Development of a Machine Vision System to Localise a Zinc Die Cast Product

ME Thesis

by

Luke Butters

Submitted in Partial Fulfilment of the

Requirements of the

Degree of Master of Engineering

School of Engineering, Computer and Mathematical Sciences

Auckland University of Technology

New Zealand

August 2019

List of acronyms

1D	1–dimensional
2D	2–dimensional
3D	3–dimensional
CCD	Charge–couple device
fps	Frames per second
PCA	Principal component analysis
PLC	Programmable logic controller
SLAM	Simultaneous localisation and mapping
ToF	Time–of–flight
USB	Universal serial bus

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Luke Butters

Auckland

August 11, 2019

Abstract

The thesis presents a system to automate a manual/repeatable process in an Auckland, New Zealand manufacturing facility using predominantly machine vision techniques. An overview of the research [3] has been accepted into The 2nd International Conference on Control and Computer Vision (ICCCV 2019).

A manual/repeatable process has previously been required in the production of zinc die cast products. Where a worker stands at the end of a conveyor and picks up incoming die cast outputs for processing.

To automate the process, a machine vision proof of concept was developed including four elements. The proposed system decides whether the incoming die cast objects are face up or down on the conveyor, determines the robot pick location and object orientation and conducts a quality control measure to check whether the correct cast is in production.

The proposed system was successfully capable of checking the cast face, determining the robot pick location and orientation along with checking for error for a set of four die cast samples provided by the company. In some cases, small levels of error were corrected for using post vision manufacturing processes including mechanical nests and custom built robot gripping tools.

Acknowledgments

I would first like to express gratitude to my supervisor Professor Reinhard Klette of the School of Engineering, Computing and Mathematical Sciences at Auckland University of Technology. He allowed me to take the project in to my own hands while steering me in the right direction. Thank you for always being a strong source of passion, enthusiasm and wisdom.

I would also like to show my appreciation to David Berry and Paul from ControlVision Limited in Avondale, Auckland. Thank you for sharing with me your time and knowledge to implement key practical elements of the research.

I would also like to acknowledge the hard work and effort provided by the electronics technicians on Level 4 WZ engineering building at Auckland University of Technology. With special thanks to Stephen Hartley, your door is always open.

I would also like to acknowledge Assa Abloy with special thanks to Marc Simkin, Bevan Rolfe and Russell King in Auckland along with Duane Kinney, Greg Schroeder and Chaz Zygmont in Connecticut. You taught me real–world automation fundamentals and openly shared your experience.

I would finally like to give a special thanks to all of my post–graduate colleagues, family and friends who have been at my side throughout my thesis endeavour. Without you, nothing would be possible.

Luke Butters

Auckland

August 11, 2019

Contents

Chapter 1

Introduction

The following introductory chapter provides a background to automation and machine vision, motivation for the work conducted along with an outline of the document. In the document outline, a brief disclosure of the proposed system is included.

1.1 Background

Automation [52] and robotics to enhance human capability have been developing strongly since the early 1900s. During World War I mankind saw one of the first recorded uses of an autonomous war boat [47] developed by Nikola Tesla [61]. The boat was seen as an unmanned torpedo which could be directed toward enemy ships or ports. This early piece of technology was the beginning of the drone which has since been developed. Regardless of the fact that the general public saw the boat as a war device, Tesla saw his creation as something greater – a self–realisation that his own body and mind could be automated. The device of the time responded to an operator from some distance. In the future it was hoped that such creations could act for themselves

to their surroundings.

By 1947, the term *automation* began to spread after Henry Ford implemented an 'Automation Department' at the Ford Motor Company. Earlier in 1922 Ford's autobiography discussed his concept for the assembly line [51]. The benefits of such an assembly line were huge and eliminated the need for employees to carry heavy loads, bend or halt work. Ford removed requirements of special training, creating jobs that majority of the people could execute along with the opportunity for hiring of immigrants. These benefits had a huge impact on the Ford company amongst many others to come. Value added by the assembly line allowed for a pay rise from $\$1.50$ per day to $\$5.00$ per day.

By 1958, companies including Siemens Simatic and Phillips developed the next generation of *relay logic*. These systems used solid state digital technology to hard–wire programmable logic controllers. Eventually by 1968, Dick Morley [42] invented the *programmable logic controller* (PLC) as we know it today using ladder logic to program instead of hard wiring. Morley, a junior engineer responsible for designing communication systems, atomic bombs and aircraft with colleagues created the single most influential tool for the manufacturing industry.

Machine vision [60] provides capability for systems to sort and inspect items, execute repeatable tasks with greater efficiency and speed than a human worker and help find defects in the manufacturing environment. Machine vision therefore is related to computer vision from computer science but has a specific relation with the manufacturing and real–world environment.

During the 1950s *2–dimensional* (2D) [40] vision systems were created based on a pixel by pixel computation of optical flow for statistical pattern recognition. *3–dimensional* (3D) machine vision came during the 1960s where Larry Roberts from Massachusetts Institute of Technology researched the extraction

of 3D points from 2D images.

In the modern time, machine vision has advanced to a stage where high–speed real–time image analysis is possible for 3D and 2D images. Many systems incorporate colour and or thermal imaging [59] to exploit the sensing of wavelengths greater than those on the visible spectrum. The colour spectrum visible to the human eye ranges from 400 nm to 700 nm, where as the thermal image ranges from 700 nm to 1 mm. Other vision systems incorporate *time–of–flight* (ToF) [12] and structured light sensors [13].

The future of machine vision [39] looks exciting for the engineering community. A rapid increase of hardware speed has allowed for higher resolution cameras to be deployed in real time. Due to low computing power, low resolution grey cameras were used in the past.

An example of an outdated machine vision system I have personally worked with was responsible for identifying if a product had successfully been ejected from its process or not. The low–resolution grey camera would monitor the ejector position and find the mean value of intensity across the image. If the image had a mean intensity bellow a threshold across all pixels in the image, it was assumed that the product had been ejected. If the mean intensity of all pixels in the image was greater than a threshold, it had failed to eject the part correctly and the machine operator was notified. Machine breakdowns are typically devastating to production profits.

With computing time now rapidly reduced, low cost PC–based vision systems can monitor their application with extremely high resolution. This increase in image resolution allows for a much broader range of image processing applications. These applications include the surveillance and monitoring of the public, safe and robust autonomous driving, the inspection of food and manufactured products to find defects, anomalies or damage. None of these

applications would be possible with the low–resolution grey image process described above.

Flexible smart cameras are becoming more and more popular. A smart camera is a machine vision product which contains all the necessary components of the system including the matrix sensor, circuitry, memory, computer processor, communication system, inputs and outputs, lens, lighting and a real time operating system. *Smart cameras* such as the Cognex Insight 7600 [34] come with patented software which allows you to plug the camera in and program using the Cognex programming environment. This allows front end automation engineers to rapidly program their machine vision system to be robust and fast without experiencing the extreme learning curve that the computer scientist would have to endure in order to program the back end system.

As discussed, programming a machine vision system is only part of the challenge. An automation engineer must develop the whole system. This includes the high–level creative process; a flow chart detailing what task each element of the system will be performing. They must select suitable hardware (including the camera and illumination) to fit the task at hand within budget along with suitable calibration techniques.

Throughout my undergraduate studies, I worked with Compac Sorting Equipment for several years. During my time at Compac I was responsible for the installation, calibration and configuration of machine vision systems. These systems were used to detect the blemishes/surface defects, colour, size, shape and density of fruits including but not limited to citrus, apples, kiwifruit, peaches and cherries. An optional add–on using infrared spectrometers to analyse a fruits internal property including brix level (percentage of solids present in the juice of a plant), internal defects and internal colour.

The machine vision systems used by Compac would take approximately

thirty images of each fruit from five angles to recreate an image map. This entire operation would occur in real time as fruit passed by the machine at 600 fruit per minute per lane.

Traditional vision techniques useful in the development of a proof of concept to localise zinc die cast products could be implemented with OpenCV. OpenCV is an open source computer vision library with a versatile range of programming functions available in Python, C++, Java and MATLAB. Originally developed by Intel. The library allows researchers to test computer vision methods in a set–up programming environment on input images or video streams. The programming library is BSD–licensed, allowing businesses to take advantage by modifying code to suit their needs.

OpenCV includes greater than 2500 algorithms which contain a mixture of traditional and contemporary computer vision algorithms. Some features include:

1. Facial recognition,
2. classify video actions of people,
3. 3D estimations,
4. object detection,
5. object and movement tracking,
6. stitching images together,
7. generate point clouds from stereo images.

A sub–set of the entire libraries functionality was required in the presented development for understanding shape properties of subject images.

1.2 Defining the Scope

Assa Abloy [50], a Swedish manufacturing company and global leader in door entry systems runs a manufacturing facility in Albany, Auckland, New Zealand. The Swedish conglomerate by sales volume is the largest lock manufacturer in the world. They have made over 150 acquisitions of other successful companies including Yale and Sargent from the USA. I had the pleasure in September of 2018 to visit Assa Abloy's Connecticut factories which are responsible for the production of Yale, Sargent and Assa Abloy products. In 2017 Assa Abloy saw a total revenue of 76.14 billion SEK (Swedish Krona), approximately $\$13$ billion NZD.

The founder of the original Abloy, Emil Henriksson invented the disc tumbler lock [56] in 1907 before the company had acquired Assa to complete the current title. Assa Abloy was finally established in 1994 after separating from Swedish security company Securitas AB. Abloy Oy, a Finnish high security lock manufacturer joined Assa.

Along with having a large global presence, Assa Abloy have solidified their position in the New Zealand market. Door entry system products that are supplied to New Zealand through Assa Abloy include but are not limited to Henderson, Interlock, Trimec, Lockwood and Assa Abloy. Currently Assa Abloy have two main facilities in Auckland. A manufacturing facility in Albany predominantly producing products such as locking mechanisms, hinges and door handles. The Mount Wellington facility is predominantly an industrial supplier for door entry systems such as automatic doors, also containing a field service and install team.

While China has become a major player in the global domain over the past forty years, countries like New Zealand and the USA have seen the majority

of their manufacturing move offshore [45]. With New Zealand having such a high minimum wage in comparison to competing nations, over time it has become less and less financially viable to manufacture goods on shore. Low wages in China have allowed for mass production of consumable goods which are sold back to nations like New Zealand and the USA. In 2014, China for the first time exported more value in products than the USA.

Zinc die casting [55, 53], a major factory process for Assa Abloy New Zealand is subject to such challenges where the production cost can be extremely high due to labour costs. The Albany production facility requires on average six machine operators (one of which is a team leader) and a main supervisor running day and night in order to keep up with customer demands. A tooling team is required, typically two to three workers continuously producing new tools for new custom products. The process requires raw zinc ingots at its input which are mined and shipped in along with six die cast machines/conveyor systems.

A major element of Assa Abloy New Zealand's manufacturing process involves *zinc die casting*. But what is zinc die casting? The process itself is reasonable simple by nature. Molten metal is forced with high pressure to fill the inside of a mold cavity (this is the die). The mold cavity can be visualised as the 3D negative hollow shape of the desired cast output. Figure 1.1 showed an example of a die cast output provided by the company where window handles were seen on the right hand side.

When the molten metal runs into the mold cavity it enters through what is called the 'sprue' and then fills the 'runners' before being forced through the 'gate' into the final fill location which takes form of the product (door handle, hinge, locking mechanism). Typically non–ferros metals are used in die casting such as zinc, magnesium, copper or aluminium. The cast tooling is machined and broken into two halves which are forced together to allow the

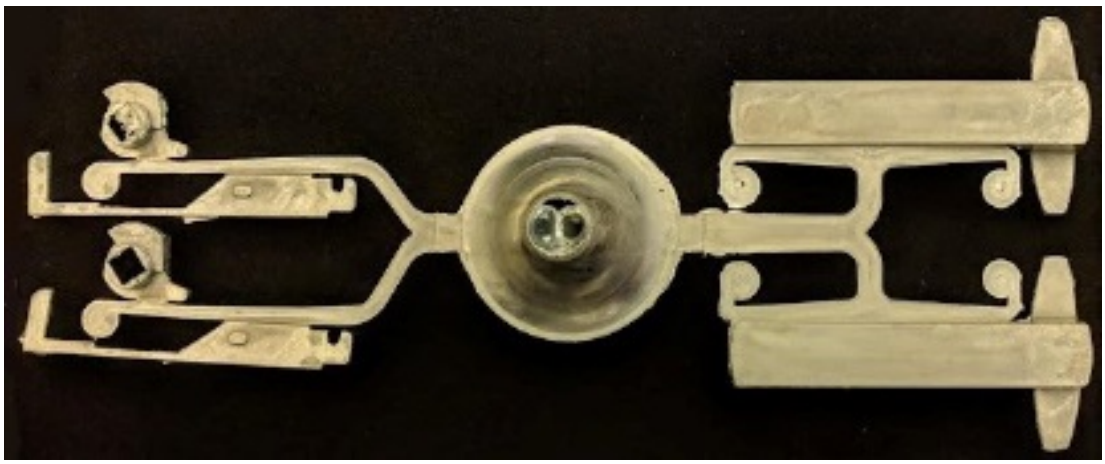

Figure 1.1: Zinc die cast output sample provided by Assa Abloy

molten metal to fill inside, typically using pneumatics.

The sprue can be seen as the circular feature in the centre of the top down view of the die cast sample from Fig. 1.1. This is where the molten metal enters the mold cavity. To either side of the circular sprue feature, the runners are the 'branch–like' features. Between the door handle product on the right hand side and runner is the gate, where molten metal is transferred from the waste section of the casting the usable product.

Gate widths are adjusted to optimise production quality. A reduction in gate width improves the ease of removing the product from the waste, while reducing quality. Gate widths are kept low in this scenario to 0.2 mm, causing the molten zinc to be sprayed into the empty mold cavity. Spray pressure must be carefully considered to ensure entire die is filled with molten zinc. This is particularly important for small or detailed objects such as the ones seen on the left hand side of Fig. 1.1. If spray pressure is to low, in some cases parts of the object will be incomplete and unworthy of sale.

In usual zinc die cast manufacturing, the gate is set around 2 mm–4 mm. Such widths are considered thick and allow the molten zinc to flow freely

through the gap. It is to large for a human or robot to break the product off the sprue runner with ease as a process of separation. Post processing techniques must be deployed involving cutting or grinding machines to separate the product from its waste.

Another unique method deployed by Assa Abloy New Zealand's die cast process that is not seen elsewhere around the world is a modulated tooling approach. Typically in die casting, a single machine tool is used throughout each cast cycle. A production line will produce one part using a single tool for hundreds or thousands of cycles until the inventory reached satisfaction. Once the correct number of products has been produced, the tool will be swapped out to produce another product if needed. Assa Abloy New Zealand on the other hand runs two tools at once, allowing for multiple products to be produced on one machine at any given time. Perhaps one tool is run for 500 cycles before being swapped for another tool for 500 cycles, but the second tool is run for 1000 cycles.

This unique method of modularisation allows for extremely fast and flexible production. The manufacturing facility operates in real time from customer order to dispatch. Within Assa Abloy New Zealand's real estate, there is little to no inventory kept. All products are produced and ready for shipment within 48 hours of customer order.

Machinery, tooling and operation costs of production are high. Die casting as a process is usually justified financially when dealing in high volumes. For Assa Abloy New Zealand, companies such as Bunnings Warehouse and Mitre 10 will order products such as door handles, hinges or locks in the quantity range of the thousands at any given time.

With swift and steady supply chain systems available today, Assa Abloy New Zealands die cast process is at risk due to the high cost of labour required.

If the company can successfully automate the repeatable manual process performed by the machine operators, the process itself will be able to remain on New Zealand shores. The die casting may no longer be financially justifiable with the current cost of labour required to remain profitable.

The idea that robots are taking away people jobs is not always completely accurate. With low skilled labour being more affordable offshore, it is fair to say that robots will only encourage manufacturing to return to countries like New Zealand and the USA where minimum wage is higher. To maintain competitiveness in the modern market, New Zealand manufacturing companies like Assa Abloy must invest in robotics and automation innovation.

Marc Simkin the Manufacturing Engineering Manager at the time from Assa Abloy New Zealand required a suitable master's student to undertake a research project which is delivered in this thesis. The project was intended to be a collaboration between Assa Abloy New Zealand, AUT University and the master's student. The main research question was – How can the position, orientation and identification of zinc cast sprues be found using predominantly machine vision techniques?

Machine operators are required throughout the die cast process. Their main role is highly repeatable by nature. Viewing a moving conveyor belt until a cast sprue appears before picking and identify. The useful products are located and removed from the cast sprue before discarding the sprue waste into the melting pot Fig. 1.2.

Finding the cast sprue position and orientation will allow for the command of a robot arm such as a 6–axis Universal Robot [49] to robustly grip the sprue from its sprue centre and rotate it to a known position. Once the cast sprue is fixed in a known position, the necessary post processing manufacturing techniques can be applied to separate the useful product output with the zinc waste

Figure 1.2: Assa Abloy employee manually dealing with a die cast output

from the sprue and runner. An identification process can prove whether on not the cast sprue being observed by the system is the expected output. Any incorrectly mismatches can be sent to recycle while the system can be notified.

1.3 Thesis Outline

This thesis provides a discussion and research based proof of concept solution to the main research question posed earlier in the literature namely – How can the position, orientation and identification of zinc cast sprues be found using predominantly machine vision techniques?

I began writing warming the reader up to the history and current status of the disciplines of automation and machine vision. A brief explanation of some of the previous automation and machine vision applications that I have been exposed and influenced by personally was noted. Following, I introduced the company who's systems the research is based; Assa Abloy New Zealand. After talking about the company's history, processes and position in New Zealand I explained the die casting process at greater depth in order to give the reader a better idea of the key element involved within the scope of the project. Once the die cast process had been revealed I explained the crucial need for automation and how this research could begin paving the way for such action. The major research question was outlined, and a brief high–level explanation of the proposed system was explained.

In Chap. 2, a comprehensive review of the relevant research to date is reflected upon. The research predominantly covers machine vision theories and techniques that have been explored and have made contribution to the research project discussed in this thesis.

Towards the outset of the literature review, the topic first focused on is the element of machine vision namely the detection of robot picking location. Hough circles [11, 41] gave insight detecting a circle of specific radius in an image. Additional research was explored to form the foundation of preprocessing including blurring [46], thresholding [48], edge detection methods [33]

and contour detection [35].

Once the research of robot picking location was conducted, current studies around finding the casts angle of orientation and pose were analysed. *Principal component analysis* (PCA) [22], using mathematics to transform a 2D image into a *1–dimensional* (1D) line with measurable gradient appears as the main element. With addition, cropping methods [37] for image reduction can be seen.

To determine if the correct cast was being produced or not, convex hull [54] along with moments [58] to find perimeter, area and centroid were investigated. Finding the convex hull of an object is to find the 2D shape that would form if an elastic band were placed around the object; completely enclosing the entire object inside, while reducing its area.

Detecting if incoming casts were facing up Fig. 1.3 towards the camera or down Fig. 1.4 away from the camera was a necessary preprocessing stage required. Cast machine outputs could appear either face up or down on the

Figure 1.3: Side view of face up die cast sample

Figure 1.4: Side view of face down die cast samples

conveyor belt.

ToF and structured light sensors offered low cost and robust methods to generate rough point clouds of the top surface of the die cast output in 3D. The shape of the point cloud when analysed can describe whether a known cast output shape is face up or down. Laser triangulation gave a simple method to detect a line or dot's 3D coordinates.

Other topics of research within machine vision were studied in order to generate an understanding for the foundations of the application over and above those that have been discussed. These include Hu moments [10], 2D vision, 3D vision along with smart cameras.

Chapter 3 discusses the proposed system. A high–level process model was created. This model gave an outline so a zinc die cast output could be picked from a conveyor exiting the die cast machine in a known, fixed orientation and position while determining if the cast was the expected cast or not.

The ToF vision system used to determine if a zinc die cast passing on the conveyor was face up or face down is discussed. Die cast machines output the

objects randomly; either face up or face down onto the conveyor belt. If a cast appears face down, a mechanical process is used to flip the cast to the face up position before continuing to the next process; namely the main vision system.

The 2D vision system used to determine the face up casts robot picking location, cast orientation along with a quality control check to determine if the correct cast is being produced is explained. For all casts, the robot grips the zinc die cast on the circular repeatable shape; the cast sprue cone. The system can robustly find the (x, y) pixel location on the image of the cast sprue cone robustly and is assumed the centre of the robot pick location. With the angle of orientation found using PCA along with robot (x, y) pickup location, the zinc die cast part can be gripped and fixed in a known pose. Once the object is fixed in a known pose, the system is able to perform the necessary post pickup manipulations.

A crucial quality control measure was used as part of the 2D vision system. The produced zinc die cast output image was split into its two respective tools and tests were performed on each to decide whether the correct products were being produced. A notification would be sent to the machine operator if any incoming object was not as expected.

The 2D vision system responsible for determining robot pick location, angle of orientation along with performing a quality control method had positive results that were presented. Data includes robot (x, y) picking coordinates, angle of orientation and the convex hull perimeter, area and centroid of each modulated tool. All values described were compared against expected values and reflected upon in the discussion section to follow.

Chapter 4 presents the discussion of results. Weaknesses and strengths of the designed method for the machine vision systems used to answer the main research question were noted. The outputs of the proposed system operated

with small errors seen at the (x, y) pick location and angle of orientation. Post vision manufacturing processes were used to compensate for the error introduced by the vision system. A specially designed vacuum gripping tool would allow for the zinc die cast output to be picked robustly. A mechanical nest was introduced to completely align the die cast output after it had been gripped and re–orientated into a known position. This corrected small errors introduced through determination of the objects orientation.

Finally, with the discussion of results considered the conclusion of the work and future possibilities of implementation are explained. The conclusion aims to answer the main research question by confirming the outcome. In addition, suggestions for the implementation phase are recommended and elements within this thesis are critiqued.

Chapter 2

Literature Review

This chapter provides a review of selected research to date in the field of machine vision. Some machine vision basics are introduced at the outset of the chapter before relevant publications are analysed to provide knowledge towards the design of the solution. Relevant research is provided in an easily accessibly manner to help in answering the major research question – How can the position, orientation and identification of zinc cast sprues be found using predominantly machine vision techniques?

2.1 Machine Vision

Foundation machine vision techniques including image smoothing, binarisation, edge detection, contour detection, convex hull and moments are explored. Such tools are useful when preparing a 2D image for assessment.

Image Smoothing

When analysing images using machine vision, noise can occur. Colour and brightness can vary at random to generate misinformation in the image. In the

case of a manufacturing environment, even physical matter can create noise in an image such as particles of dust. Image smoothing [62] can be applied to data to reduce such noise.

The Gaussian blur [18, 46] is a robust images smoothing method which applies convolution to the image with the Gaussian function, causing a reduction in contrast and blurring the edges. By smoothing the image, the important data can be considered.

Binarisation

Image binarisation [4] or thresholding is commonly used in machine vision. Fixed thresholding was proposed in [7] where a value was set between the minimum intensity (black) and the maximum intensity (white). All pixel values in the image bellow the threshold value become white, while all values above the threshold value became black.

Edge Detection

Edge detection methods [21] are used to find the pixel locations where there has been a sharp change in brightness. The relevant discontinuities are mapped and an edge image [57], representing the edges of the input image is produced.

Canny edge detection [20] is a popular edge detection algorithm [33] developed by John F. Canny. The multi–stage algorithm incorporates an upper threshold value and a lower threshold value. The output edge map only includes edges within the gradient intensity range specified.

Contours

To find the outer shape of an object with a high contrast background, contours [35] were seen to be a useful tool. A continuous line forming a loop in a grey scale image with constant intensity is considered a contour. Edge maps and binary images typically provide higher accuracy at the output of the contour detector than raw grey scale images. The perimeter of an object lying on a surface with high contrast against the background can be found by isolating the largest contour [17].

Convex Hull

The convex hull [54] represents the minimum convex set that contains a shape. In 2D space, the convex hull can be seen as a region contained by an elastic band. To operate accurately, a contour image is used at the input [36].

Paper [9] provides a method where the convex hull of a hand is used to determine its gesture. The need for heavy gloves with in built sensors to detect hand gestures could be replaced with 2D image analysis.

Moments

Image moments [58] in machine vision are often used to help characterise objects. Contour area [43] and centroid [5] can be determined. Hu moments [10] are particularly useful for shape matching where a set of seven scale, rotation and translation invariant moments are determined to describe a shape.

2.2 Circular Hough Transform

To determine the robot picking location, the centre point coordinate (x, y) of the zinc die casts cone sprue was necessary. When the zinc die cast machine output was viewed by a high–resolution camera from the top down view, the cone sprue was seen as a circle or near circle (ellipse). With high contrast using a matt black conveyor (background), a clear circle with consistent radius was distinguished in the image, representing the base of the cone sprue. The shape was repeatable across all die cast outputs and of equal shape and size.

To detect circles as generalised shapes, the circular Hough transform was presented in [11]. The circle shape was characterised as having a circle centre $(x_c, y_c$ and radius r. For a circle centre point in an image $p = (x, y)$ a circle is determined when all circle edge points relate to the centre point p with the same radius. A range of radius values must be defined with an upper and lower threshold where:

$$r_0 \leq r \leq r_1$$

The circular Hough transform outputs circles detected within the range of radius values specified.

Paper [23] describes the growing need for smart car parks in India and proposes a solution to the problem involving circle detection with the circle Hough transform. India has a rapidly growing use of auto–mobiles seen over the past decade. With large levels of congestion on the roads, infrastructure has become compromised. Circle detection was said to be beneficial as it has proven successful for the detection of cyclist helmets in Thailand, stop–go light detection and in the case of this paper; detecting if a car park is available or not. The proposed scheme involved a circle dot painted on each car park as seen in

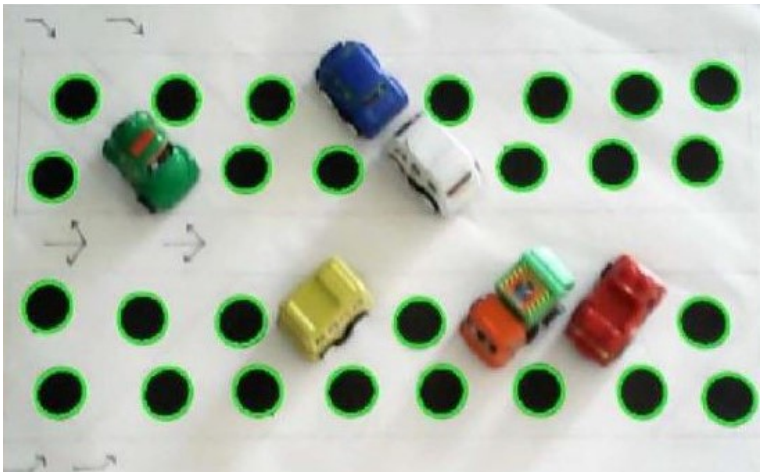

Figure 2.1: Model smart car park with with circles in position of vacant car parks [23]

Fig. 2.1.

When a car occupied a given car park, there was no circle detected for the given park. If there was no car occupying a given car park, a circle was detected. Incorporating this method into a smart car park model, allowed for major reduction in time, congestion and fuel while searching for a vacant park. The system accumulated circles on a counter, then displayed the number of vacant parks within the smart car park before the customer entered the car park. Once the customer was inside the car park, they had the ability to know where the vacant parks exist. Such systems are already in place in some countries, they used a sensor on every single car park. This newly proposed method allowed for a reduction in cost along with increased simplicity of installation and maintenance.

To detect the circles on a given car park, OpenCV and Matlab libraries were used to implement the traditional circle Hough transform. The data set used in experimentation was taken from Google images of circle blobs, along with a developed image including a model smart car park as seen in Fig. 2.1 from

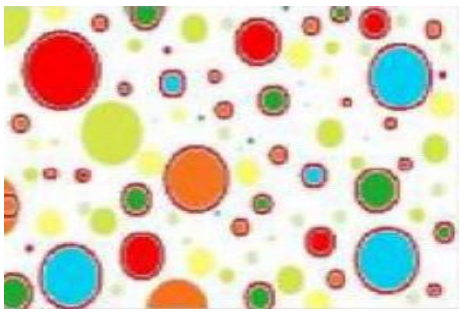

Figure 2.2: Circle image with CHT circle detection output [23]

Figure 2.3: Circle image with CHT circle detection output [23]

paper [23] where model toy cars were used in conjunction with drawn circles as car parks. The generalised Hough transform for a line was said to be represented as in Eq. (2.1). The perpendicular vector length from origin to the line was ρ, and the perpendicular angle of the line to the positive x–axis was θ:

$$\rho = x * \cos(\theta) + y * \sin(\theta) \tag{2.1}$$

When the circular Hough transform algorithm was run on Fig. 2.2 from [23] it was seen at the output that the majority of circles had been successfully detected and mapped. Some of the yellow circles were not detected. This was also seen in Fig. 2.3 from [23]. The author claimed that an average of one fifth of the circles were not detected correctly which accounted for the yellow circles. The circle Hough transform detected non–yellow occluded circles successfully.

In the case of detecting the base of the zinc die cast cone sprue, occlusions were not a concern. Only one circle was visible on the conveoyor belt under normal conditions and would not be occuded. False detections and undetected circles were a concern. It was noted by the author, that illumination should be an early consideration for all computer vision applications and may have had a large impact on the outcome of the circle detection as seen in Fig. 2.2 and 2.3 from [23]. It was important to create contrast between the circular object and its background. Therefore the correct pre–image processing techniques must be deployed in order to create sufficient edge models of all circles. This was considered when developing the method to create contrast between the background of the conveyor and the zinc die cast output.

Paper [25] introduced an accurate approach to detect circles using the Hough transform along with a local maxima concept. The author began by outlining the need for circle detection in industry and mentioned key application areas including detection of pellets in a pelletisation plant, target detection and inspection of manufactured products. For determining the picking location of the zinc die cast, the application included detecting the base of the zinc die cast cone sprue. The data set of images used in the papers experiments were based on detection for the purpose of counting pellets in a container. Pellets are formed from iron ore; making them an ideal shape to be loaded into a furnace in the manufacturing of of steel.

The author acknowledged the circular Hough transform (and its variants) as the most common technique for the detection of circles within the field of computer vision. This particular transform was from the $x--y$ plane (2D image) to the parameter space. For a circle in the $x--y$ plane, the mathematical equation was expressed in Eq. (2.2). The three parameters seen include r, a and b. These parameters were the circles radius, circle centre x position and circle

centre y position, respectively:

$$r^2 = (x - a)^2 + (y - b)^2 \tag{2.2}$$

It was seen that when the traditional method of circular Hough transform was applied to industrial images of pellets, false positives occurred. Figure 2.4 from [25] showed implementation of the traditional circular Hough transform being applied to an edge image of the displayed pellets. Instead of seeing one clear circle around each pellet, multiple false positives were seen around the pellets. This occurred because the pallets were not perfect circles on the 2D plane. The circle Hough transform naturally fit multiple circles around the odd shapes. When the traditional circle Hough transform was applied to an image of coins from top down view, there were no false positives seen. This indicated that only for perfect circles did the circle Hough transform perform ideally. In some cases the cone could be lying slightly on its side, creating an ellipse shape as opposed to a circle when seen by the camera.

To combat the false positives seen in Fig. 2.4 from [25], a method was proposed using local maxima to assign one correct positive to each non–perfect circle. The algorithm bellow described step by step how this process was executed as outlined in the paper [25]:

1. Generate edge map of the circles to be dealt with,
2. at an edge point: draw a circle with the edge point being the centre of the circle,
3. increment the value in an accumulator matrix for the coordinates which lie on the perimeter of the drawn circle,
4. repeat Step 2 and 3 for all edge points and all radii defined within radius

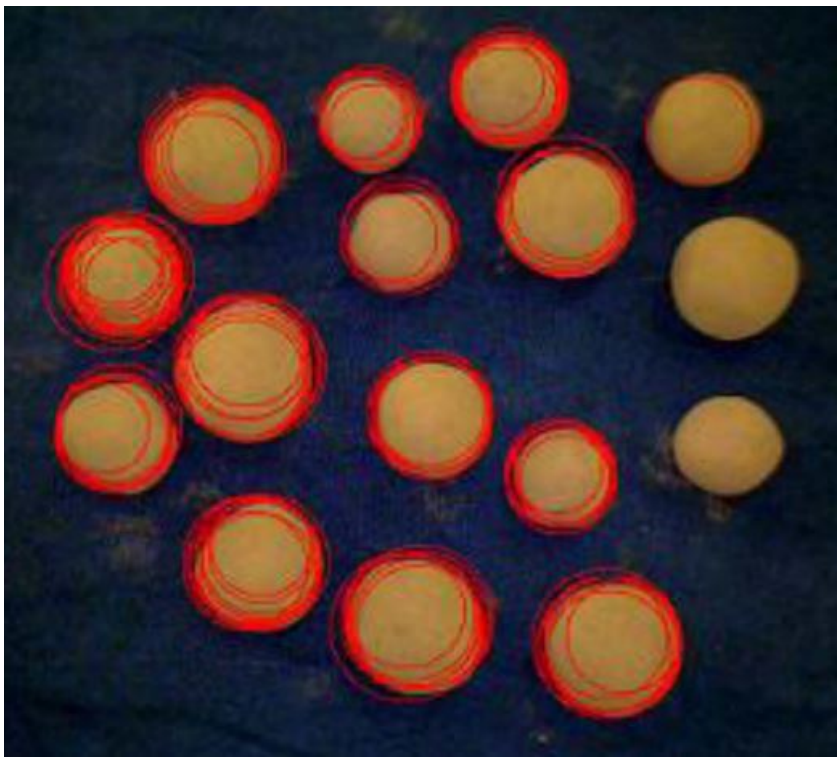

Figure 2.4: Output of traditional CHT applied to pellets [25]

range,

5. for each maxima found: compare the accumulator maxima with with the previous and next accumulators,

6. local maxima found gives circle parameters r, a and b: map these as the output.

A reduction in false positives was seen when the algorithm ran with the industrial image data set of the pellets. The paper compared performance of the traditional circular Hough transform with the proposed method. Edge and circle detection were known to be a necessary process for detecting the base of the zinc die cast sprue cone. Paper [25] brought to light the potential need for additional methods when dealing with imperfect circles.

A hands–on guide explaining how to apply the Hough Circle technique was presented by Adrian Rosebrock [19]. Once the necessary OpenCV and numerical package requirements are installed correctly into a Python programming environment, cv2.HoughCircles function could be utilised to fit circles across clearly defined circular shapes in test images. The NumPy [44] package used for numerical computation was installed.

An input image was loaded into memory and converted from colour to grey. The grey image is then used as input to the cv2.HoughCircles function, along with detection method, an accumulator value and a minimum distance value. For well defined circular shapes in images, it was seen that when the detection method was set to HOUGH_GRADIENT and the accumulator to 1.5, then the minimum distance could be modified to suit the application. This distance described the minimum distance between the center coordinates of detected circles.

Ellipse Detection

An alternative approach to detecting the centre of the repeatable zinc cone sprue was explored. Searching for an ellipse [38] as opposed to a circle could possibly allow for a more robust output for robot pick location. Naturally the repeatable circular cone feature may appear as an ellipse when observed in 2D from the top down. Where a small difference could be seen between major and minor axis.

Ellipse detection methods [24, 26] were considered. Robust extraction schemes showed that for elliptical or near elliptical shapes, of varying image resolutions, an ellipse could be fit onto the image. Once an ellipse was fit, the centre of the major and minor axis could represent the robot pick location.

The ellipse shape can be defined with five parameters. The ellipse centre (x_0, y_0), the major and minor axis a and b and the direction θ_0.

2.3 Principal Component Analysis

In order to fix the cast in a known pose, it was necessary to understand the objects angle of orientation. When the die cast was viewed on a conveyor from the top down position, the objects perimeter was captured. This object outline was posed at an unknown angle of orientation between $-180\,^\circ$ to $180\,^\circ$ of a reference position. With this angle of orientation known, the robot arm could rotate the object about the picking location to fix the zinc die cast output in a known position. PCA, a mathematical operation which reduces the dimensions of data; giving a 1D angle of direction for 2D data points.

The technical report [22] gave an introduction to PCA. The document was broken down in to three sections. Chapter 2; a recap of the necessary mathematics required to apply PCA. This included statistics; standard deviation, variance, covariance, the covariance matrix and eigenvectors and eigenvalues from matrix algebra. Chapter 3; PCA was introduced once the foundation math was understood. A step by step method was used to apply PCA to multidimensional data. The final section of the document Chap. 4 informed the reader how to utilise PCA in the application of computer vision.

The statistics introduced in Chap. 3 was designed to prepare the reader for the analysis of big data based on relationships between singular points in the data set.

Standard deviation was defined as a measure of how spread out data was within a data set. Equation (2.3) showed the function for standard deviation, where X_i represented a data point and n represented the number of data

points:

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \tag{2.3}$$

Variance was defined as the standard deviation squared. It was said to be another commonly used method of expressing the spread of a data set. Equation (2.4) showed the function for variance and its similarities to standard deviation. Covariance was defined as a measure to find how dimensions vary from the mean with respect to each other. Standard deviation and variance being single dimension measures. Covariance, a multi–dimension measure where correlations such as student height and class score could be determined. Finally, the covariance matrix was appropriate for taking covariances of more than two dimensions. The formula for calculating the variance of a dataset:

$$s^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n - 1} \tag{2.4}$$

Matrix algebra was a crucial tool available in order to find the eigenvectors and eigenvalues of a matrix.

By compressing and stretching data, the eigenvectors gave the direction of the linear transformation. Closely related to eigenvectors, eigenvalues gave the amount of scaling of the original vector once multiplied with the matrix. Compression of the data was a factor of the eigenvalue. Both eigenvectors and eigenvalues came as pairs and were both required for PCA.

Chapter 3 introduced PCA as a powerful tool for finding patterns in data of high dimensions. Another major benefit referred to image compression were dimension reduction could be applied with minimal loss of data. The chapter gave a five–step method to apply PCA to reduce dimensions of data.

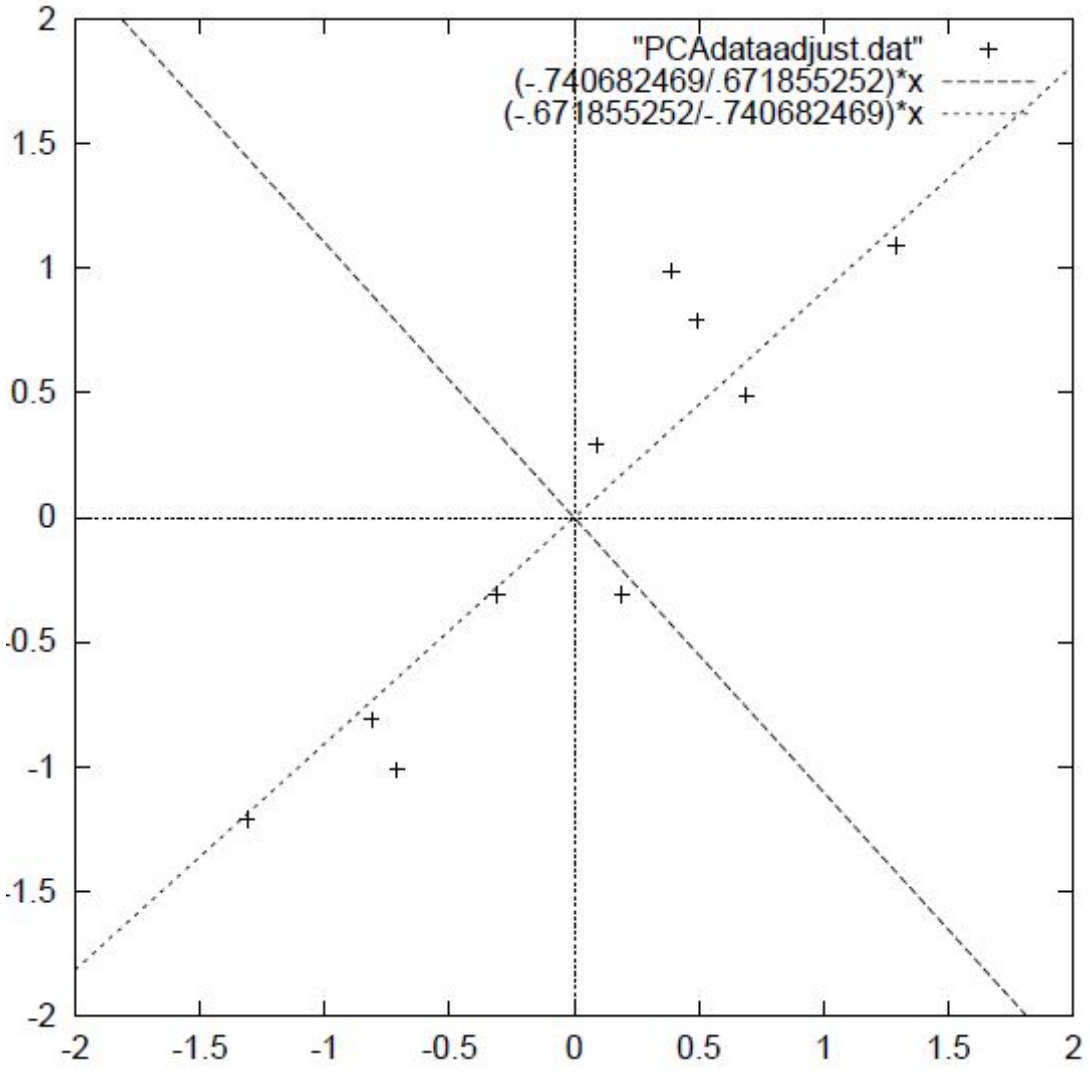

Figure 2.5: Overlayed eigenvectors taken from mean adjusted data points [22]

Firstly, a 2D data set was used which made it possible to understand what was going on graphically. Two–dimensions were additionally appropriate for an image based application. Secondly, the mean value of each dimension was subtracted. Thirdly, the covariance matrix of the 2D data set was taken. From statistics, we knew that for 2D data the covariance matrix would be 2×2. Fourth, the covariance matrix eigenvectors and eigenvalues were calculated. The eigenvectors were orthogonal. Figure 2.5 from [22] clearly showed how

well the eigenvectors could be used to describe a line of best fit for the 2D data set provided. Fifth, the largest eigenvalue corresponded to the eigenvector that was the principal component of the multi–dimensional dataset. Meaning that the line of best fit through the 2D data in Fig. 2.5 from [22] was the principal component and was the eigenvector that corresponded to the highest eigenvalue. The other eigenvector corresponding to the smaller eigenvalue could be disregarded if not needed. A feature vector was formed to confirm which eigenvectors to keep from the list. In the example case, the only options were to keep either one or both. Finally, the output could be computed by multiplying the feature vector with the mean adjusted data after the vector had been transposed.

Chapter 4 concluded the tutorial by outlining how PCA was used in the application of computer vision (related to machine vision). The knowledge throughout the section referred to facial recognition from [27].

Firstly, representation of images to apply PCA to was considered where square $N \times N$ images were represented as N^2 dimensional vectors. Secondly, an example where a new facial image was compared to a data set of twenty facial images where none of which were the same as the new facial image. An image matrix was formed. The example provided showed the technique necessary to apply PCA. It differed slightly from determining the angle of orientation of zinc die cast output. In the example, the principal component was used to find patterns when comparing an entire facial image against twenty other facial images. This same technique could be implemented to find patterns between the x and y axis of a 2D image.

2.4 3D Scanner Systems

3D scanner systems are used in machine vision to determine 3D point cloud coordinate data. This 3D data allows for the reconstruction of an object in 3D space. Two methods are explored including structured light and time of flight.

Structured Light

To ensure that all zinc die casts could robustly be assessed for robot picking location and angle of orientation, it was necessary to determine whether the cast was lying face up or face down on the conveyor. When the zinc die casts were output from the machine, they fell onto the conveyor face up or down. The casts distribution of mass across its volume, the particular machines ejection pin characteristics and the use of a quench bath to cool the die cast output among other factors all influenced whether the cast landed face up or down on the moving conveyor belt. If a rough surface point cloud could be established giving a third dimension of depth data on top of the 2D image, the cast could be determined as face up or down. Taking into consideration the z direction coordinates of the object representing the thickness of the cast at various pixels of the 2D image.

Paper [28] delivered a construction method for a 3D scanner system based on monocular vision. The proposed system was designed for low cost, simplicity, high accuracy and ease of use. Monocular vision (single pinhole camera) was used as opposed to stereo vision (dual camera set–up). The paper introduced a useful camera–image distance calibration scheme which was simple to understand using principals of geometry. Such geometric schemes could be used in conjunction with other technologies including laser line detection and image processing to construct a 3D surface point cloud of objects.

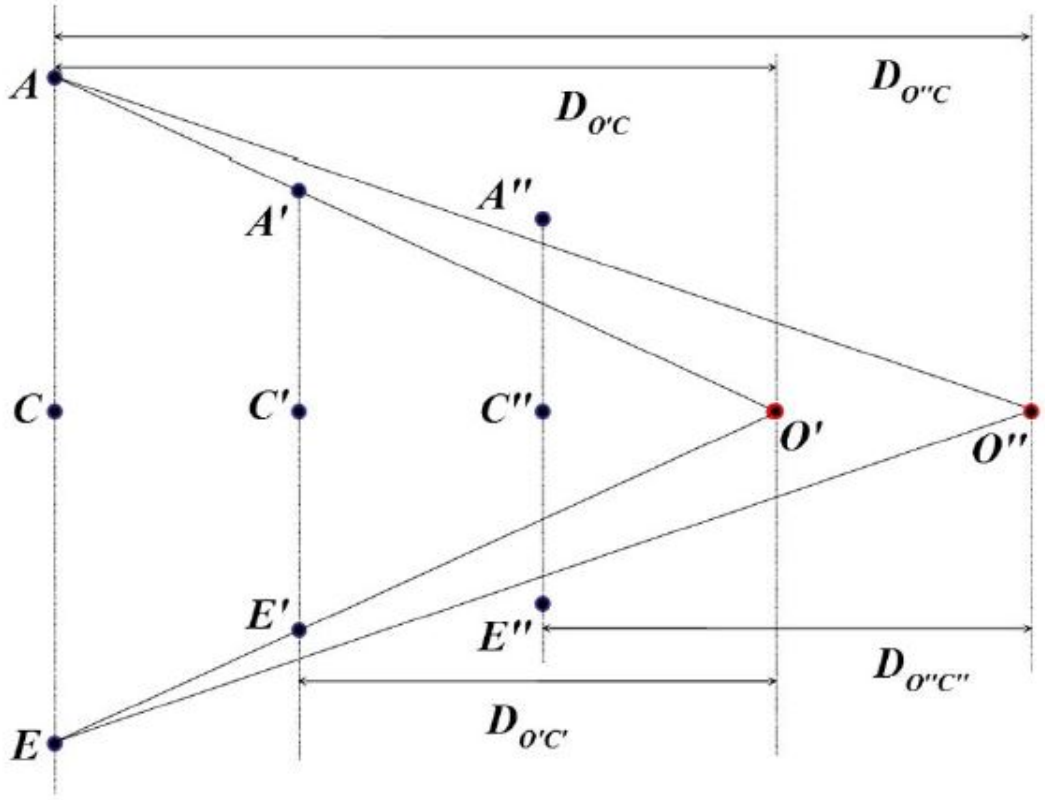

Figure 2.6: Calibration geometry for camera–image distance [28]

The paper gave a method to find the error between the real camera lens centre and a marker dot position. Knowing this quantity of error was necessary to determine how accurate a 3D surface reconstruction method really was. The size of the error dictated how appropriate the utilised method was, given the application. O' and $O"$ gave two test positions for the camera lens. Equation (2.5) gave the equations to solve the expected distance error between the real camera lens and the expected marker position:

$$\varepsilon = \frac{D_{A'E'}D_{O'C} - D_{A"E"}D_{O"C}}{D_{A'E"}D_{A"E"}} \tag{2.5}$$

Figure 2.6 from [28] gave the geometry of quantities in 2.5.

Paper [14] introduced a practical method to build a 3D surface scanning

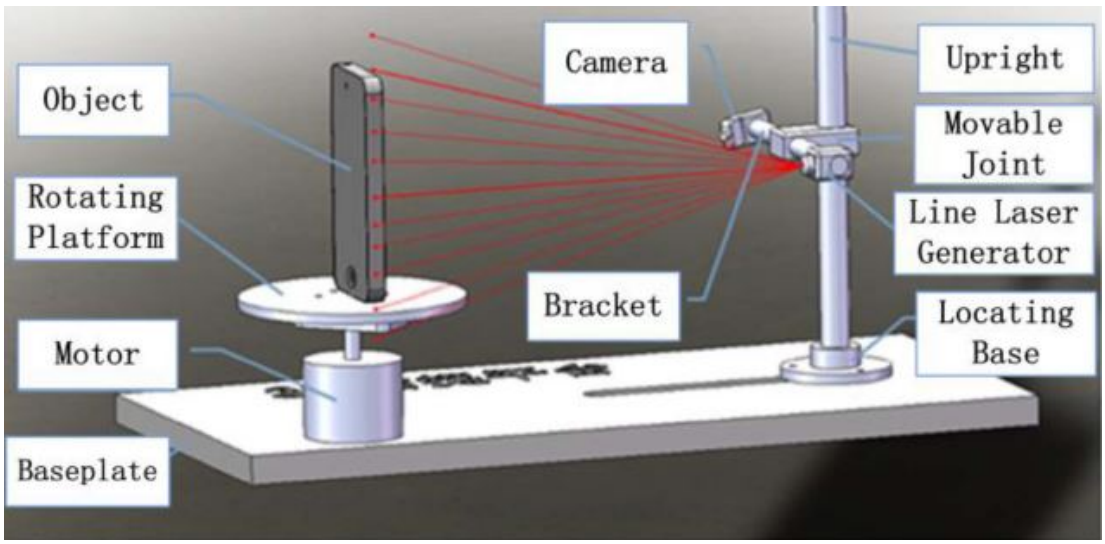

Figure 2.7: Physical set–up of 3D scanner experiment [14]

system based on monocular vision. The paper outlined the benefits of non–invasive surface reconstruction methods and gave examples of its use in application. An overview of mechanical, software, geometry and image processing design was covered. The experiment set–up involved a single camera and line laser pointing at a rotating pan. Objects to be scanned were placed on the rotating pan so the laser line strip could be projected vertically across the centre of the object while the camera acquired the images. The rotating pan was revolved at 1 rad/s while the camera acquired 30 *frames per second* (fps). The mechanical design of the system was seen in Fig. 2.7 from [14].

Camera to world coordinates are transformed by an affine transform:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0^T} & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \tag{2.6}$$

The principal of monocular vision was outlined from a geometric point of view where three coordinate systems are introduced. These three coordinate systems included:

1. Image coordinate system – containing feature extraction point P with coordinates (u, v),
2. world coordinate system – containing the projected point P with coordinates (X_w, Y_w, Z_w),
3. camera coordinate system – containing the projected point P with coordinates (X_c, Y_c, Z_c).

The relationship of the three coordinate systems could be seen in Eq. (2.6) and Eq. (2.7). $[\mathbf{R}\,\mathbf{t}]$ was simply the notation for the rotating pan, this proposed set–up would need to be modified for an object on a conveyor belt. The amplification factor f_x and f_y of the x and y–scale were used in the camera model:

$$Z_c = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0^T} & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \tag{2.7}$$

The design of the software component of the scanner system was comprised of four main elements. These four elements included:

1. Image acquisition and storage,
2. image processing,
3. feature extraction,
4. 3D reconstruction.

A camera of resolution 640 x 480 with frame rate of 30 fps was fixed and poised at the rotating object. *Universal serial bus* (USB) communications were used to deliver the image from the camera to the PC. Pre–processing and feature extraction was used to reduce noise in the image. Regardless of the laser line creating high contrast with the rest of the image, image processing techniques were implemented to detect the laser line. Variables that were considered when dealing with the image processing included:

1. Width of laser source,
2. the surface of the scanned object,
3. influence of the environment.

The width of the laser source was ideally set to a single pixel when the image was acquired by the camera, in practice this was not always the case. When the angle of incidence was large, the interpreted laser width was large. Material of the scanned object and varying colour was seen to alter the output. These variables directly affected the reflection of the laser strip upon the surface of the object. This was considered thoroughly especially when dealing with reflective zinc die casts. Lighting conditions were considered as they influenced the environment.

A geometric model of the rotating object under the projection of the laser line and camera was seen in Fig. 2.8 from [14]. In polar coordinates, the point P was modelled as $(l_{0_1P}, -\theta)$ or $(l_{0_2P}/\cos\theta, -\theta)$. These coordinate systems formed the basis for the scanning system with a rotating disk. Such considerations were important in the case of scanning a stationary object on a conveyor belt, but modification would be necessary.

The proposed method was implemented using C# and EmguCV library functions. Experiments were taken using a drinking bottle seen in Fig. 2.9

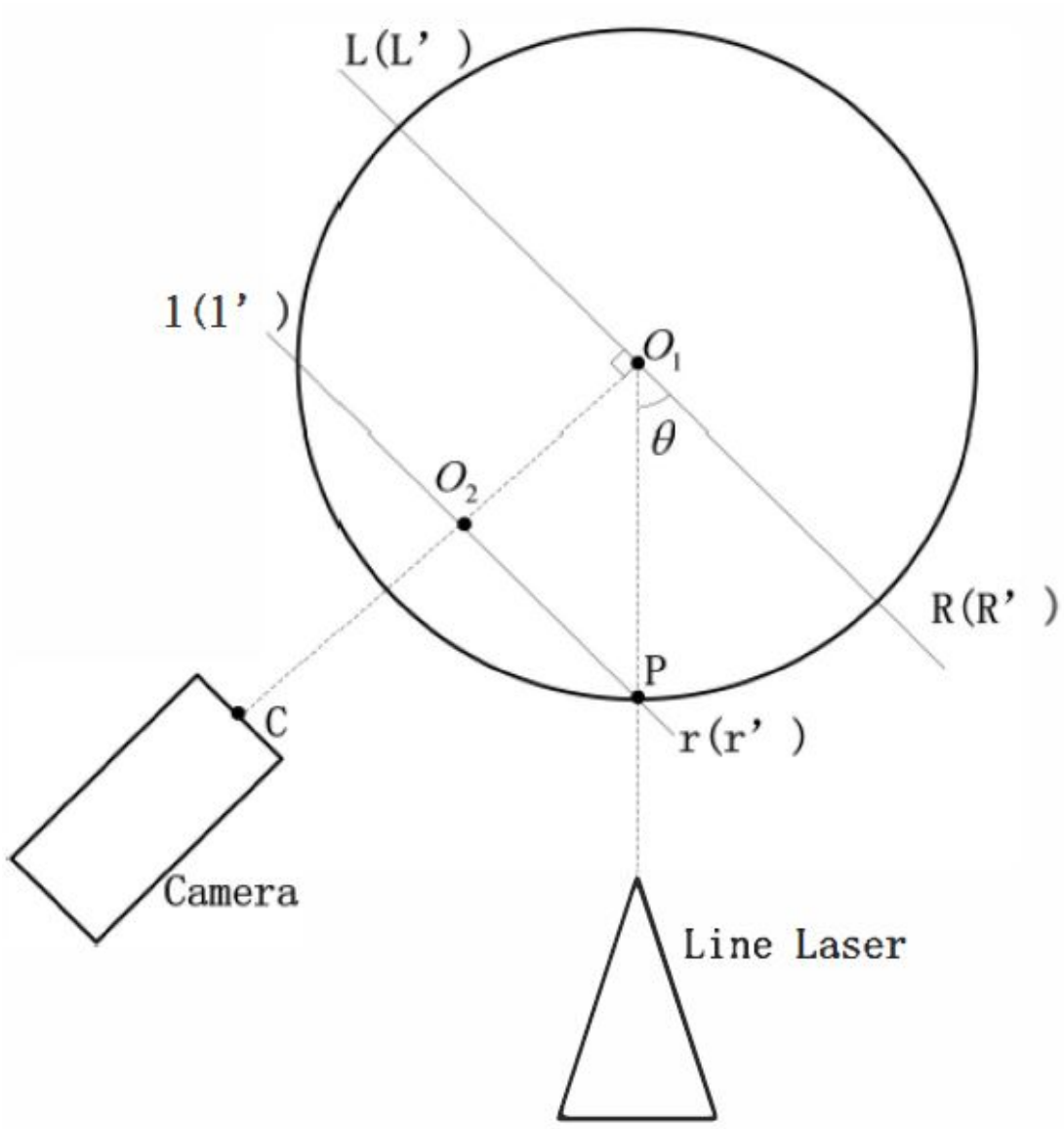

Figure 2.8: Geometric model of camera and laser position [14]

from [14]. Results proved greatly successful, operation time took less than 30 seconds and produced a highly detailed 3D scan. In the case of the zinc die cast, less data points were required for the surface point cloud and therefore would majorly reduce the run time.

Paper [13] introduces a structured light calibration algorithm used to detect

Figure 2.9: Output results of 3D scan of bottle [14]

weld seams for an arc welding robotic arm. By utilising the principal triangulation, the precise location of a weld seam is detectable. A calibration algorithm is delivered along with the structured light vision set–up including a *charge–couple device* (CCD) camera, laser line and calibration chessboard.

Paper [16] introduced a method that could be used to detect a laser spot emitted from a laser pointer. The method was structured in such a way that the detection of either a red or green laser dot was possible in real time. Laser dot detection was useful for generation of a rough surface point cloud of the zinc die cast outputs. Detecting a laser dot was the first step in detecting a laser line which could be used for structured light incorporating a laser line as a plane of light. An algorithm was introduced to demonstrate the effectiveness of detecting a laser dot.

Challenges arose when detecting the position of the laser dot including camera resolution and fps, interference with other projections, laser power and changing light conditions throughout the application. The greatest challenge

was the sensitivity of the camera chip. The camera chip was less sensitive than the human eye, the intensity of the laser dot seen by the image sensor gave misleading information. An instance occurred when the image sensor was looking at a red laser dot. A human would clearly interpret the laser dot as being red, the camera saw laser dot as white. This was caused when the camera chip was saturated with light. Its intensity value became set to the highest value, for all colours this was white.

Incorrect detection, merging of laser dots or lack of laser dot recognition occurred when the challenge of camera chip saturation was not dealt with correctly. It was recommended that the exposure of the camera was reduced so that the laser dot returned the highest value of intensity in the image sensor. When the laser dot produced an intensity value greater than 30% of the maximum, the exposure was reduced.

With the exposure settings complete, all pixels within the top 5% of the value threshold were assumed to be light intensities caused by the laser dot. If the total pixel count exceeded 100, then remaining steps were skipped to maintain real–time operation. These pixels were grouped together. With the groups established (laser dot locations), the centroids of each group were found and assumed to be the position of each laser dot.

The following list gave the steps of the algorithm used to establish laser dot locations taken from [16]:

1. Capture image from the camera,
2. convert image from RGB to HSV,
3. check the exposition of camera. If it is not suitable make sure corrections are made and go to Step 2,

4. obtain maximum value of intensity (value) channel and find the laser dot pixels,
5. if the number of pixels exceed 100, go to Step 2,
6. group the pixels,
   (a) compute diameter and centroid,
   (b) recognise laser dot colour and add to the local list,
7. update global laser dot list from local list,
8. go to Step 1.

Time of Flight

[12] gave a technical introduction to the 3D vision technique; ToF. ToF cameras were a robust solution to a wide range of computer vision challenges while having moderate to high accuracy and high frame rates. The paper contents included ToF basics, comparisons and finally the areas of industry that most gain from ToF technology.

Light is emitted in a modulated fashion by the ToF camera at near infrared wavelengths around 850nm. Modulated wave patterns are reflected off the environment in view of the camera. A phase shift is seen between the emitted light and the reflected light. This measured phase shift translates to a distance that can be expressed as a point cloud of 3D data points.

ToF technology typically uses one of two phase shift detection methods. Where the emitted light is either pulsed or modulated using a continuous sine or square wave.

During pulsed modulation, light is emitted for dt. Q_1 and Q_2 represent the electrical charge of two out of phase windows. The Basler ToF camera [31] is a

popular industry option based upon this method. Equation for distance from [12] where c equals the speed of light:

$$d = \frac{c\,dt}{2}\left(\frac{Q_2}{Q_1+Q_2}\right) \tag{2.8}$$

The continuous wave method was slightly more complex than the pulsed method. More than one sample per data point allowed for lowered system gain. Attenuation and amplification of data are reduced. The equation for distance is displayed where Q_1, Q_2, Q_3 and Q_4 are the four charge samples phase stepped at $90°$:

$$d = \frac{c}{4\pi f}\arctan\left(\frac{Q_3-Q_4}{Q1-Q2}\right) \tag{2.9}$$

A point cloud is produced by the ToF camera where every pixel on a 2D image is assigned a z distance value, producing a depth map. Grey scale images are typically used to represent the depth map where higher intensity pixels represent close by objects and low intensity pixel values represent far away objects. The depth map can alternatively be viewed as a 3D matrix of data.

The ToF camera was compared with stereo vision and structured lighting in Fig. 2.10. It was clear that ToF cameras had many benefits particularly in regards to:

1. Compactness,
2. response time,
3. poorly lit environments,
4. software complexity.

CONSIDERATIONS	STEREO VISION	STRUCTURED-LIGHT	TIME-OF-FLIGHT (TOF)
Software Complexity	High	Medium	Low
Material Cost	Low	High	Medium
Compactness	Low	High	Low
Response Time	Medium	Slow	Fast
Depth Accuracy	Low	High	Medium
Low-Light Performance	Weak	Good	Good
Bright-Light Performance	Good	Weak	Good
Power Consumption	Low	Medium	Scalable
Range	Limited	Scalable	Scalable
APPLICATIONS			
Game		X	X
3D Movies	X		
3D Scanning		X	X
User Interface Control			X
Augmented Reality	X		X

Figure 2.10: Comparing 3D vision techniques from [12]

A method for 3D reconstruction is proposed in [15] utilising the ToF camera. The technology is regarded in the paper as smart and fast, but not precise or robust. Calibration, filter and the iterative closest point algorithm are utilised to improve system robustness and accuracy.

Autonomous robots are reliant on high quality mapping of their relative environment. When a robots position in its domain has been understood, reconstruction of its environment is possible. If map data is provided, then detecting a robots position is possible. Combining the two approaches is called *simultaneous localization and mapping* (SLAM).

Stereo vision and laser scanning have been used for SLAM applications due to greater levels of accuracy and range. ToF cameras are more superior when it comes to motion detection due to greater frame rates. They can also produce quality 3D reconstruction irrespective of varying lighting and surface conditions. [15] finds a way to 'bridge the gap' by incorporating ToF into SLAM

application.

Quality of ToF 3D map output has its limitations due to error. Inherent errors within the principal of measurement are fixed. These can be caused by poor signal to noise ratio, high levels of near infrared inter–reflecting within the scene and the scattered illumination across the sensor lens. Inter-reflections can cause corner shapes to show as rounded in the 3D reconstruction. While the scattered illumination caused variation in measurements of highly lit objects. Two of three explained are unknown prior to taking the initial observation of the sensor environment.

Other foreseeable errors can be minimised when calibrated correctly. These include an error introduced by distance, amplitude and sensor chip pixel to signal generator distance. Through calibration, the three mentioned errors can be minimised. Distance measurements assume that light travels in a pure sinusoidal pattern, where true light patters are near sinusoidal.

A point cloud representing the 3D reconstruction is taken before the following steps are applied:

1. Apply filter,
2. generate 3D map,
3. accumulate and relax error,
4. apply final filter.

The initial filtering process deals with low lighting and occlusion errors. *Jump edge filtering* in [6] is used to remove jump edges caused where one shape is transitioning to another within the image. Large frame rates allow the use for the ICP algorithm [1]. Minimisation with least squares is iterated. Recording

two corresponding pairs in a list before computing the error in Eq. 2.10:

$$E(\mathbf{R}, \mathbf{t}) = \sum_{i=1}^{N} || \mathbf{m_i} - (\mathbf{Rd_i} + \mathbf{t}) ||^2 \tag{2.10}$$

Accumulated error caused by iterative ICP approach along with poor accuracy is minimised. Distances between cloud data points when correlated by an edge are minimised in Eq. 2.11:

$$E(\mathbf{R}, \mathbf{t}) = \sum_{j \to k} \sum_{i=1}^{N} || (\mathbf{R}_j \mathbf{p}_{j,i} + \mathbf{t}_j) - (\mathbf{R}_k \mathbf{p}_{k,i} + \mathbf{t}_k) ||^2 \tag{2.11}$$

The final filter step refines the 3D reconstruction by approximating patches and fitting them to neighbourhoods of point data along with reducing sparse data. Due to the high noise seen in ToF sensors, some surfaces are output with thickness in the tens of millimetres. PCA is used to find surface normals and shift pixels along the calculated normals.

[8] introduces the concept of fusing ToF point cloud data with stereo vision disparity maps to improve the 3D reconstruction of a relevant scene. Results of the data fusion show highly populated disparity images where the output has higher accuracy that using only a ToF camera.

3D stereo images are computed using pixels that are corresponding to each other across a pair of images x_l and x_r. Disparity maps $D_l(x_l\, x_r)$ are produced by finding the difference between the two pixel locations.

A *SwissRanger SR3000* ToF sensor is used for the experiment has an amplitude modulated light source as seen in Fig. 2.11.

When emitted light signals are returned back to the ToF sensor chip, the sensor camera rebuilds the observed scene. An intensity image in grey scale is produced based on the returning signals amplitude along with a range image

Figure 2.11: The SwissRanger SR3000 and its output: the intensity image I_t and the range image Z_t from [8]

where every pixel represents a depth value based on the offset of the signal phase. High speed is possible, while low quality 3D reconstruction is seen.

One fusion method explored in [8] involves fusing the ToF range measurement Z_t with a standard 2D image I_s. While the second method uses the ToF point cloud output with a stereo image pair to find disparity between ToF range Z_t and stereo image points I_l and I_r.

[29] introduces an alternative method to fuse ToF point cloud data with stereo images. ToF cameras alone are ideal for real time conditions but lack the accuracy. Where stereo vision has a higher accuracy but does not handle some conditions like low lighting. By fusing the two, a more versatile 3D image sensor may be possible.

A calibration scheme triangulating points X_{st} in 3D space is explained. Figure 2.12 depicts the geometric set–up of the ToF camera along with two CCD cameras for stereo vision.

The system proved to output depth maps within a range of 1m to an accuracy of 5mm. A 300% increase in system accuracy was seen when compared a standard ToF depth map output. In the case of complex environments with non–ideal illumination, inter–reflections and texture error was reduced 50%.

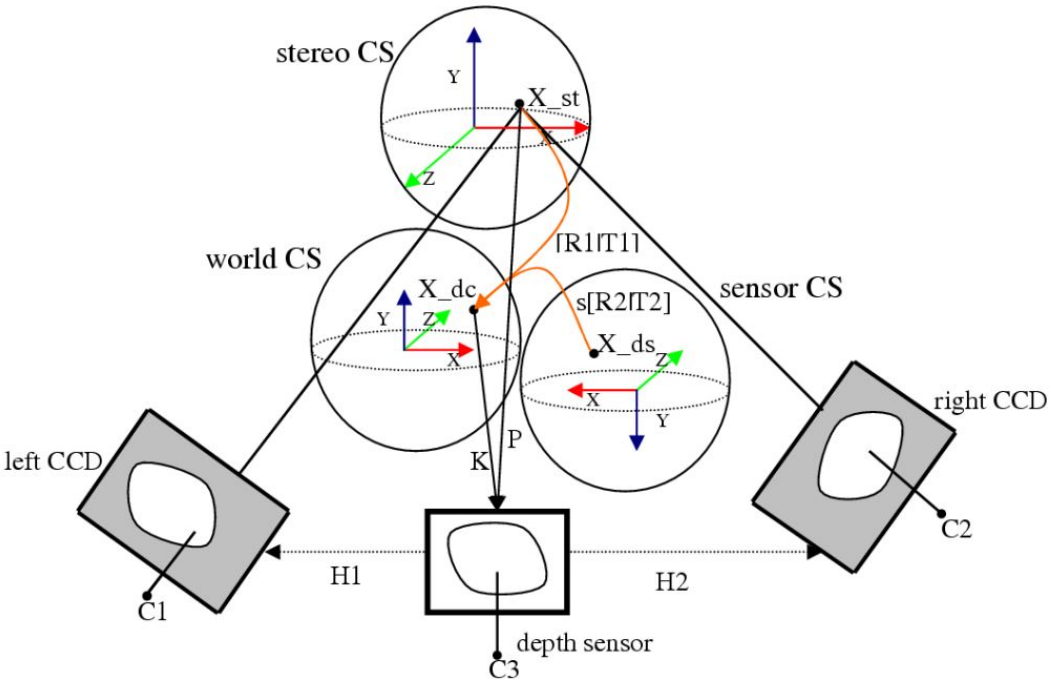

Figure 2.12: Diagram of epipolar geometry of relationship between ToF depth sensor and stereo CCD cameras from [29]

Chapter 3

The Proposed System

The following chapter delivers a method for the proposed system. The current state of the manual/repeatable system is introduced. Four main machine vision tasks are included – determination of cast face, robot pick location, object orientation along with an error checking process.

3.1 The Current State

The zinc die cast production at Assa Abloy is responsible for producing door entry products such as door handles or locking mechanisms. Zinc ingots are melted at high temperatures before being filled into a die. The die is modular and has dual tooling, where two separate products can be produced together in the same batch.

Once the die tooling has been filled with molten zinc, the hardened cast is output from the machine onto the conveyor belt. The zinc die cast output may appear face up or down and orientated in any direction across 360° on the conveyor.

Currently, each zinc die cast machine requires its own dedicated operator. When the zinc die cast output appears on the conveyor, a worker picks up the object and manually separates the products from the waste sprue Fig. 1.2. The products are placed in their assigned bins while the waste is returned to a recycle conveyor, sending the rejected zinc material back to the melting pot.

3.2 The Suggested Solution

In order to answer the main research question 'how can the position, orientation and identification of zinc cast sprues be found using predominantly machine vision techniques?', a high level process flow was created. The basis of the objective ensured that a repeatable/manual process, undertaken by a human worker could be automated predominantly using machine vision techniques. Additional supplementary mechanical/electronic processes were required in the design for further implementation by Assa Abloy engineers.

The complete automation solution was broken down into four distinct machine vision processes:

1. Determination of the cast face,
2. detection of the robot pick location,
3. determination of the object orientation,
4. a quality control measure to determine if the correct tooling combination was being used.

The first machine vision process in the solution was the determination of the cast face to decide whether incoming zinc die cast outputs were facing up or down on the conveyor belt. The machine output die casts that fell from

the ejector pins for $1.0\,\text{m}$ before landing on the moving conveyor belt. The conveyor system relocated the die cast from the machine to the next stage of the process. Depending on the weight distribution and physical size of the die cast, it fell either face up or face down on the conveyor.

An initial step to determine whether the cast appearing on the conveyor was face up or down was incorporated. A mechanical 180° flipping process was used when the incoming die cast appeared face down which allowed for the remaining machine vision techniques to be implemented robustly.

The second machine vision process was the detection of the robot pick location. This process was to be executed on face up die casts only. An (x, y) 2D coordinate was determined, referring to the centre of the die cast sprue cone base. This was a repeatable shape across all cast tooling combinations and a suitable location to grip the the object.

The third machine vision process was the determination of die cast orientation. Calculating the robot picking location alone was not enough. By finding the angle of orientation about the $x - -y$ plane when the die cast rested on the conveyor belt, the robot arm was able to re–orientate the item by pivoting its wrist. Once the piece had been successfully gripped and orientated into a fixed/known position, it was ready for manipulation.

The fourth and final process was a quality control measure. It was necessary to test whether the correct products were being manufactured in real time. Assa Abloy typically produced in the quantities of the thousands. It was crucial to detect incorrectly assigned product tools early to avoid loss of company profit. One or two tools were typically ran at any given time and therefore each tool was individually tested to check if it were correct or not. This process was implemented last as it required information acquired by the previous processes.

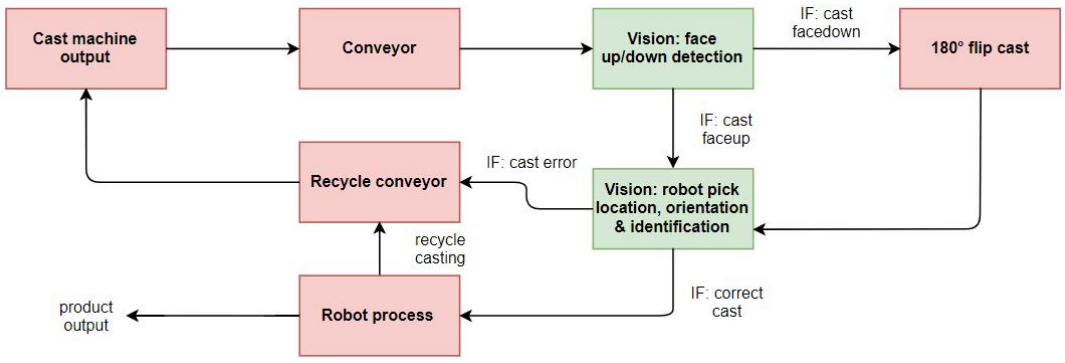

Figure 3.1: Proposed high level solution flow diagram

All four machine vision process were tied together in a high level system design as outlined in Fig. 3.1 where green boxes represented machine vision processes and red boxes represented mechanical processes. The die cast was output from the casting machine where it landed on the moving conveyor. The item was relocated out of the machine and brought to the first machine vision process where the cast face was tested as up or down. If the cast was face down, then a $180\,^{\circ}$ flipping process occurred. Next, the robot (x, y) pick location coordinates and cast angle of orientation were determined to allow for the robot arm to grip and fix the cast in a known location. The quality control measure was implemented to determine if the correct products were being produced. Finally, the robot arm could reach in to the conveyor system and pick up the part before applying the necessary manipulations. If the die cast output happened to be incorrect, the robot arm did not reach in to pick the object. The cast fell off the end of the conveyor into a recycle process. The recycle process delivered the incorrect cast back into the melting pot where it was heated into molten zinc and reused.

Figure 3.2: Physical set–up of Google Nexus camera viewing Sample 1

3.3 Camera Set-up

To create a reliable data set of images, a camera set–up was used. Images of the four sample die casts were taken. The described camera set–up was used for experiments; robot pick location, object orientation and error checking.

A Google Nexus 5X rear camera was set at 410 mm height from a bench top Fig. 3.2. The surface was covered with a 3 mm thick mat black material of dimensions 317 mm x 270 mm used to simulate the mat black conveyor. The camera was secured to a piece of window framing acting as a truss between two standing trays. During image acquisition, natural lighting was reduced and a soft white incandescent lamp was aimed at the ceiling to provide a low level diffused light across the die cast samples.

There was a 407 mm distance set between the image sensor and the surface. The object of interest lay on the surface while the single camera took images.

Lighting was fixed to low levels of incandescent light with minimal natural

light present. Curtains and blinds in the room were drawn and a $70\,\mathrm{W}$ incandescent lamp was poised at the ceiling of the room, distributing light evenly through the space. This reduced the shadows seen on the images.

The camera set–up was different for the 3D scanner experiment and can be seen Sec. 3.4.

3.4 Determination of Cast Face: Up or Down

Incoming die cast outputs could appear on the conveyor belt either in the face up or face down position. To conduct machine vision analysis of the die cast outputs, it was necessary for the objects to be lying face up on the conveyor before being assessed for robot pick location, orientation and correct cast. The circular base feature of the cone sprue seen in Fig. 1.1 appeared as either a circle or an ellipse with near equal major and minor axis when the object was in the face up position. Robust pick up by the robot gripping tool was possible when dealing with face up casts.

For any incoming die cast outputs appearing on the conveyor in the face down position, a 180° mechanical flipping process would ensure all die cast outputs appeared in the face up position for further machine vision analysis.

A Basler ToF camera [31] was set–up to produce a point cloud of the incoming die cast outputs for a rough surface reconstruction as seen in Fig. 3.3. By understanding the approximate shape of the incoming cast, a method was produced to determine if the correct cast was being produced or not.

For single tool die casts like App. A.1 and App. A.2, the expected height of the object in both the face up and down position was known. The two values representing the face up height and the face down height were discreetly different for all samples.

Figure 3.3: Physical set–up of Basler ToF camera viewing Sample 2

Using the Basler ToF camera, a rough point cloud was acquired for all four samples provided by the company in both the face up and face down positions. Using Matlab, the x and y data were discarded, only considering the z height values. Once the z height values were sorted in order, the mean of the top 20% for all tests was recorded. By taking the mean of the top 20%, around 1500 data points were included. If the true maximum value was taken (a single data point), the system would fail as a result of a single outlier.

For double tool die casts like App. A.3 and App. A.4, all face up tool height values lay in a lower region while all face down tool heights lay in an upper region. The gap between the two value ranges was distinct. As for single tool castings, the mean of the top 20% of z height values was recorded:

1. Acquire 3D point cloud using Basler ToF,
2. use CloudCompare to produce a .txt file containing the point cloud data,
3. in Matlab, isolate the z height data and find the mean of the top 20%,
4. compare against expected value with tolerance for the correct tooling.

3.5 Detecting Robot Pick Location

Once all incoming casts were known to be face up, the second machine vision process could be applied. To command a 6–axis robot arm to accurately grip the face up cast lying on the conveyor belt, the detection of the (x, y) picking coordinates were required. The base circle feature on the die cast sprue was chosen as the centre of the gripping location on the $x - -y$ plane. An ideal feature to target because it was repeatable and consistent across all possible die cast outputs. The shape was not similar to any other seen across the varying tooling combinations.

Using the camera set–up described in Sec. 3.3, a high resolution image was acquired. It was necessary to create high contrast between the back ground (conveyor belt) and the object of interest (zinc die cast). To achieve this for a data set of sample images, a 'mat' black surfaced rubber mat of dimensions $317\,\text{mm}x270\,\text{mm}$ was chosen as the background. The die cast object was placed on top in a resting stationary position. The chosen size of the background mat

exceeded the maximum length of the longest zinc die cast product within the sample set.

Detecting the repeatable circular base feature of the die cast cone sprue required a robust circle detector. From [11], we knew that a circle could be characterised by a radius r and a circle point centre (x_c, y_c). The Hough circle transform was used to find the repeatable circular feature within the image to determine the robot picking location. This particular transform was seen as advantageous due to its ability to handle varying noise data and its ease of adjustment. For Sec. 3.3, the expected radius of the repeatable circular feature of the zinc die cast sprue base was 84 pixels. From [11], the Hough circle transform required a range estimation for the radius value.

Software was developed to detect the repeatable circular feature and output the pixel coordinates of the circles centre using OpenCV in a Python programming environment. Image processing techniques were necessary to implement prior to applying the circle Hough transform to ensure robust detection. The order of operation that the developed circle detector used to output the centre gripping location for the 6–axis robot arm:

1. Input grey image,
2. blur image,
3. apply thresholding,
4. apply edge detection,
5. find the largest contour,
6. apply circle Hough transform,
7. output circle centre.

The described order allowed for a raw input image to be processed to output the gripping location in terms of pixels on the $x - -y$ plane. First, the grey image was called and immediately blurred using a Gaussian blur. This reduced noise in the image and run time. Next, the blurred grey image was binarised using thresholding to separate the background (as black) with the object (as white). Once the object was separated from the background, the Canny edge detector was applied to map all edges between the black background and the white object. The contours were then found using the edge map and the largest contour was isolated. The largest contour was the perimeter of the object. Using the map of the largest contour, circle Hough transform was applied to detect the repeatable circular feature. The circle centre was output as the robot gripping pixel location as an (x, y) coordinate.

3.6 Determination of Cast Angle of Orientation

With the robot picking location established, the 6–axis robot arm could grip the die cast around the repeatable circular sprue cone feature at a known coordinate on the $x - -y$ plane. Alone, this was enough for the 6–axis arm to successfully grip and pick up the object. Without knowing the die casts angle of orientation as it lay on the conveyor, there was no way to fix the object in a known position of reference. There was no robust way to work on the given object after it had been pickup up.

PCA as outlined in [22] proved to be a suitable approach to finding the angle of orientation of the die casts lying flat on the conveyor belt. The camera set–up in Sec. 3.3 was arranged for the sample images taken for the experiment.

OpenCV library functions were used in the Python programming environment to develop the necessary software for the detection of cast orientation.

Image processing techniques were necessary to implement prior to applying PCA. The original image was reduced using information taken from the circle detection process to simplify the challenge. This allowed for the robust detection of die cast angle of orientation. This angle value was used to re–orientate the randomly orientated object by rotating the wrist of the robot. The order of operation of the developed angle of orientation detector used:

1. Input grey image,
2. apply thresholding,
3. apply edge detection,
4. find largest contour,
5. apply PCA to the largest contour,
6. output angle of orientation.

The described order allowed for a raw input image to be processed to output the angle of orientation relative to the reference position in terms of radians. First, the grey image was called. The grey image was binarised using thresholding to separate the background (as black) with the object (as white). Once the object was separated from the background, the contours were then found using the edge map and the largest of all contours was isolated. The largest contour of all contours represented the perimeter of the object. Using the map of the largest contour, PCA was applied to determine the angle of orientation of the zinc die cast. The angle of orientation was output in terms of radians.

3.7 Quality Control Process: Checking Correct Cast

Once the robots gripping location on the $x - -y$ plane along with angle of orientation had been established for the zinc die cast output lying face up on the conveyor had been discovered, the machine vision system had one final task to complete. To ensure that the correct tooling combination had been set on the die cast machine by the machine operator, a quality control measure was implemented. Due to large customer orders in the thousands, finding an error minimise loss of profit.

Due to the nature of Assa Abloys die cast tooling arrangement, two separate tools could be used to produce die casting on the same sprue. This was a modulated approach and required the machine vision system to act accordingly. The quality control system dealt with the die cast outputs in a modulated fashion. The image was split into its two respective tools as seen in Fig. 3.4 and Fig. 3.5. Left hand tool images were rotated 180° so the x-direction was referenced consistently for all cropped images.

Once the images had successfully been split into their two separate tools, they could be analysed. The system output a binary result for each image. Either the correct or incorrect tool had been seen by the system. If either of the two tools set–up in the die cast machine were incorrect by not matching their expected signatures, then the machine operator could be notified to modify the tooling combination.

The order of process to check for the correct tooling combination of die cast outputs:

1. Crop orientated image into two separate tools about the centre of the sprue,

2. blur each image,

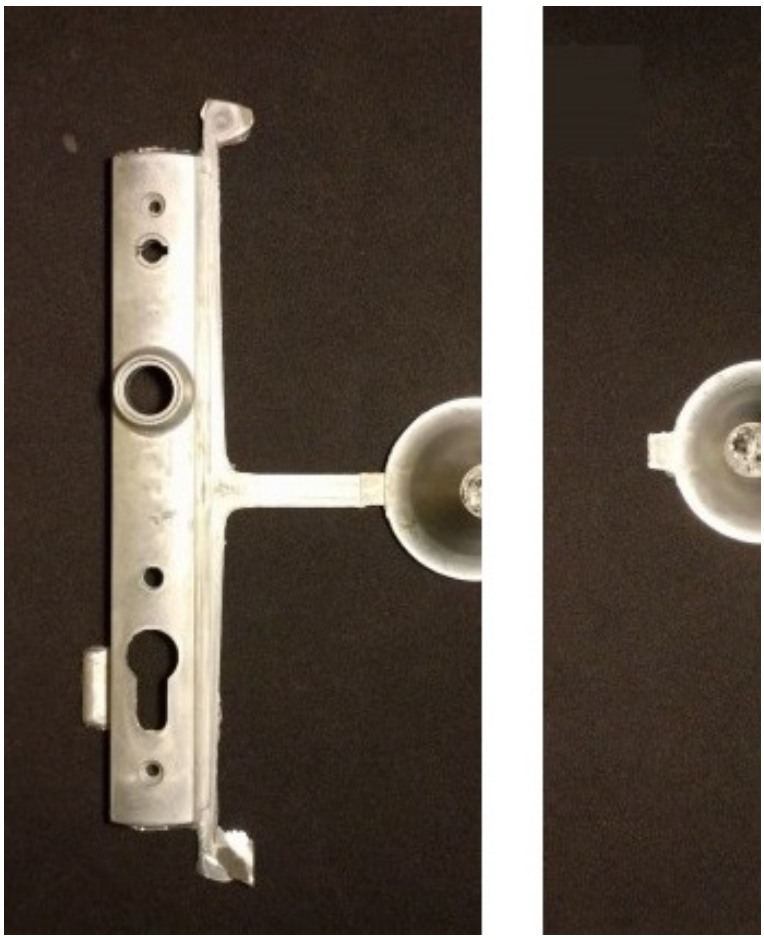

Figure 3.4: Sample 1 output split into its left and right tools

3. apply thresholding to each image,

4. apply edge detection to each image,

5. find the largest contour of each image,

6. take the convex hull for both images and compare to the expected values,

7. find the area and centroid of each images convex hull and compare with expected values.

The described order allowed for an orientated image of a zinc die cast output lying on the conveyor to be measured to check for the correct tooling.

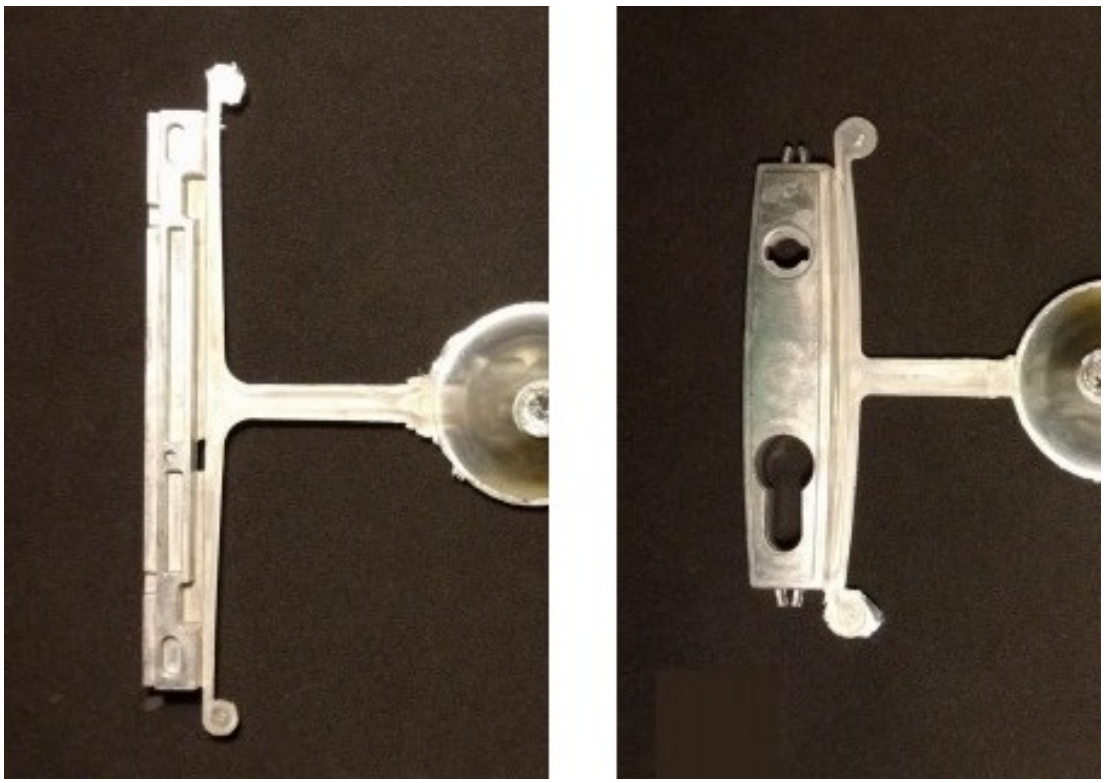

Figure 3.5: Sample 3 output split into its left and right tools

Firstly, the image was cropped [37] into two separate images at the centre of the cone sprue (robot gripping location). Once each tool had been separated, the following processes were performed in parallel for both images. Secondly, each image was blurred using Gaussian blur [46] to reduce noise. Thirdly, the image was binarised using thresholding [48] to separate the background from the object. Fourth, Canny edge detection [33] was implemented to find the edges of the binary image. Once the edges had been found, the largest contour [35] was determined and gave the perimeter of each tool. Having the perimeter map established, the convex hull [36] was taken for each image. This convex hull image was then used for a series of three tests. The outcomes of these tests were then compared with the expected values given the tool signature. The tests included the pixel distance of the convex hull perimeter, the pixel

area of the convex hull along with the centroid coordinates of the convex hull on the $x - -y$ plane. The area and centroid were found using moments [43].

Chapter 4

Evaluation of Results

The following chapter delivers the results for proof of concept using the proposed system. Accuracy and robustness is considered when analysing data and error correction methods are proposed.

4.1 Determination of Cast Face: Up or Down

To determine whether a die cast output was lying face up or down on the conveyor, a ToF camera was used to acquire a point cloud of the object. For each of the four samples provided by Assa Abloy, z_v the mean of the top 20% of z height values was recorded. Where the z represents the direction orthogonal to the $x - -y$ plane, passing through the conveyor and the camera lens.

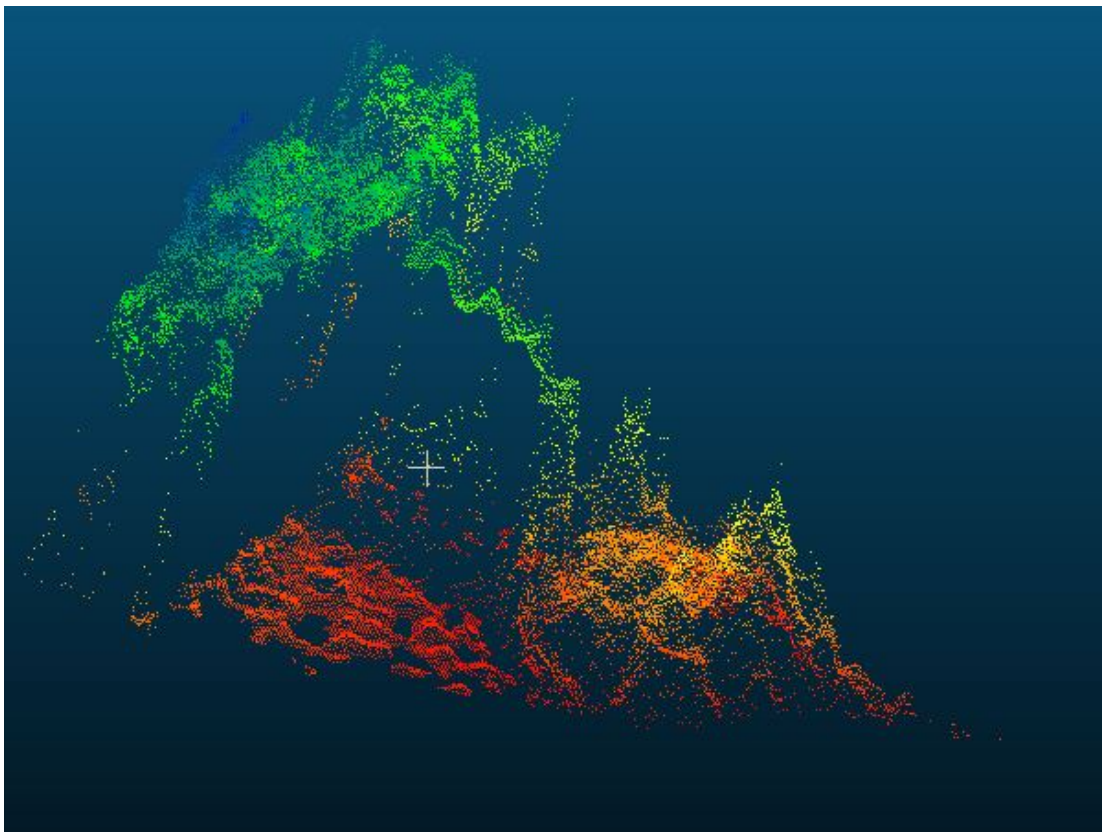

Figure 4.1: CloudCompare point cloud visualisation of Sample 3 face down

Table 4.1: Finding the mean of the top 20% z cloud data.

Cast sample	Face	No. of tools	z_v (mm)
1	Up	Single	20.14
1	Down	Single	34.21
2	Up	Single	36.33
2	Down	Single	30.21
3	Up	Double	20.44
3	Down	Double	77.00
4	Up	Double	16.84
4	Down	Double	52.80

4.1.1 System Output

Figure 4.1 showed the point cloud produced by *CloudCompare* using the Basler ToF camera. See App. A.4 for more point cloud screenshots. The green representing close values of z height to the camera and red representing far away

values of z height when viewing the object from the top–down.

Table 4.1 presented the results where z_v was the mean of the top 20% of z height values for each sample in the face up and down position.

4.1.2 Evaluation of Results

For single tooled die casts, the expected z_v value for a specific tool selection was known. While the die cast machine was producing a particular single tool output, the vision system received the information regarding its characteristics. Measured z_v values from Tab. 4.1 were compared to expected values in real time.

When the measured value of z_v came within $\pm\, 2.00$mm of the expected face up or down value, the part was assumed to be face up or down respectively. This triggered the mechanical flipping process when a cast was presented face down.

For double tooled die casts, the z_v value range for face up and down objects was distinct, regardless of the two tools selected for production. Face down die casts generally had much larger z_v values than face up casts. From Tab. 4.1, double face up sample z_v values were no greater than 20.44 mm while double face down samples were no less than 52.80 mm. A second test proved that the highest possible mean of the top 5% of the z data points was no more than 30.92 mm for the double tooled cast samples face down. Thresholding the z_v value at 40 mm allowed for a clear distinction between face up and down casts. All double tooled die casts which measured z_v as under 40 mm were assumed to be face up. While all double tooled die castings where z_v measured greater than 40 mm were assumed to be face down.

4.1.3 Ellipse Method

To determine whether the die cast output was lying face up or down on the conveyor, an additional method was considered. Instead of using a ToF camera to reconstruct the surface of the cast output and find its face, an ellipse was fit over the repeatable circular cone shape.

The ellipse was fit for both face up and down positions, so the difference in major and minor axis could indicate whether the cast lay face up or down on the conveyor. Figure 4.2 showed a 2D grey image of die cast Sample 3 facing down with an ellipse fit from [24].

The distinction between major and minor axis clearly showed that the sample was lying face down on the conveyor. When major and minor axis were similar as seen in Fig. 4.6, the system recognised the die cast as face up.

A major advantage with the ellipse method is the use of 2D vision. With robust ellipse detection, the ToF camera can be discarded. This would simplify the system, only requiring one vision sensor; a 2D camera.

Angle of orientation for the face down cast was also visible in Fig. 4.2 where the major and minor axis were perpendicular and parallel respectively. For face up die cast outputs seen in Fig. 4.6, orientation was less distinguishable.

4.2 Detecting Robot Pick Location

The outputs of the system used to detect the robot picking location by finding the centre of the repeatable circular pattern across all die cast outputs were presented. The algorithm was run with four die cast samples provided by Assa Abloy. The $x - -y$ plane picking location was output for each sample and then compared against the expected error values.

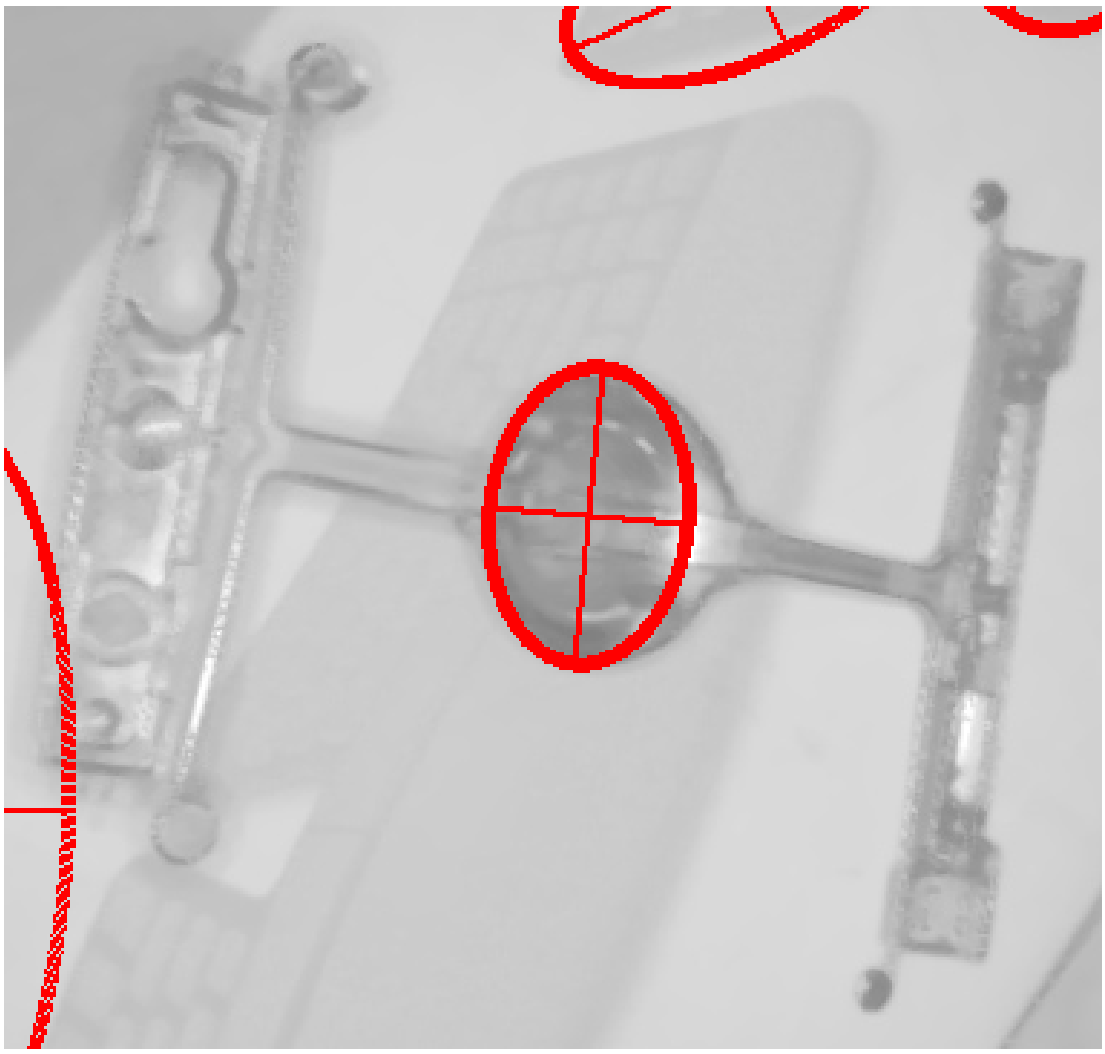

Figure 4.2: Sample 3 face down with ellipse fitting from [3]

All images were taken with the described camera set–up using the Google Nexus 5X rear camera. The images were reduced to 320x427 before being run through the algorithm.

4.2.1 System Output

In Tab. 4.2, the output results were compared to the expected results for the robot picking location on the $x - -y$ plane. The die cast Sample 1 could be

Table 4.2: Robot pick location output results for given samples.

Cast output sample	Expected x	Expected y	Output x	Output y
1	217	221	218	220
2	174	224	178	224
3	154	220	150	226
4	162	244	158	248

seen in Fig. 4.3, which clearly displayed where the Hough transform fitted its circle and then its centre of circle. Additional die cast Samples 1 – 4 can be seen in Appendix A. The x average error and total total system error:

$$X_{\varepsilon} = \frac{\sum_{i=1}^{n} \mid X_{p_{exp}} - X_{p_{out}} \mid}{n} \tag{4.1}$$

$$\varepsilon_{system} = \sqrt{X_{\varepsilon}^2 + Y_{\varepsilon}^2} \tag{4.2}$$

Using the notation,

$$\begin{aligned} x_{p_{exp}} &= x \quad \text{(expected pixel value)} \\ y_{p_{exp}} &= y \quad \text{(expected pixel value)} \\ x_{p_{out}} &= x \quad \text{(output pixel value)} \\ y_{p_{out}} &= y \quad \text{(output pixel value)} \end{aligned} \tag{4.3}$$

Eq. (4.1) derived with influence from [22] where $n = no.of\,samples$ were used to determine the average error of the x plane. The same was used to calculate Y_{ε} for the y plane. The average error was determined for both x and y plane for the *four* samples provided by the company.

By finding the average error of each axis on the $x - -y$ plane, Eq. (4.2) was used to find the average error of the system in pixels. With system output coordinate and actual values seen in Tab. 4.2, the average system error was $\varepsilon_{system} = 4.26$ pixels. With a maximum error of $\varepsilon_{max} = 7.21$ pixels.

Given the system having $1.38\,\frac{pixels}{mm}$, the total average system error was calculated as 3.08 mm. This implied that when a coordinate (x, y) was output by the system, the value of the ideal robot gripping location lay with a ± 3.08 mm radius circle 360° around.

4.2.2 Evaluation of Results

The system performed its required task of detecting the robot picking location for the 6–axis robot arm as an (x, y) coordinate in pixels. The system incurred some error and was accurate to ± 3.08 mm.

Throughout all four samples provided by the company App. A.1 A.2 A.3 A.4, the base circle feature of the cone sprue was detectable. Visually the centre of the Hough circle indicated that it was close to the true centre as seen in Fig. 4.4.

With the average error provided, the automation engineers at Assa Abloy confirmed that the error was well within the tolerance. A specialised gripping tool provided a vacuum force around the cone sprue. It allowed for errors greater than ± 10.00 mm. The concept of design could be seen in Fig. 4.5. This allowed for the cone sprue to be robustly gripped given the error specification. The tool itself was an inverted cone shape with wider walls to allow for the cone sprue to be forced into the tool. Once the cone sprue was inside the tool, the sides collapsed with pneumatic air pressure to allow for a tight grip of the object.

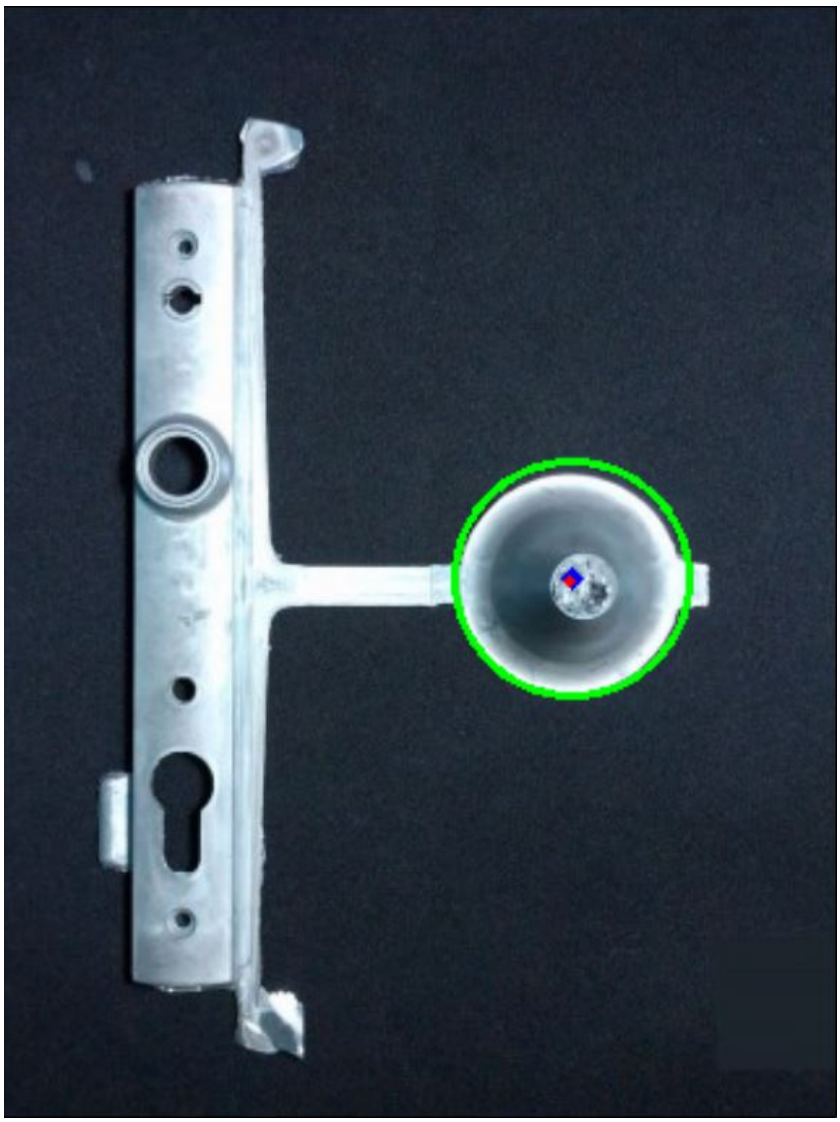

Figure 4.3: System output for die cast Sample 1

4.2.3 Ellipse Method

To determine the robot picking location, the centre of the repeatable circular cone sprue was needed. Fitting an ellipse over the circular cone shape allowed for the major and minor axis intersection to give the robot pick location.

A comparison of ellipse fitting methods including the proposed method from [24] showed how well an ellipse could be fit over Sample 1 & 3 Fig. 4.6.

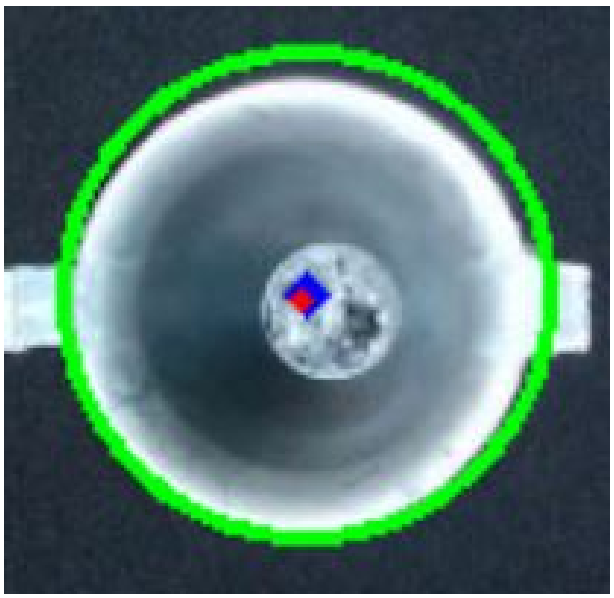

Figure 4.4: Zoomed system output for die cast Sample 1

4.3 Determination of Cast Angle of Orientation

A comparison of the determined angle of orientation of the zinc die cast outputs, with expected angles of orientation was considered. With the die cast output lying face up on the conveyor and the robot picking location known, a Python program was used to crop the image to 120 % width of the sprue circle diameter. The image was cropped in such a way to include a clear representation of the cone sprue, along with the direction of each of the cast sprue runners. Once the cone sprue had been cropped, the orientation of the perimeter of the object was taken.

All images were taken with the described camera set–up using the Google Nexus 5X rear camera and were reduced to 320x427 before executing with the algorithm. All measurements of radian angles in Tab. 4.3 were referenced from the unit circle seen in Fig. 4.7 where the 0° was on the right hand side of

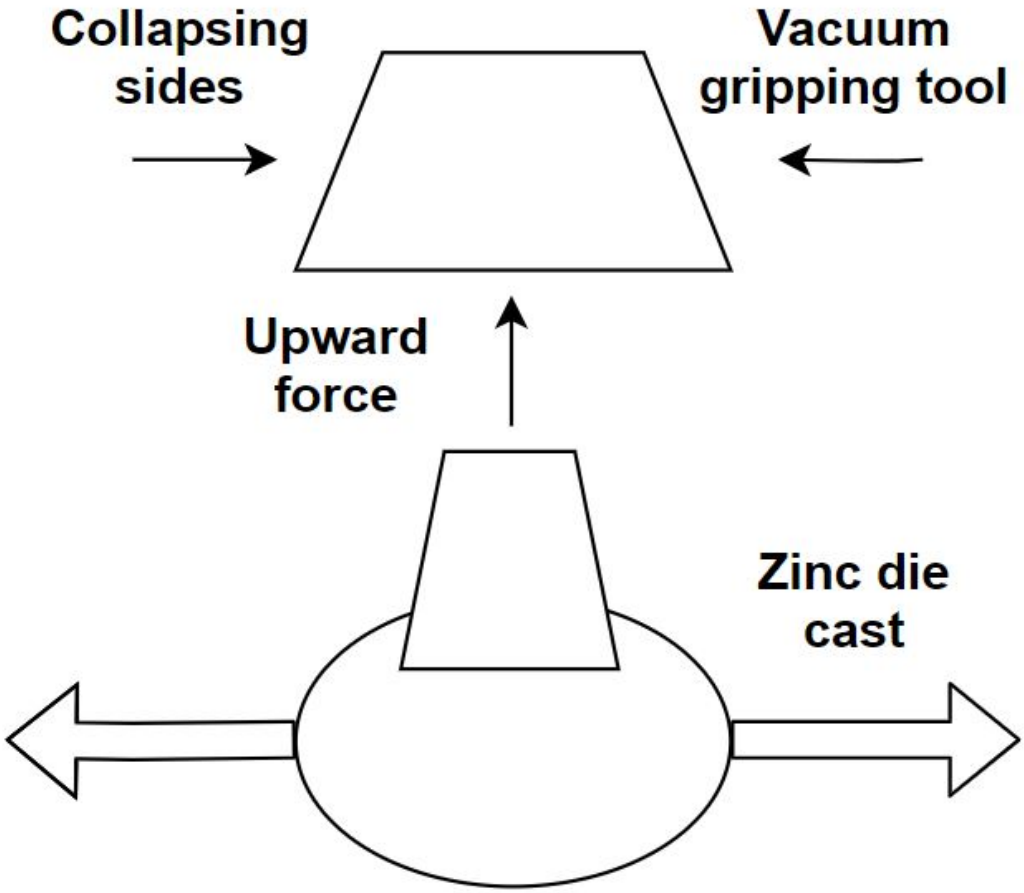

Figure 4.5: Display of vacuum gripper tool operation

the circle. Positive angles were clockwise and negative angles were counter–clockwise. Output radians were displayed from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$.

In Tab. 4.3, the four samples from App. A.5 A.6 A.7 A.8 were used for the experiment. These samples were identical to the four used in the previous experiment; where the robot picking location was determined. However each of the samples had been rotated to cover a range of orientation inputs.

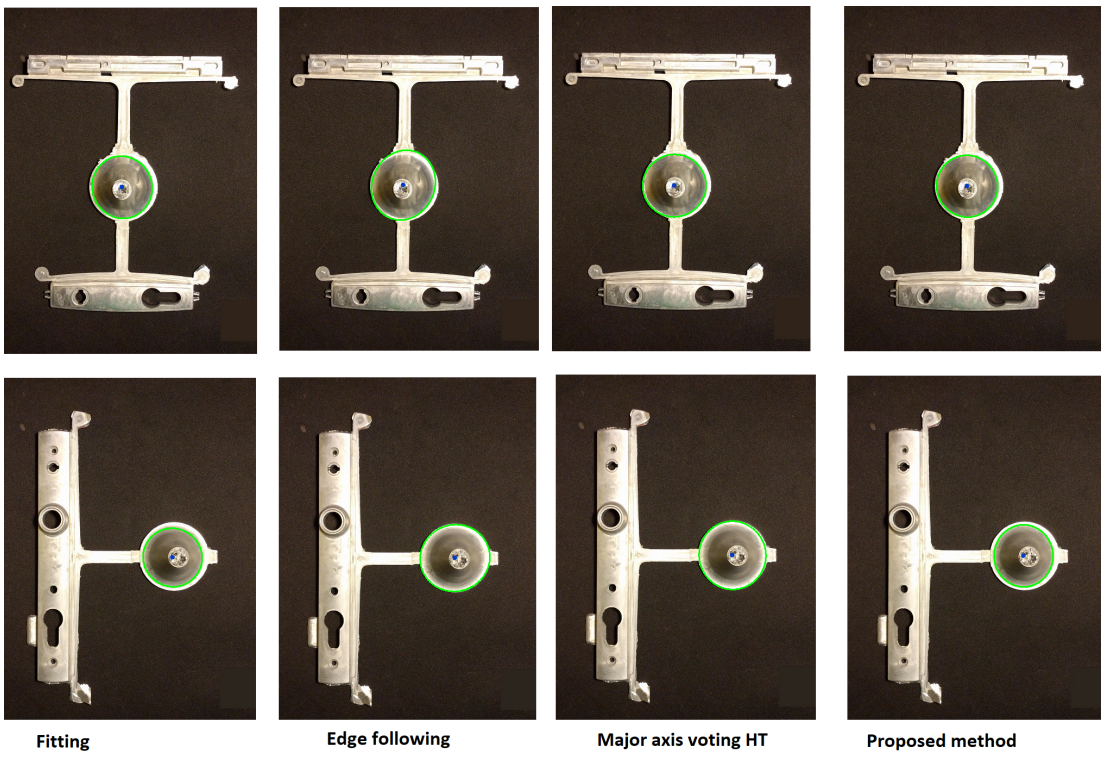

Figure 4.6: Comparing ellipse fitting algorithms on Sample 1 & 3 from [3]

Table 4.3: Die cast output orientation output results for given samples.

Cast sample	Expected (rads)	Output (rads)	Absolute error (rads)
1	0.19	-0.12	0.31
2	-0.12	-0.16	0.04
3	1.40	1.44	0.04
4	1.17	1.08	0.09

4.3.1 System Output

Table 4.3 displayed how well the algorithm handled the various cast samples. The samples yielding the lowest error were Sample 2 and 3 with an absolute error of 0.04 rads or $2.3°$. Sample 1 produced the highest absolute error at 0.31 rads or $17.8°$. The average error was found using the mean equation from [22] as 0.12 rads or $6.9°$.

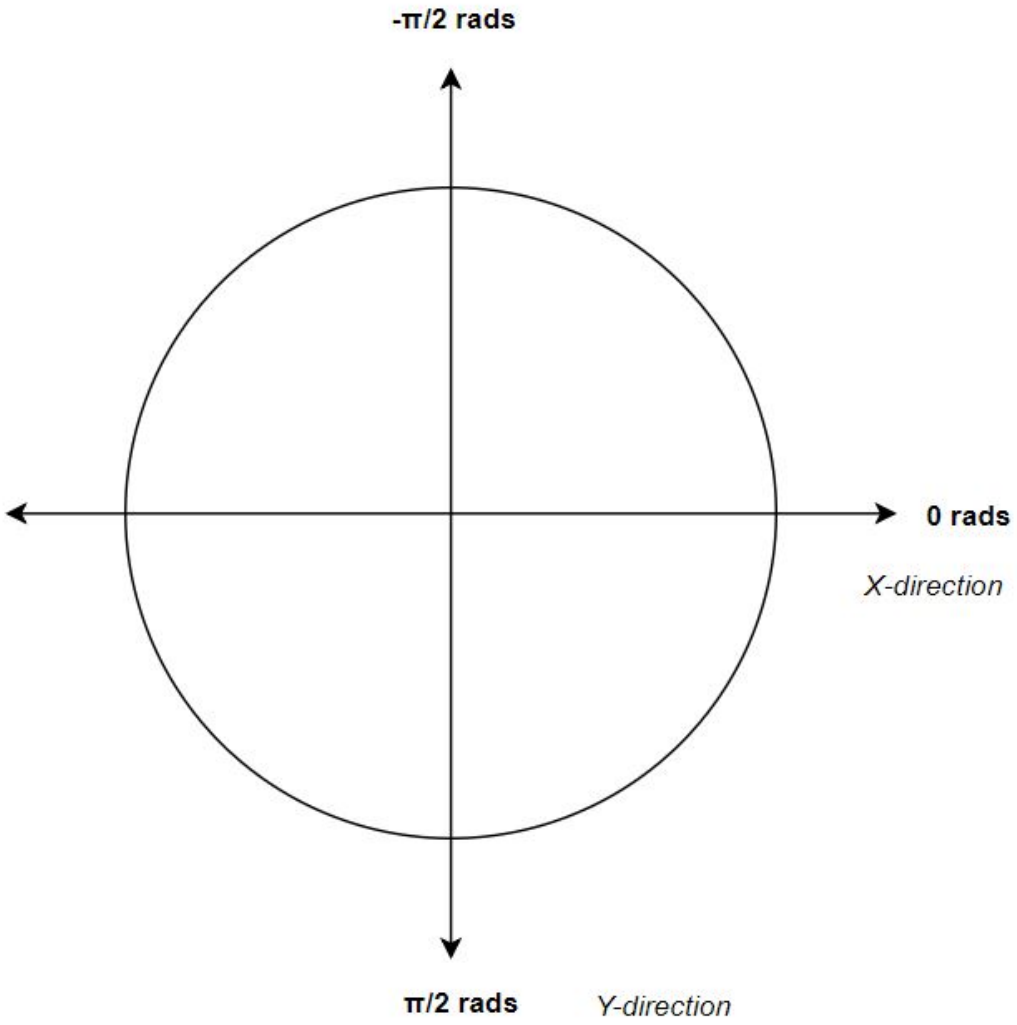

Figure 4.7: Unit circle with positive and negative angle directions

Figure 4.9 gave a visual representation of the angle of orientation for Sample 3 where the green arrow represented the predicted line parallel to the die cast runners of the zinc casting. In this example the output was almost perfect with only 2.3° of error between the expected and output angles.

Figure 4.8 gave a visual representation of the angle of orientation for Sample 1. This example behaved differently from the others where the absolute error was three times larger than Sample 4; the second highest absolute error.

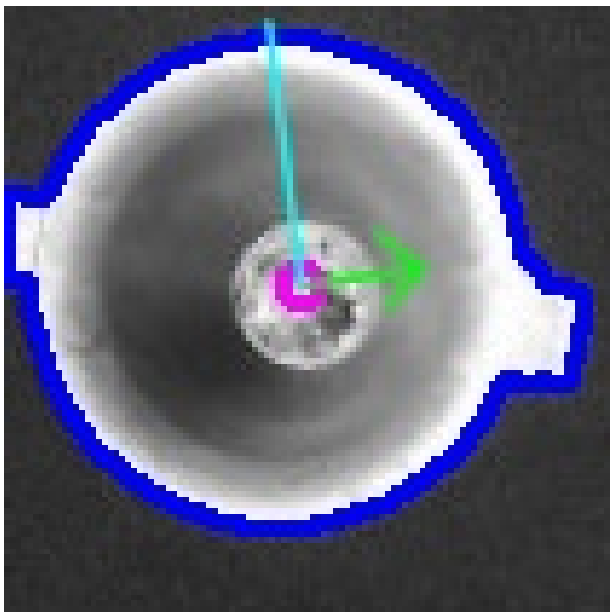

Figure 4.8: Orientation detection output of Sample 1

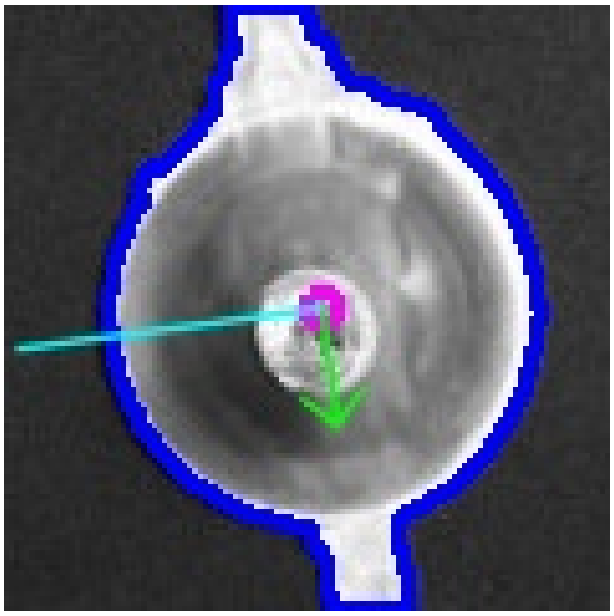

Figure 4.9: Orientation detection output of Sample 3

If this sample had been considered an outlier, the average error would have reduced from 0.12 rads or $6.9°$ to 0.06 rads or $3.2°$.

4.3.2 Evaluation of Results

The determination of cast angle of orientation was successful with a maximum error of $17.8°$ when considering the four samples seen in App. A.5 A.6 A.7 A.8. Table 4.3 showed for each sample, an absolute error value that was equal to the difference between the expected and output angles of orientation.

The absolute error at the output of Sample 1 proved three times greater than the second largest absolute error recorded. The major causes of this larger error were due to the cropped image of Sample 1 Fig. 4.8 having two distinct differences to the other cropped images including Sample 3 Fig. 4.8. Firstly, the sprue runner seen on the right hand side of Fig. 4.8 had a very small amount of contour data to deal with. Secondly, the right hand side sprue runner seen in Fig. 4.8 discontinued before leaving the frame. This could be dealt with by reducing the crop dimensions, at the expense of loosing more contour data of the other sprue runner.

All other absolute error values seen in Tab. 4.3 were between $2.3°$ and $5.2°$. These values could be accurate enough to require zero post robot gripping processes. The largest angle of error equalled $17.8°$. After the robot successfully gripped and orientated the object, a mechanical nesting process was necessary. A mechanical nest is typically used to align an object that has been gripped by a robot arm when using a vision sensor.

Figure 4.10 showed a nearly orientated object being gripped by a robot arm on the right hand side, with a mechanical nest on the left hand side. The robot arm already gripped the object and orientated as best it could using the vision system and was not completely aligned. The robot arm swept the object though the mechanical nest while applying force in the correct direction. This caused the object to leave the mechanical nest fully orientated.

Figure 4.10: Mechanical nest process using robot arm and vision

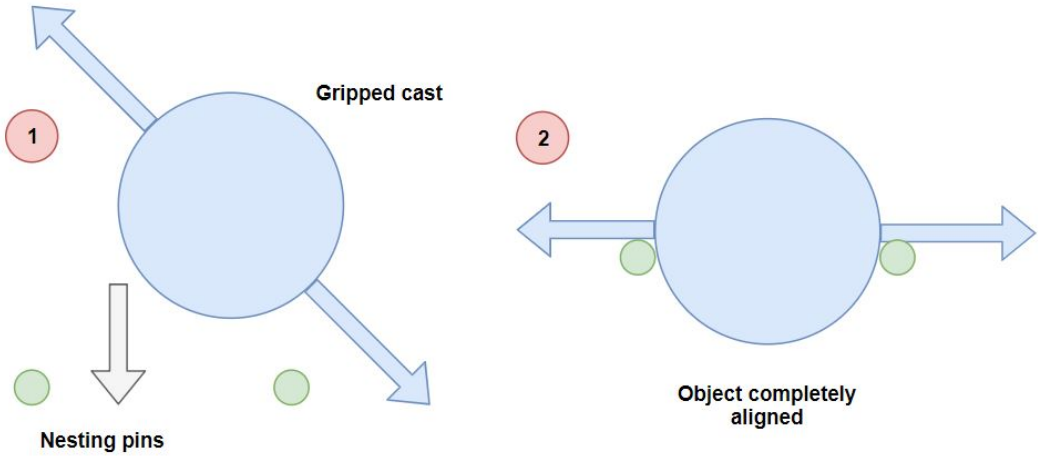

Figure 4.11: Physical steps of a mechanical nest for the die cast object

The automation engineers at Assa Abloy confirmed that a mechanical nest could successfully complete the orientation process with full alignment using a mechanical nest. Once the object had been successfully gripped and orientated as best possible, the cast was dropped into two fixed pins seen in Fig. 4.11. Allowing for the die cast runners to catch and complete the alignment process.

Table 4.4: Convex hull perimeter, area and x direction centroid measurements for die cast samples.

Cast Sample/Side	Perimeter (pixels)	Area (pixels)	x (pixels)
1/L	875	46952	93
1/R	235	3406	12
2/L	843	43609	86
2/R	265	4371	21
3/L	732	33709	90
3/R	591	24111	73
4/L	636	23036	111
4/R	606	26532	80

4.4 Quality Control Process: Checking Correct Cast

To ensure that the correct tooling had been selected during production, a quality control measure was implemented. High quantity orders meant that running the incorrect tooling combination would loose significant profits due to wasted time and resource.

An algorithm seen in Chap. 3 was used to split the 2D image into its its two separate tools as seen in Fig. 3.4 3.5. For the four samples App. A.1 A.2 A.3 A.4 provided; the perimeter length, area and x direction centroid coordinates were recorded and compared.

4.4.1 System Output

Table 4.4 showed the results of the three measurements taken on the four samples for both left and right tools. The three measurements were compared with expected values to determine if the correct cast was in production.

Two examples illustrated the output: If a random die cast output was measured using the proposed machine vision scheme and one tools measured values equalled 882 pixels for convex hull perimeter, 46871 pixels for area, and an x centroid coordinate of 89 pixels; it was assumed to be the left tool used in cast Sample 1; if an output was measured and convex hull perimeter equalled 587 pixels, area equalled 24102 pixels and x centroid coordinate equalled 69 pixels, it was assumed to be the right tool of Sample 3.

Such assumptions were made when considering a comparison with tolerances of ± 20 pixels for perimeter, ± 100 pixels for area, and ± 20 pixels for the x centroid coordinate.

4.4.2 Evaluation of Results

The proposed method proved suitable for the application. A series of comparisons based upon measured features provided future flexibility: When a new tool is added to the production inventory, a sample could be used to add the relevant information to the data matrix in very little time.

In Tab. 4.4, Sample 3 right and Sample 4 right were the same tool paired with different tools. The convex hull perimeter and x centroid proved to measure very close to one another, within 15 pixels. The area was measured to have a difference of 2421 pixels. This large variance in area was caused by a small casting overflow (tags) that in some cases fell off the object when falling from the die cast machine. The tags could be seen on Sample 4 right side in Fig. 4.12. Pixel area of the tags could be negligible. But when included in the convex hull, the difference in area was huge. By accepting ± 2500 pixels as tolerance when making convex hull area comparisons for incoming die cast outputs, the system remained robust.

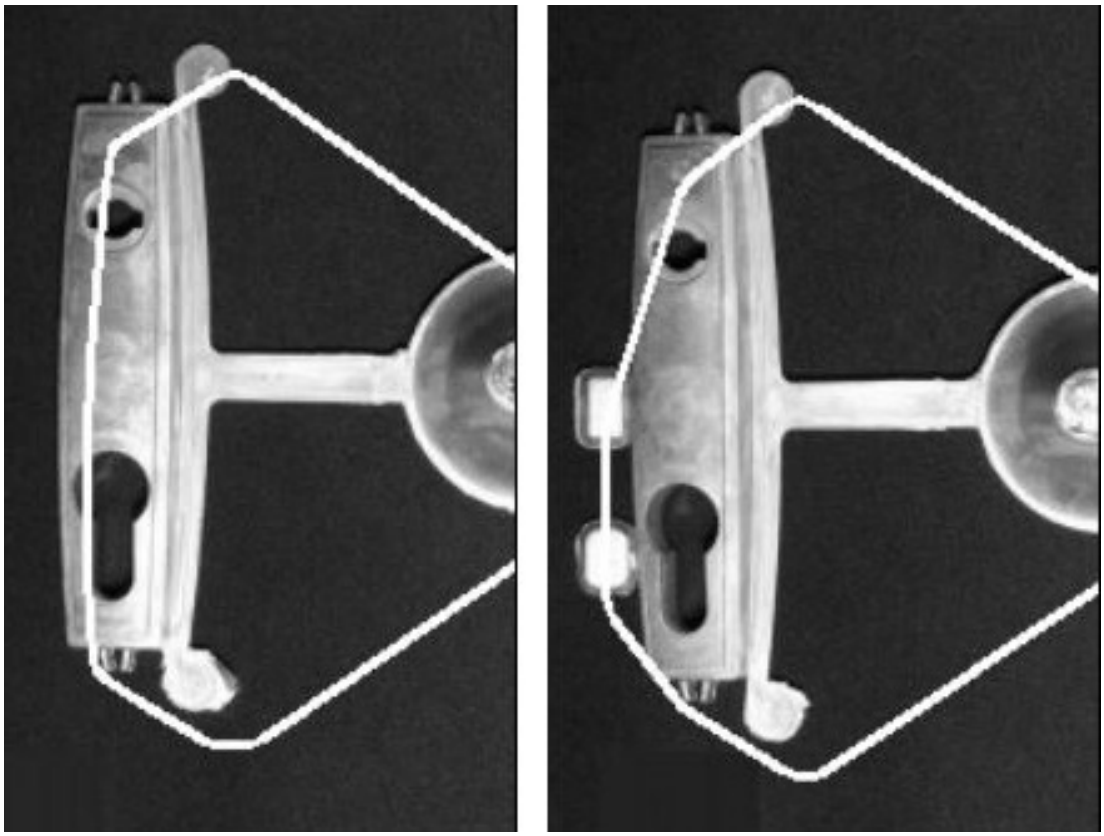

Figure 4.12: Comparison of Samples 3 right & 4 right with convex hull masks

Chapter 5

Conclusion and Future Work

The following chapter concludes the discussed work. The proposed system and analysis of results are reflected on to allow for future recommendations. The research generates a foundation for further implementation.

5.1 Conclusion

The research conducted aimed to answer the main research question – *How can the position, orientation and identification of zinc cast sprues be found using predominantly machine vision techniques?*

Assa Abloy, a local manufacturing company in Auckland, New Zealand produce high quantities of zinc products using a modulated die cast process. Zinc ingots are heated and melted into molten zinc before being injected into a machine die tooling. Once the molten zinc has set inside the die, the object is ejected onto a moving conveyor belt through a quench process. The company utilises a modulated tooling approach where two different tools can be coupled

to produce two separate products in one single die cast sample. This approach allows the company flexibility in production.

After the die cast output lands onto the conveyor belt, a worker picks up the object and removes the products from the waste zinc sprue. Once the waste has been removed, it is thrown onto a recycle conveyor before it is melted and used again. The products are placed into their respective bins.

The proposed methods introduced throughout the paper formed a proof of concept for the future implementation of a robust and time–efficient machine vision system which could take a randomly directed, face up or down die cast travelling on a conveyor and fix it in a known position. The methods produced accurate and robust results within the practical means of the company.

Generating a 3D point cloud to reconstruct a rough surface estimation of the die cast lying on the conveyor allowed for comparison with expected values to determine if a cast was lying face up or down on the conveyor. The system proved robust using the Basler ToF camera. An alternative method to detect the cast face was explored. A 2D grey image with an ellipse detector was sufficient to determine if a cast was lying face up or down on a conveyor by comparing the minor and major axis of the ellipse.

The circular Hough transform, along with other image processing techniques, proved to accurately determine the 2D robot picking location on the conveyor. The design of the vacuum gripping tool on the robot arm allowed for small inaccuracies. An alternative method involving an ellipse fit over the repeatable circular cone feature managed to robustly locate the robot picking location.

PCA gave a practical means to find the direction of a rotated die cast object lying face up on the conveyor. A mechanical nest, implemented after the robot had picked the object, allowed for complete alignment.

The quality control process delivered in the paper sufficiently identified the different tools used across the considered samples. It was noted that convex hull area was quite sensitive in nature to large variance. This indicated that the comparison tolerance was to be set to a greater value than the perimeter and x centroid.

5.2 Future Work

The research proved successful, developing a proof of concept for implementation of an automation system involving two vision sensors, a robot arm and conveyor system. Where a repeatable manual process can be handled with such system.

Additional improvements and future works are possible:

1. Larger range of samples,
2. machine learning to improve quality control process,
3. implementation with systems like Cognex smart cameras and PLC.

Throughout the thesis, the proposed system described was designed to operate accurately and robustly for the four zinc die cast samples provided by the company seen in App. A.1. The four samples were selected carefully to attempt to provide the system with a diverse range of samples that would challenge the system. Two were single and two were double tooled cast outputs.

By providing a larger range of zinc die cast samples, the system can be adjusted to deal with such a range in production. While only four samples were provided by the company, the number of possible zinc die cast output combinations is in the thousands.

The final process involved in the proposed system is the quality control process, used to check whether the correct die cast combination is being produced at a given time. A 2D vision procedure is used to conduct multiple tests on the face up die cast lying on the conveyor. Where the outcomes of the tests are compared against a table of data to isolate the correct cast which is being produced.

A machine learning process could be used in place of the proposed quality control method. Using supervised learning, a model could be trained prior to production with large numbers sample data. For each die cast tool output passing the conveyor, an image could be taken and labelled for a large dataset. With a large dataset of images representing the possible die cast tool outputs in varying position and angle of orientation, predictions could be made to a level of confidence when a zinc die cast output appears on the conveyor. This alternative quality control method would be hugely beneficial in the event that more than one die cast tool is too similar in terms of its features and no additional features can be added to distinguish the tools.

To implement such a system as proposed throughout the thesis, various methods can be considered. Smart cameras such as Cognex Insight [34] can be deployed fast with ease of programming. The smart camera can utilise the 2D machine vision processes described throughout Chap. 3. A Universal Robot (UR) [49] can be installed to communicate with the vision system and position the part as commanded. With a suitable conveyor system as described in Fig. 3.1, a PLC [2] could be used as the central control unit. All processes including the vision, robot arm and conveyor systems can be connected through the PLC.

The implementation environment needed careful consideration. Given that

the experiments presented throughout the thesis were conducted in the laboratory environment, noise conditions such as illumination were heavily controlled. The intensity, colour, diffusion and frequency of light were all set to ensure accurate and robust results.

To deploy such technology in a noisy industrial environment, a vision cabinet with controlled lighting conditions inside may be included. All vision parameters would be configured to operate for the given cabinet conditions.

Machine learning object detection may be considered to train a model that can detect the necessary shapes within the image under varying illumination. This method would require large samples of balanced training data under all environmental conditions expected within the noisy manufacturing environment and would not require the installation of a vision cabinet.

Appendix A

Appendix A: Images

The following chapter delivers relevant images to the research. A break down of sub–sections focusing on specific elements including the raw cast samples, rotated cast samples, pick up location output images and screenshots of the 3D point cloud surface reconstruction.

A.1 Cast Samples

The four zinc die cast output samples provided by Assa Abloy were photographed using the camera set–up Sec. 3.3. The 2D images produced were used as raw input images for the machine vision systems discussed in Sec. 3.5, 3.6, 3.7. The sections including the determination of robot pick location and orientation along with a quality control measure to check for error.

A.2 Rotated Cast Samples

The images provided in this section were used to supplement Sec. 3.6, 4.3; determination of the object orientation. Original images from Sec. A.1 were

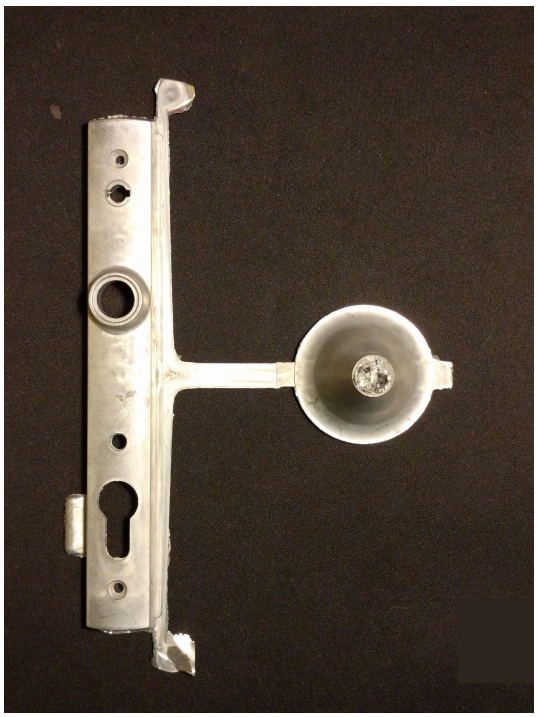

Figure A.1: Zinc die cast Sample 1

rotated by various amounts to allow for a range of input angles.

A.3 Pick Location Output Images

The following section provides the output images from the algorithm used to determine the robot pick location; results seen in Sec. 4.2. For each die cast sample, the output image of the fitted Hough circle appeared on the original image along with a zoomed version focusing on the fitted circle were presented.

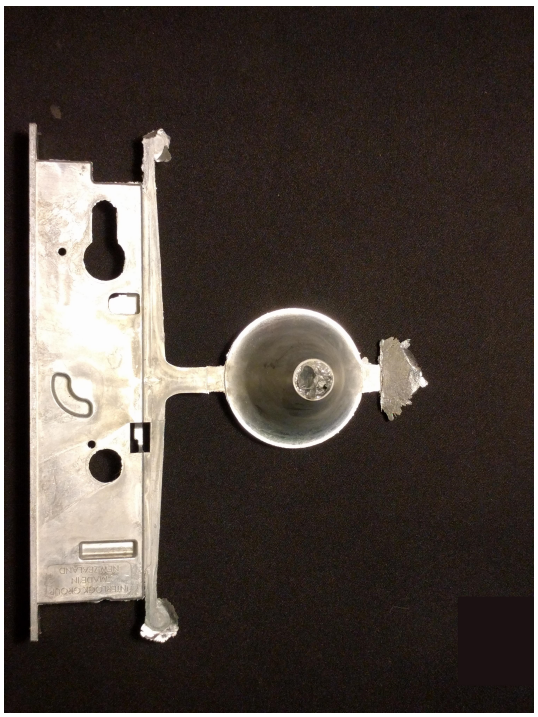

Figure A.2: Zinc die cast Sample 2

A.4 3D Point Cloud Reconstruction

To identify whether an incoming die cast appeared face up or down on the conveyor, a ToF camera was used to generate a point cloud 3D surface reconstruction of the cast. CloudCompare software was used in conjunction with the ToF sensor to generate the point clouds seen for Sec. 4.1. Each point cloud screenshot shows a side angle of the 3D reconstruction. Red pixels representing a greatest distance from the sensor.

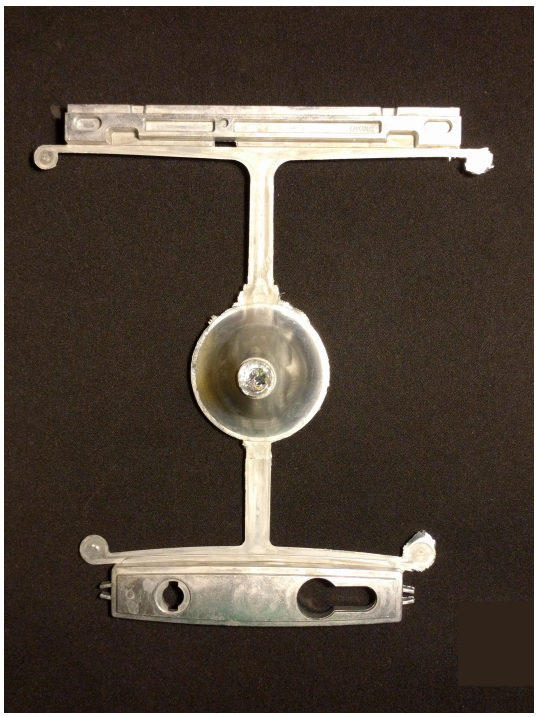

Figure A.3: Zinc die cast Sample 3

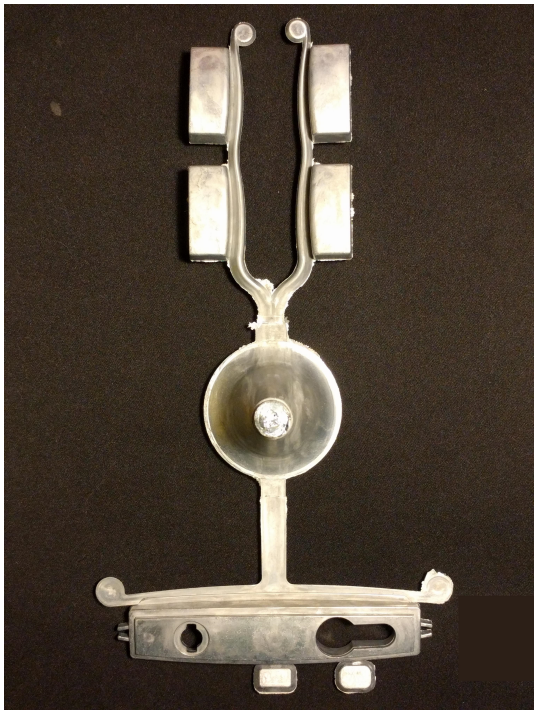

Figure A.4: Zinc die cast Sample 4

Figure A.5: Zinc die cast Sample 1 rotated

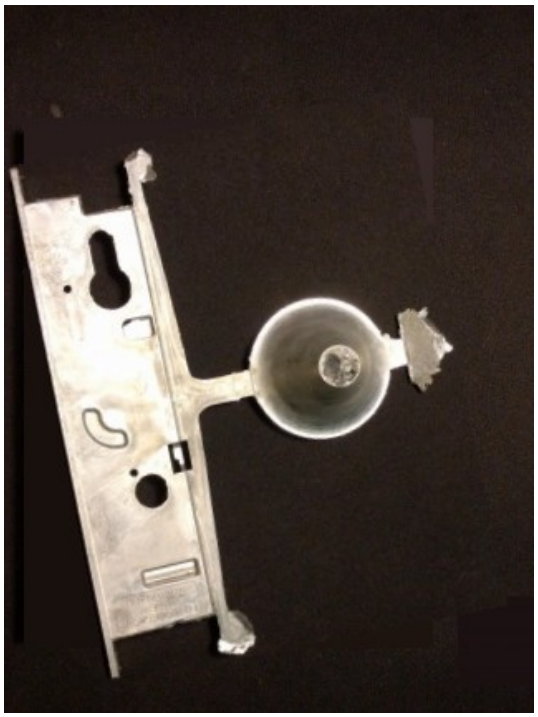

Figure A.6: Zinc die cast Sample 2 rotated

Figure A.7: Zinc die cast Sample 3 rotated

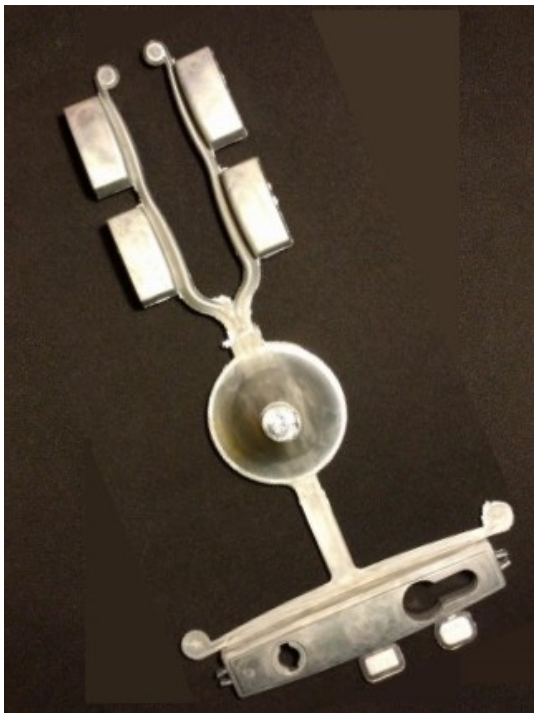

Figure A.8: Zinc die cast Sample 4 rotated

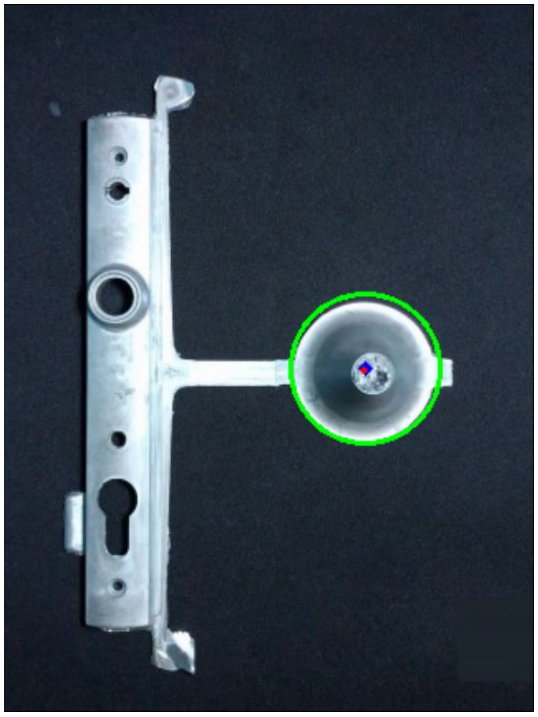

Figure A.9: System output for die cast sample 1

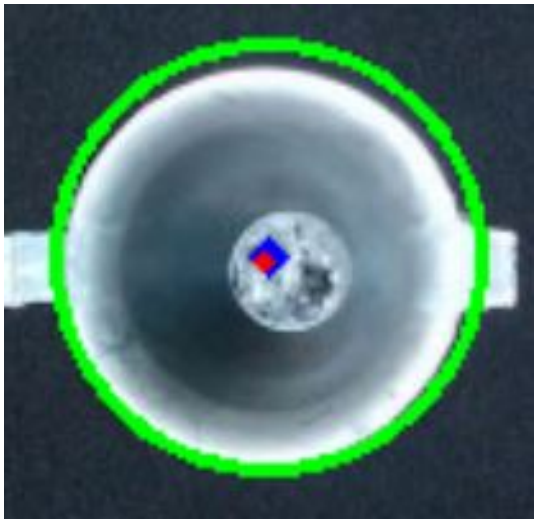

Figure A.10: Zoomed system output for die cast sample 1

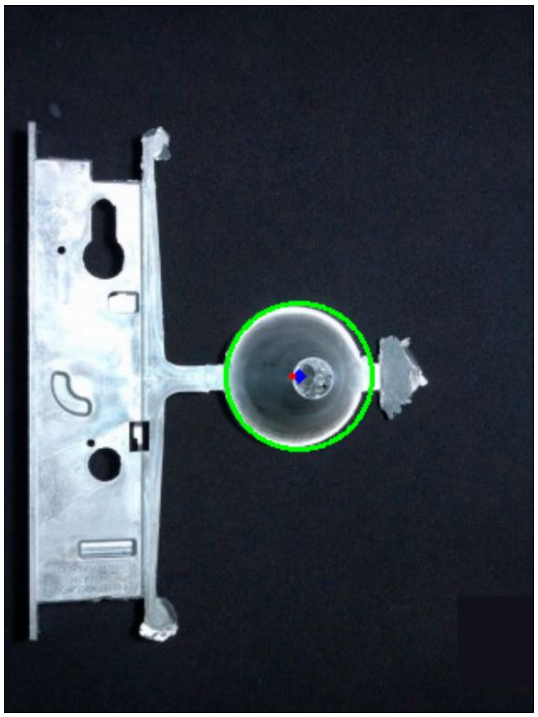

Figure A.11: System output for die cast sample 2

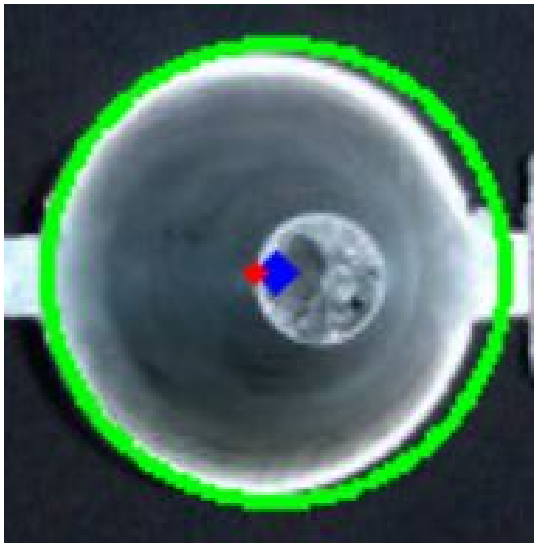

Figure A.12: Zoomed system output for die cast sample 2

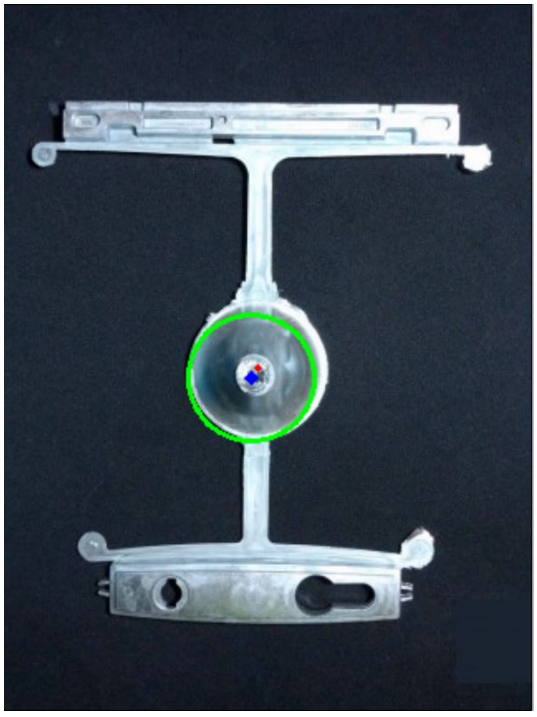

Figure A.13: System output for die cast sample 3

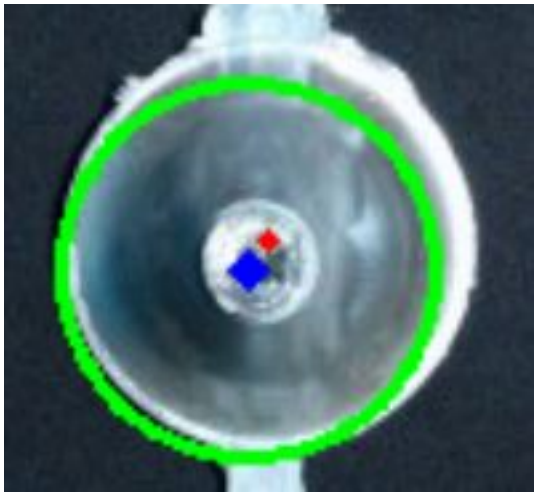

Figure A.14: Zoomed system output for die cast Sample 3

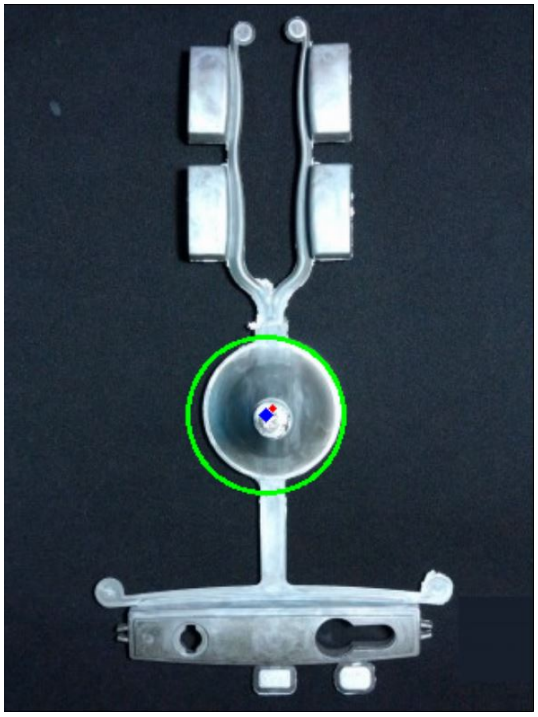

Figure A.15: System output for die cast Sample 4

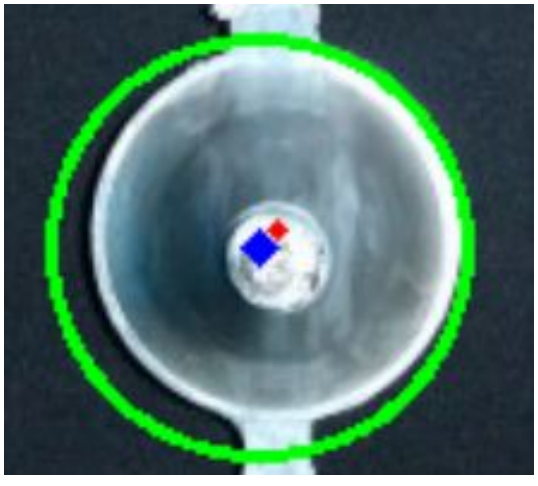

Figure A.16: Zoomed system output for die cast Sample 4

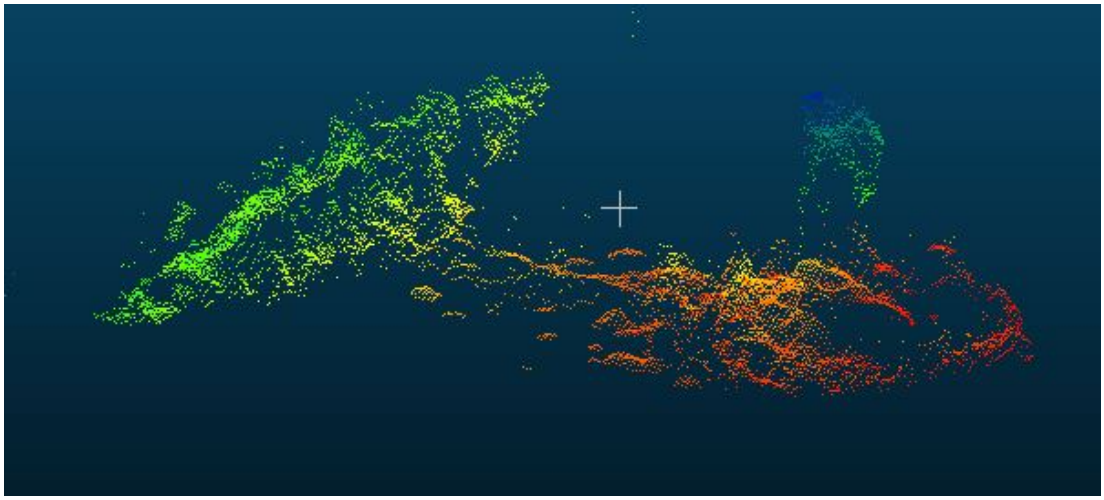

Figure A.17: Side view of point cloud 3D surface reconstruction for Sample 1 face up generated on CloudCompare

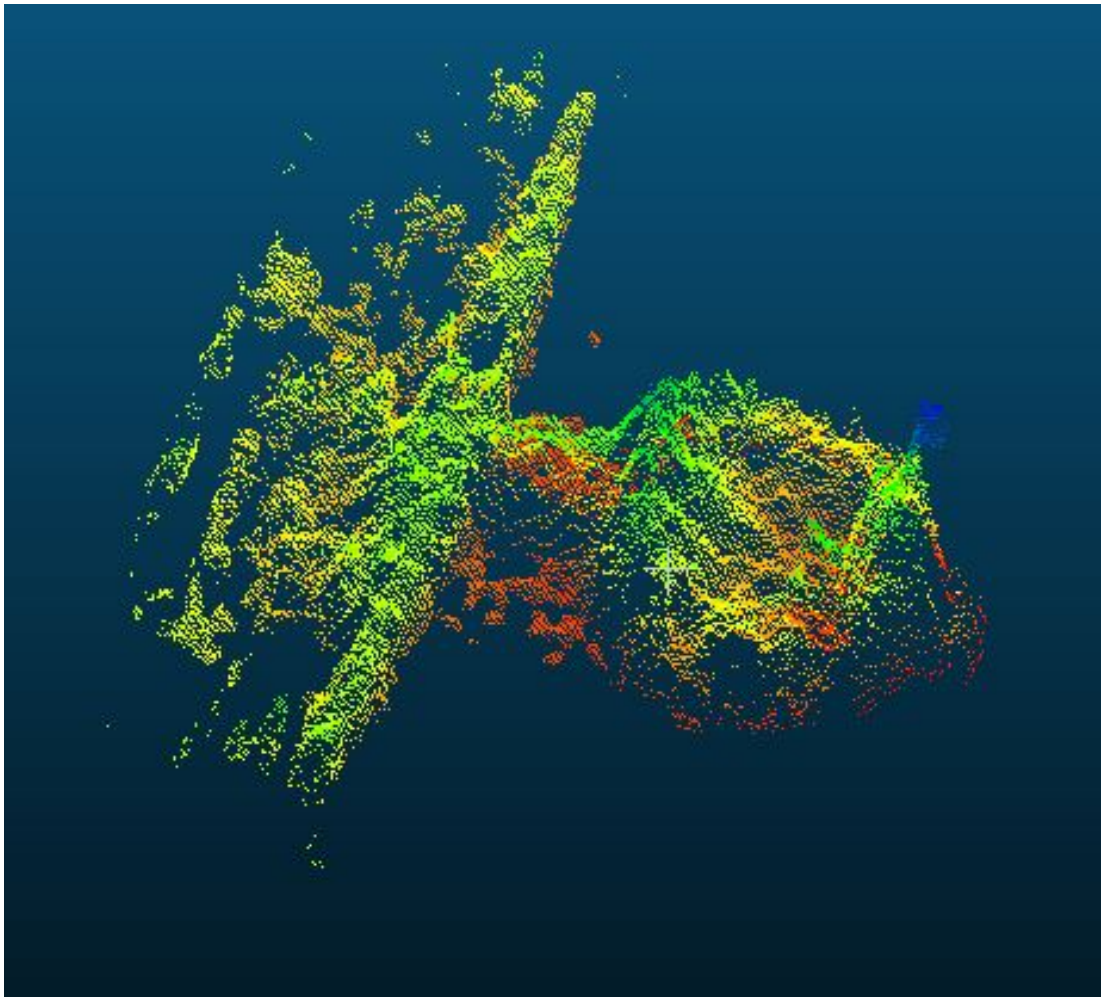

Figure A.18: Side view of point cloud 3D surface reconstruction for Sample 2 face down generated on CloudCompare

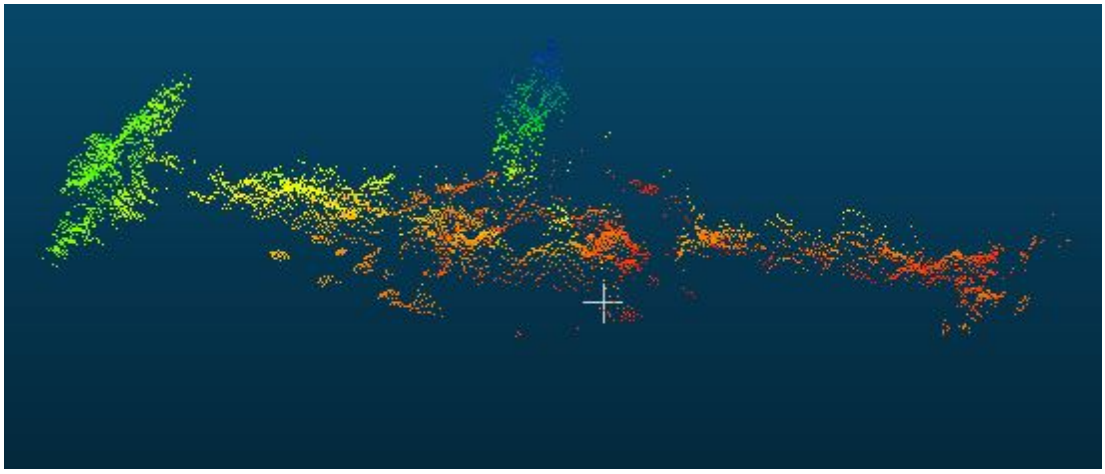

Figure A.19: Side view of point cloud 3D surface reconstruction for Sample 3 face up generated on CloudCompare

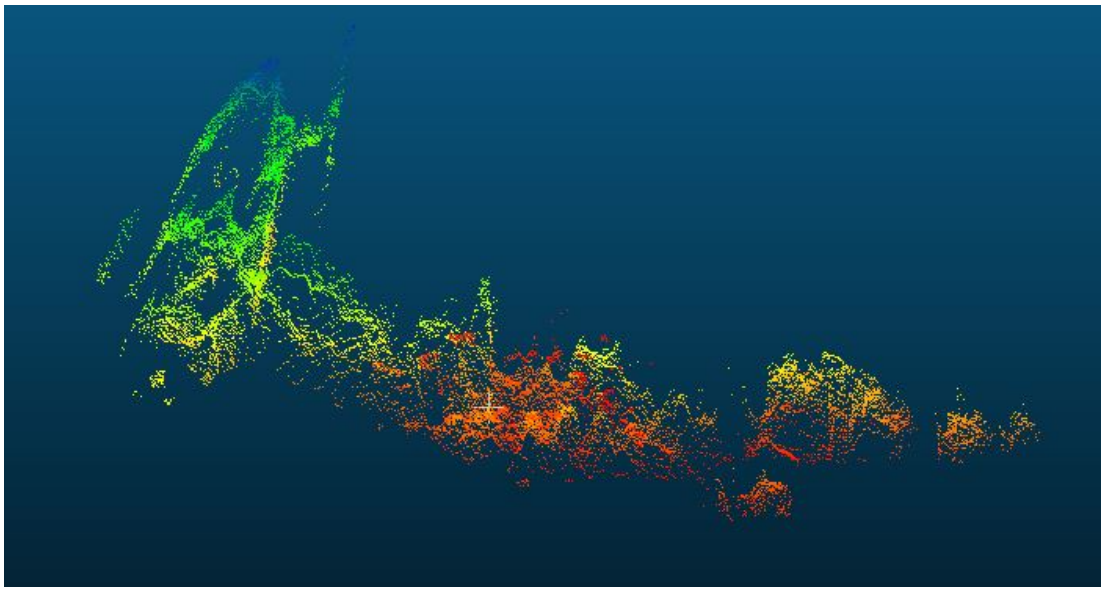

Figure A.20: Side view of point cloud 3D surface reconstruction for Sample 4 face down generated on CloudCompare

Bibliography

[1] P. J. Besl, et al., "Method for registration of 3–D shapes", Proc. *Conf. Sensor Fusion Control Paradigms Data Structures*, vol. 1611, pp. 586–607, 1992.

[2] L. A. Bryan, et al., "Programmable Controllers - Theory and Implementation Second Edition", `core.ac.uk/download/pdf/41463043.pdf`, visited 17 Dec 2018, An Industrial Text Company Publication, Atlanta, 1997.

[3] L. Butters, et al., "Using machine vision to command a 6-axis robot arm to act on a randomly placed zinc die cast product", Proc. *Int. Conf. Control Computer Vision*, 2019.

[4] N. Chaki, et al., "A Comprehensive Survey on Image Binarization Techniques", Springer, New Delhi, 2014.

[5] C. Fosu, et al., "Determination of Centroid of CCD Star Images", University FAF, Munich Germany, 2004.

[6] S. Fuchs, et al., "Calibration and registration for precise surface reconstruction with ToF cameras" Proc. *Int. Conf. Dynamic 3D Imaging DAGM*, vol. 1, 2007.

[7] N. K. Garg, et al., "Binarization techniques used for grey scale images", J. *Int. Computer Applications*, vol. 71, no. 1, 2013.

[8] S. A. Guomundsson, et al., "Fusion of stereo vision and time–of–flight imaging for improved 3D estimation", J. *Int. Intelligent Systems Technologies Applications*, no. 3/4, pp. 425–433, 2008.

[9] R. M. Gurav, et al., "Real time finger tracking and contour detection for gesture recognition using OpenCV", Proc. *Int. Conf. IEEE Industrial Instrumentation Control*, pp. 974–977, 2015.

[10] M. K. Hu, "Visual pattern recognition by moment invariants", *IRE Trans. Information Theory*, vol. IT-8, pp. 179–187, 1962.

[11] R. Klette, "Concise Computer Vision", Springer, London, 2014.

[12] L. Li, "Time of Flight Camera – An Introduction", `ti.com/lit/wp/sloa190b/sloa190b.pdf`, visited 4 Feb 2019, Texas Instruments, Dallas, 2014.

[13] W. B. Li, et al., "A calibration algorithm of the structured light vision for the arc welding robot", Proc. *Int. Conf. Ubiquitous Robots Ambient Intelligence*, pp. 481–483, 2017.

[14] Y. X. Liang, "Development of the three-dimensional scanning system based on monocular vision", Proc. *Int. Conf. Power Electronics Systems Applications*, pp. 1–5, 2015.

[15] S. May, et al., "Robust 3D–mapping with time–of–flight cameras", Proc. *Int. Conf. Intelligence Robots Systems*, pp. 1673–1678, 2009.

[16] M. Mesko, et al., "Laser spot detection", J. *Information Control Management Systems*, vol. 21, no. 1, 2013.

[17] J. R. New, et al., "Facilitating user interaction with complex systems via hand gesture recognition", Proc. *Southeastern Conf. ACM*, 2003.

[18] F. Patin, "An Introduction To Digital Image Processing", `teachme.free.fr/ImageProc.pdf`, visited 21 Dec 2018, Cenrale Supelec, Gif-sur-Yvette, 2003.

[19] A. Rosebrock, "Detectinng Circles in Images using OpenCV and Hough Circles", `pyimagesearch.com/2014/07/21/detecting-circles-images-using-opencv-hough-circles/`, visited 4 Aug 2019.

[20] M. Shah, "Fundamentals of Computer Vision", University of Central Florida, Florida, 1997.

[21] M. Sharifi, et al., "A classified and comparative study of edge detection algorithms", Proc. *Int. Conf. Information Technology: Coding Computer*, pp. 117–120, 2002.

[22] L. I. Smith, "A Tutorial On Principal Component Analysis", Department of Computer Science, University of Otago, Dunedin, 2002.

[23] J. Trivedi, et al., "OpenCV and matlab based car parking system module for smart city using hough transform", Proc. *Int. Conf. Energy Communication Data Analytics Soft Computing*, pp. 2461–2464, 2017.

[24] Z. Xu, et al., "Accurate ellipse extraction in low-quality images using a 2D accumulator array", *Pending acceptance*, 2019.

[25] V. K. Yadav, et al., "Approach to accurate circle detection: circular hough transform and local maxima concept", Proc. *Int. Conf. Electronics Communication Systems*, pp. 28–32, 2014.

[26] H. K. Yuen, et al., "Ellipse detection using the Hough Transform", Proc. *Alvey Vision*, pp. 1–8, 1988.

[27] J. Zhang, et al., "Face recognition: eigenface, elastic matching, and neural nets", J. *IEEE*, vol. 85, no. 9, pp. 1423–1435, 1997.

[28] Z. Zhang, et al., "Building 3D scanner system based on monocular vision", *Applied Optics*, vol. 51, no. 11, pp. 1638–1644, 2012.

[29] J. Zhu, et al., "Fusion of time–of–flight depth and stereo for high accuracy depth maps", Proc. *Int. Conf. IEEE Computer Vision Pattern Recognition*, pp. 1–8, 2008.

[30] Y. Zhuang, et al., "Automatic extrinsic self-calibration for fusing data from monocular vision and 3-D laser scanner", J. *Transactions Instrumentation Measurement*, vol. 63, no. 7, pp. 1874–1876, 2014.

URLs

[31] Basler - Basler ToF camera user's manual 2016, `kamery.atesystem.cz/site/assets/files/1635/aw00133807000_tof_gige_um.pdf`, visited 4 Feb 2019.

[32] Bordering - OpenCV tutorial, `docs.opencv.org/3.4/dc/da3/tutorial_copyMakeBorder.html`, visited 30 Jan 2019.

[33] Canny edge detector - OpenCV tutorial, `docs.opencv.org/3.1.0/da/d22/tutorial_py_canny.html`, visited 30 Jan 2019.

[34] Cognex in-sight 7000 specifications, `www.cognex.com/products/machine-vision/2d-machine-vision-systems/in-sight-7000-series/specifications`, visited 31 Jan 2019.

[35] Contours - OpenCV tutorial, `docs.opencv.org/3.3.1/d4/d73/tutorial_py_contours_begin.html`, visited 30 Jan 2019.

[36] Convex hull - OpenCV tutorial, docs.opencv.org/3.4/d7/d1d/tutorial_hull.html, visited 30 Jan 2019.

[37] Cropping - OpenCV tutorial, opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_core/py_basic_ops/py_basic_ops.html, visited 30 Jan 2019.

[38] Ellipse drawing - Pysource OpenCV example, pysource.com/2018/01/22/drawing-and-writing-on-images-opencv-3-4-with-python-3-tu

[39] Future of machine vision, www.qualitydigest.com/oct98/html/machfutr.html, visited 31 Jan 2019.

[40] History of machine vision, www.epicsysinc.com/blog/machine-vision-history, visited 31 Jan 2019.

[41] Hough circle transform - OpenCV tutorial, docs.opencv.org/3.4/d4/d70/tutorial_hough_circle.html, visited 31 Jan 2019.

[42] Inventor of the PLC - Dick Morley, www.automationmag.com/factory/855-the-father-of-invention-dick-morley-looks-back-on-t visited 31 Jan 2019.

[43] Moments - OpenCV tutorial, docs.opencv.org/3.4.2/dd/d49/tutorial_py_contour_features.html, visited 30 Jan 2019.

[44] NumPy - numerical computing package, numpy.org, visited 4 Aug 2019.

[45] Save American jobs by investing in robots, www.inverse.com/article/29740-robots-automation-jobs-china, visited 31 Jan 2019.

[46] Smoothing - OpenCV tutorial, opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_imgproc/py_filtering/py_filtering.html, visited 30 Jan 2019.

[47] Tesla's remote controlled boat, engadget.com/2014/01/19/nikola-teslas-remote-control-boat/, visited 31 Jan 2019.

[48] Thresholding - OpenCV tutorial, docs.opencv.org/4.0.1/d7/d4d/tutorial_py_thresholding.html, visited 30 Jan 2019.

[49] UR3 user manual, vajdamuvek.hu/files/UR3_User_Manual_en_Global.pdf, visited 14 Oct 2018.

[50] Wikipedia page on Assa Abloy, en.wikipedia.org/wiki/Assa_Abloy, visited 31 Jan 2019.

[51] Wikipedia page on the assembly line, en.wikipedia.org/wiki/Assembly_line, visited 31 Jan 2019.

[52] Wikipedia page on automation, en.wikipedia.org/wiki/Automation#20th_century, visited 31 Jan 2019.

[53] Wikipedia page on casting, en.wikipedia.org/wiki/Casting_(metalworking)#Mold_cavity, visited 31 Jan 2019.

[54] Wikipedia page on convex hull, en.wikipedia.org/wiki/Convex_hull, visited 28 Nov 2018.

[55] Wikipedia page on die casting, en.wikipedia.org/wiki/Die_casting, visited 31 Jan 2019.

[56] Wikipedia page on disk tumbler lock, en.wikipedia.org/wiki/Disc_tumbler_lock, visited 31 Jan 2019.

[57] Wikipedia page on edge detection, en.wikipedia.org/wiki/Edge_detection, visited 12 Dec 2018.

[58] Wikipedia page on image moments, en.wikipedia.org/wiki/Image_moment, visited 15 Dec 2018.

[59] Wikipedia page on infrared, en.wikipedia.org/wiki/Infrared, visited 31 Jan 2019.

[60] Wikipedia page on machine vision, en.wikipedia.org/wiki/Machine_vision, visited 31 Jan 2019.

[61] Wikipedia page on Nikola Tesla, en.wikipedia.org/wiki/Nikola_Tesla, visited 30 Jan 2019.

[62] Wikipedia page on Smoothing, en.wikipedia.org/wiki/Smoothing, visited 21 Dec 2018.

Index