing strategy. The trajectory for the motion of the robot joint is smooth and in accordance with the joint limit. They use Kalman filter to reduce the time delay between image acquisition and image processing. They also suggested a future scope that due to this delay, there is a possibility of collision of end effector with the target.

A vision information based control of robot proposed (Lai and Lin, 2013) in which an autonomous cross-floor navigation system for stair-climbing mobile robot using wireless and vision sensors is developed. Image based navigation with the help of upward and forward looking camera are used. Features like ceiling landmarks, stair lines, and skirting lines are extracted by appropriate image processing technique to guide the robot in its navigation.

Tamtsia et al. (2013) proposed a new method for controlling the rotational motion around x-axis and y-axis of the camera which is invariant to shape of the object. The invariant features are calculated by employing shifted moments. The influence of noise on the performance of servoing is reduced by applying low order shifted moments.

Zhang and Liu (2013) presented an advanced approach of IBVS control by using uncalibrated camera. The control scheme is developed by using an extended interaction matrix. The feature vector is decoupled based on estimation using Kalman filter. The analysis and comparison between

point-based and moment-based features are carried out with respect to a 4-DOF positioning task. Then, an extended interaction matrix related to the digital image, and a Kalman filter (KF)-based estimation algorithm of the extended interaction matrix without calibration. The proposed method is useful for any robot controller in dynamic environments and efficient in handling planar object with complex and unknown shape. The control scheme used in this work is resulting a better robot trajectory in Cartesian space at the expense of the performance in image feature space. The Kalman filter based algorithm is extended to realize an approximation to decoupled control scheme. Experimental results conducted on an industrial robot that provides an accurate estimation of interaction matrix. The performance is similar to traditional calibration-based method. The proposed methods is suitable for application to any robot control system in dynamic environments, and can realize instant operation to planar object with complex and unknown shape at large displacement.

Bakthavatchalam et al. (2014) introduced the idea of tunable visual features with a shifted moment for visual servoing schemes. They compute two metrics, the errors in the image space and orthogonality between the rotational motion along X- and Y-axis. The proposed method is suitable for controlling the rotational motions where all DoF are involved. This visual servoing technique is applied successfully for a symmetrical object using binary moments and a freeform planner target using photometric moments.

Tahri et al. (2015) proposed a visual servoing technique based on shifted moments. This moment provides a unique combination of visual features to control the 6-DoF of an eye-in-hand camera set up independently by the shape of the object. They computed the features from shifted centered moments which enables in controlling the rotational motion. They used two cases of shifted moments. Firstly, shifted parameters are computed from object moments in the image and secondly, shifted points are determined from object contours. Affine invariant is used to calculate the shifted points.

Jagersand et al. (2015) proposed a grasping method by using visual servoing scheme. They used image based visual servoing scheme to calculate the feasible grasping points. The use of servoing scheme is simplifying the development of stable grasp plan and enhancing the performance of grasping.

Hajiloo et al. (2016) presented an online IBVS scheme for a 6-DoF robot. They used robust model predictive control to develop the proposed scheme. The control law is designed by considering

physical limitations and visibility constraints of robot. The proposed IBVS controller avoids the inverse of the image Jacobian matrix which can solve the intractable problems for the classical IBVS controller, such as large displacements between the initial and the desired positions of the camera.

Several important literatures and their outcomes are given in table 2.1.

Table 2.1: List of some important literature

Author	Year	Important Contribution	Remark
Meer and Georgescu	2001	Proposed a gradient-based edge detection with embedded confidence that provides an improved performance for low-level vision operators.	An efficient gradient based edge detection technique is implemented
Tahri and Chaumette	2003	Determined the moment invariant of an object which is invariant to scale and rotation. These invariants are used to decouple the camera DoF that helps the robot to determine the trajectory. A new shape descriptor technique is suggested to identify the planner objects. The features invariant to scale and rotation is used for shape recognition. A visual servoing scheme is proposed where robot is capable of handling planner objects with complex and unknown shapes. These features can be used in pose estimation problem.	Moment invariant based shape descriptor is used for IBVS.
Coleman et al.	2003	Proposed a novel evaluation technique which analyzes the sensitivity of edge found through discrete second derivative operators considering angular orientation and displacement errors. The proposed method uses a finite element interpolation to the output of the second order derivative operator.	A performance evaluation scheme for edge detectors with second order derivative is presented.
Tahri and Chaumette	2005	Proposed an approach for pose estimation problem of planar objects by moment invariant method. The visual servoing scheme is developed by using iterative optimization method. The moment invariants are used to control DoF of the camera that allows to provide a large convergence domain for the system and avoid the presence of local minima.	The control law for PBVS is developed based on moment.
Brannock and Weeks	2008	Proposed a wavelet transform based edge detection technique which retains the separable filtering and the simplicity of computations. The filter is designed by considering the standard 2D wavelet transform.	A simple and low computational complex edge detection technique is developed by using wavelet transform. This technique is suitable for object detection system
Govindarajan et al.	2008	Proposed a quantitative analysis of the performance of the edge detection techniques. They proposed an edge detection technique by using a partial derivatives of Boolean function which is proved to determine the optimal edges. This localization of edge pixels is proved to be efficient.	A quality assessment method for edge detectors is presented.

Boaventura	2009	Proposed a method for performance evaluation of edge detectors like canny, sobel and Russo etc. by calculating the quality indices such as true positive, false positive and false negative.	Performance evaluation of edge detection technique is suggested.
Dawei et al.	2010	Proposed an edge detection technique by combining Hopfield neural network with mathematical morphology for defect detection in woods.	A neural network and mathematical morphology based edge detection technique is developed for defect detection.
Kaur and Garg	2011	Proposed a procedure for the application of mathematical morphology based edge detection to remote sensing images. This method enhances the noisy image and then detect edges. The author compared the proposed technique with other edge detection for evaluating the performance of the proposed technique.	Mathematical morphology based edge detection is developed for remote sensing images.
Zhang and Li	2011	Proposed an efficient edge detection technique for color image by using morphology operator. The morphological edge detection is implemented by using different structure and scale elements to hue, saturation and intensity. The proposed algorithm efficiently eliminate the noise and capable of extracting the complete edge information.	A method of color edge detection using mathematical morphology is developed.
Alonso et al.	2011	Proposed a multiscale algorithm for the unsupervised extraction of the most significant edges in SAR images. The wavelet transform is used to enhance the pixels in the image prior to edge detection.	Localization of edge pixel is performed by using wavelet transform to improve the rate of edge detection in SAR images
Wei-Fengma et al.	2012	Proposed an improved wavelet multi-scale edge detection algorithm. Morphology operator is used with the wavelet transformation to effectively filter the salt and pepper noise.	Proposed an improved wavelet multi-scale edge detection algorithm.
Jiang and Zhao	2012	Proposed an automatic Apple recognition method by using image processing technique for harvesting by a robot. The optimized Hough transform is used to recognize Apples. The Otsu algorithm is used for effective segmentation of adjacent and overlapped Apples and morphological operator is used for eliminating noise. Sobel operator is used for edge detection.	Machine vision system is used for harvesting Apple.
Jihong et al.	2012	Proposed an edge detection algorithm based on fuzzy entropy. Non-maximal image suppression with dynamic thresholding technique is used for optimizing the edge detection algorithm.	An adaptive fuzzy entropy algorithm is used for detection of edge pixels.
Du et al.	2012	A segmentation algorithm based on wavelet transform is proposed for anterior chamber Optical Coherence Tomography (OCT) images. The segmentation algorithm is using a dynamic thresholding technique. Kirsch edge detection operator is used to detect the edge of region of interest.	A segmentation algorithm is developed by employing a dynamic thresholding technique.
Junna et al.	2012	Discussed edge detection algorithm based on mathematical morphology. A soft morphological edge detection algorithm is proposed that effectively detects edge by restraining the impact of noise.	An algorithm of edge detection based on soft morphology is developed.

Xu et al.	2012	Proposed a morphological edge detection operator based on multi-structure elements for colored noisy images.	Color edge detection using mathematical morphology is developed.
Wang and Yan	2012	Proposed a vector morphological operator for detection of edges in color images. The method overcomes the drawback of traditional edge detection methods which are more sensitive to noise.	Edge detection of color image using vector morphological operators is developed.
Liu et al.	2013	Proposed an effective local feature extraction algorithm for lane detection. The lane detection technique is using morphological operator along with edge refining procedure to reduce the interference and computational cost.	Lane detection algorithm based on local feature extraction technique is developed
Patel and More	2013	Proposed an edge detection technique by using fuzzy logic to produce precise and noise free edge. The detection rate is enhanced by cellular learning automata that reduces the repetition of pixels in edge.	An Edge detection technique by employing fuzzy logic and cellular learning automata is developed.

2.5 Determination of Research Gap

Although a lot of research work in the area of vision system in general and that for robotic system in particular have been carried out by numerous authors over last few decades, there remain a lot towards the advancement of these work for enhancement of technology. After carrying out an extensive survey of the available research work, it is felt that most of these work are done for achieving some specific objective. A dedicated and integrated robotic vision system that can take care of all the necessary tasks of a robotic assembly system is call of the day.

2.6 Problem Statement

Robotic technology alone may not be sufficient enough to fulfill the needs that the present day manufacturing industries demand. In order to make the robotic system more autonomous and competitive, other technologies such as sensor technology can be of great help. Although robot itself makes use of several sensors for its major operations, use of additional external sensors can improve its performance and expand its scope to a very large extent. Therefore, it is thought that the use of an appropriate vision system can greatly improve the performance and achieve the desired objectives in the situations of parts assembly. The work presented in this thesis is focused on design and development of a robotic assembly system with the aid of vision system and necessary algorithms for making the robot capable of carrying out part assembly process autonomously.

2.7 Objective

The prime objective of the proposed work is to design and develop a robotic part assembly system with the help of vision system and necessary algorithms. The developed system should be capable of the following.

- Capturing information about the part presence.
- Identifying the parts through extraction of features such as size, shape, and orientation.
- Determining the position and points for stable grasping for desired manipulation or orientation as per the requirement of the assembly tasks.
- Navigating the robot manipulation in the environment of multiple obstacles.
- Avoiding possible collisions and in dynamic unstructured environment.
- Inspecting the end products for their completeness.

Hence, the objective of the research work are:

- to develop a robotic system with vision sensor for carrying out part assembly in industrial environment;
- to develop a set of appropriate algorithms for part measurement, part detection and robot navigation through visual sensing technology;
- to validate the developed algorithm and check their robustness for uncertainty and
- to implement the developed system for part manipulation in robotic assembly process for mechanical parts.

2.8 Summary

Over last few decades, lot of research work have been carried out by numerous researcher in the area of vision system and for robotic system in particular. Extensive survey of the available literature is performed in this chapter to understand the issues lies in this area. In the survey, feature extraction, part detection and visual servoing methods are considered. To design and develop the proposed system, it is important to finds out the important aspects of such available system. Some of the important research outcomes are identified for this research work.

Chapter 3

Materials and Methods

3.1 Overview

The work envisaged under this research programme can be split into two distinct phases. In the first phase, most of the focus remains on reviewing the literature, framing the methodology and developing algorithms for creating appropriate base for the second phase. In the second phase, setups are created for an integrated robot vision system that helps in carrying out part assembly in an industrial environment. This chapter presents the items, both hardware and software along with other necessary items required for abetting the work.

A vision guided part assembly system includes three core system such as robotic system, vision system and component handling system. The vision system and control software gives the robot exact coordinates of the components, that are distributed randomly underneath the camera field of vision and reports the information back to the robot. With this information, the robot arm can move to a desired components and pick from the conveyor belt. The conveyor, normally, stops under the camera where the position of the desired component is determined. On the other hand, a component can be picked up from the belt without stopping it when the cycle time is small. This is achieved by fitting an encoder to the conveyor and tracking the component through vision software. This functionality is usually referred as VGR.

3.2 Materials

True vision guidance in robotics is both a hardware and software issue. In order to achieve the objective of the proposed research work, it is desirable to select the typical application areas, selection of appropriate vision system (Monocular/Binocuar) and selection of additional sensors as per requirement of the tasks and assembly environment. Proper integration of sensors and their feedback system with robot controller is required. Finally, integration of MV system, other necessary sensors with robot controller is required as per the need.

In order to achieve the objective of the research work proposed the following as in Table 3.1 hardware and software are required.

Table 3.1: Materials required in the proposed research work

Sl.No.	Material	Equipment/Facility	Purpose
1	Hardware	Robot Manipulator	For Assembly of components
2		Machine Vision System	For Object Identification, Position and Orientation Information of Objects, Assisting in Robot Manipulator for carrying out the desired Assembly Tasks
4	Software	MATLAB	Modelling and Simulation of methods for desired tasks
5		LabVIEW	Modelling, Simulation and Feedback Control of robot

3.2.1 Hardware

The development of a vision guided robotic part assembly system requires detailed study about the components required. The selection of MV system is vital as the precision plays an important role. The world achieves the qualitative as well as the quantitative enhancement of the productivity and the lesser manufacturing costs by employing the vision system. The unique vision power along with the unparalleled sharpness of the vision system supports to accomplish these advantages.

The continual advances in vision options fuel improvements in manufacturing automation and quality. Several industry houses are in making of MV systems for easy integration with any robotic or automation application. Vision systems are used in almost every automation process for managing part location, inspection, and tracking/bar-coding purposes. The following hardware components are required for developing a vision based solution

- Industrial Robot
- End-effector
- Machine Vision System
- Other Sensors
- Sample Objects/Parts

A. Industrial Robot

The industrial robot used in the proposed research is Kawasaki RSO6L shown in Figure 3.1 as it is available in the laboratory and is suitable for the desired purpose. The technical specification of the robot is given in Table 3.2 [http://www.robotics.org/product-catalog-detail.cfm/Kawasaki-Robotics-USA-Inc/Kawasaki-RS06L-Robot-high-speed-6-kg-payload/productid/3650].

Figure 3.1: Industrial Robot (Kawasaki RSO6L)

Table 3.2: Technical Specification of Kawasaki RSO6L

Feature	**Specification**
Arm Type	Articulated
DOF	6
Payload Capacity	6 kg
Horizontal Reach	1,650 mm
Vertical Reach	2,982 mm
Repeatability	+ 0.05 mm
Maximum Speed	13,700 mm/s
Mass	150kg

B. End-Effector

There are several makers of end effector available grippers are not utilized exclusively with robots in any case. They can be utilized for an automated or semi-robotized assembly line. As per as the requirement the SCHUNK End-effector is selected to fulfil the criteria of this research work. The end effector used in this work is shown in Figure 3.2.

Figure 3.2: End-effector (Make: SCHUNK)

C. Machine Vision System

The selection of vision system depends on a number of factors such as the environment, the exact task, and the type of objects etc. Selection of a universal test system which will satisfy all the required criteria is not unique and are application oriented. Choice of objects to be tracked and the process environment where the system is to be installed are two major essential requirements for establishing a stable system. Figure 3.3 provides a detailed description about the basic system parameters and conditions for choosing a MV system.

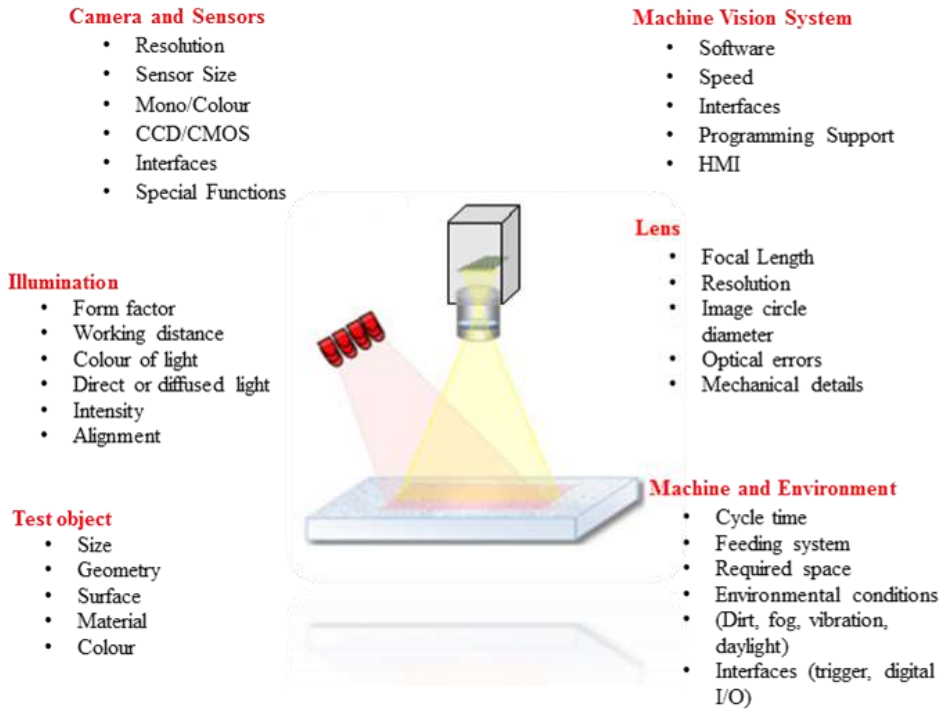


Figure 3.3: Selection Criteria for a Machine Vision System

A vision processor is a class of microprocessor which is designed to perform specific task and acted as an accelerator for vision tasks. Direct interfaces are available in a vision system. Consequently, data can be extracted from several cameras and on-chip dataflow among many parallel execution units can be performed. The vision processor used in this research work is NI PXIe 1082 as shown in Figure 3.4 and the technical specifications are given in Table 3.3 [www.ni.com/pdf/manuals/372752b.pdf]. This processor has interfaces for other sensors and data acquisition devices.

Figure 3.4: NI Machine Vision Processor (Model: NI PXIe 1082)

Table 3.3: Technical Specification of NI Machine Vision Processor

Control System	Model
PXI platforms	NI PXIe-1082
X series Multifunctional DAQ	NI PXIe-6341
Motion Controller	NI PXI-7340
1394 Host Adapter	NI PXI-8252
NI Compact RIO	NI cRIO-9074
Dual GBEthernet	NI 8234
4-Ch Universal Analog Input DAQ	NI 9212
16-Ch, 24 DI/DO DAQ modules	NI 9375
8-Ch,TTL High speed DIO DAQ	NI 9401
ATI DAQ Module	FTPS1
SCB DAQ	SCB-68A

D. Camera

The image acquisition is a complex interaction of the material and surface characteristics of the test object and the lighting used. The vision sensor (camera) plays an important role in acquiring the image and can be termed as "detector unit". An image is captured by the sensor and is processed by the vision processor. The selection of camera depends on the required accuracy and speed of the inspection as well as the application area.

There are two types of cameras available i.e. charged couple device (CCD) and complementary metal oxide semiconductor (CMOS). A CCD camera creates high quality and less noisy images. It requires analogue to digital converter as the output signal is analogue in nature. It consumes more power and costs high. But a CMOS camera does not require any converter as the output signal is digital one. CMOS type camera are more susceptible to noise, but consumes less power and is economic. The speed of a CCD camera is ranging from moderate to high, whereas CMOS camera speed is high. This choice really depends upon the particular application requirements.

Types of industrial cameras

- Line scan cameras: This camera has a single row of pixel sensors. The lines are continuously fed to a computer that joins them to each other and creates an image. This is usually performed by connecting the camera output to a frame grabber. Buffering of images as well as preprocessing of images is sometimes performed by the frame grabber. A suitable number of individual lines are captured in a very quick succession by one single such camera.
- Area scan camera: This type of camera contains a matrix of pixels that capture an image of a given scene. They are more general purpose than line scan cameras, and offer easier setup and alignment. Area scan cameras are best suited towards applications where the object is stationary.
- Intelligent cameras: This type of camera performs the complete process within it i.e., from image acquisition to store image data. After that, several interfaces and protocols are utilized to transfer the stored information to a robot. These types of cameras are mainly area scan cameras and sometimes line scan cameras.

Criterion to select an appropriate camera

The selection of camera solely depends on the type inspection task. The features to be extracted should be represented in such a way that the features can be reasonable evaluated by the software.

- Camera Type: Development of a successful application environment depends on the correct type of camera. Based on the category of task, the type of camera is selected. Area scan camera captures the image either in continuous or triggered mode whereas the seamless scanning of objects are performed by the line scan camera.
- Resolution and sensor size: Area scan cameras generate digital images comprised of pixels which correspond to some area that the camera "sees". Since a pixel is the smallest individual unit that the camera can detect, this is the baseline for selecting camera resolution.

  A general rule,

  $$\text{Resolution} = (\text{Area to be viewed}) / (\text{Size of object})$$

  In case of large size components, high resolution camera is needed. However, the size of camera solely relies on the application environment.
- Speed of frame rate: Speed of a camera defines the number of images or lines per second that the camera captures. Fast communication interfaces are needed for fast cameras. In terms of frame rate the CMOS type sensors outperform the CCD sensors.
- Quality of image: CCD sensors provide higher quality images than CMOS. The choice between CCD and CMOS determined by application requirements. For several applications, CMOS will be good enough. CCD camera yields homogenous images. These images are sensitive to lighting conditions yet inclined to spreading and blooming in case of overexposure. CMOS cameras usually have need of more light and require calibration (dark and bright image). CMOS type produces images which are more inhomogeneous in nature. But, they are robust to different lighting conditions. They do not display any spreading and blooming like CCD cameras.
- Monochrome or color sensor: Number of bits required for image data to represent distinguishes the sensor type as monochrome or color. Monochrome camera use 8-bit representation format to code the image data and are prone to lighting conditions. Color images require 24-bit and involve additional processing to distinguish different

colors of the identical brightness. The volume of data in color sensor is much larger than monochrome.

- Transmission interface: Selection of interface depends on the bandwidth required. The bandwidth (per second) required can be calculated by multiplying the bit depth, frame rate and image size. The interface must be efficient enough to communicate the image data to the processing unit reliably. Simultaneously the software used should be capable of integrating the interface to the associated camera.
- Mechanical dimensions and form factor: The dimension and form factor cannot be overlooked while selecting the camera. The camera to be used must fit into the system to be developed. As many cameras are available in the market, so the choice is wide.

Lenses

Selection of lenses is playing an important role in optimizing the use of high performance camera. The machine vision lenses can be classified into two broad categories:

- Lens for Field of view (FOV) that is much larger than camera sensor size
- Lens for Field of view that is smaller than camera sensor size.

In industrial applications, the lenses are available with either fixed focal distance or a variable focal distance. Generally, there are 3 factors that governs the lens selection such as FOV, working distance and sensor size of the camera. The focal length required for the application can be estimated as:

$$\text{Focal Length} = (\text{Magnification} \times \text{working distance}) / (1 + \text{Magnification})$$

Where, Magnification of an image acquired = ((Sensor Size of the camera))/FOV

The cameras as shown in fig. 3.4 and lens are chosen by considering the above cited criteria. For acquiring image in the present work, BASLER CSA640-70gc and 7fc model camera is used. The technical specification is given in the Table 3.4.

Table 3.4: Technical specification of Camera used

Camera model	Resolution (H x V pixels)	Sensor	Frame Rate	Mono/Color	Interface
scA640-70gc (Make: BASLER)	658 x 492	ICX424	70 fps	Color	GigE
scA640-70fc (Make: BASLER)	658 x 492	ICX424	71 fps	Color	FireWire

There are several other necessary sensors required for accomplishing several task and the sensors required for such tasks are given in the Table 3.5.

Table 3.5: Sensors associated with different activities

Activity	For Measuring	Types of Sensor
Detection	Pressure	Electromagnetic, Optical sensor, Infrared Ranging, Vision Sensor
	Range or distance	Triangulation (Range), Structured Lighting (Range), Ultrasonic sensors, LADER (Light Detecting and Ranging)
surface Measurement	Textures	Touch sensor, Proximity capacitive sensor, Proximity Sensors, Capacitive-based sensors
	Softness or Hardness	6-axis force sensor, Net-structure Proximity Sensor
Shape	Size and Shape	Ultrasonic sensor, Image sensor, Electric field sensing, Range image sensor
Properties	Roughness	Touch sensor, Limit switches, 6-axis force sensor
	Magnetic Properties	Magneto- resistive sensors
	Metallic/Non -Metallic	Capacity Proximity sensor, 3-wire DC Inductive Proximity Sensor
	Weight	Accelerometers, Slip Sensor, 6-axis force sensor, Joint-angle sensors

Figure 3.5: Vision Camera, Model- scA640-70gc (Make: BASLER)

E. Frame Grabber

Normally it is used as an element of vision system. This system helps to capture digital still frames from an analog video signal or a digital video stream. Aiding to this, it helps to store, transmit and display the captured frames. It is used as expansion cards which were used to interface cameras to computer system. The interfaces such as USB, Ethernet and IEEE 1394 ("FireWire") are used to connect the frame grabber to system. Additional operations may be accomplished as well, such as interlacing, image transformations and real time compression.

Frame grabbers of digital type accept and process digital video streams which are digital video decoder such as Camera Link, CoaXPress, DVI, GigE Vision, LVDS and RS-422. GigE vision is used in the setup which is a standard interface introduced in 2006 for high performance of industrial cameras. The GigE vision makes a framework available for transmitting high-speed video and related control data over Ethernet networks.

Figure 3.6 Frame Grabber

F. Sample Parts

The sample parts used for autonomous part assembly system development are shown in Figure 3.7, 3.8 and 3.9. These parts are considered for evaluation of algorithms developed. Some specific parts in Figure 3.10 and Figure 3.11 are also used for performance evaluation of some methods. The specification of each part is given in Table 3.6.

Figure 3.7: Part 1 (Metallic)

Figure 3.8: Part 2 (Metallic)

Figure 3.9: Part 3 (Metallic)

Figure 3.10: Part1, Part2 and Part3 after assembly (Part4)

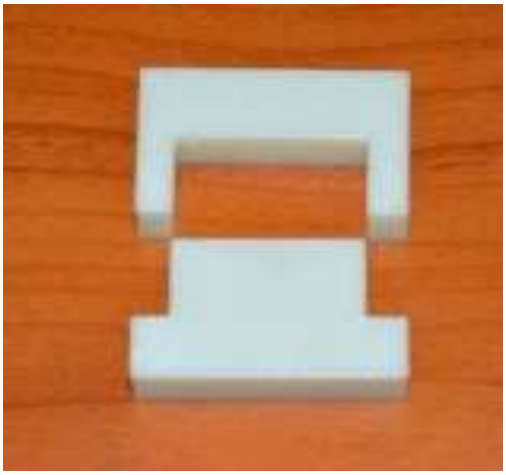


Figure 3.11: 3D Printed Parts (Part 5)

Table 3.6: Specification of Sample Parts

Sl.No.	Name	Material	Size	Single/Assembled
1	Part1	Metallic	39×39	Single
2	Part2	Metallic	33.21	Single
3	Part3	Metallic	6×15	Single
4	Part4	Metallic	34.95×22	Assembled
5	Part5	Plastic	35×19.70	Single

3.2.2 Software

The successful development of the proposed system in this thesis also depends on the use of best suited software tools. The software environment used for development of a vision based robotic system are Laboratory Virtual Instrument Engineering Workbench (LabVIEW, Make: National Instruments, Version 2013), Digi-Matrix Kawasaki Robot Library (Make: DigiMetrix GmbH, Version: v0.2.0.59) and MATLAB (Math Works, Version: 2012, 2015).

i) LabVIEW programming environment involves in computer based measurement and data acquisition. LabVIEW is commonly used in the field of data acquisition, instrument control and industrial automation. It also provides easy-to-use construction of graphical user interface. Several add-ons are available with this programming environment. This includes organization of code to standalone devices, analysis and

processing of signals. Aiding to these, control, simulation, system analysis, report creation and database connection are also embedded in this platform.

ii) The Digi-Metrix Kawasaki Robotics Library [http://sine.ni.com/nips/cds/view/p/lang/en/nid/211069/overview] is used to integrate robot controller with machine vision system through the LABVIEW programming environment. This enables in development of applications for automated test, laboratory automation, and automated assembly setup. This library eliminates the need for expertise in complex robotics programming and helps in controlling and commanding robots directly from a graphical user environment. The add-ons present in this library supports the user to perform several phases of machine control and automation i.e., from part handling and robot control to advanced measurements, inspection etc. by a LABVIEW program.

iii) Initially, the methods involved in realizing the system proposed are programmed in MATLAB. This software provides a suitable environment for multi-paradigm numerical computing environment. The efficiency of the algorithms developed are determined by this software environment.

3.3 Methods

The proposed tasks will be designed theoretically, modeled and simulated virtually using appropriate tools so far as the integration and control techniques are concerned, and finally the system will be implemented through hardware to realize the system to carry out the desired tasks. The proposed methodology is as follows:

- Development of Vision System
  - Selection of hardware
  - Interfacing the Vision System with other devices
  - Algorithm development for part detection, identification, & feature extraction.
- Motion planner development for the industrial robot
  - Development of algorithms for trajectory generation
  - Development of algorithms for motion coordination with respect to vision data
- Development of an assembly work cell with part feeding in the automation laboratory.
  - Part dispensing
  - Part feeding
  - Part picking and assembly station

- Integration of Vision System with the robot control system.
  - Selection of configuration of the embedded controller
  - Interfacing with other ancillary devices
- Modeling and Simulation in virtual environment
- Implementation of vision guided part assembly system for realizing assembly tasks and inspecting the assembly.
  - Motion programming development
  - Incorporation of vision data in the motion programming
  - Online monitoring of the assembly task

The activities involved in this thesis work are part detection, part measurement, and part recognition and navigation control. Several available methods like edge detection, shape descriptor, region growing and corner detection etc. are used.

3.3.1 Part Detection

Proximity sensor [https://en.wikipedia.org/wiki/Proximity_sensor] is used to sense the presence of neighboring objects without any physical contact. Sound waves are utilized for the detection of objects. So color and transparency is not a problem which makes it an ideal equipment to use in several applications including the long range detection. Inductive proximity sensor can be applied for close range detection of ferrous materials like iron, steel, aluminium and copper etc. Capacitive proximity sensor [http://www.ab.com/en/epub/catalogs/12772/6543185/12041221/12041231/Capacitive-Proximity-Sensing.html] is used for close range non-ferrous materials like liquid, wood, plastic and glass etc. Similarly, photoelectric proximity sensors [https://www.ia.omron.com/support/guide/43/introduction.html] is used for long range small or long range target detection of objects like silicon, plastic, paper and metal etc. Exact non-contact detection of targets is provided by the photoelectric sensors. The sensors radiate infrared or laser light. The radiation falling on the target object is interrupted or reflected back to the sensor to trigger the sensor output. It helps to track various attributes of the target objects such as shape, size, color etc. These sensors outperform other sensing methodologies in terms of operating at a larger distance, furnishing various mounting options and providing greater flexibility.

Ultrasonic sensor [http://sensorwiki.org/doku.php/sensors/ultrasound] contains a sonic transducer that radiates a series of sonic pulses and get pulses in return as it reflects back from reflecting target. The sensor sends signal to control devices after receiving the reflected signal. The sensing

range can be extended up to 2.5 m. The standard diffuse ultrasonic sensors produce a simple present/absent output, some produce analog signals, indicating distance with a 4 to 20 mA which can be converted into distance information. Photoelectric sensors are used to detect targets less than 1 mm in diameter, or from 60 m away.

3.3.2 Part Identification

The parts are recognized with respect to the shape, size and materials from the image captured by vision system through image processing techniques. It is also required to determine the exact position of the target object for robot to grasp. Some of the operations accomplished by a machine vision system in the identification process are recognizing the part, sorting of parts, and gripping the parts oriented from a conveyor.

Recognizing desired part can be termed as a labelling problem based on known object models. For example, in an automated assembly environment, the conveyor belt may contain one or more similar or different objects. To recognize the object of interest, a set of labels corresponding to known object models, the system must be capable of assigning the correct labels to the object or region in the image taken by system. In the process of recognizing objects, feature extraction methods plays an important role. Feature extraction techniques provide either a single or group of features of an image. Typically, the extracted features may be a scalar quantity providing the area or aspect ratio of an image. Similarly it can be in vector format such as coordinate or texture information of an object etc. Feature of the region of interest or the object present in the captured image is extracted. There are several feature extraction techniques like edge map, corner point detection, color histograms, mathematical morphology, moment based and connected components labelling etc. used in the subsequent chapters. These features are used to compare with the features of the stored object models. The object recognition is fully dependent on the efficiency of the feature extraction methods. Feature extraction of image is an indispensable principal step in robot control application.

Several classes of features are

- Line features are generally a set of curved line segments named as edges. These segments appear in region of image where the abrupt change of intensity values or brightness occurs. On the other hand, point features are corner points which exist at the intersection of two edge segments. Hence, at a corner point there exist two dominant and different edge directions. For extraction of such features, available edge

detection techniques like Canny edge, Sobel, Prewitt and Robert etc. are compared and a hybrid edge detection algorithm, considering fuzzy rules along with wavelet transform procedure, is developed.

- As defined above the corner points lie at the intersection of two edges. Co-variance measure of the change in intensity values at each pixel is used to detect a corner point. Corner detection methods like Harris/Plessey, Curvature Scale Space and Wang-Brady [https://en.wikipedia.org/wiki/Corner_detection] are implemented to find the corner points and the efficient detector is considered for part identification.
- Region of interest features are connecting group of pixels that are identical with respect to some constraint. Region growing technique is used to identify the exact part in an image.

3.3.3 Part Measurement

Perfect measurement of part features is helpful in accurate identification of parts and its grasping point determination. The part detection and identification is performed by extraction of features like distance and size measurement of the desired object. The distance between the robot and desired object is measured by coordinate transformation method.

Factors like boundary, area, region of interest point and center of gravity are measured for making the part recognition process autonomous. Once the correct part is identified, then the robot has to pick it for placing in desired place. For this, grasping points as points of interest are determined. These measurements are performed by mathematical morphology and moment based shape descriptor. The descriptors are some representative framework to outline a predefined shape. Accurate reconstruction of object or shape may not be possible from the descriptor. But the descriptor of different shapes must be distinguishable from each other. For tracking of a desired part in an assembly environment, the accurate measurement of part is essential.

3.3.4 Stable Grasping

A large variety of objects in an assembly work cell are manipulated by robots. Manipulation of objects is to establish a physical connection between the robot end-effector and the concerned object. In our context, this physical connection is a robotic grasp. Effectiveness of the grasping technique depends grasp planning algorithms and the shape of the object. Stable grasping is to transport an object from one place to another without changing the orientation.

3.3.5 Robot Navigation and Control

Vision-based robot control is used to control the motion of a robot and navigate it by using extracted feature information from a vision sensor. Tracking of the part from feeder to final product is carried out by vision system. Image based and position based servoing techniques are used for navigating the robot end effector to the desire position. With the help of visual data, the navigational control is used for avoiding obstacles and automatic path planning of a robot.

3.4 Summary

This chapter includes the significant materials required for the development of vision system based part assembly system. The description about the necessary accessories used for development of a vision guidance system is given. The important methods required for the successful integration of vision guidance system with a robotic system are discussed in brief.

Chapter 4

Feature Extraction and Object Detection

4.1 Overview

The present-day trend in industrial automation are intensely appealed towards flexible automation with robotic system as a major driver. The use of robotic system offer major role in handling the long run operating costs, quality improvement of final product with a growth in productivity. One of the best approach in developing automated flexible assembly system is the integration of the machine vision technology with the robotic platform. The competitiveness of industrial houses to implement assembly tasks, particularly in unstructured environment requires advanced sensor-based, intelligent systems to take care of the operations especially under uncertainties.

Vision system plays an important role in building an autonomous assembly system. The important tasks like detection and recognition of the correct part that need to be carefully handled by a MV system to initiate the process. The process of assembly comprises of several sub-processes wherein the image acquisition, analysing and increasing the image representation quality etc. do account for successful operation of the system and completion of the desired tasks. This chapter contains the feature extraction techniques and the proposed methods of object recognition in part assembly process.

4.2 Object Detection

Object detection is the method for finding and identifying objects in an image or video sequence. Human visual system can recognize multiple objects in images simultaneously with little effort, even though the image of the objects might vary to some extent in different viewpoints, different sizes and scales or may be affected by different image transformations. Objects can even be recognized when they are partially obstructed from view. This task is still a challenge for vision systems. Many approaches to the task have been implemented over multiple decades. The general object detection model system is given in the Figure 4.1. In this, the first step involves in application of image processing techniques for extraction of image features. There are several feature extraction techniques available like, edge detection, corner detection and moment invariant methods. In this chapter several available techniques are discussed along with the proposed methods. Experimental results along with quantitative analysis are presented to validate the outperforming behavior of the suggested methods.

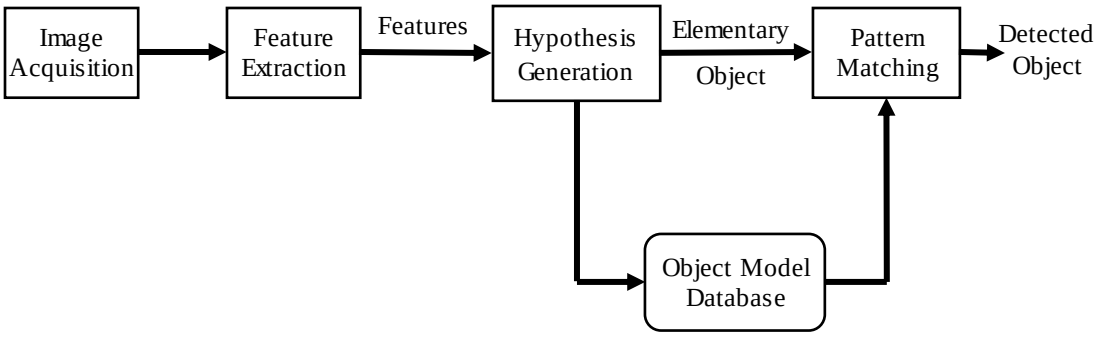


Figure 4.1: General model of an object detection system

The problem of object recognition is defined as a labelling problem which is based on known object models. Considering a case where an image containing one or more objects of interest (and background) and a set of labels agreeing to a set of models known to the system are given, the detection system has to assign correct labels to regions, or a set of regions, in the image. This problem is related to the segmentation problem i.e. segmentation is possible only when partial recognition of objects is performed. Similarly, recognition is fully dependent on segmentation.

As shown in Figure 4.1, the general model of an object detection system comprising of image acquisition, feature extraction technique, generation of hypothesis, object model database and a matching unit. The details of image acquisition are discussed in Chapter 5. The database comprises of all known object models. The information that describes the object in the database depends on the recognition method. The object representation scheme can be different from a qualitative or functional description to precise geometric representation information. A feature is an important characteristic of the object which helps in describing and recognizing the object with respect to other objects. Size, color, and shape are some commonly used features.

Feature detection techniques are applied on images to extract features and assist in establishing object hypotheses. The type of feature depends on the object type involved in recognition process. The hypothesis generator assigns likelihoods to each object based on the extracted features. This step reduces the search space for the recognizer. The object models are organized by using indexing scheme to facilitate the exclusion of dubious object candidates from likely consideration. The pattern matching block, that performs the matching procedure, compares the stored model information with the hypotheses and calculates the likelihood of the object. The system detects the

object with the highest likelihood as the accurate object. All object recognition systems use models either explicitly or implicitly and employ feature detectors based on object models used in the process. Pattern classification approaches are a good example of this approach.

An object recognition system must select appropriate tools and techniques for the steps discussed above. Many factors must be considered in the selection of appropriate methods for a particular application. The central issues that should be considered in designing an object recognition system are:

- Object representation: This plays an important role in object detection process. The representation of object must be efficient enough that the important attributes or features should be present in the model. It is the case that, geometric representation is suitable for some objects, while, generic or function features are more important for some other type of objects. Hence, object representation method must include all relevant and non-redundant information of objects.
- Feature extraction: The use of appropriate feature extraction technique makes the object detection process efficient. The selection of feature extraction method depends on the type of feature to be extracted. Most features can be extracted in two dimensional images but they are related to three-dimensional characteristics of objects. Due to the nature of the image formation process, some features are easy to compute reliably while others are very difficult. Feature detection issues were discussed in the following sections of this chapter.
- Hypotheses generation: The hypothesis generation step is basically a heuristic to reduce the size of the search space. This step uses knowledge of the object model to assign some kind of probability or confidence measure to different objects present in the image. This measure reflects the likelihood of the presence of objects based on the detected features. Hypothesis generation step is required for such application environment with large number of objects.
- Object Matching: The presence of each expected object can be verified by using the stored models.

4.3 Features Extraction Technique

The proposed robotic assembly system under vision guidance must be capable of

- Recognition of the object(s) of interest by the extraction of important features,
- determining the object location in the workspace, and
- determining the points of grasping

The recognition of objects, localization of objects and determination of coordinates of grasping points are performed by the MV system. The vision system performs the task of recognition and localization by the use of features extracted where detection of edge map of objects forms the first step. It is essential to use a suitable and efficient edge detection technique. The performance of edge detection techniques depends on factors like focal length of the camera, the distance between camera and object, the threshold value and the illumination type.

In part assembly application, part identification is an important task. This is fully dependent on the efficiency of feature extraction techniques and the algorithms involved in measurement of parts or objects. Accurate measurement is helpful in precise identification of parts and its grasping point determination. The intelligent vision system is the standard system for identification and position recognition. It provides information about the part, its position and orientation that are placed on a specific plane.

For object detection, several features such as edge, corner and shape are extracted from the object present in an image. Most of the features are related to regions or boundaries of the image. A region is a closed boundary that corresponds to the entire object or a part of it. Broadly, the feature of an image is classified as local features, global features and relational features.

Global Features

Global features [Zhang and Lu, 2004; Nagabhushana, 2005; Singh and Sharma, 2013] are the regions in image such as perimeter, area (size), Fourier descriptors or moments. These features can be determined either by considering all points within a region, or considering only points on the boundary of the region. In each case, the main aim is to find descriptors that are obtained by considering all points, their locations, intensity characteristics, and spatial relations.

Local Features

Local features [Zhang and Lu, 2004; Nagabhushana, 2005] are generally present on the boundary line in an image which represents a distinct area of a region. Curvature and related properties are commonly used as local features. The curvature is computed on a boundary or a surface. High

curvature points are commonly called corners and play an important role in object recognition. Specific shapes of a small boundary segment or a surface patch constitutes the local features. Boundary segments, corners and curvature are commonly used local features. A comparative study on corner detection methods are discussed in this chapter.

Relational Features

Relational features [Zhang and Lu, 2004; Nagabhushana, 2005] are related to the relative positions of different entities like regions, closed contours and also the local features. These features are the distance between local features and the relative orientation. This type of feature are helpful in representing complex objects which are described by several regions or set of local features.

4.3.1 Edge Detection Technique

An edge is defined as a set of connected pixels present on the boundary that separates two or more regions [Gonzalez and Woods, 2007; Park and Murphey, 2008]. Detection of edges is a set of mathematical models for identifying points in an image where intensity changes sharply or has discontinuity. In image processing, edge detection is defined as a process which concerns with the localization of significant variations of the grey level images and the determination of physical features that originated this. This plays an important role in recognition, image registration and enhancement etc. The techniques involved in edge detection are differentiation and smoothing. Smoothing yields a loss of information and differentiation is an ill-conditioned problem [Ziou and Tabbone, 1998]. The object recognition method by a vision system comprises of two processes such as edge detection and pixel localization.

This research work is intended to develop an appropriate edge detection algorithm that would be convenient for the purpose and can produce better result for detection and identification of correct object(s). The general process of object detection is shown in Figure 4.2.

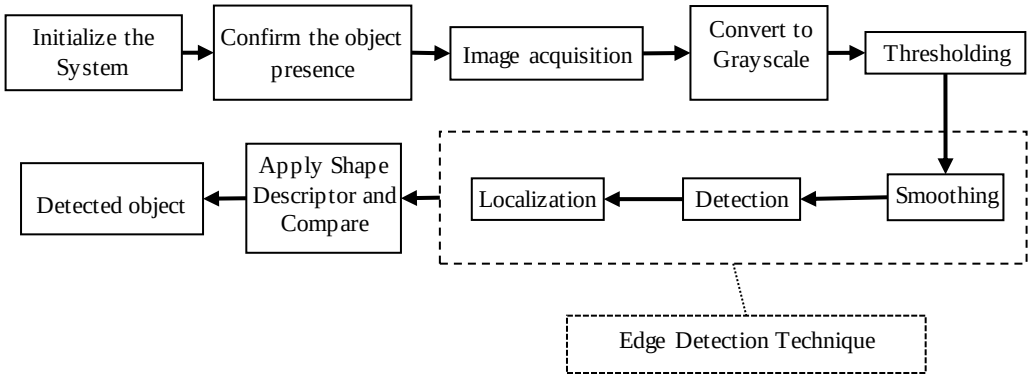


Figure 4.2: Block diagram for object identification [Balabantaray et al., 2014]

The edge map of an object in an image is determined by considering the association of a pixel with its neighbor pixels. If the gray value of a pixel is similar to that of the neighbor pixels, then it is considered that no edge point is present. If the gray levels of the neighbor pixels have a sharp variation, then that pixels is considered to be a point in the edge. The model of an edge is constructed by considering the relationship between the brightness and spatial coordinate of pixels and is shown in Figure 4.3. It is observed from this model that an ideal edge is a set of connected pixels (in vertical direction) which are located at the orthogonal step shift in gray level. This is a measure of the edge points.

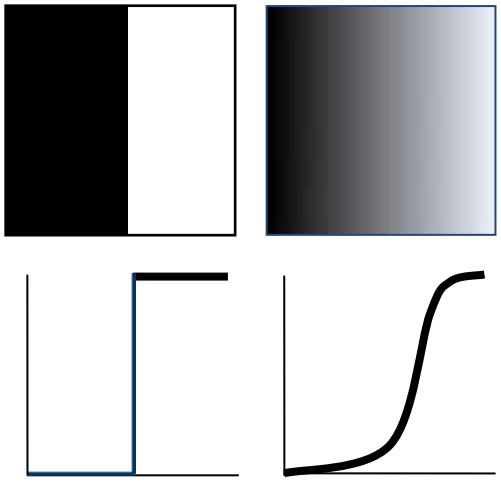


Figure 4.3: Model of an edge with respect to brightness versus spatial coordinates [Gonzalez and Woods, 2007; Balabantaray et al., 2014]

The four steps of edge detection [Gonzalez and Woods, 2007; https:// www.cse.unr.edu /~bebis/ CS791E / Notes/ EdgeDetection.pdf] are as follows:

Smoothing: It is the process suppressing noise without disturbing the true edges.

Enhancement: The quality of an image is enhanced by the application of image enhancement techniques (sharpening).

Detection: This step involves in detection of true edge pixels by eliminating the falsely detected edges. This is achieved by using thresholding technique.

Localization: This step involves in determining the location information of an edge pixel. Edge thinning and linking are the methods involved in localization of edge pixel.

Edge detection is an important problem in image processing as edge represents the boundary of an object. The regions with high intensity value are considered as edge i.e., a jump in intensity value from one pixel to the next determines presence of the edge [Ziou and Tabbone, 1998]. Edge detection technique considerably reduces the amount of data required for representation of an object in an image. It also eliminates the useless information by conserving the structural properties in an image. However, the different edge detection methods are divided into two categories i.e., gradient based edge detection and Laplacian based edge detection [Juneja and Sandhu, 2009; Maini and Aggarwal, 2009; Shrivakshan and Chandrasekar, 2012; http: // www.owlnet.rice.edu/ ~elec539 / Projects97/ morphjrks/ moredge.html]. The gradient based method finds the edges by considering the maximum and minimum values of first derivative of the image. In Laplacian based method the edges are determined by searching the presence of zero crossings in the second derivative. Edge is represented as 1-D shape of a ramp and the location of edge pixel is highlighted by taking derivative of image. The changes in the intensity can be observed where there is an edge and is shown in Figure 4.4. If there is a presence of gradient in an image (i.e., the first derivative), the curve is parabolic in nature as shown in Figure 4.5.

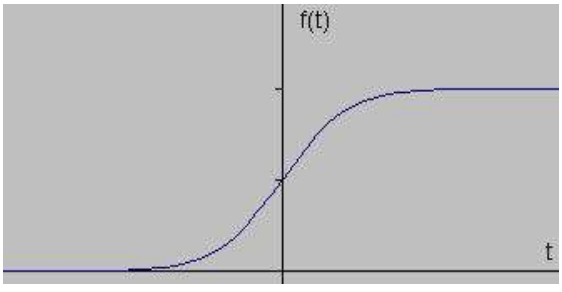


Figure 4.4: Intensity variations for an edge pixel [Gonzalez and Woods, 2007; Balabantaray et al., 2014]

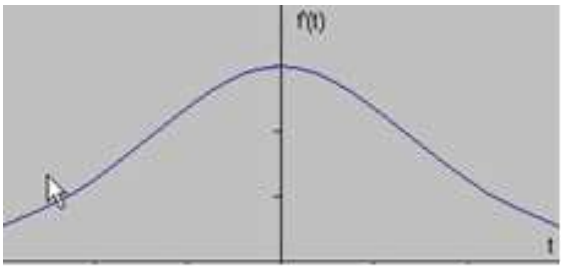


Figure 4.5: First derivative of the edge [Gonzalez and Woods, 2007; Balabantaray et al., 2014]

The first order derivative indicates the presence of maxima located at the center and considered as edge. This filtering process in called ‘‘gradient filter’’ which is a widely used filter for edge detection. A pixel is considered to be an edge pixels if the gradient value is higher than the threshold value. It is obvious that edge pixels have a higher intensity value as compared to the neighboring pixels. Therefore, a threshold is set, detection of edge pixel is performed by comparing the gradient value with the threshold value. Furthermore, edge pixels are identified from the first order derivative of the image where the edge pixel attains a maximum value. Similarly, the edge pixels are identified by the detection of pixels that passes through zero (i.e., zero crossing) in case of second derivative as shown in Figure 4.6. This technique of edge detection is called Laplacian based edge detection.

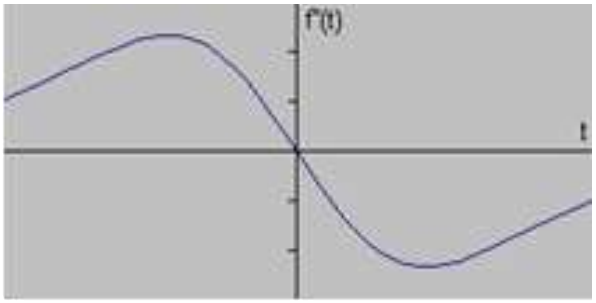


Figure 4.6: Second derivative of the edge [Gonzalez and Woods, 2007; Balabantaray et al., 2014]

The important consideration in edge detection is that, no true edge pixels are missed and false edge pixels are considered. Another consideration is proper localization of edge pixels i.e., the distance between the detected edge pixels and the true edge pixels must be minimum. The third condition is to have single response to an edge pixel.

General formulation for edge detection using masks for an image $f(x,y)$ with a gradient vector $G(x,y)$ is as follows

(i) Suppress the noise by applying the smoothing process to the input image,

$$\hat{f}(x,y) = f(x,y) * G(x,y)$$

(ii) $$\hat{f}_x = \hat{f}(x,y) * G_x(x,y)$$

(iii) $$\hat{f}_y = \hat{f}(x,y) * G_y(x,y)$$

(iv) $$magn(x,y) = \left|\hat{f}_x\right| + \left|\hat{f}_y\right|$$

(v) $$dir(x,y) = \tan^{-1}\left(\hat{f}_y \Big/ \hat{f}_x\right)$$

(vi) If $magn(x,y) > T$, then there is an edge point

A. Sobel Operator

The Sobel operator (sometimes called the Sobel–Feldman operator) performs a 2-D spatial gradient measurement on an image. This edge detection technique, a discrete differentiation operator, involves in estimating the gradient of the intensity function. This method results in a gradient vector or norm of the gradient vector for each points in the image. It considers two distinct 3×3 masks for convolution of the image. These masks are responsible for determining the gradient vector in x-direction and y-direction separately. The pair of masks are relatively smaller than the actual image and can able to produce an integer value. These masks scan the image and manipulate a square of pixels simultaneously [Sharifi et al., 2002, Aybar, 2006]. The masks are defined in Eq. 4.1.

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} \text{ and } G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \tag{4.1}$$

where, G_x is the gradient vector in x-direction and G_y is the vector in y-directions.

The value of the gradient is determined as:

$$|G| = \sqrt{G_x^2 + G_y^2}$$

The approximate gradient vector is performed by:

$$|G| = |G_x| + |G_y|$$

The spatial gradient value, that is the orientation of edge pixel, is calculated by Eq. 4.2:

$$\theta = \arctan\left(G_y / G_x\right) \tag{4.2}$$

Here, θ describes the direction of maximum variation of contrast from black to white in clockwise movement. Also, the orientation in anticlockwise is calculated. The pseudo-convolution operator is used to calculate the absolute value of gradient vector by adding the two values, which is used to scan the whole image. The pseudo-convolution operator is given in Figure 4.7.

P_1	P_2	P_3
P_4	P_5	P_6
P_7	P_8	P_9

Figure 4.7: Mask used in Pseudo-convolution operation

Now, the approximate value of the gradient vector is determined by Eq. 4.3 considering the pseudo-convolution mask.

$$\begin{aligned} |G| = &\left|\left(P_1 + 2\times P_2 + P_3\right) - \left(P_7 + 2\times P_8 + P_9\right)\right| + \\ &\left|\left(P_3 + 2\times P_6 + P_9\right) - \left(P_1 + 2\times P_4 + P_7\right)\right| \end{aligned} \tag{4.3}$$

B. Robert Cross Operator

Edge detection based on Robert cross operator [Senthilkumaran and Rajesh, 2009; Bhadauria et al., 2013; Balabantaray et al., 2014] is a differential operator that involves in approximation of the gradient vector by using a discrete differentiation process. The approximate gradient value is determined by the sum of the squares of the intensity variations in pixels that are adjacent diagonally. The edge detection is performed by applying convolution to the original image. It considers two 2×2 masks and they are calculated by the formula given in Eq. 4.4. One of the mask is the 90° rotation of the other. This method is similar to Sobel operator.

$$G_x = \begin{bmatrix} +1 & 0 \\ 0 & -1 \end{bmatrix} \text{ and } G_y = \begin{bmatrix} 0 & +1 \\ -1 & 0 \end{bmatrix} \tag{4.4}$$

These masks properly respond to edges at 45° orientation to the pixel grid. The masks yield separate gradient values in x- and y-direction. The absolute gradient value is determined by adding the gradient values in both directions and is given by

$$|G| = \sqrt{G_x^2 + G_y^2}$$

The approximate gradient value can be computed as:

$$|G| = |G_x| + |G_y|$$

The spatial gradient value is calculated by considering the angle of orientation as in Eq. 4.5.

$$\theta = \arctan\left(G_x / G_y\right) - 3\pi/4 \tag{4.5}$$

The pseudo-convolution operator, used in this edge detection technique, is the addition of the absolute gradient magnitude in both direction and is given in Figure 4.8

$\mathbf{P_1}$	$\mathbf{P_2}$
$\mathbf{P_3}$	$\mathbf{P_4}$

Figure 4.8: Pseudo-convolution operator used in Robert edge detection technique

The approximate magnitude is determined by considering the pseudo-convolution operator as given in Eq. 4.6

$$|G| = |P_1 - P_4| + |P_2 - P_3| \tag{4.6}$$

C. Prewitt operator

This operator is very similar to the Sobel operator. This operator is used for detection of horizontal and vertical edges. [Senthilkumaran and Rajesh, 2009; Bhadauria et al., 2013; Balabantaray et al., 2014]. The convolution masks used in Prewitt operator are given by Eq. 4.7.

$$G_x = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \text{ and } G_y = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \tag{4.7}$$

D. Laplacian of Gaussian Operator

The Laplacian of Gaussian (LoG) operator [Senthilkumaran and Rajesh, 2009; Bhadauria et al., 2013; Balabantaray et al., 2014] based edge detection finds the edge points by searching the zero crossings of the second derivative (as in Figure 4.6) of the intensity of the image. However, the second derivative is very sensitive to noise and it is removed by use of Gaussian filter.

The noise is eliminated by applying the convolution operator with the Gaussian filter. Consequently, noise and small structures are isolated and then removed. Pixels with zero crossing in the second order derivative and gradient value with a local maxima, are reflected in the edge map as an edge pixel. The orientation of edge is calculated with respect to the direction of pixels with zero crossing.

The result of LoG operator is denoted as $h(x, y)$ and defined as

$$\begin{aligned} h(x, y) &= \Delta^2 \left[g(x, y) * f(x, y) \right] \\ &= \left[\Delta^2 g(x, y) \right] * f(x, y) \end{aligned} \tag{4.8}$$

Where, $\Delta^2 G(x, y) = \left(\frac{x^2 + y^2 - 2\sigma^2}{\sigma^4} \right)^{-\left(x^2+y^2\right)/2\sigma^2}$ is commonly called the Mexican hat operator.

E. Canny Edge Detection Algorithm

The Canny operator detects a wide range of edges by using a multi-stage algorithm [J. Canny, 1986]. In this method, smoothing process is applied to the original image by Gaussian filter. The regions with higher intensity value are highlighted as a result of first order derivative of the smoothed image. The pixels with higher intensity value is considered as ridges in the gradient magnitude image. Non-maximal suppression is performed by tracking each pixels and assigning zero to those pixels that are not overlapped with the ridge top. The tracking process is controlled by two threshold values: T_1 and T_2 with $T_1 > T_2$. The tracking process is started only when the threshold value at a point on the ridge is greater than T_1 and continued in both horizontal and vertical directions till the value falls under T_2. This ensures that the maximum suppression of noise information. Several steps are followed to implement the Canny edge detection technique and they are as follows [Ali and Clausi, 2001; Green, 2002].

Step 1: This step involves in removal of noise information present in the original image by employing Gaussian filter. The Gaussian filter is easy to implement and fast. Smoothing process is executed through a convolution method with a smaller mask. The larger the width of the mask, the lower is the noise sensitivity and increase in localization error.

Step 2: In this step, 2-D spatial gradient value of the image is determined to find out the edge. The absolute gradient value is the measure of the strength of the edge and is determined at each point. A pair of 3x3 masks is used in convolution operator for estimating approximate gradient value in both x- and y- direction.

The gradients G_x and G_y are calculated as

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} \text{ and } G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & +2 \end{bmatrix} \tag{4.9}$$

The magnitude of the gradient is then approximated using the formula:

$$|G| = |G_x| + |G_y|$$

Step 3: In this step, the orientation of the edge pixels is determined by the known gradient values for both direction which is a trivial solution. It produces an error as when the sum of values in x-direction is equal to zero. This problem is handled by setting a restriction. When the sum of the gradient in x-direction is equal to zero, the edge orientation is considered to be 90° or 0° with respect to the gradient value in y-direction. If G_y has a value of zero, the direction of edge is equal 0°, else 90°. This is calculates by the formula below.

$$\theta = \arctan\left(G_x / G_y\right)$$

Step 4: The direction of edge is mapped to a direction that is traced in an image with respect a 3×3 image tracing map as shown in Figure 4.9.

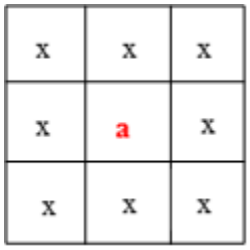


Figure 4.9: Tracing of edge direction in a 3×3 image

Considering the red marked pixel "a" for which four likely directions are available. They are considered to be 0° for pixels in horizontal direction, 45° pixels along positive diagonal, 90° for pixels in vertical direction, or 135° for pixels along the negative diagonal. So, the edge orientation has to be resolved into one of these four directions depending on which direction is close to the edge. Suppose the angle of orientation is 3°, then consider edge direction is 0°. The edge direction for all possible direction is shown in Figure 4.10.

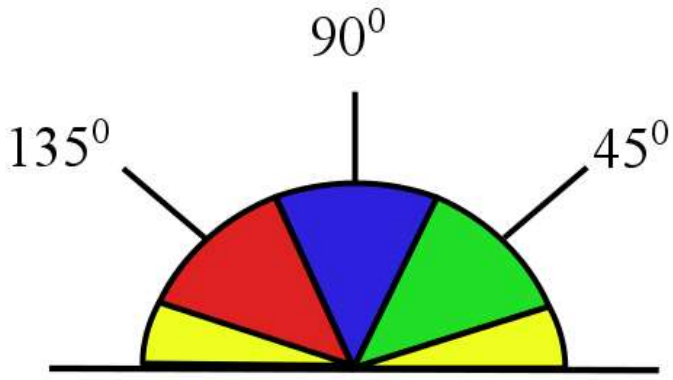


Figure 4.10: Edge orientation with respect to four corresponding neighbors [Gonzalez and Woods, 2007; Balabantaray et al., 2014]

Hence, direction of edge falls inside the yellow colored range i.e., 0° to 22.5° and 157.5° to 180° is considered as 0°. If the direction falls in the green range 22.5° to 67.5°, then edge orientation is 45°. Similarly for blue colored range edge direction is considered as 90° (Range: 67.5° to 112.5°) and for the red colored region, it is 135° (Range: 112.5° to 157.5°.

Step 5: This step involves an edge refinement procedure. The non-maximum suppression method is used by tracing the edge along the edge direction and suppress the falsely detected edge pixel. This step results a thin line in the image.

Step 6: Now, the streaking is removed by using a hysteresis method. Streaking occurs due to the fluctuating behavior of the operator which breaks the edge contour. For a single threshold value T_1, the edge falls below the set threshold due to the effect of noise. The edge above the set threshold value will result an edge as a dashed line. This is avoided by the use of hysteresis method where two threshold values are used. Pixel with value greater than T_1 is accepted as an edge pixel. Similarly, pixels that are connected to the detected edge pixel and have a value greater than T_2 are also nominated as edge pixel.

F. Wavelet Based Edge Detection

The wavelet transform acts as a zooming method in signal processing. It is used in image processing for its good behavior both time domain and frequency domain. The Fourier transform was the main mathematical tool for analyzing edges in images before the Wavelet technique. But, The Fourier transform of a signal is an independent function of location, which only provides a description of the overall regularity of signal. Since Wavelet transform can characterize and localize both time and frequency, it has been introduced for the detection and analysis of the edges. [Mallat and Hwang, 1992] mathematically explained that signals and noise had different singularities and edge structures present observable magnitudes along the scales. So analyzing an image at different scales increases the accuracy and reliability of the edge detection. Because of having this ability, Wavelet transform is an advantageous option for the image edge detection. [Brannock and Weeks, 2008] developed an edge-detection method based on the discrete wavelet transform. [Guan, 2008] has proposed a multiscale Wavelet edge detection for the lip segmentation. By using edge Wavelet transform, the edges can be detected more accurately than the other methods based on differentiation. The process of wavelet based edge detection [Xue and Pan, 2009; Weifeng and Caixia, 2012] is given in Figure 4.11.

Figure 4.11: Wavelet based edge detection process

Let us consider the mathematical background of the traditional wavelet transformation [Li, 2003; Shih and Tseng 2005] on edge detection.

For an arbitrary function $f(x,y) \in L^2(R^2)$, its Fourier transform is

$$\hat{f}(\omega,\xi) = \int_R \int_R f(x,y) e^{-(\omega x+\xi y)dxdy} \tag{4.10}$$

For two functions $f(x,y), g(x,y) \epsilon L^2(R^2)$, their convolution is

$$f * g(x,y) = \hat{f}(\omega,\xi) = \int_R \int_R f(u,v) g(x-u, y-v) dudv \tag{4.11}$$

Transforming these two into one wavelet function $\psi \epsilon L^2(R^2)$ with a condition that if it meets the following condition

$$C_\psi = (2\pi)^2 \int_R \int_R \frac{|\hat{\psi}(\omega+\xi)|^2}{\omega^2+\xi^2} d\omega d\xi \ < +\infty \tag{4.12}$$

If $\psi \epsilon L^1(R^2) \cap L^2(R^2)$ then, $\hat{\psi}(0,0) = 0$, that is $\int_R \int_R \psi(x,y) dx dy = 0$.

The continuous form of the wavelet function ψ can be expressed as

$$W_{\psi_1} f(x,y) = (f * \psi_S)(x,y) = \frac{1}{S^2} \int_R \int_R f(u,v) \psi\left(\frac{x-u}{S}, \frac{y-v}{S}\right) du dv \tag{4.13}$$

Here, S is the scaling function and the wavelet is expressed as

$$\psi_S(u,v) = \frac{1}{S^2} \psi\left(\frac{u}{S}, \frac{v}{S}\right) \tag{4.14}$$

Suppose $\theta(x,y)$ is the smooth function and it meets the following conditions

$$\begin{cases} \iint_R \theta(x,y) dx dy = 1 \\ \lim\limits_{\substack{x \to \infty \\ y \to \infty}} \theta(x,y) = 0 \end{cases} \tag{4.15}$$

The two dimensions of the wavelet function ψ can be defined as

$$\psi^1(x,y) = \frac{\partial \theta(x,y)}{\partial x} \text{ and } \psi^2(x,y) = \frac{\partial \theta(x,y)}{\partial y} \tag{4.16}$$

By considering the scaling factor $S = 2^j$, the two dimensions of wavelet can be computed by

$$\begin{cases} \psi^1_{2^j}(x,y) = \frac{1}{2^{2^j}} \psi^1\left(\frac{x}{2^j}, \frac{y}{2^j}\right) \\ \psi^2_{2^j}(x,y) = \frac{1}{2^{2^j}} \psi^2\left(\frac{x}{2^j}, \frac{y}{2^j}\right) \end{cases} \tag{4.17}$$

Eq.4.13 can be expressed as

$$W_{2^j}^1 f(x,y) = f * \psi_{2^j}^1(x,y)$$
$$= \left(f * 2^j \frac{\partial}{\partial x}\theta_{2^j}(x,y)\right)$$
$$= \left(2^j \frac{\partial}{\partial x}(f * \theta_{2^j})(x,y)\right) \quad (4.18)$$

And

$$\mathrm{W}_{2^j}^2 \mathrm{f}(\mathrm{x},\mathrm{y}) = \mathrm{f} * \psi_{2^j}^2(\mathrm{x},\mathrm{y})$$
$$= \left(\mathrm{f} * 2^j \frac{\partial}{\partial \mathrm{y}}\theta_{2^j}(\mathrm{x},\mathrm{y})\right)$$
$$= \left(2^j \frac{\partial}{\partial y}(f * \theta_{2^j})(x,y)\right) \quad (4.19)$$

Eq. 4.18 indicates that $W_{2^j}^1 f(x,y)$, is a partial derivative of the smooth function in the horizontal direction. Similarly, $W_{2^j}^2 f(x,y)$ in Eq. 4.19 is the partial derivative in the vertical direction. The value of the function changes quickly along the gradient direction for a binary function. So, the gradient reach at a maximum value in the gradient direction. For the image, the smoothed gradient function can be defined as

$$grad(f * \theta_{2^j}) = \begin{bmatrix} W_{2^j}^1 f(x,y) \\ W_{2^j}^2 f(x,y) \end{bmatrix} = 2^j \begin{bmatrix} \frac{\partial}{\partial x}(f * \theta_{2^j})(x,y) \\ \frac{\partial}{\partial y}(f * \theta_{2^j})(x,y) \end{bmatrix} \quad (4.20)$$

It can be defined that $\alpha(x,y) = arctan\left(\frac{W_{2^j}^1 f(x,y)}{W_{2^j}^2 f(x,y)}\right)$ and the gradient direction is determined by

$$\left|grad(f * \theta_{2^j})\right| = \frac{1}{2^j}\sqrt{\left|(f * \psi_{2^j}^1)(x,y)\right|^2 + \left|(f * \psi_{2^j}^2)(x,y)\right|^2}$$
$$= \frac{1}{2^j}\sqrt{\left|W_{2^j}^1(x,y)\right|^2 + \left|W_{2^j}^2(x,y)\right|^2} \quad (4.21)$$

In this method of edge detection, the gradient direction of all points in the image is determined then edge points is determined and oriented by the two dimensions of wavelets transforms. The image pixels is considered as edge point whose gradient value is maximum. To achieve the better result, dual threshold value is used. The edge detection is accomplished through applying Canny's edge detector to the image containing only maximum gradient values.

The traditional wavelet transform is then improved by [Lixia et al. 2010] for edge detection. The wavelet transform is expressed as

$$W_{\psi_1} f(x,y) = (f * \psi_S)(x,y) = \iint_R f(u,v)\psi(sx-u, sy-v)dudv \tag{4.22}$$

Here, the wavelet function used is $\psi_S(u,v) = \psi(su, sv)$. The two dimensions of wavelet are defined by considering the scaling factor $S = 2^j$ as

$$\begin{cases} \psi^1_{2^j}(x,y) = \psi^1\left(2^j x, 2^j y\right) \\ \psi^2_{2^j}(x,y) = \psi^2\left(2^j x, 2^j y\right) \end{cases} \tag{4.23}$$

The two dimensional form of wavelet is expressed as

$$\begin{aligned} W^1_{2^j} f(x,y) &= f * \psi^1_{2^j}(x,y) \\ &= \left(f * \frac{\partial}{\partial x}\theta_{2^j}(x,y)\right) \\ &= \frac{\partial}{\partial x}\left(f * \theta_{2^j}\right)(x,y) \end{aligned} \tag{4.24}$$

And

$$\begin{aligned} W^2_{2^j} f(x,y) &= f * \psi^2_{2^j}(x,y) \\ &= \left(f * \frac{\partial}{\partial y}\theta_{2^j}(x,y)\right) \\ &= \frac{\partial}{\partial y}\left(f * \theta_{2^j}\right)(x,y) \end{aligned} \tag{4.25}$$

The gradient function of the image $f(x,y)$ is defined as

$$grad(f * \theta_{2^j}) = \begin{bmatrix} W_{2^j}^1 f(x,y) \\ W_{2^j}^2 f(x,y) \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x}(f * \theta_{2^j})(x,y) \\ \frac{\partial}{\partial y}(f * \theta_{2^j})(x,y) \end{bmatrix} \tag{4.26}$$

$$\begin{aligned} |grad(f * \theta_{2^j})| &= \sqrt{|(f * \psi_{2^j}^1)(x,y)|^2 + |(f * \psi_{2^j}^2)(x,y)|^2} \\ &= \sqrt{|W_{2^j}^1(x,y)|^2 + |W_{2^j}^2(x,y)|^2} \end{aligned} \tag{4.27}$$

This improved wavelet based edge detection algorithm is very effective for detection of edge. The traditional wavelet transform is modified by considering different value for scaling factor in the proposed hybrid edge detection method which is described in the following section.

G. Mathematical Morphology Based Edge Detection

Conventional edge detection techniques are based on first-order gradient or zero-crossing value of second derivative for detecting edges. Maintaining an appropriate balance between the de-noising ability and detection accuracy is the main conflicting issue. Unlike traditional methods, mathematical morphology (MM) based edge detection techniques resolve this conflict by successfully preserving the accurate edges and restraining the noise well. MM based image processing assumes the images as a point set and carried out set operation for processing and identifying. MM based edge detection methods [Shang and Jiang, 2012; Sridhar and Reddy, 2013] are nonlinear in nature and the basic operations are based on set theories such as shift, union, intersection and complementary operations. A definite modal structuring element (SE) is adopted which move persistently in the image to extract and measure the shape in the corresponding image. There are two types of SE, such as non-flat and flat type. Multi-scale SE is also available in literature to extract the image details and to minimize the noises at different scale. Selecting appropriate scale of the structuring element can able to filter out noise and retain the image information simultaneously. These methods enable to extract more smooth edges and are insensitive to noise. These methods efficiently satisfy the real-time requirement and can be easily integrated for hardware implementation.

Let the input image be $f(x,y)$ and $S(x,y)$ be a disk structuring element, then the eroded image subtracted from the dilated image of $f(x,y)$ provides the edge magnitude of the corresponding original image.

Erosion operation of $f(x,y)$ using a non-flat SE $S(x,y)$ is denoted by $f_{erode}(x,y) = [f \Theta S](x,y)$ and is defined as

$$[f \Theta S](x,y) = \min_{(i,j)\in S} \{ f(x+i, y+j) - S(i,j) \} \tag{4.28}$$

Dilation operation of $f(x,y)$ using a non-flat SE $S(x,y)$ is denoted by $f_{dilate}(x,y) = [f \oplus S](x,y)$ and is defined as

$$[f \oplus S](x,y) = \max_{(i,j)\in S} \{ f(x-i, y-j) + S(i,j) \} \tag{4.29}$$

Edge of the image $f(x,y)$ is defined by

$$f_{edge}(x,y) = f_{dilate}(x,y) - f_{erode}(x,y) \tag{4.30}$$

Erosion operation weakens the smaller and brighter parts and darkens the underlying image whereas the dilation operation brightens the underlying image. The outer boundary of the image under consideration is extracted by the dilation operator and the inner edge is extracted by the erosion operator. The dilation minus erosion operator extracts the real edge of an image which ensures the connectivity of the image. However, these operators are affected by noises and hence combined with opening and closing operations to detect the edge of an image.

Opening operation of $f(x,y)$ using a non-flat SE $S(x,y)$ is denoted by $f_{open}(x,y) = [f \circ S](x,y)$ and is defined as

$$[f \circ S](x,y) = (f \Theta S)(x,y) \oplus S(x,y) \tag{4.31}$$

Closing operation of $f(x,y)$ using a non-flat SE $S(x,y)$ is denoted by $f_{close}(x,y) = [f \bullet S](x,y)$ and is defined as

$$[f \bullet S](x,y) = (f \oplus S)(x,y) \Theta S(x,y) \tag{4.32}$$

Edge of the image $f(x,y)$ is defined by

$$f_{edge}(x,y) = (f \circ S) \oplus S - (f \bullet S) \Theta S \tag{4.33}$$

In the above edge detector, opening and erosion can diminish the positive noise whereas the closing and dilation operation can suppress the negative noises. Hence this type of edge detector is robust to any type of noise and can retain the image details efficiently.

Here, we have utilized the later edge detection method with the symmetrical SE $S(x,y)$ to detect the edges of an image in arbitrary direction.

$$S(x,y)=\begin{bmatrix}0&0&1&0&0\\0&1&1&1&0\\1&1&1&1&1\\0&1&1&1&0\\0&0&1&0&0\end{bmatrix} \tag{4.34}$$

Apart from that the multi-structure elements, Omni-directional SE can detect types of edges and edges in all directions respectively.

H. Hybrid Edge Detection Method

The adjustment between detection of edge pixels and localization of these pixels creates a challenging situation that leads to inaccuracy. The edge detection rate increases with a variation in threshold value, but the localization of edge pixels decreases. The above mentioned edge detection methods cannot provide desired result. This is due to the factors like noise, low contrast and accurate threshold value to be used etc.

The wavelet transform is theoretical efficient for image edge detection, but it is not sensitive enough to find the edge pixels in the vertical and horizontal direction. So, it is required to enhance the wavelet transform algorithm. This method is using fuzzy logic along with wavelet transform for edge detection. Hybridizing the notion of fuzzy logic and the traditional wavelet transform with a modification in scaling factor, this algorithm outperform the existing standard edge detection methods.

Wavelet transformation is used detect edge for the multi-scale edge detection properties. The major disadvantage of traditional wavelet transform is its bad sensitive to directional properties. In order to find out the directional edge points, fuzzification process is applied before wavelet transformation. The gray scale image is fuzzified with the help of fuzzy "if-then" rules. The fuzzification process finds out the directional edge points i.e., along horizontal, vertical and diagonal directions. Wavelet transformation is applied on the fuzzified image which yields the edge of the image. Now noise removal process is applied to the output image of wavelet

transformation technique to remove the false edge points. This edge detection method consists of five steps and the flow of events is shown in Figure 4.12.

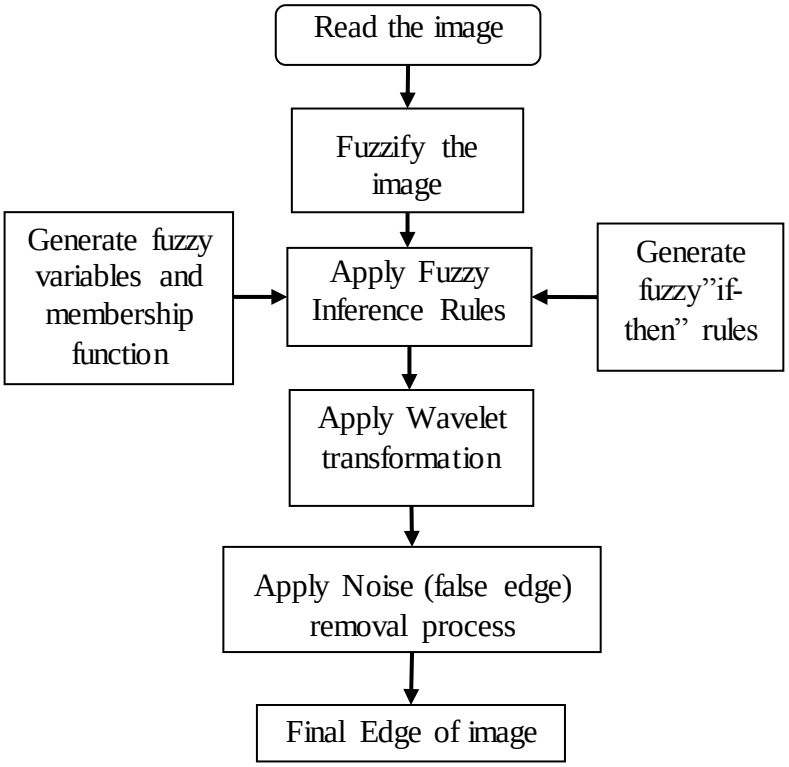


Figure 4.12: Process of hybrid edge detection method

Stepwise Procedure:

Step 1: Convert the original image to gray scale and intensity ranges from 0 to 255. The value 0 is considered as black pixel and 255 is considered as white.

Step 2: Image fuzzification:

To apply the fuzzy algorithm, the intensity level must lie in the range of 0 to 1. So, the intensity values of the pixels are converted from gray-level plane to the membership plane called fuzzification. The fuzzification method assigns membership values to each pixels by a using appropriate fuzzy based techniques. The fuzzification method can be like fuzzy clustering, fuzzy rule-based and fuzzy integration approach. In this work fuzzy-inference rule based method is used.

Step 3: Fuzzy Inference Rule:

Implementation of the proposed algorithm is using 8-bit quantization image; so, the gray levels are in the range between 0 and 255. "Black", Edge and "White" are the three linguistic variables used. Triangular membership function is used in this work as shown in Figure 4.13. The values of the membership functions are determined based on the if-then fuzzy inference rules [Patel and More, 2013] as given in Eq. 4.35 through Eq. 4.42.

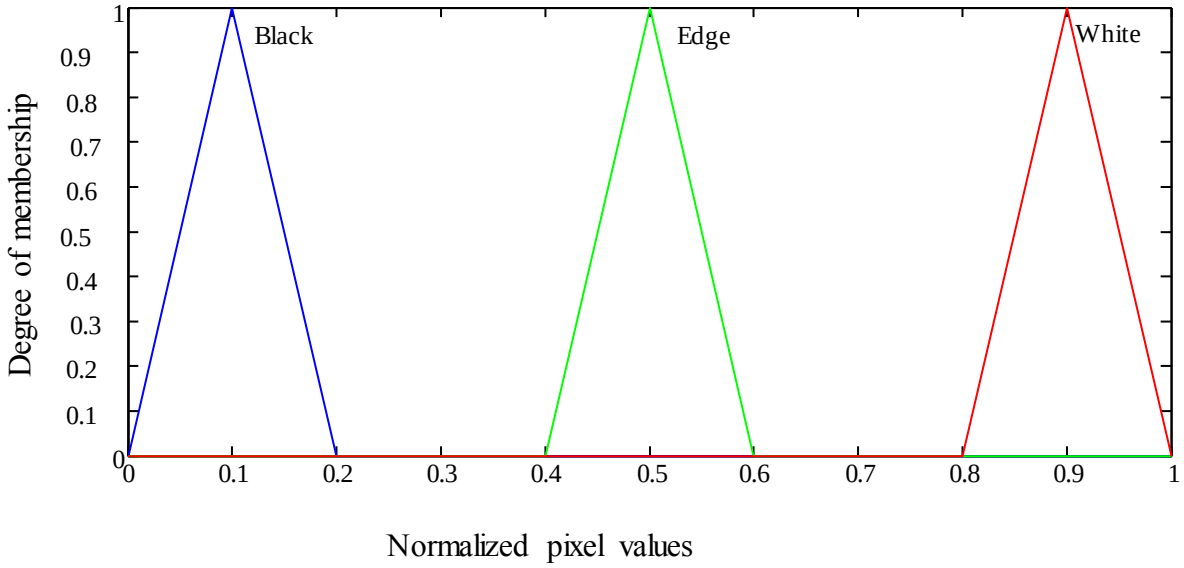


Figure 4.13: Triangular membership function

The weights of the membership for each pixels are determined by considering degree of blacks or white. These weights are assigned based on the 8-neighbored mask through scanning all the pixels. The 3x3 gray level mask is shown in Figure 4.14.

[i-1, j-1]	[i-1, j]	[i-1, j+1]
[i, j-1]	[i, j]	[i, j+1]
[i+1, j-1]	[i+1, j]	[i+1, j+1]

Figure 4.14: 8-neighbored mask

The fuzzy inference rules are developed based on the 8-neighbored mask and applied on the image to extract the edge pixels in different directions. In the rule "and" and "or" operations are performed by the minimum and maximum membership values respectively. Rules from 1 to 4 are used for determining the edges in vertical and horizontal direction around the center pixel as given in the mask. The detection of edge pixel is steered based on the Figure 4.15. The Figure 4.15 (Rule 1 to Rule 4) indicates that if the gray levels present in one line are black and the remaining are white, then, the center pixel is an edge. The rules from 5 to 8 are used for the detection of corner edge pixels depending on the weights of the pixels. The center pixel is considered to be an edge pixel when four pixels are black sequentially and the remaining four are white.

$$\text{Rule 1}\begin{cases} if\ \{(i-1,j-1)\ \&\ (i-1,j)\ \&\ (i-1,j+1)\}\ are\ whites \\ if\ \{(i,j-1)\ \&\ (i,j)\ \&\ (i,j+1)\}\ are\ whites \\ if\ \{(i+1,j-1)\ \&\ (i+1,j)\ \&\ (i+1,j+1)\}\ are\ blacks \\ \text{then the pixel is edge} \end{cases} \tag{4.35}$$

$$\text{Rule 2}\begin{cases} if\ \{(i-1,j-1)\ \&\ (i-1,j)\ \&\ (i-1,j+1)\}\ are\ blacks \\ if\ \{(i,j-1)\ \&\ (i,j)\ \&\ (i,j+1)\}\ are\ whites \\ if\ \{(i+1,j-1)\ \&\ (i+1,j)\ \&\ (i+1,j+1)\}\ are\ white \\ \text{then the pixel is edge} \end{cases} \tag{4.36}$$

$$\text{Rule 3}\begin{cases} if\ \{(i-1,j-1)\&(i,j-1)\ \&\ (i+1,j-1)\}\ are\ blacks \\ if\ \{(i-1,j)\ \&(i,j)\ \&\ (i+1,j)\}\ are\ whites \\ if\ \{(i-1,j+1)\&(i,j+1)\ \&\ (i+1,j+1)\}\ are\ white \\ \text{then the pixel is edge} \end{cases} \tag{4.37}$$

$$\text{Rule 4}\begin{cases} if\ \{(i-1,j-1)\ \&(i,j-1)\ \&\ (i+1,j-1)\}\ are\ whites \\ if\ \{(i-1,j)\ \&(i,j)\ \&\ (i+1,j)\}\ are\ whites \\ if\ \{(i-1,j+1)\ \&(i,j+1)\ \&\ (i+1,j+1)\}\ are\ blacks \\ \text{then the pixel is edge} \end{cases} \tag{4.38}$$

$$\text{Rule 5}\begin{cases} if\ \{(i-1,j)\&(i-1,j-1)\&(i,j-1)\&(i+1,j-1)\}\ are\ blacks \\ if\ \{(i-1,j+1)\&(i,j+1)\&(i+1,j+1)\&(i+1,j)\}\ are\ whites \\ if\ (i,j)\ is\ white \\ \text{then the pixel is edge} \end{cases} \tag{4.39}$$

$$\text{Rule 6}\begin{cases} if\ \{(i-1,j)\&(i-1,j-1)\&(i,j-1)\&(i+1,j-1)\}\ are\ whites \\ if\ \{(i-1,j+1)\&(i,j+1)\&(i+1,j+1)\&(i+1,j)\}\ are\ blacks \\ if\ (i,j)\ is\ white \\ \text{then the pixel is edge} \end{cases} \tag{4.40}$$

$$\text{Rule 7}\begin{cases} if\ \{(i-1,j-1)\&(i,j-1)\&(i+1,j-1)\&(i+1,j)\}\ are\ blacks \\ if\ \{(i-1,j)\ \ \&(i-1,j+1)\&(i,j+1)\&(i+1,j+1)\}\ are\ whites \\ if\ (i,j)\ is\ white \\ \text{then the pixel is edge} \end{cases} \tag{4.41}$$

$$\text{Rule 8}\begin{cases} if\ \{(i-1,j)\&(i-1,j+1)\&(i,j+1)\&(i+1,j+1)\}\ are\ blacks \\ if\ \{(i-1,j-1)\&(i,j-1)\&(i+1,j-1)\&(i+1,j)\}\ are\ whites \\ if\ (i,j)\ is\ white \\ \text{then the pixel is edge} \end{cases} \tag{4.42}$$

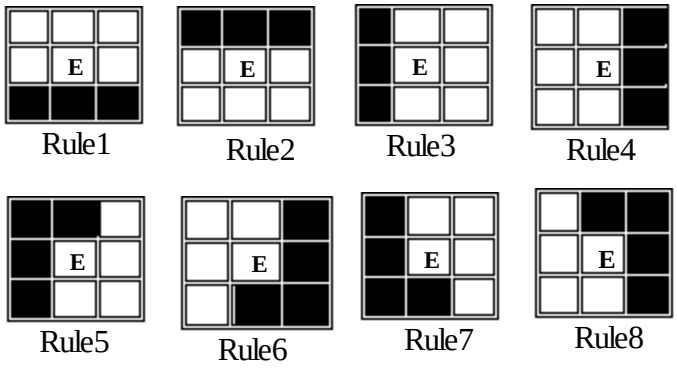


Figure 4.15: Fuzzy inference rules

The range of the pixel gray level values in the image are: Black ([0 3 5]), Edge ([130 133 135]) and White ([249 252 255]).

Step 4: Wavelet transformation:

The result from step 3, say $f(x,y)$, is now processed through wavelet transform. The continuous wavelet transform by considering the wavelet function ψ for image $f(x,y) \in L^2(R^2)$ with the scaling factor S, $\psi_S(u,v) = \psi(su, sv)$, is defined as

$$W_{\psi_1} f(x,y) = (f * \psi_S)(x,y) = \frac{1}{S^2} \iint f(u,v)\psi(sx-u, sy-v)dudv \quad (4.43)$$

The smoothing function $\theta(x,y)$ meets the conditions in Eq. 4.15. Considering Eq. 4.42, the two components of wavelet transform as in Eq. 4.18 and Eq. 4.19 along with the gradient (as in Eq. 4.20 and Eq. 4.21) is computed.

The edge points can be determined and oriented through two components of wavelet transform. At every point of the image, the direction of gradient is determined first. Then the edge point is found out whose modulus value is maximum along the gradient direction. Selection of threshold value has an impact on the clarity of edge image.

Step 5: Noise (false edge) removal:

The idea of noise removal aims to remove the pixels which have been falsely recognized as edge by the processing. 3×3 pixels mask is slid over the whole image pixel by pixel row wise and the process continues till the time whole image is scanned for unwanted edge pixels. Figure 4.16 shows p5 as falsely marked edge pixel as all the surrounding pixels i.e. p1, p2, p3, p4, p6, p7, p8 & p9 are white. Such types of falsely marked edge pixels are changed to White by the noise removal algorithm.

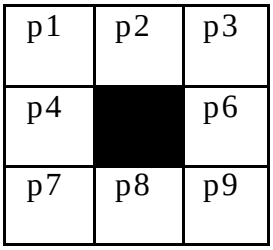


Figure 4.16: Falsely marked edge pixel

The image used often become blurred against processing because of containing noise or image information loss due to transmission. Gradient method is used for image enhancement and can make the fuzzy image clear. Images of three sample parts are used in this method. The images are fuzzified first and then the edge of image is extracted based on modified wavelet transform. The performance is analyzed quantitatively by comparing with other methods.

I. Results and Performance Analysis

This section provides the performance analysis of different edge detectors. The quantitative performance evaluation of different edge detectors are tested with the sample parts. The whole simulation work is performed in MATLAB (R2015a). The performance analysis of eight edge detector i.e., Prewitt, Sobel, Roberts, Canny, LoG, Mathematical morphology (MM) [Yu-qian et al., 2006], Wavelet and Hybrid are presented here to provide a comparison analysis. The results obtained from different edge detection methods are given in Figure 4.17 through Figure 4.21. From the visual perception it is quite clear that for each and every test image the hybrid edge detector provides a better edge detection.

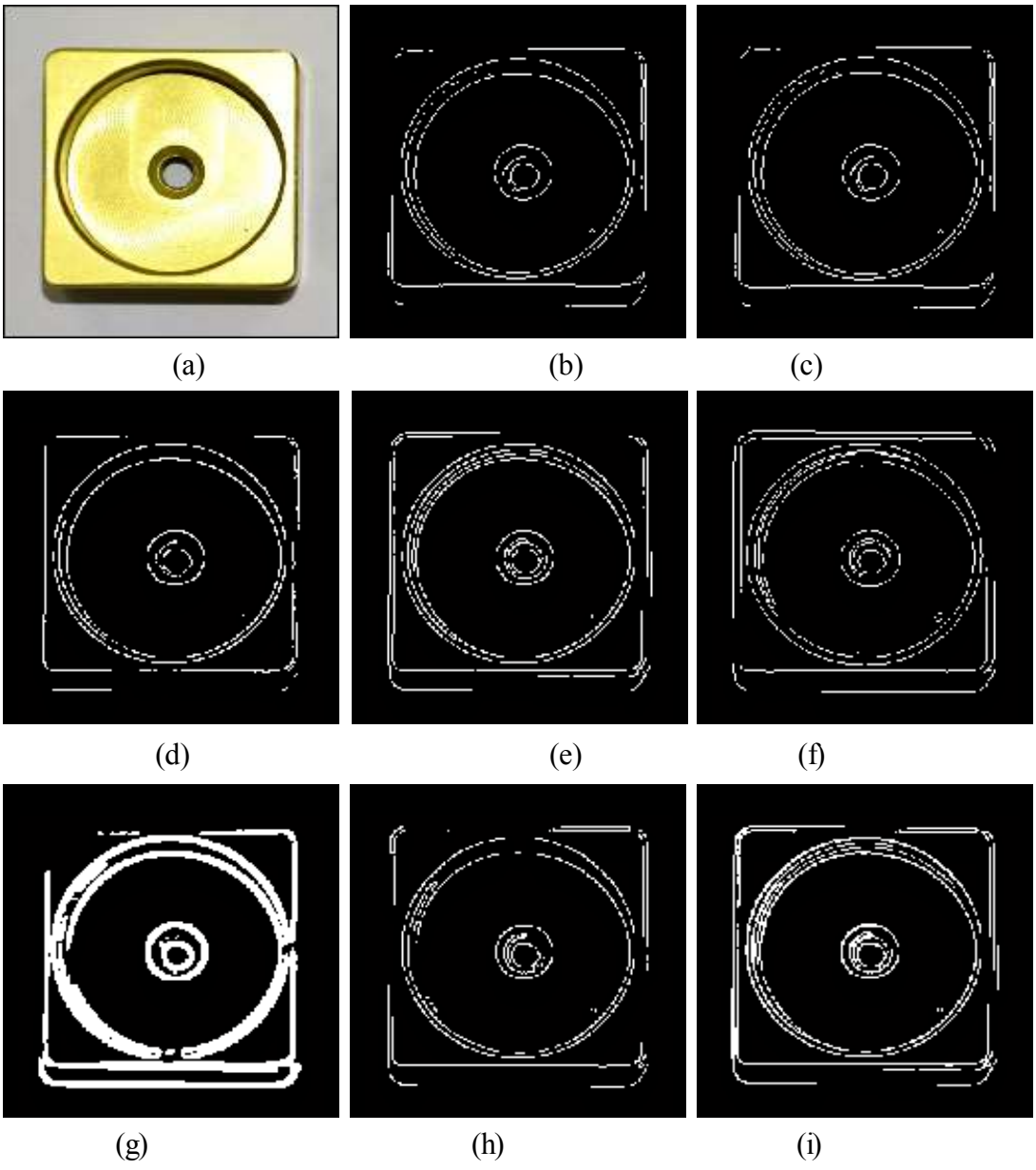


Figure 4.17: Comparison of results for image "Part1.jpg" (a) Original image, (b) Prewitt, (c) Sobel, (d) Roberts, (e) Canny, (f) LoG, (g) MM, (h) Wavelet (i) Hybrid

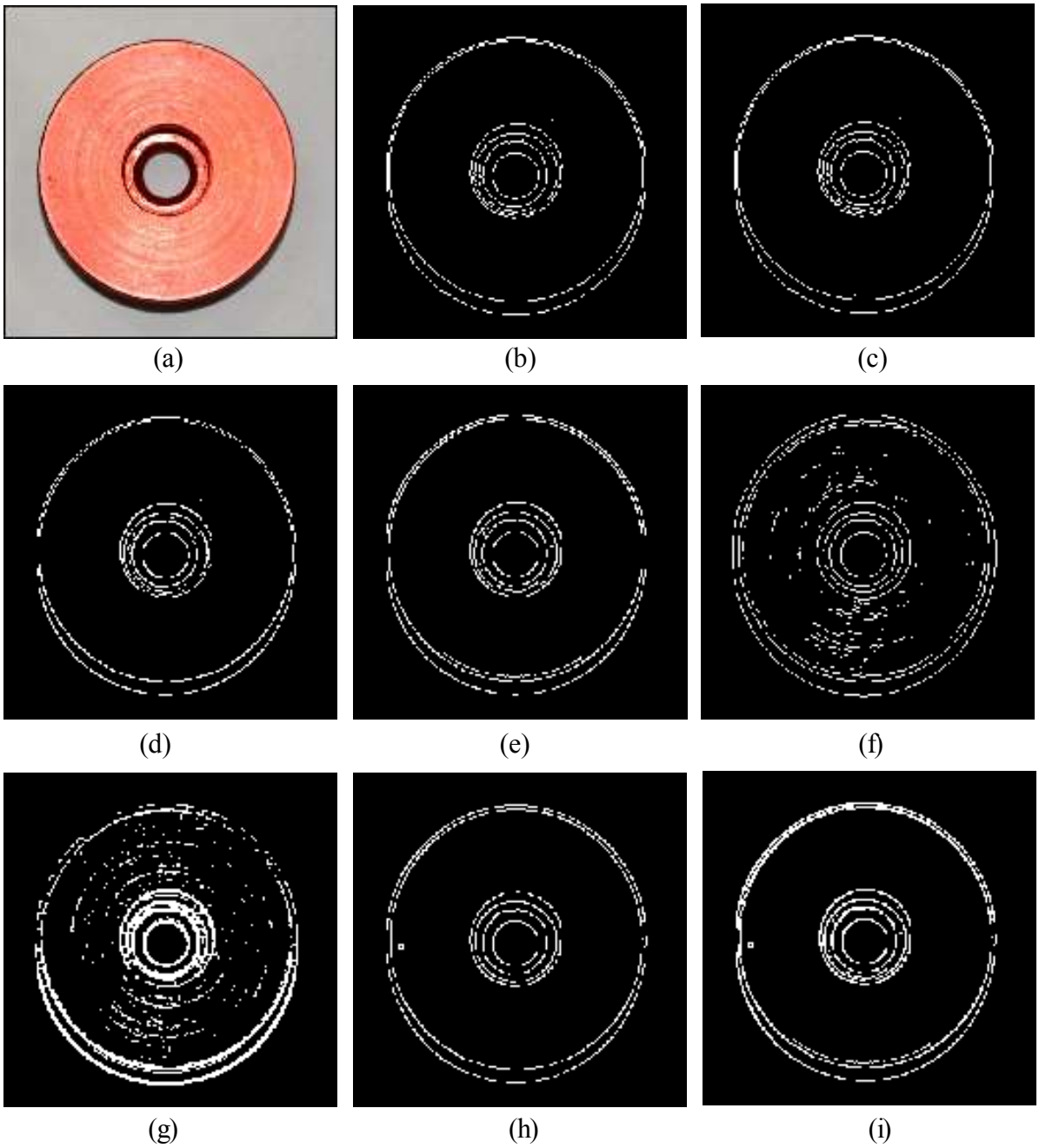


Figure 4.18: Comparison of results for image "Part2.jpg" (a) Original image, (b) Prewitt, (c) Sobel, (d) Roberts, (e) Canny, (f) LoG, (g) MM, (h) Wavelet (i) Hybrid

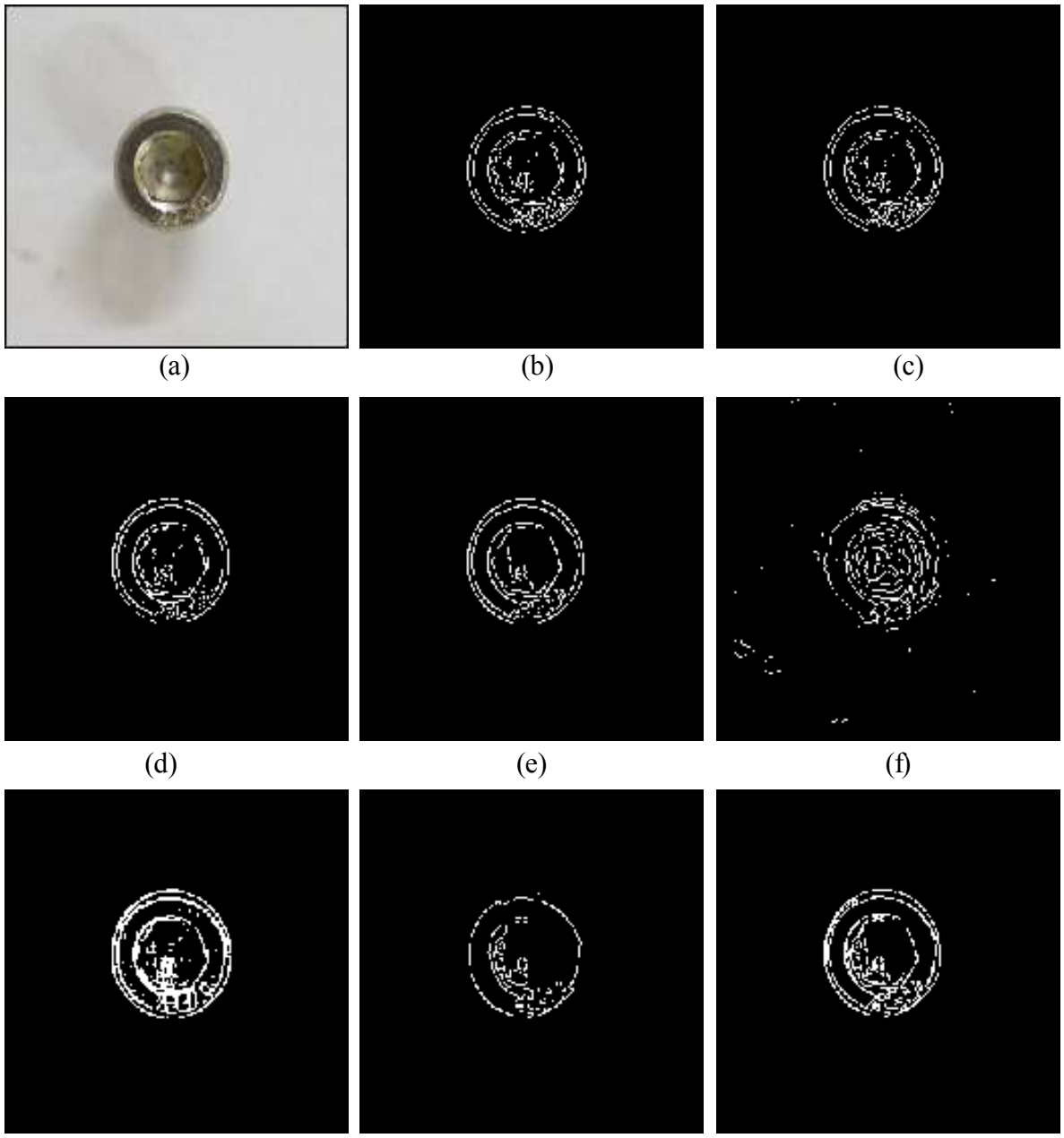


Figure 4.19: Comparison of results for image "Part3.jpg" (a) Original image, (b) Prewitt, (c) Sobel, (d) Roberts, (e) Canny, (f) LoG, (g) MM, (h) Wavelet (i) Hybrid

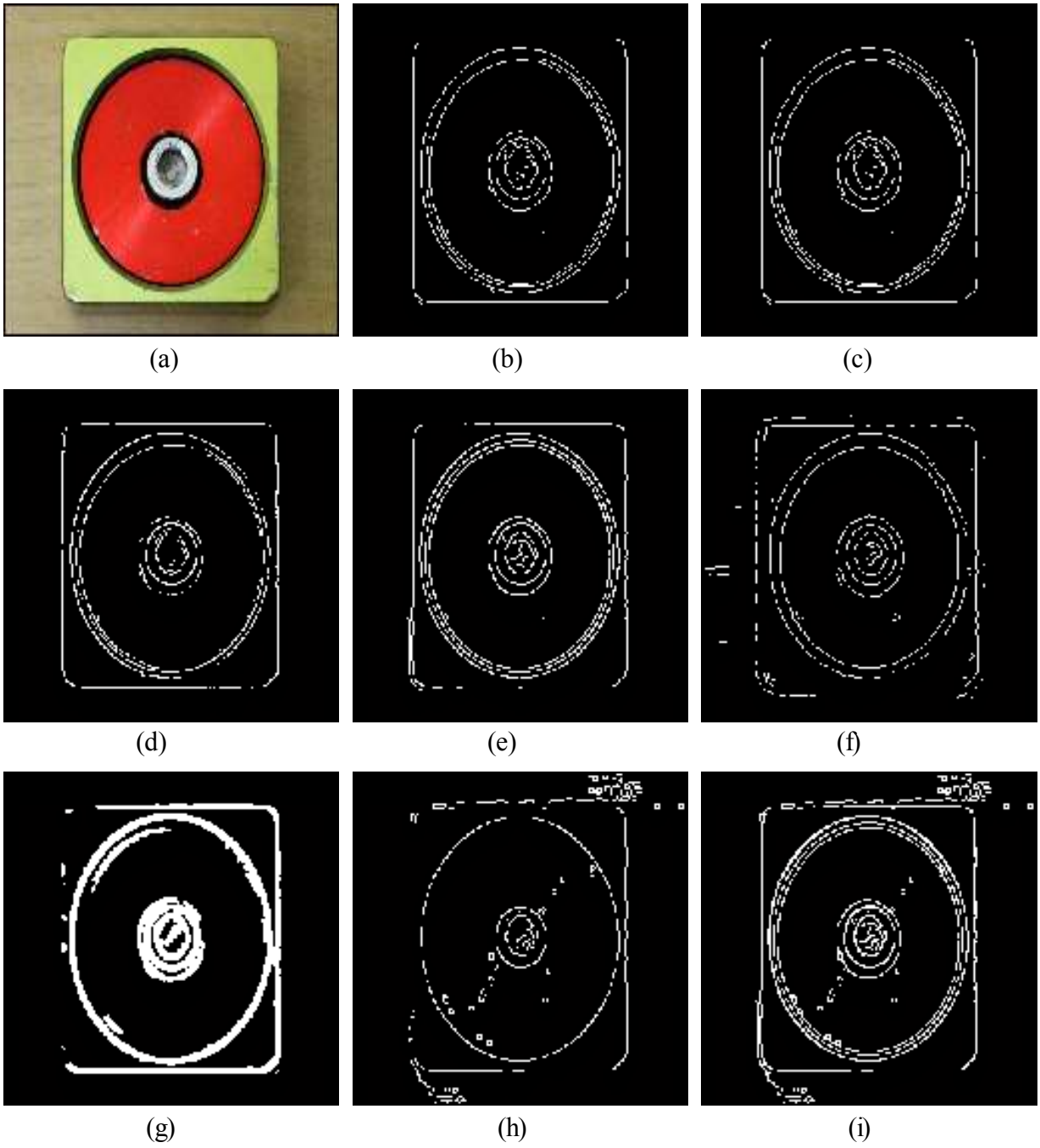


Figure 4.20: Comparison of results for image "Part4.jpg" (a) Original image, (b) Prewitt, (c) Sobel, (d) Roberts, (e) Canny, (f) LoG, (g) MM, (h) Wavelet (i) Hybrid

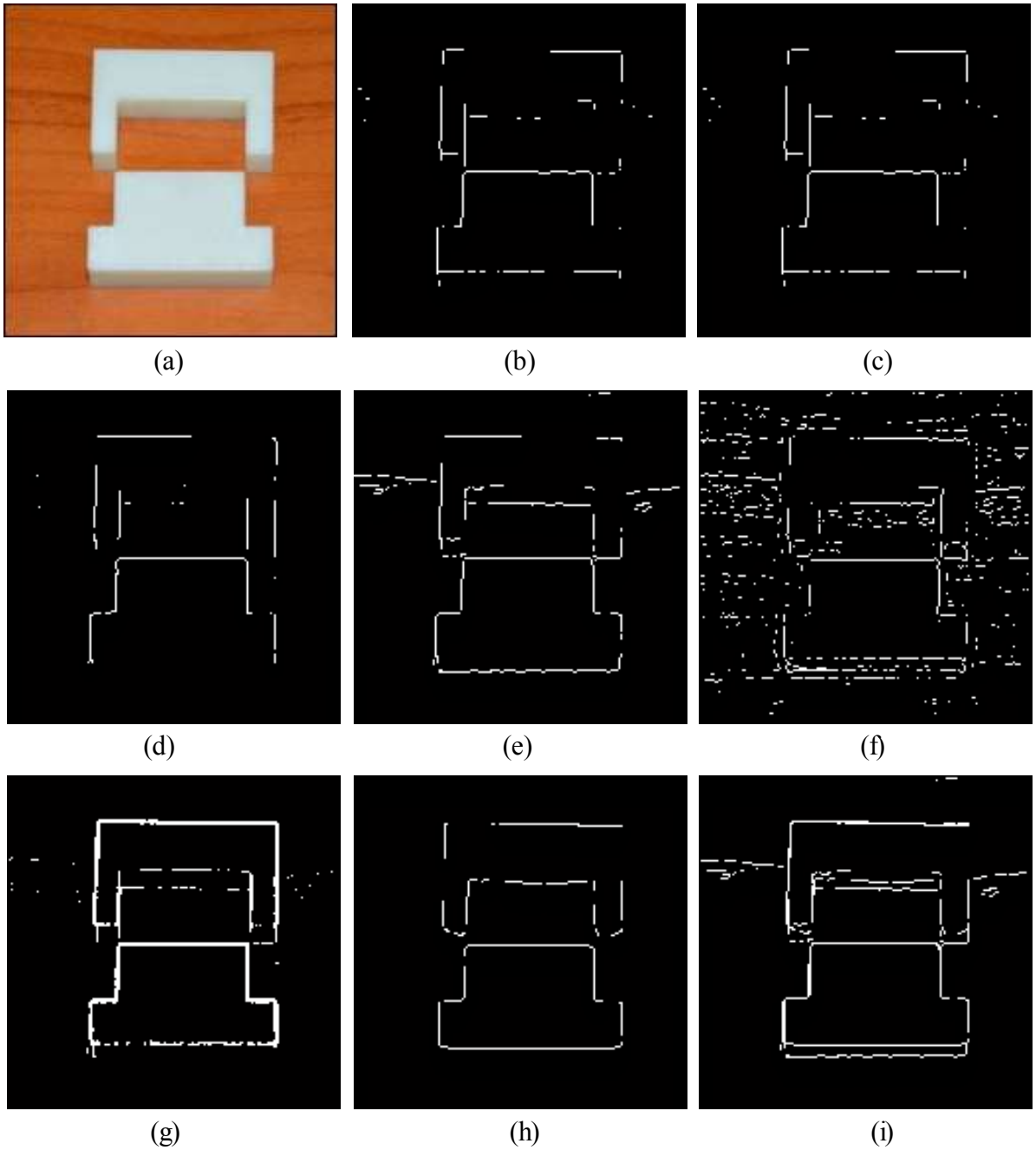


Figure 4.21: Comparison of results for image "Part5.jpg" (a) Original image, (b) Prewitt, (c) Sobel, (d) Roberts, (e) Canny, (f) LoG, (g) MM, (h) Wavelet (i) Hybrid

Besides visual analysis two significant metrics, sensitivity and specificity [Boaventura, 2009] are also presented here to characterize the accuracy of the edge detectors. These two metrics allow to recognize true edges and not to recognize the false alarms. Sensitivity of an edge detector characterizes the probability of detection of a true edge as an edge pixel. Whereas, probability of detection of a false edge as a non-edge pixel is characterized by the term sensitivity. The sensitivity of an edge detector is normally expressed in terms of true-positive rate (TP_{Rate}) and specificity in term of false-positive rate (FP_{Rate}). Eq. 4.44. Provides the mathematical formulation for true positive rate and Eq. 4.45 represents the calculation of the false-positive rate.

$$\text{Sensitivity} = TP_{Rate} = \frac{T_{positive}}{(T_{positive} + F_{negative})} \tag{4.44}$$

$$\text{Specificity} = FP_{Rate} = 1 - \left(\frac{T_{negative}}{T_{negative} + F_{positive}} \right) \tag{4.45}$$

Where, $T_{positive}$ denotes the true positive i.e. the number of edge pixels correctly detected. Eq. 4.46 provides the probability of true-positive ($PT_{positive}$) outcome of an image of size say (M x N) mathematically.

$$PT_{positive} = \text{mean}(X_{(m,n)}.Y_{(m,n)}) \tag{4.46}$$

X(m, n) denotes the probability of a pixel to be a true edge where m = 1, ..., M and n= 1, ..., N. Y(m, n) represent the probability of a pixel to be detected as edge. In the similar way the number of pixels which are erroneously classified as edge pixels are referred as false positive ($F_{positive}$) and the edge pixels that are not classified as edge pixel, called as false negative ($F_{negative}$). In order to calculate all these parameters for the comparison of different edge detectors an ideal edge map is to be necessarily defined. However defining an ideal map is unclear therefore in the present paper the edge map obtained from the hybrid edge detector is considered as an ideal edge map for the calculation of different parameters. Table 4.1 presents the comparison of percentage of correct detection (P_{cod}) and Table 4.2 presents the percentage of pixels that were not detected (P_{nod}) respectively. Percentage of false alarm P_{fal} (i.e. percentage of pixels that are erroneously detected

as edge pixels) for different edge detectors is plotted in Figure 4.22. The P_{cod}, P_{nod} and P_{fal} are computed based on the following equations.

$$P_{cod} = \frac{T_{positive}}{Max(E_I, E_D)} \times 100 \tag{4.47}$$

$$P_{nod} = \frac{F_{negative}}{Max(E_I, E_D)} \times 100 \tag{4.48}$$

$$P_{fal} = \frac{F_{positive}}{Max(E_I, E_D)} \times 100 \tag{4.49}$$

The ideal edge points of the image is presented by E_I and the number of edge points detected is defined by E_D.

Table 4.3 presents the comparison of detection error of different edge detectors. The detection error of an edge detector is expressed by the Eq. 4.50. The difference between the false-positive rate (FP_{Rate}) and true positive rate (TP_{Rate}) of concerned edge detector is termed as the detection error. [Abdou and Pratt, 1979] had proposed figure of merit (FoM), an important measure for accessing the performance measure of edge detectors. The gap between all pairs of points corresponding to the ground truth and the estimated edge map is used for the calculation of FoM. Plot of FoM for different edge detectors are presented in Figure 4.22. The mathematical formula for the calculation FoM is given in Eq. 4.51.

$$Detection\,error = \sqrt{\left(1 - TP_{Rate}\right)^2 + \left(FP_{Rate}\right)^2} \tag{4.50}$$

$$FoM = \frac{1}{Max(E_G, E_E)} \sum_{i=1}^{E_E} 1 \Big/ 1 + \alpha . d_i^2 \tag{4.51}$$

where E_G and E_E represents the number of edge points on the ideal ground truth and the calculated edge map respectively. d_i measures the distance between i^{th} edge pixel of the estimated and the ground truth edge map. α is the constant coefficient and considered as 1/9 in this work.

The range of the FoM varies in the range of [0, 1]. The optimal value of the FoM is 1 which represents that the detected edges coincides with the ground truth. The falsely detected edge points are considered as false alarm. The false alarm for different edge detectors are plotted in Figure 4.23.

Table 4.1: Comparison of Pcod using different edge detectors

Method	Part1	Part2	Part3	Part4	Part5
Prewitt	0.4942	0.5378	0.6759	0.2509	0.4519
Sobel	0.4948	0.5399	0.6809	0.2518	0.4537
Roberts	0.5341	0.6264	0.7131	0.2539	0.4176
Canny	0.6998	0.7399	0.7509	0.4417	0.6980
LoG	0.5832	0.6300	0.7201	0.6301	0.4423
MM	0.6641	0.6432	0.7426	0.3421	0.7012
Wavelet	0.6708	0.6825	0.7108	0.6782	0.5373
Hybrid	0.7292	0.7899	0.8389	0.8799	0.8645

Table 4.2: Comparison of Pnod using different edge detectors

Method	Part1	Part2	Part3	Part4	Part5
Prewitt	0.4082	0.4231	0.2732	0.4345	0.3417
Sobel	0.4079	0.4220	0.2707	0.4341	0.3408
Roberts	0.3883	0.3788	0.2357	0.4330	0.3588
Canny	0.3107	0.3412	0.1917	0.3391	0.2186
LoG	0.3637	0.3665	0.2336	0.2449	0.4213
MM	0.4171	0.3507	0.2139	0.3876	0.1932
Wavelet	0.3499	0.3440	0.2248	0.2208	0.3210
Hybrid	0.2854	0.2970	0.1499	0.1200	0.1354

Table 4.3: Comparison of detection error using different edge detectors

Method	Part1	Part2	Part3	Part4	Part5
Prewitt	0.5421	0.4019	0.2362	0.1200	0.1354
Sobel	0.5436	0.4035	0.2370	0.1310	0.1454
Roberts	0.6116	0.4066	0.2621	0.1727	0.1212
Canny	0.7225	0.7018	0.7076	0.3251	0.5325
LoG	0.5730	0.5400	0.2036	0.5463	0.1765
MM	0.5441	0.6443	0.6087	0.4423	0.5861
Wavelet	0.6910	0.6518	0.5058	0.6213	0.2562
Hybrid	0.7651	0.8653	0.7810	0.7380	0.5988

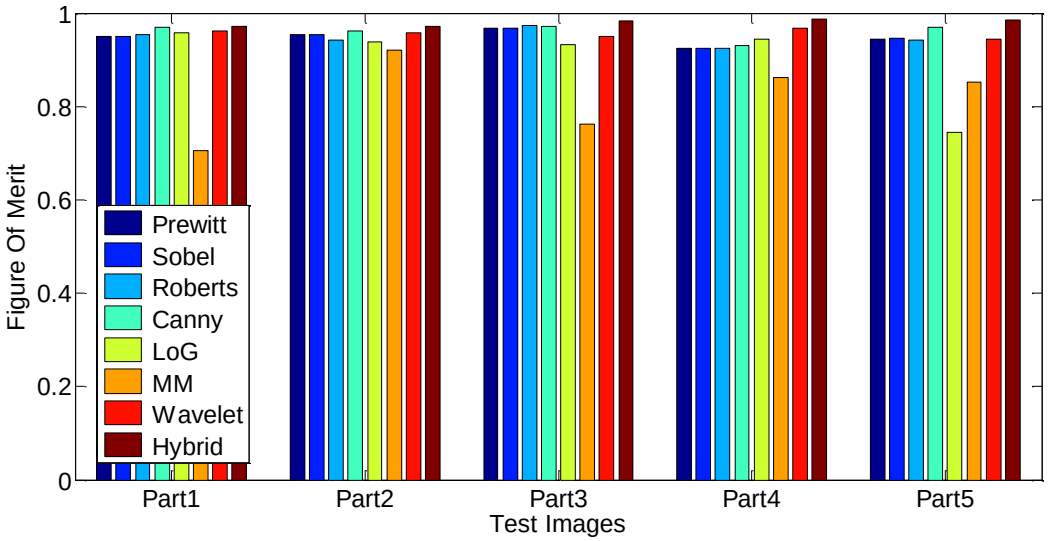


Figure 4.22: Figure of Merit for different edge detectors

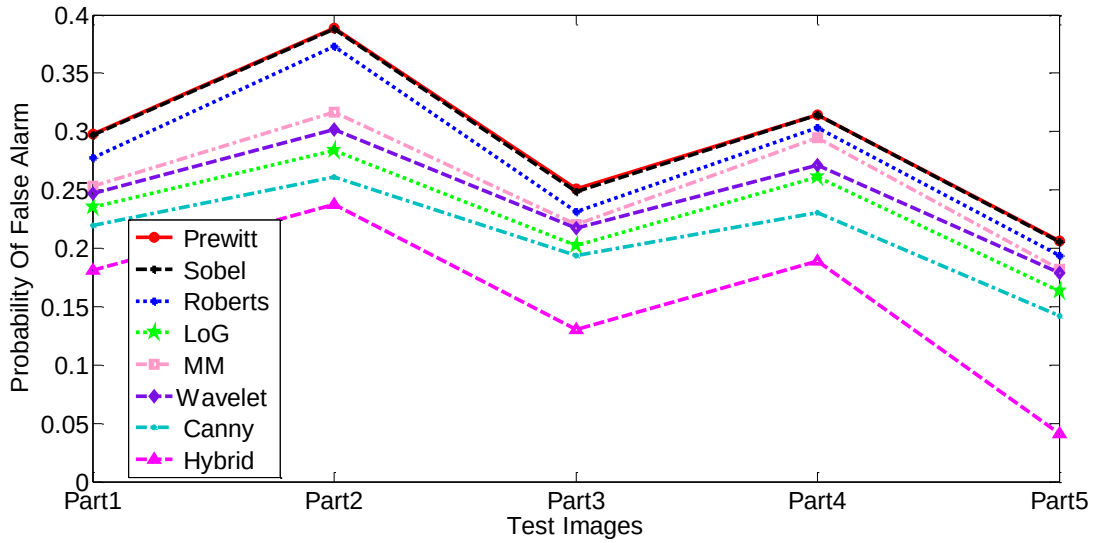


Figure 4.23: Percentage of false alarm for different edge detectors

Different objective and subjective quality measures are discussed to analyze the performance of edge detectors. In the experiment, several significant edge detectors are compared with the hybrid edge detector. Along with the visual perception of the compared edge detectors several quantitative analysis measures such as sensitivity, specificity, P_{cod}, P_{nod}, $PT_{positive}$, detection error and FoM are employed here for the analysis. From the visual analysis, the hybrid edge detector outperforms the other compared edge detectors in terms of better edge localization and edge detection. Moreover, it is also found that the quantitative analysis measures of the hybrid edge detector are better than other methods. The quantitative measures completely agree with the visual results hence can be used further. Finally the hybrid edge detector characterized by lower detection error and higher FoM is considered to be a best detector for detecting edges.

4.3.2 Corner Detection Technique

Many applications require matching of two or more images through extraction of features from them. Matching of images can be possible only the corresponding points in the image are some way related. Such points are referred to as interest points and are detected by using interest point detectors. Finding a relationship between images is then performed using only these points instead of comparing all the points in the image. This drastically reduces the computation overload of the process. Corner points are interesting as they are formed from two or more edges and edges usually define the boundary between two different objects or parts of the same object. Figure 4.24 shows part of a hypothetical system to illustrate how a corner detector might be used for object detection in a part assembly environment. This assembly line fills with parts positioned properly on the conveyor belt to ensure that parts are picked up properly by the robot. An overhead camera is used to capture image of each part as it passes under it and a system is used for processing this image further. This system enables the object detection procedure to start.

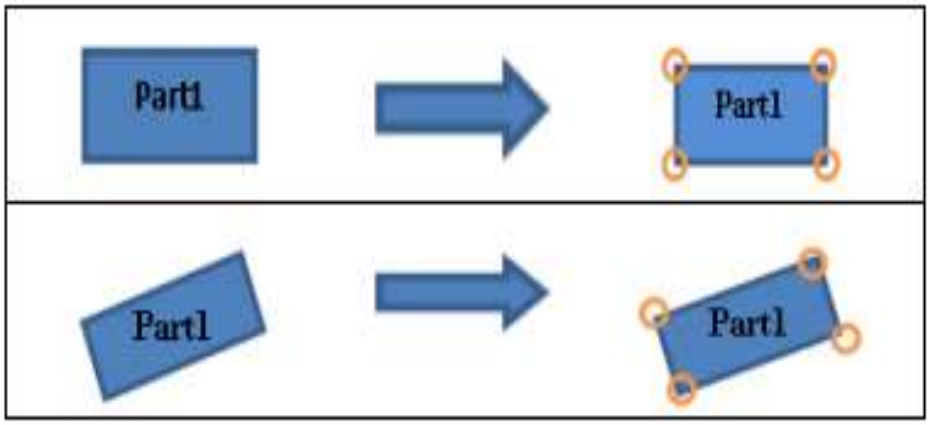


Figure 4.24: Parts of system for aligning boxes on an assembly line

Interest point detection is an approach used in computer vision system to extract certain kinds of features and infer the contents of an image. A corner can be defined as the intersection of two edges or as a point where there are two dominant and different edge directions in a local neighborhood of the point.

A. Corner Detection Algorithms

Three corner detection algorithms [Harris and Stephens, 1988; Mokhtarian and Suomela, 1998; Pei and Ding, 2007; Liu et al., 2008; Wang and Brady, 1995] are explored in detail in the following sections. They have been chosen because they are historically significant, widely used, and well suited for real-time application. In addition, these detectors can be considered as interest point detectors (interest point as corners) as they assign a measure of cornerness to all pixels in an image. General steps for detecting corners are as follows.

Corner Operator: Input to this process is the 2D image. For each pixel in the input image, the corner operator is applied to obtain a cornerness measure. The cornerness measure is simply a number indicating the degree to which the corner operator believes this pixel. The output of this step is a cornerness map.

Threshold Cornerness Map: Interest point corner detectors define corners as local maximum in the cornerness map. However, at this point the cornerness map contains many local maximum that have a relatively small cornerness measure and are not true corners. To avoid reporting these points as corners, the cornerness map is typically thresholded. All values in the cornerness map below the threshold are set to zero. Selection of threshold is application dependent and often requires trial and error experimentation. The threshold must be set high enough to remove local maximum that are not true corners and it should be that much low enough to retain local maximum at true corners.

Non-maximal Suppression: After thresholding, the cornerness map contains only nonzero values around the local maxima that need to be marked as corner points. To locate the local maxima, non-maximal suppression is applied. For each point in the cornerness map, non-maximal suppression sets the cornerness measure for this point to zero if its cornerness measure is not larger than the cornerness measure of all points within a certain distance. After nonmaximal suppression is applied, the corners are simply the non-zero points remaining in the cornerness map.

This method starts with considering a 2D gray-scale image $I(x, y)$ and denoting an image window patch by $W \in I$ centered on(x_0, y_0). The sum of square differences between W and a shifted window $W(\Delta x, \Delta y)$ is calculated as

$$S = \sum_{(x_i, y_i) \hat{I} W} \left(I(x_i, y_i) - I(x_i - Dx, y_i - Dy) \right)^2 \tag{4.52}$$

By approximating the shifted patch using a Taylor expansion truncated to the first order terms, we have:

$$S = \left[Dx, Dy\right] A \begin{bmatrix} Dx \\ Dy \end{bmatrix} \tag{4.53}$$

$$A = \begin{bmatrix} \sum_{(x_i, y_i) \hat{I} W} \left(\tilde{N}_i^h\right)^2 & \sum_{(x_i, y_i) \hat{I} W} \tilde{N}_i^h \, \tilde{N}_i^v \\ \sum_{(x_i, y_i) \hat{I} W} \tilde{N}_i^v \, \tilde{N}_i^h & \sum_{(x_i, y_i) \hat{I} W} \left(\tilde{N}_i^v\right)^2 \end{bmatrix} \tag{4.54}$$

Where, Δ_i^h and Δ_i^v represent the first order partial derivatives of image I along horizontal and vertical directions at pixel (x_i, y_i). In practice matrix A is computed by averaging the tensor product $\nabla I \cdot \nabla I^T$, ($\nabla I$ denotes the gradient image of *I*) over the window *W* with a weighting function K_ρ, i.e.

$$A = \begin{bmatrix} \sum_{(x_i, y_i) \hat{I} W} K_r(i) \left(\tilde{N}_i^h\right)^2 & \sum_{(x_i, y_i) \hat{I} W} K_r(i) \tilde{N}_i^h \, \tilde{N}_i^v \\ \sum_{(x_i, y_i) \hat{I} W} K_r(i) \tilde{N}_i^v \, \tilde{N}_i^h & \sum_{(x_i, y_i) \hat{I} W} K_r(i) \left(\tilde{N}_i^v\right)^2 \end{bmatrix} \tag{4.55}$$

Usually K_ρ is a set of Gaussian function $K_\rho(i) = \frac{1}{\sqrt{2\pi\rho}} \exp\left(-\frac{d_i^2}{2\rho^2}\right)$,

where $d_i^2 = (x_i - x_0)^2 + (y_i - y_0)^2$, ρ is the standard deviation of the Gaussian kernel and $A\rho$ is symmetric and positive semi-definite. Its main modes of variation correspond to the partial derivatives in orthogonal directions and they are reflected by the eigenvalues $\lambda 1$ and $\lambda 2$ of $A\rho$. The two eigenvalues can form a rotation-invariant description of the local pattern. Under the situation of corner detection, three distinct cases are considered. a) both the eigenvalues are small indicating that the local area is flat around the examined pixel, b) one eigenvalue is large and the other one is small that means the local neighborhood is ridge-shaped, and c) both the eigenvalues are large indicating that a small shift in any direction can cause significant change of the image at the

examined pixel. Thus a corner is detected at this pixel. The exact eigenvalue computation can be avoided by calculating the response function.

$$R\left(A_\rho\right)=\det\left(A_\rho\right)-k.trace^2\left(A_\rho\right) \tag{4.56}$$

Where det(Aρ) is the determinant of Aρ, trace(Aρ) is the trace of Aρ, and k is a tunable parameter.

B. Harris Corner Detector

The Harris corner operator [Harris and Stephens, 1988; Pei and Ding, 2007] is based on the basis of Moravec operator. The corner point is detected by the change in gray value in the image after calculating the auto-correlation matrix. Considering $I(\mathrm{x},\mathrm{y})$, gray value of the pixel(x,y), the change in intensity with a shift of(u,v), can be expressed as

$$E_{u,v}\left(x,y\right)=\sum_{\mathrm{x},y} w_{\mathrm{x},y}\left[I\left(x+u,y+v\right)-I\left(x,y\right)\right]^2 \tag{4.57}$$

where $w_{u,y}$ is the windowing function. The value of $E_{u,v}$ is 0 at nearly constant patches and large for very distinctive patches. These distinctive patches are used for corner detection.

By applying Taylor formula,

$$I\left(x+u,y+v\right)=I\left(x,\mathrm{y}\right)+I_x u+I_y v+O\left(u^2+v^2\right) \tag{4.58}$$

where, $I_x=\frac{\partial f}{\partial x}, I_y=\frac{\partial f}{\partial y}$ and $O\left(u^2+v^2\right)$ is used here to indicate that it is continuous series with second partial derivative, third partial derivative and other higher order terms.

The approximate value for first order derivative is

$$I\left(x+u,y+v\right)\approx I\left(x,\mathrm{y}\right)+I_x u+I_y v$$

So, the Harris corner derivation is as follows

$$\begin{aligned}\sum\left[I\left(x+u,y+v\right)-I\left(x,y\right)\right]^2 &\approx \sum\left[I\left(x,y\right)+uI_x+vI_y-I\left(x,y\right)\right]^2 \text{ First order approx}\\ &=\sum\left[uI_x+vI_y\right]^2\\ &=\sum\left[u^2{I_x}^2+v^2{I_y}^2+2uvI_xI_y\right]^2\end{aligned} \tag{4.59}$$

The matrix form of Eq. (4.58) is

$$\begin{aligned}\sum\left[u^2 I_x{}^2 + v^2 I_y{}^2 + 2uvI_x I_y\right]^2 &= \sum\begin{bmatrix}u & v\end{bmatrix}\begin{bmatrix}I_x{}^2 & I_x I_y\\ I_x I_y & I_y{}^2\end{bmatrix}\begin{bmatrix}u\\ v\end{bmatrix}\\ &= \begin{bmatrix}u & v\end{bmatrix}\left(\sum\begin{bmatrix}I_x{}^2 & I_x I_y\\ I_x I_y & I_y{}^2\end{bmatrix}\right)\begin{bmatrix}u\\ v\end{bmatrix}\end{aligned} \tag{4.60}$$

Now, for the small shift of (u,v), the bilinear expression is expressed as

$$E_{u,v}(x,y) \cong \begin{bmatrix}u & v\end{bmatrix} M \begin{bmatrix}u\\ v\end{bmatrix} \tag{4.61}$$

where M is a 2×2 matrix calculated from the image derivatives and defined as

$$M = \sum_{x,y} w_{x,y}\begin{bmatrix}I_x{}^2 & I_x I_y\\ I_x I_y & I_y{}^2\end{bmatrix} \tag{4.62}$$

For the simplest case, w=1 and $I_x I_y$ are the products of components of the gradients I_x and I_y. The corner response is measured by

$$R = \det M - k\left(\text{trace}M\right)^2 \tag{4.63}$$

C. Curvature Scale Space Corner Detector

The Curvature Scale Space (CSS) technique [Mokhtarian and Suomela, 1998] is suitable for recovering invariant geometric features of a planar curve at multiple scales. The curvature is defined as:

$$C(u,\sigma) = \frac{\dot{X}(u,\sigma)\ddot{Y}(u,\sigma) - \ddot{X}(u,\sigma)\dot{Y}(u,\sigma)}{\left(\dot{X}(u,\sigma)^2 - \dot{Y}(u,\sigma)^2\right)^{1.5}} \tag{4.64}$$

where,

$\dot{X}(u,\sigma) = x(u) \otimes \dot{g}(u,\sigma)$

$\ddot{X}(u,\sigma) = x(u) \otimes \ddot{g}(u,\sigma)$

$\dot{Y}(u,\sigma) = y(u) \otimes \dot{g}(u,\sigma)$, and

$$\ddot{Y}(u,\sigma) = y(u) \otimes \ddot{g}(u,\sigma).$$

$\otimes$ is the convolution operator while $g(u,\sigma)$ denotes a Gaussian function with standard deviation σ, and $\dot{g}(u,\sigma), \ddot{g}(u,\sigma)$ are the first and second derivatives of $g(u,\sigma)$ respectively. CSS algorithm comprises of the following six steps:

i. Obtain a binary edge map by applying Canny's edge detection to the gray level image.
ii. Find the edge contours, fill the gaps and find the T junctions.
iii. Compute curvature $C(u,\sigma)$ at a high scale for each edge contour.
iv. The contours having absolute curvature greater than threshold is considered as local maxima and twice as much as one of the neighboring minima.
v. Track the corners from the highest scale to the lowest scale to improve localization.
vi. Compare T junction to other corners and remove one of the two corners which are very close.

In this method, approximately 80% of the time is spent performing edge detection. Applications requiring faster corner detection can maintain a tradeoff between the performance criteria and speed by using a simpler corner detector.

D. Wang and Brady Corner Detector

The Wang and Brady detector [Wang and Brady, 1995; Harris and Stephens, 1988] considers the image to be a surface, and looks for places where there is large curvature along an image edge. The algorithm looks for places where the edge direction changes rapidly. Firstly, noise is suppressed through Gaussian function and then the edge is detected through an optimal edge detector. Then the curvature of the edge contour is calculated. The curvature is calculated as follow:

Let $n = \frac{1}{|\nabla I|}\nabla I$ be the edge normal. *t* denotes the unit tangent vector perpendicular to *n* (*t* is also sometimes referred to as the 'edge tangential').

The Laplacian is the sum of the second differentials of *I(x, y)* in orthogonal directions which is represented as;

$$\nabla^2 I = I_{xx} + I_{yy} \tag{4.65}$$

Differentiating I(x, y) in the direction of n and t using the chain rule, we have,

$$\frac{d^2I}{dn^2} = \frac{1}{\left|\tilde{N}I\right|^2}\left(I_x^2 I_{xx} + 2I_x I_y I_{xy} + I_y^2 I_{yy}\right) \tag{4.66}$$

$$\frac{d^2I}{dt^2} = \frac{1}{\left|\tilde{N}I\right|^2}\left(I_y^2 I_{xx} - 2I_x I_y I_{xy} + I_x^2 I_{yy}\right) \tag{4.67}$$

Adding Eq. 4.66 and Eq. 4.67,

$$\frac{d^2I}{dn^2} + \frac{d^2I}{dt^2} = I_{xx} + I_{yy} = \tilde{N}^2 I \tag{4.68}$$

This states the rotational invariance of the Laplacian. The total curvature of the image surface is defined to as

$$C = C_n + C_t = \left(\left(1 + I_x^2\right) I_{yy} + \left(1 + I_y^2\right) I_{xx} - 2I_x I_y I_{xy} \Big/ g^3\right) \tag{4.69}$$

where $g^3 = 1 + I_x^2 + I_y^2$ and $g \approx \left|\nabla I\right|^2$ when $\left|\nabla I\right|^2 >> 1$

Now curvature is redefined as

$$C = \frac{1}{g^3}\left(g^2 \tilde{N}^2 I - \left|\tilde{N}I\right|^2 \frac{d^2I}{dn^2}\right) \tag{4.70}$$

Now substituting Eq. 4.54 in Eq. 4.56, the curvature is derived as,

$$C» \frac{d^2I}{dt^2} / \left|\tilde{N}I\right| \tag{4.71}$$

After calculating the curvature non-maximum suppression is applied to find the pixels with large gradient magnitude. Such pixels are considered as corner score C_S, is determined by Eq. 4.72.

$$C_S = \tilde{N}^2 I - C\left|\tilde{N}I\right|^2 \tag{4.72}$$

In this case, the first term of C_S becomes the Laplacian (single-scale) blob detector.

The algorithm for this process is described as follows.

i. Obtain a binary edge map by applying Canny's edge detection to the gray level image.

ii. Find the edge contours.

iii. Compute curvature for each edge contour.

iv. Apply non-maximum suppression to find the pixels where the image gradient magnitude is large and calculate the corner score C_S.

v. Measure the Localization

vi. Suppress the false corners

E. Results and Performance Analysis

Each of the corner detectors is evaluated on images containing different corner types. It is important to remember that the intended application of the corner detector must be kept in mind when selecting a corner detector. For example, in an image alignment application the repeatability rate is critical, but detecting all true corners is of little importance as long as a sufficient number of interest points are available to allow accurate alignment of the images. However, in a part identification application failing to detect a corner may result in a different description of the part being generate leading to a misclassification of the part. The corner is detected for the object image in Figure 4.25 and the results are shown in Figure 4.26 through Figure 4.28: The comparative study is given in Table 4.4.

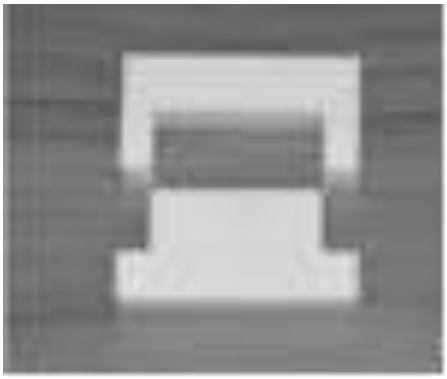


Figure 4.25: Grayscale Image of the Part

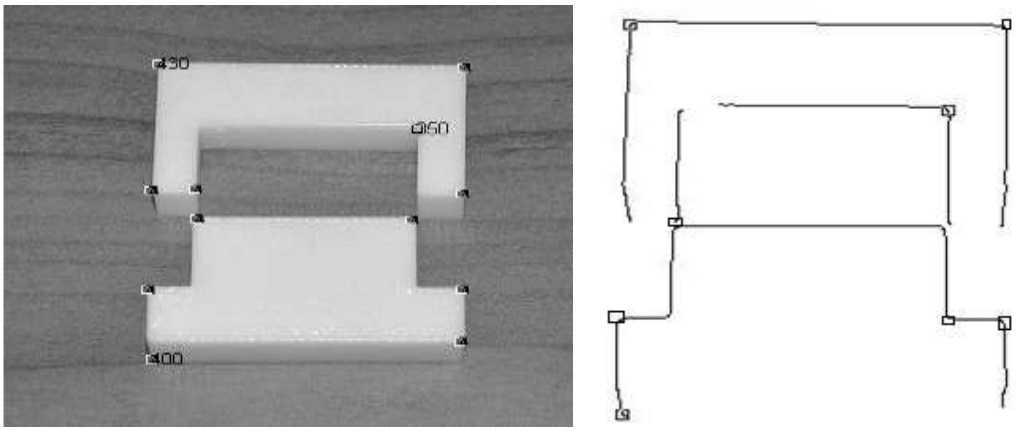


Figure 4.26: Corner detected by CSS Method

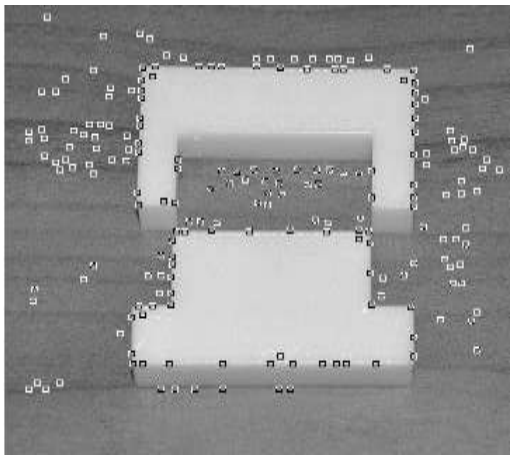


Figure 4.27: Corner detected by Harris Method

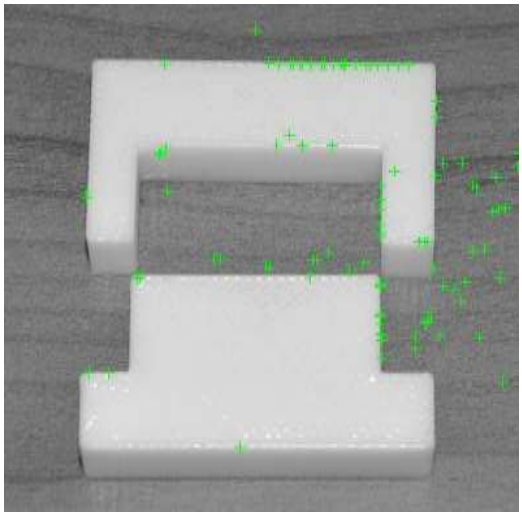


Figure 4.28: Corner detected by Wang and Brady Method

Table 4.4: Comparison of the corner detection

Corner Detector	Parameters		
	Accuracy	Speed	Localization
Harris	10 errors,16 perfect, 4 missed	73 sec	Max. error 5 pixels
Curvature Scale Space	7 errors ,21 perfect, 2 missed	43 sec	Perfect
Wang-Brady	11 errors, 14 perfect, 5 missed	57 sec	Max. error 6 pixels

For part detection in this work, the best suited corner detector is determined under a detailed analysis. If corners are to be used as features upon which subsequent processing is to be based, corners must be detected consistently. It is observed that the curvature scale space corner detection method has the best performance with respect to reliability and speed as it finds more number of true corners than false. The detection rate is good as it is excellent for all types of corners and highly robust to noise. It is highly dependent on edge detection method. The Wang and Brady method also provides better detection rate with a higher computational speed as compared to Harris method. On the other hand, Harris method provides best result only for L-junction type corners and it is not suitable for noisy images. It takes much computational time as compared to other two methods. It is easy to implement Harris method than other methods.

4.4 Mathematical Moment Based Object Detection

The parts in the assembly lines are transported to the assembly station by means of material handling devices. The captured images of the parts on the assembly station belt may be affected by geometric transformation such as scaling, rotation, translation and may be corrupted by Point-Spread-Function (PSF) blurring of camera. In order to recognize objects in such type of condition, the proposed method utilizes a set of feature vectors of the captured images which are simultaneously invariant to geometric transformations as well as to blurring environments. Moreover, the environmental noise which is inherently present in the captured images, is taken care in this method.

Object recognition is a very important and necessary image processing step in the field of computer vision. Popularly, the object recognition procedure follows four subsequent steps such as

- Image acquisition
- Image preprocessing
- Feature extraction
- Classification

Feature extraction is the most vital and principal activity in the field of object recognition. It is due to the fact that the features have the ability to represent an image via some distinctive and characteristic interest points. Consequently this process reduces the amount of data required to represent an object. Recognition of objects irrespective of their position, size, and geometrical transformation has been an active research area from last decades. The key solution to this type of problem lies in selecting object descriptors which are invariant in nature. However, whatever may be the type of object descriptors, they should take care of some of the main issues as described below.

- Selection of a feature descriptor which is invariant to image position, scale and orientation.
- Selection of a feature descriptor which is robust to noise or any other degradation present in the imagery.
- Selection of a feature descriptor which can represent the object or image compactly
- Selected feature descriptor must maintain a tradeoff between the time and space efficiency

Among a variety of alternatives, object recognition via moment invariants [Hu, 1962; Flusser, 2006; Mercimek et al., 2005], and Fourier descriptor [Generalized Fourier Descriptors with Applications to Objects Recognition in SVM Context; Object Recognition using Particle Swarm Optimization on Fourier Descriptors] are quite prominent in literature. Among all of the above mentioned object descriptors, moment invariants are extensively used in object recognition as they satisfy almost all these requirements of a good object descriptor [Khotanzad and Hong, 1990]. Therefore, this work focuses on the moment invariants to provide a suitable background for object identification.

Moment invariants are those features of the image which retain their values when the image undergoes different degradations such as shift, scaled, or rotation and blur. Moment invariants are calculated in terms of ordinary moments. Due to this excellent behavior, moment invariants have been widely used in numerous applications such as object recognition, image classification [Keyes and Winstanley, 2001], template matching [Hosny, 2010] and image registration [Flusser and Suk, 1994] etc.

The invariant features of the image are used as the input to the classification process to select the labelling of the underlying image. Hence, extraction of appropriate feature vectors is so much important in the process of object identification. However, along with the extraction of feature vectors some other factor such as deciding the optimal number of feature vector which can represent the image compactly as well what is the amount of contribution of each feature vector towards the image representation also impact on the success of the method. The majority of the existing states of art techniques in literature use an unplanned procedure for achieving such requirements. Furthermore, from the indigenous study of object identification in literature it is observed for the identification of a blurred image is performed by de-blurring that image prior to apply the image identification methods. However, de-blurring of an image without the prior knowledge about the blurring parameter is ill-posed in nature and till now it is also an open research area.

In order to take care of the above mentioned issues, the main aim of the proposed work is to utilize a new set of feature vectors of the captured images that are invariant to geometric transformations such as scaling, translation and rotation as well as to blurring environments. They are the feature vectors which are invariant to similarity transformation and blur by using orthogonal Zernike moments (ZM) [Khotanzad and Hong, 1990]. Besides this, a systematic reconstruction based approach is used for the selection of optimal order of the orthogonal ZM required in the classification problem is established.

The reason why to use orthogonal Zernike moments:

- Noise sensitivity of ZM are better than other types of moments
- ZM outperform non-orthogonal and other orthogonal moments in terms of information redundancy and image representation capability.

The rest of the discussion follows the sequence of procedures followed for the object recognition.

4.4.1 Mathematical Background of Zernike Moment

Zernike moment of order *n* with repetition l of an intensity image $f(x,y)$ is given [Urban et al., 1998; Chong et al., 2004] in Eq. 4.73

$$\mathbb{Z}_{nl} = \frac{n+1}{\pi}\sum_{x}\sum_{y} R_{nl}(\rho)\exp(jl\theta) f(x,y)\,dxdy \quad n \geq 0, |l| \leq n, n-|l| \text{ is even} \tag{4.73}$$

where $R_{nlp}(\rho)$ describes the real valued radial polynomial and is described in Eq. 4.74.

$$R_{nl}(\rho) = \sum_{s=0}^{n-\frac{|l|}{2}} (-1)^s \frac{(n-s)!}{s!\left(\frac{n+|l|}{2}-s\right)!\left(\frac{n-|l|}{2}-s\right)!}\rho^{n-2s} \tag{4.74}$$

In order to compute the ZM of an image the first and foremost step is to map the center of the image into origin and to map all of the pixel coordinates to arrange of unit circle (x^2+y^2). The pixels within that unit circle contribute in the computation of ZM whereas the pixels outside the unit circle don't contribute. Imperfect mapping scheme will introduce geometric error in the computation of ZM. Hence in order to avoid geometric error the proposed work utilizes the mapping scheme as given in [Yang and Fang, 2010] and is described in Eq. 4.75.

$$\{f(x,y)\}_{M\times M} \rightarrow \{f(u_i,v_j)\}_{[-1,1]\times[-1,1]} \tag{4.75}$$

where, $u_i = a(i+0.5)+b \quad i = 0,1,...,(M-1)$

$v_j = a(j+0.5)+b \quad j = 0,1,...,(M-1)$

$a = \frac{2c}{M-1}; b = -c; c = \frac{1}{\sqrt{2}}$

The intensity image $f(x,y)$ of size ($M\times N$) is mapped to $f(u_i,v_j)$ onto the unit circle.

4.4.2 Extraction Feature Vectors using Combined Orthogonal Zernike Moment

This section aims at creating a set of feature vectors based on orthogonal ZM which are not only invariant to geometric transformation such as rotation, scaling, translation but also insensitive to the blurring conditions [Chen et al. 2011]. The combined moment invariants of order *n* with repetition l for the digital image $f(x,y)$ is denoted as $CI(n,l)^{f(x,y)}$. Eq. 4.76 provides the mathematical formulation for $CI(n,l)^{f(x,y)}$ at $n-l+2p$.

$$CI(l+2p,l)^{f(x,y)} = e^{-jl\theta_f}\sum_{m=0}^{p}\sum_{k=0}^{p}\Gamma_{f(x,y)}^{-(l+2k+2)}C_{p,k}^{l}D_{k,m}^{l}I(l+2m,l)^{f(x,y)} \tag{4.76}$$

Where $\Gamma_{f(x,y)} = \sqrt{\mathbb{Z}_{00}}$ and $\theta_f = \arg(\mathbb{Z}_{11})$. $\mathbb{Z}_{nl}$ defines the ZM of digital image of order *n* with l repetition and is defined in Eq. 4.73. The formulation of $C_{p,k}^{l}$ and $D_{k,m}^{l}$ is described in Eq. 4.77 and Eq. 4.78 respectively. $I(l+2p,l)^{f(x,y)}$ is the set of feature vectors based on ZM invariant to blurring condition and is described in Eq. 4.79.

$$C_{p,k}^{l} = (-1)^{p-k}\frac{l+2p+1}{\pi}\frac{(l+p+k)!}{k!(p-k)!(l+k)!} \tag{4.77}$$

$$D_{k,m}^{l} = \frac{k!(l+k)!\pi}{(k-m)!(l+k+m+1)!} \tag{4.78}$$

$$I(l+2p,l)^{f(x,y)} = \mathbb{Z}_{l+2p,l} - \frac{1}{\mathbb{Z}_{00}\pi}\sum_{t=0}^{p-1}I(l+2t,l)^{f(x,y)}\sum_{j=1}^{p-1}\mathbb{Z}_{2j,0}A(l,p,t,j) \tag{4.79}$$

where the formulation for $A(l,p,t,j)$ is described in Eq. 4.80

$$A(l,p,t,j) = \sum_{k=t+j}^{p}\sum_{n=t}^{k-j}\binom{l+k}{l+n}\binom{k}{n}C_{p,k}^{l}D_{n,t}^{l}D_{k-n,j}^{0} \tag{4.80}$$

4.4.3 Image Reconstruction using the Combined Orthogonal Zernike Moment

This section describes the reconstruction of a digital image $\overset{o_{\max}}{f}(x,y)$ by utilizing the feature vectors up to a predefined order $o_{\max}$ based on combined invariant to similarity transformation and to blur using orthogonal ZM of the captured image $f(x,y)$. Eq. 4.81 provides the mathematical formulation for the estimation of $\overset{o_{\max}}{f}(x,y)$.

$$\overset{o_{\max}}{f}(x,y) = \sum_{\substack{k=0 \\ k=l+2p}}^{o_{\max}} \sum_{l} CI(l+2p,l)^{f(x,y)} R_{kl}(\rho)\exp(jl\theta) \tag{4.81}$$

When $o_{\max} \to \infty$, then the $\overset{o_{\max}}{f}(x,y)$ will approach towards $f(x,y)$. However, it is quite difficult to find out the optimal value of $o_{\max}$ for which $\overset{o_{\max}}{f}(x,y)$ will approach towards $f(x,y)$. From the literature it is quite obvious that as the moment order increases the noise sensitivity also increase. Likewise, though the lower order moment invariants are less sensitive to noise they have limited representation capability. Hence, in this work we want to develop a strategy to select the optimal order $o_{\max}$, which will maintain a proper balance between the reconstruction capability and sensitivity to noise.

4.4.4 Selection of Optimum Moment Order (o_{opt})

This discussion provides the way to choose the optimum moment order of the combined invariants for which $\overset{o_{\max}}{f}(x,y)$ will approach towards $f(x,y)$ and the reconstruction process should be less affected by noise. Selection of optimum order o_{opt} is achieved by an optimization process as described in Algorithm 4.1. Here the cost function of the optimization process is to minimize the energy function which measures the gap between the correlation between the captured and estimated image.

Algorithm 4.1: Selection of optimum moment order o_{opt}

Input: $\overset{o_{\max}}{f}(x,y)$, $f(x,y)$, $o_{\max}$ and

Output: o_{opt}

Step 1:

for i= 2 to $o_{\max}$ do

Step 2:

Find the amount of correlation between the image $\overset{i}{f}(x,y)$ and $f(x,y)$ by using the following equation

$$corr(i)=\frac{\sum_{m}\sum_{m}\left(\hat{\overset{i}{f}}(x.y)-mean(\hat{\overset{i}{f}}(x.y))\right)\left(f(x.y)-mean(f(x,y))\right)}{\sqrt{\left(\left(\sum_{M}\sum_{M}\left(\hat{\overset{i}{f}}(x.y)-mean(\hat{\overset{i}{f}}(x.y))\right)^2\right)\right)\left(\left(\sum_{M}\sum_{M}\left(f(x.y)-mean(f(x,y))\right)^2\right)\right)}} \quad 0\le corr(i)\le 1$$

Step 3:

If $\left(corr(i)\ge\varepsilon\right)\quad 0\le\varepsilon\le 1$

then $i=o_{opt}$

That means with the current moment order *i* the correlation between the captured and original image is quite high and no more improvement is needed in the reconstruction process.

Else go to step 4

Step 4:

Perform minimization of the energy function as described in the following equation to find o_{opt}

$$o_{opt}=\arg\min_{o_{opt}}\sum_{x=1}^{M}\sum_{y=1}^{M}\left(f(x,y)-\overset{o_{\max}}{f}(x,y)\right)^2$$

The above equation measure the grid point (x,y) gap between the $\overset{o_{\max}}{f}(x,y)$ and $f(x,y)$. By considering the mapping scheme as described in Eq. 4.75 we will find out the empirical solution to the above minimization problem as below.

$$o_{opt}=\Delta x\Delta y\sum_{x=1}^{M}\sum_{y=1}^{M}\left(f(x,y)-\overset{o_{\max}}{f}(x,y)\right)^2\gamma\left(o_{opt}\right)$$

where $(\Delta x,\Delta y)$ defines the pixel interval in the mapping process. $\gamma\left(o_{opt}\right)$ denotes the penalty factor and is defined as

$$\gamma\left(o_{opt}\right)=\left(1-\left(\frac{(o_{\max}+1)(o_{\max}-1)}{2}\right)\Delta x\Delta y\right)^{-1}$$

Step 5: end for

Step 6: Obtain the optimum moment order o_{opt} .

The accuracy of the o_{opt} resulted by Algorithm 4.1 is verified by experimental analysis. In the experimental analysis, the plot of normalized reconstruction error for the reconstruction of $f^{o_{\max}}(x,y)$ using the combined moment invariants $CI(l+2p,l)^{f(x,y)}$ of order zero through $o_{\max}=15$ from the captured image $f(x,y)$ is plotted in Figure 4.29. The plot is the average plot for the 5 test images such as "Part1.jpg", "Part2.jpg", "Part3.jpg", "Part4.jpg", and "Part5.jpg" as given in Chapter 3. Eq. 4.82 provides the mathematical formulation of the reconstruction error. From the plot it is found that the reconstruction error decreases slowly and attains the minimum value at around 6th order. After that the reconstruction error starts increasing. Hence, the o_{opt} from the experimental study is found to be at 6th order and is termed as $\{o_{opt}\}_{Ex}$. However, the selection of o_{opt} by using Algorithm 4.1 is found to be 7 and is termed as $\{o_{opt}\}_{Al}$. The relative error between the selections of optimum order moment can be determined by the formulation given in Eq. 4.83 and is found to be 3.32%.

$$error^2 = \frac{\sum_{x=0}^{M-1}\sum_{y=0}^{M-1}\left(f(x,y)-f^{o_{\max}}(x,y)\right)^2}{\sum_{x=0}^{M-1}\sum_{y=0}^{M-1}\left(f(x,y)\right)^2} \tag{4.82}$$

$$relativeerror\left(\{o_{opt}\}_{Ex}\{o_{opt}\}_{Al}\right) = \frac{\left\| f^{\{o_{opt}\}_{Al}}(x,y) - f^{\{o_{opt}\}_{Ex}}(x,y) \right\|}{\left\| f^{\{o_{opt}\}_{Al}}(x,y) \right\|} \tag{4.83}$$

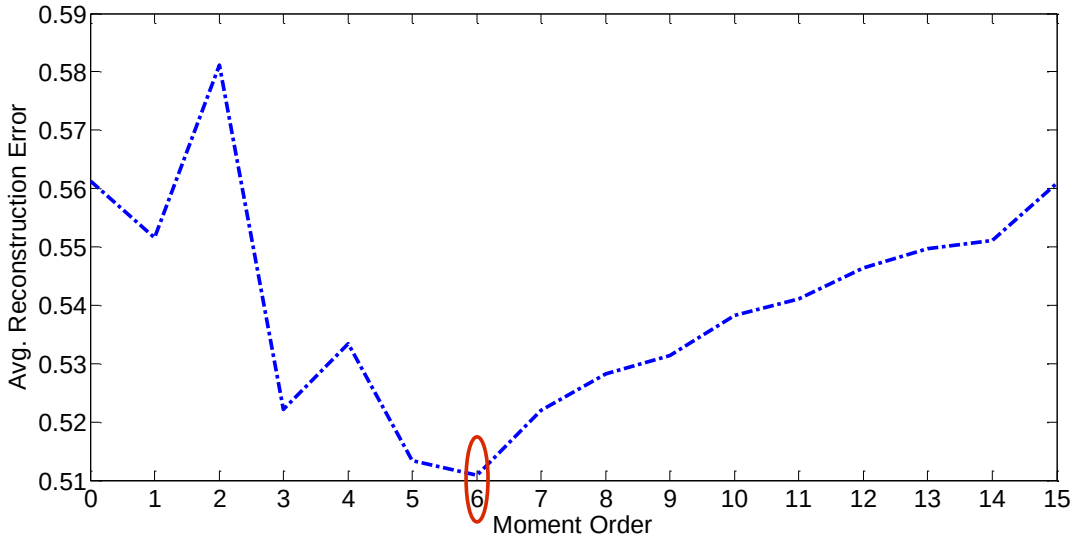


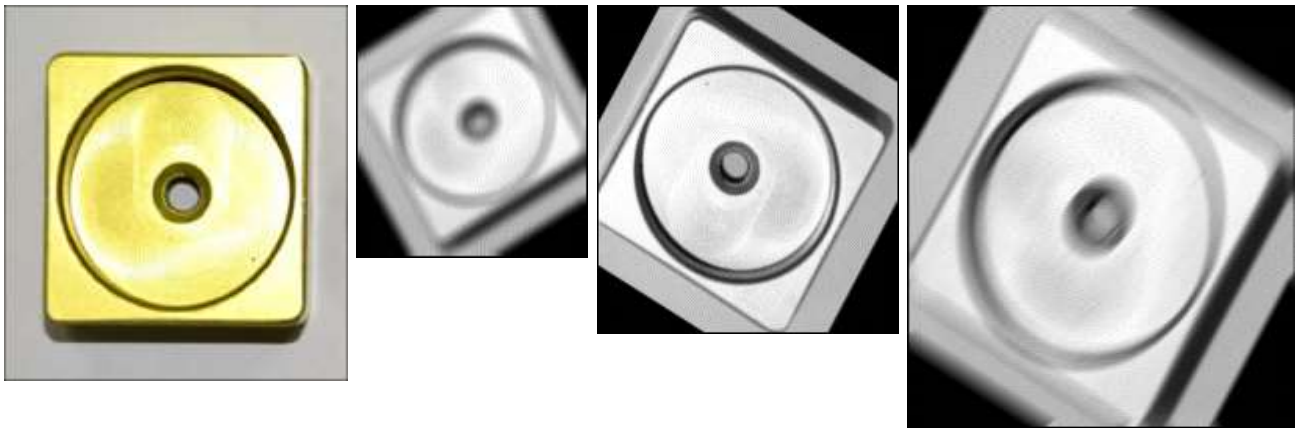
Figure 4.29: Average reconstruction error vs Moment order

4.4.5 Result Analysis for the Invariance Measure

Effectiveness and robustness of the proposed method is demonstrated through experimental results is in this section. Tests are performed in two different scenarios. The first experiment is tested with different synthetically generated test images and the second experiment is tested on the real images captured from the conveyor belt. To test the invariance of moment invariants towards the geometrical transformation and to blurring condition, test images are generated synthetically for five parts such as "part1.jpg", "part2.jpg", "part3.jpg", "part4.jpg" and "part5.jpg". The synthetic images are rotated by different angles from $\theta = 5^{\circ}$ to 150° and different scaling parameters as well as three different types of blurring kernels are also used. The parameters used for creating synthetic images are Δx, Δy, θ and S. Δx and Δy represents the translation along X-axis and Y-axis respectively. θ and S defines the angle of rotation and amount of scaling respectively. The blurring types such as average blur with size (7 x7), Gaussian blur with size (31 x 31) with standard deviation=0.1 and motion blur with parameter (20,45) are used. The detail description about the parameters used for generating the synthetic images for different sample part images are given in Table 4.5. Figures 4.30 through 4.34 provide the graphical representation of the synthetically generated images whereas Figure 4.35 shows the real captured images from the conveyor belt.

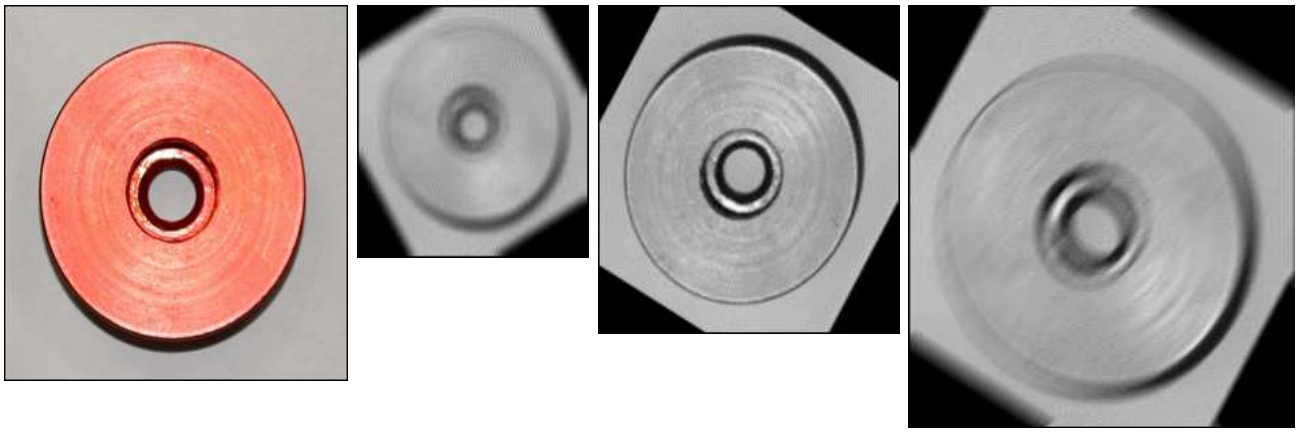
Table 4.5: Parameters used for creating synthetic images

(a) Original image	**(b)**	**(c)**	**(d)**
Part1.jpg	$\Delta x = -5, \Delta y = 7$ $\theta = 30, S = 0.5$ *Average Blur (7 x7)*	$\Delta x = -5, \Delta y = 5$ $\theta = 150, S = 0.75$ *Gaussian Blur (31 x31)*	$\Delta x = -1, \Delta y = 6$ $\theta = 60, S = 1$ *Motion Blur (20,45)*
Part2.jpg	$\Delta x = -5, \Delta y = 7$ $\theta = 30, S = 0.5$ *Average Blur (7 x7)*	$\Delta x = -5, \Delta y = 5$ $\theta = 150, S = 0.75$ *Gaussian Blur (31 x31)*	$\Delta x = -1, \Delta y = 6$ $\theta = 60, S = 1$ *Motion Blur (20,45)*
Part3.jpg	$\Delta x = -5, \Delta y = 7$ $\theta = 30, S = 0.5$ *Average Blur (7 x7)*	$\Delta x = -5, \Delta y = 5$ $\theta = 150, S = 0.75$ *Gaussian Blur (31 x31)*	$\Delta x = -1, \Delta y = 6$ $\theta = 60, S = 1$ *Motion Blur (20,45)*
Part4.jpg	$\Delta x = -5, \Delta y = 7$ $\theta = 30, S = 0.5$ *Average Blur (7 x7)*	$\Delta x = -5, \Delta y = 5$ $\theta = 150, S = 0.75$ *Gaussian Blur (31 x31)*	$\Delta x = -1, \Delta y = 6$ $\theta = 60, S = 1$ *Motion Blur (20,45)*
Part5.jpg	$\Delta x = -5, \Delta y = 7$ $\theta = 30, S = 0.5$ *Average Blur (7 x7)*	$\Delta x = -5, \Delta y = 5$ $\theta = 150, S = 0.75$ *Gaussian Blur (31 x31)*	$\Delta x = -1, \Delta y = 6$ $\theta = 60, S = 1$ *Motion Blur (20,45)*



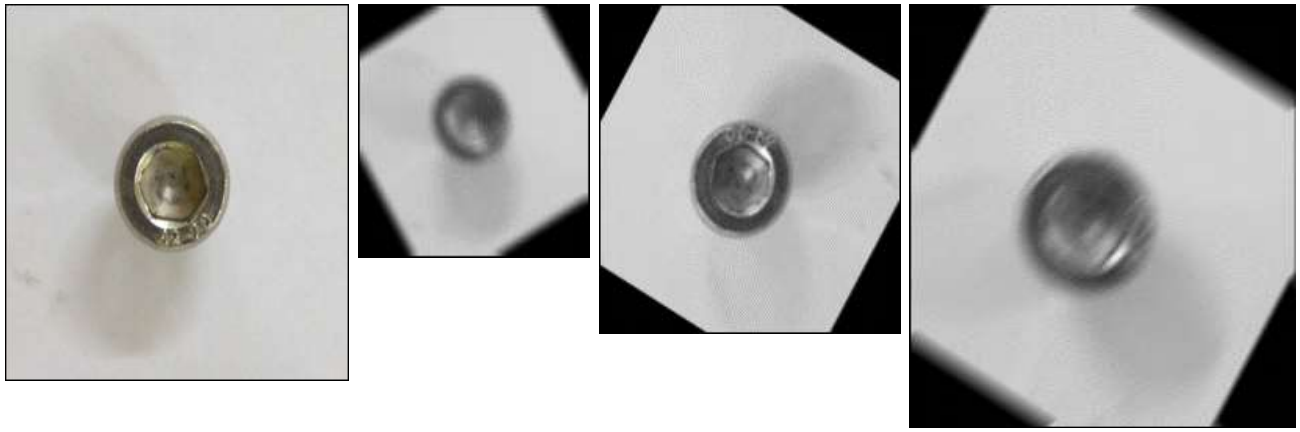
(a) (b) (c) (d)

Figure 4.30: Synthetically generated images for original image "Part1.jpg"



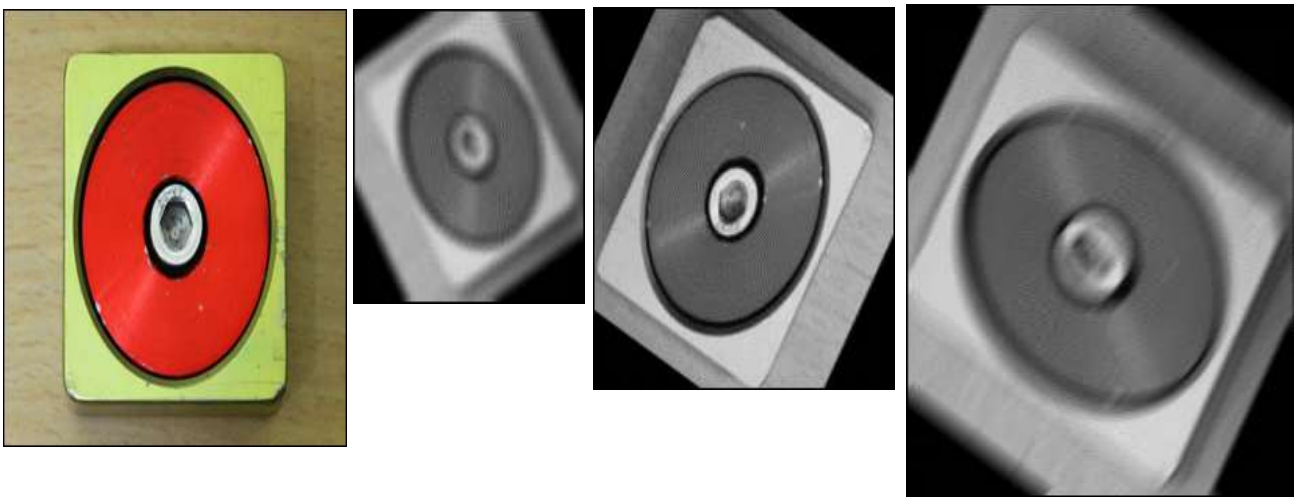
(a) (b) (c) (d)

Figure 4.31: Synthetically generated images for original image "Part2.jpg"



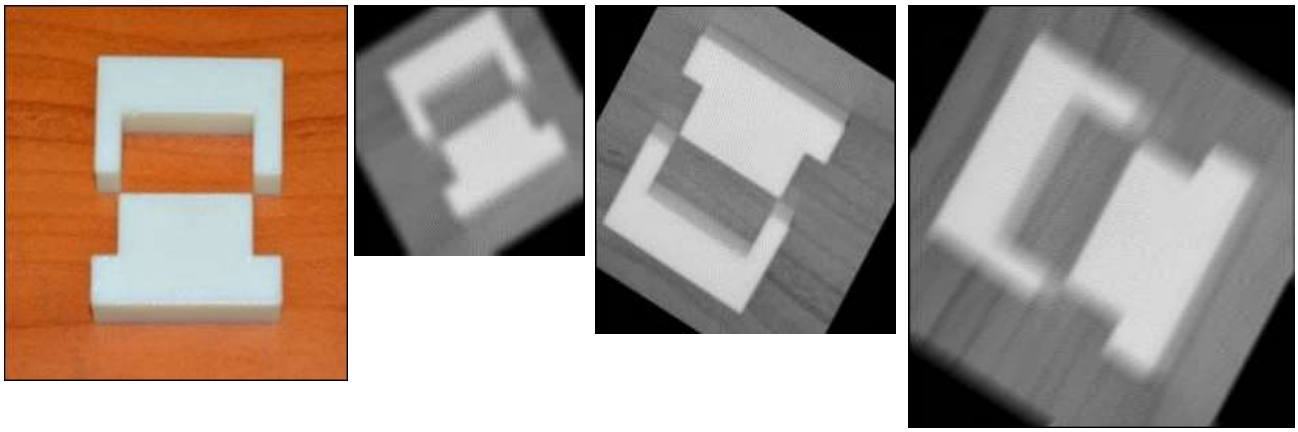
(a) (b) (c) (d)

Figure 4.32: Synthetically generated images for original image "Part3.jpg"



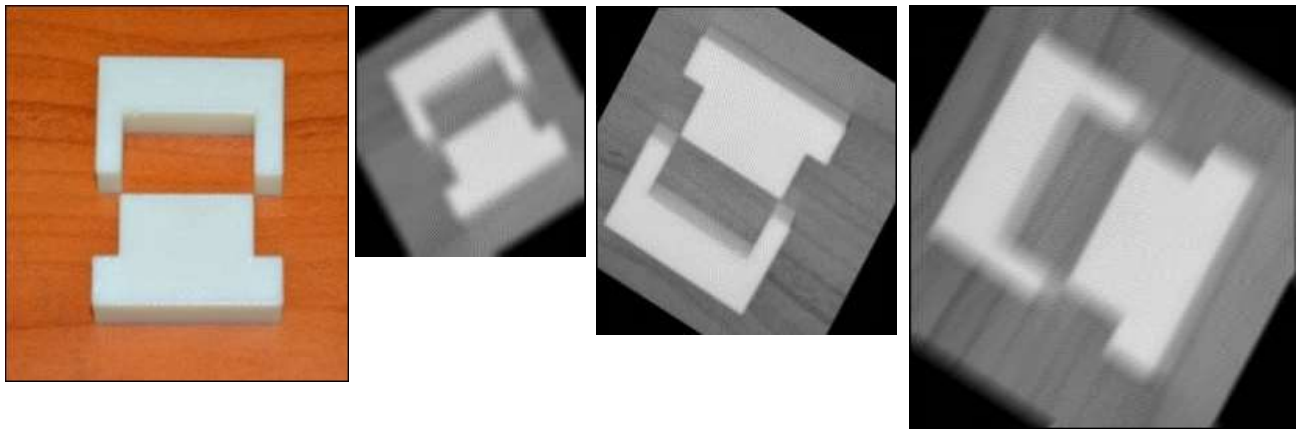
(a) (b) (c) (d)

Figure 4.33: Synthetically generated images for original image "Part4.jpg"



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

Figure 4.34: Synthetically generated images for original image "Part5.jpg"



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

Figure 4.35: Four real images from the conveyor belt

The invariants values of different order up to order o_{opt} for the synthesized images are given in Table 4.6 through Table 4.10. The invariant values for real image is given in Table 4.11. From the tabular data analysis, it is observed that irrespective of the similarity transformation and nature of degradation, there exist excellent invariance measures among the test images. Consequently, the proposed method well identifies the captured images in the conveyor belt.

Table 4.6: Combined moment invariant (order 6) for synthetic images of "Part1.jpg" image

Moment invariants	Fig. 4.30(a)	Fig. 4.30(b)	Fig. 4.30 (c)	Fig. 4.30(d)
CI(2,0)	5.56e-02	5.56e-02	5.57e-02	5.56e-02
CI(2,2)	2.36e-03	2.36e-03	2.36e-03	2.33e-03
CI(3,1)	3.62e-04	3.62e-04	3.62e-04	3.62e-04
CI(3,3)	5.98e-04	5.98e-04	5.97e-04	5.98e-04
CI(4,0)	4.41e-03	4.41e-03	4.41e-03	4.44e-03
CI(4,2)	3.30e-02	3.30e-02	3.30e-02	3.30e-02
CI(4,4)	4.56e-05	4.56e-05	4.56e-05	4.56e-05
CI(5,1)	2.01e-02	2.01e-02	2.01e-02	2.01e-02
CI(5,3)	2.67e-02	2.67e-02	2.65e-02	2.67e-02
CI(5,5)	5.23e-05	5.23e-05	5.23e-05	5.20e-05
CI(6,0)	1.33e-04	1.33e-04	1.33e-04	1.33e-04
CI(6,2)	2.56e-06	2.56e-06	2.56e-06	2.56e-06
CI(6,4)	3.12e-05	3.10e-05	3.11e-05	3.13e-05
CI(6,6)	6.57e-06	6.57e-06	6.57e-06	6.57e-06

Table 4.7: Combined moment invariant (order 6) for synthetic images of "Part2.jpg" image

Moment invariants	Fig. 4.31(a)	Fig. 4.31 (b)	Fig. 4.31 (c)	Fig. 4.31 (d)
CI(2,0)	1.56e-02	1.56e-02	1.56e-02	1.56e-02
CI(2,2)	4.67e-03	4.66e-03	4.67e-03	4.67e-03
CI(3,1)	4.56e-04	4.56e-04	4.56e-04	4.56e-04
CI(3,3)	6.45e-04	6.42e-04	6.45e-04	6.44e-04
CI(4,0)	5.89e-03	5.89e-03	5.89e-03	5.89e-03
CI(4,2)	4.23e-02	4.23e-02	4.21e-02	4.23e-02
CI(4,4)	5.17e-05	5.17e-05	5.17e-05	5.17e-05
CI(5,1)	2.81e-02	2.81e-02	2.81e-02	2.81e-02
CI(5,3)	3.07e-02	3.07e-02	3.07e-02	3.06e-02
CI(5,5)	4.11e-05	4.10e-05	4.14e-05	4.11e-05
CI(6,0)	3.78e-04	3.78e-04	3.78e-04	3.78e-04
CI(6,2)	4.10e-06	4.10e-06	4.11e-06	4.10e-06
CI(6,4)	3.42e-05	3.42e-05	3.42e-05	3.42e-05
CI(6,6)	7.80e-06	7.80e-06	7.79e-06	7.82e-06

Table 4.8: Combined moment invariant (order 6) for synthetic images of "Part3.jpg" image

Moment invariants	Fig. 4.32(a)	Fig. 4.32 (b)	Fig. 4.32 (c)	Fig. 4.32 (d)
CI(2,0)	3.11e-03	3.11e-03	3.11e-03	3.11e-03
CI(2,2)	5.45e-04	5.45e-04	5.44e-04	5.45e-04
CI(3,1)	5.78e-05	5.78e-05	5.78e-05	5.78e-05
CI(3,3)	2.10e-05	2.10e-05	2.10e-05	2.10e-05
CI(4,0)	5.19e-04	5.19e-04	5.19e-04	5.17e-04
CI(4,2)	2.45e-03	2.45e-03	2.45e-03	2.44e-03
CI(4,4)	3.67e-06	3.67e-06	3.67e-06	3.67e-06
CI(5,1)	3.90e-03	3.91e-03	3.90e-03	3.90e-03
CI(5,3)	4.32e-03	4.30e-03	4.31e-03	4.33e-03
CI(5,5)	5.31e-06	5.31e-06	5.31e-06	5.30e-06
CI(6,0)	4.61e-05	4.61e-05	4.61e-05	4.61e-05
CI(6,2)	3.22e-07	3.22e-07	3.22e-07	3.21e-07
CI(6,4)	4.51e-06	4.50e-06	4.50e-06	4.51e-06
CI(6,6)	5.56e-07	5.56e-07	5.56e-07	5.56e-07

Table 4.9: Combined moment invariant (order 6) for synthetic images of "Part4.jpg" image

Moment invariants	Fig. 4.33(a)	Fig.4.33(b)	Fig. 4.33 (c)	Fig.4.33(d)
CI(2,0)	1.16e-02	1.16e-02	1.16e-02	1.16e-02
CI(2,2)	4.35e-04	4.35e-04	4.35e-04	4.35e-04
CI(3,1)	5.08e-05	5.08e-05	5.08e-05	5.08e-05
CI(3,3)	3.16e-06	3.16e-06	3.17e-06	3.18e-06
CI(4,0)	4.44e-04	4.44e-04	4.44e-04	4.44e-04
CI(4,2)	6.17e-03	6.18e-03	6.18e-03	6.17e-03
CI(4,4)	4.50e-05	4.50e-05	4.50e-05	4.51e-05
CI(5,1)	3.90e-04	3.92e-04	3.91e-04	3.90e-04
CI(5,3)	3.32e-03	3.32e-03	3.32e-03	3.32e-03
CI(5,5)	5.46e-05	5.46e-05	5.46e-05	5.44e-05
CI(6,0)	2.23e-04	2.22e-04	2.20e-04	2.21e-04
CI(6,2)	1.12e-06	1.10e-06	1.12e-06	1.12e-06
CI(6,4)	3.45e-05	3.45e-05	3.45e-05	3.45e-05
CI(6,6)	4.58e-06	4.51e-06	4.56e-06	4.58e-06

Table 4.10: Values of combined moment invariant (order 6) for synthetic images of "Part5.jpg" image

Moment invariants	Fig. 4.34(a)	Fig. 4.34 (b)	Fig. 4.34 (c)	Fig. 4.34 (d)
CI(2,0)	4.90e-02	4.90e-02	4.90e-02	4.90e-02
CI(2,2)	5.78e-03	5.78e-03	5.78e-03	5.78e-03
CI(3,1)	2.56e-04	2.56e-04	2.56e-04	2.56e-04
CI(3,3)	5.55e-03	5.55e-03	5.55e-03	5.57e-03
CI(4,0)	5.32e-02	5.32e-02	5.31e-02	5.22e-02
CI(4,2)	3.43e-03	3.43e-03	3.43e-03	3.43e-03
CI(4,4)	1.78e-05	1.78e-05	1.75e-05	1.81e-05
CI(5,1)	1.86e-03	1.86e-03	1.86e-03	1.86e-03
CI(5,3)	5.11e-03	5.11e-03	5.13e-03	5.14e-03
CI(5,5)	8.33e-05	8.31e-05	8.36e-05	8.33e-05
CI(6,0)	5.11e-03	5.11e-03	5.11e-03	5.11e-03
CI(6,2)	6.55e-05	6.55e-05	6.55e-05	6.55e-05
CI(6,4)	7.65e-06	7.65e-06	7.65e-06	7.65e-06
CI(6,6)	1.57e-06	1.56e-06	1.58e-06	1.57e-06

Table 4.11: Values of combined moment invariant (order 6) for real images

Moment invariants	Fig. 4.35(a)	Fig. 4.35 (b)	Fig. 4.35 (c)	Fig. 4.35 (d)
CI(2,0)	7.56e-04	7.56e-04	7.56e-04	7.56e-04
CI(2,2)	2.09e-04	2.19e-04	2.09e-04	2.09e-04
CI(3,1)	4.19e-05	4.19e-05	4.19e-05	4.19e-05
CI(3,3)	6.88e-07	6.88e-07	6.84e-07	6.85e-07
CI(4,0)	3.60e-05	3.60e-05	3.60e-05	3.60e-05
CI(4,2)	5.89e-05	5.89e-05	5.86e-05	5.89e-05
CI(4,4)	5.40e-02	5.40e-02	5.44e-02	5.47e-02
CI(5,1)	2.8e-03	2.8e-03	2.81e-03	2.8e-03
CI(5,3)	5.67e-04	5.65e-04	5.60e-04	5.67e-04
CI(5,5)	8.00e-06	8.00e-06	8.03e-06	8.00e-06
CI(6,0)	7.66e-07	7.66e-07	7.69e-07	7.66e-07
CI(6,2)	3.56e-06	3.56e-06	3.56e-06	3.56e-06
CI(6,4)	5.95e-06	5.96e-06	5.96e-06	5.95e-06
CI(6,6)	8.15e-06	8.15e-06	8.15e-06	8.15e-06

4.4.6 Object Classification using Nearest-Neighbor Method

In this step the classification of an image is done by passing the k-dimensional optimal feature vectors to a nearest-neighbor (NN) classifier. The optimal feature vectors are obtained from the utilized combined invariants $CI(l+2p,l)^{f(x,y)}$ of order $o_{opt}=6$. With $o_{opt}=6$ the images can be appropriately represented by only 14 number of feature vectors. As a result, the computational burden of the whole system will reduce to a larger extent.

Figure 4.36 shows the block diagram for the above process. The main aim of the image classification is to label an unknown image I into a class c provided a database of N classified images. For which, we have first find out the 14 optimal feature vectors (using the feature vectors with $o_{opt}=6$) of image. Next step is to find out the 14 optimal feature vectors for each of the classified images present in the database. After finding out the feature vectors of the tested image and the images in the data base we have to find out the distance or the similarity measure between the test image and the images in the data base through image Euclidian distance (IMED). Unlike Euclidian distance the IMED not only measures the intensity differences but also it considers the spatial relationship among the pixels into consideration. Hence it provides a better similarity measure between image pairs. Then the unknown image I is assigned to that class of image for which the IMED is minimum.

In a N class problem, let $T_k^{(i)}$ be the set of images in the database with k=1, 2, 3, ...,14 feature vectors and labeled to class i where i=1, 2, 3, ..., N and $T_k^{(i)}=\left[T_{k1}^{(i)}\ \ T_{k2}^{(i)}\ \ T_{k3}^{(i)}\ \cdots\ T_{kt}^{(i)}\right]$.

The unknown image I is assigned to class N^* and defined as

$$N^* = \min_N \rho\left(I, T_k^{(i)}\right) \tag{4.84}$$

where $\rho\left(I, T_k^{(i)}\right)$ measures the distance between the feature vectors between the image I and $T_k^{(i)}$ and is described in Eq. 4.85.

$$\rho\left(I, T_k^{(i)}\right) = IMED\left(I, T_k^{(i)}\right) \tag{4.85}$$

$$IMED^2\left(I,T_k^{(i)}\right)=\frac{1}{2\pi}\sum_{x,y=1}^{MN}\exp\left\{-\left(\frac{\left|I_i-\left\{T_k^{(i)}\right\}_j\right|}{2}\right)\right\}\left(I^x-\left\{T_k^{(i)}\right\}^x\right)\left(I^y-\left\{T_k^{(i)}\right\}^y\right) \quad (4.86)$$

where $\left|I_i-\left\{T_k^{(i)}\right\}\right|$ defines the pixel distance.

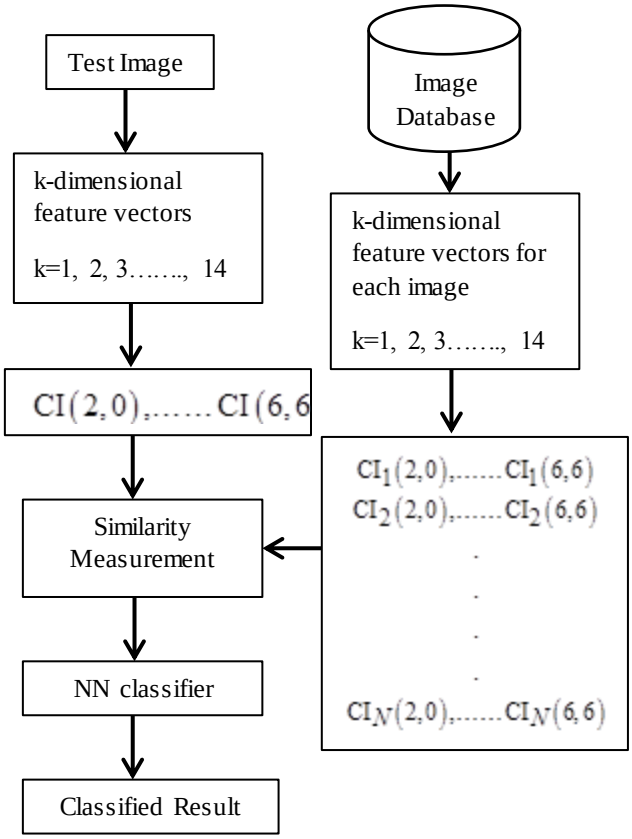


Figure 4.36: Block diagram for object classification using Nearest Neighbor method

Table 4.12: Recognition accuracy using NN classifier

Feature Type	Feature order	Number of features	Accuracy (%)
Geometric Moment	12	47	99
Legendre Moment	11	40	97
Zernike moment	9	28	98
Combined Zernike moment invariant	6	14	99

Table 4.12 provides the analysis of classification accuracy by using the NN classifier. From the tabular data analysis, it is observed that, the classification accuracy of the classifier using feature vectors based on Combined Zernike moment invariant is superior. By utilizing only 14 features the classifier achieves accuracy around 99%. Whereas, to achieve the identical accuracy 47 number of feature vectors based on geometric moment is needed. On the other hand 40 numbers of feature vectors based on Legendre Moment are used to achieve an accuracy around 97% and 28 numbers of feature vectors based on Zernike Moment are used to achieve an accuracy around 98%. From the above analysis it is quite obvious that using the proposed Combined Zernike moment invariant based feature the system achieve an excellent trade-off between the performance and computational time.

4.5 Summary

In this chapter, the basic model of object detection system is discussed. The success of the object detection system depends on the feature extraction technique as the detection is only possible by the features of the objects. The main and vital activity in object detection is the feature extraction. This is because the features have the ability to represent an image via some distinctive measures and interest points. More importantly, this process minimizes the volume of data required to represent an object. Two feature extraction methods are proposed by considering the relational features like edge contours and moments. Firstly, a hybrid edge detection technique is proposed whose performance is compared with widely used edge detection techniques. The performance of the hybrid edge detection technique is measured by quantitative analysis and found to be better than the other methods. Secondly, an orthogonal Zernike moment based object detection method is proposed in which the extracted features are invariant to scale, rotation and transformation. Besides this, an efficient weighted scheme is used based on the image representation ability of the each moment order. This method measures the feature that is best for object detection.

Chapter 5

Integration of Vision System and Robot Controller

5.1 Overview

Machine vision (MV) performs exploration of images and extracts data for governing a process or activity. It is a sub-discipline of artificial intelligence in which smart camera are used to obtain information about an environment and extract data from digital images of objects present. Vision system (VS) takes an image as input and provides some level of description about the objects in it (i.e., size, position and orientation etc.). A suitable VS with a smart camera is chosen properly. The performance of these devices depends on the proper interfacing. Important consideration for image acquisition is camera calibration method. Similarly, a robot controller (Kawasaki Robot) is selected as a facilitator in the application. The performance of robot requires a proper installation of its controller and other necessary components. Considering the system as a whole, interfacing of machine vision system with the robot controller is performed. This chapter contains the description of experimental setup to realize the desired objective. In the setup, the hardware and software components cited in Chapter 3 are used along with the interfacing procedures.

5.2 Installation of Vision Processor

The vision processor developed in the proposed work comprises of NI PXIe-1082 chassis [http://sine.ni.com/nips/cds/view/p/lang/en/nid/207346], NI PXIe-8135 embedded controller [http://sine.ni.com/nips/cds/view/p/lang/en/nid/210547; www.ni.com/pdf/manuals/373716b.pdf], NI PXIe-6341 data acquisition device [http://sine.ni.com/nips/cds/view/p/lang/en/nid/207415] and interfaces for camera and other devices. A PXIe-1082 eight-slot chassis is used for developing a complete MV system as shown in Figure 5.1.

This chassis is a high-bandwidth backplane to meet a wide range of high-performance test and measurement application needs. It accepts PXI Express modules in every slot and supports standard PXI hybrid-compatible modules in up to four slots. The chassis operating temperature range is extended to 55 °C for applications in demanding environments. By combining this chassis with the LCD monitor, mouse and keyboard, a best-in-class portable and rugged VS is developed. It also incorporates all the features of the latest PXI specification including support for both PXI and PXI Express modules and built-in timing and synchronization features.

The PXIe-8135 as shown in Figure 5.2 is a high-performance Intel Core i7-3610QE processor-based embedded controller is attached to slot no. 1 in the chassis. This contains 2.3 GHz base

frequency, 3.3 GHz (single-core, Turbo Boost mode) quad-core processor, and dual-channel 1600 MHz DDR3 memory, this controller is ideal for processor-intensive modular instrumentation, vision algorithms and DAQ applications.

Figure 5.1: Chassis of NI Machine Vision System (NI PXIe 1082)

Figure 5.2: NI PXIe-8135 Embedded Controller

The camera and robot controller connection is done through Ethernet, which is achieved by connecting with NI 8234 dual Gb Ethernet interface [http://sine.ni.com/nips/cds/view/p/lang/en/nid/205150] in slot 2 of the chassis. The NI 8234 (Figure 5.3) is a high-performance dual Gigabit Ethernet interface for PXI Express. It includes two Gigabit Ethernet ports within the single-slot PXI Express module for connecting multiple devices over Ethernet. The ports are compatible with 10BASE-T and 100BASE-TX fast networks.

Figure 5.3: NI 8234 Gb Ethernet

NI X Series multifunction data acquisition (DAQ) devices [http://sine.ni.com/nips/cds/view/p/lang/en/nid/207415] provide a new level of performance with the high-throughput PCI Express bus, NI-STC3 timing and synchronization technology, and multicore-optimized driver and application software. In slot 3 a NI PXIe-6341 DAQ (Figure 5.4) assistance is fitted which helps in accurately taking small-scale measurements. The driver software for this device is NI-DAQ is NI MAX which is used for collecting information about the devices connected to PXIe.

Figure 5.4: NI PXIe-6341 X Series Multifunction DAQ

The NI 7340 (Figure 5.5) motion controller [www.ni.com/pdf/manuals/370838b.pdf; http://sine.ni.com/nips/cds/view/p/lang/en/nid/3809] is an advanced motion control device associated with easy-to-use software tools and add-on motion VI libraries for the use with LabVIEW. The motion controller is a combination servo and stepper motor controller for PXI. This controller provides fully programmable motion control for up to four independent or

coordinated axes of motion, with dedicated motion I/O for limit and home switches and additional I/O for general-purpose functions. Arbitrary and complex motion trajectories are determined by this controller by using servo motors.

Figure 5.5: NI PXI-7340 Motion Controller

The PXI-8252 [http://sine.ni.com/nips/cds/view/p/lang/en/nid/208264] is an IEEE 1394a interface shown in Figure 5.6. The IEEE 1394a standard is ideal for high-data-rate machine vision applications with bandwidth up to 400 Mb/s. The PXI-8252 acquires data from cameras through three IEEE 1394a camera input ports. Hubs connected to any of the ports allow acquisition from cameras. It is connected in slot 6 of the chassis.

Figure 5.6: PXI-8252 IEEE 1394 Host Adapter

All the modules are connected to the chassis PXIe-1082. This acts as a processor (Figure 5.7) which can process complex vision algorithms and DAQ related methods.

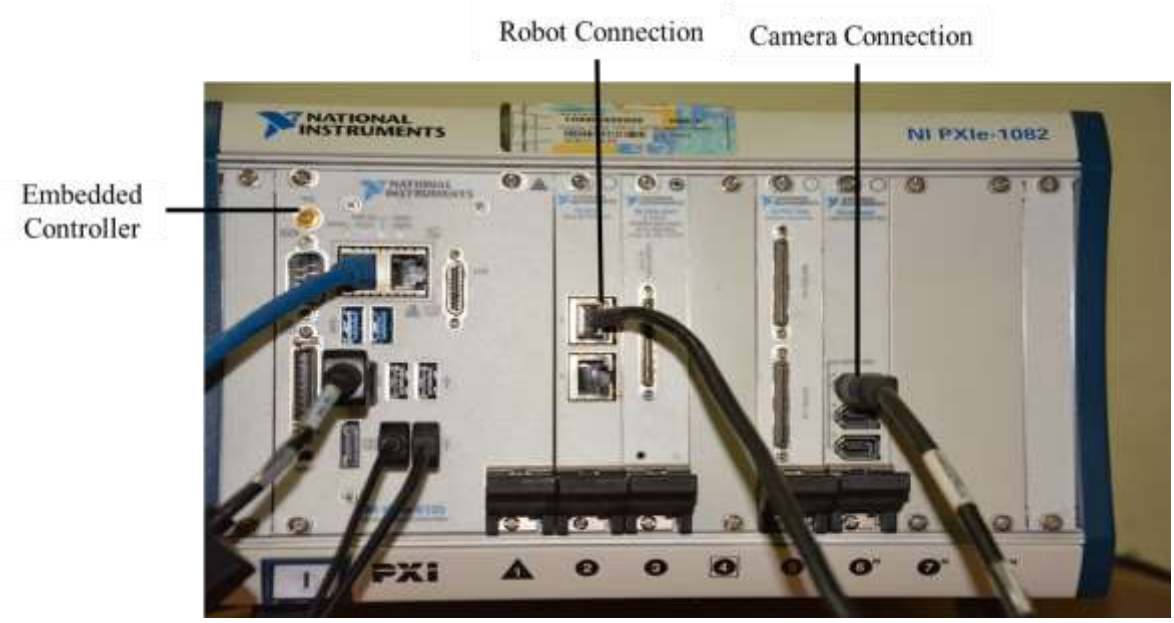


Figure 5.7: NI Machine Vision Processor

5.3 Camera Calibration and Image Acquisition

The performance of the vision system lies on the quality of image captured by camera. This depends on the camera, its configuration and acquisition method. National Instruments Measurement and Automation Explorer (NI MAX), LabVIEW and the Vision Toolkit make image acquisition a straightforward and integrated task.

5.3.1 Configuring the Camera

The use of measurement and automation explorer (NI MAX) finds the camera details connected to the system as shown in Figure 5.8. The devices and interfaces connected to PXIe chassis is shown in the NI MAX. Cameras are considered as an image acquisition device. In NI MAX, under National Instruments image acquisition card (NI-IMAQ), camera configuration details are available. The details of the chosen camera are shown in in Figure 5.9 and the camera attributes can be edited as per requirement.

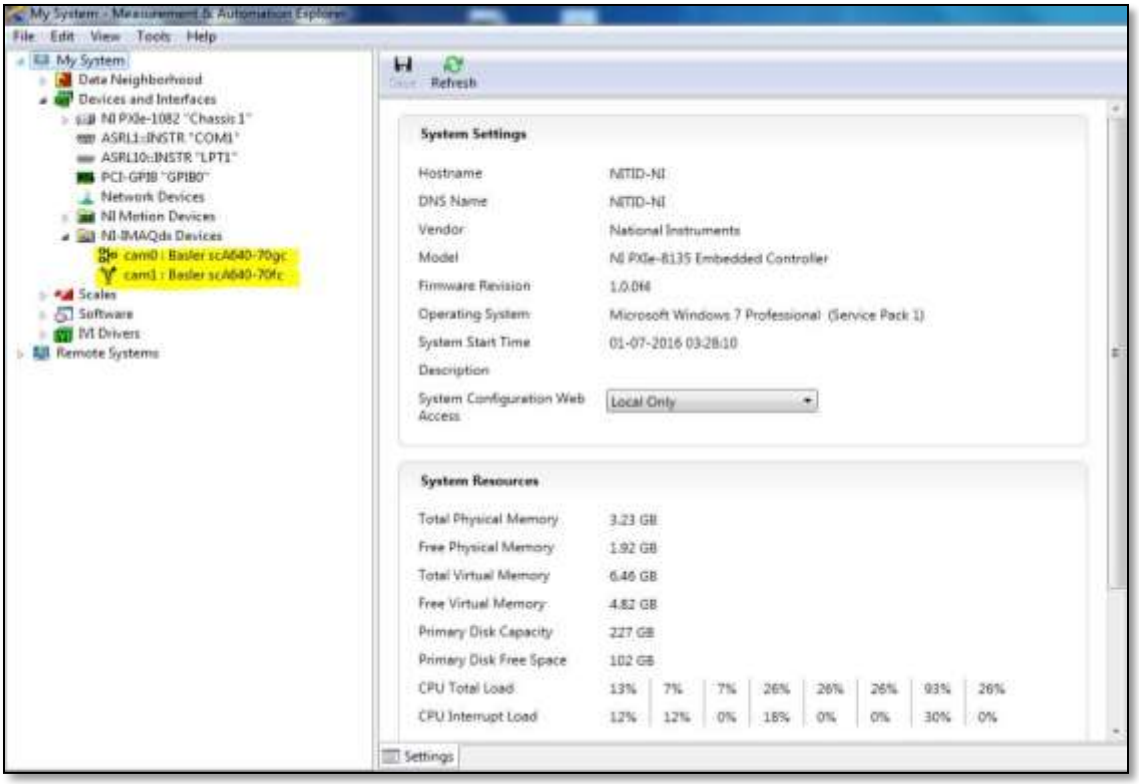


Figure 5.8: NI MAX Explorer: NI-IMAQ Devices

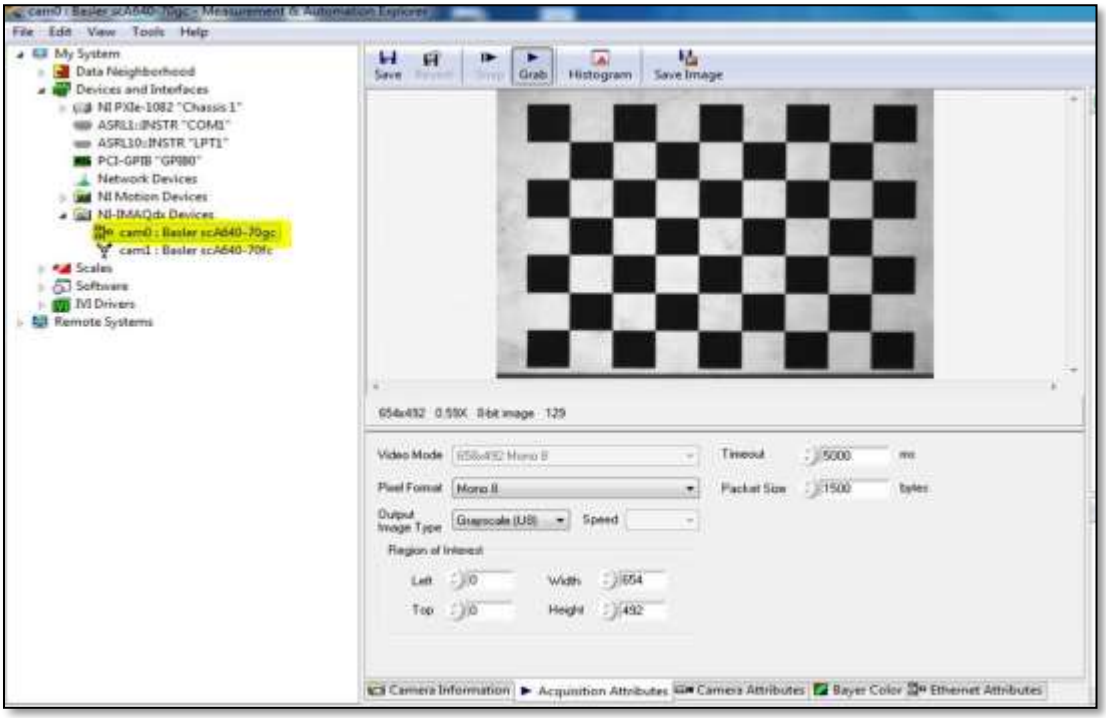


Figure 5.9: NI MAX Explorer: NI-IMAQ Devices

The camera properties can be set in this process. The configuration parameters such as brightness, auto-exposure, sharpness, white balance, saturation, gamma, shutter speed, gain and iris are set in this step. The desired properties of a camera are generally fixed during writing programs containing camera module. The Vision Acquisition module when used, it asked about the initialization of acquisition module. In this case, the camera attributes like brightness, color and contrast etc. are set. This setting helps in acquiring the best quality image.

5.3.2 Camera calibration

Detection, measurement or tracking of different objects in the real world by interfacing a stereo vision system with a robot is not a simple process. The whole process necessitates camera calibration. Camera calibration [Heikkila and Silvén, 1997; Zhang, 2000] aims at establishing a relationship between the image pixels and real world dimensions and helps to accomplish a high accuracy in terms of object detection, tracking or measurement.

Camera calibration is defined as the process of obtaining the specific parameters of a camera or a pair of stereo cameras which can help in rectifying a pair of stereo images. This is generally carried out by using images of a certain calibration pattern. The most general pattern used is a checkerboard. The calibration of camera is the first challenge in this phase. The most common method of overcoming the first stage of this challenge involves detection of a checkerboard pattern captured by a pair of stereo cameras. There are three parameters considered that influence the calibration process. These are as follows

- Intrinsic parameters of camera
- Extrinsic parameters of camera
- Distortion coefficients of camera

Accuracy in camera calibration is another aspect that needs to be taken care of. If the calibration is not accurate, the disparity map will be distorted and the 3D reconstruction of the desired scene would fail.

The accuracy of the specific parameters of the camera used to acquire the images is enhanced by implementing one of the following techniques

- Calculation of re-projection errors
- Calculation of the camera parameter estimation errors
- By plotting the relative location of the checkerboard used as calibration pattern and the camera used for image acquisition.

The first technique provides a good insight regarding the quality of calibration. The error is the difference in distance between a certain point detected in one of the calibrated images and the corresponding world point projected onto the same image. Generally the average of these re-projection errors is taken into consideration.

A good calibration is one that incurs errors less than 4 pixel value. If the stereo camera are properly calibrated, then these cameras could be used to reconstruct a three dimensional view of the acquired image. This is known as 3D scene reconstruction. This can also be done by using multiple images of the workspace using a single calibrated camera. Another alternative of such reconstruction involves using uncalibrated stereo pair of cameras that give an unscaled output.

MATLAB software provides a convenient method for calibration in the form of 'stereo camera calibrator' application. The relative position of one camera with respect to another can be found out using this application. Further it also provides an easy way to create an object containing the essential parameters of the stereo camera system. This object will be further helpful in carrying out various operations like image rectification 3D scene reconstruction and 3D location computation etc. The camera calibration is also checked in LabVIEW environment and discussed in image acquisition in next section. The process flow diagram of camera calibration used presented in Figure 5.10.

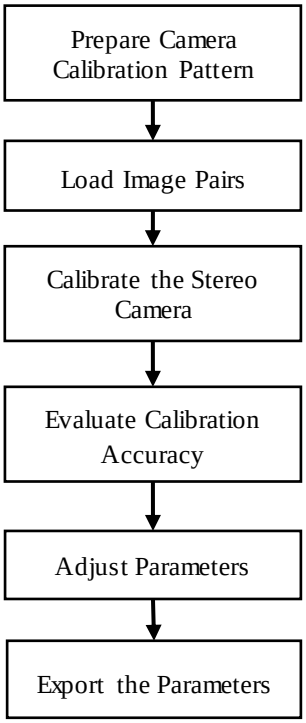


Figure 5.10: Flow diagram of camera calibration

The result of calibration step is given in Figure 5.11.

Figure 5.11: Camera calibration result

The intrinsic and extrinsic parameters are calculated during calibration process and are given in the Table 5.1 and Table 5.2 respectively.

Table 5.1: Intrinsic parameter of the camera attached to robot end effector

Focal length (pixels)	[519.9477 +/- 18.8421 550.1599 +/- 18.8377]
Principal point (pixels)	[258.3616 +/- 4.6466 208.6111 +/- 9.6353]
Skew	[5.6518 +/- 0.8261]
Radial distortion	[-0.0652 +/- 0.0261 0.0620 +/- 0.0963]

Table 5.2: Extrinsic parameters of the camera

Rotation vectors	[0.0946 +/- 0.0121	0.3800 +/- 0.0132	-1.5634 +/- 0.0030]
	[0.1952 +/- 0.0119	-0.2593 +/- 0.0115	-0.0012 +/- 0.0025]
	[0.0089 +/- 0.0109	0.2019 +/- 0.0139	-1.5661 +/- 0.0018]
	[0.2116 +/- 0.0146	0.0240 +/- 0.0095	-1.5789 +/- 0.0016]
	[0.0600 +/- 0.0088	0.2142 +/- 0.0132	-1.5533 +/- 0.0021]
	[0.1154 +/- 0.0148	0.0094 +/- 0.0105	0.0153 +/- 0.0014]
Translation vectors	[42.1001 +/- 4.3553	86.3477 +/- 8.6684	501.1429 +/- 17.6641]
	[-59.7427 +/- 3.9015	-72.6382 +/- 7.6510	431.2895 +/- 15.2148]
	[125.9464 +/- 5.8659	52.2380 +/- 11.4787	660.2275 +/- 23.7385]
	[118.8004 +/- 5.4618	58.6739 +/- 10.7327	616.3565 +/- 22.2593]
	[-125.4635 +/- 5.6810	55.8394 +/- 11.4233	656.0964 +/- 23.5766]
	[-135.2389 +/- 5.9376	-176.3618 +/- 11.7526	651.7334 +/- 23.5035]

5.3.3 Image Acquisition

Image acquisition [Lindley, 1991] in image processing can be broadly defined as the action of retrieving an image from some source. Normally, in this process a hardware-based source is used which can be passed irrespective of the processes involved. The typical image acquisition process is shown in Figure 5.12 where light source, optical device and object are used to acquire image. Image acquisition in image processing is always the first step in the workflow sequence because, without an image, no processing is possible. The image that is acquired is completely unprocessed and is the depended on the hardware setup used. The hardware setup plays an important role in maintaining the baseline for image processing. The source must be properly handled and controlled that can reproduce the same image as and when required. The reproduction of same image makes it easy to locate and eliminate the anomalous factors present in the image if any.

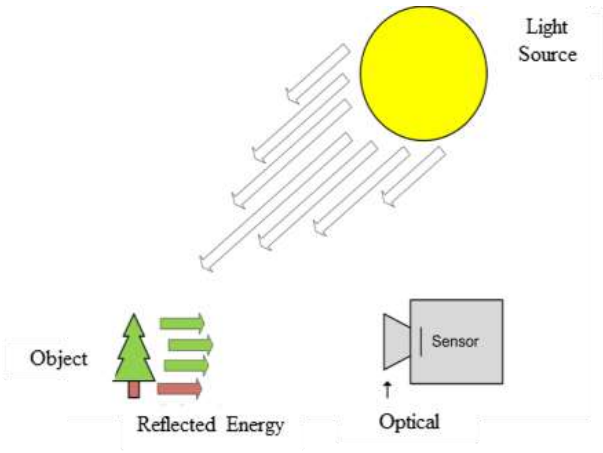


Figure 5.12: Typical image acquisition process

There are four image acquisition types supported by the Vision Toolkit such as Snap, Grab, Sequence and Still Color. In NI-IMAQ image acquisition process open IMAQ session is also used to determine a plethora of camera, image and session properties. This open session returns the type of image which is determined by either camera initialization or is set dynamically. This value of image type can be passed to IMAQ Create, reserving the correct number of bytes per pixel for the acquired image data.

In LabVIEW environment, Vision Acquisition is set by NI Vision Acquisition Express (Figure 5.13). Five tasks are performed under this condition. The first task is to selection of acquisition source in which the source of image acquisition is chosen. Second one is to select the acquisition type (as in Figure 5.14), where the single acquisition or continuous acquisition with inline processing is selected as per requirement. The third one is configuring the acquisition setting for selected source device attributes. This contains the detailed information and status of the device. Here the device is the camera. The fourth task is to configure the image logging. This task deals with the storing information of the acquired images i.e., file path, file name prefix and format of the image file. The acquisition rate is highly affected by enabling the image logging. The last one is to select the controls and indicators like image in and out, error in and out etc.

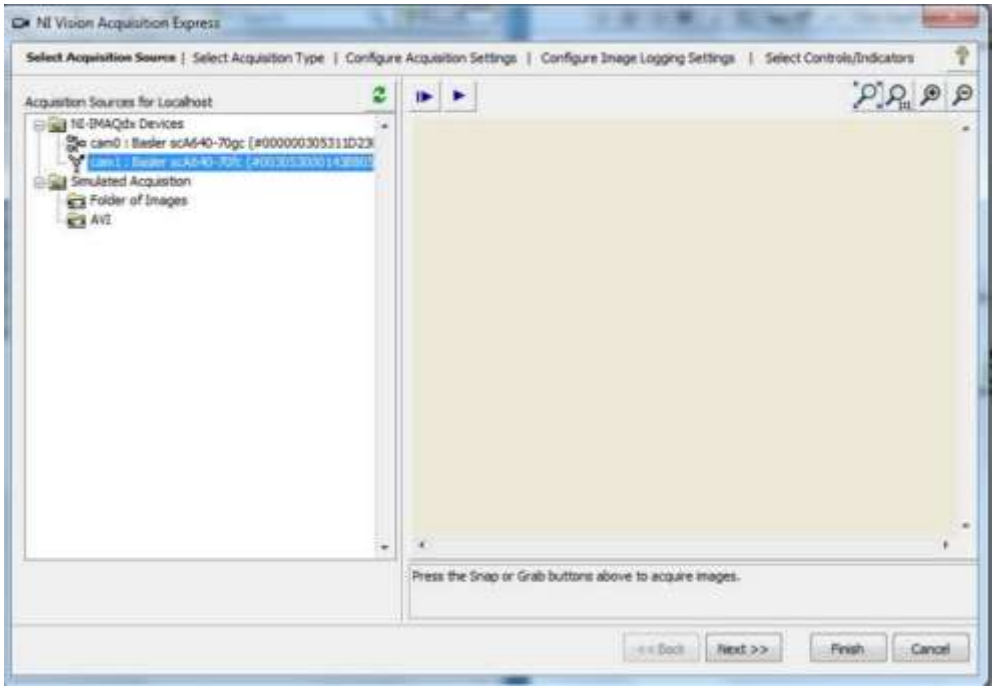


Figure 5.13: NI Vision Acquisition Express (Selection of image source)

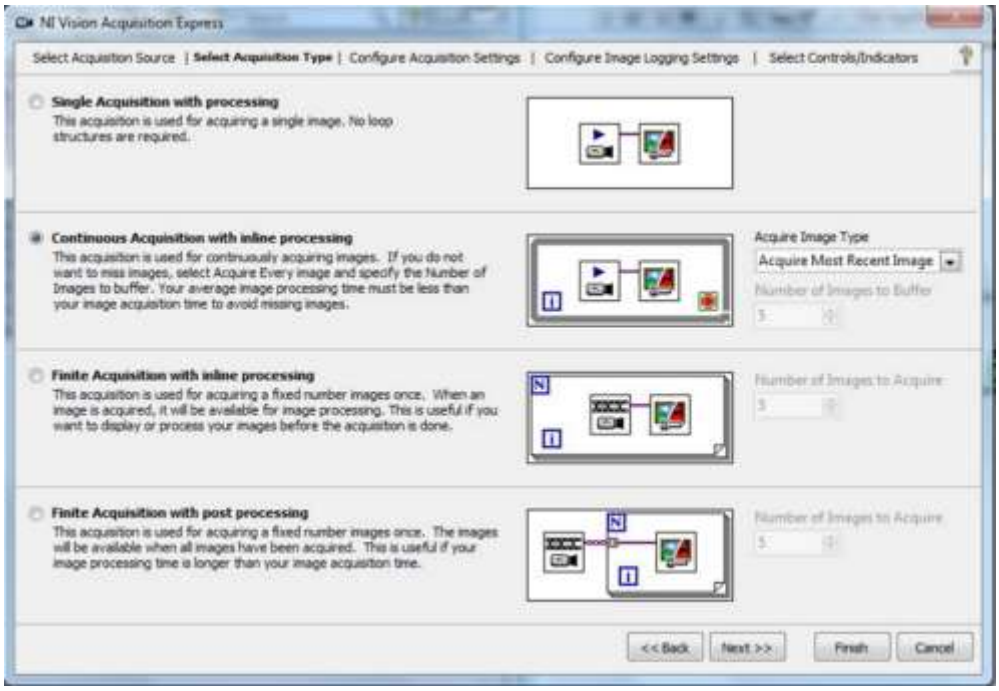


Figure 5.14: Selection of image acquisition type

5.3.4 Illumination and Thresholding

A. Illumination

The illumination is the most important part of any imaging system. Lighting environments need to be optimized and tuned for a camera to pick up things. In machine vision application, illumination provides better contrast to make object clearly visible to camera. For an effective imaging system, sufficient amount of uniform light and be from the correct angle is needed to avoid the effects of shadows.

Generally, illumination is considered at early stage during the development of a machine vision system. The illumination is dependent on the light source used in the environment.

Commonly, three basic types of illuminations are used for machine vision with respect to light source. They are fiber optic, fluorescent and LED based. Each illumination type is best suited for specific application. The LED illuminators are less intense but are reasonably inexpensive with a longer service lifetime. The fiber optic illumination is very intense, more expensive and has a shorter service life. The Fluorescents type comes in the middle of the category, offering moderate intensity with a service life to match the expected cost. The standard fluorescent tubes are used for illumination purpose. The problem arise when the object is a colored one. The color of the illumination used with monochrome cameras is not so important for monochrome objects. A red colored object appears as white when illuminated with red light. Similarly for green and blue colored objects. Therefore, it is important that the illumination has same amount of all the colors (RED-GREEN-BLUE) for clear and accurate visibility of the object.

The illumination technique is considered with respect to the type of objects used in the application. The objects can be dull or reflective, flat or complex in shape. Illumination method is broadly classified as two category as bright-field illumination and dark-field illumination.

Bright-field illumination is the most common technique of illuminating diffuse objects. The term bright-field refers to the mounting position of the illuminator. If a camera is positioned pointing towards a plain mirror, the bright-field is the area in which any reflected light is within the FOV (Field of view) of the camera. Similarly, dark-field illumination technique is used to highlight surface of the objects. It is also obvious that optimizing the illumination can reduce the use of filters required for correct analysis of image. The images resulting from poor illumination (due to non-uniform light) are difficult to segment.

B. Thresholding

Image segmentation [Russ and Woods, 1995; Kitney et al., 2005] involves in separating an image into regions (or their contours) corresponding to objects. The intensity value of pixels represents an image. Image thresholding is an effective method of partitioning an image into a foreground and background. This is a type of image segmentation which isolates the objects by converting the grayscale images into binary images. Image thresholding is most effective in images with high levels of contrast.

From a grayscale image, thresholding can be used to create binary images in which each pixel of the image is to be replaced by a black or white pixel with respect to their intensity value.

The threshold $g(x,y)$ of image $I(x,y)$ with respect to some global threshold value T is defined as

$$g(x,y) = \begin{cases} 1 \text{ if } I(x,y) \geq T \\ 0 \text{ otherwise} \end{cases} \quad (5.1)$$

There are two types of thresholding algorithms such as:

- Global thresholding algorithms
- Local or adaptive thresholding algorithms

In global thresholding, a single threshold "T" for all the image pixels is used. When the pixel values of the components and that of background are fairly consistent in their respective values over the entire image, global thresholding could be used. In adaptive thresholding, different threshold values for different local areas are used.

5.4 Robot Controller Installation

The automatic generation and execution of plans for part assembly tasks is one of the most important and complex objectives of Robotics. A Kawasaki R-series robot (RS06L) is used in the development of the proposed system. The technical specification of the robot is given in Chapter 3. The installation and setup of the robotic system is important for the success of the system suggested. The robot RS06L and the end effector (SCHUNK) is fitted perfectly and then connected with the robot controller. Some operation for robot arm control is performed in teach mode. The flow of activities for the robot controller installation is given in Figure 5.15 and the setup is given in fig. 5.16. The Kawasaki robot is controlled by a software-based system called AS language. AS language is used for communication with robots and it controls the robot according to the commands and programs. A computer system is connected to the robot controller by a terminal software KCwinTCP.

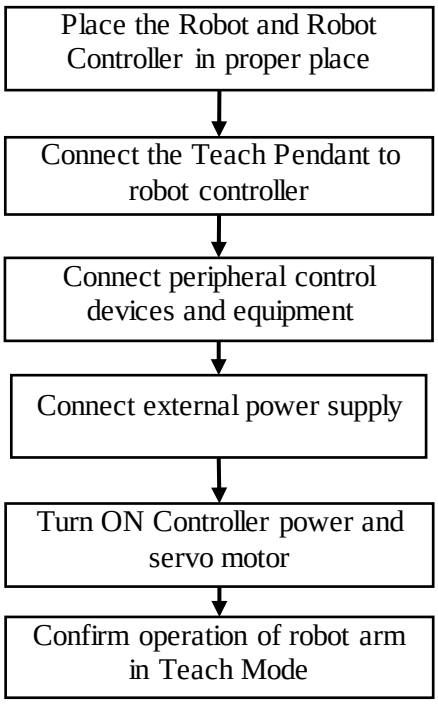


Figure 5.15: Steps for installing the robot controller

Figure 5.16: Robot System

5.5 Integration of Robot with Vision System

This step is an important step in the system development. Initially, the robot is connected to the vision processor through a terminal software KCwinTCP [http://sine.ni.com/nips/cds/view/p/lang/en/nid/211069/overview]. For this the robot address and vision system address are set. Here,

the robot address 192.168.1.1 and the system address 192.168.1.3 are mapped by the terminal software. Once they are connected, a third party software is used for controlling the robot by the LabVIEW programming environment. Digi-Matrix Kawasaki Robot Library is the third party software which provides a programming environment for giving commands to robot. This software enables the user to develop applications for automated test, laboratory automation, and automated assembly setup. This software removes the complexity in writing programs in AS language for robot control. Initially, a program is written to check the communication status between robot and vision processor as shown in Figure 5.17 and Figure 5.18.

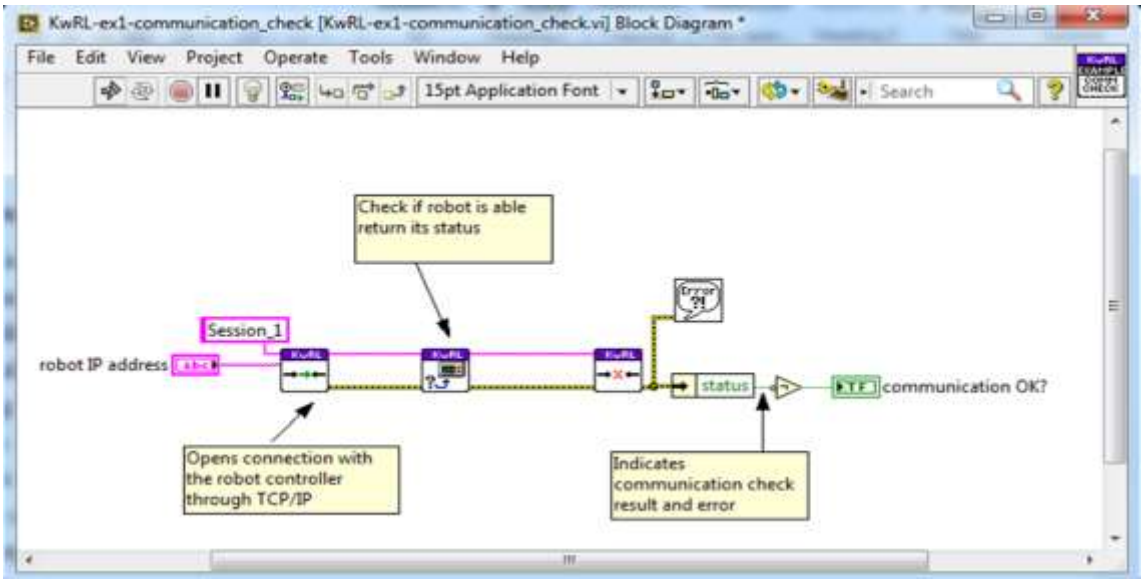


Figure 5.17: Communication check program functional module

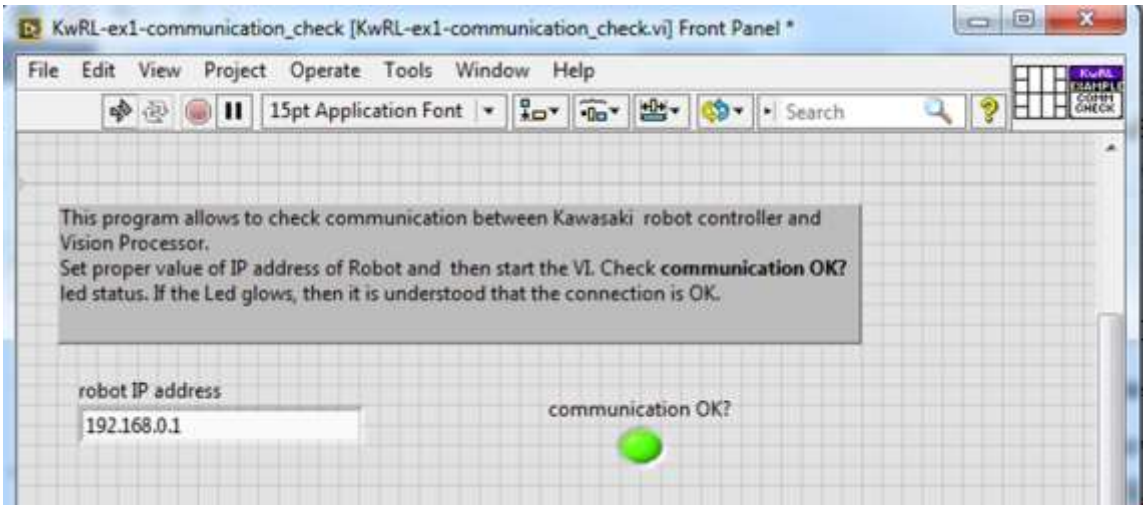


Figure 5.18: Communication check program front panel

A separate program is written to get the status of the robot and its controller. This shows whether the power is ON/OFF. This program generates an error signal if any error is occurred in the robot controller. The program is also shows the current mode of operation of controller i.e., Teach/Repeat. The front panel is shown in Figure 5.19 and its functional block diagram in Figure 5.20.

Figure 5.19: Get the status of the robot Controller (Front panel)

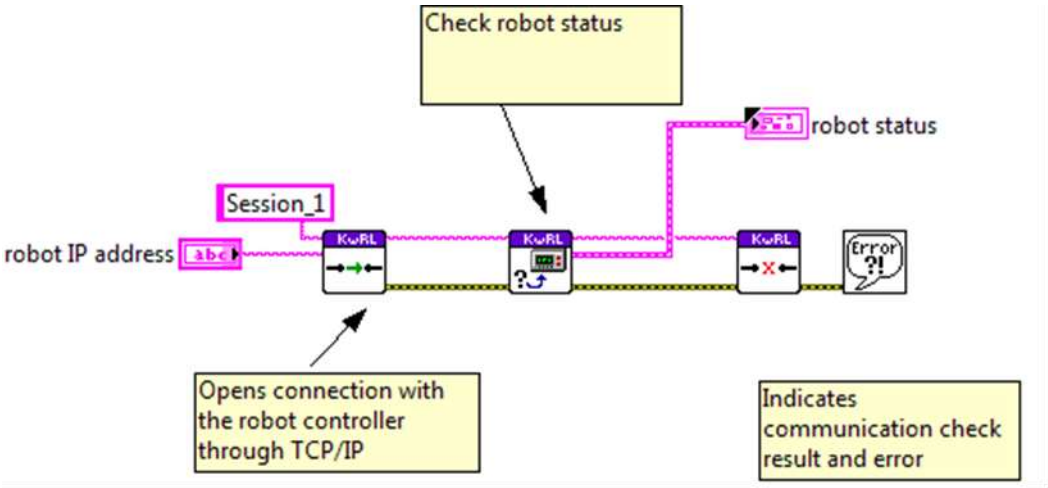


Figure 5.20: Functional Block Diagram to get the status of the robot Controller

Similarly, the servo motor is put in ON condition for robot to act. The servo ON/OFF is done by the program servo_On_Off.vi. This program makes the servo motor to ON/OFF as per the requirement. The front panel and functional block are shown in Figure 5.21 and Figure 5.22.

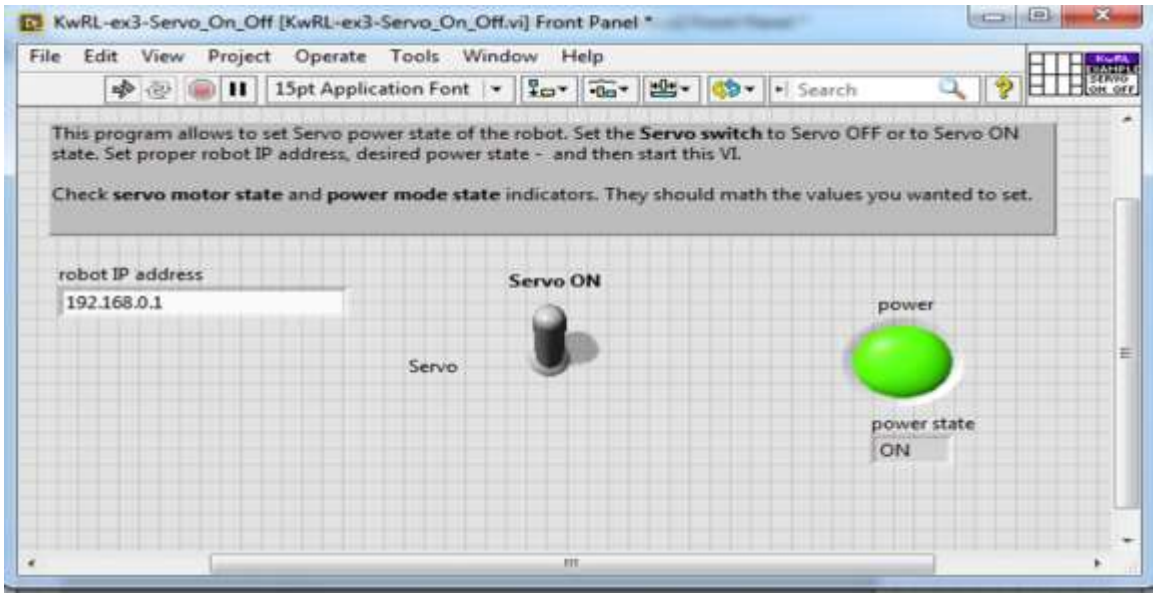


Figure 5.21: Power state of the servo motor of the robot (Front Panel)

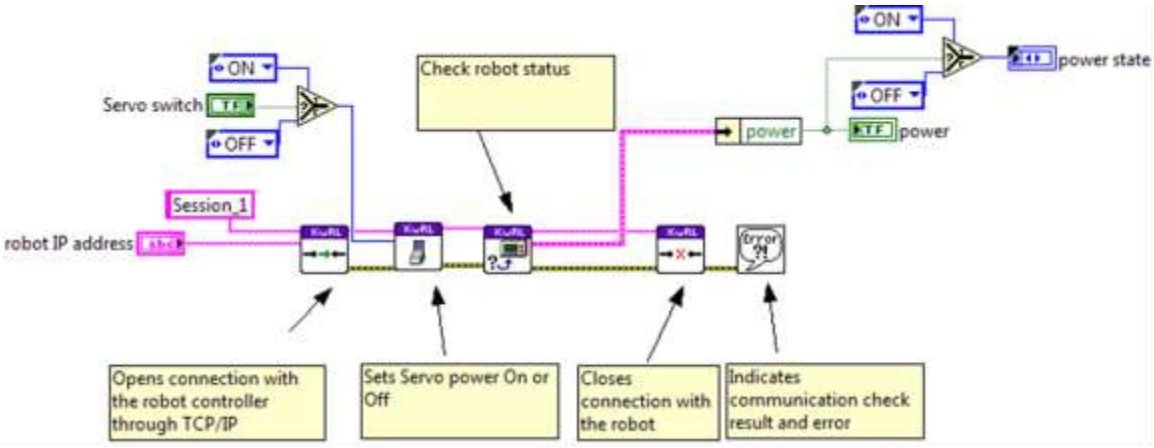


Figure 5.22: Power state of the servo motor of the robot (Functional Block)

A program digital_in_digital_out.vi allows to set control to the desired digital output line number, set the desired line state to set value. Set the line number to get control to the desired digital input or output line number and specify the line type. The front panel and block diagram are given in Figure 5.23 and Figure 5.24. Now, the robot is ready for any operation. The current position of the robot is determined by a LabVIEW program (as shown in Figure 5.25 and Figure 5.26). Actual operation by the robot is performed in program simple_move.vi as shown in Figure 5.27 and Figure 5.28.

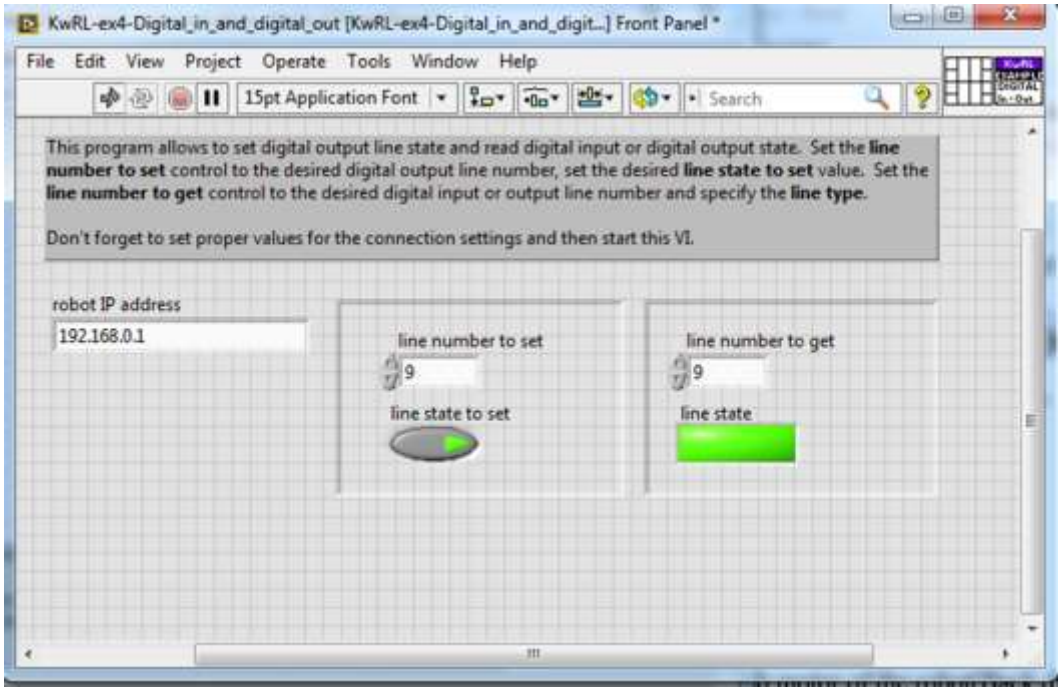


Figure 5.23: Program for setting digital in and digital out line number (Front Panel)

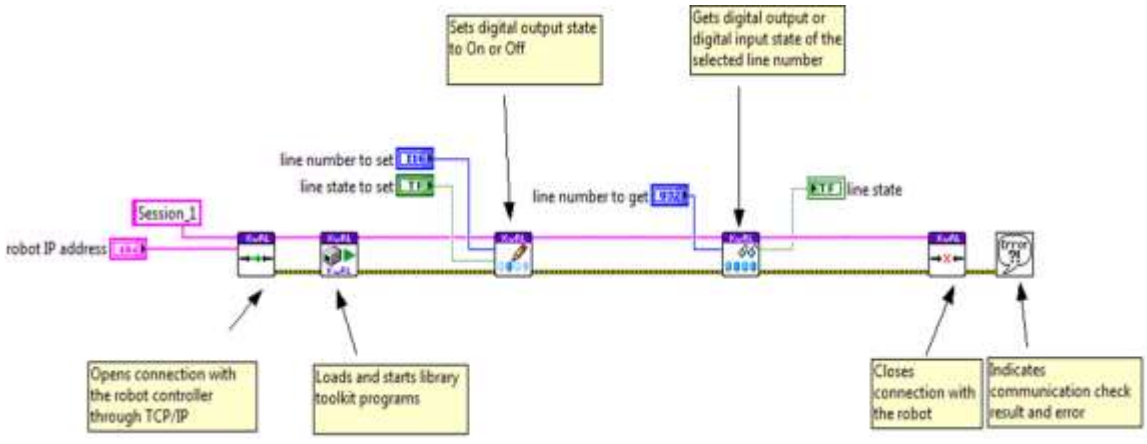


Figure 5.24: Program for setting digital in and digital out line number (Block Diagram)

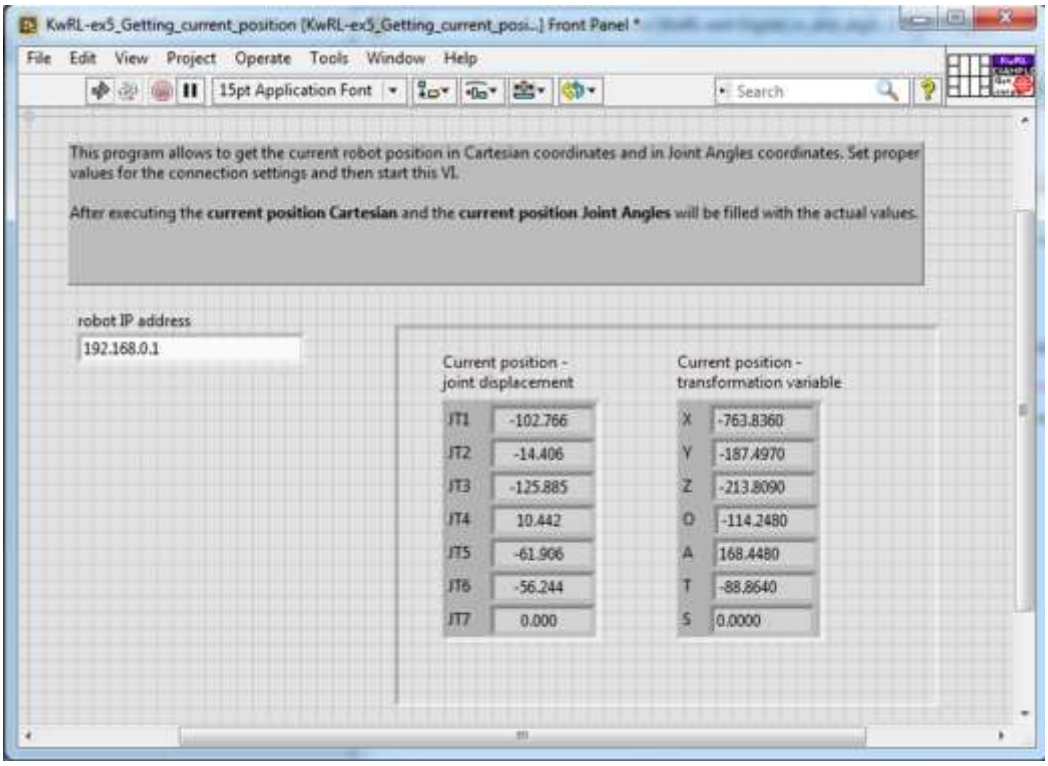


Figure 5.25: Compute current position of robot (Front Panel)

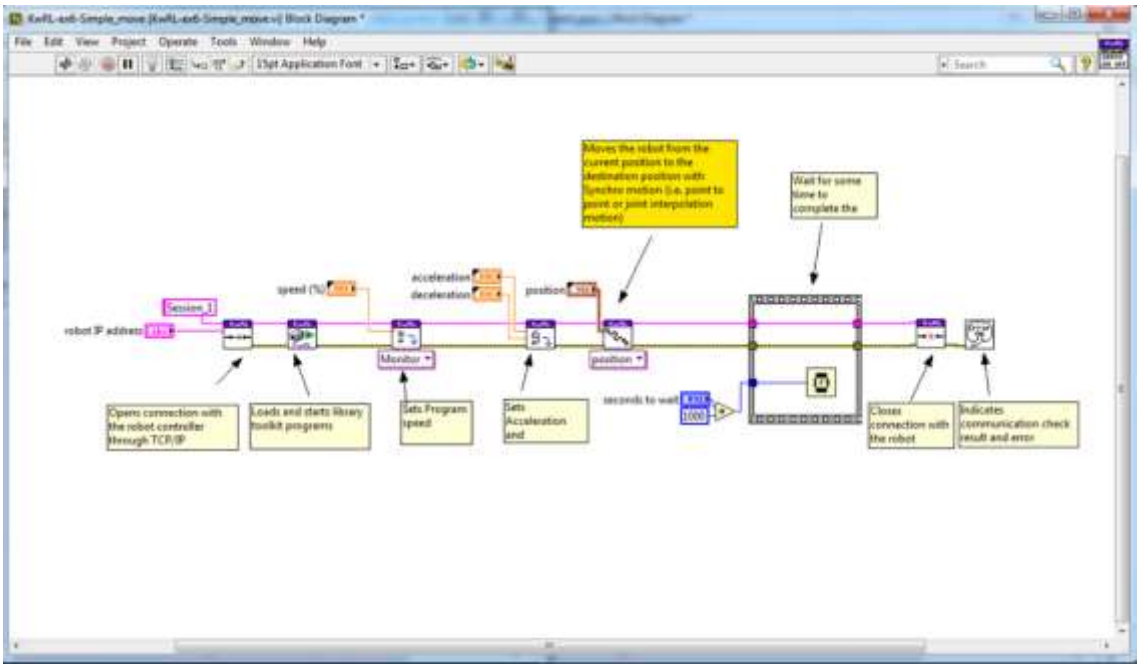


Figure 5.26: Compute current position of robot (Block Diagram)

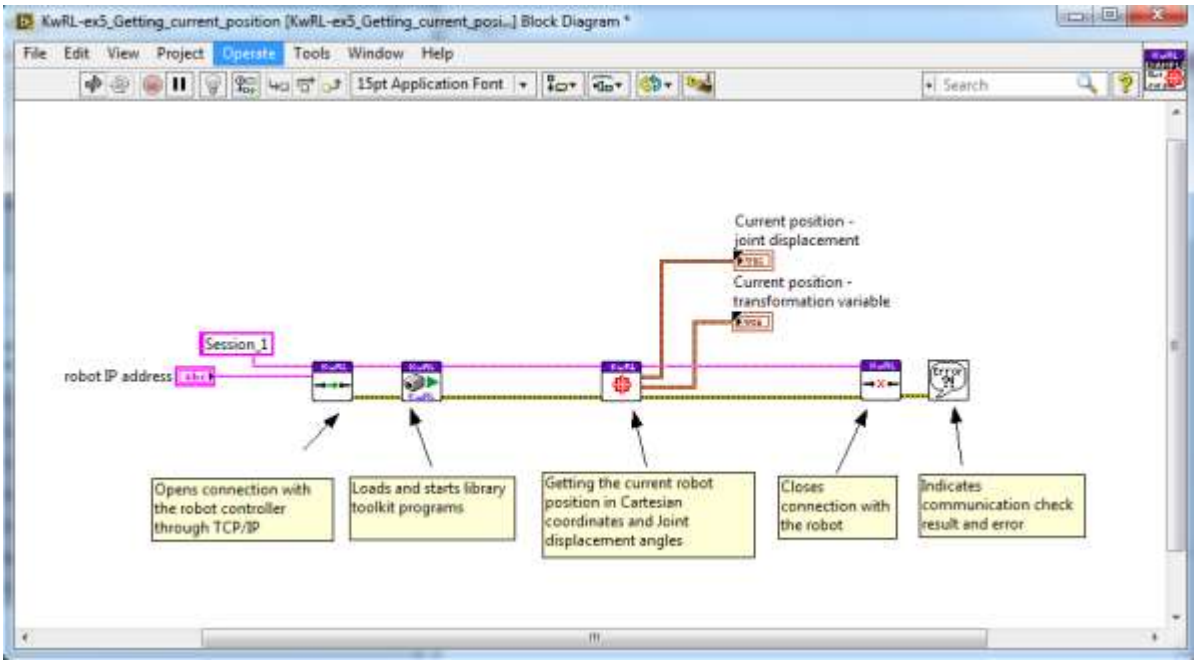


Figure 5.27: Program for a simple move of robot (Front Panel)

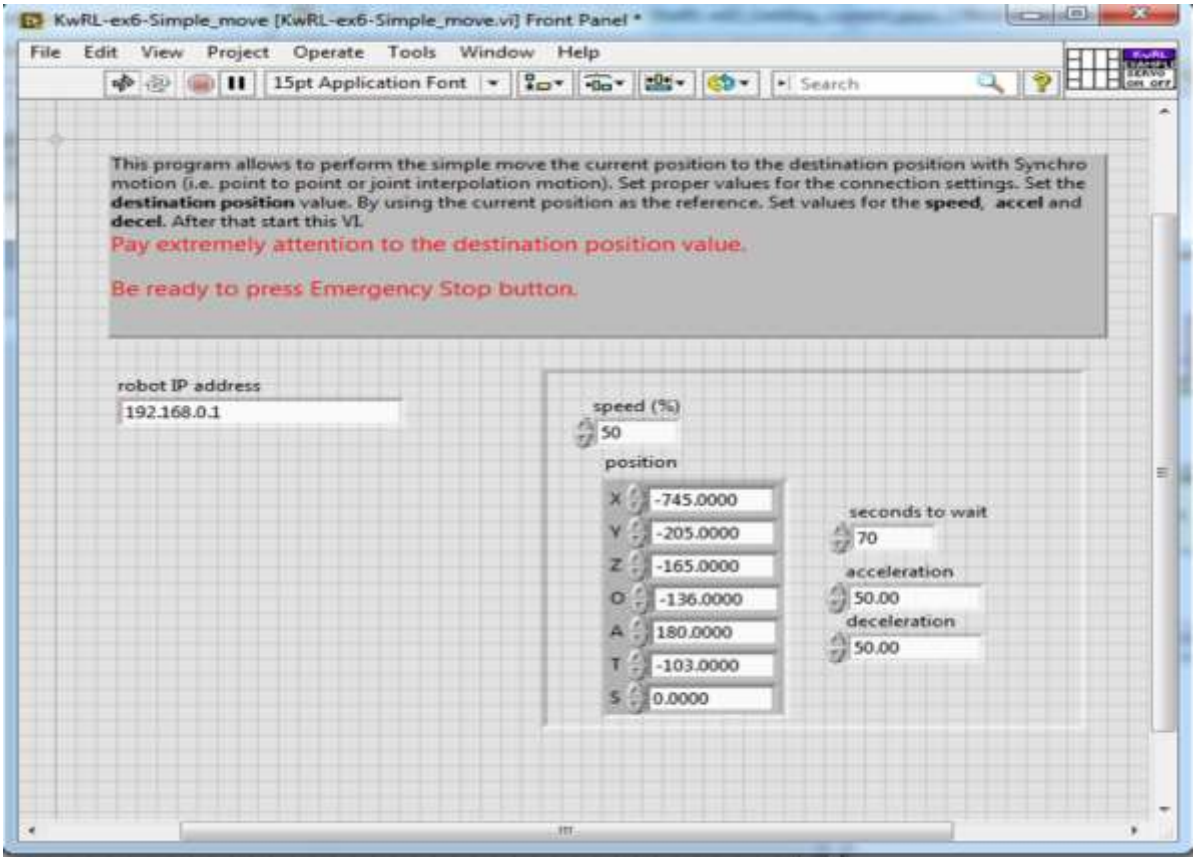


Figure 5.28: Program for a simple move of robot (Block Diagram)

5.5.1 Proposed Workspace

The technical specification of the Kawasaki robot (model RS06L) used in this work is given in the Table 3.1. The workspace of robot manipulator is defined as the set of points that can be reached by its end-effector. In other word, the work space of a robot arm is the set of positions, consisting of both a reference point and the orientation about this point, that are reachable by its end effector. A robot is designed so its end effector has unconstrained freedom of movement within its workspace. However, this workspace does have boundaries, defined in part by extreme reach allowed by the chain. The shape, size of the workspace for a robot is a primary consideration in its design.

The workspace of a linkage is defined by identifying a specific link as the work piece. Then the workspace is the set of positions that this workspace can reach. For serial open chains and platform linkages the dimension of the workspace is exactly the generic mobility F of the system, when F<=K. If the mobility F of the linkage is greater than the unconstrained freedom K, then the system is said to have redundant degree of freedom. The system developed for the application is given in Figure 5.29.

Figure 5.29: Developed System Model

The important consideration in designing the workspace is the dimension and structure (shape and volume) of the workspace. These aspects have a significant importance due to their impact on the design and manipulability of the robot. The precise knowledge about these factors are important as:

- The shape plays an important role in defining the working environment of the robot.
- The dimension determines the reachable location of the end-effector.
- The structure of workspace is important for assuring kinematic characteristics of the robot which are in relation with the interactions of the robot to its environment.

In a part assembly environment, the robot is picking part from a specific location and placing it in a desired place. So, it is understood that, the robot gripper is moving in a region specified according to the application. In this work, the movement of the gripper is restricted to some value along x,y and z directions which is given in the Figure 5.30.

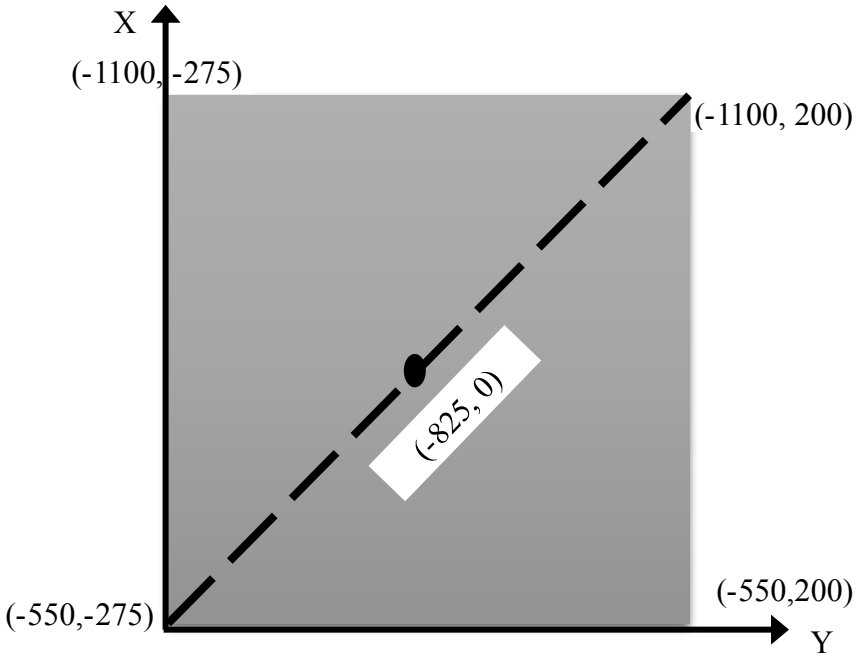


Figure 5.30: Workspace of the robot

Moreover, shape, dimensions and structure of the workspace is dependent on the robot features. The design constraints of robot workspace with respect to robot features are as follows:

- The dimensions of each link of the robot and the mechanical boundaries of each joints (both active and passive) affect the design workspace with respect to its dimension.

- The shape of the workspace depends on the geometrical structure of the robot (interference between links).
- The structure of the workspace is also governed by the dimensions of links and the structure of the robot.

The movement along x-axis is restricted to -550 to -1100. It is because, the parts are placed within this region. Similarly, it is restricted to -275 to 200 along y-axis. The downward movement along z-axis is limited to 4. These restrictions is vital for the success of the application. The model of the workspace for left-right, upward-downward and forward-backward are shown in Figure 5.31 (a), (b) and (c) respectively.

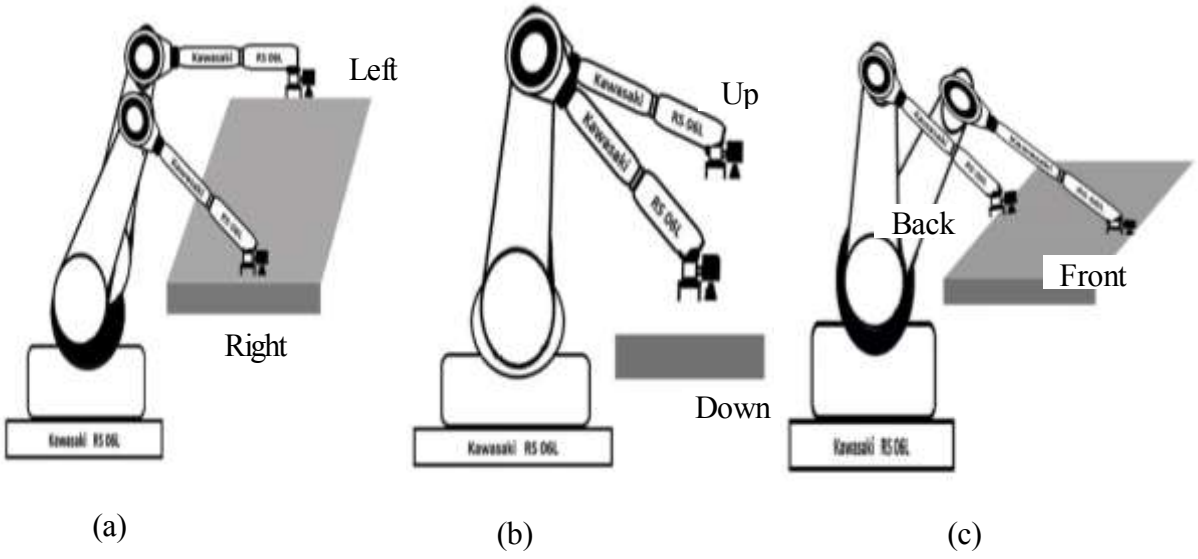


Figure 5.31: Movement of Robot gripper in the specified workspace (a) Left-right movement, (b) Upward-downward movement and (c) Backward-forward movement

There may be the case the parts are available different locations inside the workspace. It is desirable that the end effector to reach at the part in a shortest path. Considering this, program is developed in LabVIEW, which make the robot to move along the diagonal direction. The robot also moves along x-axis and y-axis direction. Suppose the desired part is available at a position just near to the center of the workspace. In this case, the robot automatically finds the shortest path by moving along the diagonal and then along x-axis or y-axis direction. Such a case is given in the Figure 5.32. The front panel diagram and functional block of the program are shown in Figure 5.33 and Figure 5.34 respectively.

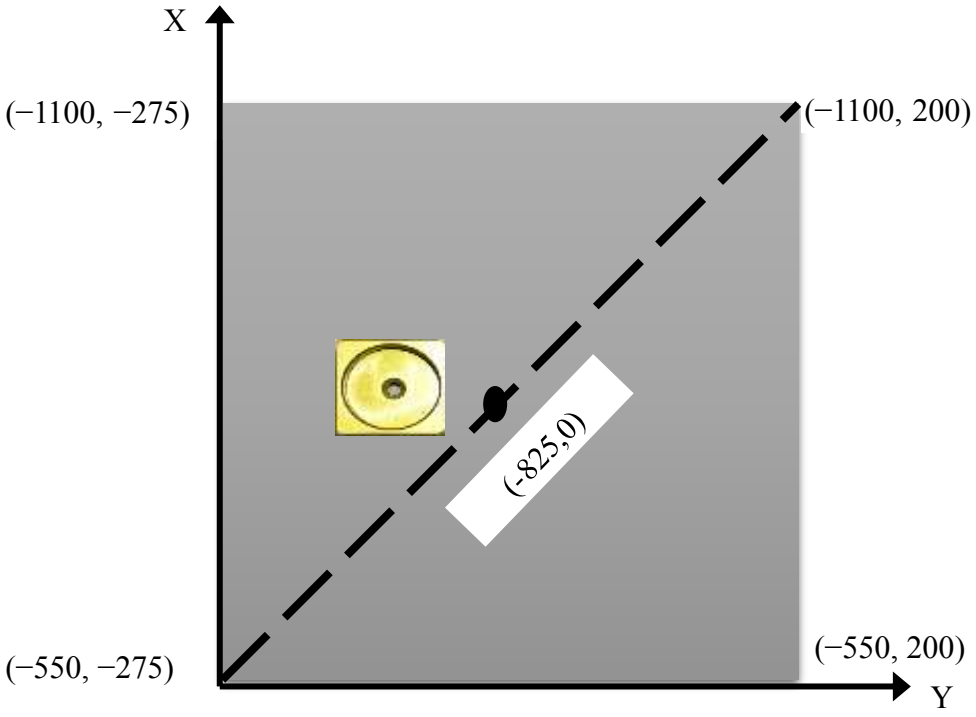


Figure 5.32: Part present in the workspace near to center

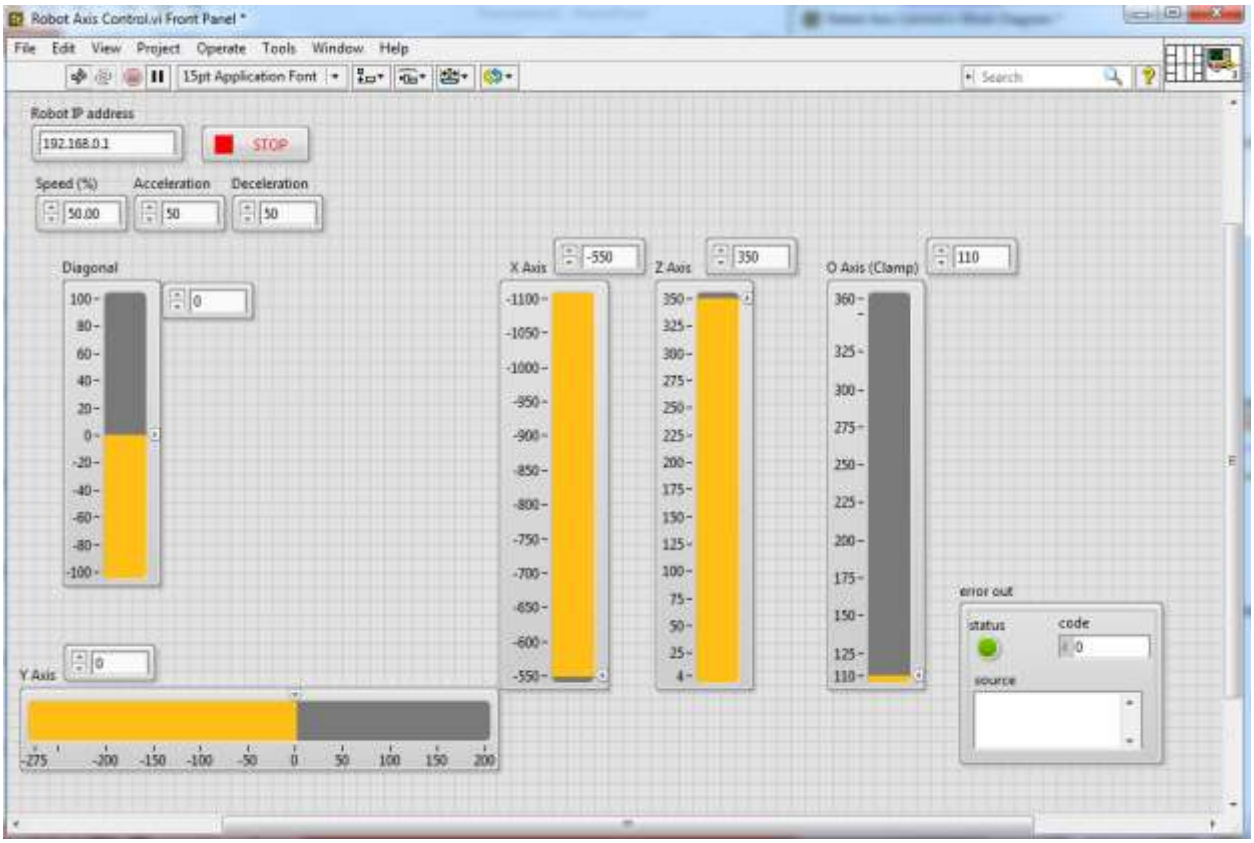


Figure 5.33: Robot axis control program (front panel)

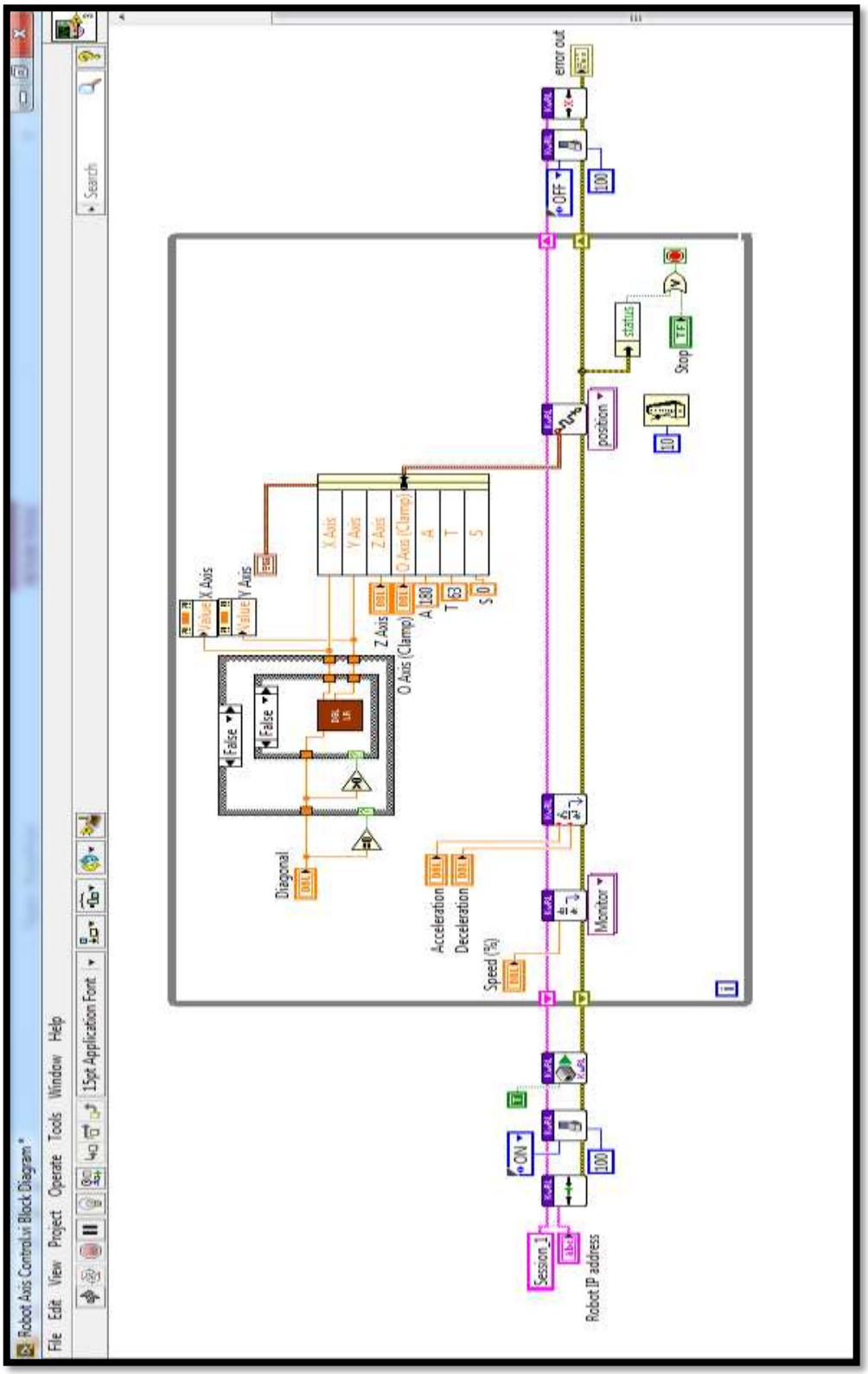


Figure 5.34: Robot axis control program (functional block)

5.5.2 Method of Integration

The success in development of the proposed system lies in effective integration of machine vision system with the robot system. Before integrating the robot system with machine vision system, it is assumed that each system is configured accurately as per the requirement in the application. Going in detail, the attached hardware like vision sensors (camera), data acquisition system, and motion controller are configured by the platform NI-Max (Measurement and Automation Explorer) provided by National Instruments. During this configuration, the camera attributes and acquisition attributes are set as required by the application. This steps are shown in Figure 5.8 and Figure 5.9. The proper calibration of camera is a key factor for acquisition of better quality images. The steps followed for calibrating the camera is shown in Figure 5.10. The calibration process is performed in the LabVIEW environment as shown in Figure 5.11. Then the robot system is configured to connect to the vision system via Ethernet port. The robot is set to be in repeat mode. For a successful connection, the robot IP and vision system IP is set. In this work, the robot IP is set as 192.168.0.1 and the vision system IP is 192.168.0.3. The status of the connection is checked by a program as shown in Figure 5.17 and Figure 5.18. This program allows to check the status of the communication between vision processor and robot controller. Similarly, the status of the robot controller is determined by executing the program as in Figure 5.19 and Figure 5.20. This program results in determining the power status, modes of operation, error status if any, status of the current task if present and accuracy of the controller. The status information of the robot are important for feeding any task to the robot controller through vision processor. The robot can be controlled by the vision system, if it is in "REPEAT" mode. The robot can be trained in "TEACH" mode. A program is executed to make the servo motor ON. This is required in advance for giving commands to the robot. This is shown in Figure 5.21 and Figure 5.22. This enables the robot to execute tasks. Now, each joints of the robot is checked by examining the digital line that connects it with the robot controller. The program shown in Figure 5.23 and Figure 5.24 is executed to check the status of each digital line. A program is executed to determine the current position of the robot in Cartesian coordinates with respect to each joint as shown in Figure 5.25 and Figure 5.26. All these programs are executed to ensure that all requisite conditions are satisfied before making the robot to move from a position to a desired position. For a simple move, the desired position coordinate is given as input to the program and also the speed, acceleration and deceleration are set. The robot moves from current position to a desired position with synchro motion i.e., point to point or joint

interpolation motion. The algorithm for performing a simple movement by the robot autonomously is given below.

Algorithm 5.1

Assumption: The robot and the machine vision system are configured properly.

Step 1. Check the status of the communication between machine vision system and robot controller.

Step 2. Get the status of the robot.

Step 3. Make the mode of the robot controller to "REPEAT"

Step 4. Make the servo motor ON

Step 5. Check the digital line for each joints of the robot

Step 6: Determine the current position of the robot with respect to joint displacement in Cartesian coordinate

Step 7. Provide the desired position coordinates value for each joint

Step 8: Move the robot to a desired position

5.6 Summary

It is important to integrate the vision system with that of the robot in order to make the robot visually active. The process of integration follows certain protocols and takes help of few algorithms for its implementation. This chapter contains the details of the process followed for integrating the two systems and finally testing the vision integrated robotic system for desired and correct operation.

Chapter 6

Visual Servoing for Robot Navigation

6.1 Overview

There is increasing demand for automation in many manufacturing scenarios especially like robotic assembly system, bin picking system and palletizing system to increase the throughputs and to reduce the manufacturing cost. This requires accurate and fast robot positioning with respect to the part to be used next. Normally, the 3D machine vision system is used for compensating the robotic positioning errors due to randomly placed parts and unknown work environment. Control of robot motion via the information extracted from the images captured by single or multiple cameras is considered as visual servoing (VS) [Agin, 1980; Chaumette and Hutchinson, 2006; Chaumette and Hutchinson, 2007; Zhao et al., 2015]. A family of closed loop control methods to regulate the degrees of freedom (DoF) of an actuated system with visual feedback is the main purpose of VS. The task in VS is to use vision system to control the pose of the robot's end effector relative to a target object or a set of target features.

Visual sensing and manipulation of robot are combined in an "open-loop" or "look and then move" procedure. The accuracy of such procedure depends directly on the visual sensor and the control strategy. The alternative approach to increase the accuracy of such sub-system is the use of visual-feedback. This visual-feedback control is called visual servoing (VS). The process of minimizing a visually-specified task by using visual feedback for motion control of a robot is called VS. This is the result from fusion of many elementary areas like image processing, kinematics, control theory and real time computation. The VS is about the use of visual feature or vision data for a guided motion of robot. The visual features are generally acquired from a vision sensor that is mounted on a robot manipulator or fixed on a position to watch the manipulator. The success of VS also depends on the control law designed based on these acquired visual features. So, the control law for servoing is designed based on the position of the camera and also the number of camera used. The basic block diagram of configuration of robot control based on visual information feedback is given in Figure 6.1. In this configuration, the robot motion is guided by the camera in motion. Alternatively, the camera can be fixed in the workspace so that it can observe the robot motion from a stationary configuration. Various types of VS control schemes and their dependencies are discussed in the next section.

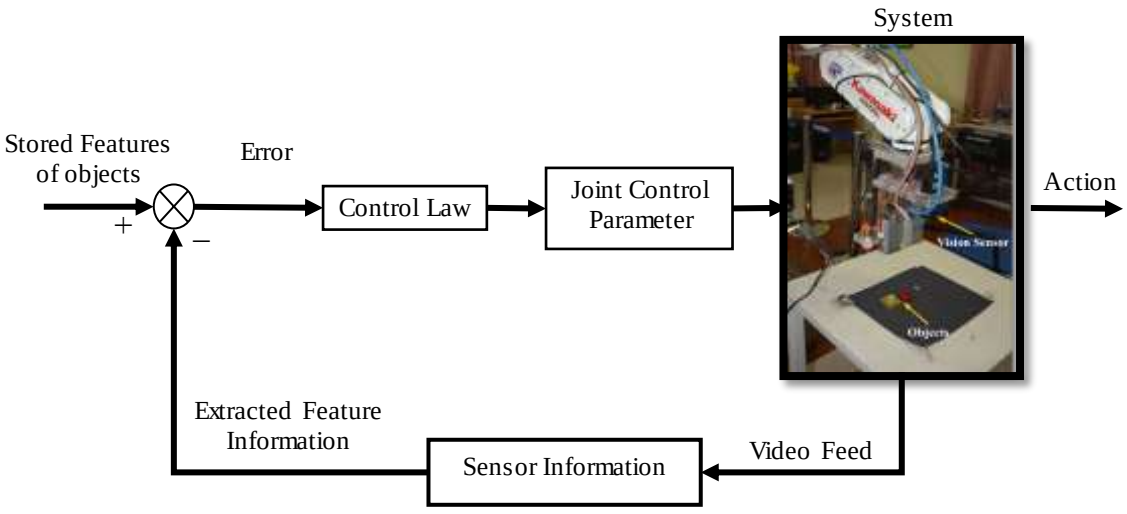


Figure 6.1: Block diagram of visual servoing scheme

6.2 Classification

Visual Servoing is, in essence, a method for robot control where the sensor used is a camera (visual sensor). There are two fundamental configurations of VS [Chaumette and Hutchinson, 2007] based on the position of camera with respect to the robot system such as eye-in-hand as in Figure 6.2 and eye-to-hand as in Figure 6.3. In eye-in hand configuration, the camera is attached to the robot end effector. This is otherwise called as end-point closed-loop control configuration. The best viewpoints of the environment can be achieved in this configuration. Such a system could be built using two robot arms, one holding a camera head, and the other doing the manipulation. But synchronization can be an important issue. This problem can be suppressed by mounting the cameras on the same arm doing the manipulation. This approach gives good overviews of object of interest by raising the manipulator high over the workspace during operation. The camera can be taken close to the object for a better precision, when picking up and manipulation is required. In this configuration, the camera is observing the relative position of the object. Such configuration is less sensitive to kinematic calibration error but it fails when the hand is obscured. The accuracy of VS using control approach depends on the camera model and 3D pose reconstruction method.

In "eye-to-hand" (or end-point open-loop control), the camera is fixed in the world co-ordinate system which observes the target and motion of the hand. This configuration requires the robot kinematics information for a guided motion of the robot. The VS control with this type of

configuration is not affected by occlusion of the hand. A clear disadvantage of fixed camera configuration is that, the workspace is uniformly sampled in the digital image. Often it is required a high resolution visual information in a few special areas, such as around an object when grasping and picking it up, or for doing fine manipulation when aligning an object for insertion.

Variable resolution results more important visual information about the object which makes the control scheme efficient. This can only be achieved in an eye-in-hand configuration. This is the factor which force to consider such configuration in development of proposed system. However, the adaptiveness is crucial to handle the scale changes as the robot moves along the optic axis of the workspace.

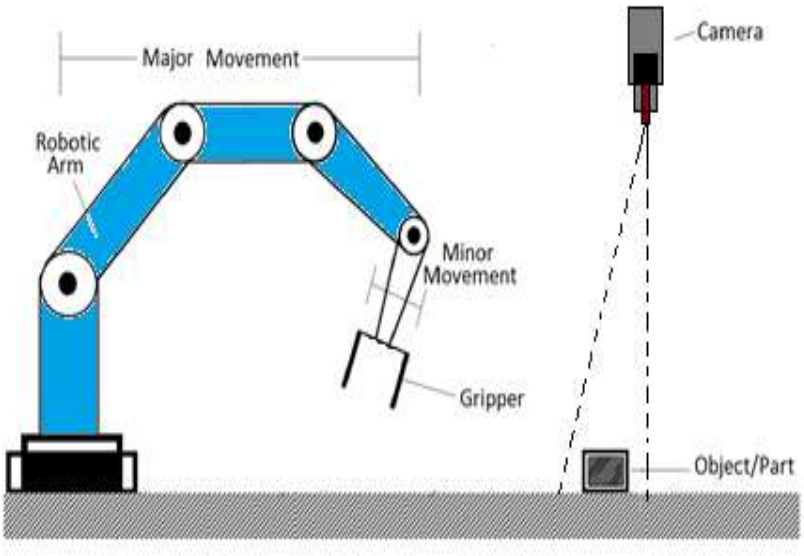


Figure 6.2: Eye-in-hand vision based robot configuration

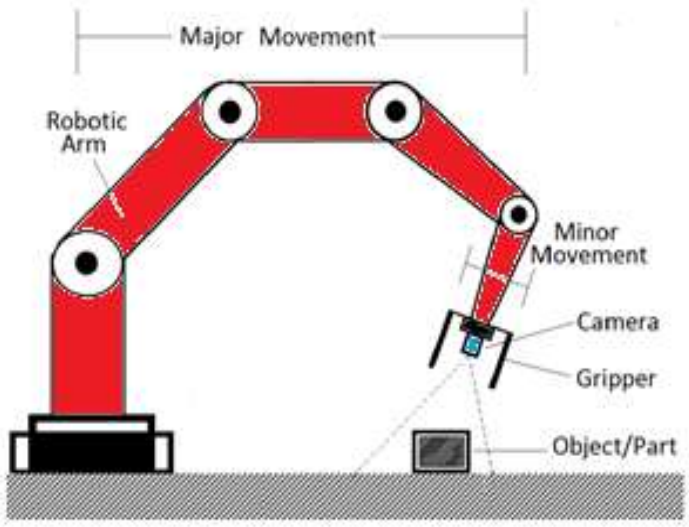


Figure 6.3: Eye-to-hand based robot configuration

The control laws are derived from the extracted visual features to provide a guided motion of robot. Based on the process of deriving the control law, the VS approach is broadly classified as image based visual servoing (IBVS) and position based visual servoing (PBVS) [Chaumette and Hutchinson, 2006; Hutchinson et al., 1996] . There is also a hybrid visual servoing technique [Corke and Hutchinson, 2000; Wang et al., 2010; Malis et al., 1999] in which control law is designed by considering both servoing controls.

6.2.1 Image Based Visual Servoing

In IBVS scheme, [Koenig et al., 2008; Bourquardez et al., 2009] the guided motion of a robot is achieved directly by the visual features. This technique was proposed by Sanderson and Weiss [Sandersom and Weiss, 1983]. In this scheme, the video stream is considered as input to the visual feature extraction methods which results visually tracked image features of the object. The visual features can be points (corner), lines (edges), statistical characteristic of the image (moments) and shapes. Then, the control law is derived based on the error between extracted and desired features on the image plane. The control law is synthesized in the image space. This depends on the point-to-point alignment, point-to-line alignment and shape alignment. The image based controller is designed by the image features towards a goal configuration that implicitly solves the original Cartesian motion planning problem [Hutchinson et al., 1996; Espiau et al., 1992]. The joint parameters are computed based on this image based control law. These joint parameters are

sufficient enough to guide the robot. This approach is inherently robust to camera calibration and target modelling errors [Espiau, 1994; Olivares-Mendez et al., 2015] which in turn reduce the computational cost.

The design of the control law involves the mapping between image space velocities and velocity in the workspace of the robot. This mapping is performed in the image space. The two major problems with IBVS are servoing to a local minima and attaining a Jacobian singularity or degeneracies [Chaumette and Hutchinson, 2007]. The robot manipulator in Cartesian space is controlled by inverse mapping from Cartesian space to joint space. This can sometimes become a problem due to the presence of singularities in the interaction matrix derived from the visual features. The arbitrary motion of the manipulator is lost in Cartesian direction which is called as "Losing a DoF". The singular position of the robot can be determined by analysing the Jacobian matrix of the manipulator. The matrix is singular, if the determinant is zero. The key factor in designing the control law is the extraction of good visual features. The extracted image points cannot be considered as good features due to the presence of singularity problem. Singularity is a function of the relative position of the target object and motion of the camera. Another factor in designing the control scheme is the relation between local minima and infeasible image features. Since control strategy is affected with respect to the image, there is no direct control over the Cartesian velocities of the hand. Thus, the trajectory of the robot workspace looks as if quite distorted in the Cartesian space, while producing image trajectories that are visually appealing. This results in achieving the higher accuracy for the motion with large rotations i.e., due to camera retreat [Corke and Hutchinson, 2001] problem. The process of IBVS control is non-linear and highly coupled, which is a major challenge in designing the controller. The important advantage of IBVS is that it reduces the computational delay, eliminates necessity for image interpretation and errors in sensor modeling and camera calibration [Hutchinson et al., 1996; Olivares-Mendez et al., 2015]. The block diagram for IBVS is given in Figure 6.4.

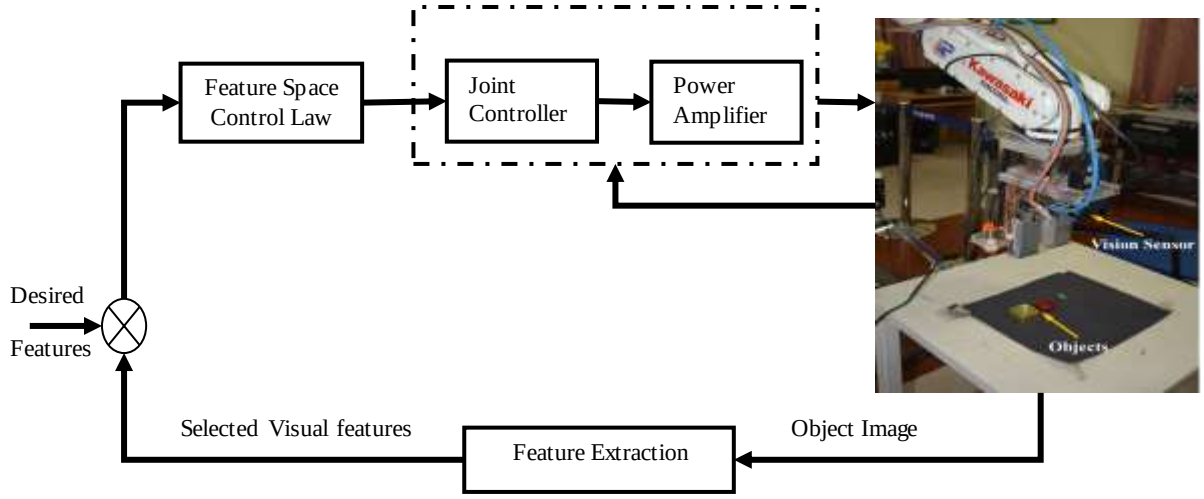


Figure 6.4: Block diagram for image based visual servoing

6.2.2 Position Based Visual Servoing

In PBVS [Chaumette and Hutchinson, 2006; Martinet et al., 1996; Wilson et al., 1996], features are extracted from the image and used in conjunction with a geometric model of the target to determine the pose of the target with respect to camera. In this control scheme, the video stream captured by the camera is fed to the features extraction method to produce visual features of the object. From these features, the relative position of the object of interest is estimated. The relative pose data is estimated from the 3D information of the object in Cartesian space. The 3D information depends on the 2D data of the object in the image and the intrinsic and extrinsic parameters of the camera. For this, a 3D camera calibration is required in order to map the 2D data of the image to the Cartesian space data. The efficacy of the PBVS control scheme requires accurate calibration of camera. Intrinsic parameters depend exclusively on the optical characteristics, namely, lens and CMOS sensor properties. The calibration of intrinsic parameters can be operated offline in the case that optical setup is fixed during the operative tasks of the robot. Extrinsic parameters indicate the relative pose of the camera reference system with respect to a generic world reference system. It is assumed that the world reference system is exactly the object frame, so that the extrinsic parameters give directly the pose of the camera with respect to the target. Obviously the extrinsic parameters are variable with robot or target motion, and an online estimation is needed to perform a dynamic look-and-move tracking task. Then, the control law is designed based on the error computed with respect to the desired relative position and the estimated

relative position of the object in the image. The joint parameters are calculated in accordance with the control law designed. The block diagram of PBVS is given in Figure 6.5.

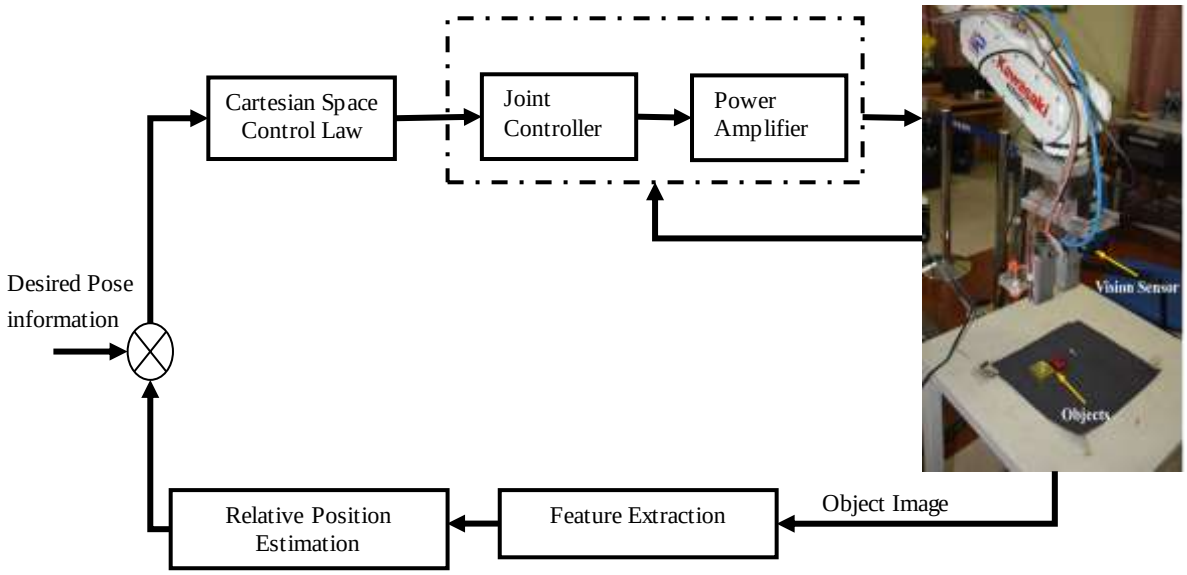


Figure 6.5: Block diagram for position based visual servoing

The main advantage of PBVS control is that it provides a stability as the joint parameters of the robot is well known in advance. The system with such control requires accurate camera calibration. This is one of the major drawback of PBVS. Errors in calibration of the vision system will lead to errors in the 3-D reconstruction and subsequently to errors during robot motion. In addition, since the control law for PBVS is defined in terms of the 3-D workspace, there is no mechanism by which the image is directly regulated. Thus, it is possible that objects of interest (including features that are being used by the visual servo system) can exit the camera's field of view.

6.2.3 Hybrid Visual Servoing

Hybrid visual servoing (HVS) [Chaumette and Hutchinson, 2006; Corke and Hutchinson, 2000; Wang et al., 2010; Malis et al., 1999; Feddema and Mitchell, 1989] control scheme uses the combination of both 2D and 3D servoing methods. This control scheme makes the use of image space (2D) information and 3D object parameters together in the control design of a dynamic look-and-move system. A linear model of camera–object interaction (affine object shape transformations) is assumed, thus reducing the size of visual representation according to active

vision requirements. This helps in maintaining the global asymptotic stability of the system. The use of IBVS helps in maintaining the field of view (FOV) constraint, while the use of PBVS guarantee the global stability of the system. This scheme is using IBVS for translational and PBVS for rotational motion control. The advantage of IBVS scheme i.e., all control parameters can be estimated in the image plane without requiring camera calibration, is retained in this scheme. This estimation is directly used in designing the control law. The system become robust as the noise effect is low. The nonlinear control strategy of PBVS scheme is used for synthesizing the control law to ensure the stability of the system. The performance of this approach is highly depended on the convolution method of combining the linear and nonlinear behavior of IBVS and PBVS scheme respectively. The major drawback of this control scheme is its computational complexity. It is higher than the other two IBVS and PBVS schemes.

6.2.4 Comparison of Visual Servoing Control Schemes

In IBVS, the control law is derived in the image space where as in PBVS, it is in Cartesian space. In image based control, the error signal is measured in the image space and mapped directly into the actuator commands [Sanderson and Weiss, 1983; Feddema and Mitchell, 1989]. This is in contrast to position-based control system systems in which extracted features are used to compute a (partial) 3-D reconstruction of the target object in the environment [Westmore et al., 1991].]. An error is computed in task space or Cartesian space. This error is used for designing the control law in both VS control schemes. The important limitation of PBVS control is the error due to the calibration of the vision system which will lead to errors in the 3-D reconstruction and subsequently to errors during task execution. However, the camera calibration is not so important in IBVS. In addition to this, as the control law for PBVS is defined in terms of the 3-D workspace, there is no mechanism by which the image is directly regulated. Thus, it is possible that objects of interest (including features that are being used by the visual servo system) can go out FOV of camera. However, there are certain problems associated with IBVS systems. The 3D reconstruction is not at all required in IBVS as it works in 2D image space. The major limitation of image based control lies in the design of control law as it requires the accurate coordination of image space velocities to velocities in the robot's workspace. This mapping depends on the singularities and conditioning number. Due to this, the trajectory planning in the robot workspace is a challenging task in Cartesian space. This performance problem with IBVS and PBVS control leads to the introduction of several hybrid control strategies. HVS methods use IBVS to control

certain degrees of freedom while using other techniques to control the remaining degrees of freedom.

Still the IBVS control approaches have seen more advantageous than other controlling strategies as it is less prone to noise, provides singularities and less computational complex [Tahri et al., 2015]. However, by the IBVS control scheme can outperforms, if the singularity problem is taken care of appropriately. This is the reason for considering IBVS approach in this work. The factors that distinguishes these servoing approaches are summerised in Table 6.1.

Table 6.1: Comparison of IBVS and PBVS schemes

IBVS	PBVS
Control law is derived in the image space	Control law is derived in the Cartesian Space
No object model is required	CAD model of the object must be known
Singularity	Not highly affected
Maintain field of view (FOV) constraint	The FOV can make an issue
Camera calibration error is not important	Camera calibration leads to errors in the 3-D reconstruction
Robust to noise	Sensitive to noise

6.3 Basic Formulation of Visual Servoing

In order to do the basic formulation of VS, $C(x, y)$ is considered be a captured image via a vision sensor with N DoF at time t. The main aim VS is to find out some control inputs $In(t) \in \mathbb{R}^N$ to the system so as to perform some desired work. An eye-in-hand configuration of a popularly used serial robot with 6-DOF is used to shift the end-effector from one position to another position in the workspace. In order to perform this task, the VS aims at achieving the relationship between the movement of robot and the variations in the characterizing properties of the image (coordinates of image, visual features).

The main aim of the VS is to design control law which will diminish the discrepancy between the desired and observed feature vector so as to achieve accurate insight of the task. Eq. 6.1 provides the control principle of the VS to minimize the error function $\left(error = D^* - C\right)$ to zero [Chaumette and Hutchinson, 2006; Chaumette and Hutchinson, 2007].

$$J = \arg\min \sum_{i=1}^{N} \left(D(i)^* - C(i) \right) \tag{6.1}$$

where, D^* is feature vector that describes $D(x, y)$, the desired image.

C is a feature vector that charecterises $C(x, y)$.

N is the number of image measurements

C is a function, i.e., $C(M, A)$, M is the image measurement and A is the potential parameters of camera

6.3.1 Design of Control Law

The current discussion describes the formulation of controlling the motion of a camera with 6-DOF i.e., a camera attached to the end effector of a 6-DoF arm. The straightforward approach to design the control law is to design a velocity controller so as to ensure that the feature vector C is approaching towards the desired vector D^*. Once the vector C is defined, it becomes easy to design the control laws. Considering a case in which goal position is fixed with a stationary target, i.e., D^* is constant. The variation in C depends only on the motion of camera. The velocity controller (VC) provides a basic correlation between the time variations of the visual features of vector C i.e., $\dot{C}(t)$ to the camera velocity C_V and is described in Eq. 6.2 [Chaumette and Hutchinson, 2006; Espiau et al., 1992]. The camera velocity C_V is a function of instantaneous as well as rotational velocities of the camera frame and is described in Eq. 6.3.

$$\dot{C}(t) = I_C C_V \tag{6.2}$$

$$C_V = (V, W) \tag{6.3}$$

Where the linear velocity of the camera frame is V and W is the angular velocity. $I_C \in \mathbb{R}^{N\times 6}$ is the interaction matrix related to C for a robot of 6-DoF.

The relationship between the camera velocity C_V and the time variation of error is presented in Eq. 6.4.

$$\dot{error}(t) = I_e C_V \tag{6.4}$$

Where $\dot{error}$ defines the time variation of the vector $error$. $\{I_e, I_C\} \in \mathbb{R}^{N\times 6}$ is the interaction matrix related to the error vector and feature vectors. Assuming the desired vector D^* to be constant, the time variation of error, $\dot{error}(t)$ can be defined as

$$\dot{error}(t) = \dot{C}(t) = I_C C_V \tag{6.5}$$

It is assume that the error can be decreased by decoupling exponentially and is described as

$$\dot{error}(t) = -\alpha\ error \tag{6.6}$$

where α is a proportional gain.

From Eq. 6.5 and Eq. 6.6, the control law can be defined for an eye-in-hand system observing a static object and is given in Eq. 6.7.

$$C_V = -\alpha I_C^{+} error \tag{6.7}$$

where, $C_V(V_X, V_Y, V_Z, W_X, W_Y, W_Z)$ is the velocity sent to the low-level robot controller and I_C^{+} is the Moore-Penrose pseudoinverse of I_C [Chaumette and Hutchinson, 2007]. In VS, it is always a challenging task to get the exact value of I_C. Hence, an approximate value of I_C^{+} is computed and denoted as I_C^{+}, and the control is defined as

$$C_V = -\alpha\ I_C^{+} error \tag{6.8}$$

This is the basic design constraint of all visual servoing techniques. The different categories of VS vary in the selection of visual features C, form of interaction matrix I_C and its approximate value. The basic steps followed for the design of a control law of an IBVS based control is given in the Figure 6.6. In which the visual features C considered are either the global feature or the local features of the image.

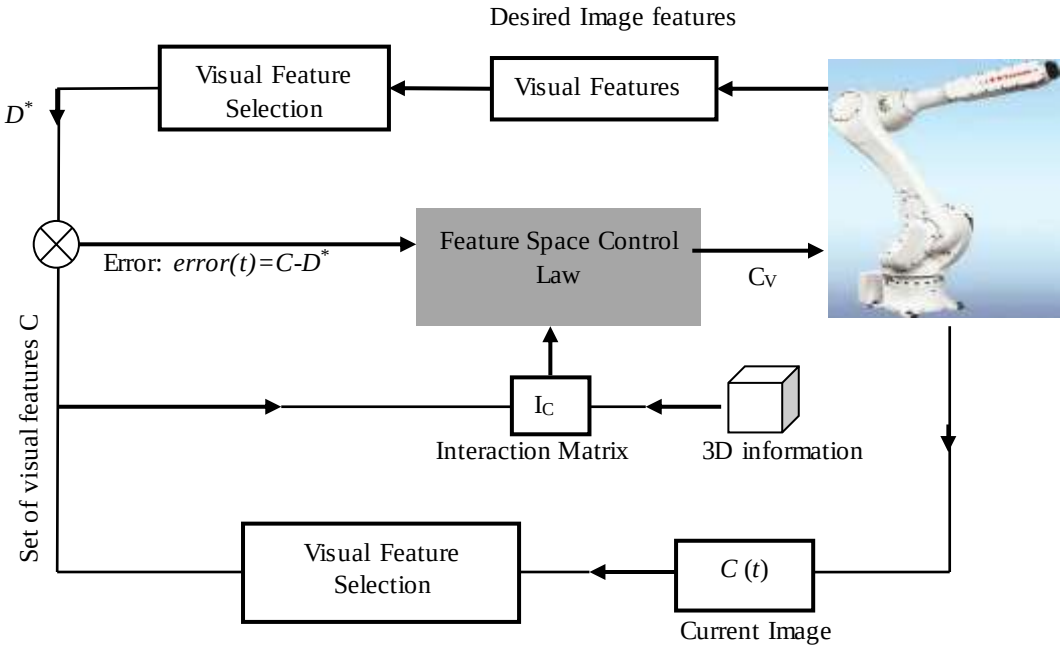


Figure 6.6: Control law for image based visual servoing

6.3.2 Selection of Features for Image Based Visual Servoing

The main issue in visual servoing scheme is extraction of useful visual features to design a suitable control law for robot motion. In chapter 4, global and local features and their impact on detection is discussed. The global features play an important role in recognizing the desired object. Global features are shape descriptor, area or size of object and image moments. Out of these, image moments are considered most suitable for visual servoing as they are invariant to scale, rotation and translation. A detailed description about image moments and their invariant behavior is given in the following section. In IBVS scheme, selection of visual features is a challenging task. Mainly there are two type of features used in visual servoing application such as local and global features [Hutchinson et al., 1996; Espiau et al., 1992]. Local feature extraction is based on simply geometry

features such as point, line, edge, corner, area or artificial marks whereas, global feature extraction is based on Eigen space method, geometric moment, stochastic transform, Fourier descriptor, optical flow, light intensity, luminance signal. The most common image Jacobian is based on the motion of points in the image [Feddema and Mitchell, 1989; Hashimoto et al., 1991].

On the other hand the global features are used as shape descriptor. They are generally classified as contour based and region based descriptor. Region based descriptor capture contour information as well as the interior content of the shape; whereas contour based descriptor contain only contour information. Global features extraction based on optical flow, luminance [Collewet and Marchand, 2011], are developed to avoid the tracking and matching in IBVS but they have limited convergence domain due to nonlinear nature. Several types of moments can be used: geometric moments [Hu, 1962], Zernike moments [Khotanzad and Hong, 1990], and Legendre moments [Teague, 1980]. Moment-based visual servoing has led to interesting results [Tahri et al., 2015; Chaumette, 2004; Mebarki et al., 2010; Tahri and Chaumette, 2005; Zhang and Liu, 2013]. Moments are the most popular descriptor among all descriptors. Moment can be easily computed from a binary or a segmented image. Moments have been initially applied for visual pattern recognition in computer vision application.

6.4 Mathematical Moment

6.4.1 Introduction to Moment

The factor affecting the robot control based on VS is the correctness of the extracted visual features. The features extraction method is fully dependent on the imaging condition and the system. The real imaging system as well as the imaging conditions is usually not perfect. So, various types of the degradations are introduced into the image during acquisition stage. Consequently, the images captured by the existing imaging systems are found to be the degraded version of the original image. The key factors that influence the degradation in the image are as follows [Flusser, 2006].

- Imaging geometry
- Lens aberration
- Wrong focus
- Motion of the scene
- Sensor error

A basic relationship between the original image $f(x, y)$ and the captured image $C(x, y)$ is given in Eq. 6.9.

$$C(x, y) = D[f(x, y)] \tag{6.9}$$

where, D is the degradation operator. Generally, the degradation operator is decomposed into two types, such as radiometric (or gray level or color) and geometric (or spatial) degradation operator. Once the video or image frames are captured by the vision sensor, it is to be analyzed for yielding measurable information of the object in the scene. Scene analysis consists of three steps. They are as follow:

- Image is segmented and objects of potential interest are identified.
- Object of interest are mathematically modeled and classified as elements of certain class from the set of predefined object class.
- Spatial relations among the objects are analyzed.

Recognition of objects of interest and patterns that are degraded in various ways have been a goal of much recent research. Basically, there are three major approaches [Flusser et al., 2009] to this problem. They are as follows:

- Brute-force
- Image normalization
- Invariant feature

In brute-force approach [Nelson and Selinger, 1997], the parametric space of all possible degradations is searched. This means, the training set of each class should contain all class representatives and their rotated, scaled blurred and deformed versions. This approach leads to extreme time complexity and is practically inapplicable. On the contrary, image normalization [Pei et al., 1995], transforms the image of the object into a certain standard form before giving it to the classifier. This is very efficient in the classification stage, but normalization requires the solving of difficult inverse problems. Generally, these problems are ill-conditioned or ill-poised. In case of blurring, normalization requires blind deconvolution. In case of spatial image, it requires registration of the image with respect to some reference frame. Recognition of objects via the implementation of invariant feature [Hu, 1962; Flusser, 2006; Mercimek et al., 2005] is more

promising and used extensively. The idea behind this approach is to describe the objects by a set of measurable quantities called invariants that are insensitive to particular degradation. This provides much discrimination power to distinguish among objects belong to other classes. This is the reason of considering this approach in the proposed work.

The invariant features I defined on the space of all admissible image functions which do not change their value under degradation D. This must satisfy the following condition for any image function $f(x,y)$, that is,

$$I[f(x,y)] = I[D(f(x,y))] \tag{6.10}$$

This is called invariance property.

The intra-class variability and noise due to the improper segmentation is considered as a weaker constraint for formulation of invariant features. So, $I[f(x,y)]$ is not different significantly from $I[D(f(x,y))]$. The desirable property of I is the discriminability power. So, it is worth mentioning that "the boarder the invariance, the less is the discriminating power". This is important to choose a proper trade-off between invariance and discrimination power for object recognition based on features. Several invariant features are generally considered for a better discrimination power and it is considered as invariance vector $(I_1, I_2, I_3, \cdots, I_n)$.

The existing invariant features can be classified into five according to the mathematical tools used with respect to the portions or characteristics of the object used for characterizing 2D objects. They are such as:

- Simple shape descriptor (Compactness, convex and elongation etc.) [Sonka et al., 2014]
- Transform coefficient features (Fourier descriptor, Hadamard descriptor etc.) [Lin and Chellappa, 1987]
- Point set invariants (Use position of points) [Mundy and Zisserman, 1992]
- Differential invariants (Use derivatives of the object boundary) [Mokhtarian and Abbasi, 2002]
- Moment invariants (Special function of image moments) [Flusser, 2006; Mercimek et al., 2005]

But the invariants are broadly classified into three [Nelson and Selinger, 1997] such as

- Local invariants
- Global invariants
- Semi-local invariants

Global invariants are calculated from the entire image. These invariants are calculated by the integration of the data acquired by projections of the image onto certain basis functions. Such invariants are more robust to noise and provide inaccurate boundary detection. The major limitation of such invariants is that, a small change to the local feature influences values of all invariants. This is the reason that global invariants are not used for partial occluded object or object out of FOV.

Local invariants are calculated from the points in the image which are dominant in the neighborhood. These invariants are computed for each detected boundary point as a function of the boundary derivatives. The invariants at any point depend only on the shape of the boundary in its immediate locality. So, the local invariants are not affected by any changes. For recognition of partial occluded objects, these invariants are suitable. But it is difficult to use practically as it is prone to discretization errors, segmentation inaccuracy and noise.

Semi-local invariants attempt to carry the positive properties of both local and global invariants. Here, the object is divided into stable parts with respect to vertices of the boundary and each part is described by global invariants. So, the entire object is characterized by a string of invariant vectors.

6.4.2 Mathematical Formulation

Generally, in image processing, the representation of object by pixels is turned on while the background is turned off. This can be achieved by backlighting or rendering the image. The object present in the image can be detected efficiently by the use of its global features. Image moment is one of the global features. The general form of image moment [The et al., 1988] is defined as

$$M_{a,b} = \sum_{x,y} x^a y^b \tag{6.11}$$

where, $M_{a,b}$ is the moment of the object present within the image and a, b are the scalar values. These scalar values help in rendering process. The object pixels x, y in the image are raised to powers of a, b which decides the pixels to be turned on. That means, this determines whether a pixel belongs to the object or not.

6.4.3 Moment in Image Analysis

Moments are the scalar quantities used to characterize an image and extract the important features. In statistics, moments are used to describe the shape of a probability density function. It is also used to measure the mass distribution of a body in rigid-body dynamics. Mathematically, it is defined as the projection of a function onto a polynomial basis.

Definition: The moment $m_{nl}^{(f)}$ of an image $f(x,y) \geq 0$, a real bound function is defined as

$$m_{nl}^{(f)} = \iint_{\mathbb{R}} P_{nl}(x,y)\, f(x,y)\, dxdy \tag{6.12}$$

where, n and l are non-negative integers and $n+l$ is the order of moment. $p_{00}(x,y), p_{10}(x,y), p_{10}(x,y)$.... are the polynomial basis functions defined on $\mathbb{R}$. The superscript $^{(f)}$ is omitted for simplicity. $f(x,y)$ is the intensity level of the image at coordinate (x,y).

The geometric moment is defined as

$$m_{nl} = \iint_{\mathbb{R}} x^n y^l f(x,y)\, dx\, dy \tag{6.13}$$

Considering the centroid of image of the object at (x_G, y_G), the central moment can be computed as

$$\mu_{nl} = \iint_{\mathbb{R}} (x - x_G)^n (y - y_G)^l\, f(x,y)\, dxdy \tag{6.14}$$

where, (x_G, y_G) defines the centroid of the object and can be defined in term of lower order moment m_{nl}. The representation of (x_G, y_G) in terms of m_{nl} is defined in eq. (6.15).

$$x_G = \frac{m_{10}}{m_{00}} \quad \text{and} \quad y_G = \frac{m_{01}}{m_{00}} \tag{6.15}$$

The geometric moment of lower order provides an intuitive meaning as follows:

Zeroth order moment,

m_{00}: the mass of the object in an image (area in case of binary image)

First order moment,

m_{10}/m_{00} and m_{01}/m_{00} : Centre of gravity or centroid of the object

Second order moment,

m_{20} and m_{02}: Mass distribution of the image with respect to coordinate axes.

The second order moment defines the orientation of the object in an image. It is also used to determine the normalized position of an image. In statistics, the higher order moments are used for skewness and kurtosis. The skewness is the deviation of the respective projection from symmetry and kurtosis is the "peakedness" of the probability density function. If the projection is symmetric with respect to origin then skewness is zero.

Skewness for horizontal projection is $m_{30}/\sqrt{m_{20}^3}$ and $m_{03}/\sqrt{m_{02}^3}$ for vertical projection.

Kurtosis for horizontal projection is m_{40}/m_{20}^2 and m_{40}/m_{02}^2 for vertical projection.

Moment invariants are those reliable moments which are insensitive to various types of geometric transformation. Several combination of moments are tried in by various researchers [Chaumette, 2004; Tahri et al., 2015, Chaumette, 2002; Tahri and Chaumette, 2003; Hoffmann et al., 2006; Mukundan and Ramakrishnan, 1998; Chong et al., 2004], those are invariant to scaling, rotation, translation, affine, projective and elastic.

Some combination of central moments of different order is presented which are mutually invariant to rotation as well as translation. The formulation some of these type of invariants are provided in Eq. 6.16 through Eq. 6.18.

$$t_0 = \mu_{02} + \mu_{20} \tag{6.16}$$

$$t_1 = \mu_{02}\mu_{20} - \mu_{11}\mu_{11} \tag{6.17}$$

$$t_2 = -\mu_{30}\mu_{12} + \mu_{21}\mu_{21} - \mu_{03}\mu_{21} - \mu_{12}\mu_{12} \tag{6.18}$$

$$t_3 = 3\mu_{30}\mu_{12} + \mu_{30}\mu_{30} + 3\mu_{03}\mu_{21} + \mu_{03}\mu_{03} \tag{6.19}$$

Another formulation for the translation & scaling invariants are proposed in [Mamistvalov, 1998] and is defined in Eq. 6.20.

$$t_s = \frac{m_{nl}}{m_{00}^{n+l+2/2}} \tag{6.20}$$

Eq. 6.20 provides the more popular type of transformation invariants that are invariant to scaling, rotation & translation (SRT) by utilizing the formulation given in Eq. 6.16 - 6.19. The transformation invariants are defined as

$$\left.\begin{aligned} t_{n1} = \frac{t_2}{t_1^{8/10}}, t_{n2} = \frac{t_3}{t_1^{8/10}}, t_{n3} = \frac{t_3}{t_2} \\ t_{n4} = \frac{t_3}{t_{00}^5},\ t_{n5} = \frac{t_2}{t_{00}^5},\ \ t_{n6} = \frac{t_1}{t_{00}^4} \end{aligned}\right\} \tag{6.21}$$

However, the main difficulty in utilizing the image moments in various VS application is how to describe the analytical form of the interaction matrix which is very essential in designing a visual servo control scheme.

6.4.4 Importance of Moment Invariants

Image moments are useful in describing a segmented image. Simple properties of image like area, centroid and orientation can be easily calculated by moment. The invariant properties of moments like translation, rotation, scaling and transformation etc. make the object detection efficient. So, it is widely used for extraction of visual features which are the base of the success of VS. The necessary condition of the VS is that the interaction matrix must not be singular. This can be achieved by an optimal selection visual feature which determines the matrix with rank 6 for designing a control law for a robot of 6-DoF. To ensure this condition, the best approach is to design a decoupled control scheme in which each DoF of camera is associated with one visual feature. Such control would make easy the determination of the potential singularities of the considered task, as well as the choice of interaction matrix.

Two main issues are observed in moments based visual servoing, first is for symmetrical object, invariants computed from odd orders central moments are zero, and invariants to translation , rotation and scale become useless because their denominator are zero. Second is sensitivity to noise increases with the increase of moment orders. However, shifted centred moments avoid the above mentioned issues. Shifted moment is used to select features to control efficiently the rotational DoF. However the selection of optimal visual features is still a key issue to solve the problem singularity in IBVS.

6.5 Visual Servoing Based Non-orthogonal Moment Invariants

It is quite obvious that the time variation of feature vector $\dot{C}$ (as in Eq. 6.2) of the captured image is related to the relative kinematic screw C_v between object and camera. By considering the feature vector based on image moment m_{nl}, Eq. 6.22 provides the link between time variations $\dot{m}_{nl}$ of moment m_{nl} as a function of C_v [Tahri and Chaumette, 2003; Tahri et al., 2015].

$$\dot{m}_{nl} = I_{m_{nl}} C_v \tag{6.22}$$

$I_{m_{nl}}$ is called the interaction matrix related to m_{nl}.
Based on this formulation, the control law is described as in Eq. 6.23.

$$C_V = -\alpha I_{m_{nl}}^{+} \left(C - D^* \right) \tag{6.23}$$

In order to analyze the above equation we need to find out the analytic description of the interaction matrix $L_{m_{nl}}$.

6.5.1 Formulation of Interaction matrix of 2D moment ($I_{m_{nl}}$)

Considering m_{nl} is 2D moment, $\dot{m}_{nl}$ as its time variation and refereeing to the work of Chaumette [Chaumette, 2004] and Tahri [Tahri and Chaumette, 2005], the $\dot{m}_{nl}$ can be derived by

$$\dot{m}_{nl} = \iint_{\mathbb{R}} \left[\frac{\partial f}{\partial x} \dot{x} + \frac{\partial f}{\partial y} \dot{y} + f(x,y) \left(\frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y} \right) \right] dx\, dy \tag{6.24}$$

where, $f(x,y) = x^n y^l$, $\dot{x}$, $\dot{y}$, $\dfrac{\partial \dot{x}}{\partial x}$ and $\dfrac{\partial \dot{y}}{\partial y}$ are linearly related to C_v.

For any point $X(x,y)$ in the image plane whose corresponding 3D point has depth $\mathbb{Z}$, we can have a relation between time variations of $X(x,y)$ to C_v as defined in Eq. 6.25.

$$\dot{X} = I_X C_V \tag{6.25}$$

where, I_X is the interaction matrix related to $X(x,y)$ and is defined as given below

$$I_X = \begin{pmatrix} -1/\mathbb{Z} & 0 & x/\mathbb{Z} & xy & (-1-x^2) & y \\ 0 & -1/\mathbb{Z} & y/\mathbb{Z} & (1+y^2) & -xy & -x \end{pmatrix} \tag{6.26}$$

From eq. (6.26), it is observed that the dynamics $\dot{X}$ with respect to C_v is not consistent. Some are linearly related to image points while some are primarily related to $\mathbb{Z}$. In many cases also some are dependent on the 2nd order presentation. The control scheme designed based on the nonlinear behavior of I_X I_x will cause a very poor robot trajectory. Hence, such type of visual features are chosen to avoid the nonlinearities that affect the formation of interaction matrix.

Assuming the observed 3D object belongs to a continuous surface, i.e., considering absence of any depth discontinuity, the depth $\mathbb{Z}$ of any 3D object point can be expressed as a continuous function of its image coordinates x and y as given in eq. (6.27)

$$1/\mathbb{Z} = \sum_{n\geq 0, l\geq 0} m_{nl} x^n y^l \tag{6.27}$$

If the object is planar, its depth can be expressed in the camera frame as

$$\mathbb{Z} = \gamma_1 X + \gamma_2 Y + \mathbb{Z}_0 \tag{6.28}$$

Now, Eq. 6.27 can be expressed as

$$1/\mathbb{Z} = Ox + Py + Q \tag{6.29}$$

where, O, P, and Q are plane parameters and are defined as $O = -\gamma_1/\mathbb{Z}_0$, $P = -\gamma_2/\mathbb{Z}_0$, $Q = 1/\mathbb{Z}_0$

Putting Eq. 6.29 in Eq. 6.27 and Eq. 6.28 it is induced that

$$\begin{cases} \dot{x} = -(Ox + Py + Q)V_X + x(Ox + Py + Q)V_Z + xyW_X - (1 + x^2)W_Y + yW_Z \\ \dot{y} = -(Ox + Py + Q)V_Y + y(Ox + Py + Q)V_Z + (1 + y^2)W_X - xyW_Y - xW_Z \end{cases} \tag{6.30}$$

From this it can be found that

$$\begin{cases} \dfrac{\partial \dot{x}}{\partial x} = -OV_X + (2Ox + Py + Q)V_Z + yW_X - 2xW_Y \\ \dfrac{\partial \dot{y}}{\partial y} = -PV_Y + (Ox + 2Py + Q)V_Z + 2yW_X - xW_Y \end{cases} \tag{6.31}$$

Now, substituting Eq. 6.30 and Eq. 6.31 in Eq. 6.24 and assuming that $f(x, y) = x^i y^j$, $\dfrac{\partial f}{\partial x} = jx^i y^{j-1}$ and $\dfrac{\partial f}{\partial y} = ix^{i-1} y^j$, Eq. 6.24 can be redefined in the form as given in Eq. 6.22 and the analytical form of $I_{m_{nl}}$ is

$$I_{m_{nl}} = \begin{bmatrix} m_{V_X} & m_{V_Y} & m_{V_Z} & m_{W_X} & m_{W_Y} & m_{W_Z} \end{bmatrix} \tag{6.32}$$

with

$$\begin{cases} m_{V_X} = -n(O\,m_{n,l} + P\,m_{n-1,l+1} + Q\,m_{n-1,l}) - O\,m_{n,l} \\ m_{V_Y} = -l(O\,m_{n+1,l-1} + P\,m_{n,l} + Q\,m_{n,l-1}) - P\,m_{n,l} \\ m_{V_Z} = (n + l + 3)(O\,m_{n+1,l} + P\,m_{n,l+1} + Q\,m_{n,l}) - Q\,m_{n,l} \\ m_{W_X} = (n + l + 3)\,m_{n,l+1} + l\,m_{n,l-1} \\ m_{W_Y} = -(n + l + 3)\,m_{n+1,l} - n\,m_{n-1,l} \\ m_{W_Z} = n\,m_{n-1,l+1} - l\,m_{n+1,l-1} \end{cases} \tag{6.33}$$

In terms of central moment μ_{nl}, the interaction matrix is denoted as $I_{\mu_{nl}}$ and is defined as [Tahri and Chaumette, 2005]

$$I_{\mu_{nl}} = \begin{bmatrix} \mu_{V_X} & \mu_{V_Y} & \mu_{V_Z} & \mu_{W_X} & \mu_{W_Y} & \mu_{V_Z} \end{bmatrix} \tag{6.34}$$

where, $\mu_{V_X} = -(n+1)O\mu_{n,l} - nP\mu_{n-1,l+1}$

$\mu_{V_Y} = -l\,O\,\mu_{n+1,l-1} - (l+1)P\mu_{n,l}$

$\mu_{V_Z} = -O\mu_{W_Y} + P\,\mu_{W_X} + (n+l+2)Q\mu_{n,l}$

$\mu_{W_X} = (n+l+3)\mu_{n,l+1} + nx_G\mu_{n-1,l+1} + (n+2l+3)y_G\mu_{n,l} - n\eta_{11}\mu_{n-1,l} - l\eta_{02}\mu_{n,l-1}$

$\mu_{W_Y} = -(n+l+3)\mu_{n+1,l} + (2n+l+3)x_G\mu_{n,l} - l\,y_G\mu_{n+1,l-1} + n\eta_{20}\mu_{n-1,l} + l\eta_{11}\mu_{n,l-1}$

$\mu_{V_Z} = n\mu_{n-1,l+1} - l\mu_{n+1,l-1}$

and $\eta_{ij} = {}^{4\mu_{ij}}\!/_{m_{00}}$

A. Selection of Feature Vectors

Based on Eq. 6.2, to control the 6-DoF of a camera, a vector of 6 visual features is required. In literature, varieties of combination of moment invariants are considered as visual features to control the DoF and to provide a sparse interaction matrix.

Usually, the first three components such as m_{V_X}, m_{V_Y} and m_{V_Z} of the interaction matrix $I_{m_{nl}}$ are utilized to control the translational DoF of a camera whereas, the rest three components such as m_{W_X}, m_{W_Y} and m_{W_Z} are used for controlling the rotational DoF. The interaction matrix $I_{m_{nl}}$ is replaced by $I^{''}_{m_{nl}}$ when the image plane is parallel to the desired position of the object plane.

a. Feature Vector to Control the Translational DoF

In literature [Tahri and Chaumette, 2015; Chaumette, 2004], various moment based visual features are used for controlling the translational DoF, But, the popular choice is the center of gravity x_G, y_G and the area of the object a to control the translational DOF.

By considering these components, the corresponding entries in the interaction matrix are defined as

$$\left.\begin{aligned} I^{''}_{x_G} &= \begin{pmatrix} -Q & 0 & Qx_G & \zeta_1 & -(1+\zeta_2) & y_G \end{pmatrix} \\ I^{''}_{y_G} &= \begin{pmatrix} 0 & -Q & Qy_G & 1+\zeta_3 & -\zeta_1 & -x_G \end{pmatrix} \\ I^{''}_{a} &= \begin{pmatrix} 0 & 0 & 2aQ & 3ay_G & -3ax_G & 0 \end{pmatrix} \end{aligned}\right\} \tag{6.35}$$

where, $\zeta_1 = \eta_{11} + x_G y_G$,

$\zeta_2 = \eta_{20} + x_G^2$ and

$\zeta_3 = \eta_{02} + y_G^2$

The above mentioned matrix is triangular and nonlinear in nature. Each visual feature does not contribute the similar dynamics to each translational DoF. Hence, in order to obtain the same dynamics for all the translational DoF, normalization process is applied. After normalization, the interaction matrix is replaced by

$$\left.\begin{aligned} I_{x_n}^{"} &= \begin{pmatrix} -1 & 0 & 0 & a_n\varepsilon_{11} & -a_n(1+\varepsilon_{12}) & y_n \end{pmatrix} \\ I_{y_n}^{"} &= \begin{pmatrix} 0 & -1 & 0 & a_n(1+\varepsilon_{21}) & -a_n\varepsilon_{11} & -x_n \end{pmatrix} \\ I_{a_n}^{"} &= \begin{pmatrix} 0 & 0 & -1 & -3y_n/2 & 3x_n/2 & 0 \end{pmatrix} \end{aligned}\right\} \quad (6.36)$$

where,

$x_n = x_G a_n, y_n = y_G a_n$,

$a_n = \mathbb{Z}^* \sqrt{a^*/a}$,

* represents the desired position

$\mathbb{Z}$ represents the depth between camera and object.

a describes the area of the object and

$\varepsilon_{11} = \varepsilon_{22} = \eta_{11} - x_G y_G/2$

$\varepsilon_{12} = \eta_{20} - \frac{x_G^2}{2}$

$\varepsilon_{21} = \eta_{02} - \frac{y_G^2}{2}$

From the normalized interaction matrix, it is concluded that variation of such features

- directly relates to the depth of image plane
- provides diagonal blocks which will allow to obtain adequate translational property of robot.
- provides identical dynamics for translational DoF.

b. Feature Vector to Control the Rotational DoF

In several research works, the object orientation θ is considered to control the rotation around the optical orientation W_Z whereas the other two remaining DoF, W_X and W_Y are controlled by two visual features $\left(R_i, R_j\right)$ invariant to SRT.

The object orientation θ is defined by

$$\theta = \frac{1}{2}\arctan\left[\frac{2\mu_{11}}{\mu_{20}-\mu_{02}}\right] \tag{6.37}$$

On the other hand, the feature vectors R_i and R_j are chosen from several combination of moments that are invariant to SRT. In literatures, variety of approaches [Tahri et al., 2015; Tahri and Chaumette, 2005; Tahri and Chaumette, 2003; Mamistvalov, 1998] have been proposed to find out the combination of such type of moment invariants.

By using the feature vectors R_i, R_j and θ to control W_X, W_Y and W_Z, the interaction matrix is described as

$$\left.\begin{aligned} I^{"}_{Ri} &= \begin{pmatrix} 0 & 0 & 0 & RiW_X & RiW_Y & 0 \end{pmatrix} \\ I^{"}_{Rj} &= \begin{pmatrix} 0 & 0 & 0 & RjW_X & RjW_Y & 0 \end{pmatrix} \\ I^{"}_{R\theta} &= \begin{pmatrix} 0 & 0 & 0 & \theta W_X & \theta W_Y & -1 \end{pmatrix} \end{aligned}\right\} \tag{6.38}$$

Let $I^{"}_{R_{ij}}$, be the matrix relating the variation of moment invariants R_i, R_j to the rotation velocity W_X, W_Y and is described in Eq. 6.39.

$$I^{"}_{R_{ij}} = \begin{bmatrix} R_iW_X & R_iW_Y \\ R_jW_X & R_iW_Y \end{bmatrix} \tag{6.39}$$

For the appropriate decoupling of W_X and W_Y, R_i and R_j should be selected in such a way that $I^{"}_{R_i}$ and $I^{"}_{R_j}$ will be as orthogonal as possible and the matrix $I^{"}_{R_{ij}}$ must be non-singular.

If the matrix $I''_{R_{ij}}$ will be non-singular, then only the combination of feature vector $\left[x_n\ y_n\ a_n\ R_i\ \mathrm{R}_j\ \theta\right]$ will produce a sparse interaction matrix I''_C and is defined in Eq. 6.40.

$$I''_C = \begin{bmatrix} -1 & 0 & 0 & a_n\varepsilon_{11} & -a_n(1+\varepsilon_{12}) & y_n \\ 0 & -1 & 0 & a_n(1+\varepsilon_{21}) & -a_n\varepsilon_{11} & -x_n \\ 0 & 0 & -1 & -3y_n/2 & 3x_n/2 & 0 \\ 0 & 0 & 0 & R_iW_X & R_iW_Y & 0 \\ 0 & 0 & 0 & R_jW_X & R_jW & 0 \\ 0 & 0 & 0 & \theta W_X & \theta W_Y & -1 \end{bmatrix} \tag{6.40}$$

From the above analysis, it is observed that the key factor in generating an efficient interaction matrix is dependent on the selection of unique combination of R_i and R_j which will provide a better sparse interaction matrix with lower condition number. However, the methods described in [Tahri et al., 2015; Tahri and Chaumette, 2005; Tahri and Chaumette, 2003; Mamistvalov, 1998] select these features randomly or by hit and trial method which is really difficult to achieve and unstable in nature. Also, the choice of the combination depends on the shape of the object and is restricted to lower order moments.

These issues motivate us to propose a method to suitably choose the combination of R_i and R_j which will guarantee a lower condition number for the interaction matrix. Consequently, the proposed method will improve the numerical stability and robustness of the system. The next section describes the method to select a unique combination of R_i, R_j from the moment of lower orders to control W_X, W_Y and the ellipse to control W_Z.

B. Selection of Unique Combination of R_i and R_j

Normally, the lower order moments are used for selection of R_i and R_j which are more stable and less sensitive to noise. In this work, a unique procedure is adopted for the selection of R_i and R_j from fifteen combinations of SRT moment invariants as in [Tahri and Chaumette, 2005] and denoted as R_{iopt} and R_{jopt} respectively. These fifteen moment invariants are given in Eq. 6.41.

$$\left.\begin{aligned} R_1 &= \frac{A_1}{A_2},\ R_2 = \frac{A_3}{A_4},\ \mathrm{R}_3 = \frac{A_5}{A_6} \\ R_4 &= \frac{A_7}{A_6},\ R_5 = \frac{A_8}{A_6},\ R_6 = \frac{A_9}{A_6} \\ R_7 &= \frac{A_{11}}{A_{10}},\ \mathrm{R}_8 = \frac{A_{12}}{A_{10}},\ R_9 = \frac{A_{13}}{A_{15}},\ \mathrm{R}_{10} = \frac{A_{14}}{A_{15}} \end{aligned}\right\} \quad (6.41)$$

where, $A_s = \dfrac{\mu_{nl}}{m_{00}^{n+l+2/2}}$ and the details are given in [Tahri and Chaumette, 2003; Mamistvalov, 1998; Mukundan and Ramakrishnan, 1998; Prokop and Reeves, 1992]

The proposed method for choosing the unique combination of feature vectors R_i and R_j is described in Algorithm 6.1.

Algorithm 6.1

Step 1: Let us consider F_a be the acquired image described by a set of fifteen feature vectors A_1 to A_{15}. Similarly, let F_d be the desired image and described by the fifteen feature vectors A_1^* to A_{15}^*

$$F_{a(initial\ image)} = \left[A_1, A_2, A_3, \cdots, A_{15}\right]$$

$$F_{d(desired\ image)} = \left[A_1^*, A_2^*, A_3^*, \cdots, A_{15}^*\right]$$

Step 2: For i=1 to length of the feature vector

Compute Hamming distance [Steane, 1996] $HM(i) = \rho\left(A_i, A_i^*\right)$

End for

Step 3: Sort $HM(i)$ //Ascending order

i.e., $\left[HM(i)\right]_{i=1 \text{ to } 15} = \left[\underset{\downarrow}{\boxed{1}}\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10\ 11\ 12\ 13\ 14\ 15\right]$

Most similar feature of the initial image as the desired image

Step 4: Now, $HM(1)$ is the most similar feature of the initial image as the desired image hence its corresponding feature vector of the initial image is selected as the feature vector R_i and its components R_iW_X and R_iW_Y are calculated.

Step 5: For i=1 to length of the feature vector-1

Compute the condition number of the matrix I''_{R_i,R_j}

$$I''_{Ri,Rj} = f(Ri,Rj) = \begin{bmatrix} RiW_X & RiW_Y \\ R_jW_X & R_jW_Y \end{bmatrix}$$

where, R_i is the already selected feature vector and R_j belongs to any one of the feature vector from the rest 14 feature vectors

Step 6: Accumulate the condition number for each combination and sort them in ascending order.

Step 7: Select that combination of $\left(R_i, R_j\right)$ as the best set of feature vector for which the condition numbers is the least and consider them as $\left(R_{iopt}, R_{jopt}\right)$

C. Feature Vector to Control W_Z

In the conventional approaches [Tahri and Chaumette, 2005; Tahri and Chaumette, 2003; Shu et al., 2007], either θ or circularity is used to control the motion of W_Z. However, in this work, ellipse of an image [Flusser et al., 2009], is considered instead of orientation. Ellipse of an image not only gives the information about the orientation of the object but also gives the information about the position of the object. Therefore, the ellipse (as in Figure 6.7) feature vector of the image is used here to control W_Z.

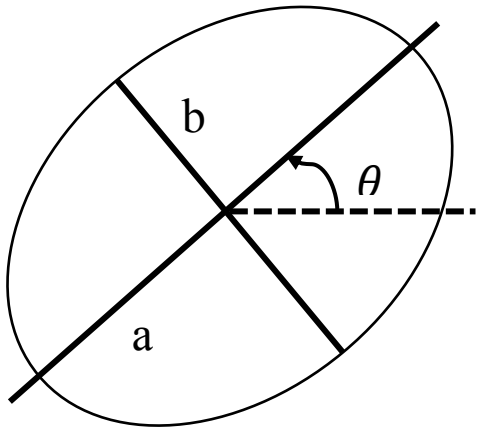


Figure 6.7: Representation of an ellipse

ξ_{a,b,θ_x} represents the shape elongation along x-axis and ξ_{a,b,θ_y} represents the shape spreadness along y-axis. The formulation for ξ_{a,b,θ_x} and ξ_{a,b,θ_y} are defined as

$$\left.\begin{aligned}\xi_{a,b,\theta_x} &= \frac{(T_1+T_2)}{m_{00}}\\ \xi_{a,b,\theta_y} &= \frac{T_1+T_2}{T_1-T_2}\end{aligned}\right\} \tag{6.42}$$

where, $T_1 = \dfrac{(\mu_{02}+\mu_{20})+\left[(\mu_{02}-\mu_{20})+4\mu_{11}^2\right]^{1/2}}{2}$ and

$$T_2 = \frac{(\mu_{02} + \mu_{20}) - \left[(\mu_{02} - \mu_{20}) + 4\mu_{11}^2\right]^{1/2}}{2}$$

where, a, b and θ are the semi major, semi minor and orientation of the ellipse.

m_{nl} and μ_{nl} defines the $(n+l)^{th}$ order moment (m_{nl}) and centroid moment (μ_{nl}) respectively.

In this case, for measuring the orientation θ is used because for any objects in which μ_{11} is zero. The value of θ vanishes as

$$\theta = \frac{1}{2}\arctan\left[\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right] \tag{6.43}$$

But for the ellipse, when μ_{11} =0

$T_1 = \mu_{20}$ and $T_2 = \mu_{02}$

D. Visual Servoing Using Proposed Method

In order to control the 6-DoF of the camera, a vector of six visual features $C = \left[x_n \; y_n \; a_n \; R_{iopt} \; \mathrm{R}_{jopt} \; \xi_{a,b,\theta}\right]$ is utilized to design the control law. The proposed method utilizes the same visual features such as x_G, y_G and a to control the DoF related translational motion i.e., the entries I_{x_n}, I_{y_n} and I_a in the interaction matrix $I_C^{"}$ as in Eq. 6.40 . The only modification lies in the formulation of $I_{R_i}^{"}$, $I_{R_j}^{"}$ and $I_{R_\theta}^{"}$ in the interaction matrix $I_C^{"}$. With these modification the interaction matrix $I_C^{"}$ is replaced by $I_P^{"}$ and is defined by Eq. 6.44

$$I_P^{"} = \begin{bmatrix} -1 & 0 & 0 & a_n\varepsilon_{11} & -a_n(1+\varepsilon_{12}) & y_n \\ 0 & -1 & 0 & a_n(1+\varepsilon_{21}) & -a_n\varepsilon_{11} & -x_n \\ 0 & 0 & -1 & -3y_n/2 & 3x_n/2 & 0 \\ 0 & 0 & 0 & R_{iopt}W_X & R_{iopt}W_Y & 0 \\ 0 & 0 & 0 & R_{jopt}W_X & R_{jopt}W_Y & 0 \\ 0 & 0 & 0 & \xi_{a,b,\theta_{W_X}} & \xi_{a,b,\theta_{W_Y}} & -1 \end{bmatrix} \tag{6.44}$$

Considering the above interaction matrix, the control law for the camera velocity of the eye-in-hand robot configuration is defined in Eq. 6.45.

$$V_C = -\lambda I_P^{"+}\left(C - D^*\right) \tag{6.45}$$

where, $\mathrm{I}_{\mathrm{P}}^{''}$ is the approximation of $I_P^{''}$, $\mathrm{I}_{\mathrm{P}}^{''+}$ is the pseudo-inverse of $\mathrm{I}_{\mathrm{P}}^{''}$ and λ is the time to convergence tuning parameter.

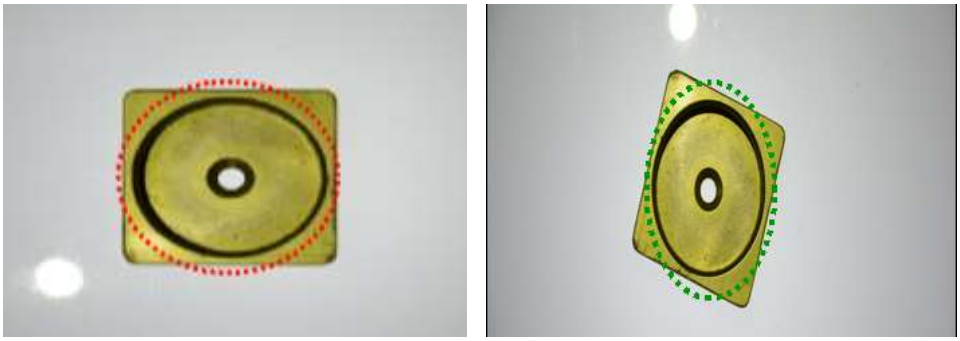
E. Result and Analysis

In this section, the control scheme is evaluated by considering that the image plane and object plane are parallel i.e., $O = P = 0$ and $Q = 2cm$ for the desired set of images. Two sets of images i.e., "Set 1.jpg" and "Set 2.jpg" are considered here. The simulation set up between the desired position of image and the initial position of image considers the rotation amount to be in the range of -30° to 30° along X, Y and Z direction whereas the translation operation is in the range of -2 to 2 along these axes. Figure 6.8 (a) and Figure 6.9(a) shows the initial images for "Set 1.jpg" and "Set 2.jpg" images respectively whereas, Figure 6.8 (b) and Figure 6.9(b) shows the desired images. Combination of six visual features constitutes the feature vector $C = \left[x_n \; y_n \; a_n \; R_{iopt} \; \mathrm{R}_{jopt} \; \xi_{a,b,\theta} \right]$ and is utilized in the formulation of control law (Eq. 6.23). The features x_n, y_n and a_n are considered for controlling the translational DoFs whereas, the appropriately chosen features vectors R_{iopt} and R_{jopt} using Algorithm 6.1 are utilized for the control of rotational DoF W_X and W_Y respectively. Similarly, ellipse of the image $\xi_{a,b,\theta}$ is utilized to control the rotational DoF W_Z. The position of ellipse for the initial images (Figure 6.8 (a) and Figure 6.9(a)) and desired images (Figure 6.8 (b) and Figure 6.9(b)) are shown in red color and green color dotted line around the images respectively. Finally, the visual features based on basic moment invariants are used to formulate the interaction matrix for the desired positions for the "Set 1.jpg" and "Set 2.jpg" are described in Eq. 6.46 and Eq. 6.47 respectively.

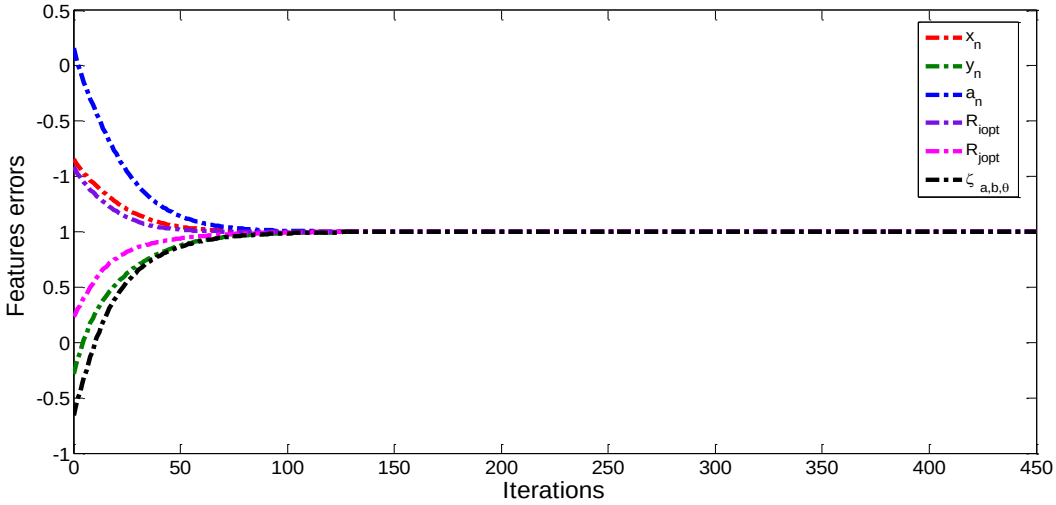
$$I_P^{''}\Big|_{Set1.jpg} = \begin{bmatrix} -1.00 & 0 & 0 & 0.01 & -1.01 & 0 \\ 0 & -1.00 & 0 & 1.01 & -0.01 & 0 \\ 0 & 0 & -1.00 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.67 & -0.21 & 0 \\ 0 & 0 & 0 & -0.14 & 0.32 & 0 \\ 0 & 0 & 0 & -0.02 & -0.01 & -1.00 \end{bmatrix} \tag{6.46}$$

$$I_P^{"}\Big|_{set2.jpg} = \begin{bmatrix} -1.00 & 0 & 0 & 0.01 & -1.08 & 0 \\ 0 & -1.00 & 0 & 1.08 & -0.01 & 0 \\ 0 & 0 & -1.00 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.50 & -0.01 & 0 \\ 0 & 0 & 0 & -0.13 & 0.29 & 0 \\ 0 & 0 & 0 & -0.20 & -0.01 & -1.00 \end{bmatrix} \tag{6.47}$$

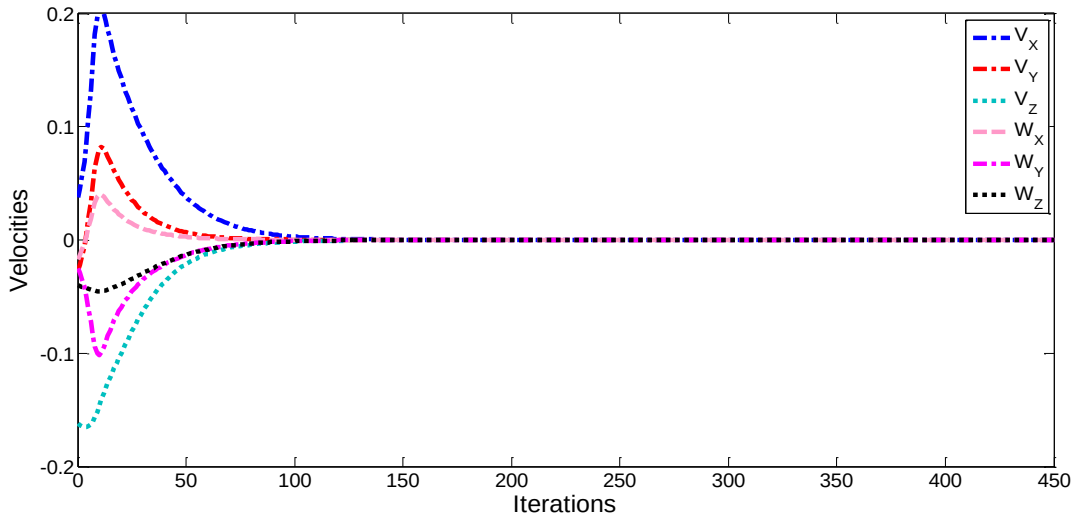
From the interaction matrices for both the image data sets, it is observed that they are sparse and block triangular in nature. The nonzero terms are present around the diagonal and produce a satisfactory decoupling behavior with a lower condition number. The condition number of the interaction matrix for image "set1.jpg" is 6.31 whereas, for image "Set2.jpg" it is found to be 7.50. The servoing is achieved using the desired value of the interaction matrix $I_P^{"}$ described in Eq. 6.46 and Eq. 6.47 for the two sets of images. The convergence of the features errors as in Figure. 6.8 (c) and in Figure 6.9 (c) for the two set of experiments. The plots for camera velocities are described in Figure 6.8 (d) and in Figure 6.9 (d) for the two sets of test images. From the above results, it is observed that plot of visual features errors and the camera velocity contains no oscillation during the period of convergence which induces stability to the control law and achieves better exponential decoupling behavior.



(a) (b)

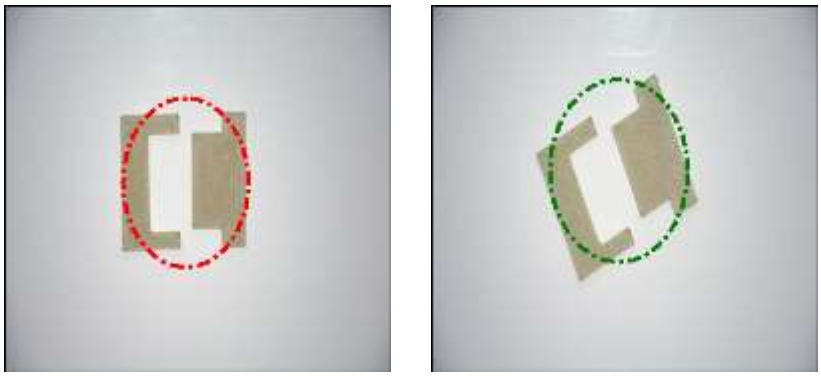


(c)

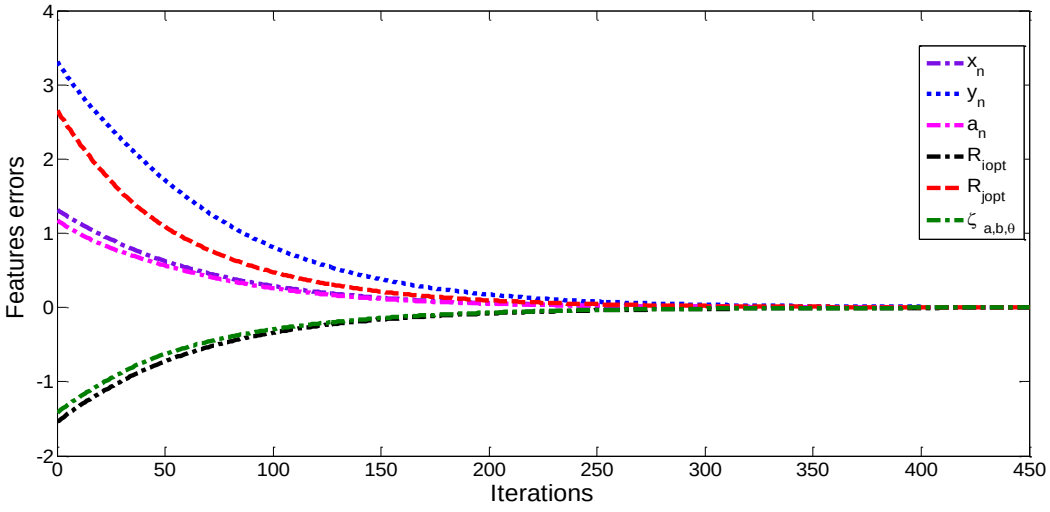


(d)

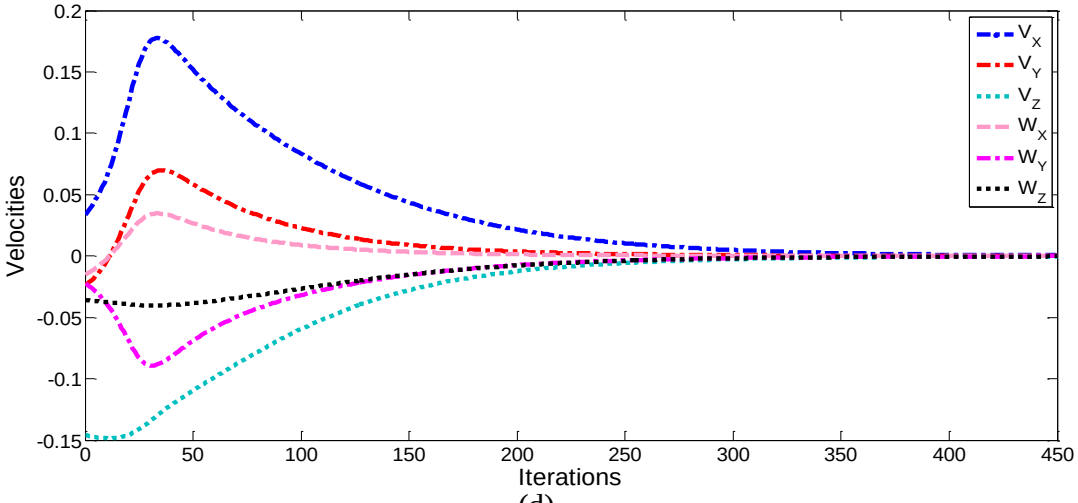
Figure 6.8: "Set 1.jpg" (a) Initial image (b) Desired image (c) Convergence of the features error in the proposed control scheme for desired image (d) Camera velocity in the proposed control scheme for desired image



(a) (b)



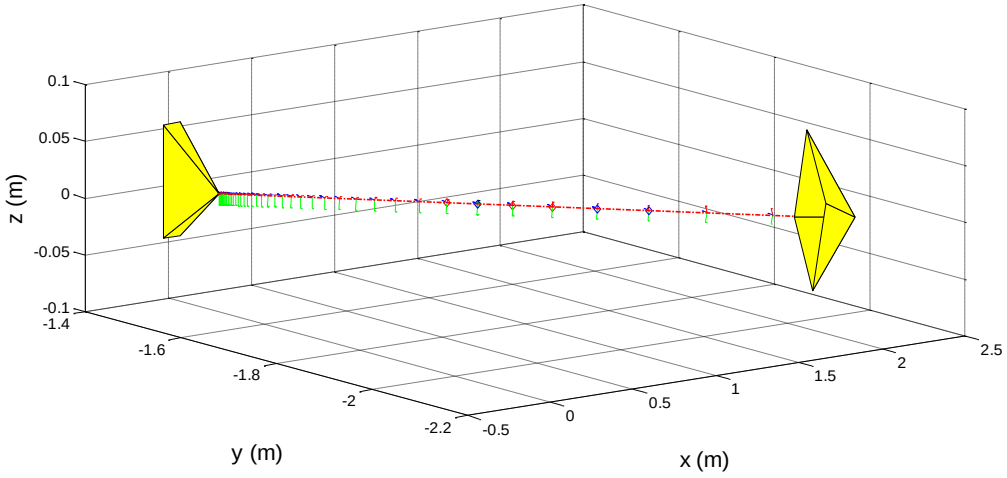
(c)



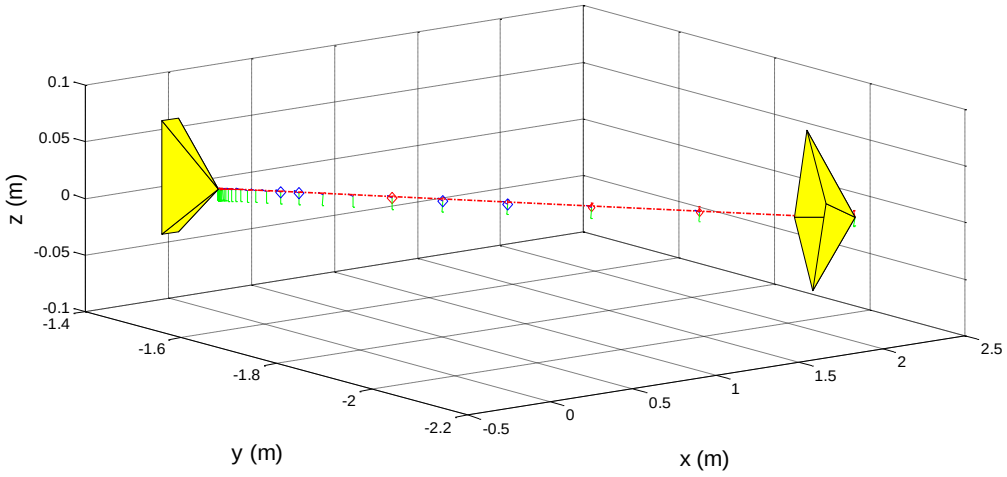
(d)

Figure 6.9. "Set 2.jpg" (a) Initial image (b) Desired image (c) Convergence of the features error in the proposed control scheme for desired image (d) Camera velocity in the proposed control scheme for desired image

Figure 6.10 (a) and (b) show the plot of camera trajectory for the images "set1.jpg" and "set2.jpg" respectively. From the plot, it is observed that irrespective of the presence of different transformation between the initial and desired position of both the images, the camera trajectories are approximately straight line.



(a)



(b

Figure 6.10: Camera trajectory for the proposed control scheme. (a) for the desired image of "Set1.jpg" (b) for the desired image of "Set2.jpg"

6.6 Visual Servoing Based on Orthogonal Moment Invariants

In the previous section, the non-orthogonal moment invariants of lower order were used for finding visual feature to control 6 DoF. These lower order moments have limited image representation capability and the representation capability of the moment increases with the increase in moment order. However, the higher order non-orthogonal moment invariants suffer from a number of disadvantages such as

- sensitivity to noise is high
- suffers from numerical instability due to the large variation of dynamic range of values for different orders.
- contains large volume of redundant information due to non-orthogonal basis function.

As compared to non-orthogonal moment invariants, higher order orthogonal moment invariants such as Legendre and Zernike moments [Teague, 1980; Teh and Chin, 1988; Flusser et al., 2009; Shu et al., 2007; Coatrieux, 2008] are less sensitive to noise and possess minimum information redundancy. These invariants are stable in presence of local deformation and they are able to achieve a near zero value of redundancy measure in a set of moment function.

However, the main disadvantage of orthogonal moment is lack of native scale invariance which can be achieved by the normalization process. Due to these efficient characteristics of the orthogonal moments, the higher order Legendre moment and Zernike moment invariants can be effectively used to provide feature vectors for designing a stable and robust control law.

6.6.1 Visual Servoing Based on Legendre Moment Invariants

A. Legendre Moment

2D, Legendre moment (LM) [Teague, 1980; Teh and Chin, 1988] of an image $f(x,y)$ is described as

$$LM_{nl} = C_{nl} \int_{-1}^{1} \int_{-1}^{1} P_n(x) P_l(x) f(x,y)\, dx\, dy \quad n,l = 0,1,2 \ldots \tag{6.48}$$

where $P_n(x)$ is defined by the Legendre polynomial as

$$P_n(x) = \sum_{p=0}^{n} a_{np} x^p$$

$$C_{nl} = \frac{(2n+1)(2l+1)}{4} \text{ and}$$

$$a_{np} = \begin{cases} \sqrt{2n+1}/2 \dfrac{(-1)^{\frac{n-p}{2}}(n+p)!}{2^n\left(\frac{n-p}{2}\right)!\left(\frac{n+p}{2}\right)!\,p!} & n-p = even \\ 0 & n-p = odd \end{cases}$$

In literature, a variety of LM invariants are already proposed which are invariants to various type of geometric transformation such as scale, rotation and translation. In [Chong et al., 2004; Zhang et al., 2011], a complete set of LM insensitive to affine translation is well presented. These invariants can be efficiently used as the feature vectors to control the DoF of a camera [Tamtsia et al., 2012; Hosny, 2005].

However, the only issue is the computational complexity in the calculation of LM invariants as described in Eq. 6.48. In order to avoid this scenario, an indirect method of computation is employed in the proposed work.

a. Indirect Method to Compute Legendre Moment Invariants

In indirect method of computation, the traditional LM is described as a linear combination of geometric moment invariants which are invariants to scaling, rotation and translation (SRT).

The $(n,l)^{th}$ order of LM can be expressed by a combination of geometric moment of same order or less as given in Eq. 6.49.

$$LM_{nl} = \frac{(2n+1)(2l+1)}{4} \sum_{i=0}^{\left\lfloor \frac{n}{2} \right\rfloor} \sum_{j=0}^{\left\lfloor \frac{l}{2} \right\rfloor} B_{n,i} B_{l,j}\, \eta_{n-2i,l-2j} \tag{6.49}$$

Where the geometric moment invariant $\eta_{p,q}$ of order (p,q) is defined by

$$\eta_{pq} = \sum_{k=0}^{p} \sum_{j=0}^{q} \binom{p}{k} \binom{q}{j} (-1)^k (\sin\theta)^{q+k-j} (\cos\theta)^{p+j-k} \mu'_{p+q-k-j,k+j} \tag{6.50}$$

The central moment μ'_{pq} of order (p,q) and geometric moment m_{pq} of order (p,q) are described in Eq. 6.51 and Eq. 6.52 respectively.

$$\mu'_{pq} = \sum_{k=0}^{p}\sum_{j=0}^{q}\binom{p}{k}\binom{q}{j}(-\bar{x})^{p-k}(-\bar{y})^{q-j}m_{kj} \tag{6.51}$$

$$m_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x,y)\,dxdy \tag{6.52}$$

B. Formulation of Interaction Matrix

The interaction matrix related to the 2D LM is given as

$$I_{LM_{nl}} = \begin{bmatrix} LM_{V_X} & LM_{V_Y} & LM_{V_Z} & LM_{W_X} & LM_{W_Y} & LM_{W_Z} \end{bmatrix} \tag{6.53}$$

Where,

$$LM_{V_X} = -(i+1)P\,LM_{i,j} - iQ\,LM_{i-1,j+1} + i\left[\left(\frac{2C_{10}}{C_{00}}-1\right)Px_G + \left(\frac{C_{10}}{C_{00}}-1\right)Qy_G\right]LM_{i-1,j} + j\left(\frac{C_{01}}{C_{00}}\right)Py_G LM_{i,j-1}$$

$$LM_{V_Y} = -jP\,LM_{i+1,j-1} - (j+1)Q\,LM_{i,j} + j\left[\left(\frac{C_{01}}{C_{00}}-1\right)Px_G + \left(2\frac{C_{01}}{C_{00}}-1\right)Qy_G\right]LM_{i,j-1} + i\left(\frac{C_{10}}{C_{00}}\right)Rx_G LM_{i-1,j}$$

$$LM_{V_Z} = (i+j+3)\left(PLM_{i+1,j} + QLM_{i,j+1} + RLM_{i,j}\right) - RLM_{i,j}$$

a. Feature Vector to Control Translational DoF

The visual features x_G, y_G and a are utilized to control the translational motion V_X, V_Y and V_Z respectively and they are expressed in Eq. 6.54.

Expressing the geometric moment in terms of LM, the following is obtained.

$$\left.\begin{aligned} x_G &= \left(\frac{C_{00}}{C_{10}}\right)\left(\frac{\lambda_{10}}{\lambda_{00}}\right) \\ y_G &= \left(\frac{C_{00}}{C_{01}}\right)\left(\frac{\lambda_{01}}{\lambda_{00}}\right) \end{aligned}\right\} \tag{6.54}$$

where,

$$\left.\begin{aligned} C_{nl} &= \frac{(2n+1)(2l+1)}{4} \\ C_{00} &= 1/4 \quad C_{10} = \frac{(2\times1+1)\times1}{4} = 3/4 \\ \lambda_{10} &= \frac{3}{4}m_{10} \quad \lambda_{00} = \frac{m_{00}}{4} \\ x_G &= \left(\frac{m_{10}}{m_{00}}\right), y_G = \left(\frac{m_{01}}{m_{00}}\right) \end{aligned}\right\}$$

In order to obtain the same dynamics for all the translational DoF, normalization process as described in Eq. 6.36 is applied and the normalized visual features x_n, y_n and a_n are used to control

V_X, V_Y and V_Z respectively. By considering these normalized visual features, the interaction matrix $I_{LM_{n,l}}$ related to V_X, V_Y and V_Z is defined in Eq. 6.55.

$$\begin{bmatrix} I''_{x_n} \\ I''_{y_n} \\ I''_{a_n} \end{bmatrix} = \begin{bmatrix} -1 & 0 & 4Rx_n & a_n\varepsilon_{11} & -a_n(1+\varepsilon_{12}) & 3y_n \\ 0 & -1 & -4Ry_n & a_n(1+\varepsilon_{21}) & a_n\varepsilon_{11} & -3x_n \\ 0 & 0 & -1 & \frac{-9}{2}y_n & \frac{9}{2}x_n & 0 \end{bmatrix} \tag{6.55}$$

It is noted that the block which computes to the translational DoF is triangular with the following conditions.

$$\left.\begin{aligned} \varepsilon_{12} &= 4\left(\gamma_{20} + \frac{3}{2}\left(\frac{C_{20}}{C_{00}} - \frac{3}{9}\right){x_G}^2\right) \\ \varepsilon_{21} &= 4\left(\gamma_{02} + \frac{3}{2}\left(\frac{C_{02}}{C_{00}} - \frac{3}{4}\right){y_G}^2\right) \\ x_G y_G &= \left(\frac{m_{10}}{m_{00}} \times \frac{m_{01}}{m_{00}}\right) \end{aligned}\right\} \tag{6.56}$$

where,

$$\gamma_{pq} = \frac{\varphi_{pq}}{\lambda_{00}}$$

$$\varphi_{pq} = C_{pq}\int_{-1}^{1}\int_{-1}^{1} p_p(x - x_G)p_q(y - y_G)\,f(x,y)\,dx\,dy$$

$$\varepsilon_{11} = 4\varphi_{11} + \frac{81}{4}x_G y_G$$

$$\varepsilon_{12} = 4\varphi_{20} + \frac{45}{2}{x_G}^2$$

$$\varepsilon_{21} = 4\varphi_{02} + \frac{45}{2}{y_G}^2$$

b. Feature vector to controlling the Rotational DoF

The rotational motion w_X and w_Y are controlled by appropriately selecting two feature vectors among the 3rd and 4th orders LM which are invariant to SRT as given in Eq. 6.57 and Eq. 6.58.

$$\left.\begin{aligned}\phi_1 &= \frac{35}{8}\eta_{3,0}\\ \phi_2 &= \frac{45}{8}\eta_{2,1}\\ \phi_3 &= \frac{45}{8}\eta_{1,2}\\ \phi_4 &= \frac{35}{8}\eta_{0,3}\end{aligned}\right\} \tag{6.57}$$

$$\left.\begin{aligned}\phi_5 &= \frac{9}{4}\left(\frac{35}{8}\eta_{4,0}-\frac{30}{8}\eta_{2,0}+\frac{3}{8}\eta_{0,0}\right)\\ \phi_6 &= \frac{105}{8}\eta_{3,1}\\ \phi_7 &= \frac{25}{4}\left(\frac{9}{4}\eta_{2,2}-\frac{3}{4}\eta_{2,0}-\frac{3}{4}\eta_{0,2}+\frac{1}{4}\eta_{0,0}\right)\\ \phi_8 &= \frac{105}{8}\eta_{1,3}\\ \phi_9 &= \frac{9}{4}\left(\frac{35}{8}\eta_{0,4}-\frac{30}{8}\eta_{0,2}+\frac{3}{8}\eta_{0,0}\right)\end{aligned}\right\} \tag{6.58}$$

Among these nine LM invariants, two visual features R_{iopt} and R_{jopt} are appropriately selected based on Algorithm 6.1 to control the DoF for W_X, W_Y related to rotational velocity. The ellipse of the image is used to control the rotational DoF W_Z. Utilizing these feature vectors the related interaction matrix is defined as in Eq. 6.59.

$$\begin{bmatrix} I''_{R_i} \\ I''_{R_J} \\ I''_{R_\zeta} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & R_{iopt}W_X & R_{iopt}W_Y & 0 \\ 0 & 0 & 0 & R_{jopt}W_X & R_{jopt}W_Y & 0 \\ 0 & 0 & 0 & \xi_{a,b,\theta}W_X & \xi_{a,b,\theta}W_Y & -1 \end{bmatrix} \tag{6.59}$$

The visual features are calculated in such a way that the combination will provide lower condition number to make the system stable. Considering both translational and rotational feature vectors, the final interaction matrix is calculated as given in Eq. 6.60.

$$I'_{LM_{n,l}} = \begin{bmatrix} -1 & 0 & -4Rx_n & a_n\varepsilon_{11} & -a_n(1+\varepsilon_{12}) & 3y_n \\ 0 & -1 & -4Ry_n & a_n(1+\varepsilon_{21}) & -a_n\varepsilon_{11} & -3x_n \\ 0 & 0 & -1 & -9y_n/2 & 9x_n/2 & 0 \\ 0 & 0 & 0 & R_{iopt}W_X & R_{iopt}W_Y & 0 \\ 0 & 0 & 0 & R_{jopt}W_X & R_{jopt}W_Y & 0 \\ 0 & 0 & 0 & \xi_{a,b,\theta}W_X & \xi_{a,b,\theta}W_Y & -1 \end{bmatrix} \quad (6.60)$$

C. Result and Analysis

In this section, the control scheme is evaluated and it is considered that the image plane and object plane are parallel i.e., $O = P = 0$ and $Q = 2cm$ for the desired set of images. The two sets of images i.e., "Set 1.jpg" and "Set 2.jpg" are considered here. The simulation set up between the desired position of image and the initial position of image considers the rotation amount to be in the range of -30° to 30° along X, Y and Z direction whereas the translation operation is in the range of -2 to 2 along these axes. Figure 6.11 (a) and Figure 6.12(a) shows the initial images for "Set 1.jpg" and "Set 2.jpg" images respectively whereas, Figure 6.11 (b) and Figure 6.12(b) shows the desired images. Combination of six visual features constitutes the feature vector $C = \left[x_n \; y_n \; a_n \; R_{iopt} \; \mathrm{R}_{jopt} \; \xi_{a,b,\theta} \right]$ and is utilized in the formulation of control law (Eq. 6.23). The features x_n, y_n and a_n are considered for controlling the translational DoFs. These features are determined from Legendre moment invariants expressed in linear combination of geometric moment invariants which are invariant to SRT. Similarly the feature vectors R_{iopt} and R_{jopt} are selected using Algorithm 6.1 are utilized for the control of rotational DoFs W_X and W_Y respectively and ellipse of the image $\xi_{a,b,\theta}$ is utilized to control the rotational DoF W_Z. The position of ellipse for the initial images (Figure 6.11 (a) and Figure 6.12(a)) and desired images (Figure 6.11 (b) and Figure 6.12(b)) are shown in red color and green color dotted line around the images respectively. Finally, the visual features based on basic moment invariants are used to formulate the interaction matrix for the desired positions are described in Eq. 6.61 and Eq. 6.62 respectively.

$$I^{''}_{LM_{nl}}\Big|_{Set1.jpg} = \begin{bmatrix} -1.00 & 0 & 0 & 0.01 & -0.08 & 0 \\ 0 & -1.00 & 0 & 0.08 & -0.01 & 0 \\ 0 & 0 & -1.00 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.90 & -0.10 & 0 \\ 0 & 0 & 0 & -0.68 & 0.37 & 0 \\ 0 & 0 & 0 & -0.20 & -0.01 & -1.00 \end{bmatrix} \tag{6.61}$$

$$I^{''}_{LM_{nl}}\Big|_{set2.jpg} = \begin{bmatrix} -1.00 & 0 & 0 & 0.01 & -1.00 & 0 \\ 0 & -1.00 & 0 & 1.00 & -0.01 & 0 \\ 0 & 0 & -1.00 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.61 & -0.25 & 0 \\ 0 & 0 & 0 & -0.18 & 0.52 & 0 \\ 0 & 0 & 0 & -0.06 & -0.03 & -1.00 \end{bmatrix} \tag{6.62}$$

From the interaction matrices for both the image data sets it is observed that they are sparse and block triangular in nature. The nonzero terms are present around the diagonal and produce a satisfactory decoupling behavior with a lower condition number. The condition number of the interaction matrix for image "set1.jpg" is 3.42 whereas, for image "Set2.jpg" it is found to be 4.02. Consequently, as compared to the previous section, the present control law utilizing feature vectors using LMIs produces more stable and robust control system.

The servoing is achieved using the desired value of the interaction matrix $I^{''}_{LM_{nl}}$ described in Eq. 6.61 and Eq. 6.62 respectively and result in the convergence of the features errors as in Figure 6.11 (c) and in Figure 6.12(c) for the two set of experiments. The plots for camera velocities are described in Figure 6.11(d) and in Figure 6.12(d) for the two sets of test images. From the above figures it is observed that plot of visual features errors and the camera velocity contains almost nil oscillation during the period of convergence and converges at a faster rate as compared to Figure 6.8(c) and Figure 6.9(c). Consequently the control law attains better stability and achieves improved exponential decoupling behavior.

Figure 6.13 (a) and (b) show the plot of camera trajectory for the images "set1.jpg" and "set2.jpg" respectively. From the plot it is observed that irrespective of the presence of different transformation between the initial and desired position of both the images, the camera trajectories are approximately straight line.

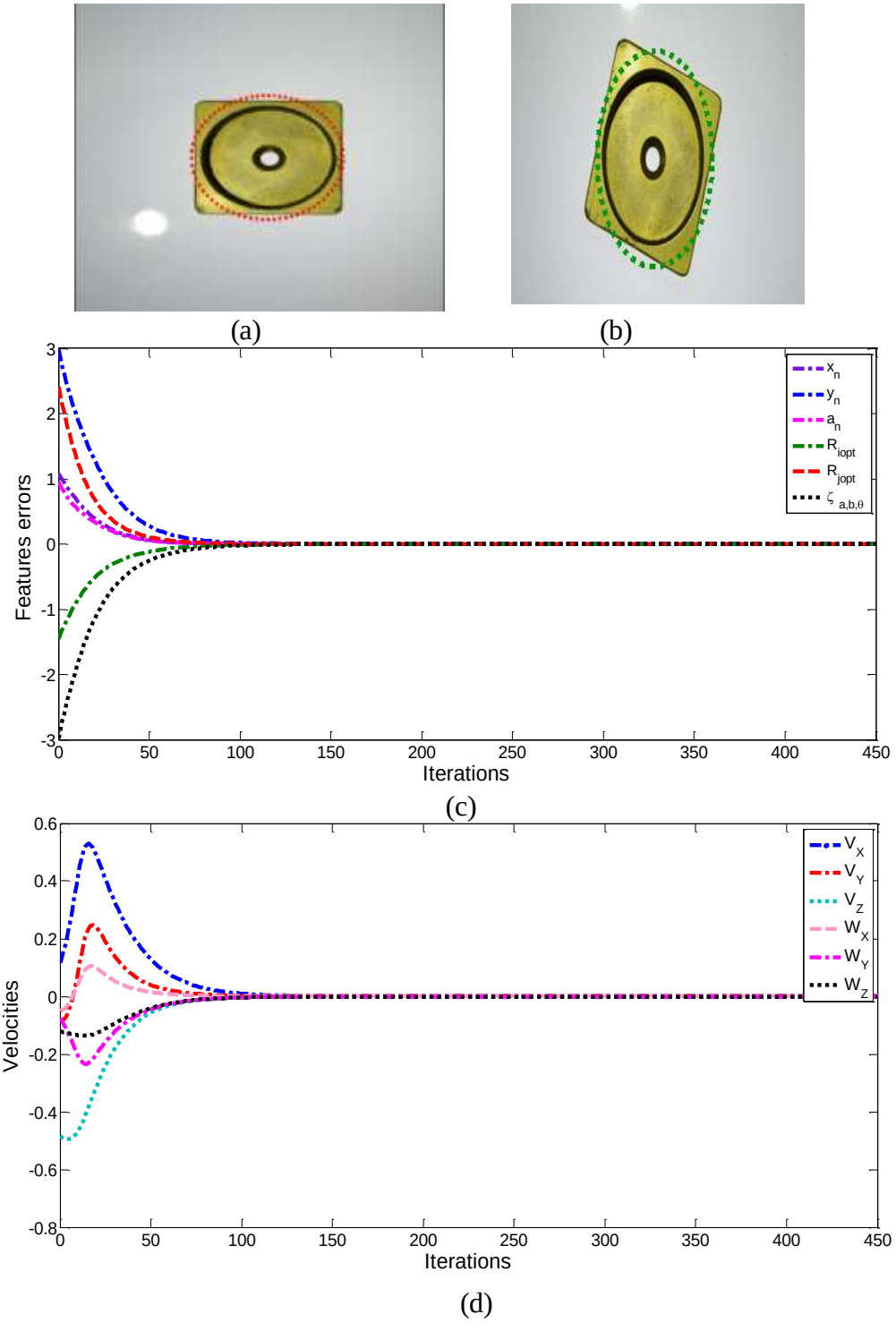
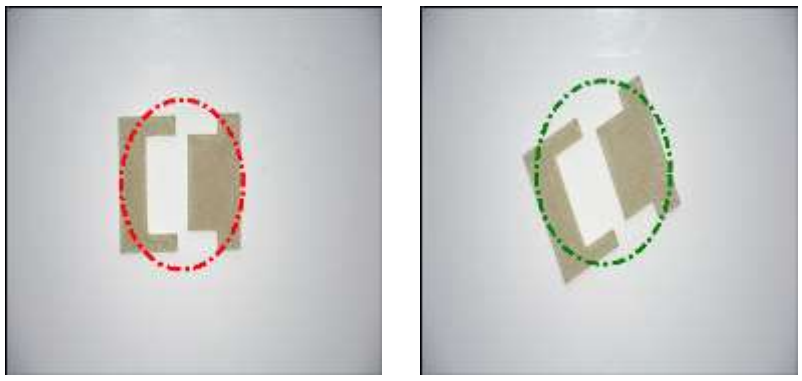
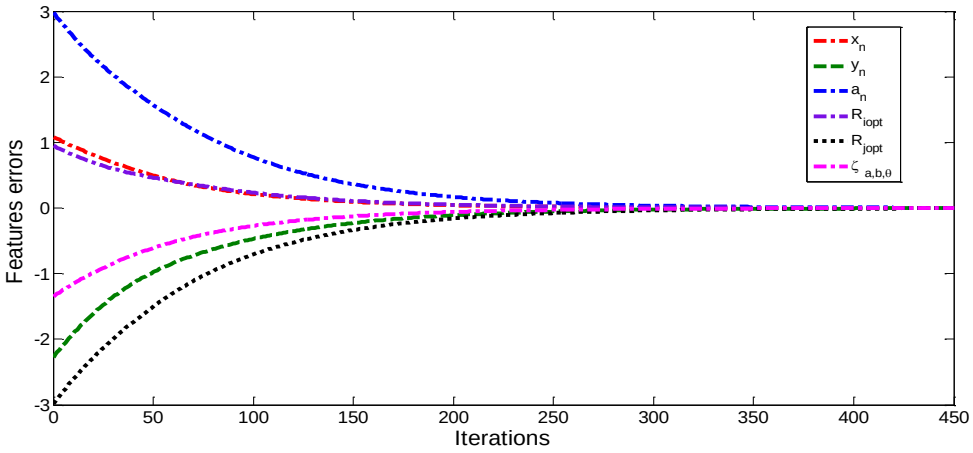


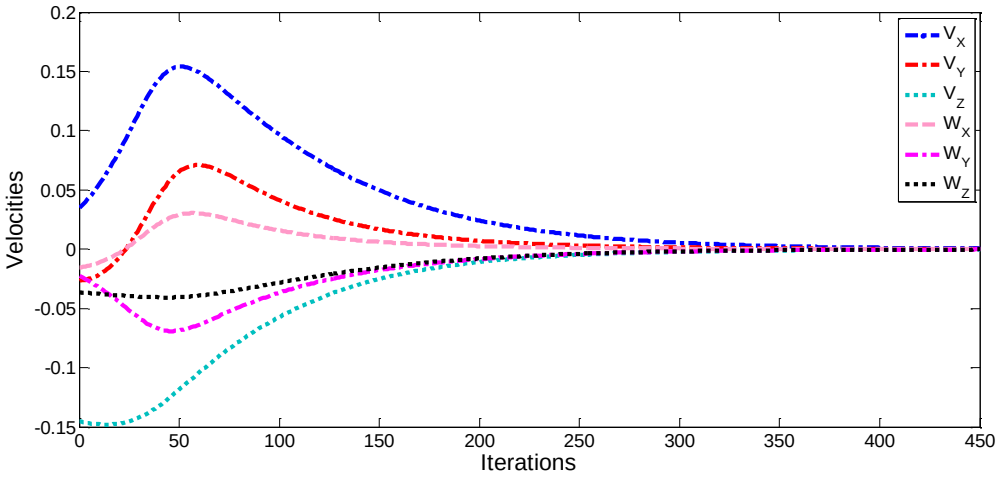
Figure 6.11: "Set 1.jpg" (a) Initial image (b) Desired image (c) Convergence of the features error in the proposed control scheme for desired image (d) Camera velocity in the proposed control scheme for desired image.



(a) (b)

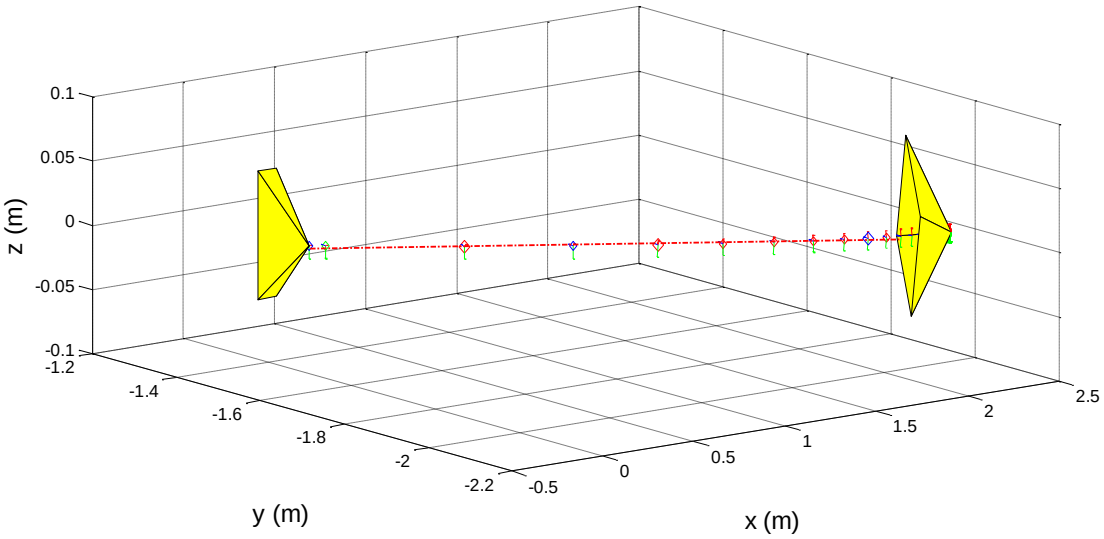


(c)

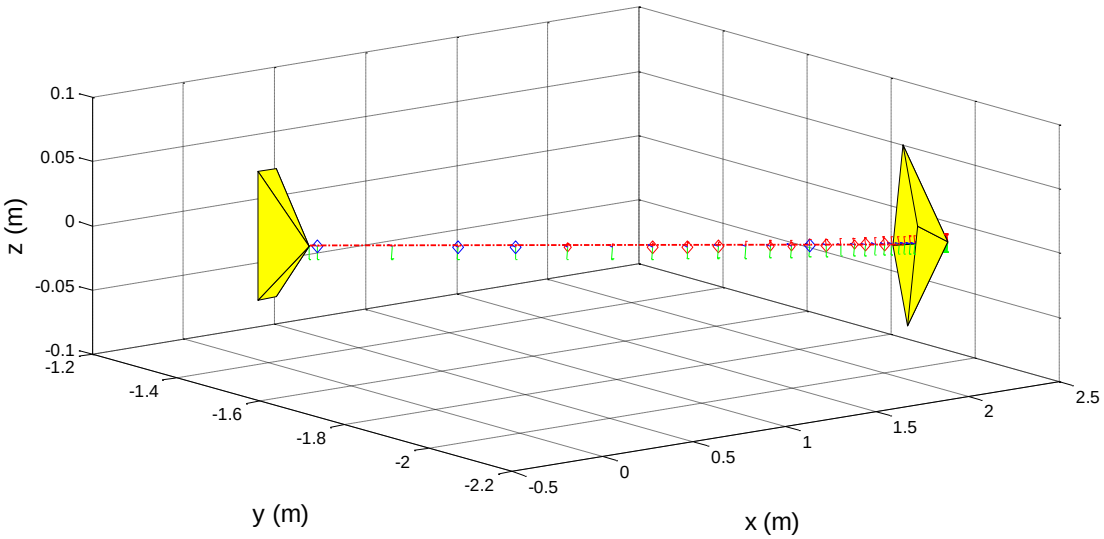


(d)

Figure 6.12: "Set 1.jpg" (a) Initial image (b) Desired image (c) Convergence of the features error in the proposed control scheme for desired image (d) Camera velocity in the proposed control scheme for desired image.



(a)



(b)

Figure 6.13: Camera trajectory for the proposed control scheme. (a) for the desired image of "Set1.jpg" (b) for the desired image of "Set2.jpg"

6.6.2 Visual Servoing Based on Zernike Moment Invariants

A. Zernike Moment

In the previous discussion, the higher order LM invariants are utilized for the design of the control law. However, as compared to LM invariants, the Zernike moment invariants (ZMI) are more robust to noise and provide better image representation capability [Khotanzad and Hong, 1990; Hosaini et al., 2013]. Hence, in this section an indirect method is used to calculate the ZMI and based on these invariants, the control law for image based visual servoing is designed.

Let $f(x,y)$ be a 2D intensity image with size $N \times N$. The formulation of Zernike moment $\mathbb{Z}_{nl}$ of order $n+l$ is given by [Teague, 1980; Teh and Chin, 1988]

$$\mathbb{Z}_{nl} = \frac{n+1}{\pi} \sum_{x} \sum_{y} f(x,y) V_{nl}^{*}(\rho,\theta)\, dx\, dy \tag{6.63}$$

Where,

$V_{nl}(\rho,\theta) = R_{nl}(\rho)\exp(jl\theta)$ and V_{nl}^{*} is the complex conjugate of V_{nl}

$$R_{nl}(\rho) = \sum_{s=0}^{n-\frac{|l|}{2}} (-1)^{s} \frac{(n-s)!}{s!\left(\frac{n+|l|}{2}-s\right)!\left(\frac{n+|l|}{2}-s\right)!} \rho^{n-2s}$$

The factorial term present in the above eq. increases the computational complexity of the Zernike moment. Hence, for the reduction of computational complexity, the indirect method of computing the Zernike moment via the Geometric moment is adopted [Saad and Rusli, 2004].

a. Indirect Method to Compute Zernike Moment Invariants Based on Zernike Moment Invariants

The magnitude of the Zernike moment of higher order are rotation invariant and to achieve the scale and translation invariant, the Zernike moments are expressed in terms of linear combination of Geometric moment which will provide invariants to SRT transformation. Eq. 6.64 provides the indirect way of representation of Zernike moment in terms of Geometric moment.

$$\mathbb{Z}_{nl} = \mathbb{Z}_{nl} = \frac{n+1}{\pi} \left\{ \sum_{\substack{|k|=l \\ n-|l|=even}}^{n} \sum_{j=0}^{q} \sum_{m=0}^{|l|} (-i)^{m} \binom{q}{j} \binom{l}{m} B_{nlk}\, \mu_{k-2j-l+m,2j+l-m} \right\} \tag{6.64}$$

where, B_{nlk} is the coefficient matrix and μ_{nl} is the central moment.

B. Formulation of Interaction Matrix

The interaction matrix related to the 2D ZMI is given as

$$I_{ZM_{nl}} = \begin{bmatrix} ZM_{V_X} & ZM_{V_Y} & ZM_{V_Z} & ZM_{W_X} & ZM_{W_Y} & ZM_{W_Z} \end{bmatrix} \quad (6.65)$$

Here, the ZM_{V_X}, ZM_{V_Y} and ZM_{V_Z} are used for controlling the translational DoF. ZM_{W_X} ZM_{W_Y} and ZM_{W_Z} are utilized for controlling the rotational DoF.

a. Feature Vectors to control translational and rotational DoF

Like the previous discussion, normalized visual features x_n, y_n and a_n are used to control the translational motion V_X, V_Y and V_Z respectively. The interaction matrix coefficient for the translational motion V_X, V_Y and V_Z are same as used in Eq. 6.36. The magnitude of moment invariants of order $2-4$ are utilized as the rotation invariant feature vectors to control W_X, W_Y, and W_Z respectively. The set of ZMIs which are invariant to SRT geometric transformation are presented in Eq. 6.64.

Three visual features R_{iopt}, R_{jopt} and R_{kopt} are chosen from the set of Zernike moments which are invariant to SRT based on Algorithm 6.1 to control the rotational DoF W_X, W_Y and W_Z. As per Algorithm 6.1 selection of R_{iopt} and R_{jopt} is based on the computation of feature vectors for which the condition number of the matrix is low. Similarly with R_{iopt} as the reference feature vector, the combination of feature vector which will provide the second lowest condition number is considered as R_{kopt}

$$\left.\begin{aligned}
\mathbb{Z}_{00} &= \frac{m_{00}}{\pi} \\
\mathbb{Z}_{11} &= \frac{2}{\pi}\left(m_{10} - i m_{01}\right) \\
\mathbb{Z}_{20} &= \frac{6}{\pi}\left(m_{20} + m_{02}\right) - \frac{3}{\pi}\left(m_{00}\right) \\
\mathbb{Z}_{22} &= \frac{3}{\pi}\left[\left(m_{20} - m_{02} - 2i m_{11}\right]\right. \\
\mathbb{Z}_{31} &= \frac{12}{\pi}\left(m_{30} + m_{21}\right) - \frac{12i}{\pi}\left(m_{03} + m_{21}\right) - \frac{8}{\pi}\left(m_{10} - i\, m_{01}\right) \\
\mathbb{Z}_{33} &= \frac{4}{\pi}\left(m_{30} - 3m_{21}\right) + \frac{4i}{\pi}\left(m_{03} - 3m_{21}\right) \\
\mathbb{Z}_{40} &= \frac{30}{\pi}\left(m_{40} + 2m_{22} + m_{04}\right) - 6\pi(m_{20} + m_{00})
\end{aligned}\right\} \tag{6.66}$$

Now, the interaction matrix can be defined as

$$\begin{bmatrix} I''_{R_i} \\ I''_{R_j} \\ I''_{R_k} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & R_{iopt}W_X & R_{iopt}W_Y & 0 \\ 0 & 0 & 0 & R_{jopt}W_X & R_{jopt}W_Y & 0 \\ 0 & 0 & 0 & R_{kopt}W_X & R_{kopt}W_Y & 1 \end{bmatrix} \tag{6.67}$$

The visual features are calculated in such a way that the combination will provide lower condition number to make the system stable. Considering both translational and rotational feature vectors, the final interaction matrix is calculated as given in Eq. 6.67.

$$I''_{ZM_{n,l}} = \begin{bmatrix} -1 & 0 & -4Rx_n & a_n\varepsilon_{11} & -a_n\left(1+\varepsilon_{12}\right) & 3y_n \\ 0 & -1 & -4Ry_n & a_n(1+\varepsilon_{21}) & -a_n\varepsilon_{11} & -3x_n \\ 0 & 0 & -1 & -9y_n/2 & 9x_n/2 & 0 \\ 0 & 0 & 0 & R_{iopt}W_X & R_{iopt}W_Y & 0 \\ 0 & 0 & 0 & R_{jopt}W_X & R_{jopt}W_Y & 0 \\ 0 & 0 & 0 & R_{kopt}W_X & R_{kopt}W_Y & -1 \end{bmatrix} \tag{6.68}$$

C. Result and Analysis

The image representation capability and robustness to noise characteristic of ZMI motivates the design of control scheme. Here, an indirect method of determining the invariants is used. The Zernike moment of order 2−4 are used for calculation of the rotational invariant feature vectors. The final interaction matrix $I^{*}_{ZM_{n,l}}$ is determined by considering the visual features computed by the indirect method and the control scheme is evaluated by considering that the image plane and object plane are parallel i.e., $O = P = 0$ and $Q = 2cm$ for the desired set of images. Two sets of images i.e., "Set 1.jpg" and "Set 2.jpg" are considered here. The simulation set up between the desired position of image and the initial position of image considers the rotation amount to be in the range of -30° to 30° along X, Y and Z direction whereas the translation operation is in the range of −2 to 2 along these axes. Figure 6.14 (a) and Figure 6.15 (a) shows the initial images for "Set 1.jpg" and "Set 2.jpg" images respectively whereas, Figure 6.14 (b) and Figure 6.15 (b) shows the desired images. Combination of six visual features constitutes the feature vector $C = \left[x_n \; y_n \; a_n \; R_{iopt} \; \mathrm{R}_{jopt} \; \mathrm{R}_{kopt} \right]$ and is utilized in the formulation of control law (Eq. 6.23). To control the translational motion V_X, V_Y and V_Z, the normalized visual features x_n, y_n and a_n are determined by the Zernike moment. From Algorithm 6.1, optimal value for R_{iopt} and R_{jopt} are selected. This selection is based on the low condition number of the matrix. Similarly, R_{kopt} is selected by considering R_{iopt} as the reference feature vector. The combination of feature vector which will provide the second lowest condition number is considered as R_{kopt}. The position of ellipse for the initial images (Figure 6.14 (a) and Figure 6.15(a)) and desired images (Figure 6.14 (b) and Figure 6.15(b)) are shown in red color and green color dotted lines around the images respectively. Finally, the visual features based on basic moment invariants are used to formulate the interaction matrix for the desired positions are described in Eq. 6.69 and Eq. 6.70 respectively.

$$I^{"}_{ZM_{nl}}\Big|_{Set1.jpg} = \begin{bmatrix} -1.00 & 0 & 0 & 0.01 & -0.58 & 0.02 \\ 0 & -1.00 & 0 & 0.58 & -0.01 & 0.02 \\ 0 & 0 & -1.00 & -0.01 & -0.02 & 0 \\ 0 & 0 & 0 & -0.43 & -0.50 & 0 \\ 0 & 0 & 0 & -0.74 & 0.17 & 0 \\ 0 & 0 & 0 & -0.05 & -0.07 & -1.00 \end{bmatrix} \tag{6.69}$$

$$I^{"}_{ZM_{nl}}\Big|_{set2.jpg} = \begin{bmatrix} -1.00 & 0 & 0 & 0.01 & 0 & 0 \\ 0 & -1.00 & 0 & 0 & -0.01 & 0 \\ 0 & 0 & -1.00 & -0.01 & -0.02 & 0 \\ 0 & 0 & 0 & -0.42 & -0.76 & 0 \\ 0 & 0 & 0 & -0.24 & 0.32 & 0 \\ 0 & 0 & 0 & 0 & 0.18 & -1.00 \end{bmatrix} \tag{6.70}$$

From the interaction matrices for both the image data sets it is observed that they are sparse and block triangular in nature. The nonzero terms are present around the diagonal and produce a satisfactory decoupling behavior with a lower condition number. The condition number of the interaction matrix for image "set1.jpg" is 2.97 whereas, for image "Set2.jpg" it is found to be 2.94. Consequently, as compared to the previous sections, the present control law utilizing feature vectors using ZMIs produces the most stable and robust control system.

The servoing is achieved using the desired value of the interaction matrix $I^{"}_{ZM_{nl}}$ described in Eq. 6.69 and Eq. 6.70 respectively and result in the convergence of the features errors as in Figure 6.14 (c) and in Figure 6.15 (c) for the two set of experiments. The plots for camera velocities are described in Figure 6.14 (d) and in Figure 6.15 (d) for the two sets of test images. It is observed in the plot of visual features errors and the camera velocity that there is no oscillation in the camera during the period of convergence and converges at a faster rate. The control law attains excellent stability and achieves enhanced exponential decoupling behavior due to the lower condition numbers of the interaction matrices.

Figure 6.16 (a) and Figure 6.16 (b) represent the plot of camera trajectory for the images "set1.jpg" and "set2.jpg" respectively. From the plot, it is observed that irrespective of the presence of different transformation between the initial and desired position of both the images, the camera trajectories are approximately straight line.

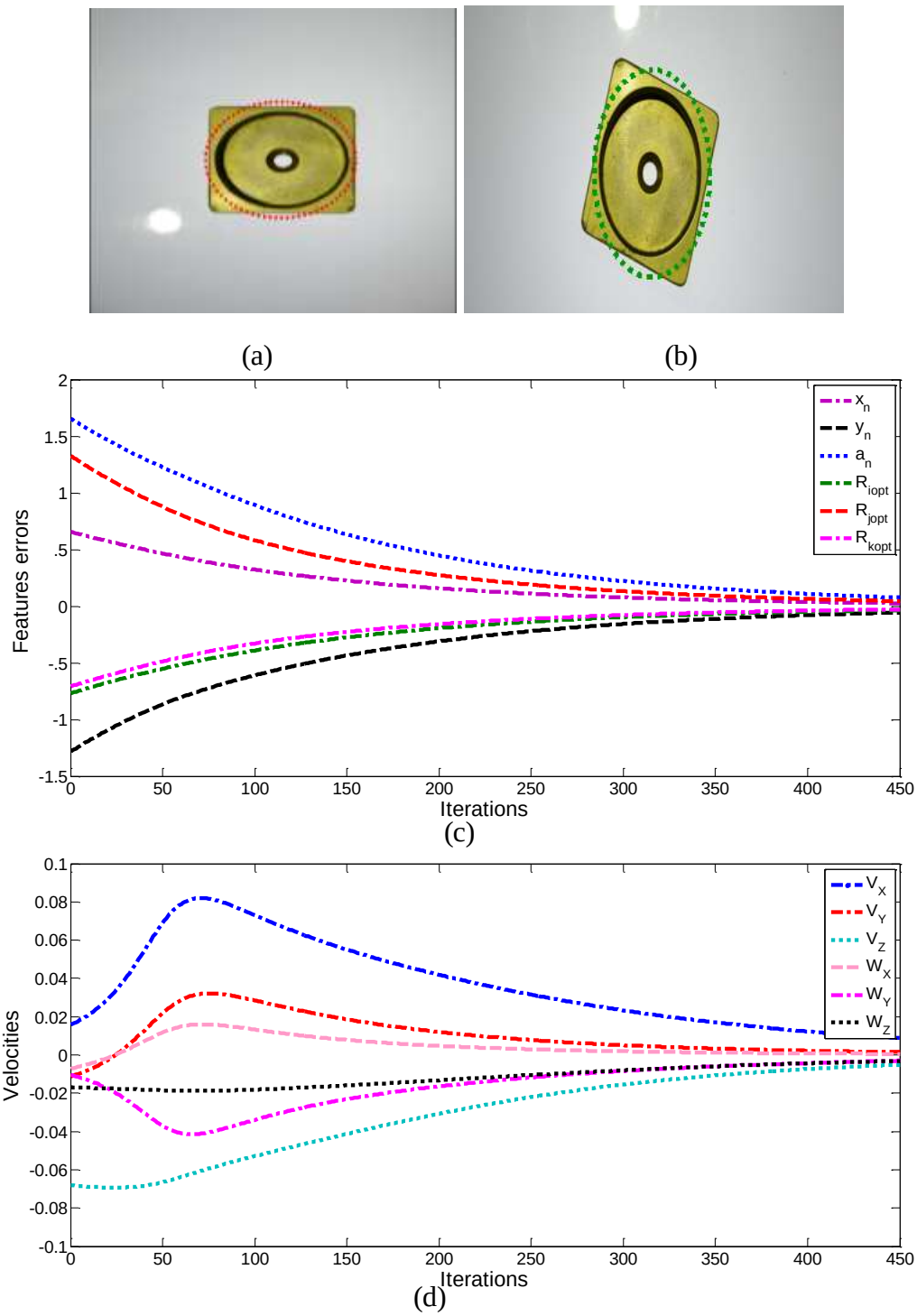
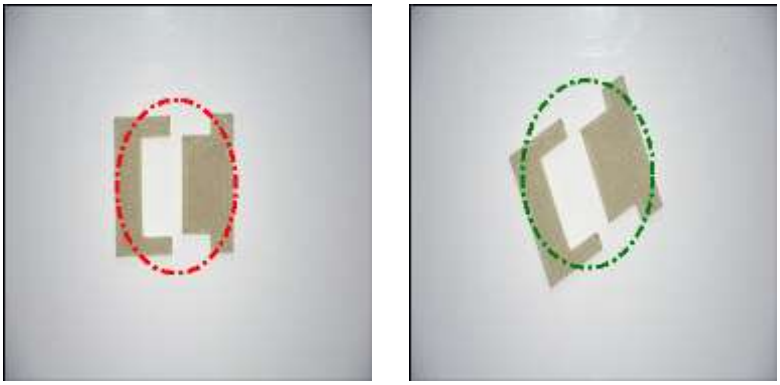
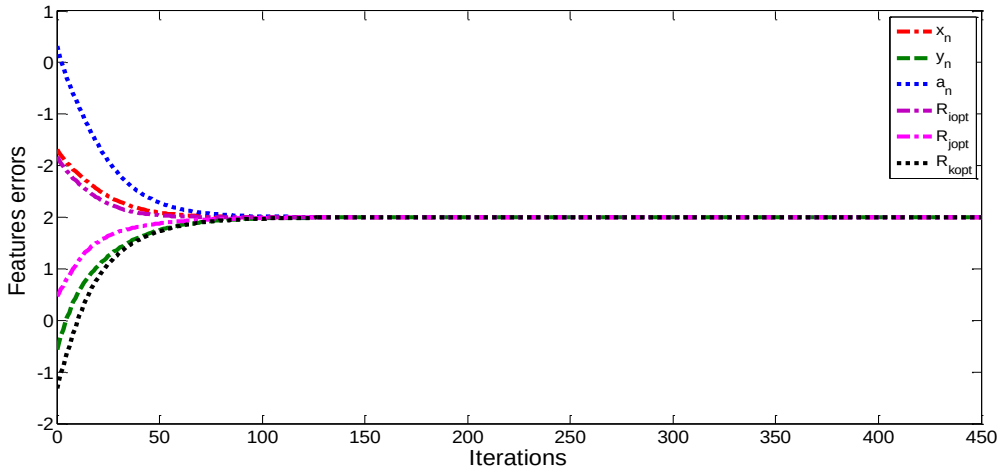


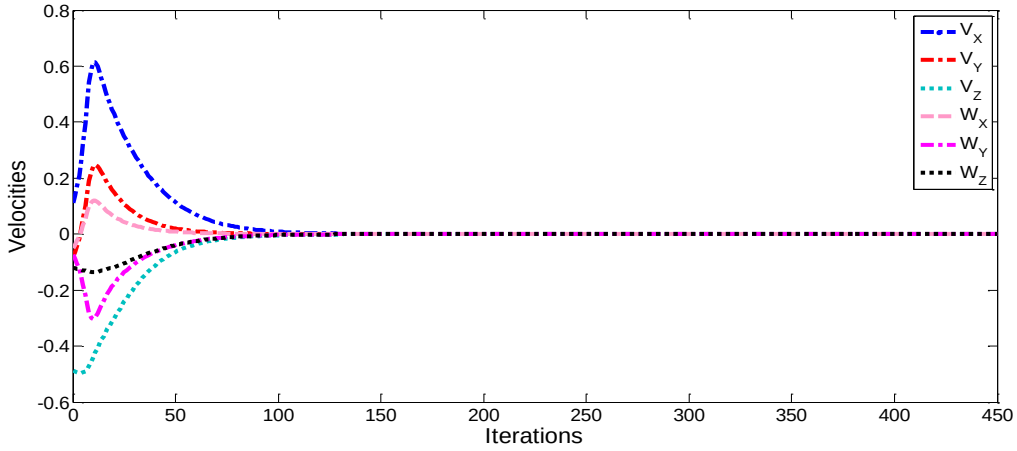
Figure 6.14: "Set 1.jpg" (a) Initial image (b) Desired image (c) Convergence of the features error in the proposed control scheme for desired image (d) Camera velocity in the proposed control scheme for desired image.



(a) (b)

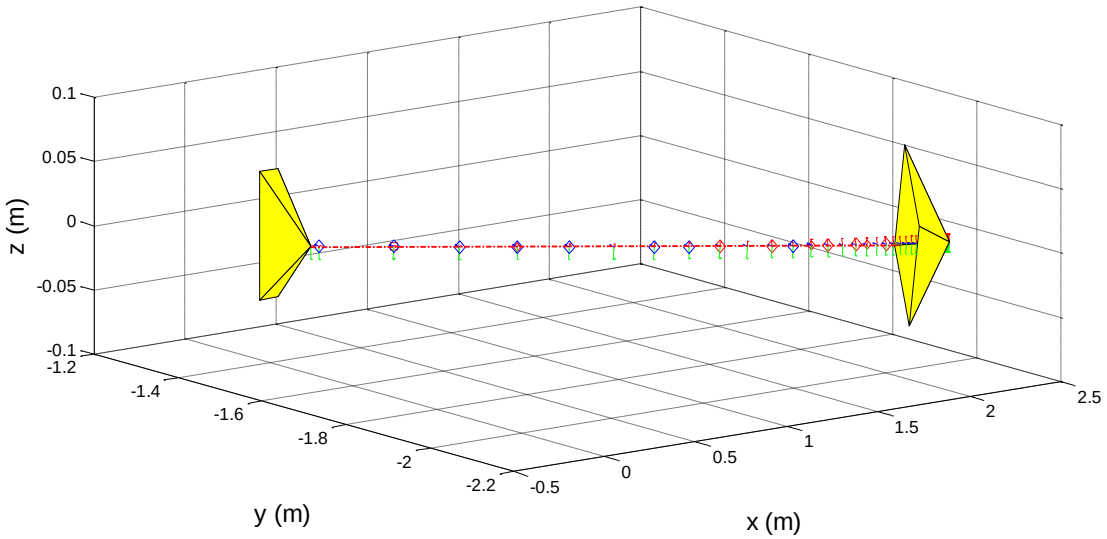


(c)

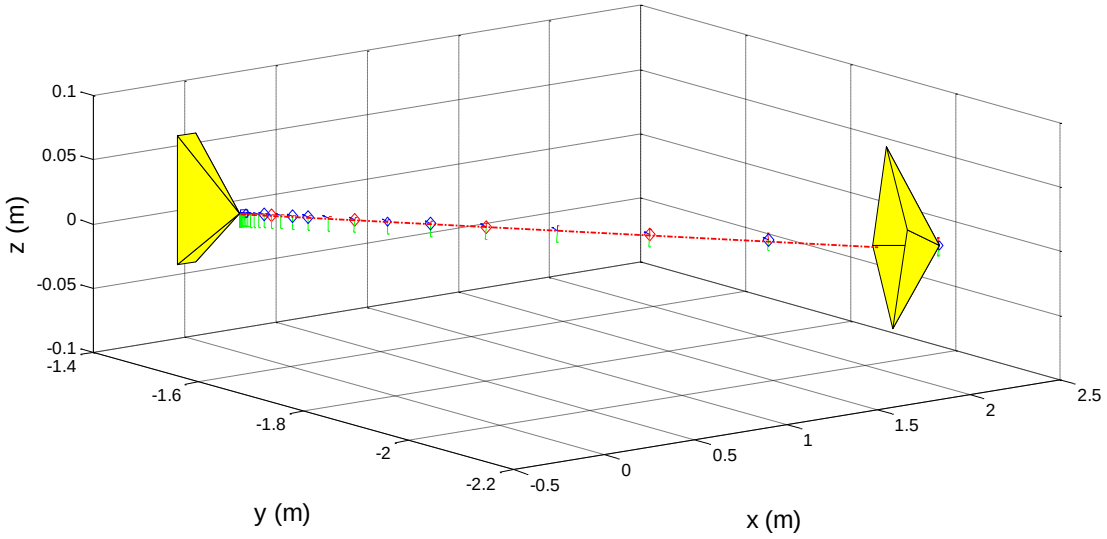


(d)

Figure 6.15: "Set 2.jpg" (a) Initial image (b) Desired image (c) Convergence of the features error in the proposed control scheme for desired image (d) Camera velocity in the proposed control scheme for desired image.



(a)



(b)

Figure 6.16: Camera trajectory for the proposed control scheme. (a) for the desired image of "Set1.jpg" (b) for the desired image of "Set2.jpg"

6.7 Summary

This chapter deals with different approaches for visual servo control i.e., IBVS, PBVS and HVS. In this work, the IBVS control scheme is used as it is having more advantages than other. The efficiency of visual servoing approach depends on the selection of visual features of object in an image. The extraction procedure of such a useful visual feature i.e., moments as a global feature is discussed. The mathematical derivation of such feature is described in detail. Their invariant behavior is presented through proper derivation. Three innovative ideas are proposed and used. Firstly, an indirect method of computation of visual features R_i and R_j for controlling rotational DoF is proposed. In general, it is selected randomly. But, in the present method, these are selected properly which makes the servoing approach efficient. An ellipse is used to calculate the visual feature R_θ makes the process less computational complex. Secondly, the Legendre moment is used to design the control law. An indirect method is followed by which the LMI are expressed in terms of geometric moment. Finally, the visual features are determined by an indirect approach to Zernike moment. These approaches outperform the other approaches.

Chapter 7

Conclusion and Future Scope

7.1 Overview

The robotic part assembly system with the aid of vision system is developed under this work. In the development of the proposed system, factors affecting such system in current scenarios and research outcomes by several researchers are considered. The important issues in manufacturing industries are precision and quality. Now-a-days, robots are used to perform wide range of tasks including pick-and-place tasks, welding, assembly, fault detection, and spray painting etc. Overall, these tasks are similar in terms of identification of parts involved and inspection of the end product ensuring a certain accuracy and repeatability with respect to a certain position or object. Common practice is the use of a specially designed system. The challenges in designing such system is to find out the position and dimensions of object features, 3D location and orientation of objects. These issues are attracting many researchers to contribute towards the advancement of such systems.

7.2 Contribution

The development of the proposed work is accomplished by devising new methods for features extraction, part recognition and robot navigation. In this work, the proposed methods are applied to five different shaped parts. The performance of the proposed methods are examined. The performance is found to be better than the available methods in terms of precision and integration view point. The contribution in the development of the proposed system can be summarized as follows.

- An edge detection technique is developed by considering the positive aspects of both wavelet transformation and fuzzy rule. A quantitative analysis is performed for performance measurement of the proposed method. It is found that this method of edge detection has superior performance as compared to Canny, Sobel, Prewitt, LoG, Robert, wavelet based and mathematical morphology based edge detection techniques.
- A combines orthogonal Zernike moment based object detection method is developed. Such moment is considered as the noise sensitivity is better than other moment invariants methods. The other reason of employing this moment is that this moment outperforms in terms of image representation capability and data redundancy when compared with other moments. The extracted features of the objects are invariant to scale, rotation and transformation. For validation of the proposed method, some images are synthetically generated by adding some form of degradation like transformation, scaling and rotation. It

is found that the combine Zernike moment invariant method extract the features of the images are same as that of the original image. Therefore, the performance of the proposed method is tested and validated properly.

- Nearest Neighbor classifier is implemented for recognizing the objects that undergoes some degradations like scale, rotation and transformations in the assembly floor. The accuracy of the Nearest Neighbor classifier is verified with moments like Geometric moment, Legendre moment, Zernike moment and combine Zernike moment. The accuracy of the classifier with the proposed combine Zernike moment is high in recognition than others.
- In assembly environment, the major concern lies on the robot navigation during part detection and pick up. The IBVS control scheme is considered in this work as it is advantageous than PBVS and HVS schemes. A visual feedback based robot navigation approach is used. There are three mathematical model of visual servoing is proposed and more specifically, the method of determining the image features that controls the translational and rotational DoF are proposed. These features will produce the interaction matrix for robot control. They are as follows.
    i. The visual servoing scheme is developed by using general moment invariants method. In the process, selection of unique combination of features in feature vector that control the rotational DoF i.e., R_{iopt} and R_{jopt} is done as in Algorithm 6.1. This results in a sparse interaction matrix with low condition number. As a novel approach am ellipse is used as reference frame which gives the information about the orientation and position of the object. This results in determining the feature $\xi_{a,b,\theta}$ that controls the rotation DoF along Z-axis.
    ii. An indirect method of determining the feature vector $C = \left[x_n \; y_n \; a_n \; R_{iopt} \; \mathrm{R}_{jopt} \; \mathrm{R}_{\xi} \right]$ is used by using Legendre moment. The Legendre moment invariants are expressed in a linear combination of geometric moment invariants which are invariant to SRT
    iii. An indirect method of designing image based visual feedback control system for robot navigation is used. The developed control scheme is robust to noise. It also provides a better image representation capability.

7.3 Conclusion

This work presents a novel approach towards the design of robotic part assembly system with the aid of vision system. The visual guidance is a rapidly developing approach to the control of robot manipulators, which is based on the visual perception of a robot. The information collected through the vision sensor is used for part recognition, part grasping, part orientation and visual servoing etc.

Chapter 1 includes a brief description of industrial vision system for robotic assembly system and part assembly system in particular. This chapter also includes the motivation and broad objective for the proposed work. In chapter 2, an extensive survey of the available literature is presented to understand the issues in the proposed area. In the survey, feature extraction, part detection and robot navigation based on visual servoing controls are considered. Chapter 3 presents the detailed system requirement for the development of vision guided robotic part assembly system. The purpose of the hardware as well as software components are discussed in details. The methods required for setting up an efficient vision based part detection system ought to be developed before developing the system. In chapter 4, feature extraction techniques for object detection are proposed and a comparative study on performance of various techniques (both proposed and available) is presented. Chapter 5 includes the process of proposed system development. In this chapter, the installation procedure of machine vision system, robotic part assembly system and their integration is discussed in details. The success of integration is tested by a proposed algorithm for the robot navigation autonomously in the specific workspace. Visual feedback based robot navigation in the desired environment is proposed in chapter 6. This chapter includes a mathematical model for the feature vectors which is used for the visual servoing and control of robot.

7.4 Future Scope

The important issues in vision guided robotic part assembly system are object detection and robot navigation. Considering object detection, feature extraction techniques need improvement. The edge features of objects in an image can be further improved by considering several other optimization techniques. Scope for the future work can be summarized for edge detection as

- The proposed edge detector can be further extended to test its performance on images corrupted with Speckle or Poisson noise.
- The proposed edge detector can be extended to cover the edge detection of colored images.

- The proposed edge detector can be extended to three dimensional images.
- The proposed edge detector can be implemented on FPGA for real time application.

The proposed moment based object detection gives out invariant features which are invariant to scale, rotation and transformation. Further improvement can be possible for shape recognition in order to achieve better outcomes from object detection system. This can be achieved by using more improved algorithms. This will help in dealing more complex objects. The extraction of more 3D points or characteristics will result in a more robust and efficient system.

The developed control schemes for robot navigation outperforms than other techniques. However, further improvement can be achieved in regards to the following issues.

- The main issue lies in determining the proper combination of visual features. This selection plays an important role in making the visual servoing technique an efficient one. The performance of the proposed procedure for finding the combination of visual features is found to be better. The procedure of selection of visual features can be improved with respect to computational complexity. Improvement can be done to reduce the presence of information redundancies.
- The formulation of interaction matrix can be improved by utilizing more appropriate optimization techniques.

References

Abdou, I.E. and Pratt, W.K., 1979. Quantitative design and evaluation of enhancement/thresholding edge detectors. Proceedings of the IEEE, 67(5), pp.753-763.

Agin, G.J., 1979. Real time control of a robot with a mobile camera. SRI International.

Agin, G.J., 1980. Computer vision systems for industrial inspection and assembly. Computer, 13(5), pp.11-20.

Al-Ghaili, A.M., Mashohor, S., Ramli, A.R. and Ismail, A., 2013. Vertical-edge-based car-license-plate detection method. IEEE transactions on vehicular technology, 62(1), pp. 26-38.

Ali, M. and Clausi, D., 2001. Using the Canny edge detector for feature extraction and enhancement of remote sensing images. IEEE 2001 International in Geoscience and Remote Sensing Symposium, 2001. IGARSS'01, Vol. 5, pp. 2298-2300.

Alonso, M.T., López-Martínez, C., Mallorquí, J.J. and Salembier, P., 2011. Edge enhancement algorithm based on the wavelet transform for automatic edge detection in SAR images. IEEE transactions on geoscience and remote sensing, 49(1), pp.222-235.

Aybar, E., 2006. Sobel edge detection method for matlab. Anadolu University, Porsuk Vocational School, 26410.

Bakthavatchalam, M., Tahri, O. and Chaumette, F., 2014. May. Improving moments-based visual servoing with tunable visual features. In 2014 IEEE International Conference on Robotics and Automation (ICRA), pp. 6186-6191.

Balabantaray, B.K., Das, B. and Biswal, B.B., 2014. Comparison of Edge Detection Algorithm for Part Identification in a Vision Guided Robotic Assembly System. In Soft Computing Techniques in Engineering Applications, Springer International Publishing, pp. 183-206.

Bao, P., Zhang, L. and Wu, X., 2005. Canny edge detection enhancement by scale multiplication. IEEE transactions on pattern analysis and machine intelligence, 27(9), pp.1485-1490.

Batchelor, B.G., 2012. Machine vision for industrial applications. In Machine Vision Handbook, Springer London, pp. 1-59.

Benjamim, X.C., Gomes, R.B., Burlamaqui, A.F. and Gonçalves, L.M.G., 2012. July. Visual identification of medicine boxes using features matching. In 2012 IEEE International Conference on Virtual Environments Human-Computer Interfaces and Measurement Systems (VECIMS), Proceedings, pp. 43-47.

Bhadauria, H.S., Singh, A. and Kumar, A., 2013. Comparison between various edge detection methods on satellite image. International Journal of Emerging Technology and Advanced Engineering, 3(6), pp.324-28.

Boaventura, A.G., 2009. Method to evaluate the performance of edge detector.

Boggs, D.W., Dungan, J.H., Paek, G.I. and Sato, D.A., Intel Corporation, 2004. PCB design and method for providing vented blind vias. U.S. Patent 6,787,443.

Bourquardez, O., Mahony, R., Guenard, N., Chaumette, F., Hamel, T. and Eck, L., 2009. Image-based visual servo control of the translation kinematics of a quadrotor aerial vehicle. IEEE Trans. on Robotics, 25(3), pp.743-749.

Brannock, E. and Weeks, M., 2008. April. In IEEE Southeast Con: A synopsis of recentwork in edge detection using the DWT. pp. 515-520.

Canny, J., 1986. A computational approach to edge detection. IEEE Transactions on pattern analysis and machine intelligence, (6), pp.679-698.

Chang, W.C. and Chu, P.R., 2010. December. An intelligent space for mobile robot navigation with on-line calibrated vision sensors. In 2010 11th International Conference on Control Automation Robotics & Vision (ICARCV), pp. 1452-1457.

Chauhan, V., Sheth, S., Hindocha, B.R., Shah, R., Dudhat, P. and Jani, P., 2011. Design and development of a machine vision system for part color detection and sorting. In Proceedings of Second International Conference on Signals, Systems & Automation (ICSSA), pp. 90-93.

Chaumette, F. and Hutchinson, S., 2006. Visual servo control. I. Basic approaches. IEEE Robotics & Automation Magazine, 13(4), pp.82-90.

Chaumette, F. and Hutchinson, S., 2007. Visual servo control, Part II: Advanced approaches. IEEE Robotics and Automation Magazine, 14(1), pp.109-118.

Chaumette, F., 2002. A first step toward visual servoing using image moments. In IEEE/RSJ International Conference on Intelligent Robots and Systems, 2002, (1), pp. 378-383.

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

Chawla, P.R. and Deb, S., 2012. A Computer-Aided Inspection Methodology for Mechanical Parts based on Machine Vision.

Chen, B., Shu, H., Zhang, H., Coatrieux, G., Luo, L. and Coatrieux, J.L., 2011. Combined invariants to similarity transformation and to blur using orthogonal Zernike moments. IEEE Transactions on Image Processing, 20(2), pp.345-360

Chong, C.W., Raveendran, P. and Mukundan, R., 2004. Translation and scale invariants of Legendre moments. Pattern recognition, 37(1), pp.119-129.

Coatrieux, J.L., 2008. Moment-based approaches in imaging part 2: invariance [A Look at...]. IEEE Engineering in Medicine and Biology Magazine, 27(1), pp.81-83.

Coleman, S.A., Scotney, B.W. and Herron, M.G., 2003, September. An empirical performance evaluation technique for discrete second derivative edge detectors. In Proceedings 12th International Conference on Image Analysis and Processing, 2003, pp. 594-599.

Collewet, C. and Marchand, E., 2011. Photometric visual servoing. IEEE Transactions on Robotics, 27(4), pp.828-834.

Corke, P.I. and Hutchinson, S.A., 2000. A new hybrid image-based visual servo control scheme. In Proceedings of the 39th IEEE Conference on Decision and Control, 2000. (3), pp. 2521-2526.

Corke, P.I. and Hutchinson, S.A., 2001. A new partitioned approach to image-based visual servo control. IEEE Transactions on Robotics and Automation, 17(4), pp.507-515.

Cubero, S., Aleixos, N., Moltó, E., Gómez-Sanchis, J. and Blasco, J., 2011. Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables. Food and Bioprocess Technology, 4(4), pp.487-504.

Dahiya, V., Trivedi, J. and Sajja, P.S., 2013, March. Innovative algorithms for vision defect identification system. In 2013 International Conference on Intelligent Systems and Signal Processing (ISSP), pp. 209-215.

Darabkh, K.A., Al-Zubi, R.T., Jaludi, M.T. and Al-Kurdi, H., 2014, February. An efficient method for feature extraction of human iris patterns. In 2014 11th International Multi-Conference on Systems, Signals & Devices (SSD), pp. 1-5.

Dario, P., Guglielmelli, E., Allotta, B. and Carrozza, M.C., 1996. Robotics for medical applications. IEEE Robotics & Automation Magazine, 3(3), pp.44-56.

Davies, E.R., 2012. Computer and machine vision: theory, algorithms, practicalities. Academic Press.

Dawei, Q., Peng, Z., Xuefei, Z. and Xuejing, J., 2010. Edge detection of wood defects in X-ray wood image using neural network and mathematical morphology. In Proceedings of the 29th Chinese Control Conference.

Diskin, Y., Nair, B., Braun, A., Duning, S. and Asari, V.K., 2013, October. Vision-based navigation system for obstacle avoidance in complex environments. In 2013 IEEE Applied Imagery Pattern Recognition Workshop (AIPR) (pp. 1-8). IEEE.

Djelal, N., Mechat, N. and Nadia, S., 2011, March. Target tracking by visual servoing. In Systems, Signals and Devices (SSD), 2011 8th International Multi-Conference on (pp. 1-6). IEEE.

Dong, Y., Li, M. and Li, J., 2013, December. Image retrieval based on improved Canny edge detection algorithm. In Mechatronic Sciences, Electric Engineering and Computer (MEC), Proceedings 2013 International Conference on (pp. 1453-1457).

Du, W., Tian, X. and Sun, Y., 2012, October. A dynamic threshold segmentation algorithm for anterior chamber OCT images based on wavelet transform. In 2012 5th International Congress on Image and Signal Processing (CISP), pp. 279-282.

Espiau, B., 1994. Effect of camera calibration errors on visual servoing in robotics. In Experimental robotics III, Springer Berlin Heidelberg, pp. 182-192.

Espiau, B., Chaumette, F. and Rives, P., 1992. A new approach to visual servoing in robotics. IEEE Transactions on Robotics and Automation, 8(3), pp.313-326.

Fareh, R., Payeur, P., Nakhaeinia, D., Macknojia, R., Chávez-Aragón, A., Cretu, A.M., Laferrière, P., Laganière, R. and Toledo, R., 2014, March. An integrated vision-guided robotic system for rapid vehicle inspection. In 2014 8th Annual IEEE Systems Conference (SysCon), pp. 446-451.

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

Fernandes, A.O., Moreira, L.F. and Mata, J.M., 2011, December. Machine vision applications and development aspects. In 2011 9th IEEE International Conference on Control and Automation (ICCA), pp. 1274-1278.

Flusser, J. and Suk, T., 1994. A moment-based approach to registration of images with affine geometric distortion. IEEE transactions on Geoscience and remote sensing, 32(2), pp.382-387.

Flusser, J., 2006, February. Moment invariants in image analysis. In proceedings of world academy of science, engineering and technology, 11(2), pp. 196-201.

Flusser, J., Suk, T. and Zitová, B., 2009. Introduction to Moments. Moments and Moment Invariants in Pattern Recognition, pp.1-11.

Fujimura, K., Konoma T., Kamijo S., 2011. Vehicle infrastructure integration system using vision sensors to prevent accidents in traffic flow. IET Intelligent Transport System, 2011, Vol. 5, Iss. 1, pp. 11–20.

Gao, X. and Jin, L., 2009, October. SwiftPost: A Vision-based Fast Postal Envelope Identification System. In SMC, pp. 3268-3273.

Gascón, R. and Barraza, M., 2012, October. Six DOF Stereoscopic Eye-in-Hand Visual Servo System BIBOT. In 2012 Brazilian Robotics Symposium and Latin American Robotics Symposium (SBR-LARS), pp. 284-289.

Gauthier, M. and Régnier, S., 2011. Robotic micro-assembly. John Wiley & Sons.

Gomes, R.B., Gonalves, L.M.G. and de Carvalho, B.M., 2008, May. Real time vision for robotics using a moving fovea approach with multi resolution. In IEEE International Conference on Robotics and Automation (ICRA), pp. 2404-2409.

Gonzalez, R.C. and Woods, R.E., 2007. Image processing. Digital image processing, 2.

Govindarajan, B., Panetta, K.A. and Agaian, S., 2008, October. Image reconstruction for quality assessment of edge detectors. In IEEE International Conference on Systems, Man and Cybernetics, 2008. SMC 2008, pp. 691-696.

Gray, J.O., Davis, S.T. and Caldwell, D.G., 2013. Robotics in the food industry: an introduction. Robotics and automation in the food industry: current and future technologies, pp.21-35.

Green, B., 2002. Canny edge detection tutorial. DOI= http://www. Pages. drexel. edu/~ weg22/can_tut. html.

Gridseth, M., Hertkorn, K. and Jagersand, M., 2015, June. On visual servoing to improve performance of robotic grasping. In 2015 12th Conference on Computer and Robot Vision (CRV), pp. 245-252.

Guan, Y.P., 2008. Automatic extraction of lips based on multi-scale wavelet edge detection. IET Computer Vision, 2(1), pp.23-33.

Hagiwara, Y., Choi, Y. and Watanabe, K., 2009, February. An improved view-based navigation by adjustable position resolution. In 4th International Conference on Autonomous Robots and Agents, 2009 (ICARA 2009). pp. 298-301.

Hamilton, R.N., 2003. Pharmacy pill counting vision system. U.S. Patent 6,574,580.

Harris, C., Stephens, M., 1988, A combined corner and edge detector. Proc. 4th Alvey Vision Conference, pp. 147-151.

Hasan, K.M. and Al Mamun, A., 2012, May. Implementation of vision based object tracking robot. In 2012 International Conference on Informatics, Electronics & Vision (ICIEV), pp. 860-864.

Hashimoto, K., Kimoto, T., Ebine, T. and Kimura, H., 1991, April. Manipulator control with image-based visual servo. In IEEE International Conference on Robotics and Automation, Proceedings. pp. 2267-2271.

Heikkila, J. and Silvén, O., 1997, June. A four-step camera calibration procedure with implicit image correction. In 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Proceedings. pp. 1106-1112.

Herakovic, N., 2010. Robot vision in industrial assembly and quality control processes. INTECH Open Access Publisher.

Hoffmann, F., Nierobisch, T., Seyffarth, T. and Rudolph, G., 2006, October. Visual servoing with moments of SIFT features. In 2006 IEEE International Conference on Systems, Man and Cybernetics. 5, pp. 4262-4267.

Hosaini, S.J., Alirezaee, S., Ahmadi, M. and Makki, S.V.A.D., 2013, September. Comparison of the Legendre, Zernike and Pseudo-Zernike moments for feature extraction in iris recognition. In 2013 5th International Conference on Computational Intelligence and Communication Networks (CICN). pp. 225-228.

Hosny, K.M., 2005. New set of rotationally Legendre moment invariants. Editorial Advisory Board e, p.177.

Hosny, K.M., 2010. Robust template matching using orthogonal Legendre moment invariants. Journal of computer science, 6(10), pp.1083.

Hrach, D., Brandner, M., Fossati, P. and Marta, S., 2005, September. Intelligent vision-sensor for robot-sensing applications. In International Workshop on Robotic Sensors: Robotic and Sensor Environments, 2005. pp. 37-42.

Hu, M.K., 1962. Visual pattern recognition by moment invariants. IRE transactions on information theory, 8(2), pp.179-187.

Huang, G.S. and Cheng, C.E., 2013, December. 3D coordinate identification of object using binocular vision system for mobile robot. In 2013 International Automatic Control Conference (CACS). pp. 91-96.

Hutchinson, S., Hager, G.D. and Corke, P.I., 1996. A tutorial on visual servo control. IEEE transactions on robotics and automation, 12(5), pp.651-670.

Huwedi, A., Steinhaus, P. and Dillmann, R., 2006, September. Autonomous feature-based exploration using multi-sensors. In 2006 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems. pp. 456-461.

Jain, R., Kasturi, R. and Schunck, B.G., 1995. Machine vision (Vol. 5). New York: McGraw-Hill.

Jerbić, B., Vranješ, B., Hrman, M. and Kunica, Z., 2005. Intelligent robotic assembly by machine vision and CAD integration. Transactions of FAMENA, 29(1), pp.17-30.

Jestin, V.K., Anitha, J. and Hemanth, D.J., 2012, March. Textural feature extraction for retinal image analysis. In 2012 International Conference on Computing, Electronics and Electrical Technologies (ICCEET). pp. 548-551.

Jia, Y., Duan, Y., Wang, D., Xue, L., Liu, Z. and Wang, W., 2010, October. Pieces Identification in the Chess System of Dual-Robot Coordination Based on Vision. In 2010 International Conference on Web Information Systems and Mining (WISM). 2, pp. 248-251.

Jiang, G.Q. and Zhao, C.J., 2012, July. Apple recognition based on machine vision. In 2012 International Conference on Machine Learning and Cybernetics. 3, pp. 1148-1151.

Jihong, Z., Jun, L. and Xianqing, L., 2012, December. An adaptive fuzzy entropy algorithm in image edge detection. In 2012 Second International Conference on Instrumentation, Measurement, Computer, Communication and Control (IMCCC). pp. 371-374.

Jin, T.S., Ko, J.P. and Lee, J.M., 2003, July. A new space and time sensor fusion for vision-based navigation by mobile robot system. In 2003 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, 2003. AIM 2003. 1, pp. 40-45.

Juneja, M. and Sandhu, P.S., 2009. Performance evaluation of edge detection techniques for images in spatial domain. International journal of computer theory and Engineering, 1(5), p.614.

Jusoh, R.M., 2006, June. Application of vision target localization for mobile robot. In 2006 4th Student Conference on Research and Development. pp. 144-146.

Karagiannis, D. and Astolfi, A., 2005. A new solution to the problem of range identification in perspective vision systems. IEEE Transactions on Automatic Control, 50(12), pp.2074-2077.

Katyal, K.D., Johannes, M.S., McGee, T.G., Harris, A.J., Armiger, R.S., Firpi, A.H., McMullen, D., Hotson, G., Fifer, M.S., Crone, N.E. and Vogelstein, R.J., 2013, November. HARMONIE: A multimodal control framework for human assistive robotics. In 2013 6th International IEEE/EMBS Conference on Neural Engineering (NER). pp. 1274-1278.

Kaur, B. and Garg, A., 2011, April. Mathematical morphological edge detection for remote sensing images. In 2011 3rd International Conference on Electronics Computer Technology (ICECT), 5, pp. 324-327.

Kellett, P. (2009). Roadmap to the Future, www.robotics.org, August 2009.

Keyes, L. and Winstanley, A., 2001. Using moment invariants for classifying shapes on large-scale maps. Computers, Environment and Urban Systems, 25 (1), pp.119-130.

Khotanzad, A. and Hong, Y.H., 1990. Invariant image recognition by Zernike moments. IEEE Transactions on pattern analysis and machine intelligence, 12(5), pp.489-497.

Kim, H.J., Kim, M., Lim, H., Park, C., Yoon, S., Lee, D., Choi, H., Oh, G., Park, J. and Kim, Y., 2013. Fully autonomous vision-based net-recovery landing system for a fixed-wing UAV. IEEE/ASME Transactions on Mechatronics, 18(4), pp.1320-1333.

Kim, J.Y., Kang, D.J. and Cho, H.S., 2001. A flexible parts assembly algorithm based on a visual sensing system. In Proceedings of the IEEE International Symposium on Assembly and Task Planning. pp. 417-422.

Kim, Y. and Hwang, D.H., 2013, October. Design of vision/INS integrated navigation system in poor vision navigation environments. 2013 13th International Conference on Control, Automation and Systems (ICCAS). pp. 531-535.

Kitney, R.I. and Smith, N., Kitney Richard I, 2005. Image segmentation method. U.S. Patent 6,839,462.

Koenig, T., Dong, Y. and DeSouza, G.N., 2008, September. Image-based visual servoing of a real robot using a quaternion formulation. In 2008 IEEE Conference on Robotics, Automation and Mechatronics. pp. 216-221.

Kress, S., 2004. Machine Vision Makes Its Mark on the Automotive Industry. Automotive Design and Production.

Kruse, D., Radke, R.J. and Wen, J.T., 2013, August. A sensor-based dual-arm tele-robotic manipulation platform. In 2013 IEEE International Conference on Automation Science and Engineering (CASE). pp. 350-355.

Labudzki, R. and Legutko, S., 2011. Applications of machine vision. Manufact Ind Eng., 2, pp.27-29.

Lai, W.M. and Lin, C.Y., 2013, October. Autonomous cross-floor navigation of a stair-climbing mobile robot using wireless and vision sensor. In 2013 44th International Symposium on Robotics (ISR). pp. 1-6.

Lee, S., Kim, J., Lee, M., Yoo, K., Barajas, L.G. and Menassa, R., 2012, August. 3d visual perception system for bin picking in automotive sub-assembly automation. In 2012 IEEE international conference on automation science and engineering (CASE). pp. 706-713

Li, H. and Yang, S.X., 2003. A behavior-based mobile robot with a visual landmark-recognition system. IEEE/ASME transactions on mechatronics. 8(3), pp.390-400.

Li, J. (2003). A Wavelet Approach to Edge Detection. PhD diss., Sam Houston State University.

Li, M., Imou, K. and Wakabayashi, K., 2010, May. 3D positioning for mobile robot using omnidirectional vision. In 2010 International Conference on Intelligent Computation Technology and Automation (ICICTA). 1, pp. 7-11.

Li, Y., Hu, Q., Wu, M. and Gao, Y., 2016. An Imaging Sensor-Aided Vision Navigation Approach that Uses a Geo-Referenced Image Database. Sensors. 16(2), p.166.

Lin, C.C. and Chellappa, R., 1987. Classification of partial 2-D shapes using Fourier descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence. (5), pp.686-690.

Lin, Y.C., Lin, C.T., Liu, W.C. and Chen, L.T., 2013, June. A vision-based obstacle detection system for parking assistance. In 2013 IEEE 8th Conference on Industrial Electronics and Applications (ICIEA). pp. 1627-1630.

Lindley, C.A., 1991. Practical image processing in C: acquisition, manipulation, storage. John Wiley & Sons, Inc..

Liu, G., Li, S. and Liu, W., 2013, November. Lane detection algorithm based on local feature extraction. In 2013 Chinese Automation Congress (CAC). pp. 59-64.

Liu, Y, Hou, M., Rao, X., Zhang, Y.: A Steady Corner Detection of Gray Level Images Based

Lixia, J., Wenjun, Z. and Yu, W., 2010, February. Study on improved algorithm for image edge detection. In 2010 the 2nd International Conference on Computer and Automation Engineering (ICCAE). 4, pp. 476-479.

Mahony, R. and Hamel, T., 2005. Image-based visual servo control of aerial robotic systems using linear image features. IEEE Transactions on Robotics. 21(2), pp.227-239.

Maini, R. and Aggarwal, H., 2009. Study and comparison of various image edge detection techniques. International journal of image processing (IJIP). 3(1), pp.1-11.

Malamas, E.N., Petrakis, E.G., Zervakis, M., Petit, L. and Legat, J.D., 2003. A survey on industrial vision systems, applications and tools. Image and vision computing. 21(2), pp.171-188.

Malis, E. and Chaumette, F., 2000. 2 1/2 d visual servoing with respect to unknown objects through a new estimation scheme of camera displacement. International Journal of Computer Vision. 37(1), pp.79-97.

Malis, E., Chaumette, F. and Boudet, S., 1999. 2½D visual servoing. IEEE Transactions on Robotics and Automation. 15(2), pp.238-250.

Mallat, S. and Hwang, W.L., 1992. Singularity detection and processing with wavelets. IEEE transactions on information theory. 38(2), pp.617-643.

Mamistvalov, A.G., 1998. N-Dimensional moment invariants and conceptual mathematical theory of recognition n-dimensional solids. IEEE Transactions on pattern analysis and machine intelligence. 20(8), pp.819-831.

Martinet, P., Gallice, J. and Khadraoui, D., 1996, May. Vision based control law using 3d visual features. In Proc. WAC. 96, pp. 497-502.

Mebarki, R., Krupa, A. and Chaumette, F., 2010. 2-d ultrasound probe complete guidance by visual servoing using image moments. IEEE Transactions on Robotics. 26(2), pp.296-306.

Meer, P. and Georgescu, B., 2001. Edge detection with embedded confidence. IEEE Transactions on pattern analysis and machine intelligence. 23(12), pp.1351-1365.

Mercimek, M., Gulez, K. and Mumcu, T.V., 2005. Real object recognition using moment invariants. Sadhana. 30(6), pp.765-775.

Ming-chien, Y. and Peng-fei, S., 1988, August. A knowledge-based vision system for identification overlapping objects. In Proceedings of the 1988 IEEE International Conference on Systems, Man, and Cybernetics. 1, pp. 658-660.

Moghadam, P., Wijesoma, W.S. and Feng, D.J., 2008, December. Improving path planning and mapping based on stereo vision and lidar. In 10th International Conference on Control, Automation, Robotics and Vision. ICARCV 2008. pp. 384-389.

Mokhtarian, F. and Abbasi, S., 2002. Shape similarity retrieval under affine transforms. Pattern Recognition, 35(1), pp.31-41.

Mokhtarian, F. and Suomela, R., 1998. Robust image corner detection through curvature scale space. IEEE Transactions on Pattern Analysis and Machine Intelligence. 20(12), pp.1376-1381.

Mukundan, R. and Ramakrishnan, K.R., 1998. Moment functions in image analysis: theory and applications (Vol. 100). Singapore: World Scientific.

Mundy, J.L. and Zisserman, A., 1992. Geometric invariants in computer vision.

Munich, M.E., Pirjanian, P., Di Bernardo, E., Goncalves, L., Karlsson, N. and Lowe, D., 2006. Application of visual pattern recognition to robotics and automation. IEEE Robotics & Automation Magazine. pp.72-77.

Murai, T. and Morimoto, M., 2013, December. A Visual Inspection System for Prescription Drugs in Press through Package. In 2013 Second International Conference on Robot, Vision and Signal Processing.

Murai, T., Morimoto, M. and Fujii, K., 2012, November. A Visual Inspection System for Prescription Drugs. In 2012 Fifth International Conference on Emerging Trends in Engineering and Technology.

Mustafah, Y.M., Noor, R., Hasbi, H. and Azma, A.W., 2012, July. Stereo vision images processing for real-time object distance and size measurements. In 2012 International Conference on Computer and Communication Engineering (ICCCE). pp. 659-663.

Nagabhushana, S., 2005. Computer vision and image processing. New Age International.

Nayar, S.K., Carnegie-Mellon University, 1990. Robotic vision system. U.S. Patent 4,893,183.

Nello, Z., 2000. Understanding and applying machine vision. NewYork, USA: Library of Congress Cataloging in Publication.

Nelson, R.C. and Selinger, A., 1997, March. Experiments on (Intelligent) Brute Force Methods for Appearance-Based Object Recognition. In DARPA Image Understanding Workshop. pp. 1197-1205.

Neto, A.M., Rittner, L., Zampieri, D.E. and Corrêa-Victorino, A., 2008, September. Nondeterministic Criteria to Discard Redundant Information in Real Time Autonomous Navigation Systems based on Monocular Vision. In 2008 IEEE International Symposium on Intelligent Control. pp. 420-425.

Nguyen, J.S., Su, S.W. and Nguyen, H.T., 2013, July. Experimental study on a smart wheelchair system using a combination of stereoscopic and spherical vision. In 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). pp. 4597-4600.

Oh, J.K., Lee, S. and Lee, C.H., 2012. Stereo vision based automation for a bin-picking solution. International Journal of Control, Automation and Systems. 10(2), pp.362-373.

Ohsaki, M., Sugiyama, T. and Ohno, H., 2000. Evaluation of edge detection methods through psychological tests-is the detected edge really desirable for humans? In 2000 IEEE International Conference on Systems, Man, and Cybernetics. 1, pp. 671-677.

Okada, K., Inaba, M. and Inoue, H., 2003, October. Walking navigation system of humanoid robot using stereo vision based floor recognition and path planning with multi-layered body image. In IEEE/RSJ International Conference on Intelligent Robots and Systems, 2003. (IROS 2003). 3, pp. 2155-2160.

Olivares-Mendez, M.A., Fu, C., Ludivig, P., Bissyandé, T.F., Kannan, S., Zurad, M., Annaiyan, A., Voos, H. and Campoy, P., 2015. Towards an Autonomous Vision-Based Unmanned Aerial System against Wildlife Poachers. Sensors. 15(12), pp.31362-31391.

Owayjan, M., Kashour, A., Al Haddad, N., Fadel, M. and Al Souki, G., 2012, December. The design and development of a lie detection system using facial micro-expressions. In 2012 2nd International Conference on Advances in Computational Tools for Engineering Applications (ACTEA). pp. 33-38.

Page, T., 2005. Design Guidelines and Rules for Effective Yield in Manufacturing & Assembly of Printed Circuit Boards Utilizing Surface Mount Technology", December 2004. Society of Manufacturing Engineers, Dearborn, MI, USA, Technical Paper: TP04PUB358. pp.1-45.

Park, J.M. and Murphey, Y.L., 2008. Edge detection in grayscale, color, and range images. Wiley Encyclopedia of Computer Science and Engineering.

Patel, D.K. and More, S.A., 2013, January. Edge detection technique by fuzzy logic and Cellular Learning Automata using fuzzy image processing. In 2013 International Conference on Computer Communication and Informatics (ICCCI). pp. 1-6.

Pei, S., Ding, J., 2007, Improved Harris' Algorithm for Corner and Edge Detections. In: Proc. ICIP, pp. 57-60.

Pei, S.C. and Lin, C.N., 1995. Image normalization for pattern recognition. Image and Vision computing. 13(10), pp.711-723.

Peña, M., López, I. and Osorio, R., 2006, September. Invariant object recognition robot vision system for assembly. In Electronics, Robotics and Automotive Mechanics Conference (CERMA'06). 1, pp. 30-36.

Pena-Cabrera, M. and Lopez-Juarez, I., 2006. Distributed Architecture for Intelligent Robotic Assembly, Part III: Design of the Invariant Object Recognition System. Manufacturing the Future, Concepts - Technologies - Visions, ISBN 3-86611-198-3. pp. 401-436.

Powell, J.R., Krotosky, S., Ochoa, B., Checkley, D. and Cosman, P., 2003. Detection and identification of sardine eggs at sea using a machine vision system. In Oceans 2003. Celebrating the Past... Teaming Toward the Future (IEEE Cat. No. 03CH37492).

Prokop, R.J. and Reeves, A.P., 1992. A survey of moment-based techniques for unoccluded object representation and recognition. CVGIP: Graphical Models and Image Processing. 54(5), pp.438-460.

Rao, U.S.N., 2013, December. Design of automatic cotton picking robot with Machine vision using Image Processing algorithms. In 2013 International Conference on Control, Automation, Robotics and Embedded Systems (CARE). pp. 1-5.

Resendiz, E., Hart, J.M. and Ahuja, N., 2013. Automated visual inspection of railroad tracks. IEEE Transactions on Intelligent Transportation Systems. 14(2), pp.751-760.

Rousseau, P., Desrochers, A. and Krouglicof, N., 2001. Machine vision system for the automatic identification of robot kinematic parameters. IEEE Transactions on Robotics and Automation. 17(6), pp.972-978.

Russ, J.C. and Woods, R.P., 1995. The image processing handbook. Journal of Computer Assisted Tomography. 19(6), pp.979-981.

Saad, P. and Rusli, N., 2004. Feature Extraction of Trademark Images Using Geometric Invariant Moment and Zernike Moment-A Comparison.

Sanderson, A.C. and Weiss, L.E., 1983. Adaptive visual servo control of robots. In Robot Vision. Springer Berlin Heidelberg. pp. 107-116.

Sansoni, G., Bellandi, P., Leoni, F. and Docchio, F., 2014. Optoranger: A 3D pattern matching method for bin picking applications. Optics and Lasers in Engineering, 54, pp.222-231.

Sardy, S., Ibrahim, L. and Yasuda, Y., 1993, October. An application of vision system for the identification and defect detection on woven fabrics by using artificial neural network. In Proceedings of 1993 International Joint Conference on Neural Networks, 1993. IJCNN'93-Nagoya. 3, pp. 2141-2144.

Segvic, S., Remazeilles, A., Diosi, A. and Chaumette, F., 2007, June. Large scale vision-based navigation without an accurate global reconstruction. In 2007 IEEE Conference on Computer Vision and Pattern Recognition. pp. 1-8.

Senthilkumaran, N. and Rajesh, R., 2009. Edge detection techniques for image segmentation–a survey of soft computing approaches. International journal of recent trends in engineering, 1(2).

Shang, J. and Jiang, F., 2012, October. An algorithm of edge detection based on soft morphology. In 2012 IEEE 11th International Conference on Signal Processing (ICSP). 1, pp. 166-169.

Sharifi, M., Fathy, M. and Mahmoudi, M.T., 2002, April. A classified and comparative study of edge detection algorithms. In International Conference on Information Technology: Coding and Computing, 2002. Proceedings. pp. 117-120.

Shi, Y., Liang, B., Wang, X., Xu, W. and Liu, H., 2012, August. Modeling and simulation of space robot visual servoing for autonomous target capturing. In 2012 IEEE International Conference on Mechatronics and Automation. pp. 2275-2280.

Shih, M.Y. and Tseng, D.C., 2005. A wavelet-based multiresolution edge detection and tracking. Image and Vision Computing. 23(4), pp.441-451.

Shimizu, S., Kato, T., Ocmula, Y. and Suematu, R., 2001. Wide angle vision sensor with fovea-navigation of mobile robot based on cooperation between central vision and peripheral vision. In 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems. 2, pp. 764-771.

Shin, M.C., Goldgof, D.B., Bowyer, K.W. and Nikiforou, S., 2001. Comparison of edge detection algorithms using a structure from motion task. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics). 31(4), pp.589-601.

Shrivakshan, G.T. and Chandrasekar, C., 2012. A comparison of various edge detection techniques used in image processing. IJCSI International Journal of Computer Science Issues. 9(5), pp.272-276.

Shu, H., Luo, L. and Coatrieux, J.L., 2007. Moment-based approaches in imaging. Part 1, basic features. IEEE Engineering in Medicine and Biology Magazine. 26(5), p.70.

Singh, C. and Sharma, P., 2013. Performance analysis of various local and global shape descriptors for image retrieval. Multimedia systems. 19(4), pp.339-357.

Sinha, A. and Chakravarty, K., 2013, October. Pose based person identification using kinect. In 2013 IEEE International Conference on Systems, Man, and Cybernetics. pp. 497-503.

Somkantha, K., Theera-Umpon, N. and Auephanwiriyakul, S., 2011. Boundary detection in medical images using edge following algorithm based on intensity gradient and texture gradient features. IEEE transactions on biomedical engineering. 58(3), pp.567-573.

Sonka, M., Hlavac, V. and Boyle, R., 2014. Image processing, analysis, and machine vision. Cengage Learning.

Sridhar, B. and Reddy, K.V.V.S., 2013, April. Qualitative detection of breast cancer by morphological curvelet transform. In 2013 8th International Conference on Computer Science & Education (ICCSE). pp. 514-517.

Steane, A.M., 1996. Error correcting codes in quantum theory. Physical Review Letters, 77(5), p.793.

Steger, C., Ulrich, M. and Wiedemann, C., 2008. Machine vision algorithms and applications. Illumination. 5(2.1), p.1.

Steiner, T.J. and Brady, T.M., 2014, March. Vision-based navigation and hazard detection for terrestrial rocket approach and landing. In 2014 IEEE Aerospace Conference. pp. 1-8.

Sternberg, S., 1985, March. Vision-guided robot for automotive assembly. In 1985 IEEE International Conference on Robotics and Automation. 2, pp. 686-690.

Tahri, O. and Chaumette, F., 2003. Application of moment invariants to visual servoing. In IEEE Int. Conf. on Robotics and Automation, ICRA'03. 3, pp. 4276-4281.

Tahri, O. and Chaumette, F., 2003. Determination of moment invariants and their application to visual servoing (Doctoral dissertation, INRIA).

Tahri, O. and Chaumette, F., 2005, April. Complex objects pose estimation based on image moment invariants. In Proceedings of the 2005 IEEE International Conference on Robotics and Automation. pp. 436-441.

Tahri, O. and Chaumette, F., 2005. Point-based and region-based image moments for visual servoing of planar objects. IEEE Transactions on Robotics. 21(6), pp.1116-1127.

Tahri, O., Tamtsia, A.Y., Mezouar, Y. and Demonceaux, C., 2015. Visual Servoing Based on Shifted Moments. IEEE Transactions on Robotics. 31(3), pp.798-804.

Tamtsia, A.Y., Mezouar, Y., Martinet, P., Djalo, H. and Tonye, E., 2012. 2D Legendre Moments-based Visual Control. In Applied Mechanics and Materials. Trans Tech Publications. 162, pp. 487-496.

Tamtsia, A.Y., Tahri, O., Mezouar, Y., Djalo, H. and Tonye, E., 2013, May. New results in images moments-based Visual Servoing. In 2013 IEEE International Conference on Robotics and Automation (ICRA). pp. 5271-5276.

Tanoto, A., Li, H., Rückert, U. and Sitte, J., 2012, October. Scalable and flexible vision-based multi-robot tracking system. In 2012 IEEE International Symposium on Intelligent Control. pp. 19-24.

Teague, M.R., 1980. Image analysis via the general theory of moments. JOSA. 70(8), pp.920-930.

Teh, C.H. and Chin, R.T., 1988. On image analysis by the methods of moments. IEEE Transactions on pattern analysis and machine intelligence. 10(4), pp.496-513.

Thuilot, B., Martinet, P., Cordesses, L. and Gallice, J., 2002. Position based visual servoing: keeping the object in the field of vision. In IEEE International Conference on Robotics and Automation, 2002. Proceedings. ICRA'02. 2, pp. 1624-1629.

Tong, J., Jiang, H. and Zhou, W., 2012, May. Development of automatic system for the seedling transplanter based on machine vision technology. In 2012 IEEE International Conference on Computer Science and Automation Engineering (CSAE). 2, pp. 742-746.

Troniak, D., Sattar, J., Gupta, A., Little, J.J., Chan, W., Calisgan, E., Croft, E. and Van der Loos, M., 2013, May. Charlie Rides the Elevator--Integrating Vision, Navigation and Manipulation towards Multi-floor Robot Locomotion. In 2013 International Conference on Computer and Robot Vision (CRV). pp. 1-8.

Tsugawa, S., 1994. Vision-based vehicles in Japan: machine vision systems and driving control systems. IEEE Transactions on industrial electronics. 41(4), pp.398-405.

Ulrich, M., Wiedemann, C. and Steger, C., 2009, May. CAD-based recognition of 3D objects in monocular images. In ICRA. 9, pp. 1191-1198.

Urban, J.P., Buessler, J.L. and Gresser, J., 1998. Neural networks for visual servoing in robotics. In Soft Computing for Intelligent Robotic Systems. Physica-Verlag HD. pp. 81-111.

Vernon, D., 1991. Machine vision-Automated visual inspection and robot vision. NASA STI/Recon Technical Report A, 92, p. 40499.

Wang, H., Brady, J.M., 1995. Real-time corner detection algorithm for motion estimation. Image and Vision Computing. 13 (9), 695-703.

Wang, J. and Yagi, Y., 2012, December. Robust location recognition based on efficient feature integration. In 2012 IEEE International Conference on Robotics and Biomimetics (ROBIO). pp. 97-101.

Wang, L. and Yan, L., 2012, December. Edge detection of color image using vector morphological operators. In 2012 2nd International Conference on Computer Science and Network Technology (ICCSNT). pp. 2211-2215.

Wang, X. and Zhang, X., 2013, August. An adaptive edge detection algorithm based on gray-scale morphology. In 2013 International Conference on Measurement, Information and Control (ICMIC). Vol. 2, pp. 1251-1254.

Wang, Y., Lang, H. and de Silva, C.W., 2010. A hybrid visual servo controller for robust grasping by wheeled mobile robots. IEEE/ASME transactions on Mechatronics. 15(5), pp.757-769.

Weifeng, M. and Caixia, D., 2012. An improved wavelet multi-scale edge detection algorithm. In 2012 International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR). pp. 302-306.

Wesley, M.A., Lozano-Perez, T., Lieberman, L.I., Lavin, M.A. and Grossman, D.D., 1980. A geometric modeling system for automated mechanical assembly. IBM Journal of Research and Development. 24(1), pp.64-74.

Westmore, D.B. and Wilson, W.J., 1991, April. Direct dynamic control of a robot using an end-point mounted camera and Kalman filter position estimation. In 1991 IEEE International Conference on Robotics and Automation. pp. 2376-2384.

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

Won, D.H., Chun, S., Sung, S., Kang, T. and Lee, Y.J., 2008, October. Improving mobile robot navigation performance using vision based SLAM and distributed filters. In International Conference on Control, Automation and Systems. ICCAS 2008. pp. 186-191.

Xiong, Y. and Quek, F., 2002, November. Machine vision for 3D mechanical part recognition in intelligent manufacturing environments. In Proceedings of the Third International Workshop on Robot Motion and Control. RoMoCo'02. pp. 441-446.

Xu, H., Zhang, Y. and Zhao, H., 2012, December. Edge detection of color image using mathematical morphology in HSV color space. In 2012 2nd International Conference on Computer Science and Network Technology (ICCSNT). pp. 2112-2116.

Xue, L.Y. and Pan, J.J., 2009, July. Edge detection combining wavelet transform and canny operator based on fusion rules. In 2009 International Conference on Wavelet Analysis and Pattern Recognition. pp. 324-328.

Yamaguchi, J. and Nakajima, M., 1990, November. A 3D shape identification system using a fiber grating vision sensor. In 16th Annual Conference of IEEE Industrial Electronics Society, 1990. IECON'90. pp. 507-511.

Yang, Z. and Fang, T., 2010. On the accuracy of image normalization by Zernike moments. Image and Vision computing. 28(3), pp.403-413.

Yanjun, Y., Zhongfan, X., Qiang, W. and Zaixin, L., 2010, October. Robot autonomous navigation based on multi-sensor global calibrated. In 2010 International Conference on Computer Application and System Modeling (ICCASM 2010). 3, pp. V3-706.

Yu-qian, Z., Wei-hua, G., Zhen-cheng, C., Jing-tian, T. and Ling-Yun, L., 2006, January. Medical images edge detection based on mathematical morphology. In 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference. pp. 6492-6495.

Zakaria, M.F., Choon, H.S. and Suandi, S.A., 2012. Object Shape Recognition in Image for Machine Vision Application. International Journal of Computer Theory and Engineering. 4(1), p.76.

Zeng, X., Huang, X. and Wang, M., 2008. Research of Invariant Moments and Improved Support Vector Machine in Micro-Targets Identification. Journal of Applied Sciences. 8(21), pp.3969-3974.

Zhang, B., Wang, J., Rossano, G., Martinez, C. and Kock, S., 2011, December. Vision-guided robot alignment for scalable, flexible assembly automation. In 2011 IEEE International Conference on Robotics and Biomimetics (ROBIO). pp. 944-951.

Zhang, H., Shu, H., Coatrieux, G., Zhu, J., Wu, Q.J., Zhang, Y., Zhu, H. and Luo, L., 2011. Affine Legendre moment invariants for image watermarking robust to geometric distortions. IEEE Transactions on Image Processing. 20(8), pp.2189-2199.

Zhang, D. and Lu, G., 2004. Review of shape representation and description techniques. Pattern recognition. 37(1), pp.1-19.

Zhang, J. and Liu, D., 2013. Calibration-free and model-independent method for high-DOF image-based visual servoing. Journal of Control Theory and Applications. 11(1), pp.132-140.

Zhang, X. and Li, J., 2011, December. A method of color edge detection using mathematical morphology. In 2011 International Conference on Computer Science and Network Technology (ICCSNT). 3, pp. 1633-1637.

Zhang, Z., 2000. A flexible new technique for camera calibration. IEEE Transactions on pattern analysis and machine intelligence. 22(11), pp.1330-1334.

Zhao, C.J. and Jiang, G.Q., 2010, July. Baseline detection and matching to vision-based navigation of agricultural robot. In 2010 International Conference on Wavelet Analysis and Pattern Recognition. pp. 44-48.

Zhao, Y.M., Xie, W.F., Liu, S. and Wang, T., 2015. Neural Network-Based Image Moments for Robotic Visual Servoing. Journal of Intelligent & Robotic Systems. 78(2), pp.239-256.

Zheng, W., Zhou, F. and Wang, Z., 2015. Robust and accurate monocular visual navigation combining IMU for a quadrotor. IEEE/CAA Journal of Automatica Sinica. 2(1), pp.33-44.

Zhu, M., Derpanis, K.G., Yang, Y., Brahmbhatt, S., Zhang, M., Phillips, C., Lecce, M. and Daniilidis, K., 2014, May. Single image 3d object detection and pose estimation for grasping. In 2014 IEEE International Conference on Robotics and Automation (ICRA). pp. 3936-3943.

Zhu, S., Shi, J. and Zhou, R., 2012, November. A defect detection experimental system for microwave heated in-place recycling of asphalt pavements. In 2012 19th International Conference Mechatronics and Machine Vision in Practice (M2VIP). pp. 344-348.

Ziou, D. and Tabbone, S., 1998. Edge detection techniques-an overview. Pattern Recognition and Image Analysis C/C of Raspoznavaniye Obrazov I Analiz Izobrazhenii. 8, pp.537-559.

https://en.wikipedia.org/wiki/Corner_detection

http://sensorwiki.org/doku.php/sensors/ultrasound

http://sine.ni.com/nips/cds/view/p/ lang/en/nid/211069/overview

http://sine.ni.com/nips/cds/view/p/lang/en/ nid/205150

http://sine.ni.com/nips/cds/view/p/lang/en/nid/207346

http://sine.ni.com/nips/cds/view/p/lang/en/nid/207415

http://sine.ni.com/nips/cds/view/p/lang/en/nid/207415.

http://sine.ni.com/nips/cds/view/p/lang/en/nid/208264

http://sine.ni.com/nips/cds/view/p/lang/en/nid/210547

http://sine.ni.com/nips/cds/view/p/lang/en/nid/3809

http://www.ab.com/en/epub/catalogs/12772/6543185/12041221/12041231/Capacitive-proximity-Sensing.html

http://www.key.net/products/cayman

http://www.kuviovision.com/machine-vision-industry

http://www.pharmaceuticalonline.com/doc/symetix-introduces-vantyx153-inspection-syste-0001

http://www.robotics.org/product-catalog-detail.cfm/Kawasaki-Robotics-USA-Inc/Kawasaki-RS06L-Robot-high-speed-6-kg-payload/productid/3650

http://www.roboticstomorrow.com/article/2015/06/vision-guided-robotics-advances/6241?

http://www.stemmer-imaging.pl/en/applications/at-the-cutting-edge-of-carrots/

http://www.visiononline.org/vision-resources-details.cfm/vision-resources/Computer-Vision-vs-Machine-Vision/content_id/4585

http://www.visiononline.org/vision-resources-details.cfm/vision-resources/ Pharmaceutical - Industry -Applications-of-Machine-Vision/content_id/1121

https:// www.ia.omron.com/support/guide/43/introduction.html

https://en.wikipedia.org/wiki/Computer_vision

https://www.sick.com/in/en/product-portfolio/vision/2d-vision/inspector/c/g114860

https://www.vrmagic.com/simulators/eyesi-surgical/

www.cse.usf.edu/.../MachineVision_Chapter15.pdf

www.ni.com/pdf/manuals/370838b.pdf

www.ni.com/pdf/manuals/372752b.pdf

www.ni.com/pdf/manuals/373716b.pdf

www.robotics.org/product-catalog.../Kawasaki- Robotics.../Kawasaki...Robot.../3650

Dissemination

Internationally Indexed Journal

1. **B K Balabantaray**, O P Sahu, N Mishra, and B B Biswal, "A Quantitative performance Analysis of Edge Detectors with Hybrid Edge Detector" Journal of Computers (JCP), Vol. 12, no. 2, pp. 165-173, 2017, ISSN: 1796:203X.
2. O P Sahu, **B K Balabantaray**, N Mishra and BB Biswal, "An Integrated Approach of Sensors to Detect Grasping Point for Unstructured 3-D Parts" Journal of Electronic Science and Technology (JEST), ISSN: 1674:862X, Vol. 9, No. 1, February 2017, DOI: 10.7763/IJET.2017.V9.950.
3. **B K Balabantaray** and B B Biswal; "Hybrid edge detection technique for part identification in robotic assembly system under vision guidance", AIMTDR, IIT Guwahati, 2014.

Other Journals and Book Chapters

1. **B K Balabantaray**, B Das and B B Biswal; "Comparison of Edge Detection Algorithm for Part Identification in a Vision guided Robotic Assembly System", Book: Soft Computing Techniques in Engineering Applications, Series: Studies in Computational Intelligence, Published by Springer International Publishing, pp 183-206, 2014.

Conferences

1. **B K Balabantaray**, O P Sahu, N Mishra and B B Biswal; "A Quantitative Performance Analysis of Edge Detectors with Hybrid Edge Detector", 2015 International Conference on Computer Systems and Instrumentation (ICCSI 2015), NTU, Singapore, 16-17th December, 2015
2. O P Sahu, **B K Balabantaray**, B B Biswal, B Patle; Part Recognition Using Vision and Ultrasonic Sensor for Robotic Assembly System, IEEE SCORed 2015, Kuala Lumpur, 13-14th December, 2015
3. N Mishra, **B K Balabantaray**, O P Sahu, and B B Biswal; "Feature based face detection in HRI", 2015 International Conference on Computer Systems and Instrumentation (ICCSI 2015), NTU, Singapore, 16-17th December, 2015
4. P Singh, T Sethi, **B K Balabantaray**, B B Biswal; "Advanced vehicle security system," 2015 IEEE International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), pp.1-6, 19-20 March 2015 doi: 10.1109/ICIIECS.2015.7193276
5. **B K Balabantaray** and B B Biswal; "Part identification in robotic assembly using vision system", Proc. SPIE 9067, Sixth International Conference on Machine Vision (ICMV 2013), 906714, UK, 16-17th November, 2013, doi:10.1117/12.2051309.
6. **B K Balabantaray**, P Jha and B B Biswal; "Application of edge detection algorithm for vision guided robotics assembly system", Proc. SPIE 9067, Sixth International Conference on Machine Vision (ICMV 2013), UK, 16-17th November 2013 doi:10.1117/12.2051303.

Index