Accurate Three-Dimensional(3D) Measurement of Highly Specular Surfaces for Quality Control Program of Large-scaled Production Line

Arpita Rajkumar Dawda

Supervisors:
Dr. Minh Nguyen
Associate Prof. Loulin Huang

In collaboration with:
Facteon Intelligent Technology Ltd.

A thesis submitted to
Auckland University of Technology
in fulfilment of the requirements for the degree of
Doctor of Philosophy (PhD)

September 2021
School of Engineering, Computer and Mathematical Sciences

Abstract

In the production line, inspection and quality control are essential to maintain the quality of the products. It assures confidence in the manufacturer and provides satisfaction to the customer. Product inspection is an indispensable tool of the modern manufacturing process. It helps in maintaining the quality of the product and reduces manufacturing costs by eliminating scrap losses. Several *non-destructive examinations* (NDE), or *non-destructive testing* (NDT), are used to analyze materials for inherent flaws such as fractures, dents or cracks. Same as the manufacturing process, the inspection process should also be automatic.

Machine vision emerged as an important new technique for industrial inspection and quality control in the early 1980s. It is proven to be an accurate and inexpensive inspection tool for high volume, labour-intensive and repetitive inspection operations in automotive and manufacturing industries. Machine vision provides the technology and methods for imaging-based automatic inspection and analysis. However, the reflective surface of the object puts some limitations on traditional methods of machine vision. Generally, the inspection of the reflective surface is performed in a dark environment, as the ambient lighting condition of the working environment makes the reflective surface look highly specular.

This research mainly focuses on overcoming the limitation of traditional machine vision methods. A novel *three-dimensional* (3D) measuring system is developed to inspect a product with a highly specular surface accurately. This technique aims to combine the concepts of stereo vision and laser triangulation for the 3D reconstruction of the product. This method provides a simple but accurate solution to inspect the reflective surface. The main advantage of this system is that it works robustly even in the presence of ambient light.

The thesis briefly explains the effect of background on the accuracy of the inspection. Also, a thorough comparison of red and blue light lasers in terms of accuracy is described. In addition, the difficulties induced by the nature of the surface in ambient lighting conditions are evaluated. An algorithm is invented to overcome these difficulties.

Along with accurate measurements, it is also essential to detect defects such as dents, bumps, cracks, and scratches present in a product. As these defects are palpable and are not visible by the camera, it is tough to detect them using vision-based inspection techniques in ambient lighting conditions. This thesis focuses on three types of defects: Dents, Bumps, and Scratches. With the proposed 3D measurement system, we can detect the defects of size 0.02mm accurately.

Artificial intelligence (AI) has many applications in the production industries. One of the applications is to inspect the products for defects. However, AI is not used

to reconstruct a 3D model of the product with reflective surfaces accurately. In this research, we propose to use machine learning-based techniques for the accurate 3D reconstruction of the product. The *one-dimensional* (1D) data of the projected laser line is used to train *machine learning* (ML) and *deep learning* (DL) models. These models are trained to detect the projected laser line accurately in the presence of ambient light. The detected laser line plays a vital role in creating an accurate 3D model of the product.

Finally, we compare different machine learning-based techniques with the above-mentioned stereo-laser technique based on accuracy.

Keywords

Three-Dimensional(3D) measurement, Quality control, Highly specular surfaces, Ambient lighting conditions, Defects detection, Stereo vision, Laser line projection, Deep learning, Machine learning, One-Dimensional(1D) convolutional neural network, Supervised regression, Recurrent neural network, Inspection.

Acknowledgment

First of all, I gratefully thank God, for his truth, mercy, grace, wisdom, and blessings.

Neil Armstrong said, "Research is creating new knowledge." This journey of discovery would not have started without the support and guidance of **Professor Reinhard Klette**. I sincerely thank him for his suggestion to pursue PhD after seeing my interest in the Computer Vision subject. His professional guidance and supervision have helped me throughout this journey. I will always be thankful to him for choosing me to work with Facteon Intelligent Technology Ltd. He always encouraged me to achieve the best of my potential, not only in academia but also in leading me to be a good person and live a better life.

I gratefully acknowledge **Dr. Minh Nguyen** for supervising this research after the retirement of Prof. Klette. His ideas, suggestions and insights have helped me in improving the quality of this research. I would forever be grateful to him for contributing to all of my publications and for being my mentor throughout this research. Additionally, I thank my secondary supervisor **Associate Professor Loulin Huang** for his guidance.

I would also like to thank **Facteon Intelligent Technology Ltd.** for providing me with the opportunity to work on their research project. My three years of work experience with Facteon has given me immense knowledge regarding research and project management in production industries. I am grateful to Facteon for their constant professional and financial support to this research.

I am thankful to **Mr. Wayne Huang** for his consistent supervision at every stage of this knowledge pursuing adventure. His inputs helped me gain an understanding of the requirements of Facteon. He has always been a helping hand during this research which made my work experience at Facteon effortless. I also appreciate the support and feedback that I received from **Mr. Ken Jakeman** and **Mr. Ray Russell** during our meetings.

This thesis would not have been done without the unconditional love and support of my dear dad, mum, brother, and the rest of my family.

Last but most, I would like to thank my friends Akash, Dhwanit, Jigar, Kavita, Meet, Mohammad, and Umangi for being my backbone and for always encouraging me to enrich my research quality.

Arpita Dawda
Auckland
March 1, 2022

Declaration

I hereby declare that this submission is own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the qualification of any other degree or diploma of a university or other institution of higher learning.

Signed:

Date: March 1, 2022

Publications

A. Dawda, and M. Nguyen, "Accurate 3D Measurement of Highly Specular Surface in Ambient Lighting Conditions", Submitted to Journal TPAMI, United States, 2021.

A. Dawda, A. Varasada, and M. Nguyen, "3-Dimensional Reconstruction of a Highly Specular or Transparent Cylinder from a Single Image", EAI Smartgift, Portugal, 2021.

A. Dawda, and M. Nguyen, "Comparison of Red versus Blue Laser Light for Accurate 3D Measurement of Highly Specular Surfaces in Ambient Lighting Conditions", ISGV, New Zealand, 2021.

A. Dawda, and M. Nguyen, "Defects Detection in Highly Specular Surface using a Combination of Stereo and Laser Reconstruction ", IVCNZ, New Zealand, 2020.

A. Dawda, M. Nguyen, and R. Klette, "Accurate 3D Measurement of Highly Specular Surface using Laser and Stereo Reconstruction ", IVCNZ, New Zealand, 2019.

M. Norouzifard, A. Dawda, A. Abdul-Rahman, H. GholamHosseini, and R. Klette, "Accurate 3D Measurement of Highly Specular Surface using Laser and Stereo Reconstruction ", IVCNZ, New Zealand, 2018.

M. Nguyen, H. Le, W. Yan, and A. Dawda, "A Vision Aid for the Visually Impaired using Commodity Dual-Rear-Camera Smartphones", M2VIP, Germany, 2018.

Contents

List of Figures

List of Tables

List of Symbols

Reflection Model	**Description**
x	Two-dimensional image coordinates.
$I(x)$	Color vector of image intensity.
$I_r(x), I_g(x), I_b(x)$	RGB elements of the color vector of image intensity.
$D(x)$	Diffuse reflection component.
$D_r(x), D_g(x), D_b(x)$	RGB elements of diffuse reflection component.
G	Specular reflection component.
G_r, G_g, G_b	RGB elements of specular reflection component.
α	Weighting factors for diffuse reflections.
β	Weighting factors for specular reflections.
λ	Wavelength of the light spectrum.
$A(\lambda, x)$	Diffuse reflectance function.
$B(\lambda)$	Specular reflectance function.
ω	Visible spectrum.
$Q(\lambda)$	Three element-vector of sensor sensitivity.
σ	Chroma of the image.
Λ	Diffuse reflection chroma.
$\Lambda_r, \Lambda_g, \Lambda_b$	RGB elements of diffuse reflection chroma.
Γ	Specular reflection chroma.
$\Gamma_r, \Gamma_g, \Gamma_b$	RGB elements of specular reflection chroma.

Artificial Intelligence	
(x_i, y_i)	Sequential train data pair.
t	Time.
w_i	Weight of kernel.
L_i	Convolution layer.
b	Bias.
α	Mini-batch variance.
σ	Scale parameter.
ϵ	Random noise.
β	Shifting parameter.
μ_B	Mini-batch mean.
C_t	A cell state of LSTM for the current timestamp.
$C_t - 1$	A cell state of LSTM for the previous timestamp.

H_t	A hidden state of LSTM for the current timestamp.
$H_t - 1$	A hidden state of LSTM for the previous timestamp.
y_i	Actual value.
$\hat{y}_i$	Predicted value.
p	Number of dataset elements.

Stereo-Laser System

b	Baseline distance.
(x, y)	Image coordinates.

Height Measurement

$A, a, B, b, C, c, D, d, x$	Scalars.
D_1	Diameter of the product.
β	Angle.
$Z = 0$	Measurement plane.
Z	Distance between the camera and the measurement plane.
h	Height of the product.

List of Abbreviations

Acronym	Definition
1D	*one dimensional*
2D	*two dimensional*
3D	*three dimensional*
ACC	*accuracy*
AI	*artificial intelligence*
ANN	*artificial neural networks*
BN	*batch normalization*
BRDF	*bidirectional reflectance distribution function*
CCD	*charge coupled device*
cm	*centimeters*
CMM	*coordinate measuring machine*
CNN	*convolutional neural network*
COPQ	*cost of poor quality*
DL	*deep learning*
DLP	*digital light processing*
DMD	*digital micro-mirror device*
DT	*decision trees*
FC	*fully connected*
FM	*feature map*
FOV	*field of view*
FPGA	*field programmable gate array*
FPR	*false positive rate*
FPS	*frames per second*
GPU	*graphics processing unit*
HDR	*high-dynamic range*
IEP	*intelligent equipment platform*
IOT	*internet of things*
IR-PMD	*infrared-phase measuring deflectometry*
KD	*k-dimensional*
KNN	*k-nearest neighbors*
LCD	*liquid crystal display*
LED	*light emitting diode*
LSDNN	*laser-stripe-detection neural network*

LSTM	*long short-term memory*
MAE	*mean absolute error*
MB/s	*mega byte per second*
MIGL	*maximum input gray level*
ML	*machine learning*
MLP	*multi-layer perceptron*
mm	*millimetres*
MP	*megapixels*
MSLE	*mean square logarithmic error*
mW	*milliWatt*
NDE	*non-destructive examinations*
NDT	*non-destructive testing*
nm	*nanometre*
NZD	*New Zealand dollars*
PMD	*phase measuring deflectometry*
R^2	*Pearson correlation*
RCS	*reflection component separation*
ReLU	*rectified linear unit*
RF	*random forest*
RMSE	*root mean squared error*
RNN	*recurrent neural network*
ROI	*region of interest*
RPM	*rotation per minute*
SGM	*semi-global matching*
SLS	*structured light system*
SNR	*signal-to-noise ratio*
SVM	*support vector machine*
TIR	*total internal reflection*
ToF	*time-of-flight*
TPR	*true positive rate*
USB	*universal serial bus*
WCS	*world coordinate system*
XGBoost	*extreme gradient boosting*

Chapter 1

Introduction

This chapter provides a brief background of machine vision for inspection in a large production line. The critical parameters of this research, such as; the working environment's ambient lighting conditions, and the specularity of the product's surface, are explained in detail. Also, a brief introduction of Facteon Intelligent Technology Ltd. is provided. The rationale of the study and motivations are as well specified in this chapter. A brief summary of our approach for the inspection in ambient light is provided. Additionally, the problem statements and research questions are presented; finally, the thesis's structure is denoted by a flow chart in Fig. 1.1.

1.1 Background

According to Campbell, "INSPECTION is a methodical examination or formal evaluation exercise. In engineering, inspection involves the measurements, tests, and gauges applied to certain characteristics in regard to an object or activity [2]." In the age of smart manufacturing, manufacturers look to gain higher product yield and reduce scrap rates by improving the accuracy of inspection processes [3]. It is also essential to measure all critical-to-quality features before the product leaves manufacturing lines. The quality results help to simulate failure modes and verify strength criteria to validate functional product design [4].

As the manufacturing process is automated, it is essential that the process of quality checking is also automated. The automation in quality checking will save time, reduce manual work and increase the accuracy of the output [5]. In automated manufacturing process, on-line non-destructive automated inspections are well suited [4]. Machine Vision is the key to automation [6]. Its ability to acquire an image, analyze it, and make an appropriate decision is extremely useful in inspection and quality control applications [2]. Accuracy, time and cost are some critical parameters to consider while choosing the inspection process. The accuracy of the results depends on the technique of inspection. However, the accuracy is compromised when machine vision techniques are used upon reflective surfaces. As the

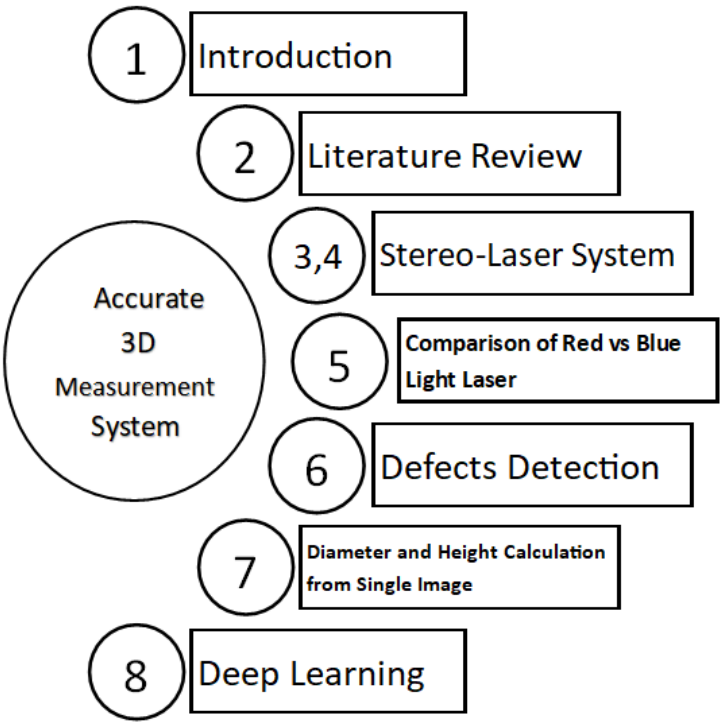


Figure 1.1: Flow chart of the thesis.

ambient light will make the surface look highly specular (mirror-like). Additionally, the manufacturing machinery causes vibrations, which can also affect the accuracy of the inspection process. In addition, to save time, defects must be detected at the same time as the 3D measurement. The ultimate goal of this research is to develop an accurate 3D measurement system that works flawlessly even in ambient light. Furthermore, the system should be fast, autonomous and cost-effective.

This research has been carried out in collaboration with Facteon Intelligent Technology Ltd. Facteon was founded in 1981 as the integral engineering division of Fisher & Paykel Appliances. "Facteon is an international specialist in designing, building and innovating world-class factory solutions [7]." Facteon is a member of Haier's industrial intelligence and automation division. It is also a member of *intelligent equipment platform* (IEP) Group. Facteon is trusted by some of the world's largest manufacturers to design, implement and maintain their production plants. Facteon expertise in applying intelligent manufacturing technologies to focus on de-

Figure 1.2: Washing machine drum.

livering a superior product to customers [7]. Automated quality control and inspection are some of the many important services that Facteon provides. This research is formed as a part of Facteon's automated inspection service. In this research, we[1] focus on inspecting products manufactured by Facteon. It manufactures different parts of consumer appliances such as drums, doors and panels, cabinets and cases, water heater cases and refrigeration foaming lines [8].

A washing machine drum and the wash plate are the products that we are going to inspect throughout this research. The drum is balanced, of high-quality construction and superior cosmetic appearance [4]. Figure 1.2 shows an image of a high-quality drum manufactured by Facteon. Quality control is done before proceeding to the next assembly stage, as the drum is composed of several parts. Most of the products manufactured by Facteon are made of stainless steel which has a specular nature of surfaces (e.g. the surface has faceted reflection, it is very reflective with light). As shown in Fig. 1.3, the wash plate of the washing machine drum has some dents which are visible in the left image. After slightly rotating the wash plate, dents are invisible in the right image. The visibility of defects depends on the

[1]The use of "we" throughout this thesis is purposeful. It is used to involve the reader with the thesis as recommended by Knuth [1].

angle of the viewing plane, camera position, and projection lights' angles. Therefore, it is not easy to perform quality inspections of highly specular surfaces using computer vision-based techniques. In ambient conditions, the shape of the specular surface causes reflection, making it difficult to inspect every tiny feature of the product [9], [10]. Hence, our focus in this research is to perform inspections in ambient lighting conditions.

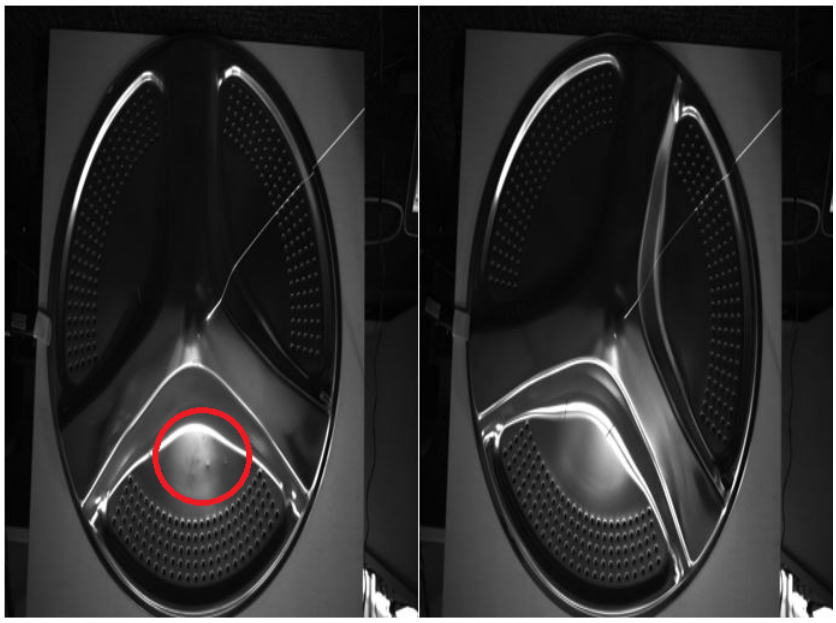


Figure 1.3: Defected wash plate with some dents, bumps and scratches [9].

The traditional machine vision concepts, such as stereo vision and laser triangulation are the base of our invented system. Research has also been done using state-of-art supervised machine learning techniques such as *convolutional neural networks* (CNN), *recurrent neural networks* (RNN), *long short-term memory* (LSTM) and traditional machine learning models such as *support vector machine* (SVM), decision tree, *K-nearest neighbors* (KNN). We have used these techniques for accurate laser line detection in ambient light, which is essential for our developed system. This is described in detail in Chapter 2.

1.2 Motivations

Morgan states the effect of quality control on an industry by, "Quality control affects a company in various ways, from productivity and profitability to customer satisfaction and public perception. In addition, quality control affects the overall operating costs of a company [12]." Inspection is needed in every stage of the manufacturing process. It helps maintain the quality of the product and reduces manufacturing costs by eliminating scrap losses [10, 11].

In this research, we will focus on a washing machine drum that is balanced, of high-quality construction and superior cosmetic appearance (See Fig. 1.2). Inspection is performed to ensure the relative position between drain holes and dimple pattern, the dimension of holes, and the drum's roundness is inset tolerance. The roundness of the drum is the most crucial feature since the drum will spin at a very high speed of up to 1000 *rotations per minute* (RPM) to drain the water out. If the drum is not perfectly round, then it will tremble and make noise while rotating at high speed. It also reduces the durability of a washing machine. Moreover, the holes should be positioned accurately in the middle of the dimple groove for the water to be drained through the holes by centrifugal force. If their relative position is not correct, the draining efficiency will decline considerably.

Reflective surfaces are the limitation of traditional machine vision inspection methods. As the products manufactured by Facteon are mostly made of stainless steel, it makes the surface of the product look reflective [10]. Also, the state-of-art techniques for the inspection of reflective surfaces do not work in ambient lighting conditions. Generally, the quality checking of reflective products is done in a semi-automatic or manual way which consumes a lot of time and human effort. Still, it is hard to achieve the accuracy of +/- 0.1 *millimeters* (mm) with manual work. Therefore, this research aims to develop an autonomous 3D measurement system that can reconstruct the accurate 3D model of the reflective surface in ambient light. This method will be beneficial, as it saves time, reduces human efforts and increases the accuracy of the output. Moreover, vision-based techniques could not be used to detect palpable defects. Therefore, a novel technique for the detection of palpable defects is also invented.

AI is now getting used in multiple fields to automate the system, to increase the efficiency and performance of the system [13]. By observing the applications of AI in numerous fields, we also tried to use AI with our developed 3D measurement system in this research. We used different models of deep learning and machine learning to detect the projected laser line in ambient light. Thus, we have developed a hybrid system by using AI in our stereo-laser system. Moreover, our proposed laser detection algorithm in a stereo-laser system, is critically analyzed by performing a

thorough comparison with different DL and ML models used for laser detection.

1.3 Problem Statements and Research Questions

Poor quality of the product can hurt a business' finances through tangible and intangible costs. Tangible cost covers things like rejection, rework and high repair costs. On the other hand, intangible costs are less visually evident and conceptual such as the loss of customer satisfaction, drop in employee morale, and a bad reputation [14]. According to [15, 16], "The traditional *cost of poor quality* (COPQ) has usually been assumed to be between four per cent and five per cent of an organization's annual revenue [16]." However, if we account for tangible and intangible costs, COPQ would rise to be 10 per cent to 25 per cent of an organization's annual revenue [16].

To reduce tangible and intangible costs, product inspection is essential. However, the product with a surface that becomes highly specular in ambient light is a limitation of many state-of-art inspection methods. "The fundamental reason is that the metallic workpiece has a much higher level of intensity variation in their reflections compared with what can be captured by the 0–255 gray-level intensity range of the camera [16]." Therefore, this research aims to inspect highly specular surfaces accurately in ambient light.

The research questions of this study are listed as follows:

1. How to get accurate 3D dimensions of the product with reflective surface in ambient light?

2. How to achieve the accuracy of $\pm$ 0.1 millimeters (mm) in the measured dimensions of the product?

3. How to detect palpable defects in the reconstructed 3D model?

4. How to make the process of 3D reconstruction automatic?

5. How to eliminate the effect of the vibrations caused by machinery on accuracy?

Overall, this research aims to develop a system that can reconstruct an accurate 3D model of any product with a reflective surface. We also aim to overcome the limitation of state-of-art techniques by reconstructing the product in the presence of ambient light.

1.4 Original Contributions

This research aims to design, implement and evaluate a novel 3D measurement system that can accurately inspect highly specular surfaces. This work is aimed at addressing the shortcoming of state-of-art inspection methods in dealing with ambient lighting conditions of the working environment. We also propose a novel fault detection method which helps the inspector in the decision-making process.

The laser line detection is a critical part of the developed system. To discover the efficiency of our system, we employed AI techniques to detect the laser line. The dataset to train supervised learning models is developed by projecting the blue light laser on different parts of the product in ambient light. After pre-processing and data cleaning, different ML regression models and 1D convolutional neural network models are trained to detect the projected laser line. The main contributions of this study are:

- Conducting a comprehensive survey of the literature covering both traditional and deep machine learning technologies for inspecting specular surfaces.
- Comparative analysis of available commercial solutions.
- Development of an accurate 3D measurement system to automatize the inspection process.
- Implementation of a novel defect detection technique for reflective surfaces.
- Evolution of a novel diameter and height Measurement technique using a single image.
- Proposing AI models for accurate detection of projected laser lines in ambient light.
- Evaluation of the developed system by performing in-depth comparison with AI models.

1.5 Structure of the Thesis

This thesis is organized into eight chapters. Chapter 2 provides a detailed review of state-of-art approaches for 3D measurement and defects detection. This is followed by an analytical and critical survey of the literature of available solutions for inspection. The theoretical concepts of deep learning and traditional machine learning regression models are also presented in this chapter.

Chapter 3 to 8 present the author's contribution to the field. These chapters include four main contributions. Chapter 3 mainly focuses on crucial details of physical setup for the novel 3D measurement system.

Chapter 4 starts with a description of the theoretical foundation of this research that exists in the area. Based on this, we developed a novel 3D measurement system. This system is called a stereo-laser system throughout the thesis. Our proposed algorithm of inspection in ambient light is also formulated in Chapter 4.

In Chapter 5, we explain a novel defects detection technique which is developed as a part of a stereo-laser system. Chapter 6 provides a detailed comparison of red vs blue light laser in stereo-laser system. Chapter 7 proposes a novel technique that measures the principal dimensions of a cylinder using a single image.

Chapter 8 describes a comprehensive analysis of deep learning and machine learning models for laser line detection.

The last Chapter 9 presents discussions; and limitations gives an overall conclusion of the thesis, and highlights the original and significant contributions. Finally, some critical future research directions are also mentioned.

Moreover, chapter-Appendix A shows the outputs obtained using SP1 stereo vision system for 3D reconstruction of highly specular surfaces in ambient lighting conditions. Appendix B presents the performance of deep learning and machine learning regression models in particular scenarios caused by laser light diffusion.

Chapter 2

Review and Theoretical Background

The chapter starts with an explanation of the reflection model, which helps understand the construction of images in ambient light. A comprehensive survey has been done on the traditional methods of 3D shape measurement. We present a brief overview of all 3D reconstruction techniques. Stereo vision and structured lighting methods are proven to be most effective compared to other surface non-contact inspection techniques in recent years. State-of-art high-dynamic range (HDR) techniques are discussed for the inspection of specular surfaces. We provide critical analysis of methods and applications in the inspection of specular surfaces and defects detection. An analytical comparison of all available commercial solutions for the inspection of specular surfaces is mentioned in this chapter. Artificial intelligence techniques have outperformed traditional techniques in different domains in recent years. We discuss the impact of the advancement of machine learning techniques on the inspection process. Moreover, a brief history of conventional machine learning and emerging deep learning approaches is also depicted in this chapter. The emerging convolutional neural networks (CNN) approaches have been used widely for solving different computer-vision problems, especially for defects detection. We summarize by outlining the effectiveness of AI in automated 3D measurement systems. Material discussed in this chapter has been published in my submitted publication [9, 10].

2.1 Reflection Model

"Reflectance models have been used for producing realistic images in the fields of computer graphics. In recent years, machine vision applications often require reflection models for image analysis and object recognition [17]." Figure 2.1 shows the dichromatic reflection model proposed by Shafer. "The medium comprises the bulk of the matter and is approximately transparent in general, while the pigments selectively absorb the light and scatter it by reflection and refraction [18]." The surface spectral reflectance of an object varies according to the geometries of illumination and viewing. Also, the material composition of the product plays a vital role. Most object's reflectance is usually decomposed into two parts: specular reflectance and

diffuse reflectance. The reflection from homogeneous materials like metals is due mostly to the specular reflectance which occurs at the interface between the object's surface and air. On the other hand, in-homogeneous materials exhibit both reflections, but diffuse reflection dominates in this case. The diffuse reflection occurs from light scattering among the pigment colorant layer of the subsurface [17, 19, 20].

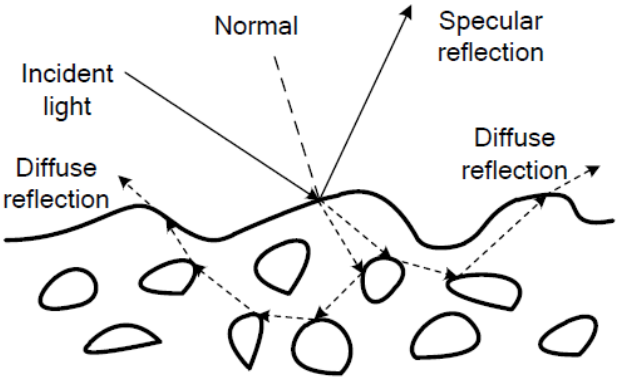


Figure 2.1: Dichromatic reflection model [18].

According to Shafer's dichromatic reflection model, reflected lights of in-homogeneous objects are a linear combination of diffuse and specular components. As a result, each pixel of the image taken by the charge coupled device (CCD) can be described as:

$$I(x) = \alpha(x) \int_{\Omega} A(\lambda, x) B(\lambda) Q(\lambda) d\lambda + \beta(x) \int_{\Omega} B(\lambda) Q(\lambda) d\lambda \tag{2.1}$$

where, $I(x)$ is the spatial parameter and the color vector of image intensity, and x is the *two dimensional* (2D) image coordinates. The weighting factors for diffuse and specular reflections are $\alpha(x)$ and $\beta(x)$, respectively. Here, λ represents the wavelength of the light spectrum, and the diffuse and specular spectral reflectance function are $A(\lambda, x)$ and $B(\lambda)$, sequentially. $Q(\lambda)$ is the three element-vector of sensor sensitivity, and the integration is done over the visible spectrum Ω.

The equation 2.1 can be simplified as:

$$I(x) = \alpha(x) D(x) + \beta(x) G \tag{2.2}$$

Here, $D(x) = \int_{\Omega} A(\lambda, x) B(\lambda) Q(\lambda) d\lambda$ and $G = \int_{\Omega} B(\lambda) Q(\lambda) d\lambda$. The diffuse reflection component is denoted by $\alpha(x)D(x)$, while $\beta(x)G$ represents the specular reflection component.

The chroma of the image in the dichromatic reflection model is defined as follows:

$$\sigma(x) = \frac{I(x)}{I_r(x) + I_g(x) + I_b(x)} \tag{2.3}$$

$\beta(x) = 0$ when only the diffuse reflection component is present in the pixel. As the chroma value is independent of the diffuse reflection weighting factor $\alpha(x)$, the diffuse reflection chroma expression is:

$$\Lambda(x) = \frac{D(x)}{D_r(x) + D_g(x) + D_b(x)}, \tag{2.4}$$

where $\Lambda = \{\Lambda_r, \Lambda_g, \Lambda_b\}$.

In the same way, when the pixels only have specular reflection component $\alpha(x) = 0$. The chroma value of the pixel is also independent of the specular reflection factor $\beta(x)$. Therefore, the specular chroma is defined as:

$$\Gamma(x) = \frac{G}{G_r + G_g + G_b}, \tag{2.5}$$

where $\Gamma = \{\Gamma_r, \Gamma_g, \Gamma_b\}$. With regards to Equations 2.4 and 2.5, Equation 2.2 can be written in terms of chromaticity:

$$I(x) = m_d(x)\Lambda(x) + m_s(x)\Gamma, \tag{2.6}$$

where $m_d(x) = \alpha(x)[D_r(x) + D_g(x) + D_b(x)]$, $m_s(x) = \beta(x)(G_r + G_g + G_b)$.

This reflectance model is used to separate the specular reflection component from the diffuse reflection component in image segmentation problems. In many 3D measurement techniques, the highlights caused by ambient light are detected in the image by applying this reflection model [18].

2.2 Traditional Approach for 3D Measurement

Over the last decade, a significant number of techniques for three-dimensional (3D) shape measurement have been proposed. These 3D measurement techniques can be classified into two categories: surface contact and surface non-contact techniques [21].

Coordinate Measuring Machine

A mechanical probe-based *coordinate measuring machine* (CMM) is one of the examples of surface contact techniques. For many years, this technique has been used to

measure the geometric shape with high accuracy. A reflective surface is not a problem for this technique. However, CMM only measures a limited number of points on the surface, and it is relatively slow compared to surface non-contact methods. Also, the surface contact method increases the chances of damage to the object [22].

For high-speed 3D inspection of objects, surface non-contact techniques are preferred. Surface non-contact techniques are further classified into active and passive techniques. Time-of-flight, laser range scanning and structured lighting are some of the examples of active surface non-contact techniques. In the active form, 3D imaging sensors generally operate by projecting electromagnetic energy onto an object followed by recording the reflected energy. On the other hand, 3D imaging sensors operate by acquiring electromagnetic energy from an object in the passive form. Photogrammetry and stereo vision are passive surface non-contact 3D measurement techniques [22, 23].

Photogrammetry

"Photogrammetry is the art and science of making measurements from photographs [24]." Photogrammetry uses one photograph or high-speed photography and remote sensing to estimate accurate 3D relative motions. "The elaboration pipeline consists basically of the following steps: camera calibration and orientation, image point measurements, 3D point cloud generation, surface generation and texture mapping [25–27]." The calibration of the camera is a crucial parameter for obtaining accurate models. The measurement procedures can be automatic or semi-automatic. Recently, photogrammetry techniques have been combined with other computer vision shape measurement techniques, to increase the measurement performance and the automation levels [28–30]. In this research, we have used the concept of photogrammetry in a calibrated environment to calculate the height of the drum.

Stereo Vision

Another surface non-contact technique is stereo vision. It is also a type of Photogrammetry technique. It is a well-known ranging method because it resembles the basic mechanism of the human eye. The computer stereo vision system uses the same principle by replacing eyes with two CCD cameras. They are displaced horizontally to obtain two different views. A disparity map is obtained by finding corresponding points in two slightly different images [31]. Stereo vision is the most accurate 3D shape measurement technique. However, it is relatively challenging to obtain an accurate disparity map for reflective surfaces [10]. This is discussed in detail in Chapter 4 and Appendix A.

Time-of-Flight

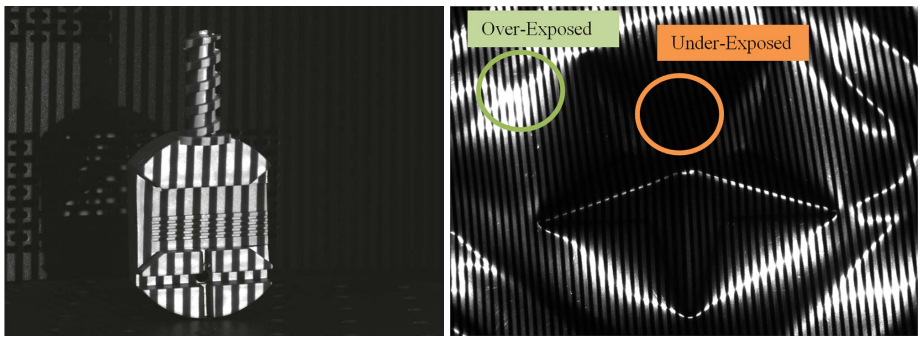
In the *time-of-flight* (ToF) technique, the time difference between the emission of a signal and its return to the sensor, after being reflected by an object, is calculated to measure the distance between the sensor and the object. The most common type of signal is light and sound [32]. Same as CMM, only a limited number of points are measured in this technique. The measurement range of large range sensors is from 15 m to 100 m. "For large measuring ranges, time-of flight sensors give excellent results. On the other side, for smaller objects, about one meter in size, attaining 1 part per 1,000 accuracy with time-of-flight radar requires very high-speed timing circuitry, because the time differences are in the pico-second range [23]." Additionally, shiny surfaces reflect little back-scattered light energy, which increases the difficulties for time-of-flight sensors [33–35].

Laser Triangulation

There are two types of laser triangulators: single-point and laser stripes triangulators. The system comprises a laser source and a camera. The camera captures the image of the projected narrow beam by the laser source. In the case of laser stripes, the laser is equipped with a cylindrical lens, which expands the light beam along one direction. Hence, a plane of light is generated, and multiple points of the object are illuminated at the same time. The location of the laser beam in the image plane is used for the measurement of the surface. The laser triangulators are very accurate and are insensitive to surface texture effects [23, 36]. The fundamentals of the laser triangulation technique and its limitations are explained thoroughly in Chapter 4.

Structured Lighting

Among these surface non-contact techniques, structured lighting is one of the most widely used 3D shape measurement techniques. It is used to perform the entire field inspection with high resolution and accuracy. It shares the active triangulation approach as mentioned above. A projector projects a sequence of coded pattern onto the object surface, and the camera captures images of the reflected pattern. The non-uniform object surface distorts the reflected pattern. The height information of the object surface is embedded in the phase distribution of the reflected pattern. Calibration is performed to understand the correspondences between the camera and the projector. The 3D coordinates of the object surfaces can be calculated by using phase to height mapping in the calibrated environment. However, when dealing with non-diffuse surfaces, the optical signal cannot be correctly retrieved. Therefore, it is usually challenging for any optical method to accurately measure shiny objects or objects with an extensive range of reflectivity variation across the surface [21].



(a) A fringe pattern image of the plastic block [21].

(b) A fringe pattern image of metallic workpiece [37].

Figure 2.2: Fringe pattern projection on different surfaces.

Figure 2.2a shows a plastic block with the projected fringe pattern image, respectively. Because of the diffuse surface of the plastic block, the camera can capture a clear image of the fringe pattern. In contrast, when a fringe pattern is projected onto a metallic workpiece, we can see in Fig. 2.2b that fringe patterns in regions with low reflection become visible while regions with strong reflection become so bright that image saturation occurs. Subject to high reflectance of the target surface, the projected patterns are usually degraded drastically in the captured images. Therefore, it is difficult to obtain 3D coordinates of a metallic workpiece with insufficient information [21].

The conventional structured light method cannot directly measure the object with shiny surfaces. There are some cost-effective options available to make the surface of the object diffuse such as spraying a thin layer of powder onto the surface or painting the object's surface. However, this supplementary step is troublesome and time consuming. Also, the accuracy is often dependent on the powder thickness and its homogeneity [38].

Table 2.1 summarizes the main characteristics of traditional shape measurement techniques. "The selection of which sensor type should be used to solve a given depth measurement problem is a very complex task, that must consider (i) the measurement time, (ii) the budget, and (iii) the quality expected from the measurement. Additionally, the sensor performance may depend on the dimension, the shape, the texture, the temperature and the accessibility of the object [23]." The strengths and weaknesses of the different techniques are specified to help the selection process. The accuracy of these techniques, when working with specular surface, is the most critical aspect of our research. The critical parameters to consider are the type of

Table 2.1: Comparison of 3D shape measurement techniques [23].

Technology	Type	Compatibility with Highly Specular Surface	Strengths	Weaknesses
Coordinate Measuring Machine	Surface contact	Yes	Highly accurate, Independent of the texture of the surface	Time-consuming, Less number of points, Can damage the product.
Photogrammetry	Surface non-contact, Passive	Not in ambient light	Simple and inexpensive, High accuracy on well-defined targets.	Low data acquisition rate, Challenging computation, Limited to well defined scenes.
Stereo Vision	Surface non-contact, Passive	Not in ambient light	Simple, fast and inexpensive, High accuracy on well-defined targets.	Challenging computation, Limited to well defined scenes, Requires high data acquisition rate.
Time-of-Flight	Surface non-contact, Active	No because reflected light scatters.	Medium measurement range, Output independent of ambient light	Medium accuracy, Slow, Measures limited number of points.
Laser Triangulation	Surface non-contact, Active	Not in ambient light	Fast and straightforward, High accuracy, Insensitive to surface texture.	Safety precautions required, Expensive, Limited range and measurement volume.
Structured Lighting	Surface non-contact, Active	Not in ambient light	High accuracy, Fast, Intermediate measurement volume.	Performance depends on ambient light, Complex, Requires safety precautions.

the method (Surface contact or Surface non-contact), the acquisition time and the safety conditions. In addition, we need to take into account if the accuracy of the measurement is affected by the ambient light conditions.

2.3 Review of Inspection Techniques for Specular Surfaces

Generally, metallic workpieces are needed to be milled in industrial applications. The surface of the workpiece becomes very bright after milling. As stated before, conventional shape measurement techniques cannot be used to measure shiny surfaces. State-of-the-art methods for 3D shiny surfaces measurement are typically referred to as high-dynamic range (HDR) techniques. In standard photographic techniques, no features are visible outside a specific range of brightness. The bright area appears pure white, and the dark region appears pure black. HDR techniques are used to reproduce a more extensive range of luminosity compared to standard photographic techniques. Below, the work done for the 3D measurement of shiny surfaces using different HDR techniques is described briefly [21].

HDR Techniques using Multiple Exposures

In this technique, a sequence of images is captured at different exposures. The brightest unsaturated intensity value is selected to combine into a single set of HDR images at each pixel. This technique was proposed by Zhang and Yau in 2009. The

exposure time is manually adjusted to acquire high-quality fringe patterns. Here, the adjustment of lens aperture is avoided as it can cause undesired motion between system components, affecting the system's calibration. "The images taken at low exposure times contain valuable information from regions on the surface with high reflectivity, whereas the images taken at high exposure times contain useful information from regions on the surface of low reflectivity [39]." We cannot determine the proper exposure time, since the used exposure time is subjectively selected. In some cases, a large number of exposure times are required to synthesize the HDR images, which makes the process time-consuming [39]. In an experiment, 23 exposures were needed to obtain a high-quality point cloud of a China vase. With a three-step phase shifting algorithm and three-frequency heterodyne phase-unwrapping algorithm, a total of 3x3x27=207 fringe images are required to synthesize the HDR images.

Another research was carried out by Liu et al. in which they used a dual-camera structured light system along with the use of multiple exposures to acquire the 3D shape of shiny objects. Here, two camera-projector monocular systems were operated from different viewing angles to fill in missing data from regions with specular highlights and diffused darkness [40].

When inspecting an unknown scenario, it is difficult to determine the initial values of the exposure time. Generally, it is recommended to use the exposure time as much as possible. A technique of using a histogram to predict one or more optimal exposure times based on the reflectivity distribution of the surface was developed by Feng et al. [43]. These optimal exposure times were used to capture raw fringe images to synthesize HDR images.

Alternatively, a technique for capturing raw fringe images by simultaneously adjusting the exposure time and the projected light intensity was proposed by Jiang et al. Instead of the brightest unsaturated intensity; this technique takes the highest modulation intensity as the criterion for selecting the pixels from the raw fringe images to synthesize the HDR images. The effect of ambient light is minimized with this approach. However, it requires five times the number of fringe images compared with the conventional phase measurement. The process becomes time consuming because of the required number of fringe patterns [41].

To solve this time-consuming problem, an auto-exposure technique to predict the required exposure time automatically according to the reflectivity of the measured object surface is presented by Ekstrand and Zhang. However, the single predicted exposure time does not always fit the surface with an extensive range of reflectivity variation [42]. Zhong et al. proposed a similar approach by selecting the optimal exposure time as large as possible. However, these techniques are limited to improve the *signal-to-noise ratio* (SNR) in the region of weak reflection [44].

Song et al. also presented an HDR based structured light approach for the precise

3D reconstruction of specular targets. The image intensity-based sinusoidal pattern is vulnerable to the reflective nature of shiny surfaces. "Therefore, the robust binary strip shifting method is adopted as the *structured light system* (SLS) coding strategy [37]." Multiple images of the target object are captured with various camera exposure settings. The estimated camera response function is used to fuse multiple-exposure images into a single image. The image with a middle exposure level is selected as the reference image and used to slightly adjust the primary fused images to make all the SLS images comply with unified lighting conditions. Later, 3D reconstruction of the specular surface is obtained based on the HDR structured light images [37].

The spatial light modulator has also been used to control the camera exposure time at individual pixels. The final exposure setting at saturated pixels is determined from multiple iterations. However, the technique developed by Jeong et al. required additional optical and control hardware [45].

Apart from post-processing the captured raw fringe images, this technique efficiently performs shape measurement by adjusting the exposure time. However, we cannot set the exposure time by using a *digital light processing* (DLP) projector to project fringe patterns. Another requirement is that the camera and projector be synchronised to precisely capture the projected grayscale values [44].

HDR Techniques by Adjusting Projected Pattern Intensities

In this technique, the *maximum input gray level* (MIGL) of projected sinusoidal fringe patterns is modified to accommodate variable ambient illumination. Waddington and Kofman proposed this technique to avoid image saturation [46–48]. Later, they improved this technique by combining the MIGL reduction and pixel-by-pixel approaches. In an improvised version, they uniformly adjusted the projected fringe pattern intensity and selected the brightest unsaturated pixels from raw fringe images captured at different illuminations to synthesize the HDR images. Similar to multiple exposure techniques, these techniques are also proven to be time consuming [49, 50]. Li and Kofman developed another technique; where the MIGL in projected patterns is modified locally according to the local reflectivity of an object surface. This is better than the uniform reduction of the MIGL for saturated regions.

Babaie et al. proposed a new method for 3D shape measurement. Here, the intensities of the projected fringe pattern are adjusted at pixel level based on the feedback from the reflected images captured by the camera [51]. Another adaptive technique was developed by Lin et al. In this technique, the projected fringe pattern's pixel-wise intensity is adjusted based on the camera response function according to the reflectivity of the surface and to the illumination of ambient light and

surface inter-reflections. Also, the required number of captured images to obtain the point cloud is less compared to other techniques [52].

"An advantage of adjusting the projected pattern intensities is that the adjustment can be performed pixel by pixel locally, rather than that of the adjustment of the camera exposure time, which is global [21]."

HDR Techniques using Polarizing Filters

Polarizing filters are used to eliminate the effect of highlights on shiny surfaces. The angle between the transmission axes of polarizers is changed. The basic idea is to limit the reflected light incident on the image sensor at a certain angle to effectively remove the intense specular light. However, it reduces the captured intensity for the whole image, making the measurement process for dark regions very difficult [53, 54]. "On the other hand, a parallel polarizer-analyzer alignment [57, 58] can maintain good fringe quality at the dark regions, but not in the bright places." As a solution, the combination of techniques using multiple exposures and using polarizing filters was developed by Feng et al [55].

In 2014, a multi-polarization fringe projection imaging system was developed by Salahieh et al. in which saturated points are eliminated, and the fringe contrast is enhanced by selecting the properly polarized channel measurements. The fringes are linearly polarized before projecting on the object. After that, the camera equipped with a pixelated polarizer array captures the projected fringe patterns. However, the hardware complexity increases if the object's optical properties are unknown as the adjustment of the polarizer depends on it [56].

HDR Techniques using Color Invariants

In this technique, the color of a point is determined by the color of the object itself and the color of the light source. It is based on the dichromatic reflection model proposed by Shafer [20]. "Diffuse reflection indicates the color information related to the object itself, whereas the specular reflection represents the color information for the light source [59]." In 2011, Benveniste and Unsalan introduced a novel color invariant to detect the red stripe, green stripe, and highlights to segment robustly the stripe from the captured image. The effects of highlights originating from the ambient light were removed by using this color invariant in segmentation. They developed methods based on binary, ternary and quaternary color-coded patterns can eliminate the effect of highlights and ambient light in the scanning process. However, the output of this technique is inevitably affected by the object itself if it contains complex textures and multiple colors [60].

Photometric Stereo Technique

This technique can recover both the 3D shape and albedo of surfaces. In this technique, multiple images at a fixed viewpoint under illuminations with different directions are obtained. This technique aims to detect specular components or to separate reflecting components [61]. Here, we require one static camera and multiple light sources. These light sources are switched on and off while capturing images [191]. However, the reflectance of the surface was required to be Lambertian. For non-Lambertian surfaces, it is challenging to estimate the reflectance robustly.

This technique aims to detect specular components or separate reflecting components by observing varying *bidirectional reflectance distribution functions* (BRDF). At an opaque surface, the reflection of light is defined by BRDF function [62]. Generally, BRDFs are complicated nonlinear functions involving seven or more variables. An efficient photometric stereo method to estimate BRDF parameters without the 3D geometric model of the object was invented in 2010 by Li et al. For surfaces with both varying specular and diffuse properties, Goldman et al. proposed a method with spatially varying BRDFs. "They modeled the surface reflectance as a linear combination of fundamental materials and recovered the shape, material BRDFs, and weight maps using the well-known Ward model [65]."

Sun et al. proposed a new algorithm based on *reflection component separation* (RCS) and priority region filling in 2017. After finding the specular pixels in the image, the reflection components are separated. However, this process changes the color information of highlight due to more significant reflection components in objects with strong specularity. The color information is restored using the priority region filling theory which is independent of polarization or image segmentation [18].

Hin-Shun and Jiaya presented a new photometric stereo technique in which an important visual cue, that is, the cast shadow silhouette of the object, was employed to optimally recover global BRDF parameters and thereby was capable of reconstructing convex object surfaces with different levels of roughness. Here, they did not apply unreliable highlight separation and specularity estimation. However, this technique requires high computing [63].

Another technique using a collimated illumination source and a plenoptic camera was developed in 2015 by Meng et al. This system realizes a single-shot shape measurement for specular surfaces. The camera captures BRDF variation of the object surface in a single image in addition to the light field information from the scene. However, their system only works for relief surfaces and will have problems to handle the surfaces with steep slopes [64].

There are multiple ways to construct HDR images without saturation from different captured images. For example, multiple cameras, color light sources, color

filters [66], and multiple light projection directions [67], have been used to handle specular reflections.

Phase Measuring Deflectometry Technique

Phase measuring deflectometry (PMD) based methods are widely popular because of their advantages: an extensive dynamic range, non-contact operation, full-field and fast acquisition, high accuracy, and automatic data processing. Zhang et al. provided a thorough review of different PMD techniques in 2017. The flow chart shows the procedure of a classical PMD technique (See Fig. 2.3). "The classical PMD technique uses the full-field fringe patterns to measure the slope information and then slope integration to obtain the 3D shape data of the specular objects [68]." A *charge couple device* (CCD) camera captures the modulated fringe patterns that appear deformed regarding the slope variation of the measured surfaces. Phase shifting algorithms are used to demodulate the phase information from captured fringe patterns. Spatial phase unwrapping and temporal phase unwrapping techniques are used to get absolute phase data. This data is later applied to calculate the slope of the measured surface using the system parameters. The slope data is integrated to reconstruct the 3D shape of the specular surface. "However, the classical PMD cannot directly measure complicated specular objects having discontinuous and/or isolated surfaces because of the procedure of slope integration [68]."

In 2018, an advanced PMD method was presented by Niu et al. Bi-telecentric lens were used to obtain 3D shape of a discontinuous specular object. Here, *liquid crystal display* (LCD) screen is used to display sinusoidal fringe patterns. The camera captures the deformed fringe patterns reflected by the measured specular surface. The advanced method proposed can calculate the depth data directly from absolute phase data [69].

Since visible light is sensitive to ambient light, Chang et al. developed *Infrared-Phase measuring deflectometry* (IR-PMD) technique which exploits IR as the light source. The discontinuous specular surfaces are directly measured from phase data by moving an IR camera. An IR projector projects sinusoidal patterns onto a ground glass which is regarded as an IR display screen. Here, IR digital display is moved to substitute the camera and the measured object's movement [70].

Shape from Multiple Images

In 2015, Bulbul et al. proposed a new technique for the reconstruction of 3D specular spherical surfaces using structure from motion software and the concept of hough transform. Firstly, a 3D point cloud is created from multiple images. "Each element of this point cloud has vertex, color RGB and normal information. From this infor-

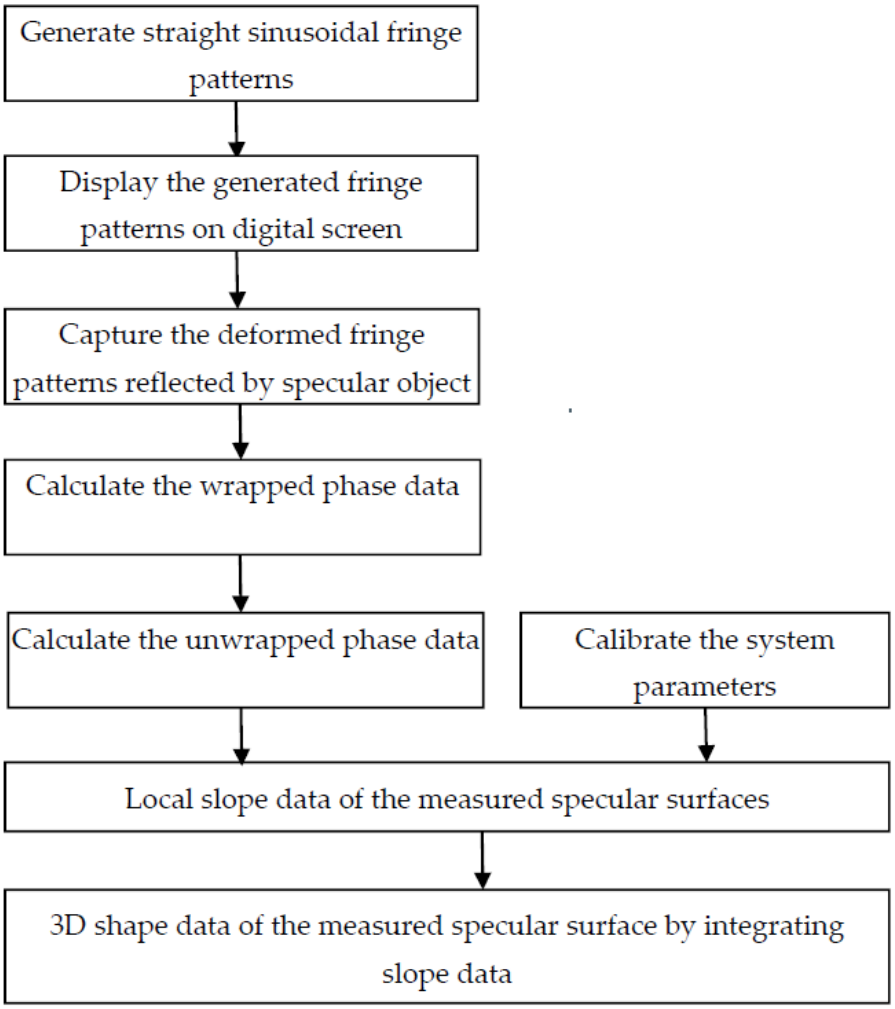


Figure 2.3: Flowchart of 3D shape measurement of specular objects by using classical PMD.

mation, a Hough Transform like process was proposed to infer the centres and radii of the spheres [71].”

Scanning from Heating

An infrared system was developed as an extension of scanning from heating technique for 3D scanning of metallic surfaces. Here, the surface of the product is locally heated by a laser source. The thermal emission of a surface is measured using an IR camera. "Moreover, the study of thermophysical properties of metals, the modeling of the thermal problems and the simulations are required to determine the characteristics of the laser (wavelength, power) [72]."

Dual Camera Fringe Projection System

In 2016, Feng et al. developed a novel dual camera fringe projection system to measure shiny objects from two different directions. "The erroneous phase obtained from a saturated pixel is corrected by the phase of its corresponding pixel in the other view which is free from the saturation problem [73]." Here, sub-pixel matching is performed for finding corresponding high light intensity areas in cameras to achieve high accuracy [73].

Recently, Xu et al. invented a technique in which a set of sinusoidal patterns with maximum intensity are projected onto the target object. The dichromatic reflection model is used to signify specular pixels in each image. The adaptive fringe projection technique is carried out to minimize the specular reflection components. Pixel inpainting is used as an alternative technique for the region that the adaptive fringe projection technique cannot handle. However, this approximation technique does not work accurately with discontinuous or irregular surfaces [74].

Table 2.2 summarizes various techniques for the inspection of highly specular surfaces. Our stereo-laser system is also mentioned for comparison purposes. These techniques are compared on the basis of achieved accuracy and complexity. The number of required projections is also specified for most of the techniques. Whether the system works in ambient light condition of working environment or not is one of the most critical parameters. The cost and complexity of the system are depended on these critical parameters. Even though the best accuracy is achieved with photometric stereo technique, the measurement volume size is really small in this case. Also, 30 projections in different directions are required. This makes the system more complicated. Therefore, we can mention that our stereo-laser system has the overall best performance in terms of achieved accuracy which is 0.02 mm within a large measurement volume size. Additionally, the system can inspect highly specular surfaces in ambient light, and it requires the lowest number of projections per scan. Therefore, the stereo-laser system provides a simple but efficient solution for the quality control program of large-scaled production lines.

Table 2.2: Techniques for the inspection of highly specular surfaces.

Technique	Reference	Number of required projections	Achieved Accuracy	Complexity	Works in ambient light
Stereo-Laser System	[10]	One projection for one scan.	0.02 mm within $600mm * 600mm * 270mm$ measurement volume size.	More simple.	Yes
HDR Techniques using Multiple Exposures	[41]	84 projections.	0.016mm within $100mm * 40mm * 31mm$ measurement volume size.	Complicated calculations.	Yes
	[37]	30 projections.	0.06 mm		
HDR Techniques by Adjusting Projected Pattern Intensities	[52]	31 projections are required if four-step phase shifting and heterodyne phase unwrapping algorithms are used.	0.012mm within $320mm * 30mm * 150mm$ measurement volume size.	Simple, because it only involves the matrix inverse operations.	No
HDR Techniques using Polarizing Filters	[43]	13 projections are required, one of which is uniform white light.	0.045 mm within $300mm * 30mm * 250mm$ measurement volume size.	More simple.	Not specified
HDR Techniques using Color Invariants	[60]	It needs 4 projections if quaternary color-coded patterns are used.	0.506 mm within $750mm * 650mm * 200mm$ measurement volume size.	Most simple.	No
Photometric Stereo Technique	[64]	30 projections in different directions.	0.005 mm within $12.5mm*10mm*1mm$ measurement volume size.	More complicated.	Yes
	[18]	Not specified.	The highlight pixel numbers are decreased by 10.1 times.	Removes specular regions.	No
Phase Measuring Deflectometry Technique	[68]	Not specified.	High accuracy, fast and automatic data processing.	Difficulties with discontinuous specular surfaces.	No
	[70]	24 projections.	0.022 mm, when measuring a step of size 6.099 mm.	Expensive and complex.	No
	[69]	12 projections.	0.0233 mm, when measuring a step of size 7.0248 mm.	Expensive and complex calculations.	No
Shape from Multiple Images	[71]	122 input images.	Accuracy is better than traditional VisualSFM.	Complex,and applicable only to spherical surfaces.	Yes
Scanning from Heating	[72]	Not specified.	Accuracy is lower compared to other techniques.	The power of the laser depends on the material of the product. This method is not applicable to steel products.	No
Dual Camera Fringe Projection System	[74]	Not specified.	0.566 mm, when trying to measure a ceramic ball of radius 34.992 mm.	Error increases when fitting too large specular reflection areas.	No

Defects Detection Techniques

The automatic detection of defects on specular surfaces is currently one of the greatest issues faced in the automotive industry. During the last decade, several systems have been developed to solve this issue. "However, these, to the best of our knowledge, have been focused solely on flat surfaces and have been unable to inspect other parts of the surfaces, namely style lines, edges and corners as well as deep concavities [78]."

In 2008, a method was proposed for real-time defect detection on highly reflective curved surfaces by Rosati et al. Here, the rays emitted from a light source are conveyed on the surface by a curved mirror. The camera collects the reflected rays from the surface. The defects appear as a shadow of different shapes and sizes in the image [75].

Another method to inspect chrome-coated highly reflective ring components was developed by Boby et al. in 2010. They compared four different algorithms: Fourier filtering, Image convolution, auto-median and single-step thresholding based on the efficiency of defect classification and speed [76].

Recently, a method based on image library matching was developed to identify defects on the highly reflective roller surface. Firstly, the initial images are preprocessed by using the accelerated optimized bilateral filtering. Next, a rough segmentation of defects is done by using the entropy and the grid gray gradient information. "Finally, a defect fine identification method based on the Hu invariant moment matching integrated with morphological classification is proposed for achieving image library matching and further quickly removing the pseudo defects [77]." Shao et al. achieved an accuracy of 98.2% by using this method [77].

Molina et al. introduced a novel approach that computes defects in less than 15 seconds using deflectometry and vision-based technologies. In the first step of the process, an image fusion algorithm is proposed to enhance the contrast between the defected region and the region with no defects. Next, a local directional blurring method is used for image background extraction. This step is repeated several times to detect defects of different sizes each time [78].

In 2016, Xu et al. presented a polarized light filtering based method for automatic visual inspection of the cylindrical smooth surface of components such as pistons. This method enhances the defect image contrast, enlarges the detection range and suppresses the flawless area noise. Due to the characteristics of a favorable light wavelength, smallness and long life, a white rectangular *light emitting diode* (LED) light source was adopted in this vision system. The picture contrast was enhanced by applying polarizers to the camera lens and the vertical light source. "A series of specific visual image processing algorithms were developed to extract the features

of surface defects based on the acquired images. The defects can then be recognized and classified into different categories based on the defect database [79]."

"The surface imperfections provoke important light rays' deviations [80]." Based on this property, Seulin et al. developed a particular lighting system to inspect specular surfaces for defects. The product is placed inside a lighting tunnel for the inspection. The binary type lighting is composed of a succession of zones of null luminous intensity and zones of maximal luminous intensity. A defect appears as a set of luminous pixels among a dark zone or a set of dark pixels among a luminous zone [80].

A new vision defect inspection system has been shown in [81]. A synthetic is reconstructed from a set of images captured under various illumination conditions. The basic algorithm is to integrate all defects in a single synthetic image. The combination of morphology technique and template matching is proposed to achieve defects inspection from the synthetic image.

Table 2.3 shows selected publications on defects detection in reflective surfaces. The technology and required hardware are mentioned to determine the complexity of the method. Here, the accuracy of the system is denoted by ACC for [77]. We can see that our method achieves the best accuracy by detecting the defects of size 0.02 mm. All the other methods have one common disadvantage that they are non-measurable processes. In contrast to our method, these methods can accurately detect the defects but cannot measure the dimension of defects. Also, it is significantly challenging to detect palpable defects using vision-based techniques. Furthermore, these techniques are not a part of the 3D reconstruction process. This adds to the time of the inspection process. However, our defects detection technique is a part of stereo-laser inspection system. Therefore, the 3D model of the product is reconstructed along with the defect detection process.

2.4 Commercial Solutions for 3D Measurement

Many commercial solutions are available in the market for inspecting reflective surfaces. Trevista Surface 50 is one of the most popular solutions. This system of Stemmer imaging combines a structured lighting arrangement with shape-from-shading technology to inspect surface defects on metal components. This system can perform an automatic inspection of objects over a 50 mm measurement field. The hardware is a piece of dome-shaped equipment that ensures optimum and even illumination from different directions and suppresses interference from ambient light [82].

Helmee designed another popular vision system using structured lighting techniques with phase measuring deflectometry. First, the object to be inspected is picked up and placed under a 180° reflective semi-transparent hemispherical dome.

Table 2.3: Defects detection techniques for highly specular surfaces.

First Author / Year	Reference	Required Hardware	Technology	Defect Size	Works in Ambient Light	System Function
Dawda 2020	[9]	Stereo Cameras, Blue light laser.	Stereo matching of blue line laser.	0.02 mm.	Yes	Detects and measures defects.
Shao 2020	[77]	Camera, LED light source, several optical lenses, *digital micro-mirror device* (DMD) and a *total internal reflection* (TIR) prism.	Bilateral filtering, Hu invariant moment matching.	ACC = 98.2%	Yes	Detects only.
Molina 2017	[78]	QEyeTunnel System.	Deflectometrty and vision-based techniques.	0.2 mm	Yes	Detects only.
Xu 2017	[79]	LED light source, hood and a camera.	A polarized light-filtering based method.	Not specified.	Yes	Only enhances the appearance of defects.
Zhang 2015	[81]	Camera and a group of light sources.	Morphology and template matching techniques.	1 mm	Yes	Detects only.
Boby 2010	[76]	Camera and halogen lamp.	Fourier filtering, image convolution, auto-median and single-step thresholding.	Not specified.	No	Detects only.
Rosati 2008	[75]	Camera, lamp and a curved mirror.	Defects appear as a shadow.	Not specified.	No	Detects only.
Seulin 2002	[80]	Camera, lighting tunnel, fluorescent tubes and light diffuser.	Structured lighting with adaptive lighting system.	Not specified.	No	Detects only.

In this system, the object is illuminated with several phase-shifted sinusoidal patterns using three projectors. This creates an illusion of texture on the smooth surface of the product. The slope variation of the surface deforms the projected fringe patterns. By analyzing the images of these deformed fringe patterns, accurate measurements are obtained [83].

Table 2.4 compares the commercial products available in the market for the inspection of surfaces. Here, only the products that work with specular surfaces are

Table 2.4: Comparison of available commercial solutions.

No.	Products	Ref.	Technology	Compatibility with Highly Specular Surface	Field of View		Accuracy	Working Distance	Time	In Ambient Light	Quantity	Cost (NZD)	Amount (NZD)
					Min	Max							
1	ATOS 5	[84]	Blue LED structured lighting	only matt finished surface	$170 * 140mm^2$	$1000 * 800mm^2$	$0.002mm - 0.06mm$	880 mm	0.2 seconds per scan	No	1	275,316	275,316
2	ATOS Triple Scan	[85]	Dual laser line triangulation with blue structured lighting	only matt finished surface	$6 * 6mm^2$	$40 * 40mm^2$	$0.01mm - 0.61mm$	-	-	No	9	226,413	2,037,717
3	Gocator 2800	[86]	Dual laser line triangulation with red laser line	-	$390mm$	$1260mm$	$0.092mm - 0.488mm$	800 mm	-	-	1	19,253	19,253
4	Gocator 3200	[87]	Blue LED structured light snapshot sensors	-	$71 * 98mm^2$	$100 * 154mm^2$	$0.06mm - 0.09mm$	110 mm	Minutes	-	2	20,324	40,648
5	Gocator 3504	[88]	Blue LED structured light snapshot sensors	-	$12.1 * 13.2mm^2$	$13 * 15mm^2$	$6.7\mu m - 7.1\mu m$	7 mm	Minutes	-	18	29,416	529,488
6	Gocator 3506	[89]	Blue LED structured light snapshot sensors	-	$27 * 45mm^2$	$30 * 45mm^2$	$0.02mm - 0.025mm$	25 mm	Minutes	-	6	24,600	442,800
7	Ensenso N30	[90]	Structured lighting with blue LED random dot pattern projector	-	-	-	-	-	-	-	-	11,232	-
8	Ensenso N35	[90]	Structured lighting with blue LED random dot pattern projector	-	-	-	-	-	-	-	-	12,943	-
9	Ensenso X30	[91]	Structured lighting with blue LED random dot pattern projector	-	-	-	-	-	-	-	-	23,7476	-
10	Stereo-Laser System	[9]	Stereo matching of blue line laser	Yes	$70 * 35cm^2$		0.02 mm	40 cm - 50 cm*	Minutes**	Yes	1	11,350	11,350
Height of Drum = 27 cm.													
*Working distance depends on the size of the product.													
**Time can be improved by using higher *frames per second* (fps) camera with *graphics processing unit* (GPU).													

considered. The products developed by ATOS only work on objects with matt finished surfaces. All products work on the concept of sheet-of-light or stereo vision. The – in the table indicates that the information is not available. The cost of all products are in *New Zealand dollars* (NZD). The quantity specifies the number of laser profilers required to inspect the washing machine drum of height 27 cm. As shown in Table 2.4, there is always a trade-off between output accuracy and the *field of view* (FOV) covered by the system for inspection. To detect palpable defects, the accuracy of the output must be high. In this case, the FOV covered by the product would be relatively small. To inspect the washing machine drum of size 27 cm, we

would require multiple units of the laser profilers (See Table 2.4). The total number of units required for each product is calculated in Table 2.4. This will significantly increase the cost of the inspection process. Moreover, most of these products do not work accurately in ambient light conditions. We would need to create a suitable working environment with no ambient light to remove the reflection problem to use these products. All commercial products are compared with our stereo-laser system in terms of type of surface, working environment, FOV, accuracy and cost. From the comparison, we can summarize that our stereo-laser system performs accurately in ambient light on specular surfaces. Moreover, our system is one of the most cost-efficient solutions compared to other available commercial solutions.

2.5 Artificial Intelligence for Inspection

Artificial intelligence (AI) is the term invented by John McCarthy in the year 1950 [92]. "AI is a bigger concept to create intelligent machines that can simulate human thinking capability and behavior [93]." The subset of AI which is known as machine learning is an established research discipline. It allows machines to learn from data without being programmed explicitly. The algorithm on which machine learning works is developed by learning from historical data [93]. From advanced web search engines, recommendation systems, understanding human speed to self-driving cars, AI has grown rapidly in every field [94]. Machine learning techniques have been used to create and develop smart factories within manufacturing industry. The creation of smart factories has inspired research that analyzes process data collected from *internet of things* (IOT) sensors, to predict product quality control in real time [95]. Supervised, unsupervised, semi-supervised, and reinforcement learning are various types of machine learning techniques. The nature and characteristics of data determine the effectiveness, performance and efficiency of a machine learning solution. "In the area of machine learning algorithms, classification analysis, regression, data clustering, feature engineering and dimensionality reduction, association rule learning, or reinforcement learning techniques exist to effectively build data-driven systems [96]." Deep learning, which is part of a broader family of machine learning methods, can intelligently analyze the data on a large scale [96]. In this section, we shall begin with the basics of a fundamental tool of DL, which is deep and conventional convolutional neural networks. Moreover, an explanation of selected traditional supervised machine learning regressors is followed by a brief review of work done for quality control.

Deep Conventional Machine Learning for Inspection

"Deep learning is part of a wider family of *artificial neural networks* (ANN) based machine learning approaches with representation learning [96]." The concept of neural networks has its root in artificial intelligence [97]. "Neural networks are a series of algorithms that mimic the operations of a human brain to recognize relationships between vast amounts of data. A "neuron" in a neural network is a mathematical function that collects and classifies information according to a specific architecture [98]." ANNs are comprised of three types of node layers: an input layer, one or more hidden layers, and an output layer. Here, each artificial neuron or a node connects with another node. Every node has an associated weight and threshold. The node is activated when the output of any individual node is above the specified threshold value. Only the activated node can send data to the next layer of the network [101]. While a single layer neural network can still make approximate predictions, the accuracy is refined by additional hidden layers [99]. The adjective "deep" in deep learning refers to the use of multiple layers in the network [100]. These multiple layers are used to extract higher-level features from the raw input progressively.

Machine learning and deep learning models are usually categorized as supervised learning, unsupervised learning and reinforcement learning [99]. "Almost all the value today of deep learning is through supervised learning [103]." The labeled datasets are required in supervised learning to make predictions. The input data needs to be labeled correctly for accurate predictions [99]. "A supervised learning algorithm analyzes the training data and produces an inferred function, which can be used for mapping new examples. This requires the learning algorithm to generalize from the training data to unseen situations in a "reasonable" way [102]." The training data has multiple pairs consisting of an input object and a desired output value.

The type of training data plays a vital role in supervised learning. In this research, we are working with sequential data, which generally does not quite fit the supervised learning framework. The reason being the training data consists of N examples of (x_i, y_i) pairs, where $x_i = (x_{i,1}, x_{i,2}, ..., x_{i,T_i})$ and $y_i = (y_{i,1}, y_{i,2}, ..., y_{i,T_i})$. "These sequences exhibit significant sequential correlation, which means nearby x and y values are likely to be related to each other [104]." Moreover, sequential supervised learning is often confused with time-series prediction. However, there are two key differences. "First in sequential supervised learning, the entire sequence $(x_1, ..., x_T)$ is available before we make any predictions of the y values, whereas in time-series prediction, we have only a prefix of the sequence up to the current time $t + 1$ [104]." Second, we are not given any y values in sequential supervised learn-

ing, and we must predict them all. On the other hand, we have the true observed y values up to time t in time-series analysis.

Supervised learning techniques are further categorized into classification analysis and regression analysis. In classification analysis, a class label is predicted for a given example as an output. On the other hand, a continuous result variable is predicted based on the value of one or more predictor variables. "The most significant distinction between classification and regression is that classification predicts distinct class labels, while regression facilitates the prediction of a continuous quantity [96]." Regression models are widely used in various real-world scenarios such as weather condition predictions, cost estimation, marketing trends analysis, time series analysis and many more. In this research, we have applied supervised learning techniques for regression analysis.

Multi layer perceptron (MLP), *convolutional neural network* (CNN) and *long short-term memory recurrent neural network* (LSTM-RNN) are the most common deep learning algorithms [96]. During this research, we have applied these algorithms to *one-dimensional* (1D) data. Therefore, instead of CNN, we will discuss *one-dimensional convolutional neural network* (1D-CNN), which works explicitly with 1D data. Furthermore, the basics of LSTM-RNN will also be explained briefly.

Multi-Layer Perceptron

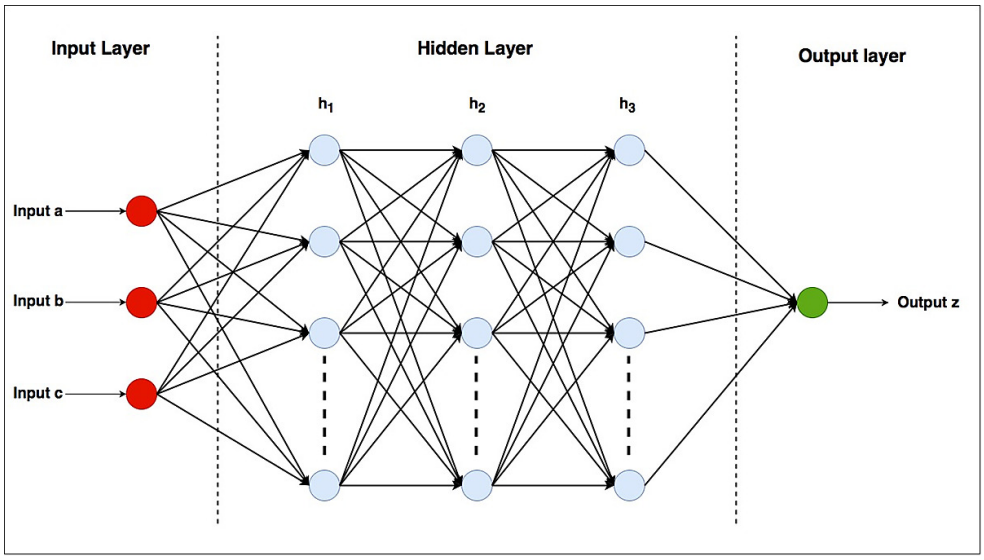


Figure 2.4: A structure of a MLP network [96].

Multi layer perceptron (MLP) is a feed-forward artificial neural network that generates a set of outputs from a set of inputs. A typical MLP is a fully connected layer that consists of at least three layers: input layer, output layer and hidden layer. MLP is a deep learning method which utilizes back-propagation technique to train the network. Generally, MLP has one or more hidden layers.

As shown in Fig. 2.4, each node in one layer connects to each node in the following layer at a certain weight [96]. Additionally, each node uses a nonlinear activation function except for the input nodes. "Each neuron in the hidden layer transforms the values from the previous layer with a weighted linear summation, followed by a non-linear activation function like the hyperbolic tan function [105]." Moreover, MLP is sensitive to weight initialization and feature scaling. Also, it requires a variety of hyper-parameters to be tuned such as the number of hidden neurons, layers and iterations [105].

One-Dimensional Convolutional Neural Network

Convolutional neural networks (CNN) have become the de-facto standard for computer and machine learning applications. Same as MLP, CNNs are also feed-forward ANNs and are trained using a back-propagation algorithm. [106]. The hidden layers of CNNs are typically consisting of convolution, pooling and fully connected layers [107, 108]. CNN was mainly designed for 2D data processing. Deep 2D CNNs with many hidden layers and millions of parameters can be trained on a massive size database to learn complex objects and patterns. "Rather than using handcrafted features, deep learning methods, specifically CNNs, provide a structure in which both (feature extraction and prediction) are performed together in a single body block, unlike other traditional methods [112]." Yet, this may not be a viable option when the training d-ata is scarce 1D signals [106]. To overcome this issue, a modified version of 2D CNNs, called 1D convolutional neural network (1D-CNN) has recently been developed [109–111]. 1D-CNNs are able to learn challenging tasks involving 1D signals with relatively shallow networks and less computational complexity. Due to their low computational requirements, 1D-CNNs are suited for real-time and low-cost applications. Therefore, 1D-CNNs are advantageous and thus preferable to their 2D counterparts in dealing with 1D signals [107].

In 1D-CNNs, sequential or time-series data is fed to the convolution input layer. The convolution layer consists of multiple kernels of the same size, followed by pooling layer, which will perform (average or max) pooling method. Both convolution and pooling layers are mainly used for features extraction from the input. The output of pooling layer is later sent to the prediction or fully-connected layer [111, 114]. "The core system consists of 1D-CNN that can merge the two major blocks: feature the extraction and prediction of a 1D signals task into a single

learning system. This prevents the computing burden and the problems arising from the choice and design of handcrafted features and maximizes prediction performance [113]."

Convolutional Layer

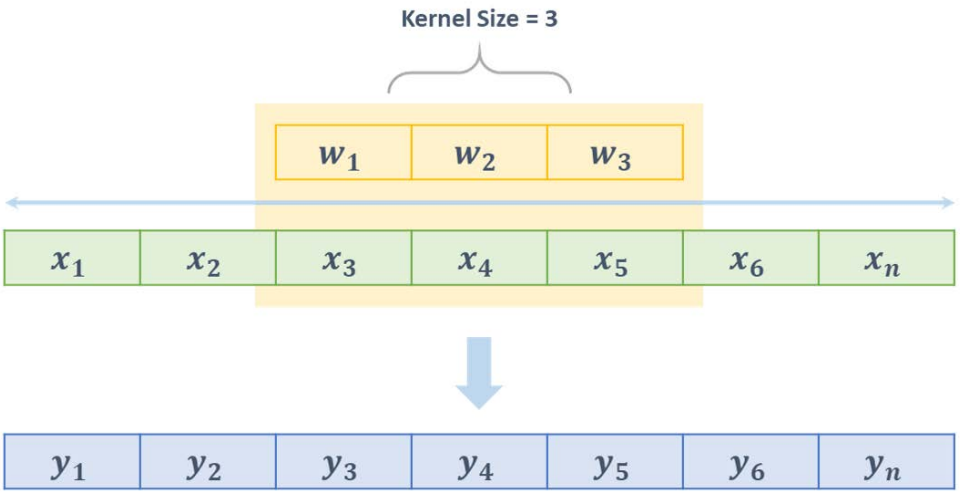


Figure 2.5: 1D convolution operation [112].

1D-CNN performs convergence over the local area of input signals to create the corresponding feature. With the use of weight sharing, 1D-CNN requires fewer parameters to converge with compared to traditional neural network models; thus, the convergence of 1D-CNN is faster. "Every kernel has unique characteristics on the *feature map* (FM) in all locations [109, 112]."

An example of 1D convolution operation is shown in Fig. 2.5. The kernel of size 3 is used, meaning all weights (w_1, w_2, w_3) will be shared by every stride of the input layer $(x_1, x_2, ..., x_n)$, and the outputs $(y_1, y_2, ..., y_n)$. The feature map value is generated by summing up the values of weights multiplied with input values. In the example shown, the feature y_4 is obtained from $y_4 = (w_1x_3 + w_2x_4 + w_3x_5)$. The number of parameters for training are reduced because of the weight-sharing function. Assuming L_i is a 1D convolution layer, the formula is demonstrated in Equation 2.7.

$$x_j^i = f(\Sigma_{i=1}^{M} x_i^{l-1} \cdot k_{ij}^l + b_j^l) \tag{2.7}$$

The output of the convolution layer is feed forwarded to the next layer as input. The Algorithm 1 shows the training procedure of 1D-CNN model. The input sig-

Algorithm 1. Training procedure for 1D-CNN model.

```
* Input: Training and validation dataset
* Output: Trained 1D-CNN model
1 Initialize: weights and biases
2 For each iteration Do:
3       Process records of the training data
4       Compare actual and predicted values
5       Calculate loss function
6       Back-propagate error and adjust weights
7       If better loss Do:
8               Save network (model, weights)
9       End If
10End For
```

nal convolution operations are carried out to get one-dimensional features. Also, various kernels extract specific characteristics from the input signals [115].

Pooling Layer Pooling layers are crucial to speeding up the CNNs [111]. Also. Pooling methods are intended to solve the issue of over-fitting. Pooling methods can be considered down-samples to minimize the number of parameters while maintaining the major features. Since there are more feature maps in the downstream sampling phase, leading to an increased data dimensionality, which does not make calculations possible [112]; therefore, each feature map is processed at this phase with either average or max pooling method [116]. The output of the max-pooling method selects the highest parameters in the default window. On the other hand, the output value of average-pooling method is based on the pre-determined pooling window.

Figure 2.6 shows an example of max-pooling operation. The length of the input layer is 6 and the pooling size is set to 2. The output values will be reduced to 3, and the maximum values for the next network layer will be selected.

Dropout Layer The dropout method can also resolve the problem of over-fitting during the training phase [117, 118]. The dropout layer is one kind of hidden layer between the input and output layers. During training, the dropout approach involves random selection of neurons. The disabled neurons have no output values or values are set to zero. The connections from disabled neurons are also removed. The neural network still functions effectively, although some neurons are temporarily disabled [119–121].

Fully Connected Layer In computer vision applications, the *fully connected* (FC) layer of convolutional neural networks is a vital component [122]. "The CNN pro-

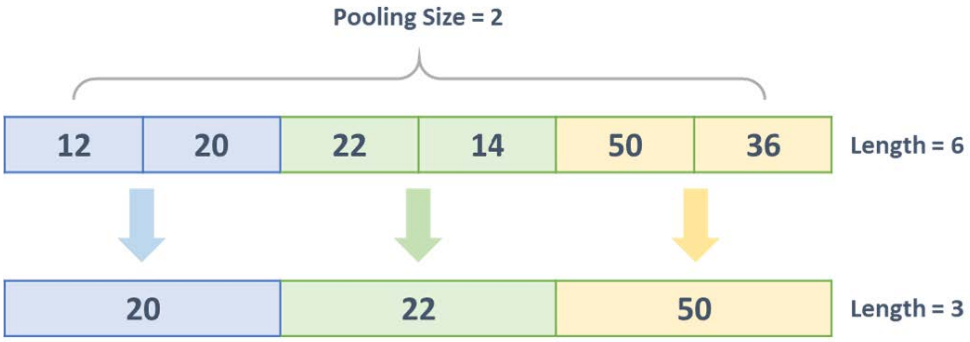


Figure 2.6: 1D max pooling operation [112].

cess starts with convolution and pooling, breaks down the input into a vector of features, then independently analyses them [124]." This process leads to the final decision which is taken by fully connected layer. The output of network's earlier layer is reshaped into a single vector. Weights are applied on the inputs taken from the feature map to determine the correct label. The output layer of FC gives the final probabilities for every label.

Batch Normalization *Batch normalization* (BN) improves neural network efficiency and stability and prevent gradient vanishing and over-fitting [125]. Normalization and standardization techniques transform the data by putting all the data points on the same scale [124]. BN is further utilized in feature maps to fix problems with internal covariance shift [126]. "The change in the internal covariance is a change in the distribution of hidden unit values that slows convergence and requires careful initialization of network parameters [127]." The previous activation output is normalized by subtracting the batch-mean then dividing by the batch standard deviation. BN works as a regularizer and eliminates the need for dropout in some cases. The neural network's stability is also increased, as BN enables us to use much higher learning rates and to take less care of initialization. Furthermore, each network-layer will learn a little more independently from other layers due to BN [128].

$$BN(x_k) = \alpha(\frac{x_k - \mu_B}{\sqrt{\sigma_{B^2} + \epsilon}}) + \beta \tag{2.8}$$

where, ϵ is a random noise for stability, α is the mini-batch variance, μ_B represents the mini-batch mean, σ represents a scale parameter, and β is the shifting parameter for the layer. Both α and β are trainable and updated in an epoch-wise manner [112].

Network Optimization Neural network's performance, accuracy and validity are determined by its hyper-parameters [129]. Optimizers are one of the important hyper-parameters used to adjust the weight and learning rate of the neural network; also, optimizers aim to minimize losses. *Stochastic gradient descent* (SGD) is one of the most influential and popular deep learning neural network optimization algorithms [130]. However, it performs computations across the entire dataset, which have been reported to be redundant and inefficient. Momentum and Nestrov acceleration are few strategies proposed to improve SGD performance [131].

"Adaptive optimizations have become popular recently as they adjust the learning rates of parameters in different scales instead of directly regulating the overall phase sizes, resulting in a smoother training process and faster convergence. The adaptive learning rate is determined by dividing the gradients by a denominator, representing the square root of the global mean for previous gradient squares [132]." Out of all adaptive optimizers, Adam optimizer is the most popular since it converges way faster than other adaptive methods. Adam combines the advantages of SGD extensions, RMSProp and AdaGrad [137–139]. "Other adaptive algorithms, such as AdaDelta [135], RMSProp [136], and AdaBound [140] were mainly proposed to replace the denominator by the square root of the exponential moving average of past gradient squares." Various studies reported that AdaGrad optimizer usually converges faster than vanilla SGD when the gradients are sparse [133]. However, its generalization performance is relatively poor [134]. The effect of different optimizers and other hyper-parameters on 1D-CNN network will be discussed more in detail in Chapter 8.

LSTM-RNN

Long short-term memory (LSTM) is an artificial *recurrent neural network* (RNN) used in the field of deep learning. Before we get details of LSTM, it is essential to understand the basics of RNN. Traditional feed forward CNNs does not work well with sequential data. RNNs have become state-of-art algorithms for sequential data. RNN is the only network with an internal memory unit to remember its input, making it ideally suited for machine learning problems involving sequential data [141]. Figure 2.7 shows the unfolded RNN model. Here, every prediction at time (h_t) is dependent on all previous predictions and the information learned from them. "However, vanilla RNNs fail to understand the context behind an input. Something that was said long before, cannot be recalled when making predictions in the present [142]." This is called the problem of vanishing gradient. This issue can be resolved by applying a slightly tweaked version of RNNs- the Long short-term memory networks [142].

LSTM allows information to persist. LSTM has an internal mechanism called

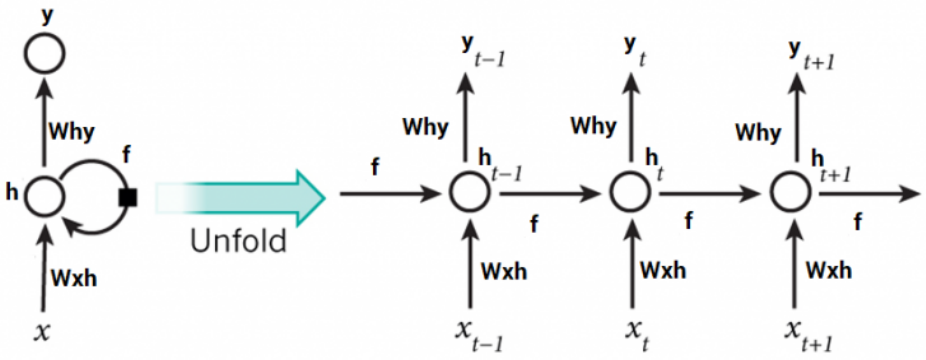


Figure 2.7: Recurrent neural network architecture [142].

gates that helps regulate the flow of information by learning which data in a sequence is necessary to keep or throw away. LSTM's cell state acts as a transport highway, it passes the relevant information down the long chain of sequences to make predictions. The information gets removed or added to the cell state via gates. There are total three gates in an LSTM cell to regulate the flow of information. A forget gate, input gate and output gate [143].

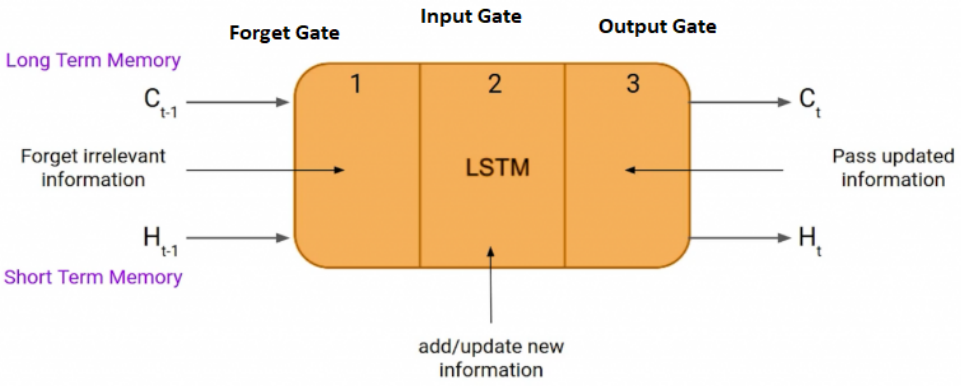


Figure 2.8: LSTM cell architecture [144].

Figure 2.8 shows the LSTM cell architecture in its simplest form. Same as RNN, an LSTM also has a hidden state. Here, H_{t-1} and H_t represent the hidden state of the previous and current timestamp, respectively. Additionally, a cell state of LSTM is sequentially represented by C_{t-1} and C_t for previous and current timestamp. In

LSTM, hidden state and cell stare are known as short term and long-term memory, respectively [144].

Now, we shall understand the roles played by these gates in LSTM architecture. Forget gate decides whether we should keep the information from the previous timestamp or forget it. The second gate, which is input gate tries to learn new information from the input to the current cell. If the information is relevant, it gets added to the cell state at the current timestamp, otherwise the information is subtracted from the cell state. At the end, the cell passes the updated information from the current timestamp to the next timestamp through output gate [144]. The topics of hyper-parameter tuning and network optimization for LSTM network are described in detail in Chapter 8.

Selected Traditional Supervised Regressors for Inspection

In this section, we discuss selected traditional regressors which we used in our studies. Due to the type of the input data being sequential, we have not used linear regressors in our research.

Support Vector Machine

Support vector machines (SVM) are a set of supervised learning methods. It constructs a hyper-plane or set of hyper-planes in a high or infinite dimension space. In general, the hyper-plane that has the largest distance to the nearest training data points of any class achieves a good separation. When this method is used for regression analysis, it is called support vector regression. The model developed by this method depends only on a subset of the training data, because the cost function of the model ignores samples whose prediction is close to their target. This feature of using subset of training points called support vectors makes the method memory-efficient [145].

This method is still effective in cases where the number of samples is lesser than the number of dimensions. However, the regularization term and the chosen kernel function are crucial parameters in this scenario.

K-Nearest Neighbors

Despite its simplicity, this method is popular in cases where the data labels are continuous. Thus, the accuracy achieved by this method is superior to linear regression methods in the case of sequential data. In *K-nearest neighbors* (KNN), the label assigned to a query point is computed based on the mean of the labels of its nearest neighbors. Here, k is the value of nearest neighbors which is specified by the user.

"This regression method utilizes uniform weights, which means each point in the local neighborhood contributes uniformly to the classification of a query point [146]." Alternatively, weights proportional to the inverse of the distance from the query point can also be applied [146].

Decision Trees

"*Decision trees* (DT) are a non-parametric supervised learning method used for regression [147]." The value of a target variable is predicted by learning simple decision rules inferred from the data features. A tree can be seen as a piece-wise constant approximation. To understand or interpret the model, trees can be visualized. Decision trees perform well even if its assumptions are somewhat violated by the true model from which the data were generated. One of the advantages of this method is that it requires little data preparation. However, it does not support missing values. Moreover, there is a chance of over-fitting by creating over-complex trees that do not generalize the data. If the dataset is not balanced prior to fitting with the decision tree, we may get biased results if some classes dominate [147].

Random Forest

Random forests (RF) is a type of ensemble method, which was developed to solve the problem of over-fitting in decision trees. Here, each tree in the ensemble is built from a sample drawn with replacement from the training set. "A random forest is a meta estimator that fits a number of classifying decision trees on various subsamples of the dataset and uses averaging to improve the predictive accuracy [148]." Additionally, at the time of nose splitting during the construction of a tree, best split is found either from all input features or a random subset of size. "The purpose of these two sources of randomness is to decrease the variance of the forest estimator [149]." The high variance problem of DT is also solved in RF by combining diverse trees. However, this slightly increases the bias [149].

Extreme Gradient Boosting

Extreme gradient boosting is mainly known as XGBoost. XGBoost is another decision-tree based ensemble machine learning algorithm. This method uses a gradient boosting framework. This algorithm is popular for its wide range of applications in regression and classification analysis. To improve the run time, XGBoost approaches the process of sequential tree building using parallelized implementation. Also, the computational performance is significantly improved because of the depth-first approach in tree pruning. Moreover, "It penalizes more complex models through both

Lasso (L1) and Ridge (L2) regularization to prevent over-fitting [150]." The need to specify the exact number of boosting iterations required in a single run is overcome because the algorithm performs cross validation algorithm at each iteration [150].

Evaluation Metrics for Quality Inspection

In order to assess the accuracy of the developed deep conventional neural networks or traditional regression models, several statistical measures have been used. In this research, *Pearson correlation* (R^2), *Root mean square error* (RMSE), *Mean absolute error* (MAE), and *mean squared logarithmic error* (MSLE) have been used as evaluation metrics. Each statistical measurement is shown in Equations 2.9, 2.10, 2.11, 2.12 as mathematical representation. The values of Pearson correlation ranges between -1 to +1. The values near to 1 show better predictions. MAE, MSLE and RMSE are ranged between 0 and ∞. The closer the forecast to 0, the better the prediction [151, 152].

$$R^2 = \frac{p(\Sigma_{i=1}^{p}\hat{y}_i y_i) - (\Sigma_{i=1}^{p}\hat{y}_i)(\Sigma_{i=1}^{p}y_i)}{\sqrt{(p\Sigma_{i=1}^{p}y_i^2 - (p\Sigma_{i=1}^{p}y_i)^2)(p\Sigma_{i=1}^{p}\hat{y}_i^2 - (p\Sigma_{i=1}^{p}\hat{y}_i)^2)}} \tag{2.9}$$

$$RMSE = \sqrt{\frac{\Sigma_{i=1}^{p}(\hat{y}_i - y_i)^2}{p}} \tag{2.10}$$

$$MAE = \frac{1}{p}\Sigma_{i=1}^{p}|\hat{y}_i - y_i| \tag{2.11}$$

$$MSLE = \frac{1}{p}\sum_{i=0}^{p}(\log(y_i + 1) - \log(\hat{y}_i + 1))^2 \tag{2.12}$$

where, y_i and $\hat{y}_i$ represent the actual value and predicted value, respectively. p is the number of dataset elements, $1 < i \leq p$.

Machine Learning for Quality Control

We summarized selected papers using traditional machine learning or conventional deep learning for quality inspection of specular surfaces in Table 2.5. Here, F1 score is an evaluation metric for machine learning classifiers. Also, TPR stands for *true positive rate*, which is also an evaluation metric for classification. All selected papers use machine learning or deep learning techniques to detect different types of defect visible on the specular surface. All papers require a huge dataset consisting of images of defects on a shiny surface.

Table 2.5: Artificial intelligence methods for quality inspection of specular surfaces.

Author, Year	Reference	Method	Output
Zhoul, 2020	[155]	CNN for classification	F1=0.7513
Yun, 2020	[95]	CNN with Data augmentation	ACC=99.69%
Zhang, 2019	[156]	CNN for classification	ACC=98.74%
Maestro-Watson, 2019	[153]	CNN with Deflectometry	ACC=88.91%
Xiao, 2019	[157]	ANN, SVM, AdaBoost for classification	MAE=0.383(ANN), MAE=0.311(SVM), MAE=0.333 (AdaBoost)
Zhou, 2019	[158]	SVM, YOLO-V3, Faster R-CNN for classification	F1=0.492(SVM), F1=0.515(YOLO-V3), F1=0.905(Faster R-CNN)
Martinez, 2017	[154]	ANN with Image fusion technique	TPR=90.16%, FPR=1.98%

After reviewing the literature, we can mention that the machine learning techniques have never been used for 3D reconstruction of the product for inspection purposes. Moreover, all the methods specified in these papers cannot measure the size of defects. Additionally, some palpable defects present on reflective surfaces are not visible in the image captured by the camera. Therefore, these image-based or vision-based techniques cannot be used in the production line for defects detection in shiny surface. In our research, we have combined conventional deep learning and traditional machine learning techniques with our stereo-laser inspection system. By combining them, we developed a hybrid method which not only reconstructs the accurate 3D model of the product but also detects defects in the product at the same time.

2.6 Summary

The specularity of the surface increases the difficulty of the inspection process. Traditionally, a person is appointed for the 3D inspection of specular surfaces in production line. This increases the probability of error and makes the process of inspection time consuming. We summarized the results of this literature review as follows:

- Traditional 3D measurement techniques have been discussed. Any traditional 3D measurement methods do not prove to be effective when it comes to shiny surfaces.
- The state-of-art HDR techniques require the product to be in a dark environment to get an accurate result. The ambient lighting conditions of the working environment are hindrance to these methods.
- Any defects detection techniques specified in this section, which either use machine learning approach or traditional approach cannot measure the size of the defect.
- The available commercial solutions can inspect the specular surfaces accurately. However, the accuracy is affected by the ambient light.
- Artificial intelligence techniques have only been used to detect visible defects in specular surfaces. They have never been used for 3D measurement of the specular product for inspection purposes.
- It is found that almost all applications of DL and ML use supervised learning with labeled image dataset.

By critically analyzing the literature, we found out the strengths and weaknesses of the traditional approach. As our method is developed by combining two traditional methods, we expect our stereo-laser system to overcome the limitations of traditional methods.

From comparison with commercial products, we can mention that our inspection system is accurate and cost-effective. Additionally, the ability of our system to work in ambient light makes it a more feasible solution.

With this review, we have tried to cover all possible areas related to the inspection of specular surfaces in ambient light. Whether it is traditional approach, state-of-art methods, commercial solutions or AI techniques, we have thoroughly analyzed each field to determine its pros and cons.

As stated before, most of the DL and ML techniques specified in this section use image datasets. Also, these techniques have not been used for 3D measurement. With our hybrid approach of using AI with our stereo-laser system for 3D measurement, we aim to improve the performance of our inspection system.

We hope that this survey not only provides a better understanding of deep 1D-CNN architectures but also facilitates future research activities and application developments in the field of traditional regression and machine learning for quality control.

Chapter 3

Setup for 3D Measurement using the Stereo-Laser System

This chapter starts with an explanation of the setup of the stereo-laser system. The crucial details of the setup, which affects the system's accuracy, are also described in detail. Additionally, the hardware components required to build this system are mentioned with their essential parameters. Halcon software is used to perform image processing tasks. A brief description of Halcon software of also provided in this chapter. Moreover, the effect of the background color on the accuracy of the 3D measurement is also demonstrated in this chapter. A portion of this chapter has been published in my publication [159].

3.1 Setup

The setup of the stereo-laser system is shown in Fig. 3.1. A laser line of blue light is projected on the product to be inspected, which is placed on the rotating table. The product is rotating at a constant speed. The encoder is embedded with the rotating table to provide the trigger signal through a synchronization cable. When the stereo cameras receive the trigger signal, the stereo images of the product with the projected blue laser line are captured. The stereo cameras, with identical focal lengths, are placed in standard epipolar geometry. In this arrangement, the stereo cameras are mounted on a plane with a displacement that is perpendicular to the cameras' optical axis [191]. The b denotes the distance between the two cameras, which is also known as the baseline distance. With a large baseline distance, we can better depth resolution at high distances. On the other hand, we can examine close objects accurately with a small baseline distance. The stereo images captured by the stereo cameras are displayed in respective monitors. The stereo images are processed to inspect the product. All experiments are performed in ambient light to develop a system, which overcomes the limitation of the traditional 3D measurement methods, and inspects a highly specular surface in ambient light.

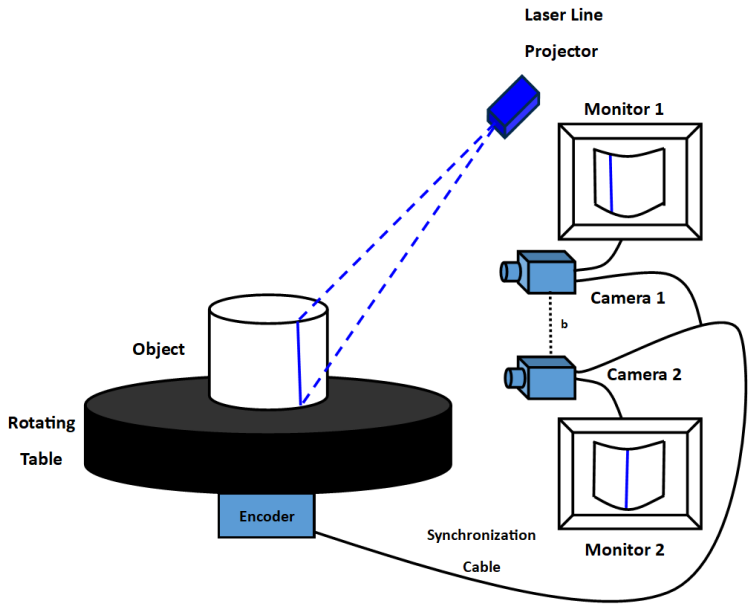


Figure 3.1: Setup of the stereo-laser system.

The important parameters of the hardware components of the system are specified below. Also, a brief introduction of the Halcon software is provided, which is used to perform all image processing tasks.

1. **Stereo cameras.** We have used a pair of Genie Nano M4020 monochrome cameras, which has 4112 $\times$3008 resolution. The speed of these stereo cameras is 9.7 fps. Both cameras have identical focal lengths. As stated above, the stereo cameras are placed in canonical stereo geometry, which means their optic axes are parallel. The LM16XC Kowa lens with a focal length of 16 mm is used with these cameras [160].

2. **Laser line projector.** The Z-laser model Z20M18H3-F-450-lp45 is used during this research. It is a blue light laser of class 2M with a 45deg fan angle, which projects a sharp narrow band of the blue light laser line. The width of the projected narrow laser line is 0.3 mm. In addition, the output power and the wavelength of this laser are 20 *milliWatt* (mW), and 450 *nanometres* (nm), respectively [161]. Due to its low output power, we can use this laser in an open environment to perform experiments. The position of the laser is the most critical parameter of the setup, which is discussed later in this section.

3. **Rotating Table.** The diameter of the rotating table is 70 cm, which is chosen according to the size of the washing machine drum, and the wash plate. Due to the low fps of stereo cameras, the rotating table is rotated at a relatively slow speed of 0.02 RPM. This speed is chosen to make sure that the stereo cameras do not drop any frames. One rotation should not take more than 4-5 seconds when using the stereo-laser system in real-world applications. Moreover, the rotating table is painted black to avoid the effect of the background on the accuracy of the output. This is explained in detail in the next section.

4. **Encoder.** The encoder is attached with the rotating table, which is used to provide the trigger signal to stereo cameras through a synchronization cable. We have used an incremental encoder Sick DBS36E, which generates 2500 pulses per revolution. Therefore, the stereo cameras capture 2500 images of the product in one rotation [162].

5. **Halcon software.** Halcon is the comprehensive standard software developed by MVTec, to perform image processing tasks for machine vision applications. "The software provides the latest state-of-the-art machine vision technologies, such as comprehensive 3D vision and deep learning algorithms [163]." The software is compatible with different operating systems, such as Windows, Linux, and macOS. It can also access entire libraries from common programming languages like C, C++, and .NET [164]. Moreover, Halcon also provides an extensive library of more than 1100 operators for blob analysis, morphology, pattern matching, 3D calibration.

Furthermore, the perimeter of the wash plate is 2103.8 mm, and the encoder produces 2500 pulses per revolution. Therefore, the distance between every two consecutive captured images is $2103.8/2500 = 0.84$ mm for this setup.

Position of the Laser

One of the most critical parameters of this technique is the position of the laser line. We projected a narrow band of a horizontal laser line onto a flat surface to understand the concept. As shown in Fig. 3.2, the laser line diffuses and appears thick and bright when it is projected directly in the centre of the image. The sharp bell-shaped curve in Fig. 3.3 shows the ideal intensity distribution of the laser. Here, the x-axis depicts the width of the *region of interest* (ROI) image. The gray value intensity of the laser is plotted on the y-axis. Compared to the ideal intensity distribution curve, the bell-shaped curve becomes wider and round, when the laser is projected directly in the centre of the image (See Fig. 3.4). As the distance from the centre increases, the intensity of the projected laser line decreases gradually. If the laser line is projected

too far from the centre of the image, then the laser line merges with the background surface due to its low intensity. The intensity distribution of the laser for this case is shown in Fig. 3.5. "Hence, we need to choose the laser position in such a way that it does not cause any reflection and does not get merged with the background in both images [159]."

Figure 3.2: Change in the intensity of laser [159].

3.2 Calibration of Stereo Cameras

After setting up the system, the first step of the experiment is to calibrate the stereo cameras, which we have carried out using the Halcon software. First, a calibration data model is developed by specifying the internal camera parameters, and the calibration object. The figure shows the image of the calibration plate used during this research. Halcon supports the calibration plate with hexagonally arranged marks, containing five finder patterns, i.e., unique mark hexagons where either four or six marks contain a hole (See Fig. 3.6). To find the calibration plate in the image and estimate its pose relative to the observing camera using Halcon software, at least one finder pattern must be completely visible in the image [165].

Each camera acquires multiple images of the calibration plate in different poses. Additionally, the calibration plate is placed directly on the measurement plane to obtain extrinsic parameters. In these images, the markers of the calibration object are extracted, and their (pixel) coordinates, together with the indices of the camera, calibration object, and calibration object pose and with a tuple containing the indices

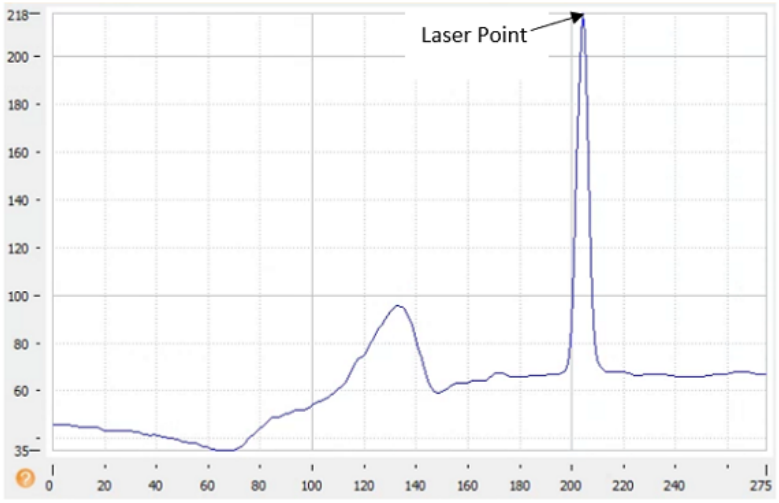


Figure 3.3: Ideal intensity distribution of the laser line.

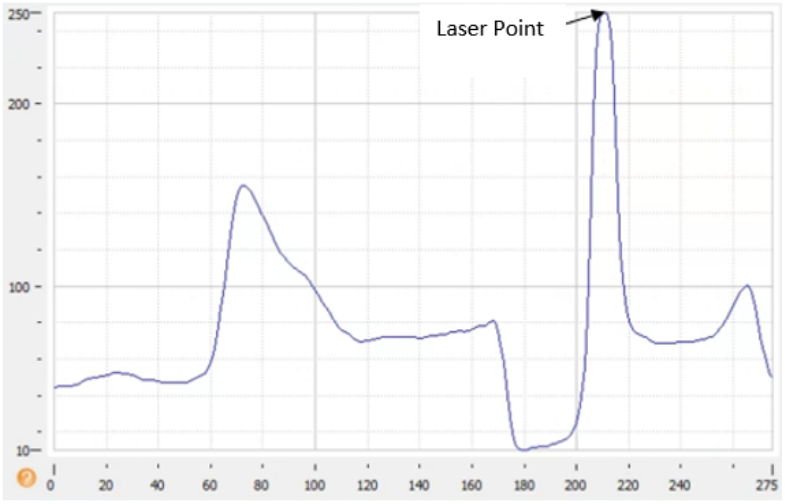


Figure 3.4: Intensity distribution for the case, when the laser line is projected in the centre of the image.

of the corresponding markers, are stored in the calibration data model. "As a result of the calibration process, only the calibration error is returned, which corresponds to the average distance (in pixels) between the back-projected calibration points and

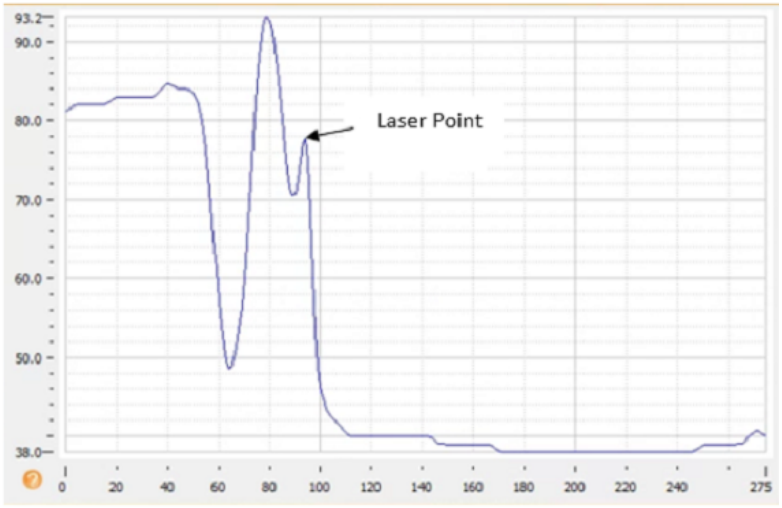


Figure 3.5: Intensity distribution for the case, when the laser line is projected too far from the centre of the image.

their extracted image coordinates. An error of up to 0.5 pixels indicates that the calibration of stereo cameras was successful [165]." The values of calibration error were 0.4 pixels and 0.3 pixels in the case of wash plate and washing machine drum, respectively. The calibration data is used to rectify stereo images before processing them further. Moreover, the world coordinates of a point are also obtained from its image coordinates using the calibration data.

3.3 Effect of Background color

Generally, the wash plate and the washing machine drum contain drainage holes. It is essential to accurately measure the dimension of each hole. However, the accuracy is affected by the color of the background. An experiment is carried out to understand this effect. A stainless steel sheet with a grid pattern of holes is created. We have captured two images of the metal sheet, one with a white background, and another with a black background, in ambient light (See Fig. 3.7).

When the color of the background is white, it is difficult to extract the edges of the holes accurately. Additionally, the holes merge with the reflected region because of the white background. A shadow appears for each hole in ambient light, which is visible because of the white background. This adds to the difficulties of extracting the edges of holes. On the other hand, we can easily extract the edges of holes, even

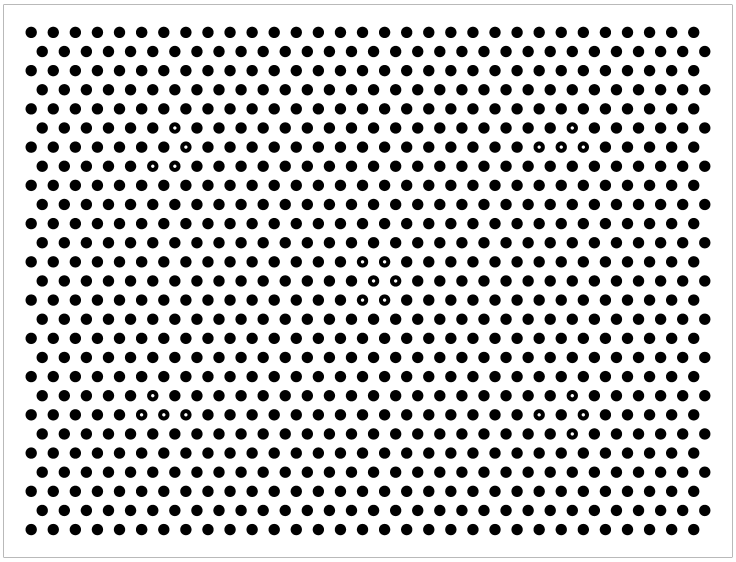
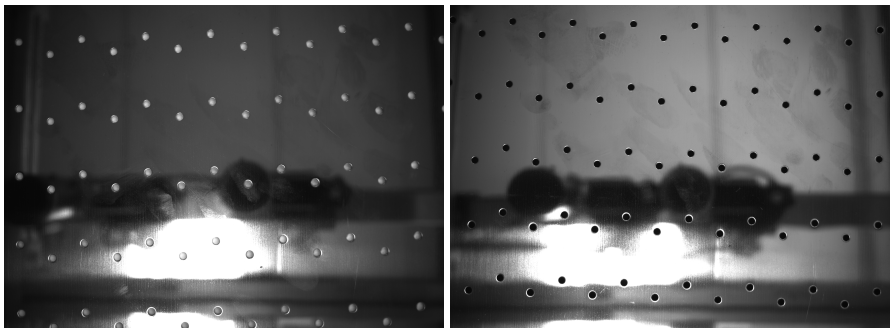


Figure 3.6: Calibration plate.



(a) Image with a white background. (b) Image with a black background.

Figure 3.7: Effect of the background color.

in the reflected region, with the black background. Moreover, we can clearly see the sharp edges of holes even in ambient light. Therefore, we have painted the rotating table with black color, to reduce the difficulties while inspecting the holes present in the wash plate. The black rotating table provides the effect of black ground, which

makes the inspection process more manageable. Furthermore, the washing machine drum is covered from the top during the experiment to avoid the same problem.

3.4 Summary

We have found out that the position of the laser and the background color are the two most crucial parameters of the setup. These parameters play a vital role when it comes to the accuracy of 3D measurement. Overall, a detailed explanation of the hardware components of the stereo-laser system is provided in this chapter. The calibration process for stereo cameras is also mentioned. Moreover, some critical parameters, such as the total number of images, the speed of the rotating table, the arrangement of stereo cameras, and the distance between two consecutive images, are provided for future reference. It is essential to keep these parameters in mind, while setting up the system for 3D measurement.

Chapter 4

Stereo-Laser System

Stereo vision and laser triangulation are among many traditional 3D measurement methods used in Computer Vision. In this chapter, we start with a detailed explanation of these two techniques. Next, we discuss the limitations imposed by the specularity of the surface on these techniques. After understanding these concepts, we discuss our proposed algorithm, which combines these two concepts to inspect the highly specular surface in ambient light. Some special cases of the intensity distribution of a laser, which occur due to the ambient light of the working environment, are described with their possible solution using the stereo-laser system. Moreover, the experimental results show the reconstructed washing machine drum and the wash plate using our stereo-laser system. An accuracy test is performed to evaluate the accuracy of the system, which is also explained in this chapter. The material of this chapter has been published in my publication [9,10].

4.1 Stereo Vision and Laser Triangulation

Stereo vision and laser triangulation are a type of surface non-contact techniques. For high-speed inspection of products in a production line, surface non-contact techniques are preferred. Here, laser triangulation is an active technique, in which 3D imaging sensors operate by projecting an electromagnetic energy signal onto an object followed by recording the reflected energy. In contrast, stereo vision is an example of the passive technique. The 3D sensors operate by acquiring electromagnetic energy from an object in this form [22,23].

Stereo Vision

According to Howard and Rogers, "Stereopsis is a term that is most commonly used to refer to the perception of depth and 3-dimensional structure obtained on the basis of the visual information deriving from two eyes by individuals with normally developed binocular vision [166]." As the eyes of humans and many animals are located at different lateral positions on the head, two slightly different images are

obtained as a result of binocular vision. These images are projected to the retinas of the eyes. The difference in the relative horizontal position of objects in the two images is referred to as binocular disparities. The visual cortex of the brain processes disparities to yield depth perception. Stereoscopic vision probably evolved as a means of survival [166].

Nowadays, Stereo Vision is a well-known ranging method because it resembles the basic mechanism of human eyes. The computer stereo vision system uses the same principle by replacing eyes with two CCD cameras. They are displaced horizontally to obtain two different views. A disparity map, which encodes the difference in horizontal coordinates of corresponding image points, can be obtained by comparing these two slightly different images. The values in the disparity map are inversely proportional to the scene depth at the corresponding pixel location [31].

As stated above, two CCD cameras replace human eyes in the computer stereo vision system; thus, they should be as identical as possible for avoiding difficulties. In canonical stereo geometry, both the cameras have identical effective focal lengths and they should be arranged in such a way that they have parallel optic axes [191].

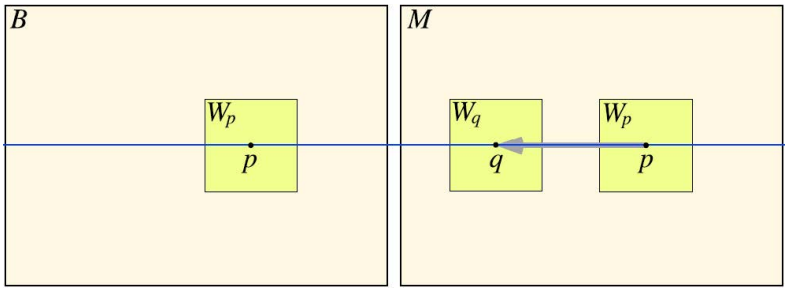


Figure 4.1: Left and right camera image for canonical stereo geometry [191].

In an ideal case, both the images have collinear row y (See Fig. 4.1). However, in practice, stereo cameras need to be calibrated, and stereo images need to be rectified before calculating the disparity. Here, the left image is the base image indicated by B, and the right image is the matched image indicated by M. The projection of point P in the base image is displayed by pixel $p(x, y)$. Now, we have to search for a corresponding pixel on the same epipolar line in the matched image M. "The two pixels are corresponding if they are projections of the same point P in the scene." The pixel $q(x + d, y)$ is the corresponding pixel of pixel p, where d is the value of the disparity, which is plotted in a disparity map. So, a disparity map encodes the difference in horizontal coordinates of corresponding image points. For any pixel P, if the value of disparity is 0, then the pre-image P is at infinity. The larger the disparity value, the closer is the pre-image to the cameras. We can generate the 3D

world coordinates for each pixel by processing the disparity map. "An increase in the image resolution is a general way to improve the accuracy of depth levels [191]."

Overall, the stereo vision technique has very accurate abilities among all the surface non-contact techniques [22]. However, a highly specular surface in ambient lighting conditions causes highlights in some parts of an object. Sometimes, the highlights caused by ambient light would be visible from the view of only one of the cameras. Therefore, some parts of the object will appear brighter in one image and darker in the other image. Moreover, specular surfaces do not have any texture. All this will increase the difficulty of finding corresponding points in stereo images. This will lead to an incorrect disparity value for the reflected regions of the surface [10].

To understand the limitation of the stereo vision technique, we have carried out an experiment using the SP1 stereo vision system, which performs stereo matching in real-time based on *semi-global matching* (SGM) algorithm. In this experiment, the stereo vision system is used to reconstruct the washing machine drum in ambient light. A detailed explanation of this experiment and its results is provided in Appendix A.

Laser Triangulation

In the laser triangulation method, a light pattern, such as a ray, a plane of light, a grid pattern, is projected under calibrated geometric conditions onto an object whose shape needs to be recovered. We have used one light source and one camera. Calibration is performed to determine the pose of the light source with respect to the camera [191].

Traditionally, there are two types of laser triangulation systems. The first is the point laser triangulation system. In this system, the laser aims a single beam at the desired location, and the system calculates the coordinates of the spot. The same process is repeated after slightly moving the point till the laser has been moved over the entire scene. From the calculated coordinates, a 3D picture of the scene can be formed. However, the data acquisition speed of point laser triangulation systems is very slow. Therefore, a line laser triangulation system is the other option to reduce the scene reconstruction time. The line laser triangulation system can also be used to find the general shape of the object and its distance from the system [167].

A narrow band of light is projected onto a 3D surface of the object in laser-based triangulation systems. This projection produces a line of illumination that appears distorted from an observation perspective other than that of the projector. A profile of the object is represented by the laser line. Figure 4.2 shows an example of laser-based triangulation geometry. The object is moved relative to the measure-

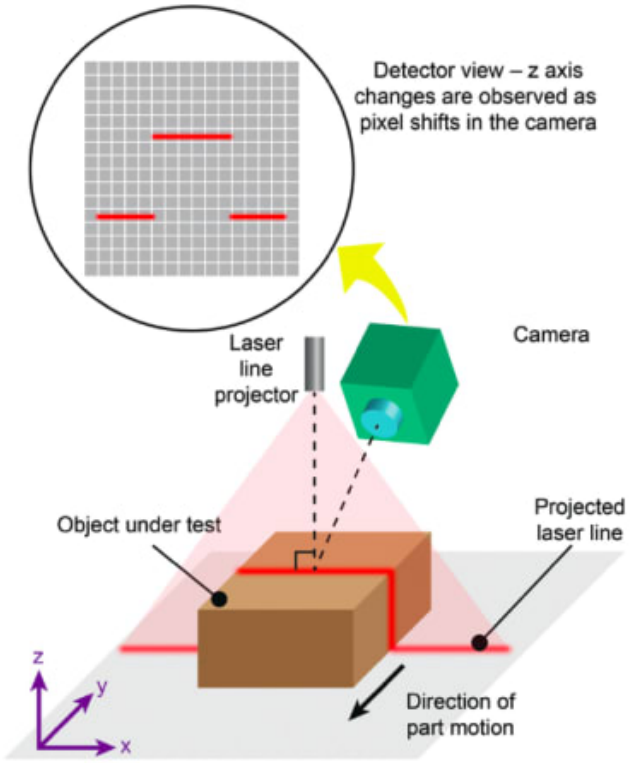


Figure 4.2: Laser-based triangulation system [168].

ment system to reconstruct the whole surface of the object [169].

A laser-based triangulation system consists of four main components: the camera, the line projector, a mechanism that moves the object, and a software system to process the captured image and accurately translate pixel offsets to height differences. For the standard geometry shown in Fig. 4.2, the height variations in the object do not produce a change in the line's y coordinate values (horizontal translation). This simplifies the calibration process and makes the system faster and more accurate. Because the camera views the object from an angle, it increases the depth of field it must accommodate to maintain focus as object height varies. Also, some parts of the line will be blocked or occluded from the camera's view in this geometry [169].

The problem of occlusion can be solved in reverse geometry where the positions of the laser and the camera are exchanged. "This setup offers increased height reso-

lution over the standard configuration, since the oblique angle of laser illumination means a given change in object height produces a more significant shift in the position of the laser line [169]." However, this geometry makes the calculation process complex as a change in height causes a change in the line's y coordinate [169].

One key parameter, which can cause measurement errors, is the power uniformity of the line beam. Due to manufacturing variations and tolerance effects, it can fluctuate which causes intensity variations at the beam edges, non-flat top intensity profile and scattering of light. The performance of the system in the absence of manufacturing errors depends on the variations in line width over its length. Width variations occur when the distance from the laser to the object under test changes [169].

The last parameter is the straightness of the laser line. In real world, the projected laser lines are rarely straight. It usually is in a bowed or slightly "S" shape. "A bowed shape is most often caused when the incident laser beam enters the beam shaping optics at a non-normal angle of incidence [169]." The straightness of the line depends on optical alignment. All these factors affect the calibration process. This system offers greater speed and accuracy than most other 3D vision technologies. However, it is a very complex system in which the results critically depend on several system parameters [169].

Moreover, the specular surface of the object diffuses the projected light pattern. Also, the defects present in the region of reflection will not be visible. In addition, it is difficult to inspect small features present in the highlighted region. Thus, we cannot reconstruct an accurate 3D model of the product with the distorted light pattern [10].

4.2 Proposed Algorithm

To overcome the above-specified difficulties, we propose a novel method that combines the concept of structured lighting and stereo vision. A narrow beam of laser light is projected onto a reflective surface. In the calibrated environment, stereo cameras capture images of the deformed laser line, which represents a surface profile of an object. We have developed an algorithm to accurately detect the laser line even in the presence of highlights. Some challenging scenarios caused by the presence of ambient light are also explained in detail. After detecting the distorted laser profiles in both images, stereo matching is performed on detected laser profiles for accurate representation in the *world coordinate system* (WCS). A solution for defects detection is also provided. This method cannot only detect defects, but it can also measure the dimension of defects accurately. The flow chart shown in Fig. 4.3 represents our approach for the 3D measurement with defects detection [10].

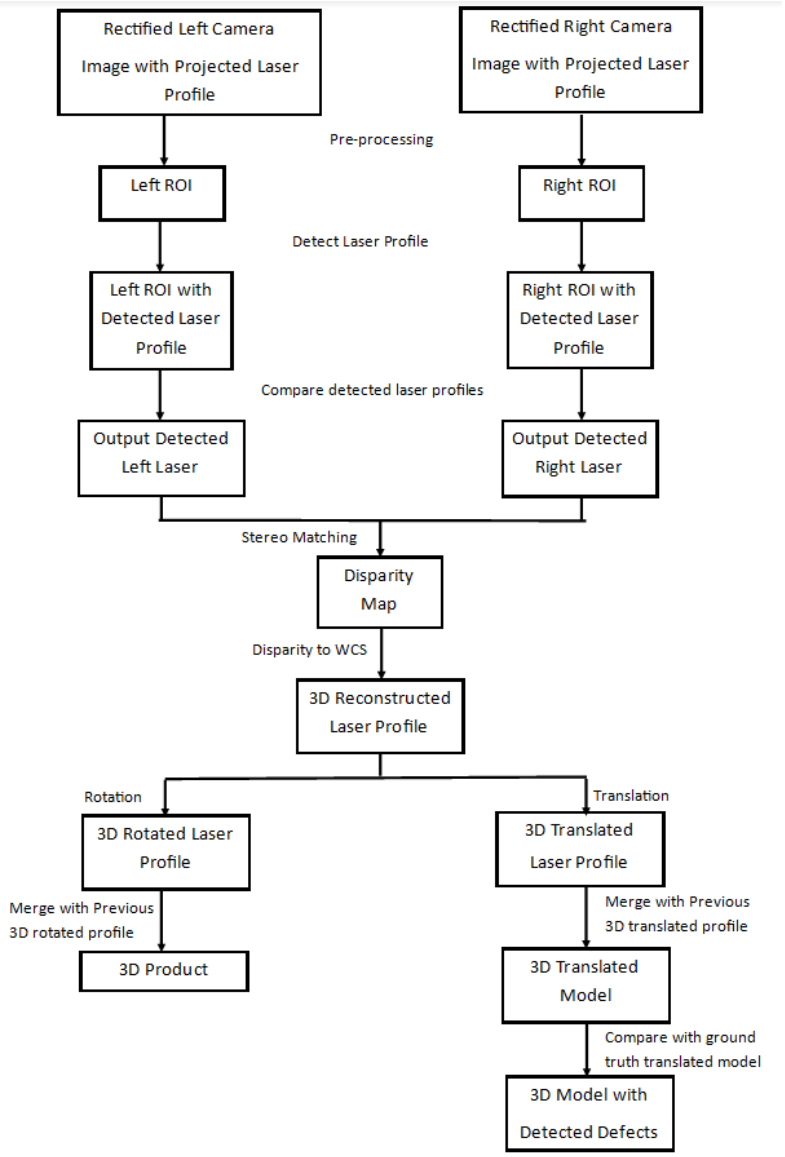


Figure 4.3: Overall approach for 3D reconstruction.

The stereo images of the product with a projected narrow band of laser line are captured in the calibrated environment. Here, the deformed laser line resembles the shape of the product.

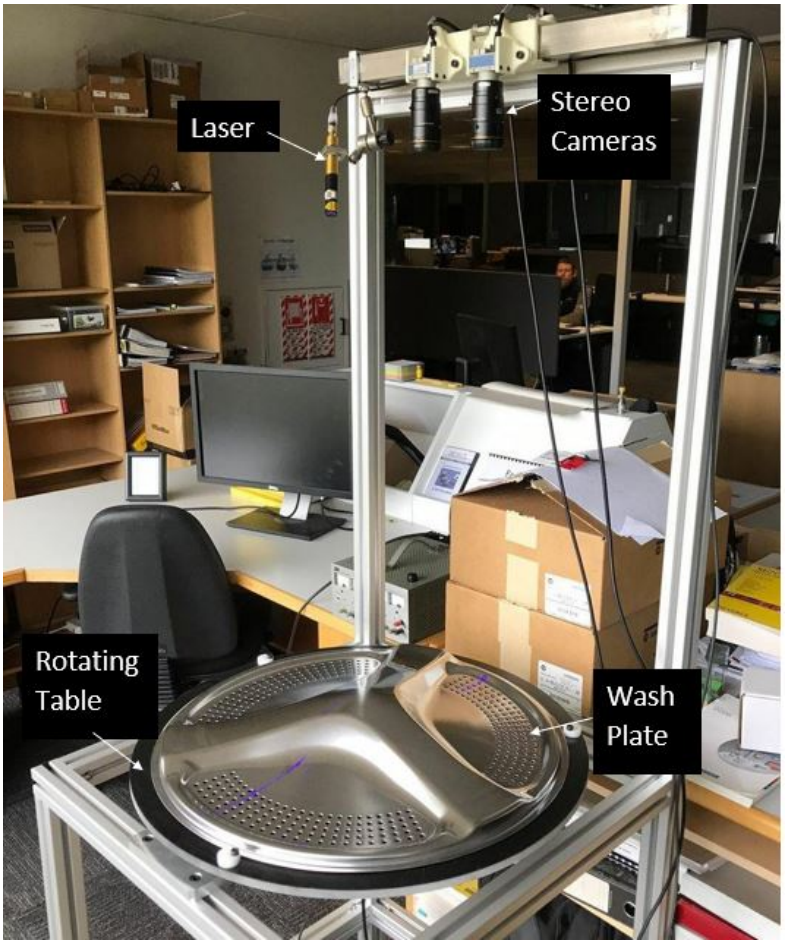
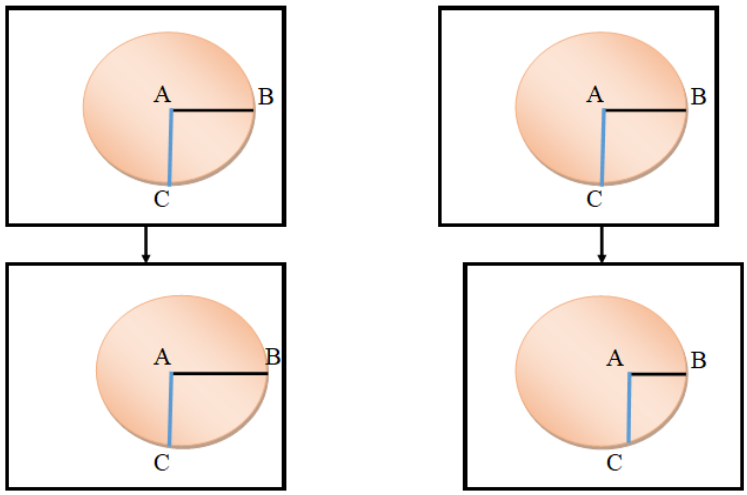


Figure 4.4: Setup of the stereo-laser system.

Pre-processing of Stereo Images

The first step after capturing the stereo images is to rectify them. In the case of repetitive patterns in the image, the matching process may be confused because many points look alike. "In order to make the matching process fast and reliable, the stereo images are rectified such that pairs of conjugate points always have identical row coordinates in the rectified images, i.e., that the search space in the second rectified image is reduced to a line." The stereo images are rectified using the internal camera parameters and the relative pose. The results of the rectification process depend on the calibration results. Therefore, it is crucial to calibrate the stereo cam-



Case 1: Product is shifted.

Case 2: Cameras and laser are shifted.

Figure 4.5: Reference image for the vibrations nullification algorithm.

eras with maximum accuracy.

Region of Interest for Laser Detection

After rectification, these images are pre-processed to nullify the effect of vibrations caused by machinery. The setup of the stereo-laser system is one unit (See Fig. 4.4). The vibrations caused by machinery generally vibrate the whole unit. The shift in the product due to vibrations is normally the same as the shift in the stereo cameras and the laser. Therefore, the effect of vibrations gets nullified in general. However, if the setup is not a single unit, there are two possible cases, which can affect the accuracy of the output. In the first case, only the product is shifted, and the stereo cameras and a laser are in an ideal position. On the other hand, the stereo cameras and a laser are shifted, but the product is steady. We have found a solution to nullify the effect of vibrations caused by machinery for both cases.

Vibrations Nullification Algorithm

Figure 4.5 is the reference image to understand the vibrations nullification algorithm. The extreme point C on the outer edge of the product is taken as a reference. This point C is obtained by detecting the projected laser line on the outer edge of

the product. AC shows the projected laser line on the product. Now, we try to detect the laser point A using the known distance AC. The detected laser point is taken as a reference point for obtaining the point B on the same row. Now, the distance AB is calculated when the product is not rotating, which means there are no vibrations. When the 3D measurement process starts, we calculate the distance AB for every stereo image pair. The calculated distance AB is compared with its ideal value. Next, the difference in distance AB in image coordinates is converted to 3D world coordinates. The average difference in distance AB is +/- 2 pixels. This could cause the error of +/- 0.3 mm in world coordinates. The computed value of difference in AB is later added to the 3D detected laser line AC. In this way, we nullify the effect of the vibrations caused by machinery for both cases mentioned above.

After computing the difference in distance AB, the stereo image pairs are cropped to create the *region of interest* (ROI) images. The height and the width of the ROI images are decided according to the shape of the product. Moreover, it is important to consider the shift in the height of the product, while determining the dimensions of the ROI images.

Laser Detection

The most crucial step of this process is to accurately detect the projected laser line in ROIs in ambient lighting conditions. Our algorithm to detect a laser line in the presence of ambient light is explained in detail. The red curve in Fig. 4.6 depicts the intensity distribution of the projected laser line, which resembles a bell-shaped curve. Here, the x-axis represents the width of the ROI image. The gray value intensity of the laser is plotted on the y-axis. The first step of the detection algorithm is to smooth the curve using a Gauss function [170]. The blue curve shows the smoothed function. Next, we find local maximums for the blue curve [171]. The pixel location of each local maximum is compared with the location of the highest gray value pixel of the red curve. "We try to find the area which is nearest to the location of the highest gray value pixel. The most immediate local maximum would be considered as the detected point of the laser line [9]." In the case of multiple values, the algorithm considers the gray value intensity at each local maximum point. The decision is made by selecting the local maximum point with the highest gray value intensity.

Special Cases

If the ROI contains highlights caused by ambient light or the product has a hole, the intensity distribution would be affected. Hence, we cannot repeat the same process for each row. We found six exceptional cases of intensity distribution, while reconstructing the wash plate and the drum.

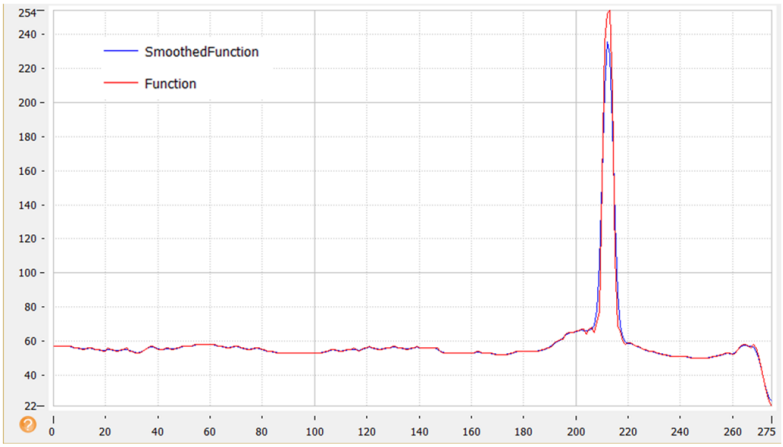


Figure 4.6: Intensity distribution for region without highlights.

Case 1

In the first case, the highlight is present in the ROI. The highlight is caused due to the reflection of the ambient light, when projected onto the specular surface of the product. However, the highlight is separate from the projected laser line in this case. Figure 4.7 shows the intensity distribution for the first case. Here, the first bell-shaped curve represents the highlight, and the second curve represents the projected laser line. "Now, if we follow the same method, then it will assume the highest intensity of highlight as a projected laser profile [10]." Our algorithm takes the location of the detected laser line in the previous row as a reference for the next row to solve this problem. "If the location of the detected laser profile is (x, y), then in the next row, we search pixel locations $(x + 1, y - 10)$ to $(x + 1, y + 10)$ for finding the pixel with the highest gray value [9]." The laser line is detected using the same method as the first row; but, in a smaller region.

Case 2

The intensity distribution for the second case is shown in Fig. 4.8. In this case, the highlight is merged with the projected laser line. It is not possible to extract the location of the laser line pixel when it is merged with the highlight. The algorithm will assume the output to be the same as the output of the previous row. If the highlight is visible in only one of the images of the stereo image pair, we can solve the problem of this case in the next step of the reconstruction process. Typically, the highlights caused by ambient lighting conditions will not be visible in both photos.

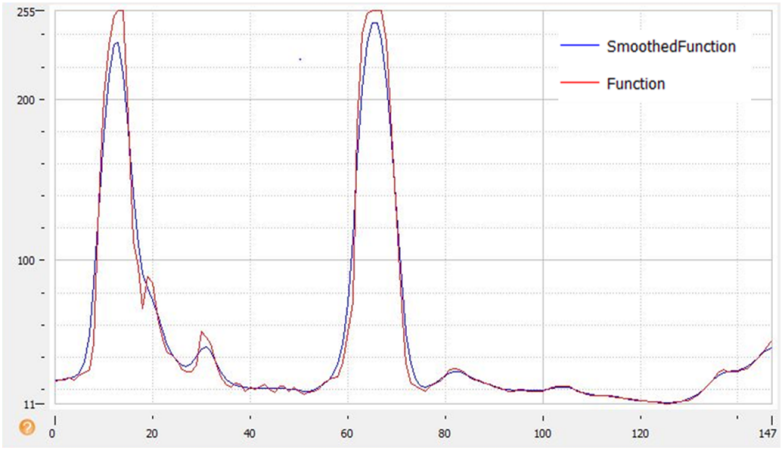


Figure 4.7: Intensity distribution for case 1.

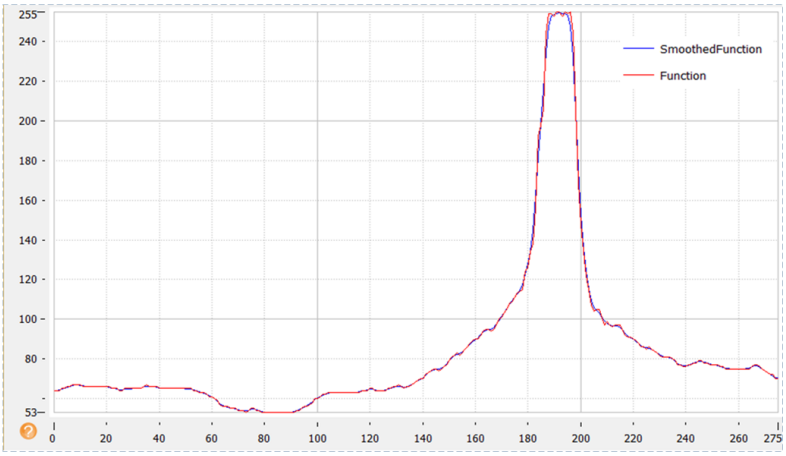


Figure 4.8: Intensity distribution for case 2.

Therefore, the highlights which are present in the left image will not be present in the right image and vice versa. For the case where the laser line is merged with the highlight in one shot, we would still be able to detect the laser line accurately in the other image. By comparing detected laser lines of both images, we can specify the region which has inaccurate output because of the reflection. Then, the inaccurate output values are replaced by the approximate output values. These approximate output values are generated from one of the images' accurate output values [10].

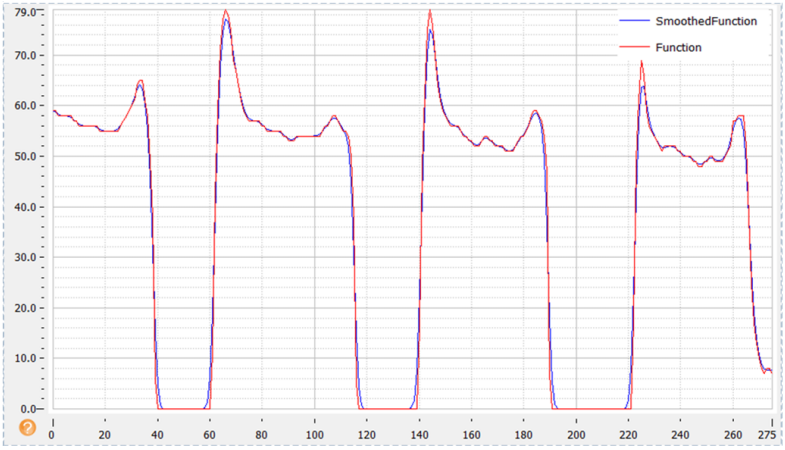


Figure 4.9: Intensity distribution for case 3.

Case 3

The third case occurs when the product has a hole. The intensity distribution in the presence of a hole is displayed in Fig. 4.9. When the algorithm searches for the peak between the pixel locations $(x + 1, y - 10)$ to $(x + 1, y + 10)$, it will consider any local maximum (small peak) in this region as a part of the laser line. To avoid this problem, we specify the minimum gray value to be considered a laser line. If the gray value of the detected peak is higher than the minimum value, then the peak is detected as a laser point. To make it easier and even more accurate, we detect the holes during the pre-processing stage using edge detection [172]. After that, we set the gray value of the detected region to be 0. This simplifies the detection process.

Case 4

When the product's shape changes drastically, the laser line may get shifted outside the search range of $y - 10$ to $y + 10$. The pixel location of the detected laser for the previous row is 160. There is no peak in the search range. However, there is a bell-shaped curve outside the search range (See Fig. 4.10). There is a possibility of getting the same intensity distribution in the presence of a hole and a highlight. In the case of a hole, the gray values in the search range will be 0. In contrast, we get some small gray values in the search range for this case. The gray values will be smaller than the minimum threshold value, so it will not be detected as a laser line. In this case, we expand the search range and repeat the process to detect the laser line.

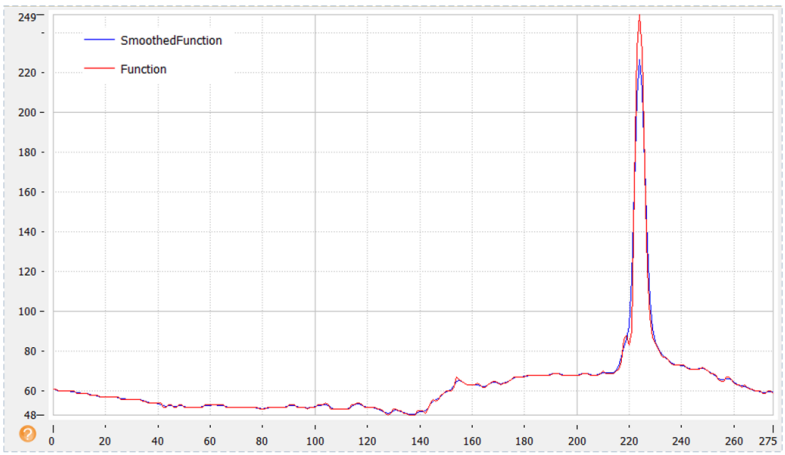


Figure 4.10: Intensity distribution for case 4.

Case 5

The following case usually appears after the second case. In this case, we get two peaks in the search region (See Fig. 4.11). One is a sharp peak, representing the laser line, and the other is the broad peak representing the highlight. The laser line is slightly merged with the highlight in this case. Here, both peaks appear in the search range. For accurate detection, the algorithm first calculates the distance between the pixel value of the detected laser point of the previous row and the pixel value of each peak. Even though the highlight has a higher gray value compared to the laser line, the sharp peak of the laser will be closer to the detected laser point of previous row. By analyzing the distances, the algorithm will select the sharp peak as the correct output.

Case 6

Due to the object's shape, sometimes we acquire the intensity distribution as shown in Fig. 4.12. We obtain two peaks of nearly the same gray values in the search range even after smoothing the curve. This case generally occurs due to the diffusion of the laser line. There is no highlight present in this case. First, the algorithm will calculate the distance between two peaks. If the distance is lesser than the threshold value, the average of both peaks' pixel locations is considered the point of a detected laser line. On the other hand, if the distance is higher than the threshold value, the peak nearest to the previous pixel location is selected.

To compare the detected laser lines, the length of both the detected laser line

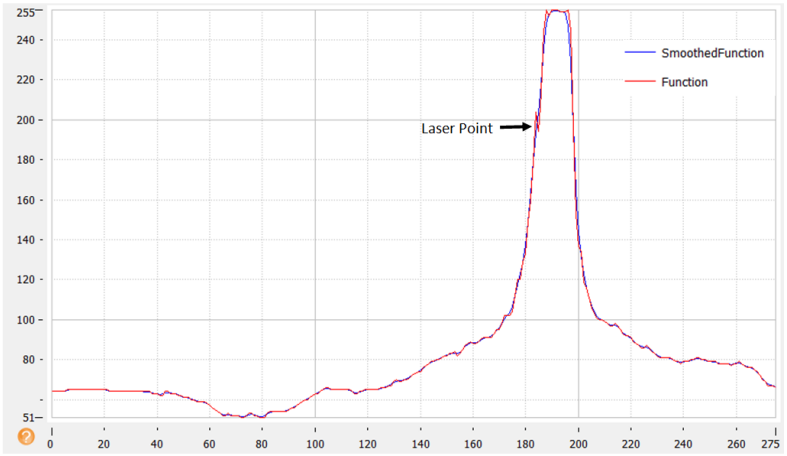


Figure 4.11: Intensity distribution for case 5.

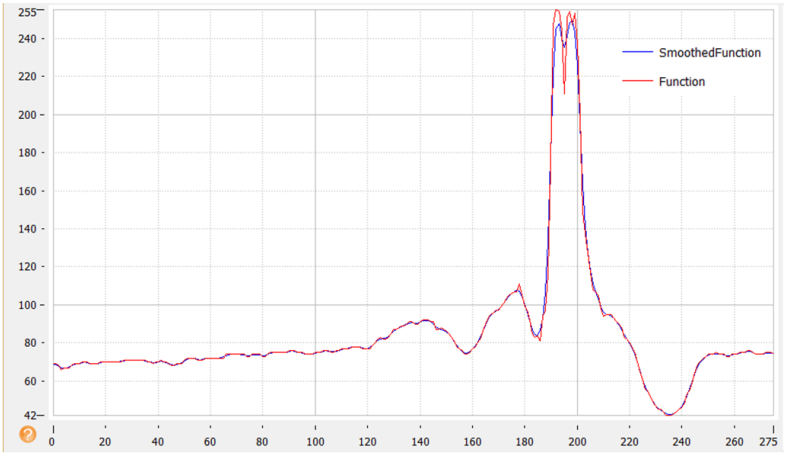


Figure 4.12: Intensity distribution for case 6.

is divided into small segments of equal lengths. The slope and the maximum difference in pixel values are calculated for each segment. We can find the inaccurately detected region by comparing the above-calculated values of every segment of one image with the corresponding values of every segment in the other image. Now, the inaccurately detected segments are replaced by the approximate segments. These approximate segments connect their previous segment with their next segment. Also, every approximate segment has the same slope value as its correspond-

ing segment of the other image. However, if the highlight is visible in both images of the stereo pair, it is impossible to detect the laser line in that region accurately. Therefore, we cannot detect defects or reconstruct that region accurately. This is the limitation of working in ambient light.

Stereo Matching and 3D Reconstruction

After detecting the laser profile in both images, the next step is to perform stereo matching only on the detected laser line. The disparity values, which define the column difference of the image coordinates of two corresponding features on an epipolar line, are calculated for the detected laser point for each row of the ROIs. The image coordinates of the detected laser profile in the left image, along with its disparity values, are used to compute the corresponding points in the world coordinate system [173].

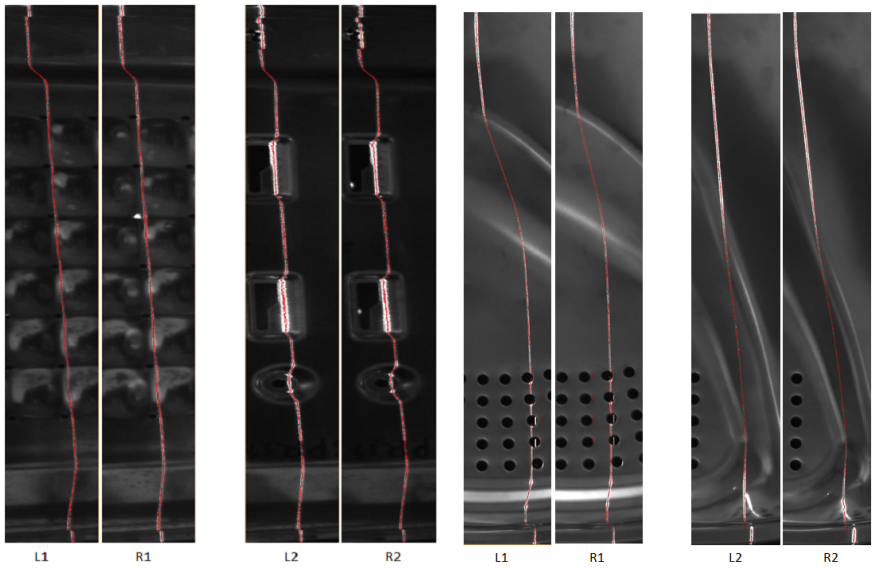
The product is rotated slightly in a clockwise direction after the reconstruction of one 3D profile. After rotation, the process of reconstructing the 3D profile is repeated. Later, the reconstructed 3D profile is also rotated slightly around the Z-axis and merged with the previously reconstructed laser profile [174, 175]. The rotation angle between every two consecutive scans is constant throughout the process. Also, the rotation angle between the location of the first scan and the current scan is identical to the angle between the first reconstructed 3D profile and the current reconstructed 3D profile. This process is repeated till the whole product is reconstructed, that is, till the product has been rotated for 360 degrees. The reconstructed product is compared with the ground truth model for accuracy evaluation. The defects detection process is also a part of the stereo-laser system, which will be explained in the next chapter.

4.3 Experiments and Results

In the experiments, we have used the stereo-laser system to reconstruct a washing machine drum and its wash plate (See Fig. 4.13). Both products are made of stainless steel, which makes its surface look reflective. The wash plate has some palpable defects, such as dents, bumps and cracks, which are also accurately detected using our method. Both products have small critical features/details which are needed to be accurately measured for customer satisfaction. By inspecting these two products, we have solved all possible critical scenarios or problems that may occur due to the product's shape. Moreover, the stereo image pairs are acquired in different weather conditions to understand the problems that may occur due to the ambient lighting condition of the working environment.

(a) A washing machine drum. (b) A wash plate.

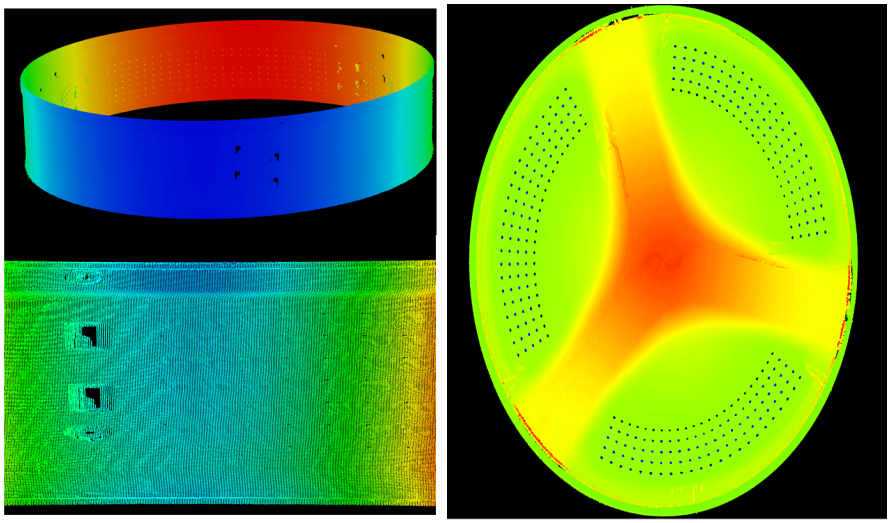
Figure 4.13: Products with highly specular surfaces.



(a) Detected laser profiles in a drum. (b) Detected laser profiles in a wash plate.

Figure 4.14: Detected laser profiles.

We have already discussed the setup of the stereo-laser system in Chapter 3. The baseline distance is 130 mm for this experiment. After calibrating the stereo cam-



(a) 3D reconstructed drum. (b) 3D reconstructed wash plate.

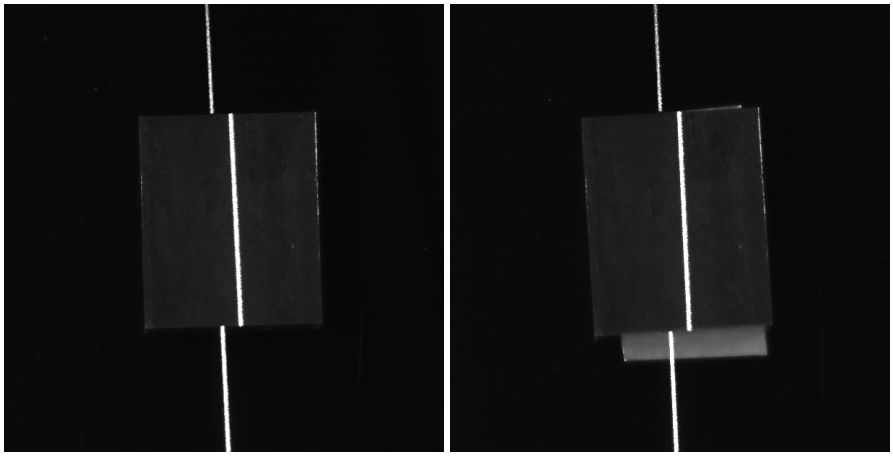
Figure 4.15: 3D model of products in WCS.

eras with utmost precision, the stereo images are rectified. As stated before, a total of 2500 stereo images are captured for the 3D reconstruction of the product. The projected laser line is detected in the ROI images using our proposed algorithm. Figure 4.14 shows two cases of the stereo ROIs with the detected laser lines for both products. Later, we perform stereo matching only on detected laser lines to reconstruct the 3D laser profile in WCS. The generated laser profile is rotated and merged with the previously reconstructed lines. This method is repeated for 2500 scans to reconstruct the accurate 3D model of the product. The rotation angle between two consecutive scans is 0.002513 radian. The output 3D model for both products is shown in Fig. 4.15.

For this experiment, the accuracy of our method is 0.02 mm. In the case of the drum, there are six scans for which the laser is not detected accurately. On the other hand, ten scans are inaccurate in the case of the wash plate. Moreover, in the case of inaccurate detection, the maximum deviation in pixel locations is 3 pixels for the drum, and for the wash plate, it is 5 pixels.

Accuracy Test

To test the accuracy of our system, a small experiment was performed. In the same setup as the above experiment, we projected a laser light on a metal cube. Later, a



(a) Metal cube. (b) Metal cube on a thin metal sheet.

Figure 4.16: Metal cube images for accuracy test.

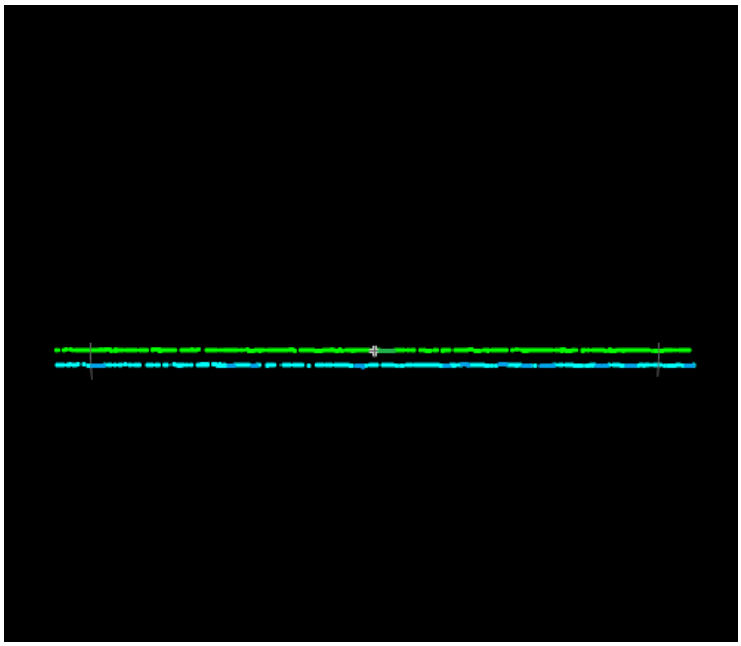


Figure 4.17: 3D reconstructed laser profiles for accuracy test.

thin sheet of height 0.038 mm was placed below the same metal cube, and we projected a laser line on this cube (See Fig. 4.16). The process mentioned above was repeated to reconstruct the projected laser line in both cases. Figure 4.17 shows the reconstructed laser profile for both cases in one model. We calculated the height difference between the two reconstructed laser profiles. The calculated height difference is 0.033 mm. Therefore, we can determine that our stereo-laser system can achieve an accuracy of 0.005 mm.

4.4 Summary

After understanding the concepts of stereo vision and laser triangulation techniques, the functionality of our stereo-laser system is explained in detail. The proposed algorithm represents the combination of stereo vision and laser triangulation methods. The proposed algorithm is divided into three parts: Pre-processing, laser detection, and 3D reconstruction, for a better understanding of the system. The exceptional cases of intensity distribution, caused due to ambient light, are described with their specific solutions. In addition, the system is capable of neutralizing the effect of vibrations caused by machinery. The performance of the developed system is evaluated by the accuracy test. The experimental results show the reconstructed 3D model of the drum and the wash plate with an accuracy of 0.02 mm. Overall, our system provides a novel, feasible, accurate and cost-effective solution for the 3D measurement of specular surfaces in ambient light.

Chapter 5

Defects Detection

The detection of defects in the product is a critical part of the inspection process. Dents, bumps, and scratches are the type of palpable defects, which could be present in the product. Due to the specularity of the surface, it is difficult to detect these defects using vision-based techniques. Defects detection is a part of our stereo-laser system. This chapter explains our proposed algorithm for automatic defects detection in detail. In this chapter, we have used the stereo-laser system to detect three types of defects (Dents, Bumps, and Scratches) present in the wash plate. The experimental results show that our system cannot only detect the defects but can also measure the size of the defects present in the product. Moreover, the experiments are performed in ambient light. Thus, we have overcome the limitation of traditional 3D measurement methods with this research. Material discussed in this chapter has been published in my publication [9].

5.1 Introduction

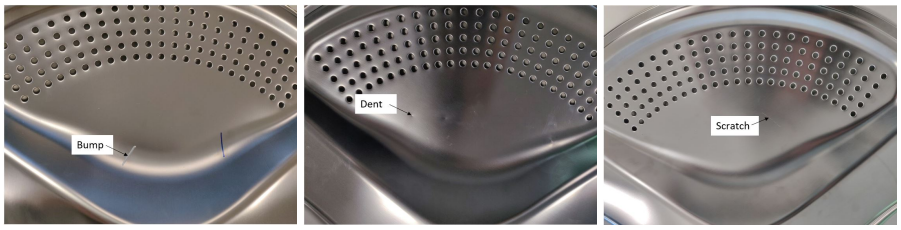
Defects Detection is an integral part of the inspection process. Along with the accurate measurements of the product, it is crucial to detect any defects present in the products. In this research, we principally focus on the detection of defects in highly specular surfaces. There are two classes of defects: Visible and Palpable. Small bumps, dents, cracks and scratches are some examples of defects [176]. Dents, bumps and scratches are the types of defects, which are present in the wash plate. As stated before, the wash plate is made of stainless steel. Thus, the surface of the wash plate is highly specular in ambient light. The visibility of the defects in the wash plate depends on the camera angle, camera position, and its point of view. Additionally, it also depends on the angles of the projection lights. "Indeed, the shape of the product causes reflection in ambient lighting conditions. The specularity of the surface makes it difficult to see every small feature of the product [9]." Therefore, we cannot use computer vision-based techniques for defects detection in reflective surfaces [9].

The comprehensive review of literature on defects detection, which is described in Chapter 2, depicts that the available defects detection methods cannot measure the dimensions of the defects. "If the accurate 3D measurement of the product is necessary for the inspection, it is crucial that the defects are detected at the same time of 3D reconstruction [9]." Therefore, the stereo-laser system is developed to detect the defects present in the product at the same time; it reconstructs the 3D model of the product for quality control. The reconstructed 3D profiles of a defective product are compared with the 3D profile of the non-defective product for defects detection [9].

5.2 Proposed Algorithm for Defects Detection

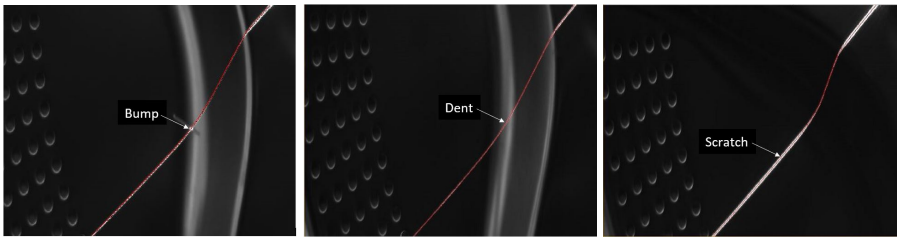
Defects detection is a part of the stereo-laser system. As described, a laser line is projected on the object, which is rotating at a constant speed. The stereo cameras capture the image of the projected laser line in ambient light. The projected laser line appears deformed, and it resembles the shape of the object. The stereo images are processed, and the projected laser line is accurately detected using the proposed algorithm of laser detection. After that, the concept of stereo matching is applied only to the detected laser line in stereo images. We get the disparity value for the laser line as an output of stereo matching. The detected laser points are transformed into WCS, using the internal camera parameters, the relative pose, and the disparity values. This process is repeated for each pair of stereo images. The 3D reconstructed laser profile is merged with the previously reconstructed laser profiles after it has been rotated according to the rotation of the product. After repeating this process for the total number of scans, the 3D model of the object is created.

Instead of rotating the 3D profile for defects detection, we translate the 3D profile in one direction for defects detection [177]. In this experiment, we translate the 3D profile along the Z-axis. Now, the translation distance between every two consecutive scans is constant. This process will create the 3D unwrapped model of the product. Next, we find the transformation between the 3D unwrapped model of the defective product and the 3D unwrapped model of the ideal product by using the registration process [178]. In this process, it is assumed that both translated models have a known approximate spatial relation. The pose of the defective product is transformed according to the outcome of the registration process [179]. Later, we compute the distances between the points of the two models. The distances are calculated using the K-dimensional (KD) tree method. In this method, the points of the defective product are organized in KD-tree. This speed up the search for the closest point in the ground truth data [180]. The output distance map clearly shows the defects present in the product. Moreover, the dimension of the defect can also be



(a) An object with bumps as a defect. (b) An object with dents as a defect. (c) An object with scratches as a defect.

Figure 5.1: Defected object [9].



(a) Detected laser profile in presence of bump. (b) Detected laser profile in presence of dent. (c) Detected laser profile in presence of scratch.

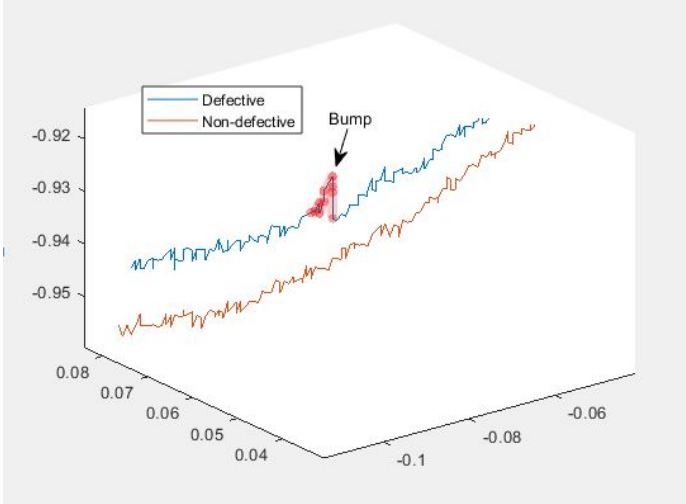
Figure 5.2: Detected laser profiles in the presence of defects [9].

calculated from the distance map.

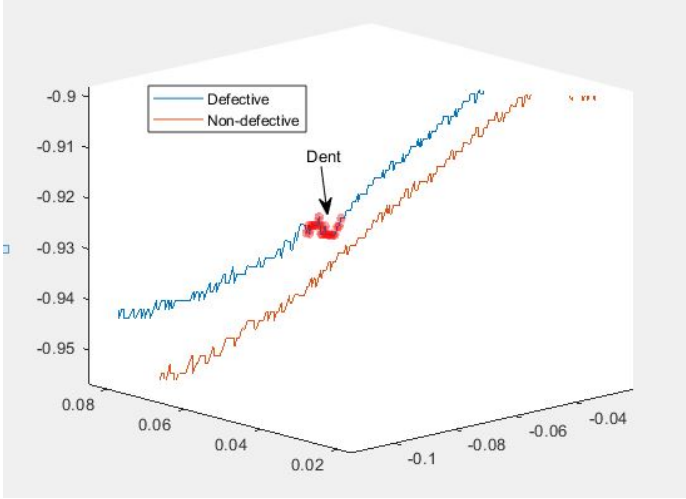
Sometimes, there is a difference in the elevation of the two products. If we compare the 3D reconstructed model with the ground truth model, this elevation difference is also detected as a defect in the product. To avoid this problem, we have proposed this new approach, which compares the translated 3D models instead of reconstructed 3D models. The product is rejected automatically, if any defects are detected in the product.

5.3 Experiments and Results

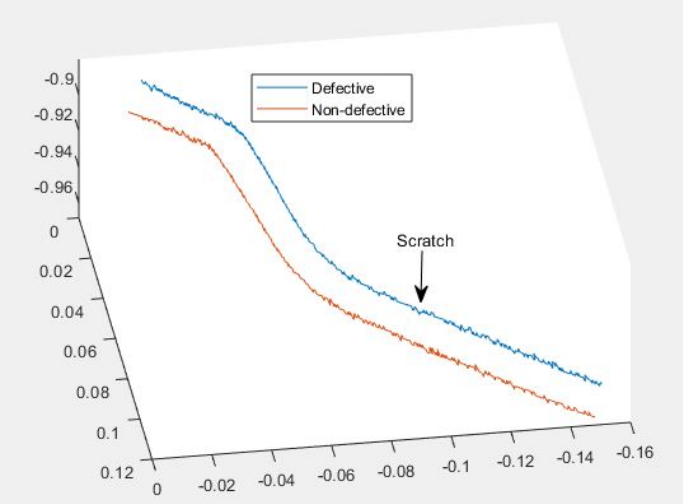
The setup for this experiment is similar to the setup of the stereo-laser system. The wash plate of the washing machine drum is used as an object. The object has three types of defects: Dents, Bumps and Scratches (See Fig. 5.1). The effect of a bump is created by using the marker. We have projected a narrow band of blue light laser at the location of the defects on the object. The stereo images of the object, with a pro-



(a) Detected Bump.

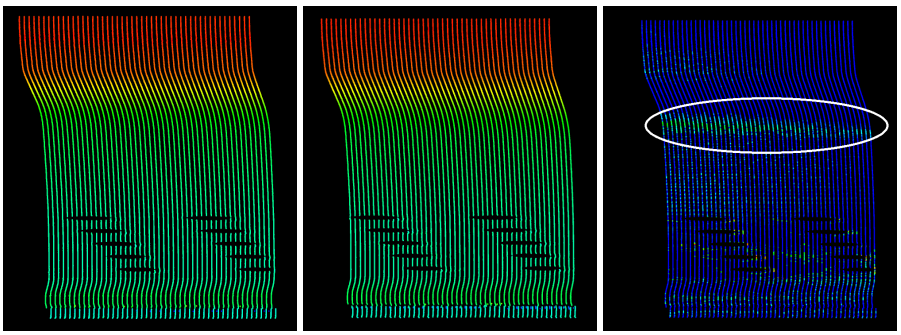


(b) Detected Dent.



(c) Detected Scratch.

Figure 5.3: 3D profiles of defective and non-defective products in WCS [9].



(a) Translated model of the defected product. (b) Translated model of an ideal product. (c) Distance map with a detected defect.

Figure 5.4: Defects Detection.

jected laser line, are captured in the calibrated environment. Figure 5.2 shows the images of the detected laser line in the presence of defects on the object. After detecting the laser line in stereo ROIs, the 3D laser profiles are obtained after performing stereo matching. We have shown the reconstructed 3D profiles in the presence of defects with the ground truth data in Fig. 5.3, to show that the stereo-laser system has accurately detected the defects [9].

The proposed defects detection algorithm is used to detect the bump, which is a line drawn by a marker (See Fig. 5.1a). The unwrapped 3D model of the product is created by combining the translated 3D laser profiles. Figure 5.4a shows the unwrapped 3D model of the defective product. For visualization purposes, we have shown only 50 scans instead of a total of 2500 scans. We have also shown an unwrapped 3D model of the ideal product in Fig. 5.4b. The distance between each consecutive scan is 0.005 points. In the distance map shown in Fig. 5.4c, we can see the green line in the highlighted region, which represents the line drawn with a marker to create the effect of a bump in the product. If any defect is present in the product, whether it is a dent or bump, it will be present in at least 10 continuous reconstructed 3D profiles. Therefore, the scattered green points present in the distance map are not considered as defects, rather they are considered as noise. The dimension of this bump can be calculated from the distance map. Therefore, we can accurately detect the dents and bumps of size 0.03 mm and 0.07 mm, respectively. However, accurate detection of scratches is proven to be more difficult compared to other palpable defects such as dents and bumps. As shown in Fig. 5.3c, "the depth of the scratch nearly equals to the noise present in the output. Therefore, the removal is noise is necessary for detection of scratches [9]."

5.4 Summary

To summarize, we cannot only detect but also measure the palpable defects of the minimum size of 0.03 mm accurately. As the experiments are performed in ambient light, the limitation of traditional 3D measurement techniques has been overcome. Along with the detection of defects, this system reconstructs the 3D model of the product for 3D measurement, at the same time. This combined approach saves time in real-world quality control processes. Unlike commercial profilers, there is no constraint such as a trade-off between the field of view and the accuracy of the output. Therefore, we can perform both tasks of 3D measurement and defects detection without making any changes in the setup. Moreover, there is no requirement for extra hardware to make the defects visible. The accuracy of our system is independent of the visibility of the defects. Thus, our approach is proven to be a simple, feasible, accurate and cost-effective solution for defects detection in highly specular surfaces.

Chapter 6

Comparison of Red versus Blue Light Laser for Stereo-Laser System

In this chapter, we have tried to examine the accuracy of our system with different colors of laser light. As the experiments are performed in ambient light, it is essential to project the correct type of laser light to achieve maximum accuracy. Two different colors (Red and Blue) of laser lights are compared in this experiment. Additionally, the commercial laser profilers that use red light laser are compared with those using a blue light laser. We have reconstructed 3D profiles of three objects of different shapes using these two laser lights. The output 3D profiles obtained using a red light laser are compared with what was achieved using a blue light laser. The results are quantitatively evaluated in terms of accuracy with the ground truth 3D model of the acquitted objects for accuracy evaluation. Moreover, we have also discussed the dependency of laser light on the specularity of the surface. The material of this chapter has already been published in my publication [159].

6.1 Red and Blue Light Laser

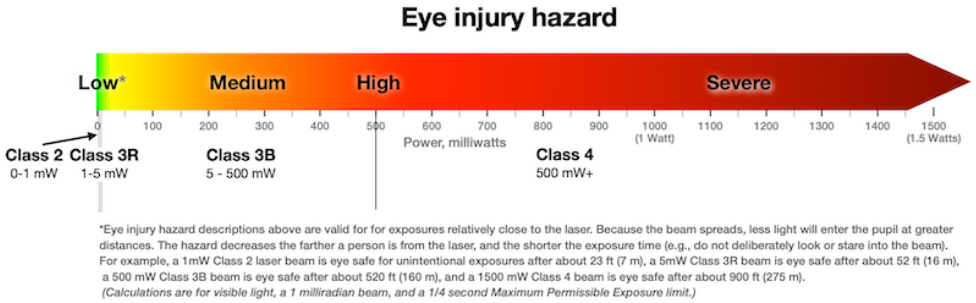


Figure 6.1: Laser classes and eye injury hazard [185].

The accuracy of the stereo-laser system mainly depends on the accurate laser

detection. To get accurate results in ambient light, the projected laser light needs to be sharp and narrow, and it should resemble the shape of the product. In this chapter, we have used red and blue colors of laser light, to evaluate the performance of the system with different colors of laser light in ambient light.

A laser emits coherent light, due to which the projected laser beam stays narrow and focused over a great distance [181]. The wavelength of the blue light laser diode is normally 450 nm, 473 nm, or 488 nm. On the other hand, the wavelengths of 638 nm, 650 nm, and 670 nm belong to the red light laser diode [182]. In the electromagnetic radiation spectrum, all these wavelengths come under the visible light region [183]. We can classify the visible-beam lasers into four classes based on their maximum output power: Class 2, Class 3R, Class 3B and Class 4 [184]. "Figure 6.1 shows that the chances of eye injury hazard increases as the lasers' power increases [159, 185]." Keeping the safety parameters in mind, we have used class 2 lasers to perform experiments in ambient light.

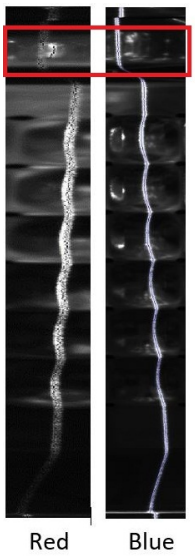


Figure 6.2: Detection of laser light in the presence of highlight [159].

One critical factor of the laser detection process is that the reflected region in the shiny surface, should not be considered as a part of the projected laser profile, because of its high intensities. In contrast to blue light laser, the red light laser has higher absorption depth. Thus, it penetrates deeper into the target surface. Moreover, the chances of the red light laser merging with the background increases, because it appears blurry and diffused in ambient light. On the other hand, the blue

light laser generates a sharply focused band when projected on the reflective surface in ambient light [159, 186]. To prove this, an example is shown in Fig. 6.2. The red light and blue light lasers, with the same output power, are projected on the washing drum. The stereo-laser system is used to detect the projected laser line. "Here, the red light laser has low intensities compared to the blue light laser [159]." Therefore, the stereo-laser system detects the highlighted region as a part of the laser line in the case of a red light laser. In contrast, the projected blue light laser is detected accurately even in the presence of the highlighted region [159]. To check the overall impact on the accuracy of the system, we have carried out few experiments, which are discussed later in this chapter.

6.2 Comparison of Commercial Solutions

Table 6.1 compares laser profilers of LMI Technologies, Micro-Epsilon, and Cognex, which work on the concept of sheet-of-light for 3D measurement. The products are compared in terms of field of view (FOV), resolution, measurement distance and clearance distance. Also, the compatibility of the products with the reflective surface is mentioned in this table. The - indicates the lack of information regarding that particular product. We can see a trade-off between the accuracy of the output and the field of view covered by the system. The products with higher resolution cover very small FOV. Therefore, multiple units of the product will be required to inspect large objects. This directly increases the cost of the inspection process. Additionally, the resolution of the red light laser profiler is lower than that of the blue light laser profiler. Moreover, the red light laser profilers cannot inspect the highly specular surfaces. Also, one common disadvantage is that all profilers mentioned in the table, do not work in ambient light conditions when working with the reflective surface. Thus, blue light laser profilers are preferred over red light laser profilers to inspect shiny surfaces [159].

6.3 Experiments and Results

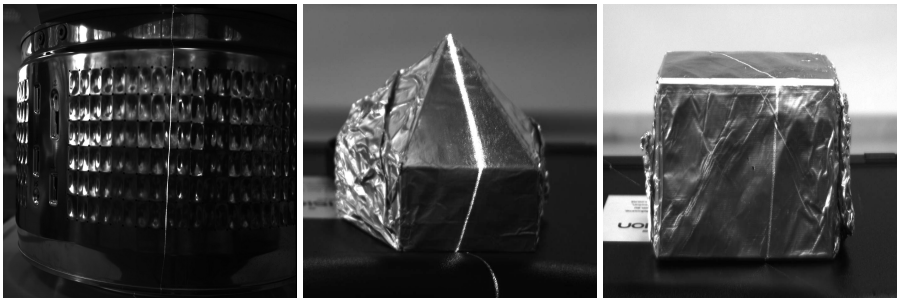
The setup of the experiment is the same as the setup of the stereo-laser system explained in Chapter 3. We have used class 2M red and blue light lasers to perform 3D measurements in ambient light. Both lasers have the same maximum output power of 20 mW. Also, the fan angle of both lasers is 45 degrees. The position of the laser is the same for both lasers, and it does not change throughout the experiment. The baseline distance between the stereo cameras is 130 mm for this experiment. In this experiment, we have projected a laser line on three different shapes of the object.

Table 6.1: Comparison of laser profilers [159].

Laser Profiler	Laser Light	Field of View (mm)	Resolution Z (mm)	Resolution X (mm)	Measurement Distance (mm)	Clearance Distance (mm)	Compatibility with Shiny Surface
Gocator 2510, 2512 [187]	Blue	13.0-14.5	-	0.008	6	17.0	Yes
Gocator 2520 [187]	Blue	25.0-32.5	-	0.013-0.017	25	47.5	Yes
Gocator 2880 [187]	Red	390-1260	0.092-0.488	0.375-1.1	800	350	No
Cognex In-Sight Laser Profiler DS910B [188]	Blue	9.4-10.7	0.001	0.0073-0.0084	8	52.5	Yes
Cognex In-Sight Laser Profiler DS910 [188]	Blue	23.4-29.1	0.002	0.0183-0.0227	25	53.5	No
Cognex In-Sight Laser Profiler DS1101 [188]	Red	64-162	0.010-0.052	0.063-0.158	220	135	No
Cognex In-Sight Laser Profiler DS1300 [188]	Red	90-140	0.016-0.265	0.088-0.410	725	180	No
Micro-Epsilon 29xx-100 [189]	Red	83.1-120.8	0.012	1280 points / profile	100	-	No

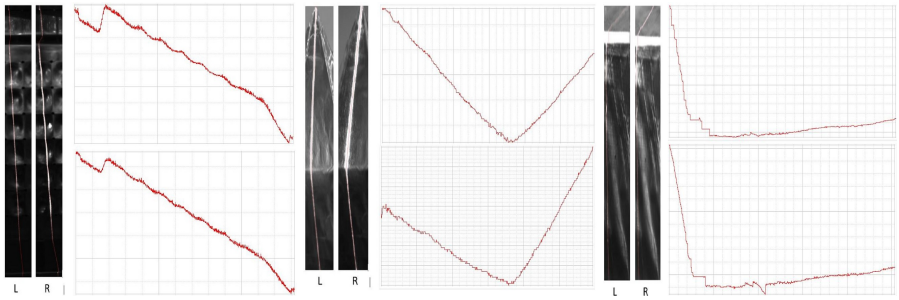
We have tried to reconstruct the projected laser profile on a washing machine drum, a prism and a cube. As we know, the drum is made of stainless steel, which makes its surface highly specular. On the other hand, we have wrapped the cube and the prism in aluminum foil to get the effect of the reflective surface. First, we have used the red light laser for 3D reconstruction. Later, the experiment is performed using the blue light laser in the same working environment. We have used the proposed algorithm of the stereo-laser system for 3D measurement tasks [159].

In a calibrated environment, the laser line is projected on three objects. Figure 6.3 shows images of the different objects with a projected laser band. The projected laser band is detected in the captured stereo images. The left and right ROI images with detected laser profiles in the case of red light laser are shown in Fig. 6.4. The process is repeated for each object using the blue light laser. The detected laser profiles in stereo ROIs are shown in Fig. 6.5. Stereo matching is performed on detected laser



(a) Drum. (b) Prism. (c) Cube.

Figure 6.3: Objects with projected laser profile [159].



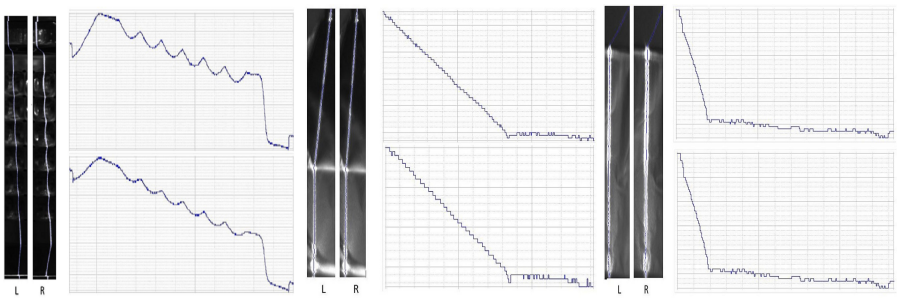
(a) Detected laser profile in drum. (b) Detected laser profile in prism. (c) Detected laser profile in cube.

Figure 6.4: ROIs with detected laser profiles using red-light laser [159].

profiles in the ROI images. The calculated disparity values are used to transform the detected laser points into world coordinates. In Fig. 6.6, we compare the output reconstructed 3D profiles by red light laser with the output reconstructed 3D profiles

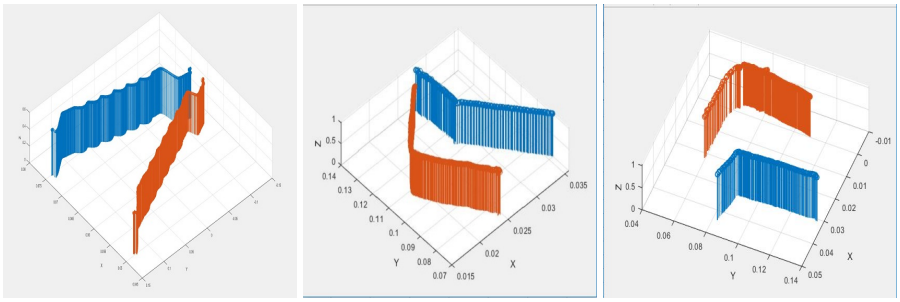
Table 6.2: Output accuracy for red and blue light laser [159].

Parameters	Object 1: Drum		Object 2: Prism		Object 3: Cube	
	Red	Blue	Red	Blue	Red	Blue
Total No. of Scans	1250	2500	500	1000	100	200
Accuracy (mm)	0.7	0.02	0.4	0.02	0.4	0.018
No. of False Positives	14	6	12	0	6	0
Maximum Deviation (Pixels)	± 5	± 3	± 3	± 1	± 3	± 1



(a) Detected laser profile in drum. (b) Detected laser profile in prism. (c) Detected laser profile in cube.

Figure 6.5: ROIs with detected laser profiles using blue-light laser [159].



(a) Reconstructed laser profiles for a drum. (b) Reconstructed laser profiles for a prism. (c) Reconstructed laser profiles for a cube.

Figure 6.6: 3D reconstructed laser profiles [159].

by blue light laser. The output accuracy is dependent on the detection of laser in the ROI images [159].

Table 6.2 summarizes the performance of the stereo-laser system with different laser lights. The total number of scans required to reconstruct the product is mentioned in the table. As the width of the red light laser is more significant than the blue light laser, the total number of possible scans with a red light laser is nearly half of the number of scans possible with a blue light laser. "Here, the number of false positive specifies how many times the reflection is falsely detected as a part of a laser line [159]." The deviation results depict that the detected red light laser is quite noisy compared to the detected blue light laser, which is quite sharp. Additionally, the number of false positives for small objects is 0, when using the blue light laser. The maximum accuracy achieved by the system with a red light laser is 0.4 mm. On

the other hand, the maximum accuracy of 0.018 mm is achieved with the blue light laser. The maximum deviation is only 1-3 pixels with the blue light laser. Therefore, the stereo-laser system can detect all small features of the object accurately when the blue light laser is projected on the object [159].

6.4 Summary

To conclude, a blue light laser is proven to be more accurate than a red light laser for the 3D shape measurement of highly specular surfaces in identical conditions. Additionally, the projected beam of red light laser diffuses and merges with the reflection caused by ambient light. On the other hand, the blue light laser does not penetrate the shiny surface. It provides a sharp narrow beam when projected onto a highly specular surface even in ambient light. Therefore, we can accurately detect the blue light laser even in the presence of reflection. Even comparing different commercial laser profilers, we can see that the blue light laser profilers have better accuracy than red light laser profilers. In addition, the red light laser profilers are not compatible in inspecting specular surfaces. Moreover, the accuracy of the output is highly improved by using the blue light laser. Thus, we recommend using a blue light laser with the stereo laser system for accurate 3D measurement of highly specular surfaces in ambient light.

Chapter 7

Diameter and Height Measurement using a Single Image

In production lines, different applications require different features to be inspected. Due to time constraints, some applications may require only essential parameters of the manufactured object to be measured during the inspection. For example, the perimeter and height are considered essential parameters for symmetric 3D geometric shape objects. This chapter proposes a simple approach to calculate the diameter and height of a highly specular or transparent cylinder from a single image captured in a calibrated environment. The dimensions are measured accurately for three different sizes cylinders. Out of three, one is made of transparent plastic material, and the other two are made of stainless steel. Moreover, the experiments are performed in ambient lighting conditions, overcoming the limitations of many standard 3D measurement methods. In the experiment, the cylinders are translated in X and Y directions with respect to the camera to check the effects of camera position on the accuracy of the results. The dimensions of the cylinders are calculated and compared for five different poses. In the end, the accuracy of the system is analyzed by comparing the results with actual dimensions. The material of this chapter has been published in my submitted publication [190].

7.1 Rationale and Significance of this Research

The detailed 3D model of the object is not required by some applications during the inspection process. We cite [190], "Only the overall shape of the object and some necessary measurements are sufficient." For a cylindrical object, the diameter and height are considered the essential parameters to be measured. We focus on cylindrical objects in this research. Generally, we require at least two images: one from the top and another from the front to calculate the diameter and height of the cylinder. To capture these two images, we need to use at least two cameras. Before capturing the images, we first need to calibrate the cameras accurately. The captured images need to be processed to measure the dimensions [190].

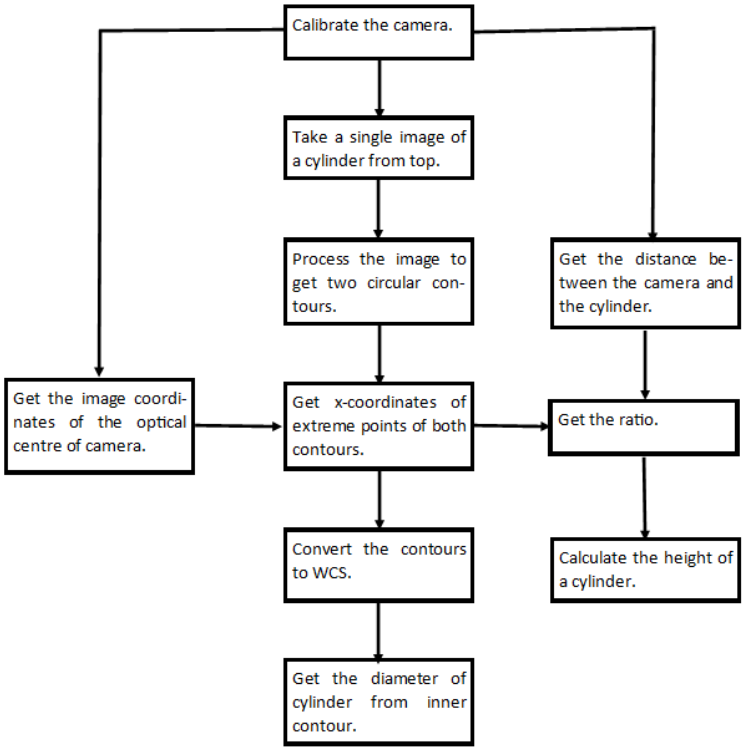


Figure 7.1: Overall approach for 3D reconstruction of the cylinder [190].

In high-scale production lines, stereo vision or laser triangulation methods reconstruct a 3D object using multiple cameras or laser respectively [165]. However, both concepts have their own sets of difficulties. To understand these difficulties, we suppose that the cylinder is in a vertical position. Two slightly different images of the cylinder are captured in a calibrated environment for stereo matching. Stereo matching is performed on these two slightly different images to generate the disparity map [191]. However, the probability of finding corresponding points accurately in slightly different images of highly specular or transparent objects is low. The difficulties are increased due to the ambient lighting of the working environment. Also, we can reconstruct only the front half of the cylinder in one go from the disparity map. Moreover, we need to rotate the cylinder by 180 degrees to reconstruct the other half. In the end, both halves are merged to reconstruct the whole cylinder. However, it is challenging to merge both halves of the cylinder accurately.

Another popular method of inspection uses the concept of laser triangulation. In

laser triangulation, a thin luminous straight laser line is projected on to the surface of the cylinder. The camera captures images of the projected line. "To reconstruct the whole surface of an object, the object must be moved relative to the measurement system, i.e., the unit built by the laser line projector and the camera [165]." If the cylinder is rotated at regular intervals for reconstruction, it is very challenging to calibrate a laser and a camera with respect to a rotating positioning system. On the other hand, if we use a linear positioning system, then only half of the cylinder would be reconstructed in one scan. Again, the same problem of merging two halves of the cylinder accurately arises. Also, the optical signal cannot be correctly retrieved if the surface of the cylinder is reflective. "Therefore, it is usually challenging for any optical method to accurately measure shiny objects or objects with a broad range of reflectivity variation across the surface [190]."

The time constraint during industrial inspection does not allow us to use the expensive process of finding corresponding points such as stereo vision or to do laser triangulation [165]. The flow chart in Fig. 7.1 represents our suggested approach for the 3D reconstruction of a cylinder using a single camera. "The same approach could also be applied to calculate the dimensions of other solid shapes such as cubes, cuboids, or prisms [190]." However, the border of the bottom part of the object must be visible in the image [190].

7.2 Single Camera Calibration and Ellipse Fitting

Single Camera Calibration

"To perform metric measurements accurately in the *world coordinate system* (WCS), the only prerequisite is that the camera has been calibrated. The camera is calibrated for a specific plane in WCS for which measurements are obtained [165]." In our case, we have placed the calibration plate beside the cylinder to define the measurement plane. The calibration plate with hexagonally arranged marks has been used in this research. Out of five finder patterns present in the plate, at least one of them should be completely visible in the image for estimating the relative camera pose [190].

For calibration, we have captured multiple images of a calibration plate in different poses in a specified plane. Here, the specified plane is the measurement plane, which is defined as the plane $Z = 0$ of the WCS. "The captured images, along with the internal camera parameters and the description of calibration plate, work as inputs for calibration. The output comprises the internal and external camera parameters [165]." External parameters are determined from the pose of the calibration plate, which is placed directly on the measurement plane. The external camera parameters describe the relationship between the measurement plane and the cam-

era [165].

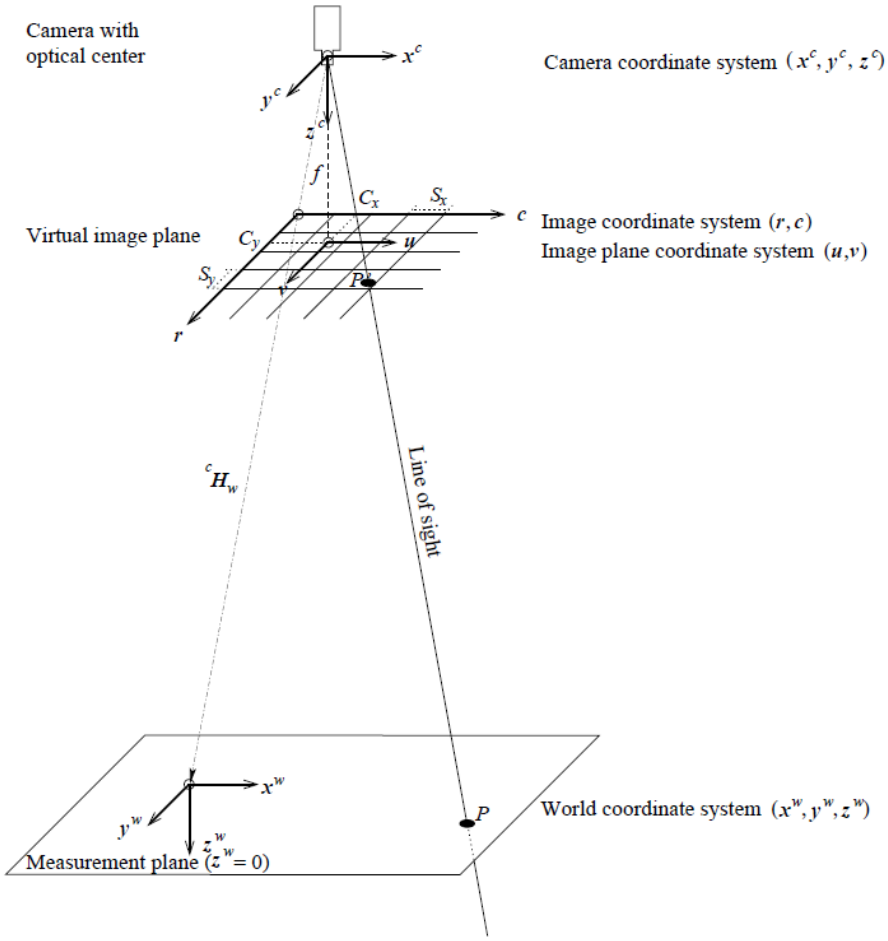


Figure 7.2: Transformation of image points in WCS [165].

Figure 7.2 shows how an image point is transformed into world coordinates. The goal is to determine the world coordinates of the corresponding point in the measurement plane from the image coordinates of one point. "For this, the line of sight, i.e., a straight line from the optical centre of the camera through the given point in the image plane, must be intersected with the measurement plane [165]." First, the image coordinates are transformed into camera coordinates and later into world coordinates using the calibration data. The process of measuring dimensions accurately after calibrating the camera is described in the next section.

Ellipse Fitting

Ellipse extraction is an essential task in the vision-based application as this geometric shape frequently occurs [192,193]. The process of ellipse extraction is difficult for highly specular or transparent objects, because low contrast affects accurate extraction of edge contours [194]. "There are three main categories of ellipse extraction methods: Fitting algorithms, Hough-transform techniques, and edge contour following methods [195]." An ellipse equation is fitted to a sequence of points in the first category. The second category performs ellipse extraction by considering it as a peak-seeking problem in parameter space. In the last category, arc-segments are extracted and grouped into an elliptic hypothesis [195].

In this research, we use the fitting-based edge extraction method, as it is proven to be the most accurate among other available methods [196–199]. This method fits an ellipse equation to a sequence of points and minimizes the error between the given data and an ellipse equation [200–203]. "There are two types of ellipse-fitting algorithms, algebraic fitting which solves the minimization problem by exploring the algebraic equation, or geometric fitting which minimizes geometric distances of given points to the fitted elliptic curve [195]."

For this research, we use Fitzgibbon's approach for ellipse fitting. This method incorporates the ellipticity constraint into the normalization factor by minimizing the algebraic distance. The benefit of using this approach is that it provides useful results under all different noise and occlusion conditions. Also, this direct least-squares fitting method is specifically designed for ellipses. It guarantees that the output will be an ellipse [204].

7.3 Measurement of the Diameter of a Cylinder

In contrast to conventional 3D reconstruction techniques, the camera is looking at the cylinder from the top instead of the front in our approach. We take a single image of the cylinder from the top after calibrating the camera. Figure 7.3 shows an example of the image of the cylinder used for 3D reconstruction. Among the three cylinders used for this research, the cylinder shown in the image is the largest one. "This right circular cylinder is made of stainless steel, which makes its surface highly reflective. Here, the cylinder is positioned in a way that the bottom and the top of the cylinder is perpendicular to the optical axis of the camera [190]." We can noticeably see two elliptical contours in the image. The camera is calibrated for the bottom plane. Therefore, the edge of the inner ellipse is extracted to measure the diameter of the cylinder. In Fig. 7.3, the image of the cylinder with detected elliptical contours is shown. "It is essential to extract the edges of elliptical contours

Figure 7.3: An image of the cylinder from the top with extracted edges [190].

accurately. We need to make sure that the reflection caused by ambient lighting does not affect the accuracy of edge extraction [190]."

Now, we fit an ellipse through the inner contour, which gives us the diameter of the cylinder in the image coordinate system. As stated before, the Fitzgibbon et al. approach is used for ellipse fitting [204]. The diameter in metric units is calculated by projecting this extracted inner contour in WCS. We cite [190], "The world coordinates are determined for each point of the detected elliptical contour in the measurement plane. A circular contour will be formed in the measurement plane. We fit a circle through this circular contour using an algebraic approach." This approach minimizes the algebraic distance between the contour points and the resulting circle [206]. The diameter of this circular contour in plane $Z = 0$ corresponds to the diameter of the cylinder. The diameter is calculated in metric units using the calibration data [190].

7.4 Measurement of the Height of a Cylinder

In this section, the method to measure the height of the cylinder is described in detail. The setup to measure the height is the same as before. First, the image of the cylinder is processed to extract the edges of two ellipses, which presents the top and the bottom of the cylinder. After extracting the edges, we fit an ellipse through these contours using Fitzgibbon's approach [205]. Figure 7.3 shows the cylinder with two

extracted edges of elliptical contours. We cite [190], three possible cases: (1) the projection centre of the camera and the cylinder's centre line are collinear. (2) the cylinder is horizontally shifted. (3) the cylinder is vertically shifted.

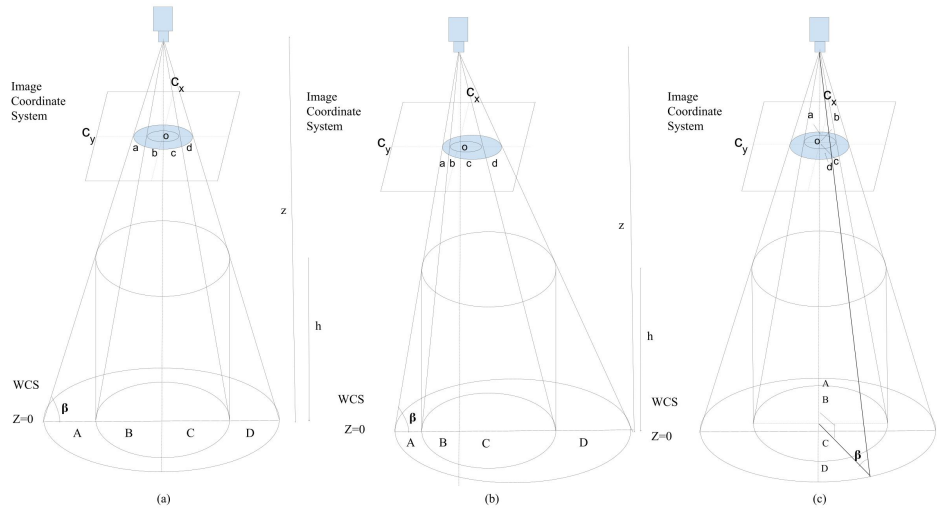


Figure 7.4: Three different cases for height measurement [190].

Case 1, Case 2 and Case 3 are illustrated in Fig. 7.4, (a), (b) and (c), respectively. As we can see in Fig. 7.4 (a), we get two concentric elliptical contours in the image plane for Case 1. Figure 7.4 (b) shows the second case. Here, the centre of the camera is not collinear with the centre line of the cylinder. Therefore, we get two elliptical contours which are not concentric. However, the centres of these two elliptical contours share the same horizontal axis. In Fig. 7.4 (c), the third case where the cylinder is vertically displaced is shown. Similar to Case 2, the centre of the camera is not collinear with the centre of the cylinder, and we get two non-concentric elliptical contours in Case 3. Instead of horizontal axis, the centres of these two elliptical contours share the same vertical axis. Now, our method to measure the height of the cylinder is described in detail below [190].

To calculate the height, it is essential to find the ratio of $a : b : c : d$ (See Fig. 7.4). Here, a and d are the distance in pixels between the left extreme points and the right extreme points of both contours, respectively. In cases 1 and 2, b and c are the distance between the optical centre and the left and the right extreme points of the internal elliptical contours, respectively. In case 3, b and c are the distance between the optical centre and the top and the bottom extreme points of the internal elliptical

contours, respectively. Also, the distance in pixels between the top extreme points and the bottom extreme points of both contours are a and d, respectively. To find the ratio of $a : b : c : d$ for case 1 and 2, it is essential to know the X - coordinates of the horizontal extreme points of the contours and of the optical centre of the camera in the image plane. Contrastingly, in case of vertical displacement, we need to know Y- coordinates of the vertical extreme points of the contours and of the optical centre of the camera in the image plane. The image coordinates of the optical centre of the camera are calculated during the calibration. We can get this parameter by extracting the intrinsic parameters of the camera. After obtaining all the image coordinates, we can calculate the ratio $a : b : c : d$. This ratio will now correspond to the ratio of $A : B : C : D$ in the world coordinate system. A and D represent the height of the cylinder in measurement plane [190].

We already know that,

$$B + C = D_1 \tag{7.1}$$

where, D_1 is the diameter of the cylinder which is calculated before.

From Fig. 7.4,

$$A = ax,\ B = bx,\ C = cx,\ D = dx \tag{7.2}$$

Substituting in Equation 7.1,

$$bx + cx = D_1 \tag{7.3}$$

Here, x is an unknown factor which converts from image coordinate system to world coordinate system:

$$(b + c)x = D_1 \tag{7.4}$$

$$x = D_1/(b + c) \tag{7.5}$$

The value of x is substituted in Equation 7.2 to calculate A, B, C and D. We also need to know the distance between the camera and the measurement plane to measure the height of the cylinder. While calibrating the camera, we place the calibration plate directly onto the measurement plane. Therefore, we can find out the distance between the camera and the measurement plane by determining the pose of the calibration plate. The pose of the calibration plate is obtained during the calibration process. By accessing the external parameters of the camera, we can measure the distance between the camera and the measurement plane denoted as Z [190].

Now,

$$\tan\beta = Z/(A + B) = h/A, \tan\beta = Z/(C + D) = h/D \tag{7.6}$$

In Fig. 7.4 (c), the angle β is shown slightly to the right to understand the image better. In reality, the angle β is the angle of the right triangle formed by the camera,

the bottom extreme point of the outer ellipse and the projection of the centre of the camera in the measurement plane.

$$h = (A/(A+B)) \cdot Z, h = (D/(C+D)) \cdot Z \tag{7.7}$$

Here, h is the height of the cylinder. We can cross-check the results by using the $D/(C+D)$ ratio for height calculation. By measuring the cylinder's height and diameter, we can accurately reconstruct the cylinder from a single image.

7.5 Experiments and Results

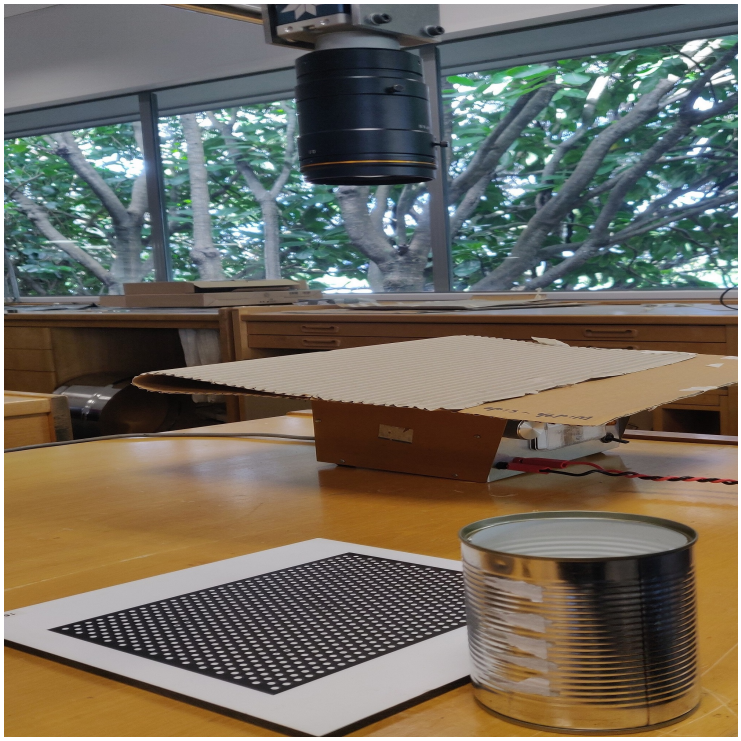


Figure 7.5: Setup for experiment [190].

Genie Nano M4020 monochrome camera, which has $4,112 \times 3,008$ resolution, has been used throughout this research. HALCON is the software used to perform image processing tasks. Firstly, we calibrate the camera by using the HALCON calibration plate with hexagonally arranged marks. As stated before, the calibration

plate is placed beside the cylinder (See Fig. 7.5). This specifies the bottom of the cylinder as the measurement plane. For large objects, we can also place the calibration board inside the cylinder. The camera is mounted at a distance Z from the top of the cylinder [190].

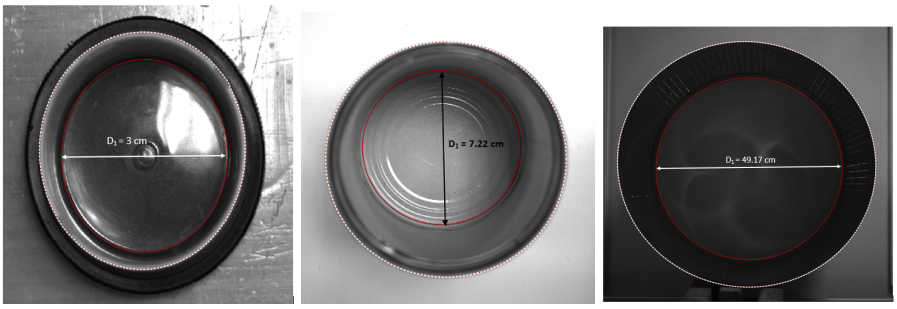
To test this concept, we have used three cylinders of different dimensions. Among them, two cylinders are made of stainless steel, and one is made of transparent plastic material. Here, a can and a washing machine drum are used to study the case of highly specular surfaces. A small plastic container of cylindrical shape is used as a transparent object. Figure 7.6 shows the images of three cylinders used for this research. We have performed all experiments in ambient lighting conditions of the working environment. "A single image of the cylinder with the calibration board is captured from the top. The camera is calibrated by using the image of the calibration board as a reference. Now, the cylinder is translated in X and Y directions with respect to the camera." We have calculated the dimensions of the cylinder for five different positions [190].

(a) Transparent Cylinder. (b) Can. (c) Drum.

Figure 7.6: Cylinders [190].

The input images of the objects used for this research is shown in Fig. 7.7. In Fig. 7.7a, the case when transparent Cylinder 1 is in leftmost position with respect to the camera is shown. Figure 7.7b represents the top-most position of Cylinder 2 (a can) with respect to the camera. Figure 7.7c depicts the collinear axes case for Cylinder 3 (a drum).

The edges of the two elliptical contours are extracted by applying thresholding and edge detection algorithms. After extracting the edges, Fitzgibbon's approach is applied to fit an ellipse through these contours [205]. "The image coordinates of the extreme points of the two elliptical contours are obtained from elliptical contours. Also, the projection of the optical centre of the camera in the image coordinate system is obtained as a result of calibration." The diameter of the inner ellipse in the



(a) Cylinder 1 in left-most position. (b) Cylinder 2 in top-most position. (c) Cylinder 3 in centre position.

Figure 7.7: Cylinders with extracted elliptical contours [190].

world coordinate system depicts the diameter of the cylinder. In Fig. 7.7, the diameters of all three cylinders for the given case are mentioned [190].

Table 7.1: Comparison of output dimensions [190].

Objects	Cylinder 1			Cylinder 2			Cylinder 3		
Positions	Height (cm)	Diameter (cm)	Accuracy (%)	Height (cm)	Diameter (cm)	Accuracy (%)	Height (cm)	Diameter (cm)	Accuracy (%)
Actual Dimensions	3	3	-	12.5	7.3	-	37	49.5	-
Left-most Position	2.79	2.8	94	11.58	7.22	94.94	36.71	49.26	99.39
Right-most Position	2.87	3	97.83	12.38	7.21	98.93	36.85	48.98	99.22
Centre Position	2.89	3	98.16	12.43	7.22	99.24	36.92	49.17	99.53
Top-most Position	2.69	2.9	96.83	12.49	7.22	99.54	35.74	48.79	97.72
Bottom-most Position	2.75	2.9	95.5	12.37	7.22	98.93	36.98	49.24	99.67

Now, the crucial parameters for height calculation are the distance between the camera and the measurement plane and the image coordinates of the optical centre of the camera. We have already obtained both parameters as a result of calibration. In this experiment, the distance between the camera and the measurement plane is 35.194 *centimeters* (cm) for Cylinder 1 and 2 and 106.062 cm for Cylinder 3.

The height of the cylinder is measured using the method mentioned earlier. "Table 7.1 compares the results for five different positions of the cylinders in either X or Y direction with the actual parameters for Cylinder 1, 2 and 3. The actual heights of cylinders 1, 2 and 3 are 3 cm, 12.5 cm, 37 cm, respectively. 3 cm, 7.3 cm, 49.5 cm are the diameters of all cylinders sequentially." Output accuracy is also calculated in the table. Results show that the accuracy of output is independent of the position of the camera. Furthermore, the accuracy of the output predominantly depends on the accuracy of the calibration. The edge detection and ellipse fitting step for the

two elliptical contours is another crucial factor in terms of output accuracy [190].

7.6 Summary

To summarize, this research provides a simple, fast, cost-effective and feasible approach for the measurement of essential parameters of a highly specular or transparent cylinder from a single image. To do so, the only prerequisite is that the top and the bottom of the cylinder are visible and perpendicular to the optical axes of the camera. Also, it is crucial to accurately extract the two elliptical contours which represent the top and the bottom of the cylinder. The accurate calibration of the camera is another essential factor, as all essential parameters required for measurements are obtained from the output of the calibration. Thus, we can state that the accuracy of output is inversely proportional to the calibration error. For small calibration error, we can achieve the highest accuracy. By analyzing the results, the accuracy of the output is independent of the position of the cylinder with respect to the camera or on the size of the cylinder. Another advantage is that the research could be performed in ambient light, which saves time, cost and human effort. Additionally, we overcome the limitation of traditional 3D measurement methods by performing this experiment in ambient light.

Moreover, this research could also be used for calculating the dimensions of other types of solid shapes such as cubes, cuboids or prisms. Also, we can measure the volume and the surface area of the product from measured dimensions. Owing to the time efficiency of the designed solution, the method is applicable in the high-scale industrial production and inspection process.

Chapter 8

Artificial Intelligence for Laser Detection

Active vision sensors are an essential part of 3D shape measurement techniques. An accurate detection of projected laser light is crucial for all active vision sensors. As we know, the specularity of the surface increases the difficulty of active vision sensors. Additionally, active vision sensors tend to have lower accuracy in ambient light. Moreover, the detection of laser is the most essential part of our stereo-laser system as well. Keeping the advantages of AI in mind, this chapter aims to develop a novel machine learning based algorithm for laser detection in ambient light. Data collection is the first step of developing any machine learning based algorithm. The chapter starts with a thorough description of our dataset used for supervised learning models. Next, we have proposed different models using deep conventional networks, such as 1D-CNN, LSTM and MLP. The effect of hyper-parameters on these models is thoroughly analyzed. Furthermore, we have also applied machine learning based traditional regressors to detect the laser. The evaluation metrics such as R^2, RMSE and MAE are used to assess the accuracy of developed models. In the end, an in-depth comparison of all developed supervised learning models with our stereo-laser system is provided.

8.1 Data Description

As stated before, the accuracy of our stereo-laser system depends on the detection of a laser. Accurate detection of the projected laser line in ambient light is the most crucial part of our system. After reviewing the literature on laser detection, we found only one method developed by Zhao et al. [207], which used the concept of neural networks for laser detection. To train the *laser-stripe-detection neural network* (LSDNN) developed by Zhao et al., a large dataset of labeled images is created by projecting a laser line in a dark environment. The LSDNN network classifies the pixels in the image between different labels of real laser stripe, background, ambient light and pseudo-light stripe. Here, pseudo-light stripes are formed due to the surface reflection capabilities of some metals, which are present in the environment [207].

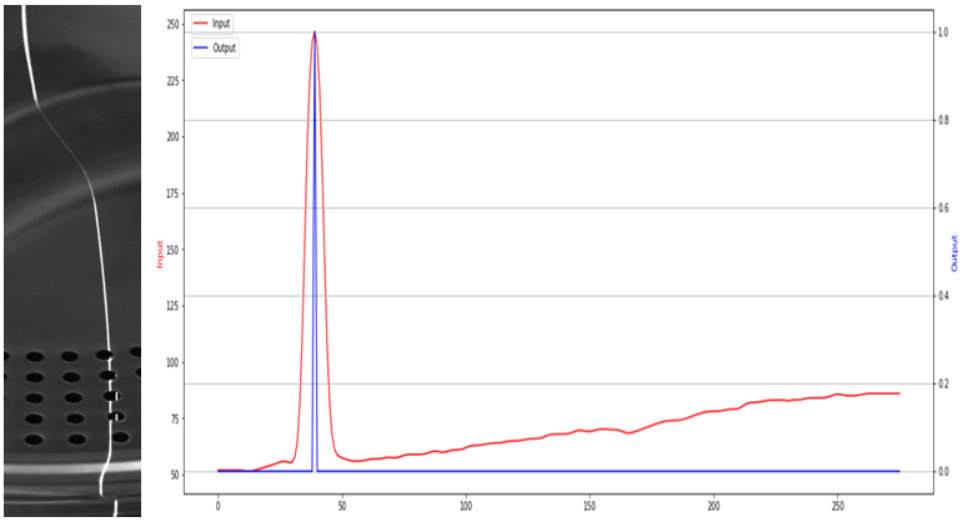
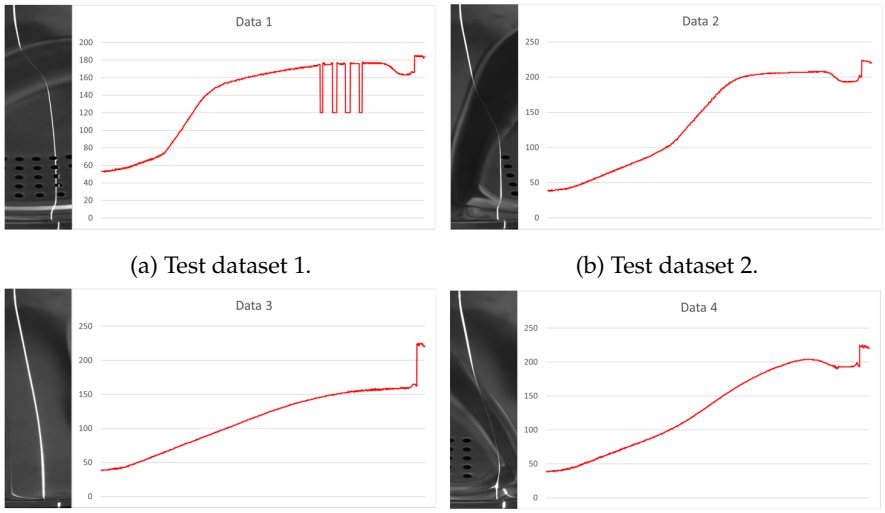


Figure 8.1: Example of an input sequence.

In contrast to [207], we have used 1D data, which shows the intensity distribution for the laser line, to train different supervised learning models. The setup is similar to the one shown in Chapter 3. A laser line is projected on the wash plate in ambient lighting conditions. The laser line of approx 30 cm in length, is projected on different parts of the wash plate, to cover the difficulties caused by various shapes and lighting conditions. As stated before, the wash plate is made of stainless steel, which makes its surface highly specular. A camera captures an image of the product with the projected laser line. The *region of interest* (ROI) is obtained by the method specified in Chapter 4. The ROI is a cropped image which only focuses on the laser line. Figure 8.1 shows an example of an ROI and the smoothed intensity distribution for the first row of the ROI image. The actual intensity distribution curve is smoothed using the Gauss function for better performance [166]. The vertical axis shows the intensity value at each pixel. On the other hand, the horizontal axis shows the length of the sequence, which is the same as the width of the ROI image.

The smoothed intensity distribution of each row in the ROI image is used as an input 1D sequence for machine learning models. If the size of the ROI image is $A \times B$, A presents the length of each input sequence, and B depicts the number of input sequences in one image. In Fig. 8.1, the location of the peak of the output sequence depicts the output label for the corresponding input sequence. The peak shows the centre of the projected laser. Therefore, each input sequence has one output label. In addition to being 1D, our input data is also sequential, as the output of the current

sequence also depends on the output of the previous sequence. Based on our input and output data, many-to-one model architecture will be required to detect the laser.



(a) Test dataset 1.

(b) Test dataset 2.

(c) Test dataset 3.

(d) Test dataset 4.

Figure 8.2: Test datasets.

For our research, we have used 22 ROI images of size 276 $\times$ 1734 pixels. These images cover all exceptional cases of intensity distribution mentioned in Chapter 4. The dataset consists of a total $22 \times 1734 = 38,148$ sequences of length 276. Out of 22 ROI images, we randomly selected 12 for training, 6 for validation, and 4 for testing. Therefore, we have 20,808 sequences for training, 10,404 sequences for validation, and 6,936 sequences for testing. Figure 8.2 shows the four test datasets, i.e. the four images of the projected laser line. We have taken our stereo-laser system as a standard for laser detection. Thus, we have manually labeled each input sequence with the output label generated using the stereo-laser system.

All machine learning based models are programmed by Jupyter Notebook in Anaconda3 (Python3.8). A computer equipped with a 3.6 GHz intel core i7-7700, 16GB of RAM, and an NVIDIA GeForce GTX 1060 graphic card, is used to test all models. Some deep learning frameworks used in this research, their advantages and disadvantages, are summarized in Table 8.1; all these deep learning frameworks are open-source [211].

Table 8.1: Deep learning frameworks and libraries.

Framework/ library	Language	Processor	Benefits	Drawbacks
TensorFlow [208]	Python, C++, CUDA	CPUs/ GPUs/ TPUs	High performance library, operates at large scale datasets, really compatible with industry.	Only Nvidia GPUs are supported.
Keras [209]	Python	CPUs/GPUs	Fast prototyping and easily extensible library, works seamlessly with CNTK, Theano, TensorFlow and Auto-Keras [210]; easily accessible deep learning tools.	It cannot be used as an independent framework, not good enough for data processing.
Scikit-learn [211]	Python	CPUs	Efficacious for statistical modelling techniques such as classification and regression, clustering on medium-scale data based on supervised and unsupervised learning algorithms.	Not efficient with GPU.

8.2 Deep Conventional Networks for Laser Detection

One-Dimensional Convolutional Neural Network

One-Dimensional convolutional neural network (1D-CNN) provides a combination of feature extraction and prediction into one learning body block. This attribute of 1D-CNN maximizes the prediction performance of the network by solving the problem of handcrafted feature selection in other networks [112]. In addition to being relatively easier to train, 1D-CNNs achieves cutting-edge performance with minimal computational complexity. Therefore, we have applied 1D-CNN, the most suitable deep learning approach for 1D data, to accurately predict the centre of the projected laser line in ambient light.

The structure of the 1D-CNN model and its parameter settings are obtained through several experiments. Figure 8.3 presents the structure of 1D-CNN, which incorporates convolution layers, pooling layers and a full connection layer. The number of convolution layers and pooling layers is decided through multiple trials. Each convolution layer comprises the appropriate number and size of convolution kernel to perform one-dimensional convolution operations. To compute the appropriate number of convolution layers, we have trained three models with 3, 5 and 6

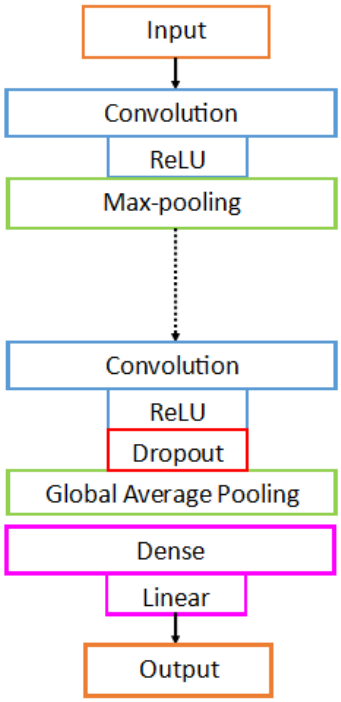


Figure 8.3: The structure of 1D-CNN for laser detection.

convolution layers respectively. The model with three convolution layers use 128, 256 and 512 convolution kernels of size 12×1 in layers 1, 2, and 3 respectively. The second model has five convolution layers. Five convolution layers use 64 kernels of size 4×1 (Conv1), 128 of size 4×1 (Conv2), 256 of size 8×1 (Conv3), 512 of size 8×1 (Conv4), and 768 of size 8×1 (Conv5). The last model with six convolution layers uses 64 convolution kernels of size 2×1 in Conv1, 128 convolution kernels of size 2×1 in Conv2, 256 convolution kernels of size 2×1 in Conv3, 512 convolution kernels of size 4×1 in Conv4, 768 convolution kernels of size 4×1 in Conv5, and 1024 convolution kernels of size 6×1 in Conv6 layer. *Rectified linear unit* (ReLU) is the activation function for the convolution layers. The nonlinear activation function ReLU makes the output neurons with a negative input value zero, which speeds up the calculation and reduces the interdependence among parameters.

The pooling layer is appended to the convolution layer. Except for the last pooling layer, 2 × 2 max-pooling operation is carried out by other pooling layers. The global average pooling layer is appended to the last convolution layer to minimize

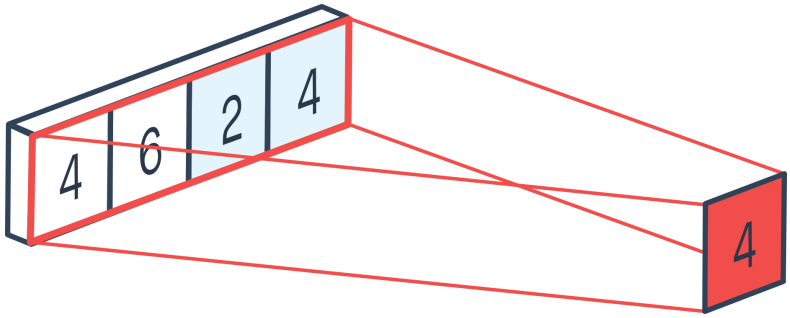


Figure 8.4: 1D global average pooling [212].

over-fitting by reducing the total number of parameters in the model. Global average pooling computes a single average value for each of the input channels (See Fig. 8.4) [212]. A dropout operation with a ratio of 0.4 is performed before the global average pooling layer. The dropout layer solves the problem of over-fitting and enhances the generalization ability of the network. It randomly selects and deletes neurons from the model, to form a random subset of a neuron. The features extracted from the above convolution layers are fed to the dense layer with a linear activation function. The output layer contains one neuron to predict the centre of the laser.

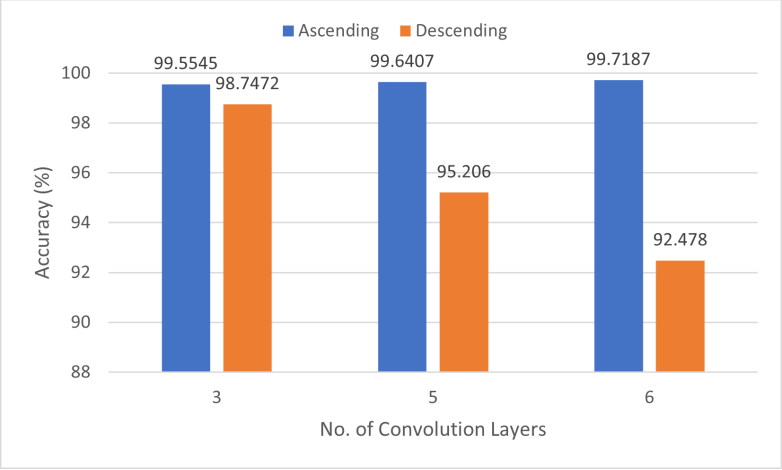


Figure 8.5: The influence of the order of the convolution layers on the accuracy of the models with the different number of convolution layers.

There are several hyper-parameters to train the proposed models, such as learning rate, filter size, batch size and an optimizer. The learning rate was set to 0.0001 for all deep learning models. In addition, *ReduceLROnPlateau* class of Keras framework was applied while training, which reduces the learning rate when the loss function has stopped decreasing. The number and size of the convolution kernel also influence the accuracy of the model. To understand this, we have trained above mentioned three models in ascending and descending order. In ascending order, the number of convolution kernels gradually increases with the increment of the size of the convolution kernel. On the other hand, the number and size of the convolution kernel reduce in descending order. The experimental results shown in Fig. 8.5 depicts that the 1D-CNN model has higher accuracy when the number of convolution kernels increases with the size of the convolution kernel for the different number of layers. Furthermore, the accuracy in ascending order increases with the increase in the number of layers. In contrast, the accuracy of the model is inversely proportional to the number of convolution layers in descending order.

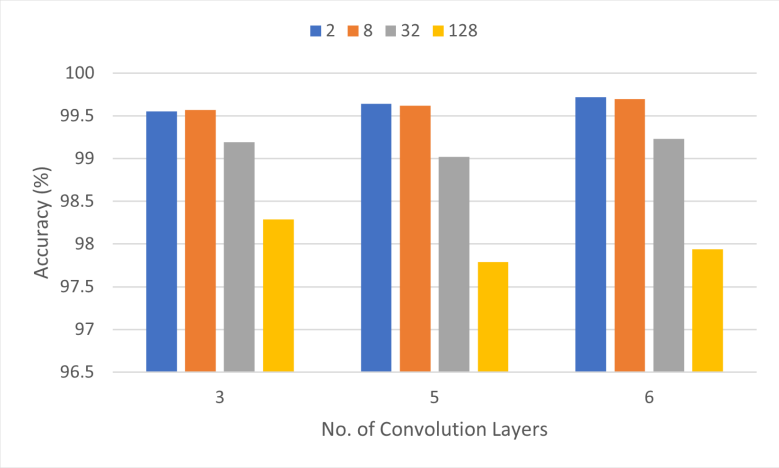


Figure 8.6: The influence of different batch sizes on model accuracy.

Another critical factor to consider while training the model is the batch size. It impacts not only the training speed, but also the accuracy. An extensive batch size expedites the training process, but large memory space is required in a computer. Although the operation speed is slower for small batch sizes in the training process, some noise is produced, which prevents the training process from falling into local optimal [213]. In this research, four different batch sizes (2, 8, 32, 128) are selected to check the effect of different batch sizes on models with the different numbers of

Table 8.2: Architecture of the 1D-CNN for laser detection.

Network Layer Type	Filter Size	No. of Filters	Activation	Output Characteristic
Convolutional	2	64	ReLU	275×64
Max-pooling	2	-	-	137×64
Convolutional	2	64	ReLU	136×128
Max-pooling	2	-	-	68×128
Convolutional	2	64	ReLU	67×256
Max-pooling	2	-	-	33×256
Convolutional	4	64	ReLU	30×512
Max-pooling	2	-	-	15×512
Convolution layer	4	64	ReLU	12×768
Max-pooling	2	-	-	6×768
Convolutional	6	1024	ReLU	1×1024
Dropout	0.4	-	-	1×1024
Global average pooling	-	-	-	1×1024
Dense	-	1	Linear	1×1

convolution layers. The number and size of convolution kernels are in ascending order in the selected models. Figure 8.6 shows the experimental results, which indicate that the accuracy obtained with a batch size of 2 and 8 is nearly the same in each model. The accuracy of more than 99.5% is achieved by all models with the batch size of 2 and 8. The accuracy with the batch size of 32 is also above 99% for all models. The model with six convolution layers performs with the highest accuracy of 99.71 % with batch size 2. With the further increment in batch size, the accuracy decreases for all models. The lowest accuracy of 97.78% is achieved by five layers model for the batch size of 128.

1D-CNN neural network with six convolution layers, where the number and size of convolution kernels are in ascending order, trained with the batch size of 2, is chosen for further comparison with different machine learning models. Table 8.2 shows the specific network structure parameters of this network. This network has been optimized by different optimizers. The performance and outcomes are discussed in the next sub-section.

Long Short-Term Memory

Long short-term memory (LSTM) is a popular *recurrent neural network* (RNN) with sequential data. RNNs are proven to be more accurate with sequential data compared to CNNs. An LSTM model is developed to check the accuracy of the proposed 1D-CNN model for laser detection. The structure of the developed LSTM model with

its network parameters are shown in Fig. 8.7.

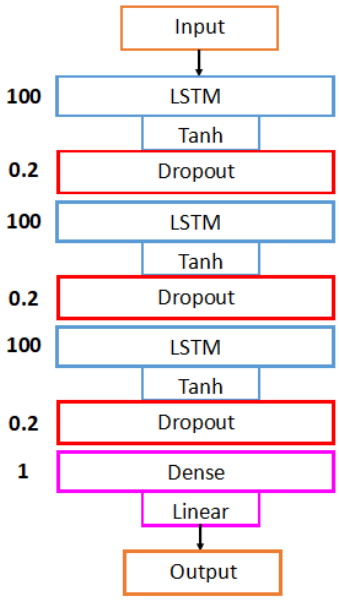


Figure 8.7: The structure of the LSTM network for laser detection.

We have selected different optimizers, such as Adam, Adamax, RMSprop, Nadam and SGD, to optimize the LSTM and 1D-CNN models. The optimized models are tested on test dataset 1. Figures 8.8 and 8.9 represent the loss functions (MAE) of LSTM and 1D-CNN models, while training the models with different optimizers.

As can be observed in Fig. 8.8, the LSTM network does not converge to 0 when optimized with an SGD optimizer. Not only the LSTM network converges faster with the RMSprop optimizer algorithm, but the final loss is also the lowest. Despite being slow to converge, the Adam optimizer was found to be the best optimizer for the 1D-CNN network with the lowest final loss (See Fig. 8.9). The 1D-CNN network optimized with RMSprop converged fastest, in only 26 epochs. In terms of accuracy, 1D-CNN optimized by Adam have shown better performance compared to other optimization algorithms. Additionally, the network with the RMSprop optimization algorithm is proven to be the fastest converging network. Moreover, the training dataset loss function is not much lower than the validation dataset loss function, meaning that network did not over-fit during the training process.

Table 8.3 compares the performance of LSTM and 1D-CNN networks by the obtained MAE, RMSE, MSLE and R^2 results. The RMSE value of the proposed 1D-

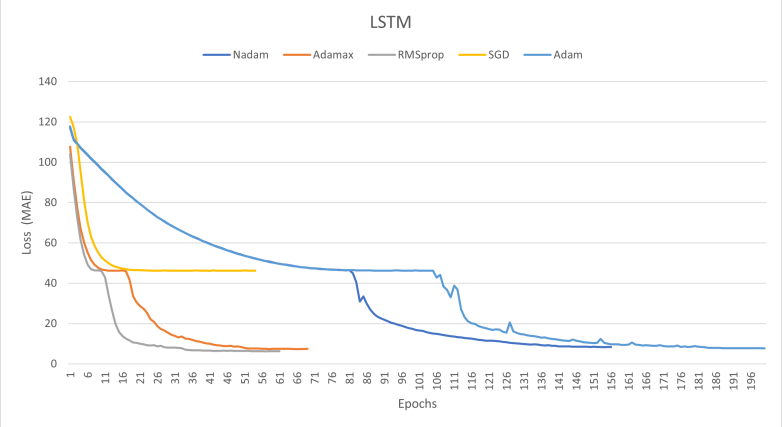


Figure 8.8: Loss functions (MAE) while training the LSTM model.

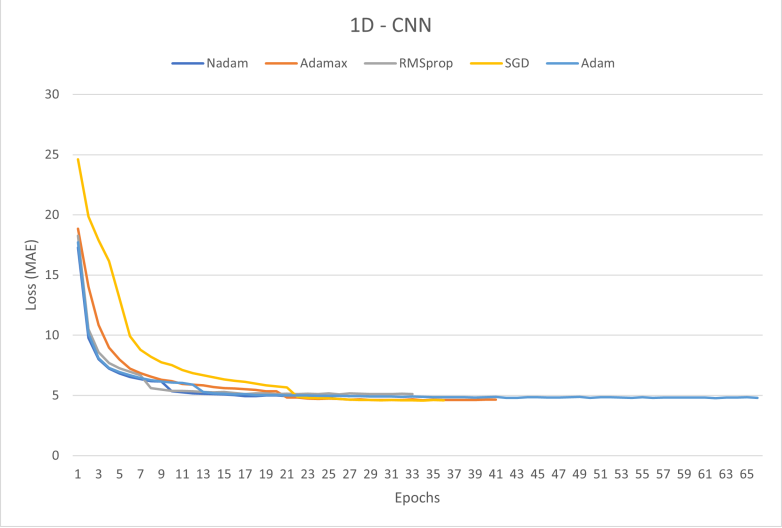


Figure 8.9: Loss functions (MAE) while training the 1D-CNN model.

CNN-Adam model is lower than the other models. Except for the LSTM-RMSprop model, the RMSE value of other LSTM networks is higher than the other models. In the case of the 1D-CNN network, there is a minor difference in the R^2 value of different optimizers, which means the performances of all networks are approximately equal. Figure 8.10 represents the values of RMSE of each optimizer and show that the 1D-CNN-Adam network outperforms other networks. Overall, the 1D-CNN

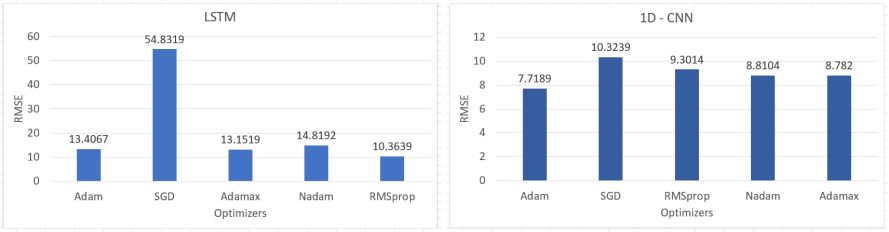
Table 8.3: Comparison of different optimization algorithms.

Optimizer	Evaluation	LSTM	1D-CNN
Adam	MAE (Loss)	6.5943	1.9071
	RMSE	13.4067	7.7189
	MSLE	0.0664	0.0494
	R^2	0.9646	0.9972
Adamax	MAE (Loss)	5.0784	2.437
	RMSE	13.1519	8.782
	MSLE	0.0639	0.0614
	R^2	0.9756	0.996
RMSprop	MAE (Loss)	2.4799	2.4581
	RMSE	10.3639	9.3014
	MSLE	0.067	0.0662
	R^2	0.9908	0.994
Nadam	MAE (Loss)	8.27	2.2368
	RMSE	14.8192	8.782
	MSLE	0.074	0.0558
	R^2	0.9624	0.9955
SGD	MAE (Loss)	47.8285	2.7563
	RMSE	54.8319	10.3239
	MSLE	0.4753	0.0827
	R^2	0.0218	0.9941

network is more accurate compared to the LSTM network for laser detection.

8.3 Traditional Supervised Regressors for Laser Detection

The developed 1D-CNN and LSTM networks are compared with the six different traditional regressors that are based on machine learning and deep learning. The



(a) LSTM. (b) 1D-CNN.

Figure 8.10: Bar chart of RMSE.

six regressor models are MLP, SVM, KNN, DT, RF, and XGBoost. The specific parameters of MLP network models are selected with the same selection method as the parameters of the 1D-CNN and LSTM models. GridSearchCV (10-fold cross-verification parameters) is used to select several parameters that affect the performance of SVM, RF, DT, and KNN models. Experiments are carried out for four test datasets i.e. the data obtained from the four different images of the projected laser line. The parameters of each model are as follows:

1. **Multi layer perceptron model.** The first, second, and third layers are the whole connective layer with 100, 200 and 100 neurons, respectively. The activation function is Relu. The fourth layer is the output layer with one neuron and linear activation function. The learning rate is set to adaptive and the optimizer is Adam. The number of maximum iterations is set to 500. This model optimizes the squared loss using stochastic gradient descent.

2. **Support vector machine model.** Epsilon-support vector regression is applied to train the model. The GridSearchCV (10-fold cross-verification parameters) is adopted. Gaussian kernel (RBF) is selected as the kernel function of SVM. The penalty factor C is determined to be 125.

3. **K-nearest neighbors model.** The GridSearchCV (10-fold cross-verification parameters) is used for the KNN model to determine the best K value of 10. The uniform weights are applied to all neighbors.

4. **Decision trees model.** The mean squared error function is used to measure the quality of a split. It minimizes the L2 loss using the mean of each terminal node [147]. The maximum depth of the tree is chosen to be 200.

5. **Random forest model.** The GridSearchCV (10-fold cross-verification parameters) is adopted. One-hundred decision trees are used to construct the random forest model. The maximum depth of the random forest tree is 50.

6. **Extreme gradient boosting model.** The *gbtree* tree-based booster is applied to the XGBoost model. The learning rate of 0.1 is selected. The maximum depth of the model is 50 in our case.

Table 8.4: Accuracy of eight different models.

Method	Accuracy (%)				Average Accuracy (%)
	Test Data 1	Test Data 2	Test Data 3	Test Data 4	
1D-CNN	99.7187	98.8857	99.6468	98.8885	99.2849
LSTM	99.0784	94.146	99.5527	96.8494	97.4066
MLP	85.8226	62.0334	69.4292	69.4292	71.6786
SVM	99.0175	53.8798	99.0423	91.0539	85.7484
KNN	99.379	72.6727	99.8257	97.5882	92.3317
DT	99.2566	51.1068	86.4157	71.6032	77.0956
RF	99.5366	71.5846	93.3604	87.7297	88.0528
XGBoost	98.8086	76.0858	94.9553	87.3825	89.3081

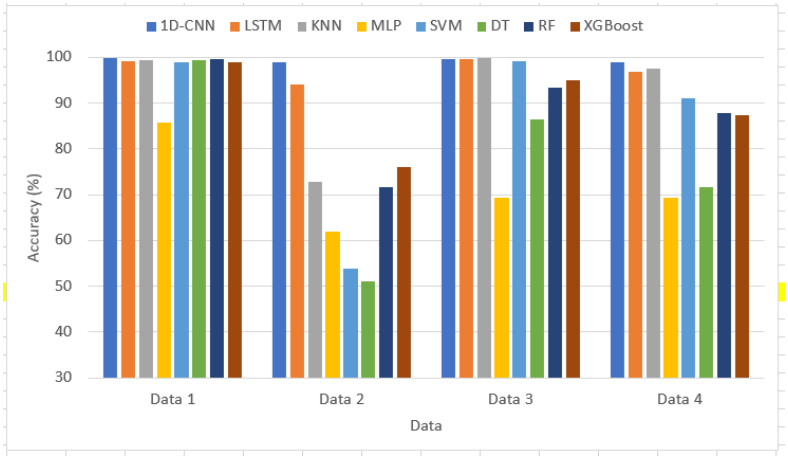


Figure 8.11: Accuracy of different models for different test datasets.

Each model is cross-validated with 10-fold when tested on different test datasets, to better evaluate the performance of the model. Experimental results of various laser detection methods are shown in Table 8.4. Figure 8.11 shows the accuracy of the eight models, including the best 1D-CNN and LSTM models, for the four test datasets. Except for the KNN model, other machine learning based methods perform worse than the 1D-CNN and LSTM models.The leading cause of poor prediction quality for the SVM model is that the SVM model does not perform well in

datasets with many feature points. The SVM model considers the length of the input sequence as the number of feature points, which severely affects the prediction accuracy. RF and DT models are sensitive to noise, which quickly leads to over-fitting.

Compared to other datasets, every model has lower accuracy for test dataset 2 (See Fig. 8.11). DT model has achieved the lowest accuracy of 51.10% for the test dataset 2. The highest accuracy, in the case of test dataset 2, is 98.89%, accomplished by the 1D-CNN model. On the other hand, each model has obtained the highest accuracy for test dataset 1. The reason behind this could be that the train dataset has more images, which are kind of similar to test dataset 1. Additionally, the 1D-CNN model outperforms other models for all test datasets, with the difference between the highest accuracy and the lowest accuracy being only 0.83%. Overall, the deep learning method is superior to the machine learning methods, mainly because machine learning cannot learn non-linear relations in complex intensity distribution datasets. Furthermore, machine learning methods are also unable to interpret sequential data.

8.4 Results and Discussions

Figure 8.12 shows the average accuracy of different models for laser detection. The average accuracy of the 1D-CNN model is 99.28%, which is the highest achieved accuracy. The LSTM network has the second-best performance with an accuracy of 97.41%. Out of all traditional regressors, the average accuracy of KNN network(97.41%) is higher than XGBoost (89.31%), RF (88.05%), SVM (85.75%), and DT (77.10%). Even though MLP is a deep learning based method, the MLP model has performed poorly with an average accuracy of 71.68%.

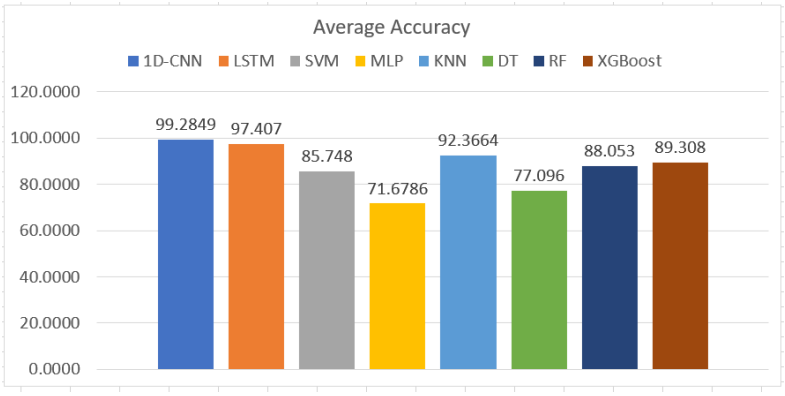
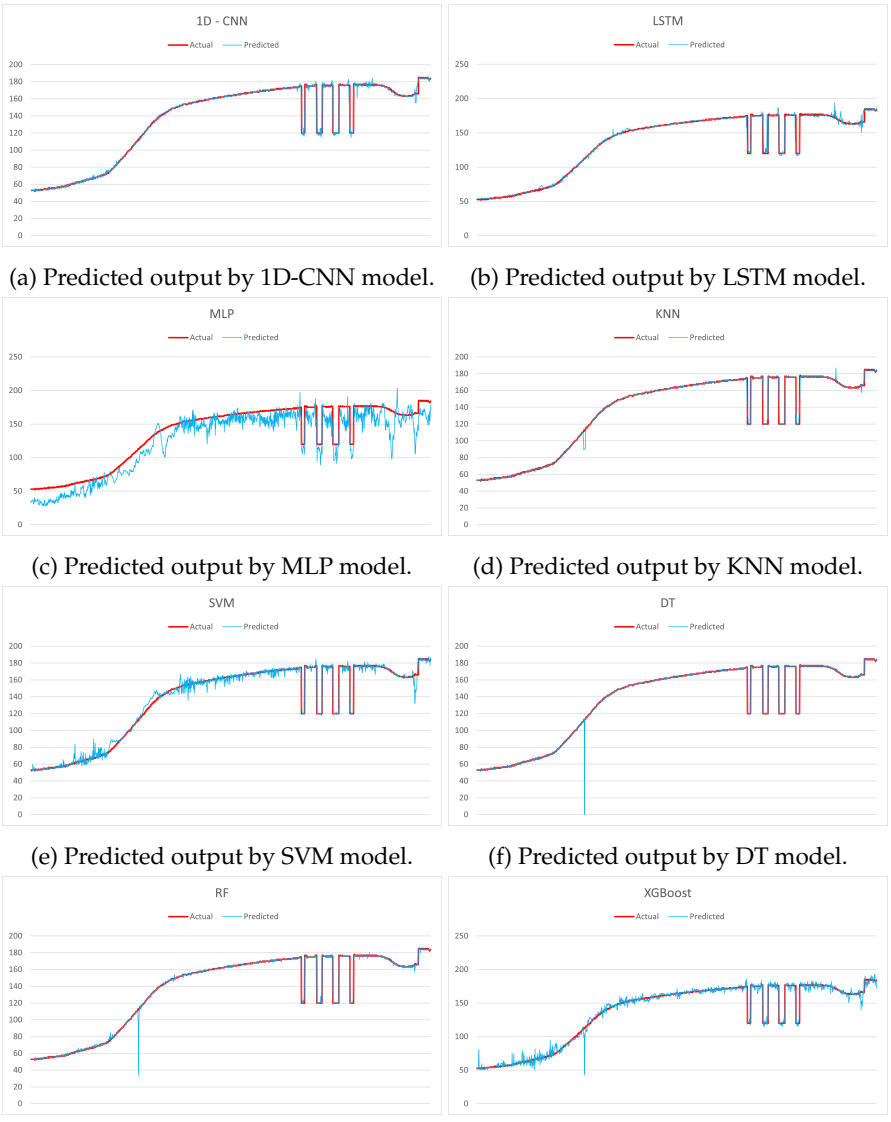


Figure 8.12: Average accuracy of different models.



(a) Predicted output by 1D-CNN model. (b) Predicted output by LSTM model.

(c) Predicted output by MLP model. (d) Predicted output by KNN model.

(e) Predicted output by SVM model. (f) Predicted output by DT model.

(g) Predicted output by RF model. (h) Predicted output by XGBoost model.

Figure 8.13: Predicted outputs by different models.

As the highest accuracy is achieved for test dataset 1, the prediction results of each model for test dataset 1 are shown in Fig. 8.13. The predicted outputs by different models for other datasets are presented in Appendix B. According to the

accuracy, we have selected the 1D-CNN network optimized by Adam optimizer and the LSTM network optimized by the RMSprop optimizer. The MLP model has the worst result compared to other models. The detected laser line obtained by XGBoost and SVM model is not smooth, which can affect the accuracy of 3D measurement. The fluctuation of more than 2-3 pixels is considered inaccurate in the laser detection process. The KNN, DT and RF models show an incorrect detection by one big fluctuation in the dark region. Therefore, the accuracy of these models is affected by the dark region. The results show that the discontinuity of the surface does not affect the accuracy of the 1D-CNN and LSTM models. These models have the best performance even in the presence of holes in the surface. In addition, the developed 1D-CNN model is not affected by the specularity of the surface.

To conclude, the highest average accuracy achieved model 1D-CNN can be combined with our stereo-laser system to develop a hybrid system. Once trained, the model can accurately detect the projected laser line on any shape and material of the surface. Additionally, the data is obtained in ambient light; thus; the network is trained to overcome the difficulties caused by ambient light. The prediction results obtained of all models specified in this chapter, for the exceptional cases mentioned in Chapter 4, are shown in Appendix B. The prediction time of the trained 1D-CNN model is 1-2 seconds. Therefore, this model is a novel, time-efficient, and cost-efficient way to detect the laser line projected by any active vision sensors. In future, we will focus on achieving high accuracy for not only the different datasets but also for the special cases. Thus, we will be able to detect the projected laser line accurately in ambient light.

8.5 Summary

In this chapter, we describe the use of artificial intelligence techniques to detect the laser from its 1D intensity distribution data. DL and ML based different models are tested to predict the output for four test datasets. The datasets are obtained by projecting a laser light on a highly specular wash plate in ambient light. This is an inexpensive laser detection solution compared to commercial products available in the market. From analyzing the results, we can conclude that the 1D-CNN, LSTM and KNN networks can provide even better results if the datasets are obtained in a dark environment. However, by detecting the centre of the laser accurately for the specular surface in ambient light, this research aims to overcome the limitation of traditional and commercial 3D measurement methods. The results of this chapter are summarized as follows:

1. For all test datasets, the 1D-CNN achieves the best accuracy. However, when

compared with the output of stereo-laser system, the fluctuations present in the predicted result of the 1D-CNN affects the accuracy of 3D measurement.

2. Even though the RNNs are more popular than CNNs with the sequential data, we achieved better average accuracy for the 1D-CNN than the LSTM network.

3. The accuracy of 1D-CNN increases, as the number of convolution layers increases, In addition, the 1D-CNN has higher accuracy, when the number of convolution kernels increases with the size of the convolution layers, which means in ascending order.

4. The smaller batch sizes show better performance than the larger batch size in 1D-CNN.

5. Despite being slow in the training process, the 1D-CNN optimized by Adam optimizer has the best results, compared to other optimizers.

6. Only the LSTM network optimized by the RMSprop optimization algorithm has a similar performance to the 1D-CNN model. Other optimizers perform poorly; SGD with the highest RMSE.

7. Out of all the traditional regressors, KNN has shown the best performance with an average accuracy of 92.33%, despite the type of data being sequential.

8. The KNN, DT and RF networks show a considerable fluctuation in the predicted results for the dark region, where the projected laser light merges with the background due to its low intensity.

Chapter 9

Discussion and Conclusions

This chapter discusses the key contribution and achievements of this thesis. The answers to the research questions, posed in Chapter 1, are addressed in detail. Next, the limitations of our stereo-laser system and the developed AI models for laser detection are mentioned in detail. The chapter concludes with an overview of future directions.

9.1 Discussion

We discuss the research questions with their proposed solutions, which satisfies the requirements of Facteon Intelligent Technology Ltd. for quality inspection of products with reflective surfaces in ambient light.

1. How to get accurate 3D dimensions of the product with the reflective surface in ambient light?

   To get accurate 3D measurements of the product with the highly specular surface, we have developed the stereo-laser system, by combining the concept of stereo vision and laser triangulation. Both stereo vision and laser triangulation separately have limitations when it comes to reflective surfaces. However, the stereo-laser system combines both methods. Like the laser triangulation technique, we project a narrow band of laser light on the product, rotating at a constant speed. Instead of one camera, we have used stereo cameras in standard epipolar geometry to capture stereo images of the product in a calibrated environment. Additionally, the proposed algorithm is developed to detect the projected laser line accurately even in the presence of ambient light. The concept of stereo vision is used to perform stereo matching on the detected laser profile in stereo ROIs. The disparity values of the projected laser profile are used to generate the 3D profile in WCS. By merging the 3D profiles of the rotating product, the output 3D model of the product is created for quality control.

2. How to achieve the accuracy of $\pm$ 0.1 mm in the measured dimensions of the product?

   The most critical part of the research is to achieve the accuracy of 0.1 mm in the presence of ambient light. The accuracy of the proposed system depends on many factors, such as the position of the laser, the calibration of stereo cameras, the type of the laser light, the pre-processing of stereo images to nullify the effect of vibrations, and the accuracy of the proposed laser detection algorithm. During this research, we have closely analyzed each factor affecting the accuracy of the system. We need to project the laser on the object in such a way that the laser line does not get diffused or reflected, and it should also not get merged with the background. In this thesis, we have explained the experiments performed using red and blue light lasers. After examining the results, we concluded that the accuracy achieved using the blue light laser is higher, as the blue light laser does not penetrate the surface. We have also performed an accuracy test to check the maximum accuracy that we can achieve by using this system. Results show that the maximum achievable accuracy by the system is 0.005 mm.

3. How to detect palpable defects in the reconstructed 3D model?

   The detection of defects is an essential part of any inspection process. The palpable defects are the limitation of vision-based defects detection techniques. The stereo-laser system is developed to not only perform 3D measurements of the product, but also detects the defects present in the object at the same time. Instead of merging each 3D laser profile to create the 3D model, the 3D profiles are translated along the Z-axis to create the unwrapped model of the product. The unwrapped 3D model of the defective product is compared with the unwrapped model of the non-defective product. The distance map is calculated as a result of the comparison, which shows the defects present in the product. The elevation difference is not detected as a defect with this technique. The experimental results show that the bump and dent of size 0.07 mm and 0.03 mm respectively are detected accurately.

4. How to make the process of 3D reconstruction automatic?

   After detecting the defects, the stereo-laser system decides to either keep the product or reject it based on the size of the defects. Additionally, the system starts the 3D reconstruction process, as soon as it receives the first pair of stereo images. Therefore, the process of the 3D measurement is fully automatic.

5. How to eliminate the effect of the vibrations caused by machinery on accuracy?

It is essential to nullify the effect of the vibrations caused by machinery. Generally, the setup of the stereo-laser system is one single unit. Therefore, the effect of vibrations gets automatically nullified. However, the stereo-laser system has a pre-processing stage, in which the effect of vibrations can be nullified if present. To do so, the stereo images are captured in a steady position in the calibrated environment. The extreme point of the product in the image is taken as a reference point. The laser point is detected for the same row of the extreme point. The distance between these two points is calculated for a steady position. When the system starts the operation of 3D measurement, the reference point, and the distance between the reference point and the laser point is calculated for each stereo image pair. The shift in the extreme point is compensated by adding a border or by cropping a border of the image. The change in the distance is compensated, when we create the 3D laser profile from disparity values of the projected laser line. This way, we try to indemnify the effect of vibrations on the output accuracy.

9.2 Summary of Remarks

This novel work [9,10,159,190] utilizes stereo matching and laser triangulation to inspect the product with the highly specular surface in ambient light. With recent advancement in AI, the concepts of deep learning and machine learning are also combined with the developed system, to create a hybrid system. The stereo-laser system developed during this research is proven to be the only system, which performs 3D measurement of reflective surfaces in ambient light. By doing so, the system overcomes the limitation of traditional and state-of-art shape measurement techniques. Unlike machine vision based defects detection methods mentioned in Chapter 2, the system measures the size of the defect along with the detection of defects.

The primary research contributions of the thesis can be summarized as follows:

- A novel system is developed for the 3D inspection of specular surfaces in ambient light. This system is simple, automatic, feasible and cost-efficient solution for 3D measurement in ambient light. This system achieves the accuracy of 0.005 mm.

- A comprehensive survey related to the traditional methods, state-of-art techniques, and commercial solutions for shape measurement is provided in this research. The above-mentioned techniques are extensively compared with the developed system, to get a better understanding of the limitations of these techniques.

- The special cases of intensity distribution are mentioned in this research, which exhibits the effect of ambient light. The proposed system is capable of detecting the laser line accurately even in the presence of ambient light. This is the main contribution of this research.
- A novel defects detection algorithm is proposed, which can also measure the size of the defects. The palpable defects of the minimum size of 0.03 mm can be detected accurately with this system.
- The effect of vibrations caused by machinery is also nullified by this system.
- A simple technique is suggested to measure the height and the diameter of the cylindrical objects from a single image.
- The basics of supervised learning for regression analysis are learned to utilize later for laser detection.
- A novel idea for laser detection is proposed, which uses the 1D intensity distribution of the projected laser, instead of the image data, as an input.
- The 1D data to train the DL and ML models is created and labelled manually from the intensity distribution of the laser.
- Different DL and ML models are compared conscientiously to evaluate the performance of each model. The performance of each model is also examined for special cases of intensity distributions caused by ambient light. The 6 layer 1D-CNN model optimized with Adam optimizer has the overall best average accuracy for different test datasets. On the other hand, the KNN model has the best performance in special cases of intensity distributions.

9.3 Limitations of the Stereo-Laser System

This section is divided into two parts. Besides the remarks that have been made hitherto, we found four limitations of the stereo-laser system during our research, which are described in detail below. Later, we explain the limitations of AI, when it comes to laser detection in ambient light.

Limitation 1

The first limitation is that we cannot reconstruct a region with highlights if the highlights are visible in both left and right camera images. Figure 9.1 shows one such

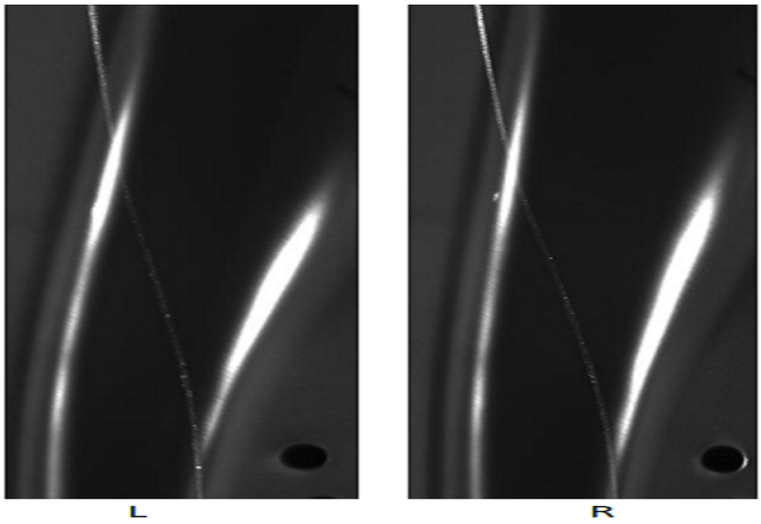


Figure 9.1: Stereo image pair for limitation 1.

case. Here, the projected laser line is merged with the highlight in both camera images. Therefore, we cannot detect the laser line in this region with highlight. If the defect is present in this region, it will not be detected. To avoid this, we need to block light from specific directions. Another solution is to use blue light filters for stereo cameras. The quality of stereo images will be improved, and the effect of ambient light will be reduced by using the blue light filters with stereo cameras.

Limitation 2

The second limitation occurs due to the setup of the experiment. In the case of the wash plate, the stereo cameras capture images from the top. Figure 9.2 shows an example of the ROI image, where the stereo cameras are looking at the product from the top. The projected laser line in the circled region of the image is not visible by the stereo cameras. Therefore, we cannot reconstruct that border of the product with this method. Moreover, we will not be able to detect any defects present in that region. This limitation is only present in the case of the wash plate. In the case of a drum, we can accurately reconstruct the 3D model of the whole product.

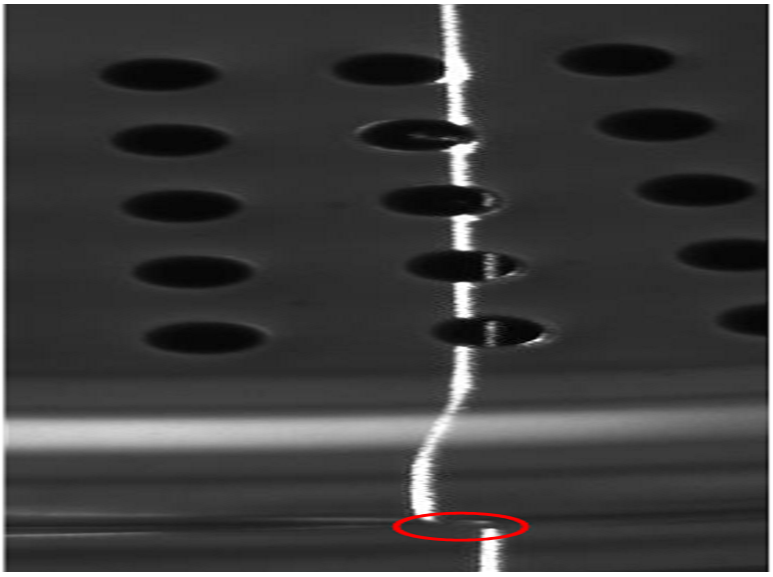


Figure 9.2: An image for limitation 2.

Limitation 3

The third limitation of this technique is concerned with scratch detection. There are two scratches on the wash plate (See Fig. 9.3). As we can see, scratch 1 is vertical; thus, when a vertical laser line is projected on the product, it merges with the scratch 1. Any displacement in the laser line is not visible, as the width of the laser line is more significant than the width of the scratch [9]. Therefore, it is challenging to detect scratch 1 when the laser line is projected directly on scratch 1. However, scratch 2 is detectable with this method [9].

Limitation 4

The last limitation is a surprising fact that we found out during our research. If we write on the product using a black marker, we will not be able to measure that region accurately. As we can see in Fig. 9.4, the projected laser line is absorbed by the black paint of a marker. Therefore, our algorithm considers it a drilled hole because the gray value for that region is below the threshold value, and the laser light is invisible. This is the last limitation of this system.

Figure 9.3: An object with scratches.

Figure 9.4: An image for limitation 4.

Limitations of Artificial Intelligence Methods

There are few limitations of supervised learning models when they are used for laser detection in ambient light, which are mentioned below:

- The accuracy requirement of the inspection process in the production line is 0.1 mm. The performance of DL and ML models in ambient light shows that the models require more train data for the special cases of intensity distribution to get desired output accuracy.
- The trained DL and ML models are not generalized enough, to use the models trained with the laser data for the wash plate data, to test on the data obtained from projecting the laser on a drum.
- The limitations of the stereo-laser system are also the limitation of AI models because the models are trained on the intensity distribution data of the projected laser line.
- As the width of the ROI is not constant for different products, we need to train the models with the train datasets corresponding to the products.

9.4 Future Directions

In future, we are planning to use blue light filters for stereo cameras. This may help in the case where the highlight is merged with the projected laser line. It may also help with the first limitation of the stereo-laser system. Additionally, we can improve the time factor for this method by installing multiple units of stereo cameras and a laser. Each unit will reconstruct some part of the product. Also, the Halcon software's processing time for one scan is higher than other open-source coding platforms. We can make a significant difference in time by switching to another platform. Another way to improve the processing time is by combining the AI approach of laser detection with the stereo-laser system. For this, the DL and ML models need to be trained with more data, especially for the special cases of intensity distribution, as the motive of this research is to perform 3D measurements of highly specular surfaces in ambient light.

Appendix A

SP1 Stereo Vision System for 3D Reconstruction of Reflective Surfaces

In the starting phase of this research, we tried the traditional 3D measurement techniques to check their effectiveness when it comes to reflective surfaces. This chapter focuses on the concept of stereo vision for the 3D reconstruction of the reflective surface. As stated before, stereo vision is proven to be the most accurate and simple method to perform a 3D measurement in ambient light. However, like every other traditional method, shiny surfaces are also a limitation of the stereo vision method. To prove this, we carried out few experiments with the SP1 stereo vision system. This system is a product of Nerian Vision Technologies, which performs stereo matching by applying the idea of semi-global matching. We have reconstructed a washing machine drum and a ceramic artifact with the SP1 system. The chapter starts with an explanation of the setup of the system, followed by the calibration process of stereo cameras, and the experimental results are shown at the end.

A.1 SP1 Stereo Vision System

Nerian Vision Technologies developed the stand-alone SP1 stereo vision system to perform stereo matching in real-time. This system has an integrated *field programmable gate array* (FPGA), which makes it capable of processing stereo images from two industrial cameras in real-time. Alternatively, the SP1 can also process the stereo image data transmitted from a client computer. A dense disparity map is produced as an output, which can be described as a mapping of image points from the left camera image to corresponding image points in the right camera image. The 3D location of the corresponding scene points can be reconstructed from the disparity map by using the disparity-to-depth matrix, which is generated at the time of the calibration of stereo cameras. As stated above, the system uses a *semi-global matching* (SGM) algorithm to perform stereo matching in real-time [214]. SGM algorithm has been developed for accurate, pixel-wise matching at low run time [215]. It combines the concepts of global and local stereo methods [216].

Setup

The first step of the experiment is to set up the SP1 system. As illustrated in Fig. A.1, the *universal serial bus* USB stereo cameras are connected with the SP1 through a USB hub, which is mandatory because the SP1 only features one USB port. The synchronization cable is used to provide the trigger signals to both cameras because only the frames with a matching timestamp will be processed by the SP1. A client computer is connected to the SP1's Ethernet port, to receive and display processing results. The Sp1 requires the network to support data rates of at least 25 MB/s without packet drops, to transmit high bandwidth data to the client computer. Moreover, NVCom application is used to receive and display disparity maps from the SP1 [214].

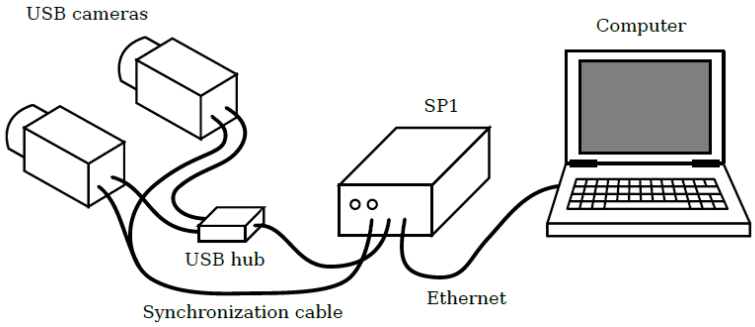


Figure A.1: Setup of the SP1 stereo vision system [214].

Calibration of SP1 Stereo Vision System

The SP1 stereo vision system is configured through a web interface to calibrate the connected stereo cameras. We have used a pair of 1.3 *megapixels* (MP) Point gray Chameleon 3 cameras. As the concept of stereo vision resembles the functionality of our eyes, the stereo cameras have to be in standard epipolar geometry to perform stereo matching. In this arrangement, stereo cameras with identical focal lengths are mounted on a plane with a displacement that is perpendicular to the cameras' optical axis [191]. The distance between both cameras is referred to as baseline distance. The depth resolution at high distances improves with a large baseline distance. On the other hand, a small baseline distance allows for the observation of close objects [214]. Figure A.2 shows the setup of the stereo cameras in standard epipolar geometry. We have chosen the baseline distance of 15 cm for this experiment. Moreover, the stereo cameras capture the stereo images of 640 $\times$480 pixels resolution at

the speed of 15 *frames per second* (fps).

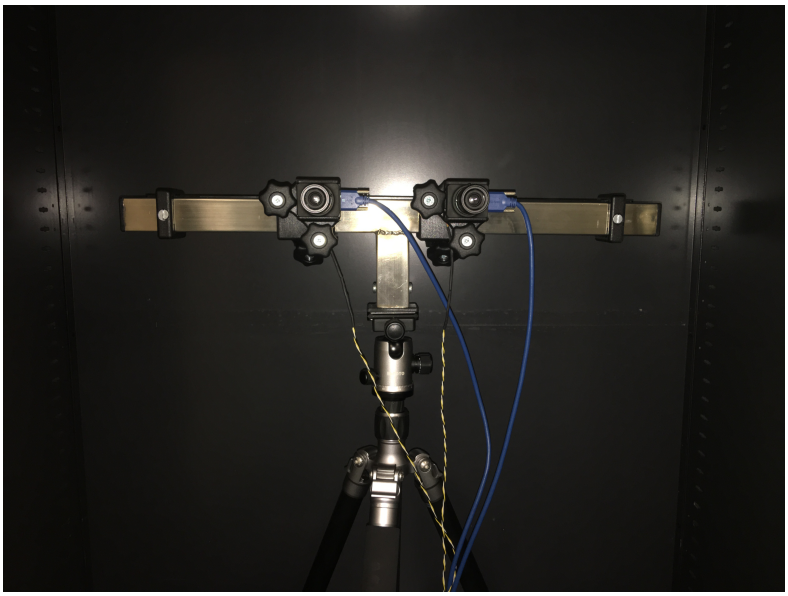


Figure A.2: Stereo cameras in standard epipolar geometry

Intrinsic and extrinsic camera parameters of a given camera configuration are specified by the camera calibration process. The parameters are required to rectify stereo images, and also to create a disparity map. The (effective) focal length, dimensions of the sensor matrix, sensor cell size, radial distortion parameters, coordinates of the principal point, or the scaling factor are some of the intrinsic parameters, calculated as a result of the calibration process. "Extrinsic parameters are those of the applied affine transforms for identifying poses (i.e. location and direction) of cameras in a world coordinate system [191]." Figure A.3 shows an image of the calibration board supported by the SP1 system. The pattern of an asymmetric grid of black circles on a white background is printed on a flat panel to create the calibration board [214]. The stereo cameras capture multiple image pairs of a calibration plate at several different positions. As a result of the calibration process, intrinsic and extrinsic parameters are calculated by the SP1 stereo vision system.

A.2 Experiments and Results

To check the concept of stereo vision with reflective surfaces, we have captured stereo images of the washing machine drum made of steel, and of an artifact made

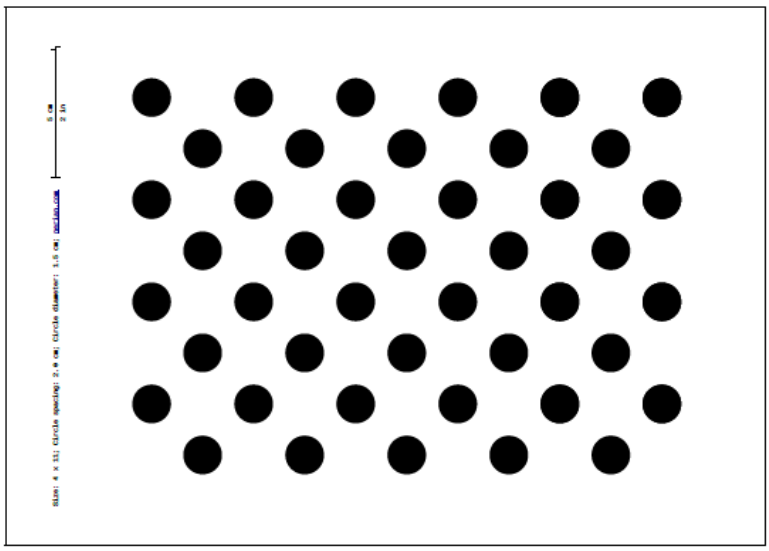
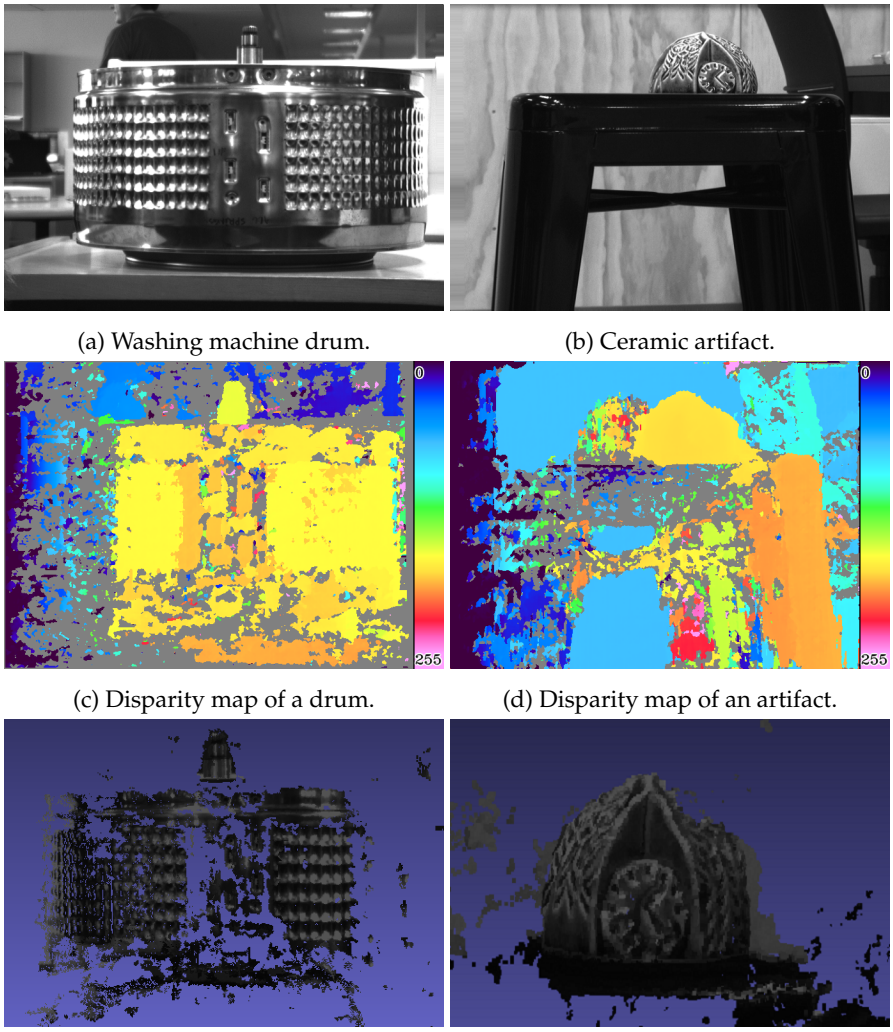


Figure A.3: Calibration plate [214].

of ceramic, in ambient light. The stereo images are rectified before the stereo matching process. Therefore, the corresponding pixels in both images should only differ in their horizontal coordinates. The disparity map thus only encodes a horizontal coordinate difference. The SP1 produces a disparity map from the perspective of the left camera as a result of stereo matching. The obtained disparity maps for the drum and the ceramic artifact are shown in Fig. A.4c and A.4d. The SP1 can also generate a 3D point cloud of the product from the obtained disparity map (See Fig. A.4e, A.4f).

The 3D point cloud of both the products shows unsatisfactory results. The poor quality of the 3D point cloud of the drum could be due to the less texture on the surface of the drum. Additionally, the reflection of the background in the drum also affects the accuracy of the system. In the case of the ceramic artifact, the top-right portion is not reconstructed, as it appears like a bright shiny surface in the input image pair (See Fig. A.4b). The other parts of the artifact, which are not illuminated by the ambient light, are reconstructed accurately. Moreover, the presence of texture on the artifact also improves the accuracy of the reconstructed product.



(a) Washing machine drum. (b) Ceramic artifact.

(c) Disparity map of a drum. (d) Disparity map of an artifact.

(e) 3D reconstructed image of a drum. (f) 3D reconstructed image of an artifact.

Figure A.4: Experimental results.

A.3 Summary

To summarize, it is difficult to reconstruct the reflective surface in ambient light using the concept of stereo matching. In addition, the surface with no texture adds to the difficulties for the SP1 stereo vision system. Moreover, only the front part

of the product is reconstructed at a time. We need to merge that with the reconstructed back part to cover the whole product. However, the poor quality of the reconstructed 3D point cloud makes the process of merging two point clouds strenuous. Furthermore, the SP1 stereo vision system might perform well with textured shiny surfaces in a dark environment. However, we can not use the concept of stereo matching to perform 3D measurements of reflective surfaces in ambient light.

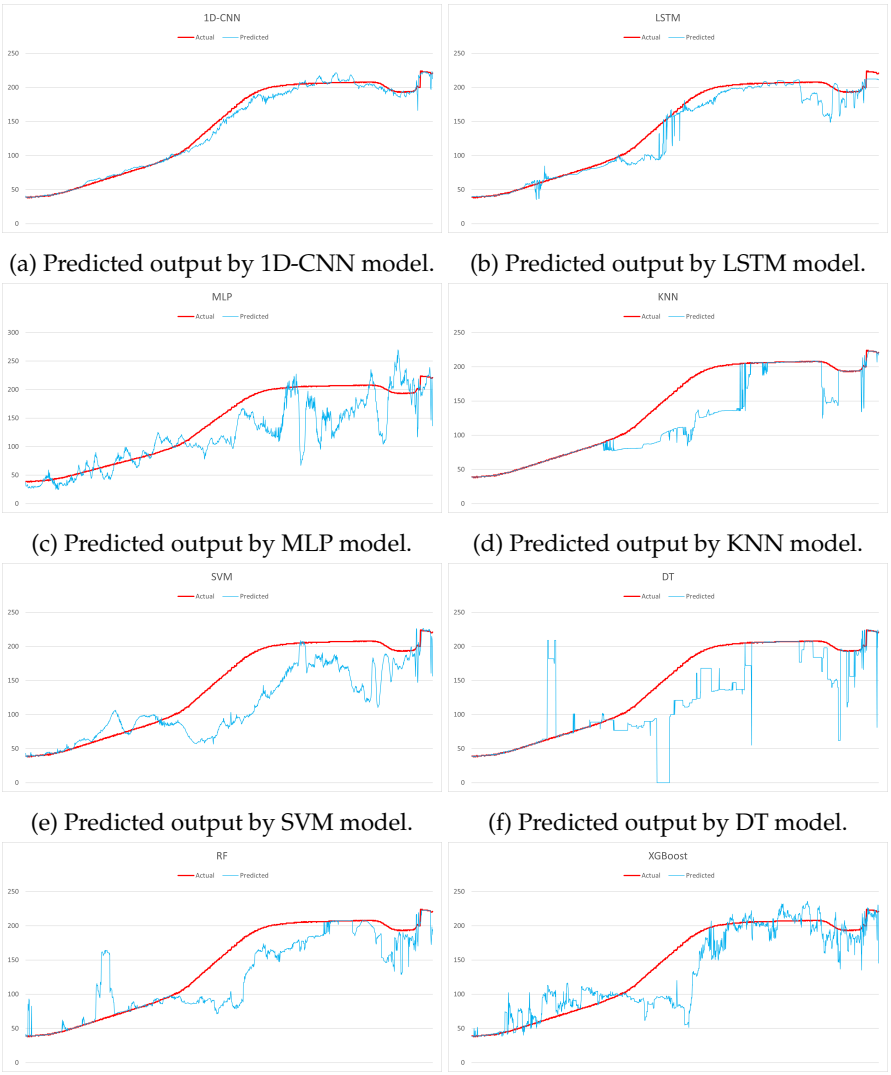
Appendix B

Laser Detection in Ambient Light: Special Cases

This chapter focuses on the special cases of intensity distribution, which occur when the laser line is projected on the highly specular surface in the ambient light. The specularity of the surface causes highlights in ambient light. Additionally, the projected laser line scatters and creates a pseudo-light-stripe in ambient light. The pseudo-light-stripes have a similar intensity distribution to the actual projected laser line. Therefore, the detection of laser in ambient light is a critical problem in computer vision. The algorithm of our stereo-laser system accurately detects the laser for the special cases specified in Chapter 4. The effectiveness of AI models for laser detection in ambient light is tested in this chapter. This chapter starts with the prediction results of AI models for the test datasets 2, 3 and 4. The comprehensive analysis of AI models specified in Chapter 8 for exceptional cases of intensity distribution is provided in this chapter.

B.1 Prediction Results of AI Models for Different Test Datasets

The results predicted by AI models for test dataset 1 are shown in Chapter 8. The AI models are further tested with test datasets 2,3, and 4 to evaluate the performance of the models. Except for the 1D-CNN model, the other model's performance is unsatisfactory for test dataset 2 (See Fig. B.1). On the other hand, the 1D-CNN, LSTM, KNN and SVM models produce adequate results for test dataset 3. For test dataset 4, even the 1D-CNN network shows some fluctuations, which are considered inaccurate predictions. The performance of all other models, except for KNN, is substandard. Moreover, the MLP model cannot be used in real-world applications, due to its consistent poor performance. To summarize, only the 1D-CNN and KNN models show satisfactory results for laser detection on the highly specular surface.



(a) Predicted output by 1D-CNN model.

(b) Predicted output by LSTM model.

(c) Predicted output by MLP model.

(d) Predicted output by KNN model.

(e) Predicted output by SVM model.

(f) Predicted output by DT model.

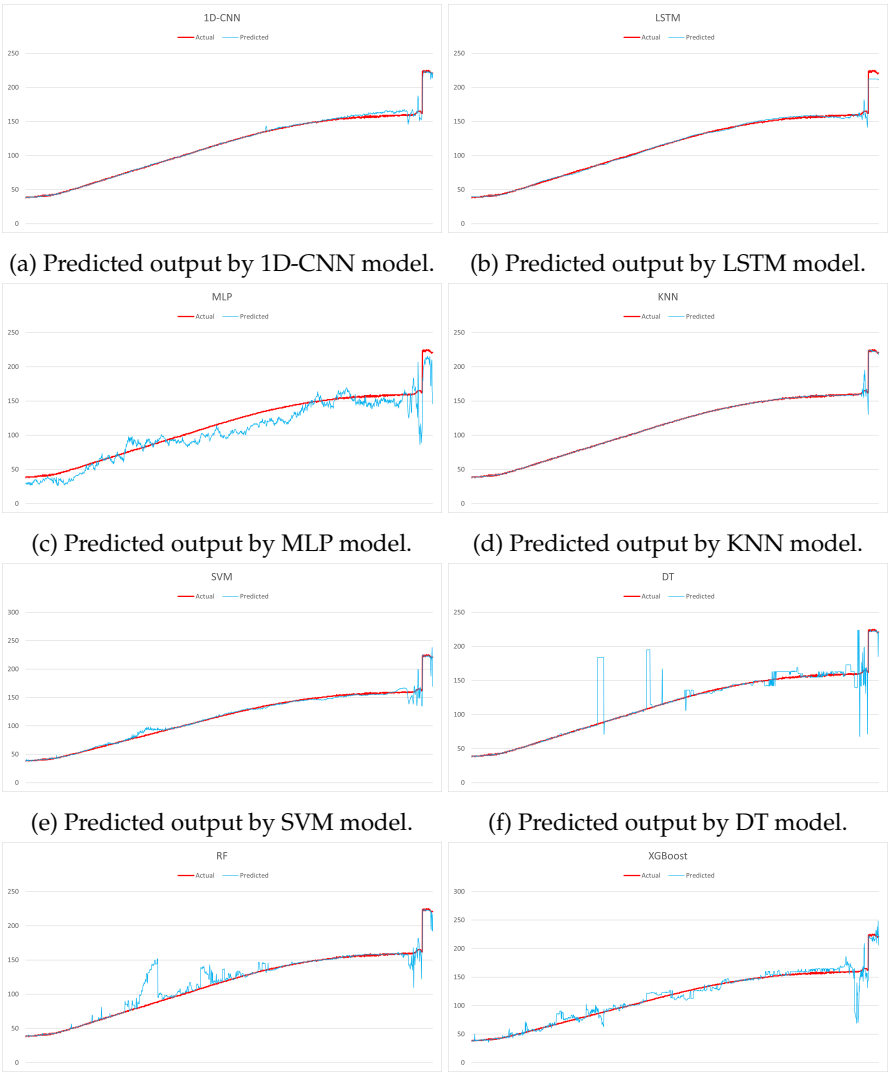
(g) Predicted output by RF model.

(h) Predicted output by XGBoost model.

Figure B.1: Predicted outputs by different models for test dataset 2.

B.2 Performance of AI Models in Special Cases

The ideal intensity distribution of the laser light is a bell-shaped curve. However, the shape of the object can sometimes change the intensity distribution of the laser.



(a) Predicted output by 1D-CNN model. (b) Predicted output by LSTM model.

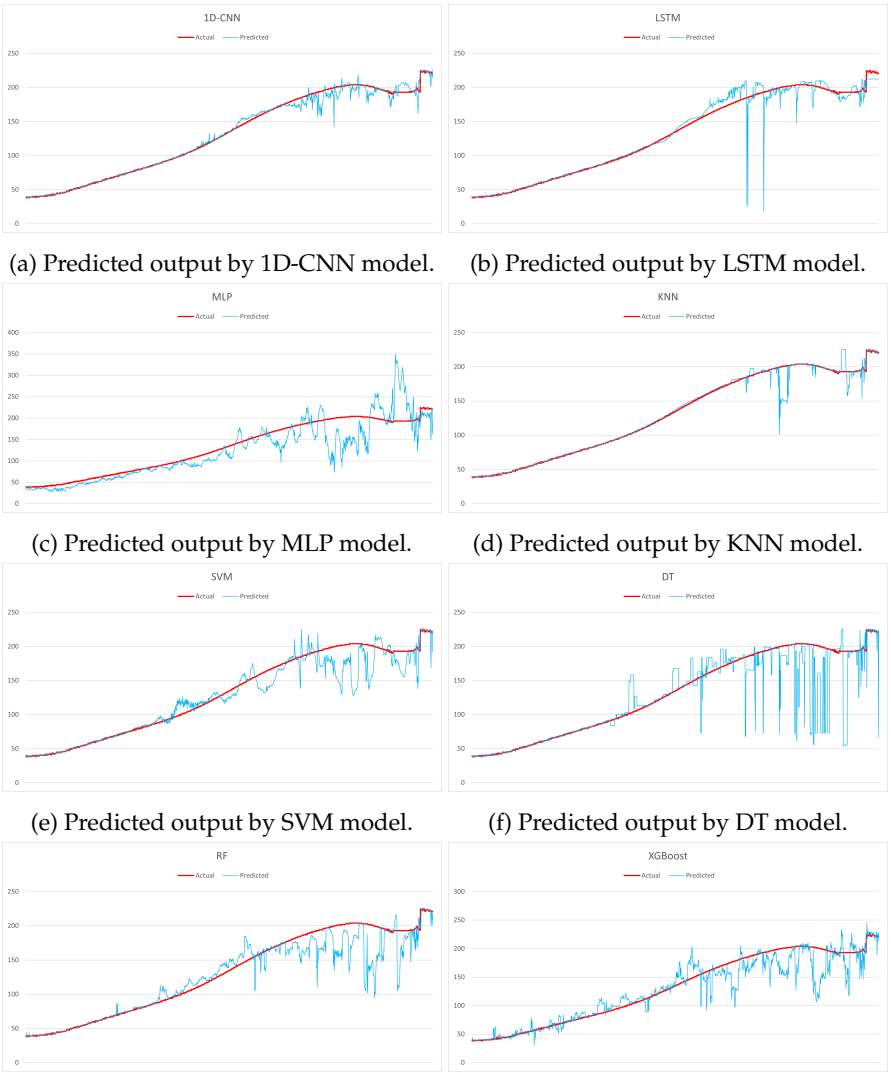
(c) Predicted output by MLP model. (d) Predicted output by KNN model.

(e) Predicted output by SVM model. (f) Predicted output by DT model.

(g) Predicted output by RF model. (h) Predicted output by XGBoost model.

Figure B.2: Predicted outputs by different models for test dataset 3.

In addition, the camera also captures the reflection of the background, which means the intensity distribution is also affected by the color of the background objects. We found a total of six exceptional cases of intensity distribution when the laser light is projected on a shiny surface in ambient light. To avoid the problems caused by



(a) Predicted output by 1D-CNN model.

(b) Predicted output by LSTM model.

(c) Predicted output by MLP model.

(d) Predicted output by KNN model.

(e) Predicted output by SVM model.

(f) Predicted output by DT model.

(g) Predicted output by RF model.

(h) Predicted output by XGBoost model.

Figure B.3: Predicted outputs by different models for test dataset 4.

reflective surface in ambient light, generally, commercial products perform 3D measurement tasks in a dark environment. We have developed the stereo-laser system to overcome the limitation, which considers all exceptional cases while detecting the laser. The AI models specified in Chapter 8 are also trained with the intensity

distribution of laser light in ambient light. In this section, we cover the test of the AI models (1D-CNN, LSTM, SVM, KNN, RF, DT, MLP, XGBoost) for six special cases of the intensity distribution.

Figure B.4 shows the ideal case of the intensity distribution. The ground truth shows the output of the stereo-laser system. The intensity distribution of each special case is fed into the AI network as an input sequence. The predicted output value of the centre of the laser is obtained as a result. Except for the MLP model, other models have accurately detected the laser in the ideal case.

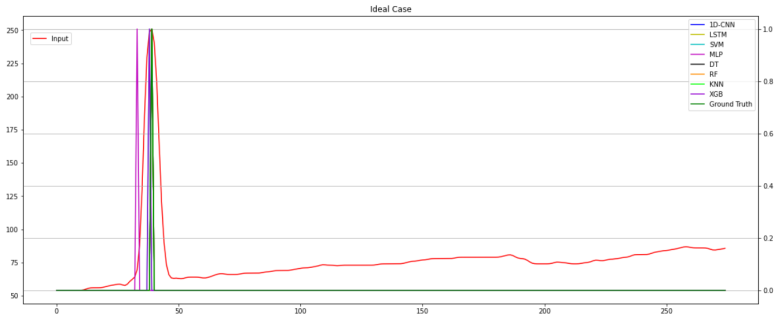


Figure B.4: Ideal case.

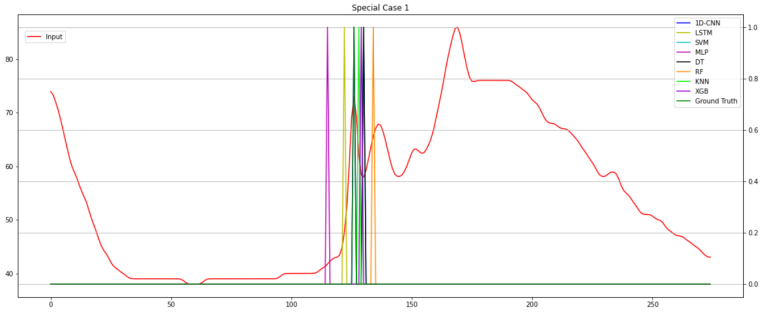


Figure B.5: Special case 1.

Special case 1 occurs when a highlight is present in the ROI near the projected laser line. The predicted value by the KNN model is closest to the actual value (See Fig. B.5). The output of LSTM, MLP and RF models are considered inaccurate. However, it is essential to note that the predicted results are near to the actual projected laser. Therefore, the models are accurate enough to not predict the high intensity value of highlight as laser light.

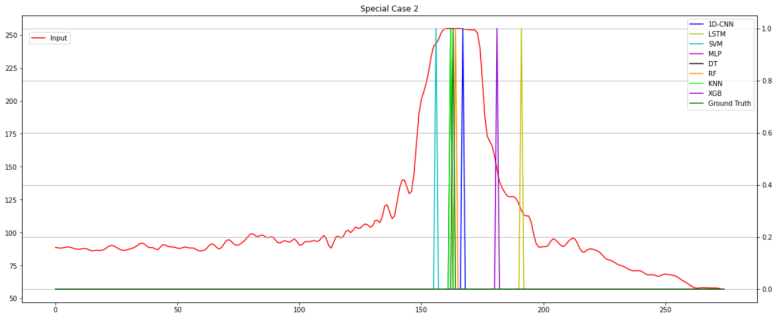


Figure B.6: Special case 2.

In special case 2, the highlight is merged with the projected laser line, it creates a broad bell-shaped curve as shown in Fig. B.6. The KNN, DT and RF networks predict the accurate results in special case 2. On the other hand, the predicted results by other models are relatively far from the actual value. In addition, the MLP model predicts the value which is greater than the length of the sequence. Any predicted values, which are less than zero or greater than the maximum length of the input sequence are converted to 0. The results for special case 3 are displayed in Fig. B.7. Special case 3 depicts the absence of a projected laser line, which happens when the product contains holes. The results show that all models work accurately in the case of the discontinuous surface.

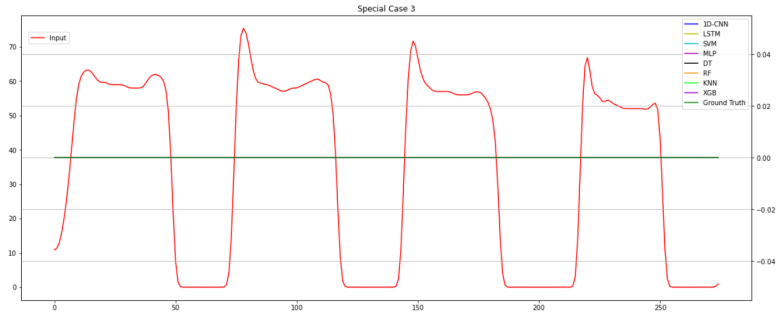


Figure B.7: Special case 3.

Special case 4 presents the jump in the height of the product. The projected laser line gets shifted to either right or left by a considerable distance, compared to the previously detected location of the laser. Figure B.8 shows that the KNN and DT models have the perfect prediction in special case 4. In contrast, the XGBoost model

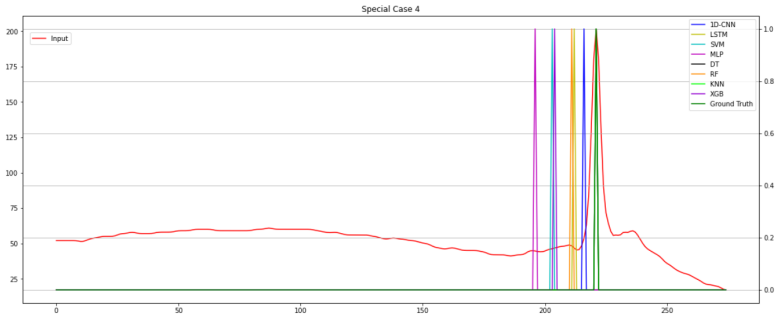


Figure B.8: Special case 4.

has performed poorly in this case.

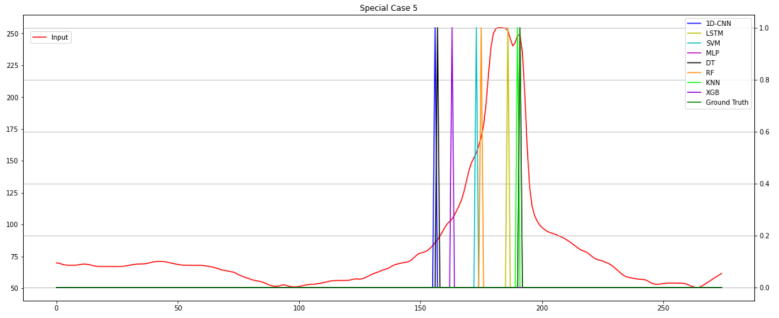


Figure B.9: Special case 5.

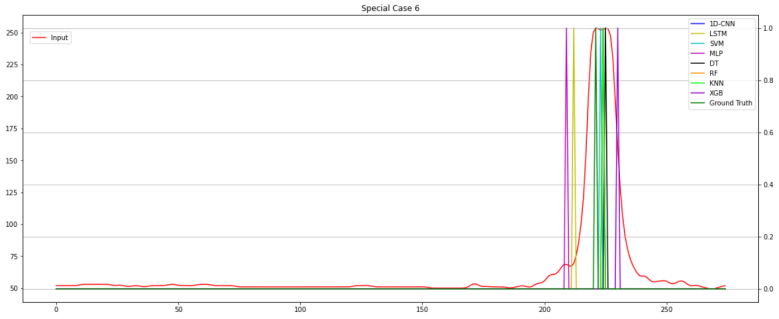


Figure B.10: Special case 6.

Special case 5 is almost similar to the case 3. The projected laser line is slightly

Table B.1: Accuracy of eight different models in special cases.

Method	Predicted Values							Average Accuracy (%)
	Ideal Case	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	
1D-CNN	38	126	167	0	216	156	224	97.28
LSTM	39	122	191	0	212	186	212	97.79
MLP	33	115	0	0	196	191	209	40.48
SVM	38	130	156	0	203	173	223	98.41
KNN	39	128	162	0	221	190	225	99.96
DT	39	130	162	0	221	157	225	97.43
RF	39	134	164	0	211	175	224	98.83
XGBoost	38	129	181	0	204	163	230	96.70

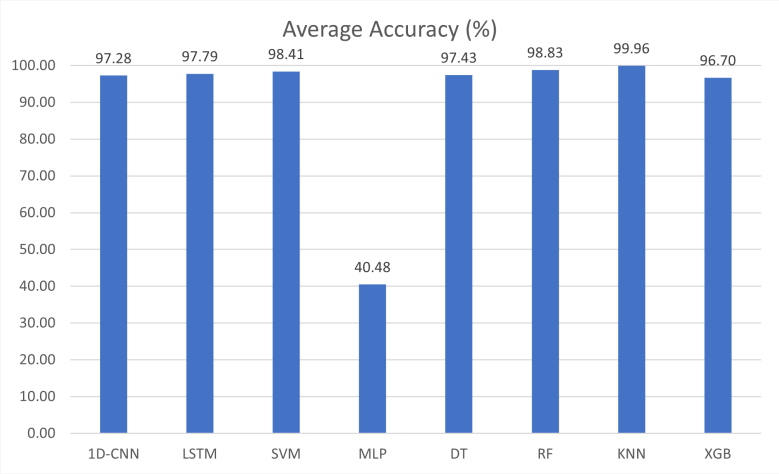


Figure B.11: Average accuracy of different models in special cases.

merged with the highlight. As we can see in Fig. B.9, the peak of the highlight is closer to the peak of the laser. Where the LSTM model detects the peak of the highlight as a laser, the KNN model accurately detects the peak of the laser light. Instead of a sharp bell-shaped curve, if the intensity distribution of the laser is similar to Fig. B.10, an average of the locations of the two peaks is considered the centre of the laser. The predicted results of KNN, SVM, 1D-CNN and DT models are considered accurate in this case. Table B.1 summarizes the performance of different models for laser detection in ambient light. The KNN model achieves the highest average accuracy of 99.96%, followed by the RF and SVM models with 98.83% and 98.41% (See Fig. B.11). The lowest achieved accuracy is 40.48% by the MLP model. Other models

have also obtained an average accuracy of more than 96%.

B.3 Summary

To culminate, the KNN model is generalized enough to make accurate predictions for special cases of the intensity distribution. Even though the 1D-CNN has shown overall better average accuracy compared to other models for different test datasets, the average accuracy of 1D-CNN for special cases is unsatisfactory. The accuracy of 1D-CNN can be improved by feeding the network with more examples of special cases. However, the network will take longer to train with more data. The overall performance of all networks can be enhanced by training them with the diverse dataset, which covers the difficulties added by shapes, materials, the color of the background, and ambient light. From this research, we can conclude that AI models can be introduced to detect the projected laser light in real-world applications.

****

Bibliography

[1] D. E. Knuth, T. L. Larrabee, P. M. Roberts: Mathematical Writing, Mathematical Association of America, ISBN: 088385063X, 1996.

[2] F. Campbell. "Inspection of metals: understanding the basics", ASM International, 2013, see https://www.asminternational.org/documents/10192/22833166/05372_Sample_BuyNow.pdf/7537be39-e632-4361-8a7c-d958b5dfe103#:~:text=methods\%20are\%20visual\%2C\%20microscopy\%2C\%20liquid,the\%20remainder\%20of\%20this\%20book (accessed November 13, 2020) .

[3] "Automated quality control & inspection", Facteon Intelligent Technology Ltd., see https://facteon.global/products-services/machinery-factory-automation-solutions/automated-quality-control-and-inspection/ (accessed November 13, 2020).

[4] "High Quality Appliance Drum Line", Facteon Intelligent technology Ltd., see https://facteon.global/products-and-services/equipment/appliance-drum-lines/ (accessed November 13, 2018).

[5] E. Asoudegi, and Z. Pan. "Computer vision for quality control in automated manufacturing systems." Computers & industrial engineering, 21(1-4), 141-145, 1991.

[6] "Machine Vision", see https://en.wikipedia.org/wiki/Machine_vision (accessed March 13, 2019).

[7] "We build factories for the internet age", Facteon Intelligent Technology Ltd., see https://www.facteon.global/ (accessed November 13, 2020).

[8] "Flexible & Intelligent manufacturing solutions", Facteon Intelligent technology Ltd., see https://facteon.global/products-andservices/equipment/ (accessed July 11, 2020).

[9] A. Dawda and M. Nguyen. "Defects Detection in Highly Specular Surface using a Combination of Stereo and Laser Reconstruction." In 2020 International Conference on Image and Vision Computing New Zealand (IVCNZ), IEEE, December 2020.

[10] A. Dawda, M. Nguyen, and R. Klette. "Accurate 3D Measurement of Highly Specular Surface using Laser and Stereo Reconstruction." In 2019 International Conference on Image and Vision Computing New Zealand (IVCNZ) (pp. 1-6), IEEE, December 2019.

[11] "INSPECTION PRODUCTION AND OPERATIONS MANAGEMENT", Wisdom IT Services India Pvt. Ltd., see https://www.wisdomjobs.com/e-university/production-and-operations-management-tutorial-295/inspection-9633.html (accessed June 13, 2019).

[12] R. Morgan. "What Are the Four Ways in Which Quality Can Affect a Company?", Hearst, see https://smallbusiness.chron.com/four-ways-quality-can-affect-company-21930.html (accessed March 15, 2021).

[13] R. Brown. "Where is Artificial Intelligence Used Today?," Medium, https://becominghuman.ai/where-is-artificial-intelligence-used-today\\-3fd076d15b68 (accessed November 13, 2020).

[14] "How the cost of poor quality can affect a business", Bit Service Company, https://www.bitservice.ca/blog/how-the-cost-of-poor-quality-can-affect-a-business (accessed March 15, 2021).

[15] J. DeFeo. "The Tip of the Iceberg", Quality Progress Vol. 34, No. 5 (May 2001): 31-32, see http://asq.org/data/subscriptions/qp/2001/0501/qp0501defeo.pdf. (accessed March 15, 2021).

[16] G. Freeman. "How Poor Quality is Costing Your Organization", Intelex, see https://blog.intelex.com/2018/08/17/poor-quality-costing-organization/#_edn1 (accessed March 15, 2021).

[17] S. Tominaga. "Dichromatic reflection models for a variety of materials." Color Research & Application 19, no. 4 (1994): 277-285.

[18] X. Sun, Y. Liu, X. Yu, H. Wu, and N. Zhang. "Three-dimensional measurement for specular reflection surface based on reflection component separation and priority region filling theory." Sensors 17, no. 1 (2017): 215.

[19] G. Klinker,S. Shafer, and T. Kanade. "The measurement of highlights in color images." Int. J. of Computer Vision 2, 7-32 (1988).

[20] S. Shafer. "Using color to separate reflection components." Color Res. Appl. 10, 210-218 (1985).

[21] H. Lin, J. Gao, G. Zhang, X. Chen, Y. He, and Y Liu. "Review and comparison of high-dynamic range three-dimensional shape measurement techniques." Journal of Sensors, 2017.

[22] F. Chen, G. Brown, and M. Song. "Overview of three-dimensional shape measurement using optical methods." Optical Engineering, vol. 39, no. 1, pp. 10–22, 2000.

[23] G. Sansoni, M. Trebeschi, and F. Docchio. "State-of-the-art and applications of 3D imaging sensors in industry, cultural heritage, medicine, and criminal investigation." Sensors 9, no. 1 (2009): 568-601.

[24] Photogrammetry 2019, see https://en.wikipedia.org/wiki/Photogrammetry (accessed March 15, 2019).

[25] F. Remondino, and S. El-Hakim. "Image-based 3D modelling: A review." Photogramm. Rec. 2006, 21, 269-291.

[26] F. Remondino, and S. El-Hakim. "Critical overview of image-based 3D modeling." Workshop 3D laser Scanning and Photogrammetry for Building Information Modeling, Trondheim, Norway, September 2008; pp. 340-351.

[27] F. Remondino. "From point cloud to surface: the modeling and visualization problem." Workshop on Visualization and Animation of Reality based 3D Models, Tarasp-Vulpera, Switzerland, February 24-28, 2003.

[28] S. El-Hakim, J. Beraldin, and F. Blais. "A comparative Evaluation of the Performance of passive and Active 3-D Vision Systems." SPIE Proc. Conf. Digital Photogrammetry, St. Petersburg, Russia, June 25-30, 1995; 2646, pp. 14-25.

[29] F. Remondino, A. Guarnieri, and A. Vettore. "3D modeling of close-range objects: photogrammetry or laser scanning?" Proc. SPIE Videometrics VIII, Istanbul, Turkey, July 12-23, 2004; 5665, pp. 216-225.

[30] P. Ronnholm, E. Honkavaara, P. Litkey, H. Hyyppa, and J. Hyyppa. "Integration of laser scanning and photogrammetry." IAPRS 2007, XXXVI part 3/W52, 355-362.

[31] "Stereo Vision 2017", The MathWorks, Inc., see http://au.mathworks.com/discovery/stereo-vision.html (accessed March 15, 2019).

[32] "Time-of-Flight Principle, 2019", Terabee, see https://www.terabee.com/time-of-flight-principle/ (accessed March 15, 2019).

[33] D. Nitzan. "Three-Dimensional Vision Structure for Robot Applications." IEEE T. Pattern. Anal. 1988, 10, 291-308.

[34] T. Nielsen, F. Bormann, S. Wolbeck, H. Spiecker, M. Burrows, and P. Andersen. "Time-of-light analysis of light pulses with a temporal resolution of 100ps." Rev. Sci. Instrum. 1996, 67, 1721-1724.

[35] J. Beraldin, F. Blais, L. Cournoyer, R. Rodella, F. Bernier, and N. Harrison. "Digital 3D imaging system for rapid response on remote sites." Proc. Second Int. Conf. on 3-D Digital Imaging and Modeling (3DIM '99), Ottawa, Canada, October 1999; pp. 34-43.

[36] F. Blais, M. Rioux, and J. Beraldin. "Practical considerations for the design of a high precision 3D laser scanner system." Proc. SPIE 1988, 959, 225-246.

[37] Z. Song, H. Jiang, H. Lin, and S. Tang. "A high dynamic range structured light means for the 3D measurement of specular surface." Optics and Lasers in Engineering 95 (2017): 8-16.

[38] D. Palousek, M. Omasta, D. Koutny, J. Bednar, T. Koutecky, and F. Dokoupil. "Effect of matte coating on 3D optical measurement accuracy." Optical Materials, vol. 40, pp. 1–9, 2015.

[39] S. Zhang and S. Yau. "High dynamic range scanning technique." Optical Engineering, vol. 48, no. 3, Article ID 033604, 2009.

[40] G. Liu, X. Liu, and Q. Feng. "3D shape measurement of objects with high dynamic range of surface reflectivity." Applied Optics, vol. 50, no. 23, pp. 4557–4565, 2011.

[41] H. Jiang, H. Zhao, and X. Li. "High dynamic range fringe acquisition: a novel 3-D scanning technique for high-reflective surfaces." Optics and Lasers in Engineering, vol. 50, no. 10, pp. 1484–1493, 2012.

[42] L. Ekstrand and S. Zhang. "Autoexposure for three-dimensional shape measurement using a digital-light-processing projector." Optical Engineering, vol. 50, pp. 895–900, 2011.

[43] S. Feng, Y. Zhang, Q. Chen, C. Zuo, R. Li, and G. Shen. "General solution for high dynamic range three-dimensional shape measurement using the fringe projection technique." Optics and Lasers in Engineering, vol. 59, pp. 56–71, 2014.

[44] K. Zhong, Z. Li, X. Zhou, Y. Li, Y. Shi, and C.Wang. "Enhanced phase measurement profilometry for industrial 3D inspection automation." International Journal of Advanced Manufacturing Technology, vol. 76, no. 9–12, pp. 1563–1574, 2014.

[45] J. Jeong, D. Hong, and H. Cho. "Measurement of partially specular objects by controlling imaging range." in Optomechatronic Computer-Vision Systems II, 671808, vol. 6718 of Proceedings of SPIE, October 2007.

[46] C. Waddington and J. Kofman. "Sinusoidal fringe-pattern projection for 3-D surface measurement with variable illuminance." in Proceedings of the International Symposium on Optomechatronic Technologies (ISOT '10), pp. 1–5, October 2010.

[47] C. Waddington and J. Kofman. "Analysis of measurement sensitivity to illuminance and fringe-pattern gray levels for fringe-pattern projection adaptive to ambient lighting." Optics and Lasers in Engineering, vol. 48, no. 2, pp. 251–256, 2010.

[48] C. Waddington and J. Kofman. "Modified sinusoidal fringe pattern projection for variable illuminance in phase-shifting three-dimensional surface-shape metrology." Optical Engineering, vol. 53, no. 8, Article ID084109, 2014.

[49] C. Waddington and J. Kofman. "Saturation avoidance by adaptive fringe projection in phase-shifting 3D surface-shape measurement." in Proceedings of the International Symposium on Optomechatronic Technologies (ISOT '10), pp. 1–4, IEEE, Ontario, Canada, October 2010.

[50] C. Waddington and J. Kofman. "Camera-independent saturation avoidance in measuring high-reflectivity-variation surfaces using pixel-wise composed images from projected patterns of different maximum gray level." Optics Communications, vol. 333, pp. 32–37, 2014.

[51] G. Babaie, M. Abolbashari, and F. Farahi. "Dynamics range enhancement in digital fringe projection technique." Precision Engineering, vol. 39, pp. 243–251, 2015.

[52] H. Lin, J. Gao, Q. Mei, Y. He, J. Liu, and X. Wang. "Adaptive digital fringe projection technique for high dynamic range three-dimensional shape measurement." Optics Express, vol. 24, no. 7, pp. 7703–7718, 2016.

[53] Y. Yoshinori, M. Hiroyuki, N. Osamu, and I. Tetsuo. "Shape measurement of glossy objects by range finder with polarization optical system." Gazo Denshi Gakkai Kenkyukai Koen Yoko, vol. 200, pp. 43–50, 2003.

[54] S. Umeyama and G. Godin. "Separation of diffuse and specular components of surface reflection by use of polarization and statistical analysis of images." IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 5, pp. 639–647, 2004.

[55] S. Feng, Y. Zhang,Q. Chen, C. Zuo, R. Li, and G. Shen. "General solution for high dynamic range three-dimensional shape measurement using the fringe projection technique." Optics and Lasers in Engineering, vol. 59, pp. 56–71, 2014.

[56] B. Salahieh, Z. Chen, J. J. Rodriguez, and R. Liang. "Multi-polarization fringe projection imaging for high dynamic range objects." Optics Express, vol. 22, no. 8, pp. 10064–10071, 2014.

[57] T. Chen, H. Lensch, C. Fuchs, and H.-P. Seidel. "Polarization and phase-shifting for 3D scanning of translucent objects." in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '07), pp. 1–8, IEEE, June 2007.

[58] R. Liang. "Short wavelength and polarized phase shifting fringe projection imaging of translucent objects." Optical Engineering, vol. 53, no. 1, Article ID 014104, 2014.

[59] R. Tan, K. Nishino, and K. Ikeuchi. "Separating reflection components based on chromaticity and noise analysis." IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 10, pp. 1373–1379, 2004.

[60] R. Benveniste and C. Unsalan. "A color invariant for line stripe-based range scanners." Computer Journal, vol. 54, no. 5, pp. 738– 753, 2011.

[61] R. Woodham. "Photometric method for determining surface orientation from multiple images." Optical Engineering, vol. 19, pp. 1–22, 1992.

[62] F. Nicodemus, J. Richmond, J. Hsia, I. Ginsberg, and T. Limperis. "Geometrical considerations and nomenclature for reflectance." in Radiometry, pp. 94–145, 1977.

[63] C. Hin-Shun and J. Jiaya. "Efficient photometric stereo on glossy surfaces with wide specular lobes." in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08), pp. 1–8, IEEE, Anchorage, Alaska, USA, 2008.

[64] L. Meng, L. Lu, N. Bedard, and K. Berkner. "Single-shot specular surface reconstruction with gonio-plenoptic imaging." in Proceedings of the IEEE International Conference on Computer Vision (ICCV '15), pp. 3433–3441, IEEE, Santiago, Chile, 2015.

[65] G. Ward. "Measuring and modeling anisotropic reflection." SIGGRAPH Computer Graphics, vol. 26, no. 2, pp. 265–272, 1992.

[66] Q. Hu, K. Harding, X. Du, and D. Hamilton. "Shiny parts measurement using color separation." in Two- and Three-Dimensional Methods for Inspection and Metrology III, vol. 6000 of Proceedings of SPIE, pp. 125–132, November 2005.

[67] R. Kowarschik, P. K¨uhmstedt, J. Gerber, W. Schreiber, and G. Notni. "Adaptive optical three-dimensional measurement with structured light." Optical Engineering, vol. 39, no. 1, pp. 150–158, 2000.

[68] Z. Zhang, Y.Wang, S. Huang, Y. Liu, C. Chang, F. Gao, and X. Jiang. "Three-dimensional shape measurements of specular objects using phase-measuring deflectometry." Sensors 17, no. 12 (2017): 2835.

[69] Z. Niu, N. Gao, Z. Zhang, F. Gao, and X. Jiang. "3D shape measurement of discontinuous specular objects based on advanced PMD with bi-telecentric lens." Optics express 26, no. 2 (2018): 1615-1632.

[70] C. Chang, Z. Zhang, N. Gao, and Z. Meng. "Improved infrared phase measuring deflectometry method for the measurement of discontinuous specular objects." Optics and Lasers in Engineering 134 (2020): 106194.

[71] A. Bulbul, M. Grogan, and R. Dahyot. "3D reconstruction of reflective spherical surfaces from multiple images." In IRISH MACHINE VISION & IMAGE PROCESSING Conference proceedings 2015. 2015.

[72] O. Aubreton, A. Bajard, B. Verney, and F. Truchetet. "Infrared system for 3D scanning of metallic surfaces." Machine vision and applications 24, no. 7 (2013): 1513-1524.

[73] S. Feng, Q. Chen, C. Zuo and a. Asundi. "Fast three-dimensional measurements for dynamic scenes with shiny surfaces." Optics Communications, 382, 18-27, 2017.

[74] F. Xu, Y. Zhang, and L. Zhang. "An effective framework for 3D shape measurement of specular surface based on the dichromatic reflection model." Optics Communications 475 (2020): 126210.

[75] G. Rosati, G. Boschetti, A. Biondi, and A. Rossi. "Real-time defect detection on highly reflective curved surfaces." Optics and Lasers in Engineering, 47(3-4), 379-384,2009.

[76] R. Boby, P. Sonakar, M. Singaperumal, and B. Ramamoorthy. "Identification of defects on highly reflective ring components and analysis using machine vision." The International Journal of Advanced Manufacturing Technology, 52(1-4), 217-233,2011.

[77] W. Shao, P. Peng, Y. Shao, and A. Zhou. "A Method for Identifying Defects on Highly Reflective Roller Surface Based on Image Library Matching." Mathematical Problems in Engineering, 2020.

[78] J. Molina, J. Solanes, L. Arnal, and J. Tornero. "On the detection of defects on specular car body surfaces." Robotics and Computer-Integrated Manufacturing 48 (2017): 263-278.

[79] L. Xu, Z. Yang, Z. Jiang, and Y. Chen. "Light source optimization for automatic visual inspection of piston surface defects." The International Journal of Advanced Manufacturing Technology 91, no. 5 (2017): 2245-2256.

[80] R. Seulin, N. Bonnot, F. Merienne, and P. Gorria. "Simulation process for the design and optimization of a machine vision system for specular surface inspection." In Machine Vision and Three-Dimensional Imaging Systems for Inspection and Metrology II, vol. 4567, pp. 129-140. International Society for Optics and Photonics, 2002.

[81] Z. Liu, H. Ukida, P. Ramuhalli, and K. Niel. "Integrated Imaging and Vision Techniques for Industrial Inspection." Springer, 2015.

[82] "Imaging micron-sized defects on reflective surfaces, 2019", STEMMER IMAGING Ltd., see https://www.stemmer-imaging.com/en-gb/news/2016-04-imaging-micron-sized-defects-on-reflective-surfaces/ (accessed Jul. 1, 2020).

[83] "Novel inspection system characterizes shiny reflective parts, 2019", Endeavor Business Media, LLC., see https://www.vision-systems.com/home/article/16736120/novel-inspection-system-characterizes-shiny-reflective-parts (accessed Jul. 29, 2020).

[84] "ATOS 5, 2020", Precise Industrial 3D Metrology, see hhttps://www.gom.com/en/products/high-precision-3d-metrology/atos-5 (accessed Jul. 8, 2020).

[85] "Industrial 3D Scanning Technology, 2020", Precise Industrial 3D Metrology, see https://www.gom.com/en/products/ (accessed Jul. 8, 2020).

[86] "Gocator 2800 Series, 2019", LMI TECHNOLOGIES INC., see https://lmi3d.com/series/gocator-2800-series/ (accessed Jul. 8, 2020).

[87] "Gocator 3200 Series, 2019", Industrial Control, see https://www.industrialcontrol.com/gocator-3200 (accessed Jul. 8, 2020).

[88] "GoMax 3504, 2019", LMI TECHNOLOGIES INC., see https://lmi3d.com/g3504/ (accessed Jul. 8, 2020).

[89] "Introducing Gocator 3506: 3D Smart Snapshot Sensor,2019", LMI TECHNOLOGIES INC., see https://lmi3d.com/resource/introducing-gocator-3506-3d-smart-snapshot-sensor/ (accessed Jul. 8, 2020).

[90] "N 30 / N 35 ,2020," Ensenso., see https://www.ensenso.com/portfolio-item/n3x/ (accessed Jul. 8, 2020).

[91] "X 30 /X 36,2020", Ensenso., see https://www.ensenso.com/portfolio-item/x3x/ (accessed Jul. 8, 2020).

[92] S. Kumar. "Advantages and Disadvantages of Artificial Intelligence, 2019," see https://towardsdatascience.com/advantages-and-disadvantages-of-artificial-intelligence-\\182a5ef6588c (accessed Jul. 8, 2020).

[93] "Difference between Artificial intelligence and Machine learning, 2019," see https://www.javatpoint.com/

difference-between-artificial-intelligence-and-machine-learning (accessed Jul. 8, 2020).

[94] "Artificial Intelligence, 2020," see https://en.wikipedia.org/wiki/Artificial_intelligence (accessed Jul. 8, 2020).

[95] J. Yun, W. Shin, G. Koo, M. Kim, C. Lee, and S. Lee. "Automated defect inspection system for metal surfaces based on deep learning and data augmentation." Journal of Manufacturing Systems 55 (2020): 317-324.

[96] I. Sarker. "Machine learning: Algorithms, real-world applications and research directions." SN Computer Science 2, no. 3 (2021): 1-21.

[97] J. Chen. "Neural network, 2020," see https://www.investopedia.com/terms/n/neuralnetwork.asp (accessed Jul. 8, 2020).

[98] W. Thompson, H. Li, and A. Bolen. "Artificial intelligence, machine learning, deep learning and beyond, 2020," https://www.sas.com/en_nz/insights/articles/big-data/artificial-intelligence-machine-learning-deep-learning-and-beyond.html (accessed Jul. 8,2020).

[99] IBM Cloud Education. "Deep learning, 2020," see https://www.ibm.com/cloud/learn/deep-learning (accessed Jul. 8, 2020).

[100] "Deep Learning, 2020," see https://en.wikipedia.org/wiki/Deep_learning (accessed Jul. 8, 2020).

[101] IBM Cloud Education. "Neural Networks, 2020," see https://www.ibm.com/cloud/learn/neural-networks (accessed Jul. 8, 2020).

[102] "Supervised learning, 2020," see https://en.wikipedia.org/wiki/Supervised_learning (accessed Jul. 8, 2020).

[103] J.Brownlee. "What is Deep Learning?, 2019," see https://machinelearningmastery.com/what-is-deep-learning/ (accessed Jul. 8, 2020).

[104] T. Dietterich. "Machine learning for sequential data: A review." In Joint IAPR international workshops on statistical techniques in pattern recognition (SPR) and structural and syntactic pattern recognition (SSPR), pp. 15-30. Springer, Berlin, Heidelberg, 2002.

[105] "Neural network models, 2020," see https://scikit-learn.org/stable/modules/neural_networks_supervised.html (accessed Jul. 8, 2020).

[106] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. Inman. "1D convolutional neural networks and applications: A survey." Mechanical systems and signal processing 151 (2021): 107398.

[107] Y. LeCun, Y. Bengio, and G. Hinton. "Deep learning." Nature 2015, 521, 436–444.

[108] Y. Bengio, I. Goodfellow, and A. Courville. "Deep Learning." MIT Press: Cambridge, MA, USA, 2017.

[109] S. Huang, J. Tang, J. Dai, and Y. Wang. "Signal status recognition based on 1DCNN and its feature extraction mechanism analysis." Sensors 2019, 19, 2018.

[110] H. Wang, Z. Liu, D. Peng, and Y. Qin. "Understanding and Learning Discriminant Features based on Multi-Attention 1DCNN for Wheelset Bearing Fault Diagnosis." IEEE Trans. Ind. Inform. 2019, 16, 5735–5745.

[111] S. Huang, J. Tang, J. Dai, Y. Wang, and J. Dong. "1DCNN Fault Diagnosis Based on Cubic Spline Interpolation Pooling." Shock Vib. 2020, 2020.

[112] M. Ragab, S. Abdulkadir, N. Aziz, Q. Al-Tashi, Y. Alyousifi, H. Alhussian, and A. Alqushaibi. "A Novel One-Dimensional CNN with Exponential Adaptive Gradients for Air Pollution Index Prediction." Sustainability 12, no. 23 (2020): 10090.

[113] X. Zhao, J. Solé-Casals, B. Li, Z. Huang, A. Wang, J. Cao, T. Tanaka, and Q. Zhao. "Classification of Epileptic IEEG Signals by CNN and Data Augmentation." In Proceedings of the ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020; pp. 926–930.

[114] A. Haidar, and B. Verma. "Monthly rainfall forecasting using one-dimensional deep convolutional neural network." IEEE Access 2018, 6, 69053–69063.

[115] L. Eren, T. Ince, and S. Kiranyaz. "A generic intelligent bearing fault diagnosis system using compact adaptive 1D CNN classifier." J. Signal Process. Syst. 2019, 91, 179–189.

[116] J. Zhao, X. Mao, and L. Chen. "Learning deep features to recognise speech emotion using merged deep CNN." IET Signal Process. 2018, 12, 713–721.

[117] M. Lv, W. Xu, and T. Chen. "A hybrid deep convolutional and recurrent neural network for complex activity recognition using multimodal sensors." Neurocomputing 2019, 362, 33–40.

[118] B. Jeon, N. Park, and S. Bang. "Dropout Prediction over Weeks in MOOCs via Interpretable Multi-Layer Representation Learning." arXiv 2020, arXiv:2002.01598.

[119] Q. Fu, D. Niu, Z. Zang, J. Huang, and L. Diao. "Multi-Stations'Weather Prediction Based on Hybrid Model Using 1D CNN and Bi-LSTM." In Proceedings of the 2019 Chinese Control Conference (CCC), Guangzhou, China, 27–30 July 2019; pp. 3771–3775.

[120] Y. Gal, and Z. Ghahramani. "A theoretically grounded application of dropout in recurrent neural networks." In Advances in Neural Information Processing Systems; Barcelona, Spain, 2016; pp. 1019–1027.

[121] J. Xiong, K. Zhang, and H. Zhang. "A Vibrating Mechanism to Prevent Neural Networks from Overfitting." In Proceedings of the 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC), Tangier, Morocco, 24–28 June 2019; pp. 1737–1742.

[122] R. Fukuoka, H. Suzuki, T. Kitajima, A. Kuwahara, and T. Yasuno. "Wind Speed Prediction Model Using LSTM and 1D-CNN." J. Signal Process. 2018, 22, 207–210.

[123] S. Soni, S. Dey, and M. Manikandan. "Automatic Audio Event Recognition Schemes for Context-Aware Audio Computing Devices." In Proceedings of the 2019 Seventh International Conference on Digital Information Processing and Communications (ICDIPC), Trabzon, Turkey, 2–4 May 2019; pp. 23–28.

[124] J. Zhao, X. Mao, and L. Chen. "Speech emotion recognition using deep 1D & 2D CNN LSTM networks." Biomed. Signal Process. Control 2019, 47, 312–323.

[125] S. Shao, P. Wang, and R. Yan. "Generative adversarial networks for data augmentation in machine fault diagnosis." Comput. Ind. 2019, 106, 85–93.

[126] S. Abdoli, P. Cardinal, and A. Koerich. "End-to-end environmental sound classification using a 1D convolutional neural network." Expert Syst. Appl. 2019, 136, 252–263.

[127] A. Courville. "Recurrent Batch Normalization." arXiv 2016, arXiv:1603.09025

[128] A. Meliboev, J. Alikhanov, and W. Kim. "1D CNN Based Network Intrusion Detection with Normalization on Imbalanced Data." arXiv 2020, arXiv:2003.00476.

[129] P. Tsirikoglou, S. Abraham, F. Contino, C. Lacor, and G. Ghorbaniasl. "A hyperparameters selection technique for support vector regression models." Appl. Soft Comput. 2017, 61, 139–148.

[130] Z. Zhang. "Improved adam optimizer for deep neural networks." In Proceedings of the 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS), Banff, AB, Canada, 4–6 June 2018; pp. 1–2. Sustainability 2020.

[131] A. Botev, G. Lever, and D. Barber. "Nesterov's accelerated gradient and momentum as approximations to regularised update descent." In Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, Alaska, USA, 14–19 May 2017; pp. 1899–1903.

[132] S. Ruder. "An overview of gradient descent optimization algorithms." arXiv 2016, arXiv:1609.04747.

[133] J. Duchi, E. Hazan, and Y. Singer. "Adaptive subgradient methods for online learning and stochastic optimization." J. Mach. Learn. Res. 2011, 12, 2121–2159.

[134] S. Reddi, S. Kale, and S. Kumar. "On the convergence of adam and beyond." arXiv 2019, arXiv:1904.09237.

[135] M. Zeiler. "Adadelta: An adaptive learning rate method." arXiv 2012, arXiv:1212.5701.

[136] T. Tieleman, and G. Hinton. "Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude." Coursera Neural Netw. Mach. Learn. 2012, 4, 26–31.

[137] D. Kingma, and J. Ba. "Adam: A method for stochastic optimization." arXiv 2014, arXiv:1412.6980.

[138] A. Wilson, R. Roelofs, M. Stern, N. Srebro, and B. Recht. "The marginal value of adaptive gradient methods in machine learning." In Advances in Neural Information Processing Systems; The MIT Press: Long Beach, CA, USA, 2017; pp. 4148–4158.

[139] N. Shazeer, and M. Stern. "Adafactor: Adaptive learning rates with sublinear memory cost." arXiv 2018, arXiv:1804.04235.

[140] L. Luo, Y. Xiong, Y. Liu, and X. Sun. "Adaptive gradient methods with dynamic bound of learning rate." arXiv 2019, arXiv:1902.09843.

[141] N. Donges. "A Guide to RNN: Understanding Recurrent Neural Networks and LSTM Networks, 2021," see https://builtin.com/data-science/recurrent-neural-networks-and-lstm (accessed Aug. 8, 2021).

[142] P. Srivastava. "Essentials of Deep Learning : Introduction to Long Short Term Memory, 2017," see https://www.analyticsvidhya.com/blog/2017/12/fundamentals-of-deep-learning-introduction-to-lstm/ (accessed Jul. 8, 2020).

[143] M. Phi. "Illustrated Guide to LSTM's and GRU's: A step by step explanation, 2018," see https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation\\-44e9eb85bf21 (accessed Jul. 8, 2020).

[144] S. Saxena. "Introduction to Long Short Term Memory (LSTM), 2021," see https://www.analyticsvidhya.com/blog/2021/03/introduction-to-long-short-term-memory-lstm/ (accessed Jul. 8, 2021).

[145] "Support Vector Machines, 2020," see https://scikit-learn.org/stable/modules/svm.html#scores-probabilities (accessed Jul. 8, 2020).

[146] "Nearest Neighbors Regression, 2020," see https://scikit-learn.org/stable/modules/neighbors.html#regression (accessed Jul. 8, 2020).

[147] "Decision Trees, 2020," see https://scikit-learn.org/stable/modules/tree.html#regression (accessed Jul. 8, 2020).

[148] "RandomForestRegressor, 2020," see https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html#sklearn.ensemble.RandomForestRegressor (accessed Jul. 8, 2020).

[149] "Ensemble methods, 2020," see https://scikit-learn.org/stable/modules/ensemble.html#forest (accessed Jul. 8, 2020).

[150] "XGBoost Algorithm, 2020," see https://towardsdatascience.com/https-medium-com-vishalmorde-xgboost-algorithm-long-she-may-rein\\-edd9f99be63d (accessed Jul. 8, 2020).

[151] H. Song, J. Dai, L. Luo, G. Sheng, and X. Jiang. "Power transformer operating state prediction method based on an LSTM network." Energies 2018, 11, 914.

[152] R. Fu, Z. Zhang, and L. Li. "Using LSTM and GRU neural network methods for traffic flow prediction." In Proceedings of the 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), Wuhan, China, 11–13 November 2016; pp. 324–328.

[153] D. Maestro-Watson, J. Balzategui, L. Eciolaza, and N. Arana-Arexolaleiba. "Deep learning for deflectometric inspection of specular surfaces." In The 13th International Conference on Soft Computing Models in Industrial and Environmental Applications, pp. 280-289. Springer, Cham, 2018.

[154] S. Martínez, C. Vázquez, J. García, and J. Ortega. "Quality inspection of machined metal parts using an image fusion technique." Measurement 111 (2017): 374-383.

[155] Q. Zhou, R. Chen, B. Huang, W. Xu, and J. Yu. "DeepInspection: Deep learning based hierarchical network for specular surface inspection." Measurement 160 (2020): 107834.

[156] Z. Zhang, B. Zhang, and T. Akiduki. "Specular reflection Surface Defects Detection by using Deep Learning." In Proceedings of the 2019 3rd International Conference on Information System and Data Mining, pp. 6-10. 2019.

[157] W. Xiao. "Defect classification on a painted specular surface using machine learning." In CONFERENCE IN INTERACTION TECHNOLOGY AND DESIGN, p. 99.

[158] Z. Zhou, Q. Lu, Z. Wang, and H. Huang. "Detection of micro-defects on irregular reflective surfaces based on improved faster R-CNN." Sensors 19, no. 22 (2019): 5000.

[159] A.Dawda and M.Nguyen, "Comparison of Red versus Blue Laser Light for Accurate 3D Measurement of Highly Specular Surfaces in Ambient Lighting Conditions." In 2021 International Symposium On Geometry and Vision (ISGV), IEEE, January 2021.

[160] "Genie Nano gige camera, 2018," A Teledyne Technologies company, see `https://www.teledynedalsa.com/en/products/imaging/cameras/genie-nano-1gige/` (accessed Jul. 8, 2018).

[161] "ZM18, 2018," see `https://z-laser.com/en/product/laser-modules/zm-laser-family/zm18/` (accessed Jul. 8, 2018).

[162] "Incremental encoders, 2020," Sick AG, see `https://www.sick.com/nz/en/encoders/incremental-encoders/dbs36-core/dbs36e-s3ck02500/p/p360058?ff_data=JmZmX2lkPXAzNjAwNTgmZmZfbWFzdGVySWQ9cDM2MDA1OCZmZl90a\\XRsZT1EQlMzNkUtUzNDSzAyNTAwJmZmX3F1ZXJ5PSZmZl9wb3M9NCZmZl9vcmln\\UG9zPTQmZmZfcGFnZT0xJmZmX3BhZ2VTaXplPTI0JmZmX29yaWdQYWdlU2l6ZT0\\yNCZmZl9zaW1pPTk0LjA=` (accessed Jul. 8, 2020).

[163] "HALCON – THE POWER OF MACHINE VISION, 2021," MVTec software GmBH, see `https://www.mvtec.com/products/halcon` (accessed Jul. 8, 2021).

[164] "Halcon, 2021," see `https://www.matrix-vision.com/halcon-en.html` (accessed Jul. 8, 2021).

[165] Solution Guide III-C 3D Vision, Machine Vision in 3D World Coordinates, Version 18.05, MVTec Software GmbH, M′unchen (2018).

[166] I. Howard and B. Rogers, "Binocular vision and stereopsis." New York: Oxford University Press, 1995.

[167] J. Morris, "Design of an active stereo vision 3d scene reconstruction system based on the linear position sensor module." 2006.

[168] "3D Machine Vision using Laser Triangulation, 2021," Tech Briefs Media Group, see https://www.techbriefs.com/component/content/article/tb/supplements/ptb/features/applications/22464?start=1 (accessed Jul. 8, 2021).

[169] "Understanding laser-based 3D triangulation methods, 2015," PennWell Corporation, Tusla, see https://www.vision-systems.com/articles/print/volume-20/issue-6/features/understanding-laser-based-3d-triangulation-methods.html (accessed Jul. 8, 2019).

[170] "smooth_funct_1d_gauss, 2020," MVTec Software GmbH, see https://www.mvtec.com/doc/halcon/12/en/smooth_funct_1d_gauss.html (accessed Feb. 14, 2020).

[171] "local_min_max_funct_1d, 2020," MVTec Software GmbH, see https://www.mvtec.com/doc/halcon/13/en/local_min_max_funct_1d.html (accessed Apr. 17, 2020).

[172] "edges_image, 2020," MVTec Software GmbH, see https://www.mvtec.com/doc/halcon/13/en/edges_image.html (accessed Apr. 17, 2020).

[173] "disparity_to_point_3d, 2020," MVTec Software GmbH, see https://www.mvtec.com/doc/halcon/11/en/disparity_to_point_3d.html (accessed Apr. 17, 2020).

[174] "hom_mat3d_rotate, 2020", MVTec Software GmbH, see https://www.mvtec.com/doc/halcon/13/en/hom_mat3d_rotate.html (accessed Feb. 14, 2020).

[175] "union_object_model_3d, 2020", MVTec Software GmbH, see https://www.mvtec.com/doc/halcon/13/en/union_object_model_3d.html (accessed Feb. 14, 2020).

[176] T. Czimmermann, G. Ciuti, M. Milazzo, M. Chiurazzi, S. Roccella, C.Oddo, and P. Dario. "Visual-Based Defect Detection and Classification Approaches for Industrial Applications—A SURVEY." Sensors, 20(5), 1459, 2020.

[177] "hom_mat3d_translate, 2020," MVTec Software GmbH, see https://www.mvtec.com/doc/halcon/13/en/hom_mat3d_translate.html (accessed Feb. 14, 2020).

[178] "register_object_model_3d_pair, 2020," MVTec Software GmbH, see https://www.mvtec.com/doc/halcon/13/en/register_object_model_3d_pair.html (accessed Feb. 14, 2020).

[179] "rigid_trans_object_model_3d, 2020," MVTec Software GmbH, see https://www.mvtec.com/doc/halcon/13/en/rigid_trans_object_model_3d.html (accessed Feb. 14, 2020).

[180] "distance_object_model_3d, 2020," MVTec Software GmbH, see https://www.mvtec.com/doc/halcon/12/en/distance_object_model_3d.html (accessed Feb. 14, 2020).

[181] "Laser, 2020," see https://en.wikipedia.org/wiki/Laser (accessed Feb. 14, 2020).

[182] "Laser pointer, 2020", see https://en.wikipedia.org/wiki/Laser_pointerColors_and_wavelengths (accessed Feb. 14, 2020).

[183] "Visible spectrum, 2020", see https://en.wikipedia.org/wiki/Visible_spectrum (accessed Feb. 14, 2020).

[184] "Laser classes chart, 2020", see https://www.laserpointersafety.com/laserclasses.html (accessed Feb. 14, 2020).

[185] "Eye injury hazard, 2020," see https://www.laserpointersafety.com/resources/Diagrams/Arrow---eye-injury-hazard-for-laser-classes\\-0800-v2.png (accessed Feb. 14, 2020).

[186] "Displacement, 2020," Bestech Australia, see https://www.bestech.com.au/blogs/displacement/3-reasons-to-use-blue-over-red-lasers\\-in-displacement-measurement/ (accessed Feb. 14, 2020).

[187] "Gocator 2500 Series, 2019," LMI TECHNOLOGIES INC., see https://lmi3d.com/products/gocator-3d-laser-sensors/gocator-2500-series (accessed Feb. 14, 2020).

[188] "In-sight Laser Profiler, 2019," Cognex Corporation, see https://www.cognex.com/en-nz/products/machine-vision/3d-laser-profilers/in-sight-laser-profiler (accessed Feb. 14, 2020).

[189] "Laser scanners for 2D/3D profile measurements, 2019," Micro-Epsilon, see https://www.micro-epsilon.com/2D_3D/laser-scanner/ (accessed Feb. 14, 2020).

[190] A. Dawda, A. Varasada, and M. Nguyen. "3-Dimensional Reconstruction of a Highly Specular or Transparent Cylinder from a Single Image." EAI Smartgift, Portugal, 2021.

[191] R. Klette. "Concise Computer Vision." Springer, London (2014).

[192] V. Pˇatrˇaucean, P. Gurdjos, and R.G.V. Gioi. "Joint A Contrario Ellipse and line detection." IEEE Trans. Pattern Analysis Machine Intelligence, vol. 39, pp. 788–802, 2017.

[193] J. Ren, H.M. Owais, T. Song, and D. Lin. "Towards fast and accurate ellipse and semi-ellipse detection." Proc. IEEE Int. Conf. Image Processing, pp. 743–747, 2018.

[194] J.N. Ouellet and P. H´ebert. "Precise ellipse estimation without contour point extraction." Machine Vision Applications, vol. 21, pp. 59–67, 2010.

[195] Z. Xu, S. Xu, C. Qian, and R. Klette." Ellipse Extraction in Low-Quality Images." In 2019 16th International Conference on Machine Vision Applications (MVA) (pp. 1-5). IEEE (2019).

[196] K. Kanatani, Y. Sugaya, and Y. Kanazawa. "Ellipse Fitting." In: Guide to 3D Vision Computation, pp. 11–32, 2016.

[197] K. Kanatani, Y. Sugaya, and Y. Kanazawa. "Ellipse Fitting for Computer Vision: Implementation and Applications." Morgan and Claypool, Williston, VA, 2016.

[198] V. Kovalevsky. "Modern Algorithms for Image Processing." (Chapter 13), Apress, Springer Science + Business Media Finance, Delware, 2019.

[199] T. Masuzaki, Y. Sugaya, and K. Kanatani. "High accuracy ellipse specific fitting." Proc. Pacific-Rim Symposium Image Video Technology, pp. 314–324, 2014.

[200] D.K. Prasad, M.K. Leung, and C. Quek. "ElliFit: an unconstrained, noniterative, least squares based geometric ellipse fitting method." Pattern Recognition, vol. 46, pp. 1449–1465, 2013.

[201] Y. Wang, Z. He, X. Liu, Z. Tang, and L. Li. "A fast and robust ellipse detector based on top-down least-square fitting." Proc. British Machine Vision Conf., pp. 156.1–156.12, 2015.

[202] S. Mulleti and C.S. Seelamantula. "Ellipse fitting using the finite rate of innovation sampling principle." IEEE Trans. Image Processing, vol. 25, pp. 1451–1464, 2016.

[203] W. Chojnacki, M.J. Brooks, A. Hengel, and D. Gawley. "On the fitting of surfaces to data with covariances." IEEE Trans. Pattern Analysis Machine Intelligence, vol. 22, pp. 1294–1303, 2000.

[204] A. Fitzgibbon, M. Pilu, and R. Fisher. "Direct least square fitting of ellipses." IEEE Trans. Pattern Analysis Machine Intelligence, 21(5):476–480 (1999).

[205] "fit_ellipse_contour_xld, 2019," See `www.mvtec.com/doc/halcon/12/en/fit\_ellipse\_contour\_xld.html` (accessed Jul. 8, 2019).

[206] "fit_circle_contour_xld, 2019," See `www.mvtec.com/doc/halcon/12/en/fit_circle_contour_xld.html` (accessed Jul. 8, 2019).

[207] C. Zhao, J. Yang, F. Zhou, J. Sun, X. Li, and W. Xie. "A Robust Laser Stripe Extraction Method for Structured-Light Vision Sensing." Sensors 20, no. 16 (2020): 4544.

[208] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng. "TensorFlow: Large-scale machine learning on heterogeneous systems." J. CoRR, arXiv:1603.04467, 2016.

[209] F. Chollet. "Keras: Deep Learning for humans, 2020," GitHub, see `github.com/fchollet/keras` (accessed Jul. 8, 2020).

[210] H. Jin, Q. Song, X. Hu. "Efficient neural architecture search with network morphism." arXiv preprint arXiv:1806.10282. 2018.

[211] A. Shaleynikov. "10 best frameworks and libraries for AI," See dzone.com/articles/progressive-tools10-best-frameworks-and-libraries (accessed Jul. 8, 2020).

[212] "Global average pooling 1D, 2020," See https://peltarion.com/knowledge-center/documentation/modeling-view/build-an-ai-model/blocks/global-average-pooling-1d (accessed Jul. 8, 2020).

[213] C. Chen, Z. Liu, G. Yang, C. Wu, and Q. Ye. "An Improved Fault Diagnosis Using 1D-Convolutional Neural Network Model." Electronics 10, no. 1 (2021): 59.

[214] Nerian Vision Technologies. "SP1:Real-Time 3D Stereo Vision Through FPGA Technology, 2017," See https://nerian.com/products/sp1-stereo-vision/ (accessed Jul. 8, 2017).

[215] H. Hirschm¨ uller, "Stereo processing by semiglobal matching and mutual information", IEEE Trans. Pattern Analysis Machine Intelligence, 30(2): 328–341, 2008.

[216] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms", Int. J. Computer Vision, 47(1-3):7–42, 2002.