



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

공학박사 학위 논문

구조화된 광원을 갖는 단일비전을 이용한 3차원 측정

**THREE-DIMENSIONAL MEASUREMENT USING
MONOCULAR STRUCTURED LIGHT VISION**

울산대학교 대학원

기계공학과 항공우주공학전공

TRAN THI TRANG

**THREE-DIMENSIONAL MEASUREMENT USING MONOCULAR
STRUCTURED LIGHT VISION**

A thesis submitted in partial fulfillment of the requirement for the Degree of
Doctor of Philosophy to Department of Aerospace
Engineering, University of Ulsan, Korea

By
TRAN THI TRANG

May, 2018

구조화된 광원을 갖는 단일비전을 이용한 3 차원 측정

**THREE-DIMENSIONAL MEASUREMENT USING
MONOCULAR STRUCTURED LIGHT VISION**

지도교수 하철근

이논문을 공학박사 학위 논문으로 제출함

2018 년 5 월

울산대학교 대학원

기계공학과 항공우주공학전공

**THREE-DIMENSIONAL MEASUREMENT USING
MONOCULAR STRUCTURED LIGHT VISION**

This certifies that the dissertation of

TRAN THI TRANG is approved by

Committee Chairman: Prof. JICHUL SHIN

Committee Member: Dr. HYOUNG SIK CHOI

Committee Member: Dr. JONG TAI JANG

Committee Member: Dr. SIBOK YU

Committee Member: Prof. CHEOLKEUN HA

Department of Aerospace Engineering

University of Ulsan, Korea

May, 2018

TRAN THI TRANG 의
공학박사 학위 논문을 인준함

심사위원장	신 지 철	인
심사위원	최 형 식	인
심사위원	장 종 태	인
심사위원	유 시 복	인
심사위원	하 철 근	인

울 산 대 학 교 대 학 원
기계공학과 항공우주공학전공

2018 년 5월

Acknowledgements

First of all, I would like to thank my advisor, Prof. CheolKeun Ha, for providing me great opportunity to study and work in the Flight Dynamics and Control Lab, University of Ulsan. I deeply appreciate Prof. Ha for his insightful academic advices, enthusiastic encouragement and thoughtful guidance throughout the study. It was my great honor to be Prof. Ha's student.

I would like to express my gratitude to committee members Prof. JICHUL SHIN, Dr. HYOUNG SIK CHOI, Dr. JONG TAI JANG, Dr. SIBOK YU for evaluation and giving me beneficial suggestions and comments throughout this thesis.

I would like to thank Prof. Nguyen Quoc Binh, Prof. Xuan Nam Tran, Prof. Nguyen Van Thuan at Le Quy Don Technical University who made me desire more knowledge. Without their warm encouragement, this thesis might have been impossible.

I am thankful to my friends, my colleagues, all members of Flight Dynamics and Control Lab, University of Ulsan, for their warm encouragement and friendly environment they created for me during the time I am in Korea.

Finally, I would like to thank my parents, my brother Tran Van Nam, my sister-in-law To Khanh Hoa, my husband Dr. To Xuan Dinh, and my daughter To Xuan Sa Rang for their patience, understanding and support during my study. Last, but not least, with much regret, I would like to thank my grandparents and my parents-in-law who passed away before I could express my gratitude.

Ulsan, May 2018

Tran Thi Trang

Contents

Acknowledgements.....	vi
Contents	vii
List of Figures	x
List of Tables	xii
Abbreviations.....	xiii
ABSTRACT.....	xiv
Chapter 1 INTRODUCTION.....	1
1.1 Overview	1
1.2 Research objectives	2
1.3 Dissertation outline	3
Chapter 2 FUNDAMENTAL TECHNIQUES IN MONOCULAR VISION SYSTEMS	4
2.1 Introduction	4
2.2 Camera calibration	5
2.3 Extrinsic calibration of camera and structured light by using a rectangle	11
2.3.1 Camera and rectangle object calibration.....	13
2.3.2 Extrinsic calibration of the structured multi-line light laser and calibration plane	14
2.3.3 Extrinsic calibration of the camera and structured multi-line laser light.....	17
2.4 Experimental results.....	17
2.4.1 Computer simulations	17
2.4.2 Real data.....	20
2.5 Chapter summary	22
Chapter 3 CIRCULAR OBJECT DETECTION AND TRACKING USING DYNAMIC FUSION OF COLOR AND SHAPE	24
3.1 Introduction	24
3.2 Circular object detection and tracking	25
3.2.1 Feature Space	25
3.2.2 Feature Estimator	27
3.2.3 Fusion of color and shape for irregular moving object detection and tracking	33
3.3 Experimental Results.....	34
3.3.1 Installation and reinstallation after occlusion confirmation.....	35

3.3.2	Performance of chasing a moving target	36
3.3.3	Performance of object detection	36
3.3.4	Performance w.r.t the changing of illumination	37
3.3.5	Performance w.r.t the changing of view point.....	37
3.4	Chapter summary	38
Chapter 4 SLIPPAGE MEASUREMENT USING MONOCULAR STRUCTURED LINE LIGHT VISION		
39		
4.1	Introduction	39
4.1.1	Related works.....	39
4.1.2	Proposed system overview.....	40
4.2	Slippage measurement.....	43
4.2.1	Target object pose estimation	44
4.2.2	Slippage measurement	49
4.3	Experimental results.....	50
4.3.1	Target object increment distance measurement.....	51
4.3.2	Slippage estimation with different target objects.....	52
4.4	Chapter summary	53
Chapter 5 GAP AND FLUSH MEASUREMENT USING MONOCULAR STRUCTURED MULTI-LINE LIGHT VISION AND ITS APPLICATION IN VIHECLE ASEMBLY		
55		
5.1	Introduction	55
5.2	Gap and flush measurement system	57
5.3	Image Processing Module	59
5.3.1	Selection of the ROI.....	59
5.3.2	Laser stripe detection and classification	61
5.3.3	Feature circle extraction.....	63
5.4	Proposed gap and flush measurement method	65
5.5	Experimental results.....	68
5.5.1	Camera-laser module performance confirmation	69
5.5.2	Gap and flush measurement in real vehicle assembly	75
5.6	Chapter summary	77
Chapter 6 MONOCULAR STRUCTURED LIGHT VISION-BASED SURFACE MEASUREMENT AND ITS APPLICATION IN UAV LANDING SURFACE DETERMINATION		
78		
6.1	Introduction	78

6.2	Laser profile extraction	79
6.3	Proposed ground surface estimation	80
6.4	Experiment Results	82
6.5	Chapter summary	88
Chapter 7	CONCLUSION	89
PUBLICATIONS.....		90
REFERENCES		92

List of Figures

Fig. 2-1. Camera pinhole model	6
Fig. 2-2. Image plane	6
Fig. 2-3. Sample calibration image	7
Fig. 2-4. Corners detection and drawing.....	8
Fig. 3-1 Example of color illumination changes over time.....	26
Fig. 3-2 Example of shape change.....	27
Fig. 3-3 (a) A sample image (906×918). (b) Contour detection. Each color represents a different contour. A total of 16 contours are extracted in 4 ms. (c) The remaining contours in vector T after complete circle detection. (d) A total of 60 contour segments are approximated.	28
Fig. 3-4 Example of installation and reinstallation after occlusion.	35
Fig. 3-5 Trajectory of the center of an object in the image	35
Fig. 3-6 Chasing a moving target.	36
Fig. 3-7 Chasing Comparison between object detection and measurement approaches. (a) using proposed feature fusion; (b) using color information; (c) using shape information	36
Fig. 3-8 Video sequence contains drastic illumination changes. (a) Initialize model; (b) Tracking using proposed online feature fusion; (c) Tracking using color information alone fails due to illumination changes.	37
Fig. 3-9 Video sequence contains view point changes. (a) Initialize model; (b) Tracking using online feature fusion; (c) Tracking using shape information alone fails due to view point changes.	38
Fig. 4-1 Video Examples of objects with rectangular features	40
Fig. 4-2 The non-contact slippage measurement system using monocular structure line light vision.	40
Fig. 4-3 The configuration of the monocular-structured light vision-based slippage estimation apparatus.	41
Fig. 4-4 A flowchart of slippage estimation using monocular structured line vision.	42
Fig. 4-5 The extracted target object lines and structured line detection. (a) Original image; (b) curve detection; (c) detection of high curvature points; (d) object lines and structured line detection.	42
Fig. 4-6 The vision sensor model of the monocular structured light slippage measurement.	44
Fig. 4-7 The real testbed system.	50
Fig. 4-8 Slippage measurement and measurement error. (a) Slippage estimation; (b) measurement error.	51
Fig. 4-9 Slippage measurement with gripper force $F_d = 47.092[N]$ and $M = 2.3[kg]$	51
Fig. 4-10 Slippage measurement with a gripper force of $F_d = 46.36[N]$ and $M = 2.1[kg]$	53
Fig. 4-11 Slippage measurement with a gripper force of $F_d = 46.16[N]$ and $M = 1.7[kg]$	53
Fig. 5-1 Gap and flush measurement in vehicle assembly. (a) Portable gap and flush measurement system [28]; (b) Inline gap and flush measurement robot system [29].	57
Fig. 5-2 The external and internal view of a gap and flush measurement performed by the camera-laser module.....	57
Fig. 5-3 An example of acquired images of laser reflection on car body parts with different gain values [26]. (a) gain value 0; (b) gain value 10; (c) gain value 40.....	57
Fig. 5-4 Acquired image of laser stripes on a real vehicle.....	59
Fig. 5-5 The selected ROI (small rectangle) within the acquired image.	59
Fig. 5-6. Contour point detection image of the ROI after noise elimination.	62
Fig. 5-7. Laser stripes on flat measured object. a) No gap and flush; b) Deformation of laser stripes in the presence of a gap and flush.....	63

<i>Fig. 5-8 Internal view of a gap and flush on a flat measurement object in the image plane.</i>	63
<i>Fig. 5-9 Sliding circle.</i>	63
<i>Fig. 5-10 Gap and flush measurement using feature circles.</i>	63
<i>Fig. 5-11 A real camera-laser module setup.</i>	65
<i>Figure 5-12 Camera and structured multi-line light calibration using rectangle model.</i>	65
<i>Fig. 5-13 Pseudo code of the feature circle position 3D mapping algorithm for a gap and flush measurement</i>	67
<i>Fig. 5-14 Camera-laser module performance confirmation test.</i>	69
<i>Fig. 5-15 Sample pattern</i>	70
<i>Fig. 5-16 Camera-laser module accuracy confirmation w.r.t changing of measurement position at the same working distance. (a) S1 measurement error; (b) S2 measurement error; (c) S3 measurement error.</i>	71
<i>Fig. 5-17 Camera-laser module precision confirmation w.r.t changing of measurement position at the same working distance. (a) S1 measurement error; (b) S2 measurement error; (c) S3 measurement error.</i>	72
<i>Fig. 5-18 Camera-laser module accuracy confirmation w.r.t changing of the measurement working distance.</i>	74
<i>Fig. 5-19 Camera-laser module precision confirmation w.r.t changing of the measurement working distance.</i>	74
<i>Fig. 5-20 Three proposed calibrated camera-laser modules attached to the end-effector of the robot arm in a vehicle assembly application.</i>	75
<i>Fig. 5-21 Gap and flush measurement using proposed system in real vehicle assembly.</i>	76
<i>Fig. 5-22 The gap and flush measurement software GUI.</i>	76
<i>Figure 6-1 The plane surface angle measurement setup; (a) Experiment testbed, (b) monocular structured laser vision module.</i>	83
<i>Figure 6-2 Real 3D reconstruction testbed.</i>	83
<i>Figure 6-3 The complex surface captured by the camera.</i>	83
<i>Figure 6-4 3D surface reconstruction.</i>	84
<i>Figure 6-5 Sample pattern.</i>	86
<i>Figure 6-6 3D surface reconstruction measurement. (a) S1 measurement error; (b) S2 measurement error; (c) S3 measurement error.</i>	87

List of Tables

<i>Table 5-1 Current traditional gap and flush measurement sensors specifications and proposed gap and flush measurement sensor requirements.</i>	<i>68</i>
<i>Table 5-2 Camera-laser module performance confirmation</i>	<i>68</i>
<i>Table 6-1 Angle between the flat surface and horizontal estimation</i>	<i>84</i>

Abbreviations

CMOS	Complementary metal–oxide–semiconductor
CCD	Charge-coupled device
2D	Two dimensional
3D	Three dimensional
GA	Genetic algorithm
LA	Learning automata
HT	Hough transform
RHT	Randomized Hough transform
AHT	adaptive Hough transform
POSE	Position and attitude
GigE	Gigabit Ethernet
UAV	Unmanned aerial vehicle

ABSTRACT

Industrial robots have been developed for several decades. Nowadays, robots equipped with 3D vision system are being widely deployed in increasing numbers into variety industries, including food, pharmaceutical, vehicle assembly, aerospace, etc. 3D vision system in industrial robots requires high levels of repeatability, modest in hardware, high accuracy and real-time performance. Toward this end, in this work, a 3D vision system using monocular camera and structured light is developed. The proposed 3D measurement sensor projects laser lines onto the target objects that are observed by the high-resolution camera. The measurement is initiated when the operator brings the device closer to the surface until it is within operating range. During the process, the line features are digitized by using proposed approach, the desired calculations are made, the non-conforming images are discarded, and the remaining images are used to perform the 3D measurement. The measurement system can deal with complex surface in noisy industrial environment and achieve higher specifications compared with current 3D measurement sensors. The usefulness of the proposed system has been demonstrated using real tests with accurate know-size patterns and a real inline vehicle assembly system in Korea.

Chapter 1

INTRODUCTION

1.1 Overview

In recent years, 3D imaging sensors for the acquisition of three dimensional (3D) measurements have created a considerable degree of interest for a number of applications. The miniaturization and integration of the optical and electronic components used to build them have played a crucial role in the achievement of compactness, robustness and flexibility of the sensors. First, laser reflection [1-2] and ultrasonic (or light) reflection [3-5] are popularly used in non-contact measurements. Laser emitters and receivers are needed in laser reflection-based approaches. Using laser technology has the advantage of speed; however, if the measured object has poor reflectivity, the system will work poorly or not at all. The ultrasonic technique has virtually the same problem in measuring objects with poor reflectivity. Inadequate object reflectivity is the source of non-contact measurement errors and failure. Second, newer manufacturing measurement technique involves widely studied field of laser vision, which provides far better reproducibility of measurement values than conventional measurement systems. In laser-vision systems, the principle of laser triangulation is commonly used [6-7]. Third, monocular vision, stereo vision are also developed for accurate 3D points estimated [8-11]. In vision system, image processing software plays a key role in determining the overall system performance. In order to obtain geometrical information of the target objects, the feature points need to be extracted precisely and efficiently. Many feature points detection algorithms have been developed and are commonly used such as Scale Invariant Feature Transform (SIFT) [12], Speed Up Robust Features (SURF) [13], Histogram of Oriented Gradients (HoG) [14], Shape Context [15], Maximally Stable Extremal Region (MSER) [16], Geometric Blur [17], Pyramid Histogram Embedding of Other Features (PHOW) [18]. However, these methods require a

great deal of computational time, greatly limiting the system performance and rendering them unsuitable for vehicle manufacturing systems in which the gross time of robot hand manipulation and measurement time per measurement point should be less than a few seconds.

In this work, a high-speed accurate non-contact 3D measuring structure using a combination of CMOS camera and laser is studied. Both camera and laser generator have some noises in their outputs, and like any non-contact 3D measurement sensors, they depend strongly on the target object material, color, etc., and we found that the noise in laser lines and wrong estimates of system calibration affect strongly the final results, especially in heavy noise industrial environment. This work proposes a fast, accurate, and robust 3D measurement method in industrial environments. The proposed method is carried out in three steps: adaptive detection of the laser stripes; accurate measurement system calibration using rectangle and finally intelligent 3D measurement.

1.2 Research objectives

The 3D measurement procedure proposed in this work has three main goals which distinguish it from previous works:

Accuracy: The proposed procedure is designed to be accurate. This is an indispensable requirement for machine vision applications which need to take action based on the information inferred from the images, in this case, from the 3D reconstruction calculated from the laser stripe projection.

Robustness: The proposed procedure is designed to be applied under heavy noise conditions where previous methods would not succeed.

Speed: The proposed procedure is designed to be applied under real-time constraints at a high frame rate. This makes it possible to use the 3D reconstruction during the industrial manufacturing.

Accuracy, robustness, and speed are requirements common to most applications, not only in industrial environments, which makes the proposed procedure applicable to many other types of applications.

1.3 Dissertation outline

The remainder of the thesis is organized as follows. Fundamental techniques in monocular structured light vision systems are introduced in Chapter 2. Chapter 3 presents circular object detection and tracking using dynamic fusion of color and shape. Slippage measurement using monocular structured-line light vision and its application in slip-margin feedback control gripper system is illustrated in Chapter 4. Gap and flush measurement using monocular structured multi-line light vision and its application in car manufacturing is provided in Chapter 5. In Chapter 6, Monocular structured light vision-based surface measurement is illustrated. Finally, the conclusions of the study and the recommendations discussed as future works are presented in Chapter 7.

Chapter 2

FUNDAMENTAL TECHNIQUES IN MONOCULAR VISION SYSTEMS

2.1 Introduction

Recently, structured light based-vision mounted on mobile robots has become very common for large fields of robot navigation tasks [19-21]. They provide real time accurate measurements in various fields at a fixed height above the ground plane, and enable robots to perform more confidently a wide range of tasks by fusing image data from the camera and structured light mounted on robots [22-24]. In order to effectively use the data from the camera and structured light, it is crucial to know their relative position and orientation from each other and, their relationship is vital to creating metric depth estimates. The calibration can be decomposed into two types, internal calibration parameters and external calibration parameters. The internal parameters calibration, for example the calibration intrinsic matrix of a camera [25-26], affects how the camera sensor samples the scene, the structured light calibration [27-29]. The external calibration parameters are orientation and position of the camera relative to the structured light coordinate system [30-34].

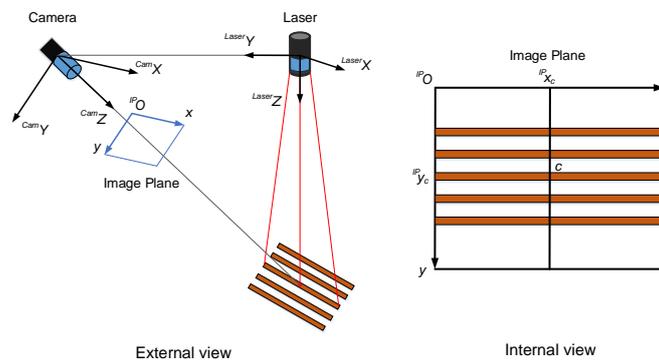


Fig. 2-1. Structured multi-line light based vision system.

In this work, we propose a method for extrinsic calibration of a camera and structured line

light, identifying the transformation from the camera coordinate system to the structured light coordinate system. There are several types of structure light, such as point light, cross line light, parallel line light. The parallel line light is chosen in this work because of its larger field of projection and giving more information about the projected surface in comparison with other types. The extrinsic calibration method employs a planar calibration pattern which is visible by the camera and the structure light. For each different pose of the planar pattern, the method constrains the extrinsic parameters by registering the line light on the planar pattern with the estimated calibration plane from the camera image. It is important also to differentiate this work from the others that at first may appear similar. There has been a great deal of work on calibration for laser scanners, which make use of the visible position of the laser point or a laser stripe by using a checkerboard. In this work, we consider an extrinsic calibration of a camera with structured multi-line light where the projected light lines are visible to the camera, as can be seen in Fig.2-1, by using only one rectangle. This calibration does not require complicated calibration patterns, so that we can use any known size rectangular feature for calibration. The simple calibration pattern has better influence in feature points extraction process than complex patterns. Moreover, the proposed method can be applied not only in extrinsic calibration of camera and structured single-line light but also structured multi-line light. The experiments show that the proposed calibration method can achieve high accurate calibration parameters at high frequency. The measurement accuracy achieves at ~ 0.0908 mm. This calibration can be applied to a very common measurement sensor such as sensor for inspection, surface estimation sensor, 3D scanner, and etc.

2.2 Camera calibration

We describe a camera using pin-hole model which is a linear model.

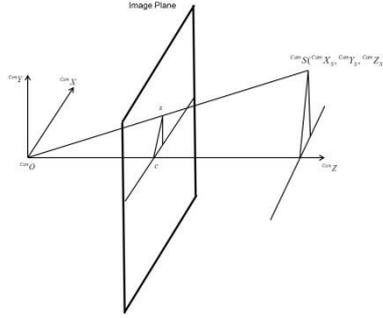


Fig. 2-2. Camera pinhole model

In the camera coordinate system, using the triangle similarity, we obtain

$$\frac{Cam X_s}{Cam X_S} = \frac{f_x}{Cam Z_S} \quad (2-1)$$

$$\frac{Cam Y_s}{Cam Y_S} = \frac{f_y}{Cam Z_S}$$

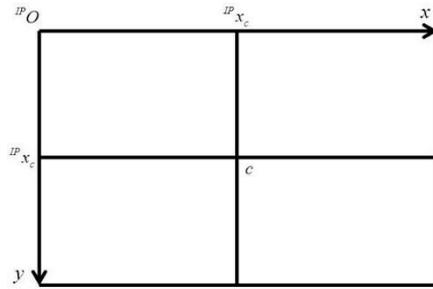


Fig. 2-3. Image plane

In the image plane,

$$IP X_s = Cam X_s + IP X_c = \frac{f_x}{Cam Z_S} Cam X_S + IP X_c \quad (2-2)$$

$$IP y_s = Cam Y_s + IP y_c = \frac{f_y}{Cam Z_S} Cam Y_S + IP y_c \quad (2-3)$$

$$w \begin{bmatrix} IP X_s \\ IP y_s \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & IP X_c \\ 0 & f_y & IP y_c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Cam X_S \\ Cam Y_S \\ Cam Z_S \end{bmatrix} \quad (2-4)$$

Let us denote $c_x = IP X_c; c_y = IP y_c; Cam M = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}; IP S = \begin{bmatrix} IP X_s \\ IP y_s \\ 1 \end{bmatrix}; Cam S = \begin{bmatrix} Cam X_S \\ Cam Y_S \\ Cam Z_S \\ 1 \end{bmatrix}$

We can rewrite equation (4) as follow

$${}^{IP}S = w^{-1} [{}_{Cam}^{IP}M_{3 \times 3}, \mathbf{0}_{3 \times 1}]^{Cam}S \quad (2-5)$$

Physical focal length F

$$f_x = F \times s_x \quad (2-6)$$

$$f_y = F \times s_y \quad (2-7)$$

with s_x and s_y are pixel pitch, the number of pixels per unit distance in the x and y direction, respectively, the pixel pitch unit is usually dot per millimeter.

$${}_{Cam}^{IP}M = \begin{bmatrix} F \times s_x & 0 & {}^{IP}x_c \\ 0 & F \times s_y & {}^{IP}y_c \\ 0 & 0 & 1 \end{bmatrix} \quad (2-8)$$

In principle, we can use any appropriately characterizes object as a calibration object, yet the practical choice is a regular pattern such as chessboard, circle, three-dimensional objects.

In this report, we used a known size chessboard.

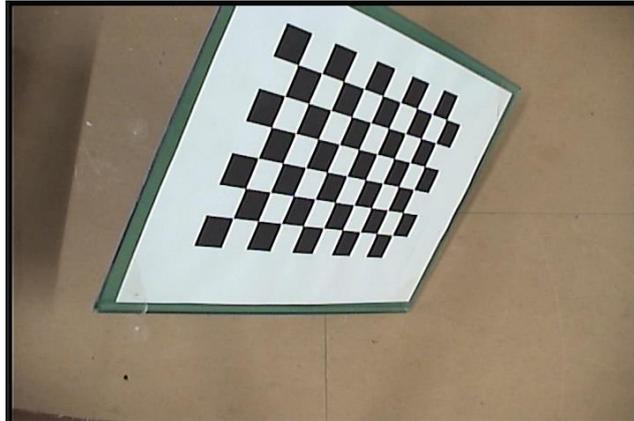


Fig. 2-4. Sample calibration image

The first task is to define and draw the chessboard corners by using Harris corner detector. The detected corners are shown in Fig.2-5 (we used only internal corners).

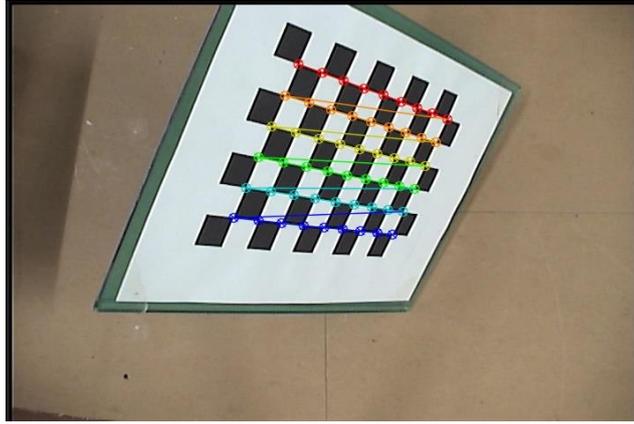


Fig. 2-5. Corners detection and drawing

How many parameters are attempted to be solved in camera calibration?

For intrinsic parameters, CCD camera intrinsic matrix which has four parameters: f_x, f_y, c_x, c_y (stay the same between views); CCD distortion parameters: Radial (k_1, k_2, k_3), tangential (p_1, p_2), and they are can be solved separately. In principle, three corners points in known pattern yielding six pieces of information are all that is needed to solve for our five distortion parameters. Thus, we need only one view of a chessboard to computer the distortion parameters (we can use much for robustness). For extrinsic parameters, we need to know where the chessboard. This will require three rotation parameters (α, β, ϖ) and three translation parameters (T_x, T_y, T_z) which we need to solved for each view.

Assuming that we have K images of the chessboard and N corners obtained in each image. K images of the chessboard provides $2NK$ constraints (we multiply by 2 because each point on the image has both an x and y coordinate). As mentioned above, we can use only one image for separate distortion calculation. We now have 4 intrinsic parameter and $6K$ extrinsic parameters. Then we have the following constraint

$$2NK \geq 6K + 4 \quad (2-9)$$

Equivalently, $K(N - 3) \geq 2$.

Because $K > 0$, thus $N > 3$. And no matter how many corners are detected on a plane, there are only four corners' worth of information per chessboard view, then, the equation can give

us only four corners of information, i.e. $K(4-3) \geq 2$ or $K \geq 2$, equivalently. This implies that 2 images of a 3-by-3 chessboard (counting only internal corner) are the minimum that could solve calibration problem. Practically, for high-quality results, we used 10 images of 6-by-9 chessboard.

The physical transformation part between object coordinate (or world coordinate) and camera coordinate is the sum of the effects of some rotation $R = {}^{Cam}_{Tg}R$ and some translation $T = {}^{Cam}_{Tg}T$

. In homogeneous coordinates, we can combine these within a single matrix as follows

$${}^{Cam}_{Tg}W = \begin{bmatrix} {}^{Cam}_{Tg}R & {}^{Cam}_{Tg}T \\ 0 & 1 \end{bmatrix} \quad (2-10)$$

Writing this out, we have:

$${}^{Cam}_{Tg}W = \begin{bmatrix} {}^{Cam}_{Tg}R_{3 \times 3} & {}^{Cam}_{Tg}T_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \quad (2-11)$$

Then the relation between image plane and target object coordinate (or world coordinate) can be expressed as follows

$${}^{IP}S = w^{-1} \begin{bmatrix} {}^{IP}_{Cam}M_{3 \times 3} & \mathbf{0}_{3 \times 1} \end{bmatrix} {}^{Cam}_{Tg}W {}^{Tg}S \quad (2-12)$$

We can choose the object plane so that $Z = 0$

$${}^{IP}S = w^{-1} {}^{IP}_{Cam}M \begin{bmatrix} r_1 & r_2 & r_3 & T \end{bmatrix} \begin{bmatrix} {}^{Tg}X_S \\ {}^{Tg}Y_S \\ 0 \\ 1 \end{bmatrix} = w^{-1} {}^{IP}_{Cam}M \begin{bmatrix} r_1 & r_2 & T \end{bmatrix} \begin{bmatrix} {}^{Tg}X_S \\ {}^{Tg}Y_S \\ 1 \end{bmatrix} \quad (2-13)$$

Let denote $H = w^{-1} {}^{IP}_{Cam}M \begin{bmatrix} r_1 & r_2 & T \end{bmatrix}$ a homography matrix that maps a planar object's points onto the imager. Obviously, H is now a 3×3 matrix.

$${}^{IP}S = H {}^{Tg}S \quad (2-14)$$

Reading off these equations, we obtain

$$\begin{aligned} h_1 &= w^{-1} {}^{IP}_{Cam}M r_1 & \text{or} & & r_1 &= w {}^{IP}_{Cam}M^{-1} h_1 \\ h_2 &= w^{-1} {}^{IP}_{Cam}M r_2 & \text{or} & & r_2 &= w {}^{IP}_{Cam}M^{-1} h_2 \\ h_3 &= w^{-1} {}^{IP}_{Cam}M T & \text{or} & & T &= w {}^{IP}_{Cam}M^{-1} h_3 \end{aligned} \quad (2-15)$$

Rotation vector are orthogonal to each other by construction, thus r_1 and r_2 are orthogonal.

Since the scale is extracted it follows that r_1 and r_2 are orthonormal, i.e. the rotation vector dot product is 0 and the magnitudes are equal. We have the following constraints

$$r_1^T r_2 = 0 \quad (2-16)$$

$$\|r_1\| = \|r_2\| \quad \text{or} \quad r_1^T r_1 = r_2^T r_2 \quad (2-17)$$

Substituting (2-15) into (2-16) and (2-17), we have

$$h_1^T {}_{Cam}^{IP} M^{-T} {}_{Cam}^{IP} M^{-1} h_2 = 0 \quad (2-18)$$

$$h_1^T {}_{Cam}^{IP} M^{-T} {}_{Cam}^{IP} M^{-1} h_1 = h_2^T {}_{Cam}^{IP} M^{-T} {}_{Cam}^{IP} M^{-1} h_2 \quad (2-19)$$

We set $B = {}_{Cam}^{IP} M^{-T} {}_{Cam}^{IP} M^{-1}$, writing this out, we have

$$B = {}_{Cam}^{IP} M^{-T} {}_{Cam}^{IP} M^{-1} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{12} & B_{22} & B_{23} \\ B_{13} & B_{23} & B_{33} \end{bmatrix} = \begin{bmatrix} \frac{1}{f_x^2} & 0 & \frac{-c_x}{f_x^2} \\ 0 & \frac{1}{f_y^2} & \frac{-c_y}{f_y^2} \\ \frac{-c_x}{f_x^2} & \frac{-c_y}{f_y^2} & \frac{c_x^2}{f_x^2} + \frac{c_y^2}{f_y^2} + 1 \end{bmatrix}$$

Using the B -matrix, both constraints have the general form $h_i^T B h_j$ in them. Let's multiply this out to see what the components are. Because B is symmetric, it can be written as one six-dimensional vector dot product. Arranging the necessary elements of B into the new vector b , we have:

$$h_i^T B h_j = v_{ij}^T b = \begin{bmatrix} h_{i1} h_{j1} \\ h_{i1} h_{j2} + h_{i2} h_{j1} \\ h_{i2} h_{j2} \\ h_{i3} h_{j1} + h_{i1} h_{j3} \\ h_{i3} h_{j2} + h_{i2} h_{j3} \\ h_{i3} h_{j3} \end{bmatrix}^T \begin{bmatrix} B_{11} \\ B_{12} \\ B_{22} \\ B_{13} \\ B_{23} \\ B_{33} \end{bmatrix}^T \quad (2-20)$$

We can rewrite our two constraints as follows

$$\begin{bmatrix} v_{12}^T \\ (v_{11} - v_{22})^T \end{bmatrix} b = 0 \quad (2-21)$$

If we collect K images of chessboards together, then we can stack K of these equations together:

$$Vb = 0 \quad (2-22)$$

where V is a $2K$ -by-6 matrix. when $K \geq 2$ then this equation can be solved for our b .

The camera intrinsic are then pulled directly out of our closed-form solution for the B -matrix:

$$f_x = \sqrt{\lambda / B_{11}} \quad (2-23)$$

$$f_y = \sqrt{\lambda B_{11} / B_{11} B_{22} - B_{12}^2} \quad (2-24)$$

$$c_x = -B_{13} f_x^2 / \lambda \quad (2-25)$$

$$c_y = (B_{12} B_{13} - B_{11} B_{23}) / (B_{11} B_{22} - B_{12}^2) \quad (2-26)$$

$$\lambda = B_{33} - (B_{13}^2 + c_y (B_{12} B_{13} - B_{11} B_{23})) / B_{11} \quad (2-27)$$

The extrinsic parameters

$$\begin{aligned} r_1 &= w_{Cam}^{IP} M^{-1} h_1 \\ r_2 &= w_{Cam}^{IP} M^{-1} h_2 \\ r_3 &= r_1 \times r_2 \\ T &= w_{Cam}^{IP} M^{-1} h_3 \end{aligned} \quad (2-28)$$

2.3 Extrinsic calibration of camera and structured light by using a rectangle

This section provides how to solve the extrinsic camera and multi-line laser calibration problem by using a rectangle, as can be seen in Fig. 2-1. Suppose that the camera is perfectly calibrated by using approach in section 2.2. We first proposed a solution to estimate the relation between camera coordinate system and object coordinate system, followed by a solution to find the relative pose including position and orientation of the laser with respect to the object coordinate system. Finally, the extrinsic calibration of camera and structured multi-line laser light parameters are refined.

There are four coordinate systems in the model: camera, image, laser and the rectangle in one way, or object coordinate systems in another.

The camera coordinate system Σ_{Cam} is a coordinate system permanently fixed to the camera with its origin ${}^{Cam}O$ at the camera origin and the ${}^{Cam}Z$ -axis coincides with the camera optical axis, pointing to the light of sight, follows the standard way to decide camera coordinate system.

The image plane is perpendicular with the ${}^{Cam}Z$ -axis. The image center is on the ${}^{Cam}Z$ -axis and the distance between the image center and the camera origin ${}^{Cam}O$ is equal to focal length. In the image coordinate system Σ_{IP} , the image origin ${}^{IP}O$ is at the top left of the image plane, x-axis and y-axis parallel to the image width and height direction, respectively as shown in Fig.1.

The laser coordinate system Σ_{Laser} is defined with an origin ${}^{Laser}O_L$ at the center of the laser generator. The laser generates N laser planes. Suppose that we place a virtual plane (P) perpendicular with the ${}^{Laser}Z$ -axis and the distance between the origin ${}^{Laser}O_L$ and plane (P) is d_λ , then we obtain N equidistant, equal and parallel line segments on (P). The line segments length is 2λ and the distance between two consecutive line segments is ε . Fig. 2-1 shows an example of structured multi-line laser with $N=5$. The generated center plane is assumed to lie on the ${}^{laser}Z$ ${}^{laser}X$ plane as can be seen in Fig. 2-1.

The plane contains rectangle is placed randomly so that the rectangle is visible to both camera and laser. The rectangle coordinate Σ_{Tg} origin is the top left of the rectangle and the ${}^{Tg}Z$ axis is perpendicular to the rectangle plane, ${}^{Tg}X$ axis and ${}^{Tg}Y$ axis coincide with the rectangle edges as showed in Fig. 3.

In real cases, the camera can show lens distortion modeled as a 5-vector parameter containing radian and tangent distortion coefficients. We assume that in the remainder of the paper the lens distortions are already eliminated in the images, or the camera has no significant lens

distortion so that we can neglect the distortion completely. The image center is the pixel at which optical axis penetrates the camera sensor.

The physical transformation part between laser coordinate and camera coordinate is the sum of the effects of some rotation ${}_{Laser}^{Cam}R$ and some translation ${}_{Laser}^{Cam}T$.

Our goal in this paper is to develop an approach to find the extrinsic parameters ${}_{Laser}^{Cam}R$ and ${}_{Laser}^{Cam}T$ which define the relative position including position and orientation of the laser with respect to the camera coordinate system.

In front of the calibration system, a planar pattern consisting of the rectangle, a calibration plane, is placed. The calibration plane and rectangle are visible to both the camera and the laser. Without losing of generality, the calibration plane is ${}^{Tg}Z=0$ in the object coordinate system. With an arbitrary position and attitude of the calibration plane, two pictures are taken with and without laser lines alternately. Then we change the position and attitude and take two pictures again by the same way. This step is repeated N_c times.

2.3.1 Camera and rectangle object calibration

In this section, the laser-turned-off-images are processed. The relative pose between camera and the rectangle is defined by some rotation ${}_{Tg}^{Cam}R$ and some translation ${}_{Tg}^{Cam}T$ which can be represented as following

$${}_{Tg}^{Cam}W = [{}_{Tg}^{Cam}R \quad {}_{Tg}^{Cam}T] = [r_1 \quad r_2 \quad r_3 \quad t] \quad (2-29)$$

Writing this out, we have:

$${}_{Tg}^{Cam}W = \begin{bmatrix} {}_{Tg}^{Cam}R_{3 \times 3} & {}_{Tg}^{Cam}T_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \quad (2-30)$$

Then the relation between image plane and target object coordinate can be expressed as follows

$${}^{IP}S = W^{-1} [{}_{Cam}^{IP}M_{3 \times 3}, \mathbf{0}_{3 \times 1}] {}_{Tg}^{Cam}W {}^{Tg}S \quad (2-31)$$

The calibration plane ${}^{Tg}Z = 0$, so we can rewrite equation (5) for the calibration plane

$${}^{IP}S = W^{-1} {}_{Cam}^{IP}M [r_1 \ r_2 \ r_3 \ t] \begin{bmatrix} {}^{Tg}X_S \\ {}^{Tg}Y_S \\ 0 \\ 1 \end{bmatrix} = W^{-1} {}_{Cam}^{IP}M [r_1 \ r_2 \ t] \begin{bmatrix} {}^{Tg}X_S \\ {}^{Tg}Y_S \\ 1 \end{bmatrix} \quad (2-32)$$

And the problem to be solved is to determine the relative position between camera coordinate and target object coordinate, i.e. we need to find the rotation ${}_{Tg}^{Cam}R$ and translation ${}_{Tg}^{Cam}T$.

In order to estimate the relative position between the calibration plane and the camera, firstly we need to find the image coordinate of the rectangle vertices ${}^{IP}p_i$ with $i = \{1, 2, 3, 4\}$. Then the four vertices image coordinates are used to calculate the four vertices camera coordinates ${}^{Cam}P_i$ which will be used for calculation of the relative between camera coordinate and target coordinate. In this paper, the Harris corner detector is used to estimate the rectangle four vertices.

The transformation from camera coordinate to target coordinate is defined as following

$${}_{Tg}^{Cam}W = [{}_{Cam}^{Tg}R \quad {}_{Cam}^{Tg}T] = [{}_{Cam}^{Tg}r_1 \quad {}_{Cam}^{Tg}r_2 \quad {}_{Cam}^{Tg}r_3 \quad {}_{Cam}^{Tg}t] \quad (2-33)$$

The relative pose between the target object and camera coordinate can be calculated by using Least Square method

$$\begin{bmatrix} {}_{Cam}^{Tg}r_1, {}_{Cam}^{Tg}r_2, {}_{Cam}^{Tg}t \end{bmatrix} = {}_{Tg}P {}^{Cam}P^T ({}^{Cam}P {}^{Cam}P^T)^{-1} \quad (2-34)$$

$${}_{Cam}^{Tg}r_3 = {}_{Cam}^{Tg}r_1 \times {}_{Cam}^{Tg}r_2 \quad (2-35)$$

2.3.2 Extrinsic calibration of the structured multi-line light laser and calibration plane

In this section, the relative pose between the laser and calibration plane is considered. The laser points are extracted by subtracting two images captured in sequence that the laser is turned on in the first frame while the laser is turned off in the second frame. This image containing laser points is then converted from RGB format to Gray format. Then, thresholding method is

applied to obtain sets of laser points which have much higher intensities than the background.

Finally, the laser lines are estimated using Least Square algorithm.

In the laser coordinate, the laser plane (LP_i) equation can be represented as following

$$(0 \quad n_i \quad t_i) \begin{bmatrix} \text{Laser } X \\ \text{Laser } Y \\ \text{Laser } Z \end{bmatrix} = 0 \quad (2-36)$$

The laser line ${}^{\text{Laser}}_P l_i$ projected on the virtual plane (P) is the intersection of laser plane (LP_i) and the virtual plane (P)

$${}^{\text{Laser}}_P l_i = (LP_i) \cap (P) \quad (2-37)$$

The relative pose between the laser and calibration plane is defined by some rotation ${}^{\text{Laser}}_{T_g} R$ and some translation ${}^{\text{Laser}}_{T_g} T$ which can be represented as following

$${}^{\text{Laser}}_{T_g} W = [{}^{\text{Laser}}_{T_g} R \quad {}^{\text{Laser}}_{T_g} T] = [{}^{\text{Laser}}_{T_g} r_1 \quad {}^{\text{Laser}}_{T_g} r_2 \quad {}^{\text{Laser}}_{T_g} r_3 \quad {}^{\text{Laser}}_{T_g} \mathbf{t}] \quad (2-38)$$

Writing this out, we have:

$${}^{\text{Laser}}_{T_g} W = \begin{bmatrix} {}^{\text{Laser}}_{T_g} R_{3 \times 3} & {}^{\text{Laser}}_{T_g} T_{3 \times 1} \\ 0_{1 \times 3} & 1 \end{bmatrix} \quad (2-39)$$

We set new coordinate system ${}^\Sigma P$, the origin is at the top left of the laser line image on (P), x-axis and y-axis parallel to the ${}^{\text{Laser}} X$ and ${}^{\text{Laser}} Y$ direction, respectively. Suppose that there is a point ${}^{\text{Laser}} Q$ in laser coordinate system, and its projection on the virtual plane (P) is ${}^P q$, we obtain the following equation

$${}^P q = hM {}^{\text{Laser}} Q \quad (2-40)$$

With $M = \begin{bmatrix} d_\lambda & 0 & \lambda \\ 0 & d_\lambda & 2\varepsilon \\ 0 & 0 & 1 \end{bmatrix}$ and h is scale.

If point Q lies on the calibration plane, we can rewrite (19) as following

$${}^p q = hM \begin{bmatrix} {}^{Laser}_{Tg} r_1 & {}^{Laser}_{Tg} r_2 & {}^{Laser}_{Tg} t \end{bmatrix} \begin{bmatrix} {}^{Tg} X_Q \\ {}^{Tg} Y_Q \\ 1 \end{bmatrix} \quad (2-41)$$

Let denote $\Delta = [\Delta_1 \ \Delta_2 \ \Delta_3] = hM \begin{bmatrix} {}^{Laser}_{Tg} r_1 & {}^{Laser}_{Tg} r_2 & {}^{Laser}_{Tg} t \end{bmatrix}$. Since the rotation vectors ${}^{Laser}_{Tg} r_1$

and ${}^{Laser}_{Tg} r_2$ are orthonormal, we obtain two constraints

$$\Delta_1^T \Omega \Delta_2 = 0 \quad (2-42)$$

$$\Delta_1^T \Omega \Delta_1 = \Delta_2^T \Omega \Delta_2 \quad (2-43)$$

$$\text{where } \Omega = \begin{bmatrix} \frac{1}{d_\lambda^2} & 0 & \frac{-\lambda}{d_\lambda^2} \\ 0 & \frac{1}{d_\lambda^2} & \frac{-2\varepsilon}{d_\lambda^2} \\ \frac{-\lambda}{d_\lambda^2} & \frac{-2\varepsilon}{d_\lambda^2} & \frac{\lambda^2 + 4\varepsilon^2}{d_\lambda^2} + 1 \end{bmatrix}$$

Because Ω is symmetric, the two constraints can be rewritten as

$$\begin{bmatrix} \eta_{12}^T \\ (\eta_{11} - \eta_{22})^T \end{bmatrix} \Phi = 0 \quad (2-44)$$

$$\text{where } \eta_{ij} = \begin{bmatrix} \Delta_{i1} \Delta_{j1} \\ \Delta_{i1} \Delta_{j2} + \Delta_{i2} \Delta_{j1} \\ \Delta_{i2} \Delta_{j2} \\ \Delta_{i3} \Delta_{j1} + \Delta_{i1} \Delta_{j3} \\ \Delta_{i3} \Delta_{j2} + \Delta_{i2} \Delta_{j3} \\ \Delta_{i3} \Delta_{j3} \end{bmatrix},$$

$$\Phi = [\Omega_{11} \ \Omega_{12} \ \Omega_{22} \ \Omega_{13} \ \Omega_{23} \ \Omega_{33}]^T.$$

If we collect N_c images of laser lines together, N_c of equations can be stacked together

$$\eta \Phi = 0 \quad (2-45)$$

where η is a $2N_c$ -by-6 matrix, usually we chose $N_c \geq 3$ then equation (2-21) can be solved.

The extrinsic parameters of the multi-line laser light and the calibration plane can be calculated as following:

$$\begin{bmatrix} {}^{Laser}_{Tg} r_1 & {}^{Laser}_{Tg} r_2 & {}^{Laser}_{Tg} t \end{bmatrix} = h^{-1} M^{-1} \Delta \quad (2-46)$$

$${}^{Laser}_{Tg} r_3 = {}^{Laser}_{Tg} r_1 \times {}^{Laser}_{Tg} r_2 \quad (2-47)$$

2.3.3 Extrinsic calibration of the camera and structured multi-line laser light

The extrinsic parameters of camera and the structured multi-line laser light are finally calculated as follows

$${}^{Laser}_{Cam}W = {}^{Laser}_{Tg}W {}^{Tg}_{Cam}W \quad (2-48)$$

where ${}^{Laser}_{Cam}W = [{}^{Laser}_{Cam}R \quad {}^{Laser}_{Cam}T]$ with ${}^{Laser}_{Cam}T$ is translation vector, ${}^{Laser}_{Cam}R$ is the Euler rotation in sequence of $Z \rightarrow X \rightarrow Y$ and the corresponding rotation angle is $Yaw \rightarrow Pitch \rightarrow Roll$ in one way, or $\gamma \rightarrow \alpha \rightarrow \beta$ in another way, then the rotation matrix ${}^{Laser}_{Cam}R$ will be

$${}^{Laser}_{Cam}R = \begin{bmatrix} \cos(\beta) & 0 & -\sin(\beta) \\ 0 & 1 & 0 \\ \sin(\beta) & 0 & \cos(\beta) \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & \sin(\alpha) \\ 0 & -\sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} \cos(\gamma) & \sin(\gamma) & 0 \\ -\sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2-49)$$

2.4 Experimental results

In this section, we verify the accuracy of the proposed calibration system through computer simulations and real experiments using a camera and a five-line laser generator. The test platform was implemented in C/C++.

2.4.1 Computer simulations

In the computer simulations, the placement of the camera relative to the laser is simulated by the camera position and orientation in laser coordinate. The camera resolution is set as 2590×1942 and the focal length is $16mm$. The calibration pattern plane is a checker board containing a pattern square $20mm \times 20mm$. The position of the plane is chosen properly so that the pattern can appear entirely on the image plane. The image points are added Gaussian noise with mean 0 and standard deviation 0.5 pixel. Uniform noise of $\pm 5mm$ is added to into the laser points. We compare the estimated extrinsic parameters with the setup known parameters. The error for camera orientation is measured by computing the angle between the estimate and the true

orientation, and the error for camera position is measured by computing the distance between the estimate and the true camera position.

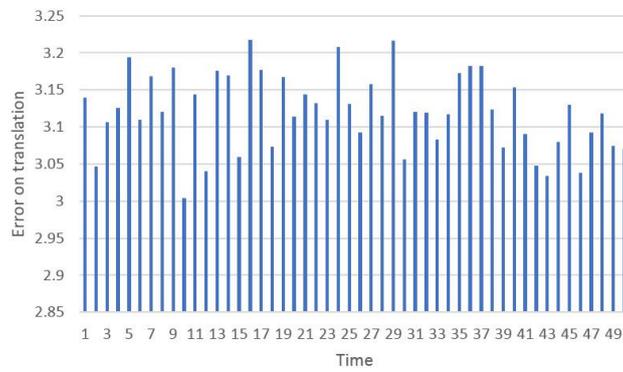


Fig. 2-6. The camera translation error

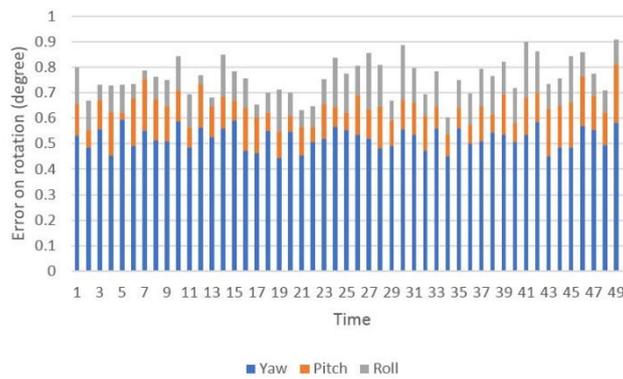


Fig. 2-7. The camera orientation errors.

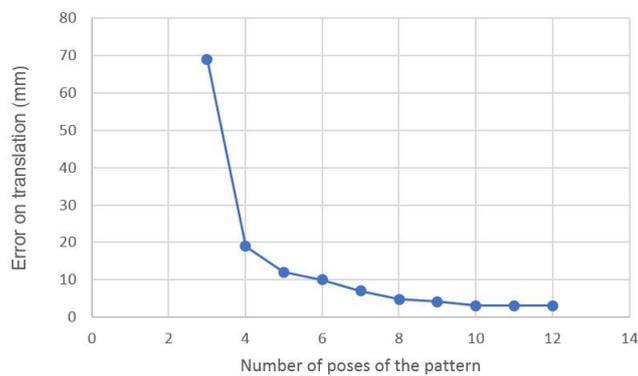


Fig. 2-8. Translation error versus the number of calibration pattern poses using for estimation.

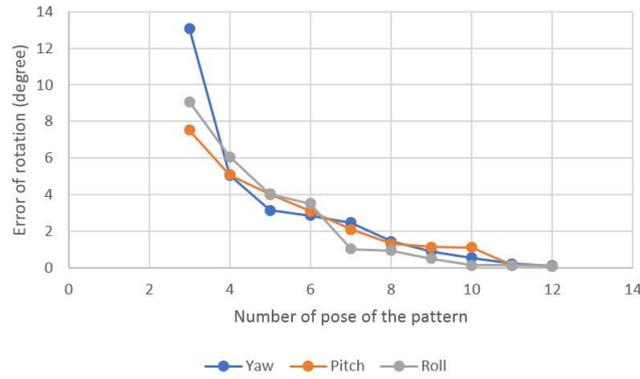


Fig. 2-9. Rotation error versus the number of calibration pattern poses using for estimation.

Performance w.r.t the changing of checkerboard pose. This experiment is performed for different pose of the checkerboard to examine its influence in the proposed calibration performance. In the experiment, we run 50 independent trials and then compare the average result with the known parameters. For each trial, 10 checkerboard poses are used. The checkerboard pose is generated as follows: The calibration pattern plane is initially parallel to the image plane; a rotation axis is randomly selected on the plane; the rotation angle θ is chosen randomly from 10° to 80° ; the plane is rotated around the rotation axis with angle θ ; the plane position is chosen properly so that the pattern can appear entirely on the image plane. The camera orientation and position errors are shown in Fig.2-6 and Fig.2-7, respectively. According to the experiment results, the maximum error in the camera position is less than 1° and the maximum error in camera position is less than 3.25 mm when the calibration pattern pose is changed.

Performance w.r.t the number of checkboard poses. This experiment evaluates the impact of the number of checkerboard poses on the calibration performance. The number of poses is varied from 3 to 12. For each experiment, the checkerboard poses are independently conducted with $\theta = 60^\circ$. The translation error and rotation error are shown in Fig.2-8 and Fig.2-9, respectively. The experiment result shows that the error of translation and rotation decreases

while increasing the number of pattern pose. The best result is obtained when 12 poses of calibration pattern as can be seen in Fig. 2-8 and Fig.2-9.

2.4.2 Real data

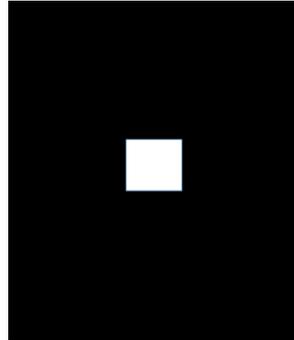


Fig. 2-10. Calibration pattern.

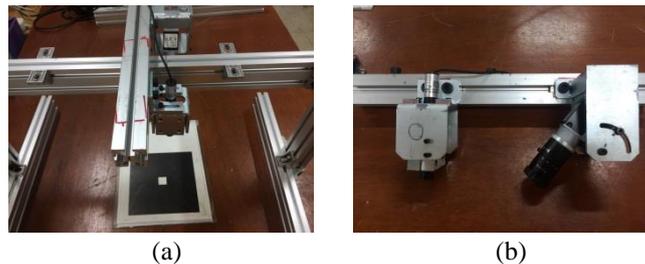


Fig. 2-11. Real extrinsic calibration testbed. (a) Extrinsic calibration of camera and structured multi-line laser using rectangle, (b) camera and structured multi-line laser system.



Fig. 2-12. Extrinsic calibration of the camera and calibration plane images. (a) original image of rectangle object taken by the camera when laser is turn off, (b) rectangle corners detection.

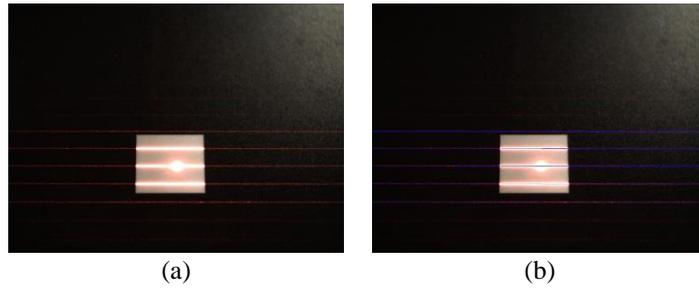


Fig. 2-13. Extrinsic calibration of the structured multi-line light laser and calibration plane images. (a) original image of laser lines projected on the calibration plane taken by the camera, (b) laser lines detection.

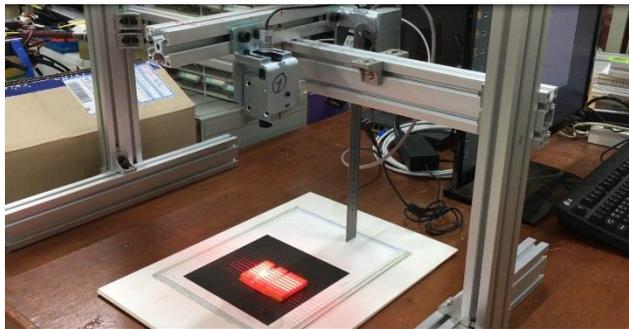


Fig. 2-14. Real 3D reconstruction and gap/flush measurement system using estimated ${}^{Laser}T_{Cam}$ and ${}^{Laser}R_{Cam}$.

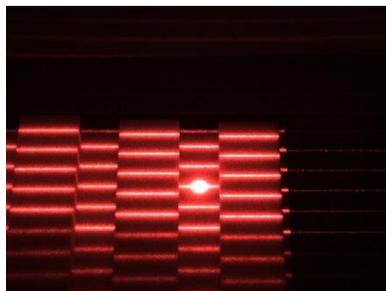


Fig. 2-15. Original image of flat panels captured by the camera.

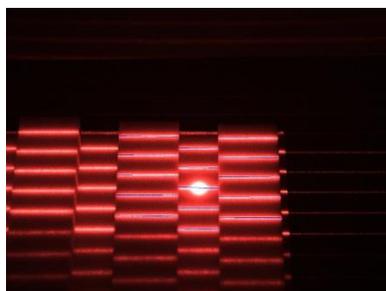


Fig. 2-16. Laser lines detection.

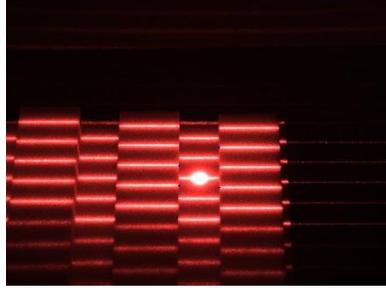


Fig. 2-17. Gap/flush measurement points.

The proposed method is tested on the real camera-laser module consisting of a focus lens, filter, a GigE camera with a CMOS image sensor and a 5-line laser. The real calibration pattern and the camera-laser module are shown in Fig. 2-10 and Fig. 2-11., respectively. First, we run proposed method to estimate the extrinsic parameter ${}_{Cam}^{Tg}T$ and ${}_{Cam}^{Tg}R$ by using a rectangle, as can be seen in Fig.2-12. Fig. 2-12 (a) shows the image of the rectangle pattern captured by the camera. Detected rectangle vertices are shown in Fig.2-12 (b) and marked by blue crosses. Second, the extrinsic parameter ${}_{Tg}^{Laser}T$ and ${}_{Tg}^{Laser}R$ are estimated as shown in Fig.2-13. In Fig.2-13(b), the detected laser lines are marked in blue color. Finally, ${}_{Cam}^{Laser}T$ and ${}_{Cam}^{Laser}R$ are calculated by using equation (30). The estimated ${}_{Cam}^{Laser}T$ and ${}_{Cam}^{Laser}R$ are used to reconstruct and measure gap/ flush between known-size flat panels. Fig.2-15 shows the original image of flat panels captured by camera. In Fig.2-16 and Fig.2-17, the detected laser lines and gap/flush measurement points are marked in blue color lines and crosses, respectively. The average gap/flush measurement error is around 0.0908 mm.

2.5 Chapter summary

In this section, we present the extrinsic calibration of camera and structured light by using a plane containing a rectangle. The proposed calibration method requires only one known-size rectangle on the calibration plane and a few poses of planar pattern which is visible for both the camera and the structured light range, and then a geometric constraint on the extrinsic

camera parameters is imposed. In order to obtain the extrinsic calibration parameters of camera and multi-line structure light, first, a solution to estimate the relation between camera coordinate and object coordinate is proposed. Second, a solution to find the relative poses including position and orientation of the laser with respect to the object coordinate system. Finally, the extrinsic calibration of camera and structured multi-line laser light parameters are refined.

In next sections, industrial systems applied monocular camera calibration and monocular camera and structured light calibration to measure object position and attitude will be presented.

Chapter 3

CIRCULAR OBJECT DETECTION AND TRACKING USING DYNAMIC FUSION OF COLOR AND SHAPE

3.1 Introduction

Circle detection is fundamental to many computer vision applications such as eye detection and tracking [34-37] automatic product inspection in industrial vision applications [38], circular traffic sign recognition [39-41] and ball detection in games [42-43]. The common main challenges of circle detection are changes in viewpoint, perspective, and noise, and time constraints in particular. This problem has received considerable attention, mainly in an attempt to improve the detection accuracy and reduce the amount of time consumed. Hough transform (HT)-based [44], approaches are commonly used. This approach is simple, yet it has massive computation and memory requirements, requires a great deal of time and is sensitive to noise. In an effort to address these problems, many more complicated methods based on HT were proposed. One efficient approach is randomized Hough transform (RHT) [45], which improves time efficiency and reduces the storage requirements of the original HT by avoiding the implementation of a computationally expensive voting process for all edge image pixels, which includes exploiting their geometric properties. The direction information of the edge points was used in adaptive HT (AHT) [46] and modified HT [47] and HT was also combined with hypothesis filtering [48] to detect circles. A novel approach based on the RHT, the iterative randomized Hough transform (IRHT), was proposed [49]. In particular, it achieves better results with complex images and noisy environments. Nonconventional HT principle-based approaches were also proposed, including genetic algorithm (GA)-based approaches [50-52] or circle detection using learning automata (LA) [53]. These methods detect circles quite accurately; however, although they can detect a circle in a multiple-circle case, they generally require a lot of time and are too computationally expensive. The computational complexity of

these algorithms is derived from their attempts to deal with edge detected image pixels whose numbers increase significantly in nature and complicated environments. Visual tracking is an active and developing field, and many trackers have been developed. Elaborate methods such as tracking using Eigen space matching techniques [54], convolving images with feature detectors [55] or the recent method based on sparse coding [56-58] in which static images and obtained images are represented by sparse coding and the target is tracked by the sparse coding comparison are too computationally complex and therefore far too expensive for a real-time system. Our goal in this paper is to find a simple, fast tracker for real-time systems that runs with inexpensive cameras and does not require calibrated lenses. The main contributions of this work include the proposal of a new method for robust circular object tracking based on feature fusion, which can be performed at a frame rate of 30 Hz.

3.2 Circular object detection and tracking

The proposed tracking strategy includes three main steps to perform the object chasing. First, the features are extracted from input frames. Then, in each feature space, the estimator will find the predicted position of a target. Finally, the dynamic fusing of information from different feature spaces will validate the target position

3.2.1 Feature Space

3.2.1.1 Color Space

Theo Gevers and W.M. Smeulders presented a comparison between different color models for color-based object recognition. The choice of color model depends on their robustness against varying illumination and changes in object surface orientation. In this work, we need to use a color model that is robust to a viewpoint change. The HSV (Hue, Saturation, Value) color space is most appropriate because the HSV color model has two of its own strong principal advantages. Firstly, the H and S components are related to the way in which a human

perceives color; therefore, the colors in this model can be clearly defined by human perception. Secondly, the H component (color information) is disassociated from the V component (the brightness of color) and S component (how concentrated the color is). Thus, the tracker can trust the H component for any purpose of object detection and recognition. In this paper, we use the hue image from the H channel in the HSV space for future histogram calculation.

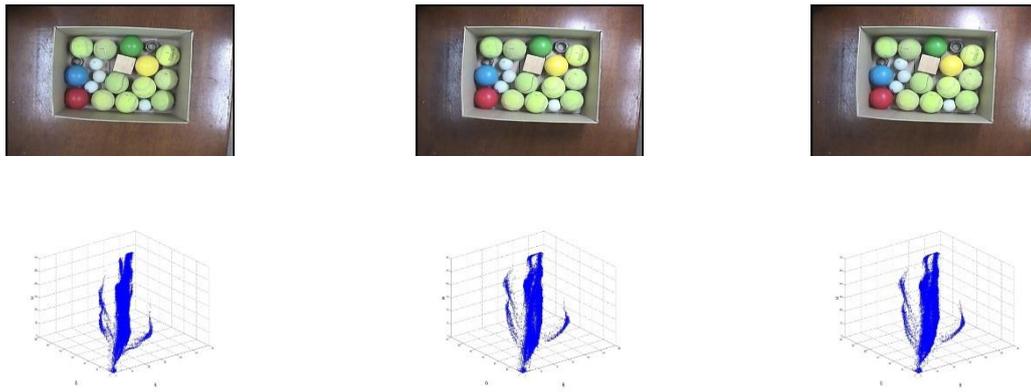


Fig. 3-1 Example of color illumination changes over time.

3.2.1.2 Shape Feature Space

To achieve real-time performance, a fast circle detector that has high detection rates and produces accurate results was proposed. The method makes use of the continuous set of edge points, or contour segments, instead of the huge edge points as traditional detectors do. First, the detector is based on isoperimetric inequality to extract complete circles, which shows a significant property of the circle compared with other curves, by considering the set of detected contours in the image as a complete circle candidate set. The isoperimetric quotient of each candidate indicates whether or not it is presented on the image. The proposed algorithm then computes the contour in a given image, which is then converted into contour segments at high curvature points. By calculating the reinforcement of each virtual circle generated from three selected points on the contour segment, we can define which circles are really represented in the image.

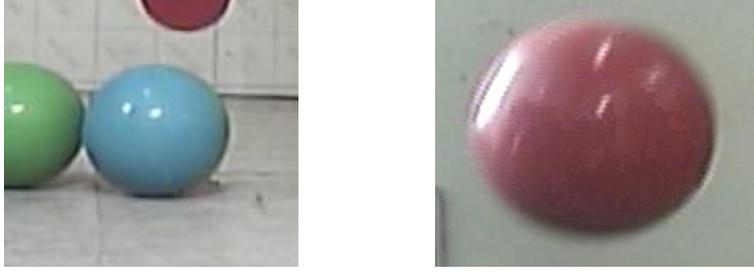


Fig. 3-2 Example of shape change.

3.2.2 Feature Estimator

3.2.2.1 Color Space Estimator

The HSV color space separates hue (color) from saturation and brightness. In this paper we take the 1D histogram from the H (hue) channel in the HSV space. Initially, the model H histogram, M_d , of the target object is sampled inside the tracking window. During operation, M_d is used as a model, or, in other words, a look-up table, to convert incoming image pixels into a corresponding probability for the model image. Let $b_{x,y}$ denote the probability of pixel (x, y) after the histogram back-projection.

For all pixels inside the tracking window

- i. Compute the zero moment

$$M_{00} = \sum_x \sum_y b_{x,y} \quad (3-1)$$

- ii. Find the first moment for x and y

$$M_{10} = \sum_x \sum_y x b_{x,y} \quad (3-2)$$

$$M_{01} = \sum_x \sum_y y b_{x,y} \quad (3-3)$$

- iii. The possible object localization is determined as follows

$$x_c = \frac{M_{10}}{M_{00}} \quad ; \quad y_c = \frac{M_{01}}{M_{00}} \quad (3-4)$$

Record (x_c, y_c) and M_{00} for the potential tracking window in the next frame, where the window size is set to

Window width:

$$s = 2 * \sqrt{\frac{M_{00}}{256}} \quad (3-5)$$

Window length:

$$l = 1.2 * s \quad (3-6)$$

3.2.2.2 Shape Space Estimator

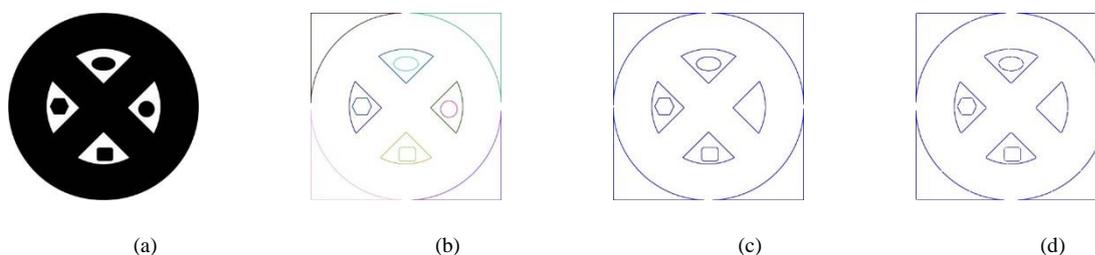


Fig. 3-3 (a) A sample image (906×918). (b) Contour detection. Each color represents a different contour. A total of 16 contours are extracted in 4 ms. (c) The remaining contours in vector \mathbf{T} after complete circle detection. (d) A total of 60 contour segments are approximated.

In circle detection, the expectation is to find triplets of (x_c, y_c, R) which completely describe a circle's center x-axis coordinate, y-axis coordinate and radius, respectively. To overcome the complex computations and time-consuming aspect of circle detection in DTS, a method based on the contour CCDA is proposed. The general idea is to detect the contours of an image, extract the complete circles and convert the remaining into a set of circle candidates, generate virtual circles based on the candidates and then validate the circles that are presented in the image. A general outline of the proposed algorithm is presented in Algorithm 1 and each step will be described in detail in the following sections.

Algorithm 1. Steps of CCDA.

- 1) Detect contours and extract complete circle by using isoperimetric inequality.
- 2) Convert the remaining contours into contour segments and make circle candidates.

- 3) Join the circle candidates to detect remaining circles.
- 4) Output the valid circles.

Complete Circle Detection

The first step of the CCDA is to detect contours in the image. Each contour, ct_i , is stored in a contour vector, $CT = \{ct_1, ct_2, \dots, ct_{N_1}\}$, then complete circles are extracted from CT to generate two contour vectors, $C = \{c_1, c_2, \dots, c_{N_{11}}\}$, which contains complete circle contours, and $T = \{t_1, t_2, \dots, t_{N_{12}}\}$, which contains the remaining contours. N_1 , N_{11} , and N_{12} represent the number of original detected contours, complete circular contours, and remaining contours, respectively, where $N_1 = N_{11} + N_{12}$. To achieve this, a well-known contour detection algorithm derived from [24] is first employed. Complete circles are then detected in the contour image according to the isoperimetric inequality in Algorithm 2.

Algorithm 2. Steps for complete circle detection based on isoperimetric inequality.

- 1) Utilize the result of reference paper [29] to calculate the square of each contour area in the contour vector, CT .

Let n_i , S_i , and P_i denote the number of edge points in the boundary, ct_i , the square of the region inside the boundary, ct_i , and the perimeter of the boundary, ct_i , respectively and let (x_j, y_j) denote the coordinates of point pt_j .

For all edge points, $pt_j (j \in [0, n_i - 1])$, on a boundary, ct_i :

$$S_i = \frac{1}{2} \sum_{j=1}^{n_i} x_{j-1}y_j - x_jy_{j-1} \quad (3-7)$$

- 2) Compute the perimeter of each boundary.

For all edge points, $pt_j (j \in [0, n_i - 1])$, on the boundary, ct_i :

$$P_i = \sum_{j=1}^{n_i} \sqrt{(x_j - x_{j-1})^2 + (y_j - y_{j-1})^2} \quad (3-8)$$

where pt_n is an additional point to close the polygon, $pt_n = pt_0$.

3) Compute the ratio between the candidate area and perimeter.

It is well known that, in Euclidean geometry, the isoperimetric inequality show the privileged roles of the Euclidean circles. The isoperimetric inequality states, for the length P of a closed curve and the area S of the planar region that it encloses, that

$$4\pi S \leq P^2$$

and that equality holds if and only if the curve is a circle. However, in image plane, the isoperimetric inequality changes according to the circle definition. In this paper, according to our experimental results, we pre-set a threshold isi to distinguish complete circles from other closed curves.

If a closed curve has $S / P^2 \geq isi$, it is determined as a complete circle (usually the threshold is set to 0.073).

4) Find the target position.

$$x_{c_i} = \frac{1}{n_i} \sum_{j=1}^{n_i} x_j; \quad y_{c_i} = \frac{1}{n_i} \sum_{j=1}^{n_i} y_j \quad (3-9)$$

$$R_i = \sqrt{\frac{2 \times S_i}{P_i}} \quad (3-10)$$

where $(x_{c_i}, y_{c_i}), R_i$ are the circular contour, complete circle, center and radius, respectively.

Store all complete circles in the contour vector, $C = \{c_1, c_2, \dots, c_{N_{11}}\}$, and the remaining contours in vector $T = \{t_1, t_2, \dots, t_{N_{12}}\}$.

Contour Segments Detection

Having obtained a contour vector $T = \{t_1, t_2, \dots, t_{N_{12}}\}$, the second step is to validate the circle candidates. Unlike traditional circle detectors [15-19] which work on a set of potential edge

pixels to generate candidate circles, the CCDA follows a proactive approach and works by first identifying contour corners and then removing all the detected corners to generate a new contour set, in other words, a candidate circle set. Contour corners are defined by the proposed algorithm [31]. Starting with a contour, the algorithm operates by first picking two external points and connecting them with a line; it then searches to find the farthest point from the drawn line and adds this point to the approximation. The process is repeated, adding the next most distant point to the accumulated approximation until all of the points are less than the indicated distance. Satisfied point coordinates are stored in a corner vector, $V = \{v_1, v_2, \dots, v_{N_{i2}}\}$, with $v_i = \{v_{i1}, v_{i2}, \dots, v_{in_i}\}$ being a corner vector of contour t_i and n_i being the number of corners on the contour, t_i .

The contour extraction process is implemented as follows. Given a list of corner points, $v_i = \{v_{i1}, v_{i2}, \dots, v_{in_i}\}$, on each contour, t_i , walk over the contours, t_i , and detach the sub-contours from the contour at the corner points. Finally, the new contour segment set is generated from T and stored in a contour vector, $ST = \{st_1, st_2, \dots, st_{N_{i2}}\}$, with $st_i = \{st_{i1}, st_{i2}, \dots, st_{in_i}\}$ being the contour segment vector of contour t_i (as shown in Fig. 3-3d).

Remaining circle detection

The CCDA detects the remaining circle based on the following ideas. A circle candidate set is first created from the contour segments, ST , and then the circles actually represented in the image are defined. For a set of contour segments ST , to be a potential candidate, it must have redundant contours; therefore, certain criteria should be released. For the purposes of our current implementation, we do not consider very short contours that are canceled as they may represent noise. Using these criteria, M satisfied contour segments are extracted to make circle candidates, and they are stored in a vector, $RC = \{r_1, r_2, \dots, r_M\}$, with r_i and M being the i^{th} circle candidate and the number of circle candidates of the contour, r_i , respectively, where

$M \leq N_{12}$. Following the technique in [19], for each extracted contour segment, i , we select three edge points, $p_{k1}(x_1, y_1)$, $p_{k2}(x_2, y_2)$, and $p_{k3}(x_3, y_3)$, and generate a virtual circle, r_i , that passes through the selected points. The virtual circle, r_i , is defined by the three parameters, x_i , y_i , and R_i , where (x_i, y_i) is the center coordinate and R_i is its radius. The virtual circle can be modeled as follows:

$$(x - x_i)^2 + (y - y_i)^2 = R_i^2, \quad (3-11)$$

where x_i and y_i can be calculated by the following equations

$$x_i = \frac{\det(A)}{4((x_2 - x_1)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_1))} \quad (3-12)$$

$$y_i = \frac{\det(B)}{4((x_2 - x_1)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_1))} \quad (3-13)$$

where

$$A = \begin{bmatrix} x_2^2 + y_2^2 - (x_1^2 + y_1^2) & 2(y_2 - y_1) \\ x_3^2 + y_3^2 - (x_1^2 + y_1^2) & 2(y_3 - y_1) \end{bmatrix} \quad (3-14)$$

$$B = \begin{bmatrix} 2(x_2 - x_1) & x_2^2 + y_2^2 - (x_1^2 + y_1^2) \\ 2(x_3 - x_1) & x_3^2 + y_3^2 - (x_1^2 + y_1^2) \end{bmatrix} \quad (3-15)$$

The radius is calculated using

$$R_i = \sqrt{(x_i - x_d)^2 + (y_i - y_d)^2} \quad (3-16)$$

with $p_d(x_d, y_d) \in \{p_1, p_2, p_3\}$.

After the virtual circles are generated, the next step is to validate whether the circles really exist in the image. First, the midpoint circle algorithm (MCA) is used to determine the required points for drawing circles on the image. The MCA input arguments are the center coordinate, (x_i, y_i) , and the radius, R_i . It starts drawing a curve at point $(R_i, 0)$ and proceeds upwards and left by using integer additions and subtractions. It is important to ensure that it does not

consider points lying outside the image plane. Assuming that the number of required points representing virtual circle r_i in the image is N_{ci} , the reinforcement, $\alpha(r_i)$, implies the matching error between the pixels on the virtual circle, r_i , and the pixels that actually exist in the image.

Let $E(x_j, y_j)$ denote a function that verifies the pixel existence at (x_j, y_j) .

$$E(x_j, y_j) = \begin{cases} 1 & \text{if } (x_j, y_j) \text{ is an edge point} \\ 0 & \text{otherwise} \end{cases} \quad (3-17)$$

Then $\alpha(r_i)$ can be calculated as follows

$$\alpha(r_i) = \frac{\sum_{j=1}^{N_{ci}} E(x_j, y_j)}{N_{ci}} \quad (3-18)$$

A value close to $\alpha(r_i)$ implies a better candidate. There are two ways to obtain the optimal solution: either one virtual circle (circle candidate) generates a matching reinforcement $\alpha(\cdot)$ under the pre-established limit (typically 0.1 as has been suggested [19], or it takes the highest probability action at the end of the process.

3.2.3 Fusion of color and shape for irregular moving object detection and tracking

The possible target positions calculated by different estimators are combined to validate the final target position by assigning higher priority to the model's more likely features. The similarity of each statistical distribution with the model is evaluated using the Bhattacharyya coefficients. Due to the fact that the Bhattacharyya coefficients are closely related to Bayes error and their properties have been previously illustrated [2], the Bhattacharyya coefficients represent a near optimal solution. Let $q = \{q_u\}_{u=1\dots m}$ denote the model histogram, which is determined in the tracking initialization step. At frame k , the candidate histogram is $p(k) = \{p_u(k)\}_{u=1\dots m}$. The distance between two m-bin histograms is defined as

$$d(k) = d(q, p(k)) = \sqrt{1 - \rho(k)} \quad (3-19)$$

The Bhattacharyya coefficient $\rho(x)$ is given by

$$\rho(k) = \rho(q, p(k)) = \sum_{u=1}^m \sqrt{p_u(k)q_u} \quad (3-20)$$

The higher Bhattacharyya coefficient reflects the higher contribution to the target position decision. Let us denote possible positions found in the color space and shape feature space as (x_c, y_c) and (x_s, y_s) , respectively, and the corresponding Bhattacharyya coefficients as ρ_c, ρ_s . The final target position is determined as follows:

$$x = \frac{\rho_c}{\rho_c + \rho_s} x_c + \frac{\rho_s}{\rho_c + \rho_s} x_s \quad (3-21)$$

$$y = \frac{\rho_c}{\rho_c + \rho_s} y_c + \frac{\rho_s}{\rho_c + \rho_s} y_s \quad (3-22)$$

The possible window size in any feature space that provides the largest Bhattacharyya coefficient is chosen as the adaptive window. In addition, the statistical representations of the model determined at the tracking initialization stage needs to be updated since the target and/or the camera are moving over time. Moreover, in the case of complete occlusion or image losses, the calculation becomes irrelevant. In such situations, it is better to rely on the evolution model. To decide when the model needs to be adaptively updated in a certain feature space, the proper thresholds T_h and T_l ($T_h > T_l$) were empirically validated. In particular, this occurred every 30 frames when computing the Bhattacharyya coefficients, and comparing their values to the thresholds. In any feature space, the model needs to be updated if its Bhattacharyya coefficient belongs to the closed interval T_l, T_h . The object is declared to be lost when any Bhattacharyya coefficient is below a given threshold T_l .

3.3 Experimental Results

This section presents the experiments conducted in the object detection and tracking system. The system sends the images from its PTZ camera to the station (PC). The data is processed on the station and the desired orientations are sent back to control the PTZ camera. The test

platform was implemented in C, operating with an AMD XP processor running at 2.80 GHz. The overall visual detection, tracking, and control system runs with a frame rate of 30 Hz.

3.3.1 Installation and reinstallation after occlusion confirmation

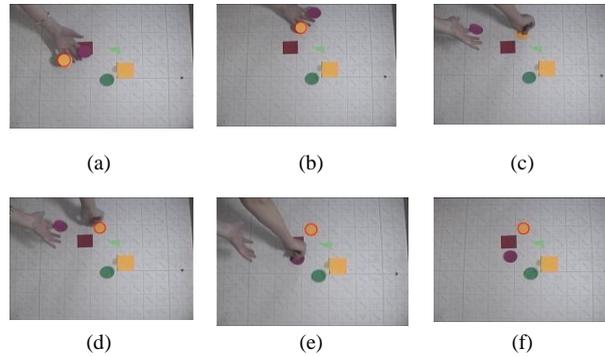


Fig. 3-4 Example of installation and reinstallation after occlusion.

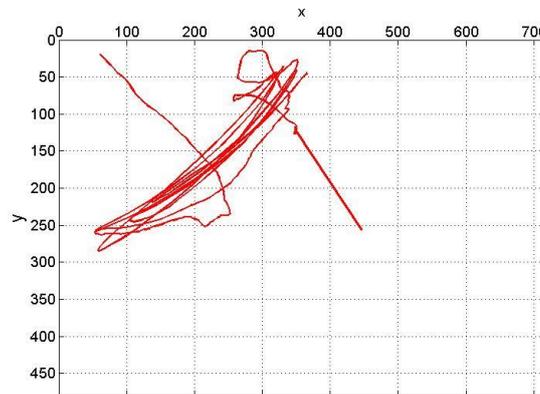


Fig. 3-5 Trajectory of the center of an object in the image

The first experiment is performed while the camera is static; the frame sequence is about 1000 frames, and the target is a yellow circular object. The experiment is performed to examine the proposed system installation and reinstallation after occlusion.

First, the system is in DTS, or in other words, whole image processing, and when the target appears in the camera view, it is detected automatically by CCDA (as seen in Fig. 3-4a, 3-4b), where the detected circle is marked in the red overlay. The system decides the detected circle is the target object, learns the target histogram model, and switches to TKS mode, or in other words, interest region processing. When a loss is detected, the system goes back to DTS mode, as shown in Fig. 3-4c. The system successfully reinitializes in Fig. 3-4d, and DTS succeeds in

finding the target after loss detection. TKS then continues to follow the target object. In the experiment, there are 4 total occlusions, 1 time fail in tracking, and 5 successful recoveries from the loss. The target moving trajectory is shown in Fig. 3-5.

3.3.2 Performance of chasing a moving target

This experiment evaluates the proposed system performance in chasing a moving target. In this experiment, we describe a longer sequence of about 10,000 frames, in which the target is moving, and the system follows the target to keep it in the CCV.



Fig. 3-6 Chasing a moving target.

Fig. 3-6 shows an example of chasing target objects. In these cases, the system detects the object successfully, marks it in the light red overlay, and the camera continues following the target object to keep it in the CCV. In the experiment, we assume that the target object moves at slow speeds, less than the maximum speed of the camera.

3.3.3 Performance of object detection

The third experiment is conducted to evaluate object detection performance of the proposed system. The test is performed in object detection and measurement of bin picking system. In this test, we compare the results between the proposed approach and approaches using color or shape information.

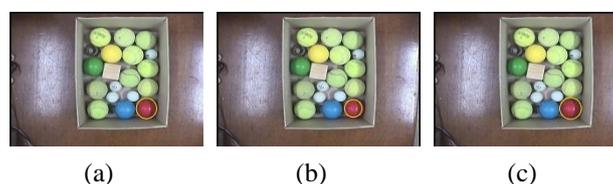


Fig. 3-7 Chasing Comparison between object detection and measurement approaches. (a) using proposed feature fusion; (b) using color information; (c) using shape information

While the three methods have comparable performances, the proposed method using feature fusion has the best performance as can be seen in Fig. 3-7.

3.3.4 Performance w.r.t the changing of illumination

This experiment is performed while changing illumination through video sequence. The proposed system detects and tracks a target object.

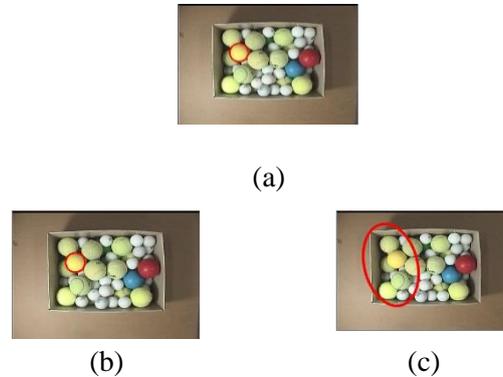


Fig. 3-8 Video sequence contains drastic illumination changes. (a) Initialize model; (b) Tracking using proposed online feature fusion; (c) Tracking using color information alone fails due to illumination changes.

The video sequence contains the target, a yellow ball, shown in Fig. 3-8(a). The tracking result shows that using color cue alone is not enough to provide good result, and that the tracker fails to track the target object when the illumination changes drastically as shown in Fig. 3-8(c). The shape feature still preserves well the structure of the target object in this case, and the tracker that employs both shape and color features maintains good performance as shown in Fig. 3-8(b).

3.3.5 Performance w.r.t the changing of view point

This experiment is conducted while changing view point through video sequence. The view point changes cause flattening effects which make circular objects appear elliptic. The experiment results show that the proposed approach can detect and track a target object successfully while changing view point.

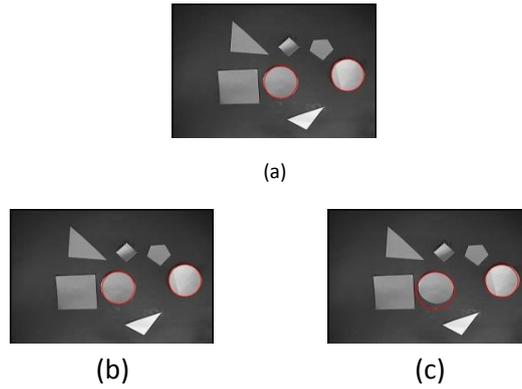


Fig. 3-9 Video sequence contains view point changes. (a) Initialize model; (b) Tracking using online feature fusion; (c) Tracking using shape information alone fails due to view point changes.

The video sequence presents challenges in viewpoint changes, shown in Fig. 3-9, in which the target objects are the two circles. Using shape feature only is not enough to track the target accurately as can be seen in Fig. 3-9(c). However, the fusion of features still tracks the target successfully as shown in Fig. 3-9(b).

Overall, the online fusion of color and shape feature provides a robust approach for object tracking. Color information compensate for shape changes due to viewpoint changes, partial occlusion, while shape information compensates for color variance in illumination changes condition.

3.4 Chapter summary

In this chapter, we proposed a robust object position detection and tracking using dynamic fusion of color and shape. The statistical model of the object is built and updated in different feature spaces, and by fusing possible target positions in the feature spaces, the final target position is validated. The experiment results demonstrate robust object detection and tracking performance under various scenarios, including perspective changes, drastic illumination changes, and occlusions. In next sections, object position and attitude measurement in 3D coordinate system approaches will be presented.

Chapter 4

SLIPPAGE MEASUREMENT USING MONOCULAR STRUCTURED LINE LIGHT VISION

4.1 Introduction

4.1.1 Related works

Tactile receptors distributed on the human skin allow human to sense grasped objects slippage so that human can easily control their hand contact force to prevent the objects from sliding. Imitating skilled human behaviors, many types of tactile sensor have been developed by using electrical resistive, capacitive, electromagnetic or ultra-sonic component, piezoelectric, optical component, strain gauges, etc. [66-68]. These sensors structure are complex, they require numerous sensing elements and complicated wiring. Apart from that, vision-based sensors or optical sensors are also developed for tactile sensing [69-71]. In order to dealing with solid objects, the vision-based tactile sensor usually consists of a CCD camera, source light emitter, a transparent acrylic plate, and touchpad [72-74]. These sensors obtain slippage based on the movements of dots printed on the surface of the touchpad captured by the CCD camera. In [75], an approach estimating slippage of grasped flexible object, an elastic object, is proposed. A feature point is drawn on the apex of the elastic sphere and the authors used deformation of the contact area measured by a camera through a transparent plate when an elastic object slides on a rigid plate.

These above contacted sensors can obtain a variety of tactile information such as contact region, slippage, and contact force accurately, but many crucial issues remain unresolved. Firstly, these sensors are very expensive, and their structures are used to be complex, require many sensing elements and complicated manufacturing. Secondly, these sensors still meet difficulty in obtaining various types of information simultaneously for example determining the grasped object position, orientation in 3-dimensional space. Thirdly, the sensor surface of these

sensors can easily be damaged especially the apex contacted with grasped objects, leading to inaccuracy measurement.

4.1.2 Proposed system overview



Fig. 4-1 Video Examples of objects with rectangular features

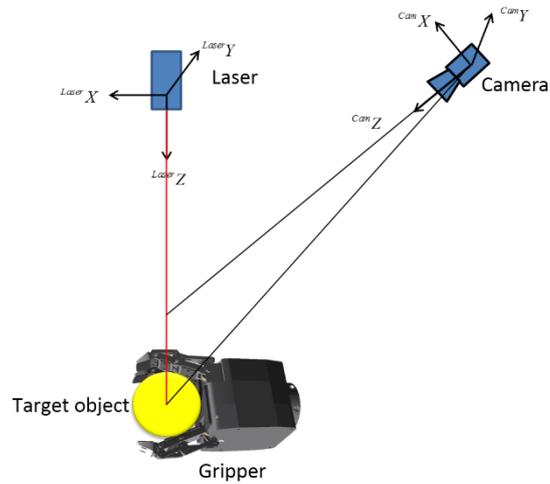


Fig. 4-2 The non-contact slippage measurement system using monocular structure line light vision.

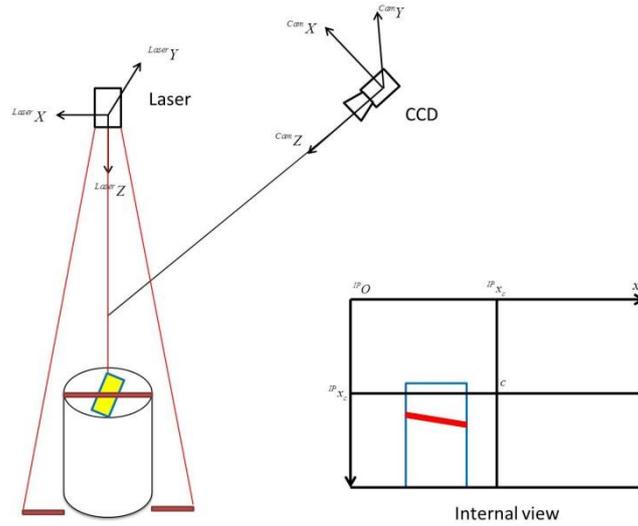


Fig. 4-3 The configuration of the monocular-structured light vision-based slippage estimation apparatus.

Currently, contact and tactile sensors are commonly used to measure the slippage of an object. There are many drawbacks of contact and tactile sensors, including that they are complex, expensive, easily damaged, and often fail in determining the object pose. Non-contact vision sensors using laser or ultrasonic technology have also been recently studied. However, these sensors cannot obtain the object pose and depend strongly on the reflectivity of the object. In order to solve these remaining problems herein, a slippage measurement system consisting of a structured light source and a camera is proposed, as can be seen in Fig.4-2. The proposed approach explores both light reflection information and object feature information. Rectangular features are very common in objects that are mechanically grasped and manipulated, and are used as grasping objects in this study, as shown in Fig.4-1. Therefore, the capability to perform object pose estimation by recognizing rectangular features is of great significance. In fact, due to the huge size of the target object, and due to limitations in the monocular camera field of view (FOV) and grasping gripper, the monocular vision-based system usually cannot capture the entire rectangular feature on the target object. Only a partial rectangular feature appears in the captured images. Thus, in this paper, a 3D object pose reconstruction approach using partial object features is proposed. Then, the slippage is calculated by estimating the object pose

change through an image frame sequence. The configuration of the vision-based slippage estimation procedure is illustrated in Fig.3. The light line forms a line on the target object. This line and the object parts are then captured by the camera to determine the target object pose.

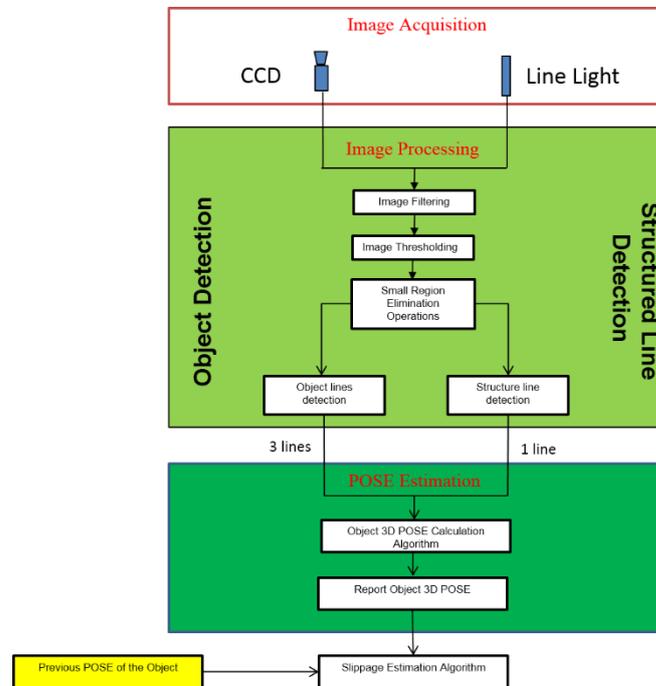


Fig. 4-4 A flowchart of slippage estimation using monocular structured line vision.

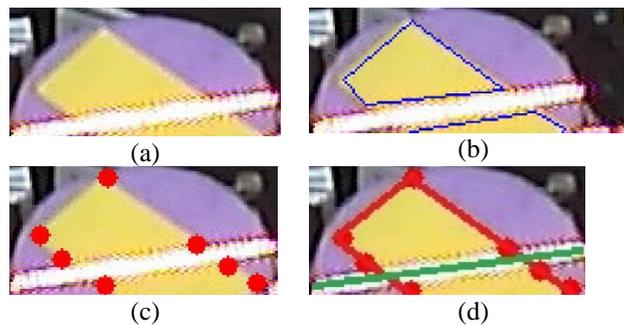


Fig. 4-5 The extracted target object lines and structured line detection. (a) Original image; (b) curve detection; (c) detection of high curvature points; (d) object lines and structured line detection.

The object slippage estimation flowchart is shown in Fig.4-4. First, the image sequence passes through a Gaussian filter. Secondly, a thresholding method is applied to convert the RGB image into a binary image. Small regions are then eliminated, as they may represent noise. After that, edges in the image are detected using the Canny edge detection algorithm, as can be

seen in Fig.4-5(b). Due to noise, illumination change, perspective, etc., the rectangular feature will not always appear as a rectangle, and the rectangle is divided into two regions by the structured line light. In the obtained edge image, a high curvature points detection algorithm is used to detect curve vertices, which are marked by a red color in Fig. 4-5(c). Due to the target model and system construction, there are 3 object lines and 1 structured light line. These lines are extracted by taking middle lines using a line fitting algorithm. Fig. 5(d) shows detected object lines (marked with the color red) and the structured line (marked in the color green). Let us denote 3 object lines l_1, l_2, l_4 , and a structured line l_5 , and suppose that the hidden line is l_3 ; further assume that l_3 is imaginary. The equations of the lines are of the form

$$\begin{aligned} A_j x + B_j y + C_j &= 0 \\ \text{Subject to } j &= \{1, 2, 3, 4, 5\} \end{aligned} \quad (4-1)$$

In the homogeneous coordinate, these lines are represented by homogeneous vectors $[A_j, B_j, C_j]^T$. These lines will be used for the calculation of the object pose.

Finally, slippage is estimated by calculating the difference between the object pose in consecutive frames.

4.2 Slippage measurement

There are two main steps in the proposed non-contact slippage measurement approach: target object pose determination and slippage measurement. The non-cooperative target object pose is determined based on the information of the structured light line and the partial rectangle feature. Then, the slippage is measured by calculating the difference in object pose through a frame sequence.

4.2.1 Target object pose estimation

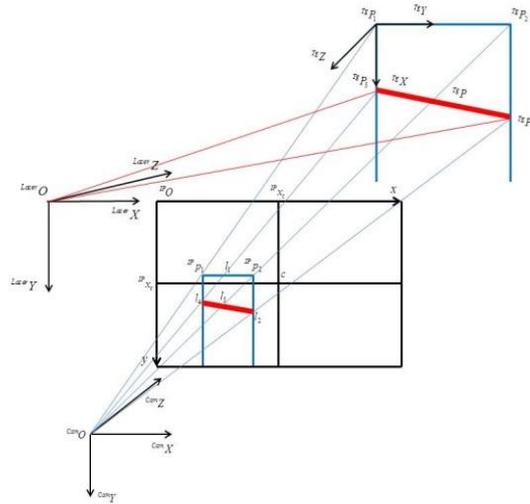


Fig. 4-6 The vision sensor model of the monocular structured light slippage measurement.

The model of the monocular structured light slippage measurement is shown in Fig.4-7. There are four coordinate systems in this model: the camera, image, laser, and object coordinate systems. The CCD camera coordinate system ${}^{\Sigma}Cam$ is a right-hand coordinate system permanently fixed to the camera with its origin ${}^{Cam}O$ at the camera origin and the ${}^{Cam}Z$ - axis coinciding with the optical axis of the camera, pointing to the line of sight, following the standard parameterization of the camera coordinate system. The camera is described by a pin-hole model. Let ${}^{\Sigma}IP$ denote the image plane coordinate system, whose coordinate are ${}^{IP}P = [{}^{IP}x, {}^{IP}y, 1]^T$, where the x-axis and y-axis are along the top and left side of the image plane, respectively. The origin ${}^{IP}O$ of ${}^{\Sigma}IP$ is at the top left vertex of the image plane and the center of the image plane is ${}^{IP}C = [{}^{IP}x_c, {}^{IP}y_c, 1]^T$. The laser system ${}^{\Sigma}Laser$ is a right-hand coordinate system in which the laser line is assumed to lie on the ${}^{XZ}Laser$ plane and the origin is at the laser origin. ${}^{\Sigma}Tg$ is the target object coordinate system, whose coordinate are ${}^{Tg}P = [{}^{Tg}X, {}^{Tg}Y, {}^{Tg}Z, 1]^T$. The origin ${}^{Tg}O$ of ${}^{\Sigma}Tg$ is at the top left vertex of the target object.

Suppose that there is a point ${}^{Cam}S = [{}^{Cam}X_s, {}^{Cam}Y_s, {}^{Cam}Z_s]^T$ in the camera coordinate system, and its image in the image plane is ${}^{IP}s = [{}^{IP}x_s, {}^{IP}y_s, 1]^T$. In the camera coordinate system, using triangle similarity, we obtain

$${}^{Cam}S = w_{Cam} {}^{IP}M^{-1} {}^{IP}s, \quad (4-2)$$

where ${}^{IP}M$ is the camera intrinsic matrix and ${}^{IP}M$ can be exactly obtained in the CCD camera calibration process.

The physical transformation between the object coordinates and camera coordinates is the sum of the effects of some rotation R and some translation t . In homogeneous coordinates, we can combine these within a single matrix, as follows

$${}^{Cam}W_{Tg} = \begin{bmatrix} {}^{Cam}R & {}^{Cam}t \\ 0 & 1 \end{bmatrix} \quad (4-3)$$

Writing this out, we have:

$${}^{Cam}W_{Tg} = \begin{bmatrix} {}^{Cam}R_{3 \times 3} & {}^{Cam}t_{3 \times 1} \\ 0_{1 \times 3} & 1 \end{bmatrix} \quad (4-4)$$

Then, the relation between the image plane and target object coordinates (or world coordinates) can be expressed as follows.

$${}^{IP}s = w^{-1} \begin{bmatrix} {}^{IP}M_{3 \times 3} & 0_{3 \times 1} \end{bmatrix} {}^{Cam}W_{Tg} S \quad (4-5)$$

We can choose the object planes as shown in Fig. 5, such that ${}^{Tg}Z = 0$ as follows.

$${}^{IP}S = w^{-1} {}^{IP}M_{Cam} [r_1 \ r_2 \ r_3 \ t] \begin{bmatrix} {}^{Tg}X_S \\ {}^{Tg}Y_S \\ 0 \\ 1 \end{bmatrix} = w^{-1} {}^{IP}M_{Cam} [r_1 \ r_2 \ t] \begin{bmatrix} {}^{Tg}X_S \\ {}^{Tg}Y_S \\ 1 \end{bmatrix} \quad (4-6)$$

The remaining problem to be solved is to determine the relative position between camera coordinates and target object coordinates, i.e. we need to find the rotation ${}^{Cam}_{Tg}R$ and translation

${}^{Cam}_{Tg}t$.

The non-cooperative target object size is unknown; and only a partial target object appears in the image plane. In order to estimate the object pose in the camera coordinate system, we need to reconstruct the target object.

Suppose that the target width and height are rh and h , respectively, and its vertices in the target object coordinate system are given by

$$\begin{cases} {}^{Tg}P_1 = [0, 0, 1]^T \\ {}^{Tg}P_2 = [0, rh, 1]^T \\ {}^{Tg}P_3 = [h, 0, 1]^T \\ {}^{Tg}P_4 = [h, rh, 1]^T \end{cases} \quad (4-7)$$

where ${}^{Tg}P_3$ and ${}^{Tg}P_4$ are virtual points. The image coordinates and camera coordinates of these four points are respectively represented as follows.

$$\begin{cases} {}^{IP}P_i = [{}^{IP}x_{p_i}, {}^{IP}y_{p_i}, 1]^T \\ {}^{Cam}P_i = [{}^{Cam}X_{P_i}, {}^{Cam}Y_{P_i}, {}^{Cam}Z_{P_i}]^T \end{cases} \quad \text{subject to } i = \{1, 2, 3, 4\} \quad (4-8)$$

where the image coordinates of the target object vertices ${}^{IP}p_i$ are the vector product (or cross product) of the two lines [15], which is represented by homogeneous vectors, as described in Section 2.

$$\begin{cases} {}^{IP}p_1 = l_1 \times l_4 \\ {}^{IP}p_2 = l_1 \times l_2 \\ {}^{IP}p_3 = l_4 \times l_3 \\ {}^{IP}p_4 = l_3 \times l_2 \end{cases} \quad (4-9)$$

Because P_1, P_2, P_3 are not on a same line, the matrix $\begin{bmatrix} {}^{Tg}P_1 & {}^{Tg}P_2 & {}^{Tg}P_3 \end{bmatrix}$ is non-singular, so we have:

$$w_4 {}^{IP}p_4 = \begin{bmatrix} {}^{IP}p_1 & {}^{IP}p_2 & {}^{IP}p_3 \end{bmatrix} \text{diag}(w_1, w_2, w_3) \begin{bmatrix} {}^{Tg}P_1 & {}^{Tg}P_2 & {}^{Tg}P_3 \end{bmatrix}^{-1} {}^{Tg}P_4 \quad (4-10)$$

Substituting (7) into (10) we obtained:

$$w_4^{-1} [-w_1, w_2, w_3]^T = \begin{bmatrix} {}^{IP}p_1 & {}^{IP}p_2 & {}^{IP}p_3 \end{bmatrix}^{-1} {}^{IP}p_4 \quad (4-11)$$

The physical transformation between the laser projector coordinate system and camera coordinate system is the sum of the effects of some rotation ${}_{Laser}^{Cam}R$ and some translation ${}_{Laser}^{Cam}T$, which were exactly calculated in the camera and laser calibration process [16]. In the homogeneous coordinates, we can combine these within a single matrix as the following.

$${}_{Laser}^{Cam}W = \begin{bmatrix} {}_{Laser}^{Cam}R & {}_{Laser}^{Cam}T \end{bmatrix} = \begin{bmatrix} {}_{Laser}r_1 & {}_{Laser}r_2 & {}_{Laser}r_3 & {}_{Laser}T \end{bmatrix} \quad (4-12)$$

Suppose that the image ${}^{IP}p_5$ of P_5 is the intersection of lines l_5 and l_4 and the image ${}^{IP}p_6$ of P_6 is the intersection of lines l_5 and l_2 . The camera coordinates of P_5 and P_6 can be calculated as follows

$${}^{Cam}P_j = [1, 0, 0] (\Delta^T \Delta)^{-1} \Delta^T {}_{Laser}^{Cam}T {}^{IP}M^{-1} {}^{IP}p_j \quad (4-13)$$

where $\Delta = \begin{bmatrix} {}^{IP}M^{-1} {}^{IP}p_j & -{}_{Laser}r_1 & -{}_{Laser}r_3 \end{bmatrix}$ subject to $j = \{5, 6\}$

In the camera coordinate system, the target object plane is determined as follows:

$$[m, n, p] \begin{bmatrix} {}^{Cam}X \\ {}^{Cam}Y \\ {}^{Cam}Z \end{bmatrix}^T = 1 \quad (4-14)$$

Hence, the normal vector of the target object plane is:

$$[m, n, p] = [1, 1, 1] \begin{bmatrix} {}^{Cam}P_1 \\ {}^{Cam}P_2 \\ {}^{Cam}P_3 \end{bmatrix}^{-1} \quad (4-15)$$

We also have:

$$\begin{bmatrix} {}^{Cam}P_1 \\ {}^{Cam}P_2 \\ {}^{Cam}P_3 \end{bmatrix} = {}^{IP}M^{-1} \begin{bmatrix} {}^{IP}p_1 \\ {}^{IP}p_2 \\ {}^{IP}p_3 \end{bmatrix} \text{diag}(w_1, w_2, w_3) \quad (4-16)$$

Combining (14), (15), (16) we have the constraint

$$[1, 1, 1] \left(\begin{bmatrix} {}^{IP}p_1 \\ {}^{IP}p_2 \\ {}^{IP}p_3 \end{bmatrix} \text{diag}(w_1, w_2, w_3) \right)^{-1} {}^{IP}M \begin{bmatrix} {}^{Cam}X \\ {}^{Cam}Y \\ {}^{Cam}Z \end{bmatrix}^T = 1 \quad (4-17)$$

By using fusion of the rectangular constraints in (11) and the constraint in (17), we obtain

$$w_4 = [1, 1, 1] \left(\begin{bmatrix} {}^{IP}p_1 \\ {}^{IP}p_2 \\ {}^{IP}p_3 \end{bmatrix} \text{diag}(b_{11}, b_{21}, b_{31}) \right)^{-1} {}^{IP}[1, 0, 0] (\Delta^T \Delta)^{-1} \Delta^T \begin{matrix} {}^{Cam}T \\ {}^{Laser} \end{matrix} {}^{IP}p_j \quad (4-18)$$

where $[-b_{11}, b_{21}, b_{31}]^T = \begin{bmatrix} {}^{IP}p_1 \\ {}^{IP}p_2 \\ {}^{IP}p_3 \end{bmatrix}^{-1} {}^{IP}p_4$

Since we can choose ${}^{IP}p_4$ which lies on l_2 , the camera coordinate system of the virtual point

${}^{Cam}P_4$ can be calculated. The couple directions $\left(\overline{{}^{Cam}P_4 {}^{Cam}P_3}, \overline{{}^{Cam}P_4 {}^{Cam}P_6} \right)$

and $\left(\overline{{}^{Cam}P_3 {}^{Cam}P_5}, \overline{{}^{Cam}P_3 {}^{Cam}P_4} \right)$ are couple conjugate points with respect to the absolute conic

which represented by the 3×3 symmetric Ω_∞ [15]. By fusing these above constraints with the

constraint that ${}^{Cam}P_3$ lies on the line passing through ${}^{Cam}P_5$, whose direction is $\overline{{}^{Cam}P_4 {}^{Cam}P_6}$,

the ${}^{Cam}P_3$ coordinate can be obtained and w_1, w_2 can be calculated. The camera coordinates of

these vertices ${}^{Cam}P_1$ and ${}^{Cam}P_2$ can be calculated by using (2). The target object now can be re-constructed in the camera coordinate system with the information of the vertices ${}^{Cam}P_1$, ${}^{Cam}P_2$, ${}^{Cam}P_3$, and ${}^{Cam}P_4$.

4.2.2 Slippage measurement

Because the length of a line is constant in different coordinate systems, we can calculate the target object size in the camera coordinate system. The real size of the triangular components are:

$$rh = \left\| {}^{Cam}P_1 {}^{Cam}P_2 \right\| \quad (4-19)$$

$$h = \left\| {}^{Cam}P_1 {}^{Cam}P_3 \right\| \quad (4-20)$$

The relative position between the target object and camera coordinate can be calculated by using the least squares method

$$\begin{aligned} \begin{bmatrix} {}^{Cam}r_1, {}^{Cam}r_2, {}^{Cam}t \end{bmatrix} &= \begin{bmatrix} {}^{Cam}P_1, {}^{Cam}P_2, \dots, {}^{Cam}P_n \end{bmatrix} \begin{bmatrix} {}^{Tg}P_1, {}^{Tg}P_2, \dots, {}^{Tg}P_n \end{bmatrix}^T \\ &\left(\begin{bmatrix} {}^{Tg}P_1, {}^{Tg}P_2, \dots, {}^{Tg}P_n \end{bmatrix} \begin{bmatrix} {}^{Tg}P_1, {}^{Tg}P_2, \dots, {}^{Tg}P_n \end{bmatrix}^T \right)^{-1} \end{aligned} \quad (4-21)$$

$${}^{Cam}r_3 = {}^{Cam}r_1 \times {}^{Cam}r_2 \quad (4-22)$$

We can calculate the center of target object in consecutive frames in the camera coordinate system as follows

$${}^{Cam}P_c = \frac{1}{4} \left[\sum_{i=1}^4 {}^{Cam}X_{P_i}, \sum_{i=1}^4 {}^{Cam}Y_{P_i}, \sum_{i=1}^4 {}^{Cam}Z_{P_i} \right] \quad (4-23)$$

The slippage between the initial position and k^{th} position is finally obtained as:

$${}^{Cam}Slip = \left\| {}^{Cam}P_{c,0} - {}^{Cam}P_{c,k} \right\| \quad (4-24)$$

4.3 Experimental results

In this section, we verify the accuracy of the proposed system through an experiment using a CCD camera and a line laser sensor. The test platform was implemented in C/C++. The target object is a cylinder which has a rectangular feature on its top. The monocular vision system is applied in the range of 300mm ~1000mm which is a sufficient FOV to capture the image.



Fig. 4-7 The real testbed system.

We apply the proposed approach into the slip-margin feedback control gripper system. The three-degree-of-freedom (DoF) gripper system consisting of a CCD camera, structured light, and force sensor, grasps a target object. The incipient slippage, which occurs on the contact surface between the grip fingers and grasping object when the object is pressed and slid, is calculated by estimating the differences of the object moving distance and the increment distance between the current object position and the initial position between two consecutive frames. Then, the grip force is immediately controlled by a direct feedback of the estimated slip margin. The experiment testbed system is shown in Fig. 4-8.

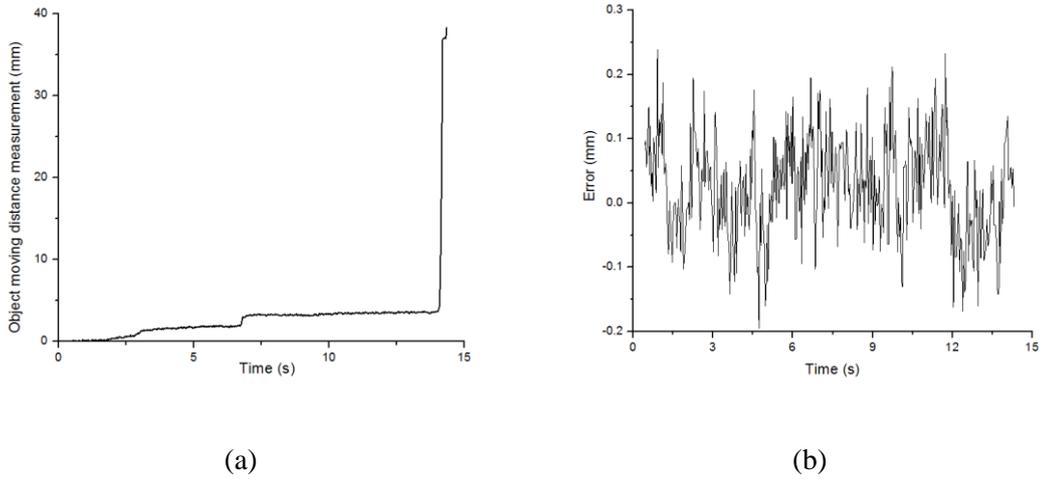


Fig. 4-8 Slippage measurement and measurement error. (a) Slippage estimation; (b) measurement error.

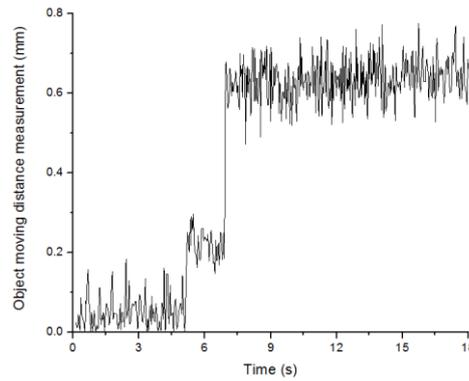


Fig. 4-9 Slippage measurement with gripper force $F_d = 47.092[\text{N}]$ and $M = 2.3[\text{kg}]$

4.3.1 Target object increment distance measurement

In this experiment, the target object is pressed by a slight force so that a known slippage occurs. The proposed system measures the slippage and Fig. 4-9(a) shows the distance increment between the current object position and the initial position, with corresponding measurement errors shown in Fig. 4-9(b). When the difference of the object moving distance between two consecutive frames, or the slippage is greater than a pre-set threshold (0.5 mm is selected

in our experiment), then the system will increase the grip force to prevent the object from sliding.

As can be seen in Fig. 4-9, the proposed system can measure different slippages in real time and the measurement error is less than 0.2 mm.

4.3.2 *Slippage estimation with different target objects*

In the second experiment, we applied different object sizes and different forces for measuring slippage. When slippage occurs, the vision system automatically estimates object movement and sends a signal to the gripper controller via a COM port. The signal information contains the object movement distance between 2 consecutive frames. Fig. 4-10, 4-11, and 4-12 show the measured object moving distance of the target object applied to different objects with different gripper forces. The physical transformation between the target initial position and position at the k^{th} time increment, in some translation T_k , and the rotations around ${}^{Tg}X$, ${}^{Tg}Y$, ${}^{Tg}Z$ are denoted by α , β , φ , respectively.

In Fig. 4-10, when the gripper force is $F_d=47.092[N]$ and the target object has a weight of $M=2.3[kg]$, there are two times at which slippage occurs. In the first slippage event, the target object slides 0.2mm from the initial position, with an average of $T=(0,0,0.2\text{mm})$, $\alpha=0$ degree, $\beta=0$ degree, $\varphi=1$ degree. At the instant of the second event, the target object slides 0.4mm, and the slippage compared with the initial position is recorded as 0.6mm, with an average $T=(0,0,0.6\text{mm})$, $\alpha=0$ degree, $\beta=0$ degree, $\varphi=2$ degree.

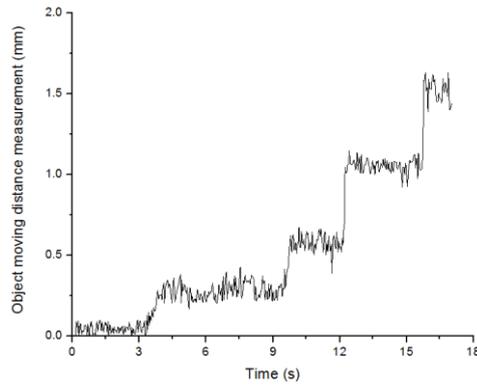


Fig. 4-10 Slippage measurement with a gripper force of $F_d=46.36[N]$ and $M=2.1[kg]$

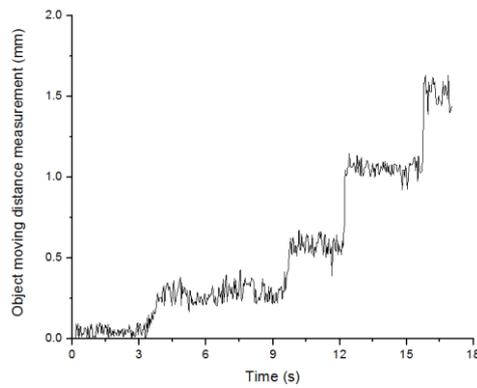


Fig. 4-11 Slippage measurement with a gripper force of $F_d=46.16[N]$ and $M=1.7[kg]$.

When the gripper force is $F_d=46.36[N]$ and the target object weight is $M=2.1[kg]$, and when $F_d=46.16[N]$ and the target object weight is $M=1.7[kg]$, there are four times at which slippage is recorded during an 18 second period, as can be seen respectively in Fig. 11 and Fig. 12. The slippage measurement system can recognize the slippage with an error of less than 0.2 mm, as confirmed through experiments.

4.4 Chapter summary

In this chapter we proposed a system consisting of a structured light and a camera to estimate the object POSE and grasp slippage. In order to estimate the object pose and slippage through an image frame sequence, a method to detect laser stripes as well as object feature points is proposed. Based on the detected laser stripe and object feature points, the object pose is reconstructed with respect to the 3D coordinate system. Then, slippage is calculated by estimating the object pose change through the image frame sequence.

To verify the proposed non-contact slippage estimation sensor, experiments have been performed. The experimental results confirmed that the sensor can be successfully applied on the slip-margin feedback control gripper system in real-time at 30Hz with a slippage measurement error of less than 0.2 mm.

In a real environment, due to the huge size of the target object, the limitations of the camera FOV, the grasping gripper, etc., the camera cannot capture the entire object. In such cases, the proposed non-contact sensor can measure the object pose and slippage simultaneously using only partially visible object features. Furthermore, no tactile sensing elements are required in the proposed system. The system is light, simple in structure, and cheap, all characteristics which make it suitable for a compact industrial robot. In many robot systems, a camera is already equipped for object grasping and other applications. The proposed slippage estimation system can be easily implemented in such cases by adding a single inexpensive laser.

Chapter 5

GAP AND FLUSH MEASUREMENT USING MONOCULAR STRUCTURED MULTI-LINE LIGHT VISION AND ITS APPLICATION IN VEHICLE ASSEMBLY

5.1 Introduction

In automotive manufacturing plants, gap and flush measurement solutions are used to contain, correct and control the fit and finish of exterior body panels. The gap and flush of parts most visible to the automobile consumer such as gaps between front and rear doors, etc. are automatically measured, recorded and displayed. The gap and flush are crucial aesthetic aspects for any automobile manufacturer, because gaps and flushes of car bodies contribute a great deal to the positive perception of a car. Especially for premium producers, accurate gap and flush measurements are a key criterion for quality. Currently, vehicles are designed with flowing contours, angled panels, edges with large radii and closure seals in the gaps. Mechanical devices often struggle to deal with such complex surfaces. The measurement methodology must be consistent with design criteria and the repeatability of measurements must be very highly accurate. Currently evolving requirements in vehicle assembly usually necessitate an accuracy of measurement of and a repeating accuracy of and still on development. To satisfy these demanding requirements, considerably diverse approaches are being studied and developed.

In this paper, a high-speed accurate non-contact measuring structure using the fusion of a complementary CMOS camera and multi-line laser is studied. The camera and laser vision system commonly explores the strong difference in intensity of the laser beam reflection and image background to extract the laser point locations. However, this method cannot typically be applied in the assembly of vehicle parts that are poorly reflective, e.g. black car body, especially in a high noise industrial environment. Therefore, we propose a multi-line laser-vision-

based gap and flush measurement system developed to solve these issues and achieve high performance measurements. In this system, a GigE high speed CMOS camera with double rate technology is employed to capture the laser stripes generated from low-power generator reflex mirrors from the object surface to ensure real-time measurements. We also contribute new image processing algorithms for adaptive laser stripe extraction and classification by automatically selecting a laser profile region and enhancing the intensity difference between the laser region and background. Then, the laser stripes are mapped into the 3D world coordinate for gap and flush estimation, using our proposed gap and flush measurement approach. The proposed gap and flush measurement system can deal with complex surface with high accuracy, satisfying the demanding gap and flush measurement requirements in vehicle assembly.

The main contributions of this work are as follows:

- New image processing algorithms for adaptive laser stripe extraction and classification by automatically selecting a laser profile region and enhancing the intensity difference between the laser region and background is proposed.
- Propose circle feature detection method to deal with complex surface gap and flush measurement.
- Propose gap and flush measurement method for simple objects as well as complex objects, especially gap and flush measurement method used in vehicle assembly.
- The measurement system and measurement method are simplified so that the measurement system can work real-time and be suitable for industrial applications.

5.2 Gap and flush measurement system



Fig. 5-1 Gap and flush measurement in vehicle assembly. (a) Portable gap and flush measurement system [78]; (b) Inline gap and flush measurement robot system [79].

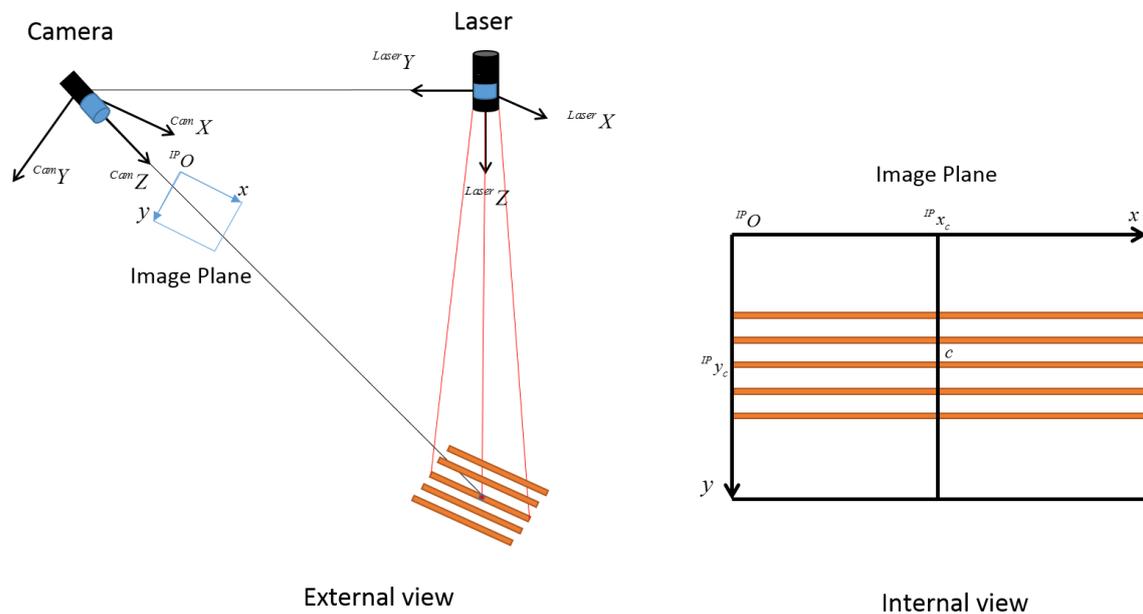


Fig. 5-2 The external and internal view of a gap and flush measurement performed by the camera-laser module.

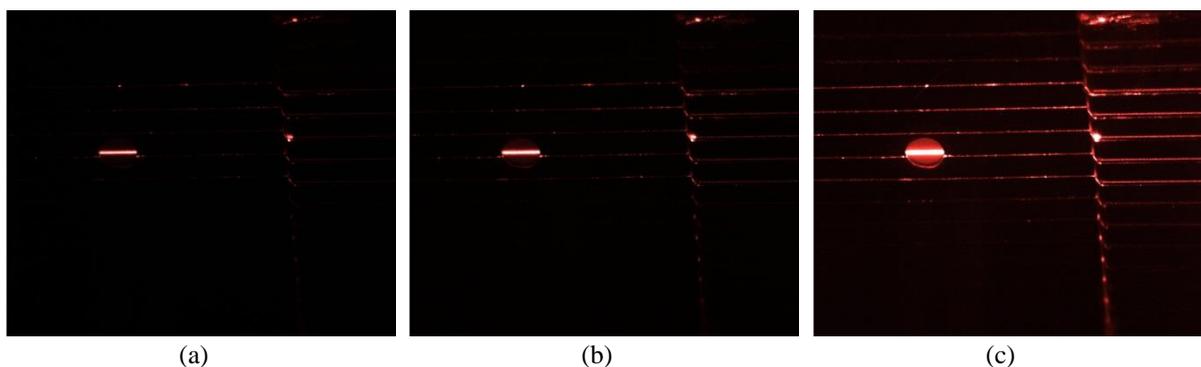


Fig. 5-3 An example of acquired images of laser reflection on car body parts with different gain values [26]. (a) gain value 0; (b) gain value 10; (c) gain value 40

Gap and flush measurement instruments, especially handheld devices, are currently popular in automobile manufacture. These devices consist of cameras and a laser generator, usually a line laser, e.g., the LaserGauge in [80] uses a one-line laser for measurement while CALIPRI

in [81] makes use of a three-line laser. The measurement is performed by moving the devices by hand to the measured objects, carried out on a non-contact basis within the measurement distance and angle range throughout the measurement process, as shown in Fig.5-1(a). For inline robot measurements [81], laser-camera modules are attached to robot hands, and each module contains a laser projector and a camera, as shown in Fig.5-1 (b). These systems provide a high accuracy of measurement values, with a highest accuracy of $\pm 100\mu\text{m}$ and precision of $\pm 100\mu\text{m}$ with a 0.01 mm resolution.

According to the authors' knowledge, there are no domestic, robotic, inline gap and flush measurement robot systems in use. Hence, in this work, we develop camera-laser modules, each consisting of a CMOS camera and multi-line laser, which will be attached to an inline robot for gap and flush measurement in vehicle assembly. During the measurement process, the multi-line laser of each module projects laser stripes onto the surface of the measured object, the car body parts, and the camera captures the stripes produced by the intersections of the laser planes with the target surface. The acquired image is then transferred to the computer through a LAN card. An image processing module of the gap and flush measurement software installed on the computer analyses the deformation of laser stripes to measure the gap and flush. Finally, the gap and flush measurement values are displayed on the GUI of the measurement software.

There are three main coordinate systems in each proposed camera-laser module: the camera, image, and laser coordinate systems. Following the standard way to decide camera coordinate system, the camera coordinate system ${}^{\Sigma}Cam$ is a right-hand coordinate system permanently fixed to the camera with its origin ${}^{Cam}O$ at the camera origin. The ${}^{Cam}Z$ axis coincides with the camera optical axis, pointing along the light of sight. The camera coordinate of a point P is represented by ${}^{Cam}P = [{}^{Cam}X_p, {}^{Cam}Y_p, {}^{Cam}Z_p]^T$. Denote image plane coordinate system ${}^{\Sigma}IP$, where coordinate is represented by ${}^{IP}p = [{}^{IP}x, {}^{IP}y, 1]^T$. The x-axis and y-axis are respectively

along the top and left side of the image plane. The origin ${}^{IP}O$ of ${}^{\Sigma}IP$ is at the top left vertex of the image plane and the center of the image plane is ${}^{IP}c = [{}^{IP}x_c, {}^{IP}y_c, 1]^T$. The laser system ${}^{\Sigma}Laser$ is defined according to the right-hand coordinate system with the laser line assumed to lie on the ${}^{Laser}XZ$ plane with the origin the laser origin and the laser coordinate of a point P is ${}^{Laser}P = [{}^{Laser}X_p, {}^{Laser}Y_p, {}^{Laser}Z_p]^T$. The system overview is shown in Fig.5-2.

5.3 Image Processing Module

Image processing plays a crucial role in the gap and flush measurement system for car manufacturing. In this work, we contribute an image processing module that performs region of interest (ROI) selection, laser profile extraction, classification, and extraction of feature points.

5.3.1 Selection of the ROI

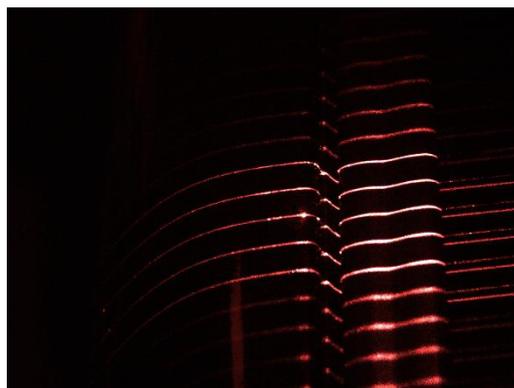


Fig. 5-4 Acquired image of laser stripes on a real vehicle.

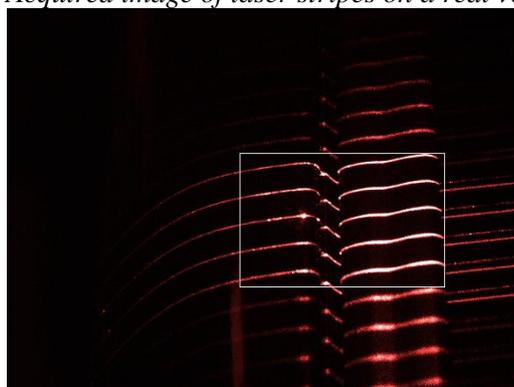


Fig. 5-5 The selected ROI (small rectangle) within the acquired image.

In order to extract the laser stripes region, the threshold method is commonly used. Let $I(x_i, y_i)$ indicates the intensity of point (x_i, y_i) and ρ the pre-set threshold. The point (x_i, y_i) belongs to the laser stripe if and only if it satisfies the threshold condition $I(x_i, y_i) > \rho$. However, if the object is poorly reflective, e.g. a black car body, or a high noise industrial environment, etc. then the difference between the laser stripe and the background is not clear, as can be seen in Fig. 5-3a. Thus, a pre-set threshold is inefficient in such cases. We propose an automatic method to extract the laser profile region in one way and ROI in another. The automatic selection of the ROI is an important procedure for reducing the computation load of the measurement process. This procedure selects a ROI that contains an image of laser stripes. Initially, the camera gain value is sufficiently increased such that the obtained laser stripes are definitively clear, a procedure that can be performed during the calibration process. Obviously, when the gain is increased, the noise is also increased, as can be seen in Fig.5-3. Hence, the ROI selection algorithm should extract the region of main laser stripes and ignore redundant regions that might represent noise. First, the acquired image is converted from an RGB image into a binary image. Second, a search window of dimensions $m \times n$ is set at a certain position, such as the center of the image. Then, moving the search window to the densest high intensity region. For all pixels (x, y) inside the search window, the search window center position (x_{sw}, y_{sw}) is calculated as follows

$$x_{sw} = \frac{\sum_x \sum_y xI(x, y)}{\sum_x \sum_y I(x, y)} \quad (5-1)$$

$$y_{sw} = \frac{\sum_x \sum_y yI(x, y)}{\sum_x \sum_y I(x, y)} \quad (5-2)$$

Repeat moving the search window until convergence.

Fig.5-4 and Fig.5-5 describe the selection of ROI on the real image captured from car assembly in industrial environment.

5.3.2 *Laser stripe detection and classification*

The proposed extraction of laser stripes is most often carried out in two steps. First, the laser peaks are detected, normally by considering each row or column of the image as independent signals. Then, a laser peak linking process is applied, followed by a smoothing process. Considering only each row of the image, or column of the image to detect the laser [28] has a notable drawback. The laser profile of each of row or column signal corresponds to a Gaussian profile; thus, the obtained point with maximum luminance of the signal should correspond to the center of the laser line. However, the process can fail in the presence of noise, since the maximum luminance could correspond to spurious points. The proposed ROI selection in subsection 3.1 that acquires only the part of the image containing main laser stripes greatly reduces noise, but does not eliminate it completely. To reduce computational time while achieving highly accurate of laser stripe detection, we developed an efficient method to extract the laser stripes in the obtained image of the ROI. First, the image of the ROI is enhanced from the intensity difference between the laser region and background, using the following steps:

- Step1: Convert the two-dimensional image intensity matrix into a one-dimensional array
- Step 2: Arrange this array in descending order
- Step 3: Choose a threshold in accordance with initial conditions (satisfying that 5%-10% of intensity values that are maxima in the image are greater than the threshold)
- Step 4: Threshold the ROI image using selected threshold.

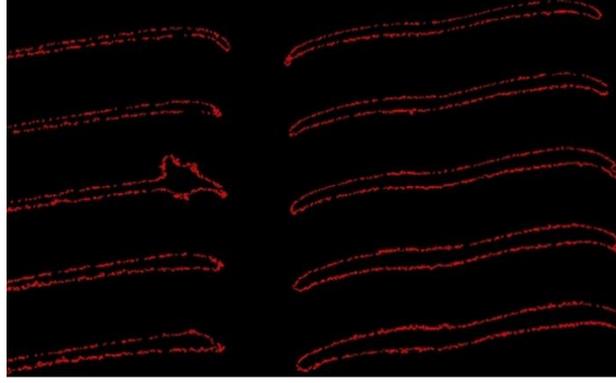


Fig. 5-6. Contour point detection image of the ROI after noise elimination.

Secondly, the contour detection algorithm is applied to detect contours of the image contained in the ROI. Then, all small regions are eliminated, as they could represent noise, using the following function:

$$\begin{cases} \Gamma(C_j) \geq \kappa & I(x_{jk}, y_{jk}) = 1 \\ \Gamma(C_j) < \kappa & I(x_{jk}, y_{jk}) = 0 \end{cases} \quad (5-3)$$

where $\Gamma(\bullet)$ is a contour area calculation function, C_j is the j^{th} contour, $I(x_{jk}, y_{jk})$ indicates the intensity of point (x_k, y_k) of the contour C_j , and κ is a threshold value. Fig.5-6 shows image of detected contours after noise elimination. After the small regions are eliminated, the remaining contours are used to estimate the laser curves. For each contour, a laser basic curve is estimated by taking the center points of the contour in every image column. Then, the basic curve formed from sub-pixelled points is approximated to a higher polynomial to get a smooth curve function.

After laser curves detection, we must classify these laser curves. Usually, a multi-line laser generator uses a color-coded system with multiple lasers of different color to classify the laser lines. However, this method is complex, computationally time consuming, and cannot be applied in this work using the mono color multi-line laser. Therefore, we propose an approach to classify the laser stripes by segmenting the image into different parts at which the laser lines deformed by the gap and flush on the object surface. First, we cluster the detected curves into N_s sets deformed by the gap and flush. Each set contains δ curves, where δ equals to N_c , and

N_c is the number of lines generated by the multi-laser lines generator. Then, the curves are scanned in ascending order in y , and the i^{th} laser curve is assigned to the i^{th} laser stripe.

5.3.3 Feature circle extraction

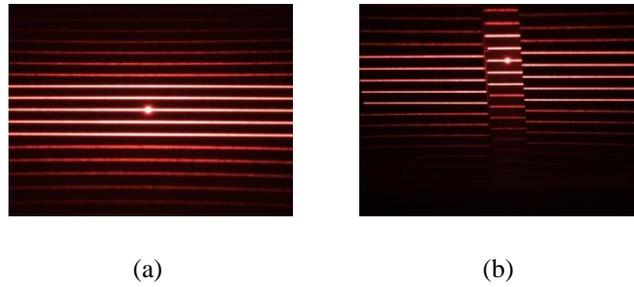


Fig. 5-7. Laser stripes on flat measured object. a) No gap and flush; b) Deformation of laser stripes in the presence of a gap and flush.

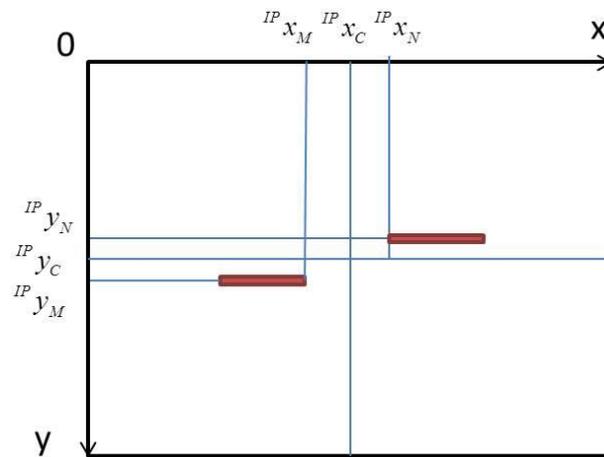


Fig. 5-8 Internal view of a gap and flush on a flat measurement object in the image plane.

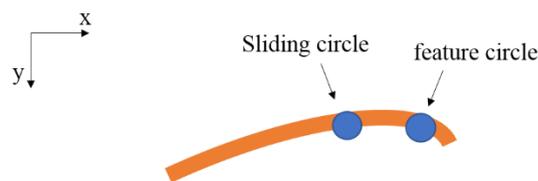


Fig. 5-9 Sliding circle.

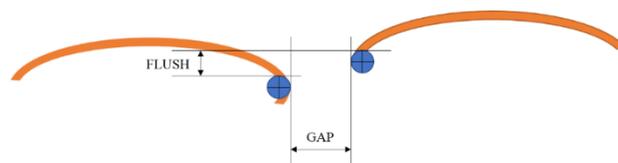


Fig. 5-10 Gap and flush measurement using feature circles.

The gap and flush are measured using feature points on the extracted laser stripes. These feature points contain geometrical information of the target object. For the flat and/or simple measured objects, the common way to select feature points is to select the beginning points and ending points of laser stripes. This method is illustrated in Fig. 5-7 and Fig.5-8. For every laser stripe, suppose that the detected ending point of the laser stripe on the left panel is M and the detected beginning point of the laser stripe on the right panel is N, as shown in Fig. 5-8. M and N are selected as feature points for the gap and flush measurement. However, current vehicles are designed with flowing contours, angled panels, edges and complex surfaces, such that the above method for a flat measurement object usually cannot be applied. In this paper, in order to guarantee a fast and effective gap and flush measurement, instead of finding feature points, we proposed a new sliding circle method to find feature circle as shown in Fig.5-9. The known-radius sliding circle slides on the detected laser curves. Then, the position at which a circle factor calculation function is maximized is selected as feature circle. The circle factor is calculated as follows:

$$\beta(C_i) = \frac{\sum_{h=1}^{N_s} E(s_h)}{N_s} \quad (5-4)$$

where β is the circle factor calculation function, C_i is the circle at sliding position i on the laser stripe curve, and N_s is the total number of pixel points on the sliding circle.

$$E(s_h) = \begin{cases} 1 & \text{if the test pixel } s_h \text{ is an edge point} \\ & \text{on the laser stripe contour} \\ 0 & \text{otherwise} \end{cases}$$

The gap and flush are measured by estimating the deformation of laser lines in the image plane using the detected feature circle, as shown in Fig.5-10.

5.4 Proposed gap and flush measurement method

Since the vision system is calibrated and the relative position between camera and laser is fixed, we can measure the gap and flush by estimating the deformation of laser lines in the image plane using the detected feature circle, as shown in Fig.5-10.



Fig. 5-11 A real camera-laser module setup.

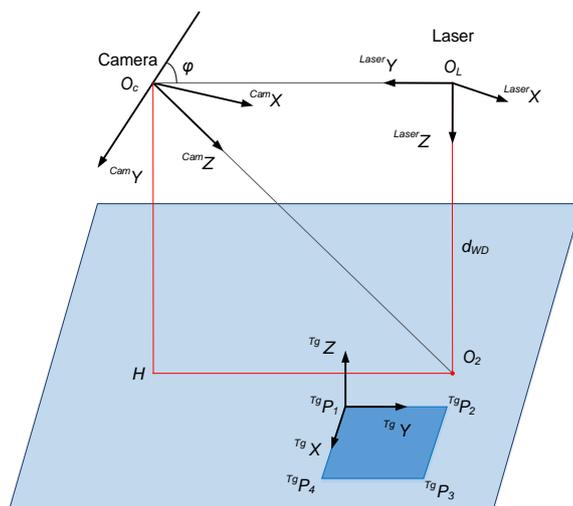


Figure 5-12 Camera and structured multi-line light calibration using rectangle model.

The extrinsic camera and multi-line laser is calibrated by using approach presented in Chapter 2. The calibration uses a rectangle, as can be seen in Fig. 5-11 and Fig.5-12. Suppose that the camera is perfectly calibrated as presented in section 2.2. We first estimate the relation between camera coordinate system and object coordinate system, followed by finding the relative pose including position and orientation of the laser with respect to the object coordinate system. Finally, the extrinsic calibration of camera and structured multi-line laser light parameters are refined.

Our next step is mapping the feature circle center in one way, feature point in another, into 3D coordinate system, which is laser coordinate system. Using the information in the calibration step, we can express the estimated curves function in the laser coordinate system. The physical transformation between the laser projector coordinate system and camera coordinate system is the sum of the effects of a rotation ${}_{Laser}^{Cam}R$ and a translation ${}_{Laser}^{Cam}T$, which were exactly calculated in camera and laser calibration process. In homogeneous coordinates, we can combine these within single matrix as follows:

$${}_{Laser}^{Cam}W = \begin{bmatrix} {}_{Laser}^{Cam}R, & {}_{Laser}^{Cam}T \end{bmatrix} = \begin{bmatrix} {}_{Laser}^{Cam}r_1 & {}_{Laser}^{Cam}r_2 & {}_{Laser}^{Cam}r_3 & {}_{Laser}^{Cam}T \end{bmatrix} \quad (5-5)$$

The relationship between camera coordinate system and laser coordinate system can be illustrated by

$${}^{Cam}P = {}_{Laser}^{Cam}W {}^{Laser}P \quad (5-6)$$

Suppose that the detected laser point ${}^{IP}p$ is the image of a real laser point ${}^{Laser}P$, and ${}_{Cam}^{IP}M$ is the camera intrinsic matrix; then, using camera pinhole model we have the following equation

$$w {}_{Cam}^{IP}M^{-1} {}^{IP}p = {}^{Cam}P \quad (5-7)$$

In laser coordinate system, each image plane is represented as follows:

$$\begin{aligned} & \begin{bmatrix} 0 & 0 & n_j & q_j \end{bmatrix} {}^{Laser}P = 0 \\ & \text{subject to } n_j^2 + q_j^2 \neq 0 \\ & \text{and } j = 0, \dots, N_c - 1 \end{aligned} \quad (5-8)$$

By solving equations (5), (6) and (7), the laser coordinate of laser point ${}^{Laser}P$ can be calculated as follows:

$${}^{Laser}P(j) = \Phi \left(\Lambda^T \Lambda \right)^{-1} \Lambda^T \begin{bmatrix} {}^{Cam}T \\ {}^{Laser}T \\ 0 \end{bmatrix} \quad (5-9)$$

where $\Phi = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, and $\Lambda_j = \begin{bmatrix} {}^{IP}M^{-1} {}^{IP}P & -\frac{{}^{Cam}r_1}{{}^{Laser}r_1} & -\frac{{}^{Cam}r_2}{{}^{Laser}r_2} & -\frac{{}^{Cam}r_3}{{}^{Laser}r_3} \\ 0 & 0 & n_j & q_j \end{bmatrix}$

```

Initialize j=0 ;
for (j=0;j<Nc;j++)
{
//Searching feature circle position
Initialize circle factor  $\beta = \beta(C_0)$ ,  $i = pos(j) = 0$ ;
while (i < size(laser_stripe))
{
if ( $\beta(C_i) > \beta$ )
{
 $\beta = \beta(C_i)$ 
 $pos(j) = i$ ;
}
i++;
}
//Mapping to 3D coordinate
 ${}^{IP}x_p = pos(j) - x$  ;//x coordinate of feature circle center
 ${}^{IP}y_p = pos(j) - y$  ;// y coordinate of feature circle center
 ${}^{Laser}P(j) = \Phi \left( \Lambda^T \Lambda \right)^{-1} \Lambda^T \begin{bmatrix} {}^{Cam}T \\ {}^{Laser}T \\ 0 \end{bmatrix}$ ;
}

```

Fig. 5-13 Pseudo code of the feature circle position 3D mapping algorithm for a gap and flush measurement

The pseudo code of the feature circle position 3D mapping algorithm for a gap and flush measurement is illustrated in Fig.5-13.

Then, the gap(j) and flush(j) of each laser stripe are measured by calculating the differences in the 3D positions between the two feature circles on the right side and left side of the laser stripe, as illustrated in Fig.5-10.

Finally, the total gap G and flush F are measured using the following equations:

$$G = \frac{\sum_{j=0}^{N_c} gap(j)}{N_c} \quad (5-10)$$

$$F = \frac{\sum_{j=0}^{N_c} flush(j)}{N_c} \quad (5-11)$$

5.5 Experimental results

Table 5-1 Current traditional gap and flush measurement sensors specifications and proposed gap and flush measurement sensor requirements.

Number	Content	Traditional measurement sensor	Proposed measurement sensor
1	Gap measurement error	$\pm 100\mu\text{m}$	$\pm 50\mu\text{m}$
2	Flush measurement error	$\pm 100\mu\text{m}$	$\pm 50\mu\text{m}$
3	Repeated measurement error	$\pm 200\mu\text{m}$	$\pm 100\mu\text{m}$
4	Working distance	$220 \pm 48(\text{mm})$	$220 \pm 48(\text{mm})$
5	Gross measurement time per measurement point	4 (seconds)	1 (second)

Table 5-2 Camera-laser module performance confirmation

S1		S2		S3	
Accuracy confirmation					
Ground truth values in mm					
$G = 10.164$	$F = 0.94$	$G = 10.172$	$F = 2.13$	$G = 10.188$	$F = 5.21$
Mean of measurement error in mm					
$\tilde{G}_a = 0.015$	$\tilde{F}_a = 0.006$	$\tilde{G}_a = 0.017$	$\tilde{F}_a = 0.0089$	$\tilde{G}_a = 0.0201$	$\tilde{F}_a = 0.0089$
Maximum of measurement error in mm					
$\tilde{G}_a = 0.0476$	$\tilde{F}_a = 0.0127$	$\tilde{G}_a = 0.042$	$\tilde{F}_a = 0.0199$	$\tilde{G}_a = 0.05$	$\tilde{F}_a = 0.0238$
Precision confirmation					
Mean of measured values in mm					
$\bar{G} = 10.16$	$\bar{F} = 0.935$	$\bar{G} = 10.173$	$\bar{F} = 2.126$	$\bar{G} = 10.192$	$\bar{F} = 5.213$
Mean of measurement error in mm					
$\tilde{G}_p = 0.0146$	$\tilde{F}_p = 0.0036$	$\tilde{G}_p = 0.0172$	$\tilde{F}_p = 0.0079$	$\tilde{G}_p = 0.0199$	$\tilde{F}_p = 0.0081$
Maximum of measurement error in mm					
$\tilde{G}_p = 0.0459$	$\tilde{F}_p = 0.0111$	$\tilde{G}_p = 0.0459$	$\tilde{F}_p = 0.0157$	$\tilde{G}_p = 0.051$	$\tilde{F}_p = 0.0209$

In this section, we first verify the accuracy of the proposed gap and flush measurement procedure and apparatus through experiments. Then, three laser-camera modules are implemented in a realistic vehicle assembling experiment, where each module includes a high-resolution camera module and a five-line laser generator. The camera module includes a focus lens, filter, and a GigE camera with a CMOS image sensor. The camera resolution is set as 2590×1942 and the focal length of 16 mm . The GigE camera transfers images at a speed of 10 frames per second through a LAN cable with a length of up to 100 m.

5.5.1 Camera-laser module performance confirmation

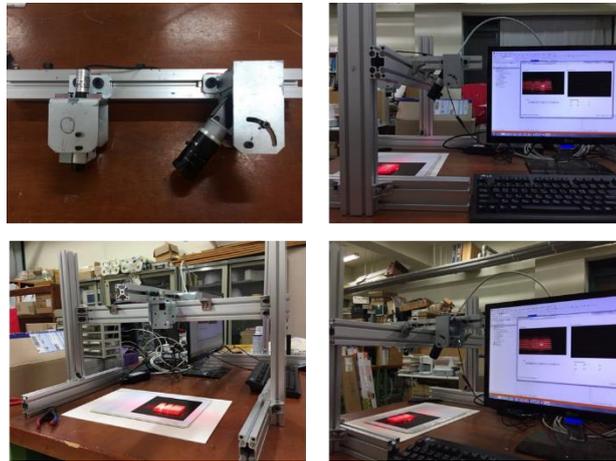


Fig. 5-14 Camera-laser module performance confirmation test.

As mentioned in section 2, we summarize the current gap and flush measurement sensor specification in Table 5-1. Requirements for the proposed gap and flush measurement system also indicated in Table 5-2. The real experiment setup is described in Fig 5-14. The ending points of laser stripes on the left panel and the beginning points of laser stripes on the right panel are selected as feature points.

5.6.1.1 Camera-laser module performance w.r.t changing of measurement position at the same working distance

This experiment is performed for different gap and flush values to examine the proposed camera-laser module measurement performance. The measurement working distance is kept at around 220 mm.

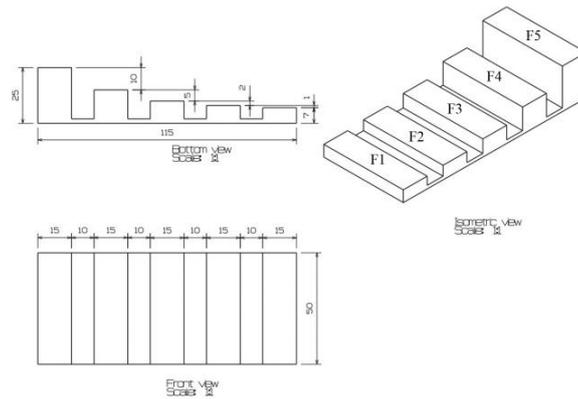
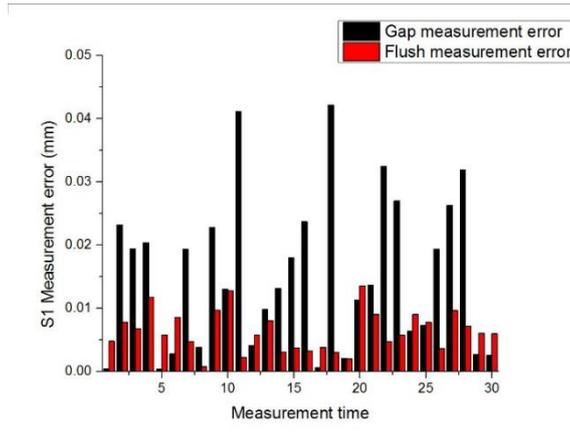
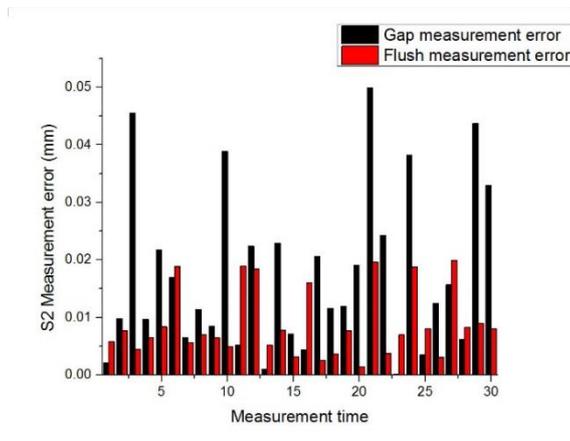


Fig. 5-15 Sample pattern

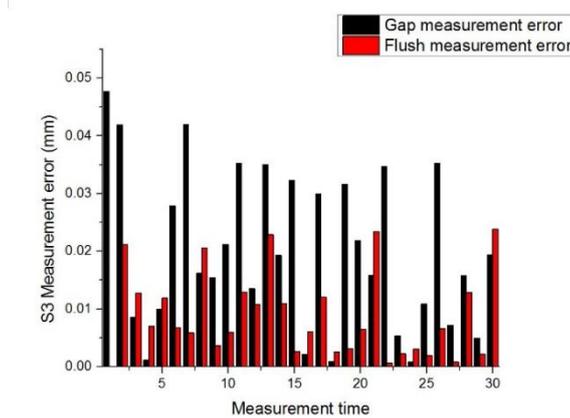
A known-size sample pattern, where the sample drawing is presented in Fig.5-15, is placed at various positions. The sample contains five flat planes F1, F2, F3, F4 and F5. Let denote the position at which we measure the gap and flush between F1 and F2, F2 and F3, F3 and F4 are S1, S2, S3, respectively. The proposed camera-laser module projects the laser lines on the measurement positions and measure the gap \hat{G} and flush \hat{F} . In order to examine the proposed camera-laser module measurement accuracy and precision, at each position, the gap and flush measurement test is performed independently 30 times.



(a)

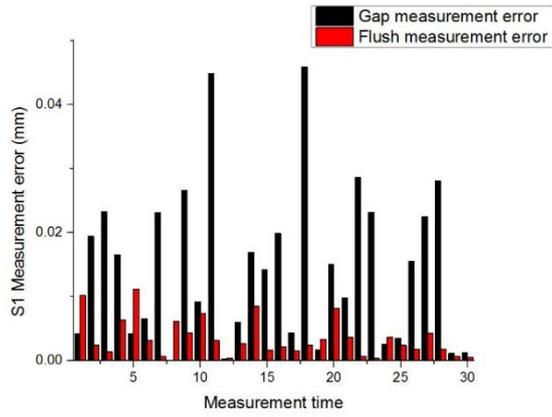


(b)

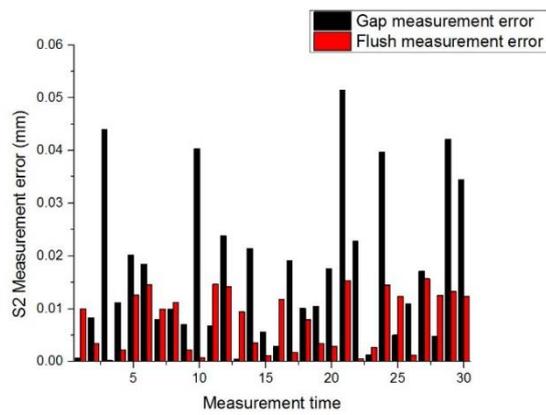


(c)

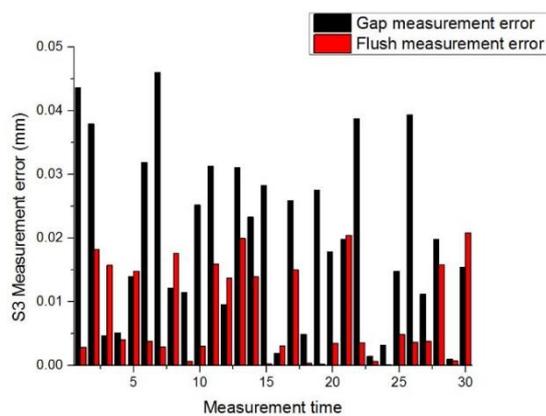
Fig. 5-16 Camera-laser module accuracy confirmation w.r.t changing of measurement position at the same working distance. (a) S1 measurement error; (b) S2 measurement error; (c) S3 measurement error.



(a)



(b)



(c)

Fig. 5-17 Camera-laser module precision confirmation w.r.t changing of measurement position at the same working distance. (a) S1 measurement error; (b) S2 measurement error; (c) S3 measurement error.

First, the measured gap \widehat{G} and flush \widehat{F} are compared with the ground truth values gap G and flush F to examine the proposed camera-laser module measurement accuracy. The ground truth values of gap and flush at position S1, S2, S3 are shown in Table 5-2. The measurement error of gap \widetilde{G}_a and flush \widetilde{F}_a are computed by

$$\widetilde{G}_a = |G - \widehat{G}| \quad (5-12)$$

$$\widetilde{F}_a = |F - \widehat{F}| \quad (5-13)$$

The measurement error of gap \widetilde{G}_a and flush \widetilde{F}_a are shown in Fig. 5-16. The measurement error varies from 0.006mm to 0.05mm. The maximum of measurement error is 0.05mm with the maximum of mean of measurement error is 0.0201mm, as can be seen in Table 5-2. The measurement error satisfies the requirement that the measurement error be less than or equal to 0.05mm. The maximum between two repeated measurement values is 0.09536 mm. The repeated measurement error satisfies the requirements shown in Table 5-1 that the repeated measurement error be less than 0.1mm.

Second, the measured gap \widehat{G} and flush \widehat{F} are compared with the means of measurement values \overline{G} and flush \overline{F} to examine the proposed camera-laser module measurement precision. The measurement error of gap \widetilde{G}_p and flush \widetilde{F}_p are computed by

$$\widetilde{G}_p = |G - \overline{G}| \quad (5-14)$$

$$\widetilde{F}_p = |F - \overline{F}| \quad (5-15)$$

The measurement error of gap \widetilde{G}_p and flush \widetilde{F}_p are shown in Fig. 5-17. Fig. 5-17 shows that the measurement error ranges from 0.0036mm to around 0.05mm. The maximum of measurement error is 0.05mm with the maximum of mean of measurement error is 0.0199mm, as can be seen in Table 5-2. The measurement error satisfies the requirement in Table 5-1 that the measurement error be less than 0.05mm. The maximum between two repeated measurement values is 0.09537 mm. The repeated measurement error satisfies the requirement that the repeated measurement error be less than 0.1mm

According to the above experimental results, the proposed measurement system achieves improved specifications with a higher accuracy and precision than current specifications for commercial gap and flush measurement systems shown in Table 5-1.

5.6.1.2. Camera-laser module performance w.r.t changing of measurement working distance

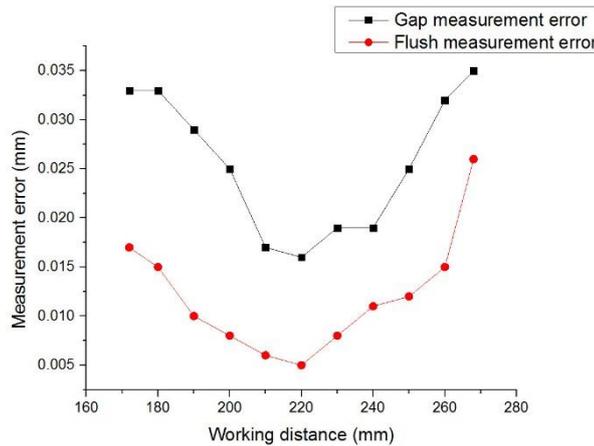


Fig. 5-18 Camera-laser module accuracy confirmation w.r.t changing of the measurement working distance.

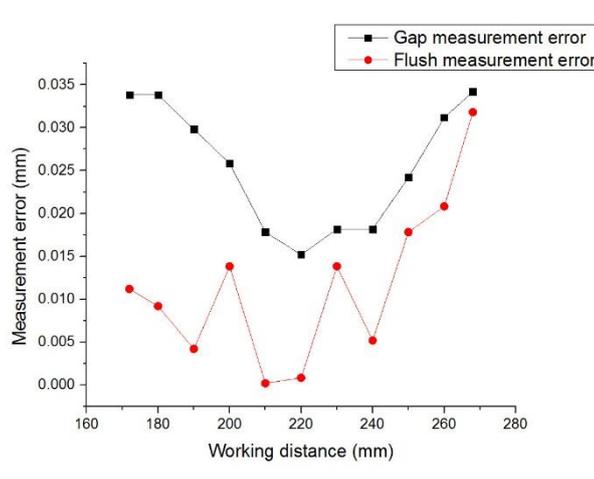


Fig. 5-19 Camera-laser module precision confirmation w.r.t changing of the measurement working distance.

This experiment evaluates the proposed camera-laser module performance at different working distance. The working distance is varied in the range of 220 ± 48 mm. The gap and flush are kept unchanged during the test. The gap and flush ground truth value are 10.166mm and

0.86mm, respectively. The measured gap and flush are compared with the ground truth values, mean of measured values to examine the measurement accuracy and precision, respectively.

The measurement accuracy and precision are shown in Fig. 5-18 and Fig. 5-19, respectively. The proposed gap and flush measurement system achieves the lowest error at the working distance around 220mm. The measurement error is less than 0.05mm and the difference between any two measurement values is less than 0.07mm. According to the experimental results, the proposed measurement system satisfies requirements shown in Table 5-1.

The experimental results confirm that measurement time at any working distance is confirmed less than 0.22 second through the experiment results.

According to the above experimental results, the proposed measurement system achieves improved specifications in comparison with the current commercial gap and flush measurement systems specifications shown in Table 5-1.

5.5.2 *Gap and flush measurement in real vehicle assembly*



Fig. 5-20 Three proposed calibrated camera-laser modules attached to the end-effector of the robot arm in a vehicle assembly application.

In this section, we applied our measurement system in a real vehicle assembly application. The poorest reflectivity car body, car body in black color, is selected. Three calibrated camera-laser modules are attached to the end-effector of the robot arm, as can be seen in Fig.5-20. The three calibrated camera-laser modules are moved to the top, left, and right, respectively. The vehicle is designed with flowing contours complex surfaces, so we used feature circles for gap and flush measurement.

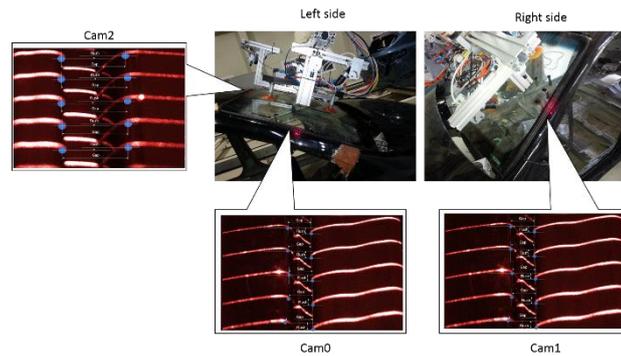


Fig. 5-21 Gap and flush measurement using proposed system in real vehicle assembly.

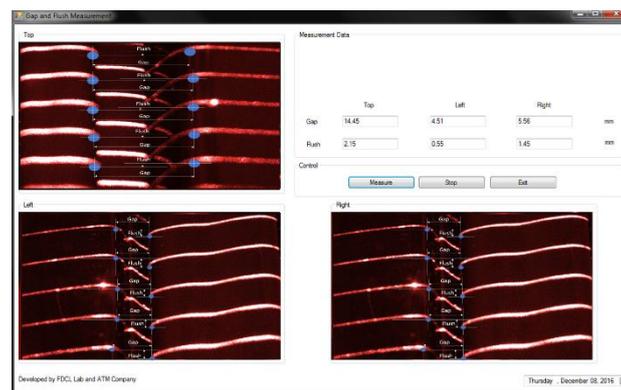


Fig. 5-22 The gap and flush measurement software GUI.

During the measurement process, at each module, the multi-line laser projects laser stripes onto the surface of the target object, and the camera captures the laser stripes produced by the intersections of the laser planes with the surface of the car body parts. The acquired image is then transferred to the computer through a LAN card. The image processing module of the gap and flush measurement software installed on the computer analyses the deformation of laser stripes to measure the gap and flush, as can be seen in Fig.5-21. Fig.5-22 shows the measurement software GUI. The top left, bottom left and bottom right panels show the image captured by the measurement modules at the top, left, and right position, respectively. We can measure three positions of the car body at the same time by clicking “Measure” in the bottom of the top right panel. The output data is shown in the measurement data box. In the experiment, the feature points are detected accurately. Section 5.1 also proved that the proposed gap and flush measurement between feature points operates well given high accurate and precise

measurement data. Thus, we do not have ground truth of the gap and flush between vehicle panels, but the measurement results are quite reasonable.

5.6 Chapter summary

In this chapter, we propose a vision-based non-contact gap and flush measurement system for a vehicle assembling system. The proposed approach makes use of the fusion of a multi-line laser generator and a high-resolution camera. In order to measure gap and flush, the multi-line laser of the camera-laser module projects laser stripes onto the surface of the measured object and the camera captures the stripes produced by the intersections of the laser planes with the target surface. The acquired image is then transferred to the computer through a LAN card. The image processing module of the gap and flush measurement software installed on the computer analyses the deformation of laser stripes to measure the gap and flush and displays the results on the GUI of the measurement software.

By using proposed image processing algorithms for adaptive laser stripe extraction and classification, circle feature detection, the gap and flush measurement system can deal with poor reflective objects with complex surface. The experimental results showed that the proposed gap and flush measurement system is an improvement over current gap and flush measurement sensors, achieving an improved higher specification. Moreover, the proposed system produces accurate measurements in a short time, 0.22 seconds, and satisfies vehicle assembly system requirements. The proposed measurement system can be applied for both portable and inline robot gap and flush measurement system. In many industrial robots, a camera is already equipped for measurement and other applications. The proposed gap and flush measurement module can be easily implemented in such cases by adding a single inexpensive multi-line laser. The usefulness of the proposed system is also demonstrated using real tests in the inline vehicle assembly system in Korea.

Chapter 6

MONOCULAR STRUCTURED LIGHT VISION-BASED SURFACE MEASUREMENT AND ITS APPLICATION IN UAV LANDING SURFACE DETERMINATION

6.1 Introduction

Landing on an unknown surface is an important problem for UAVs [87]. Real-time computer vision systems have been successfully implemented for UAVs landing [88-91]. There are two main challenging cases to be considered in this paper. Consider the first case of UAV landing on a sloping surface. Imagine a flying quadrotor equipped with an expensive payload in mission is forced to execute an emergency landing on unknown surface. The payload damage may be avoided if the quadrotor has been able to measure the ground surface geometry information so that it could design and execute a safe landing trajectory. The second case is that a quadrotor must execute 3D unknown surface reconstruction in missions, e.g. lunar mission. In both hypothetical scenarios, there is a desire to implementing UAV without prior knowledge of the surface's information by using onboard sensing. We contribute image processing algorithms for adaptive laser stripe extraction and classification by automatically selecting laser profile region and enhancing the intensity difference between the laser region and background. This approach is ideally suited for indoors or shade areas. The extracted laser stripes are then mapping onto 3D world coordinate to estimate the surface geometry information by using proposed 3D mapping approach. The proposed vision system has several advantages compared with current similar function systems. In vision-based system relying on feature detection, the main drawbacks are computationally expensive, difficult to implement in real-time and known pattern must be presented at the target surface, e.g. the landing site in UAV landing application. These problems are free in the proposed system since the image processing computational cost is minimized to permit real-time execution and no external patterns or features are required at

the surface; the only requirement is that the laser projections on the surface are detectable. The proposed system is also different from other laser-camera-based vision systems using laser beams which may look similar at the first sight. The laser-beam-based systems usually use laser dots and simple brightness detection to find laser projection on the surface. Thus, the common light dots, e.g. sunlight dots, can easily be a big problem for these systems. This problem is avoided in the proposed system by using automatic laser region extraction and laser profile classification. Additionally, the modest hardware is designed in the proposed system; there are only single camera and laser generator.

6.2 Laser profile extraction

Laser stripes extraction is commonly carried out in two steps. First, the laser peaks are detected, normally by considering each row or column of the image as independent signals. Then, a laser peaks linking process is applied following by smoothing process. However, carrying out by considering each row of the image, or column, to detect laser [8] has a big drawback. The laser profile of each of row, or column signal corresponds to a Gaussian profile, thus, the obtained point with maximum luminance of the signal should correspond to the laser center. However, the process can be fail by noise since the maximum luminance could correspond to spurious points. To save time consumption, the contour detection algorithm is applied to detecting contours. Then, all the small regions are eliminated as they may represent noise by using the following function:

$$\begin{cases} \Gamma(C_j) \geq \kappa & I(x_{jk}, y_{jk})=1 \\ \Gamma(C_j) < \kappa & I(x_{jk}, y_{jk})=0 \end{cases} \quad (6-1)$$

Where $\Gamma(\cdot)$ is contour area calculation function, C_j is the j^{th} contour, $I(x_{jk}, y_{jk})$ indicates intensity of point (x_k, y_k) of contour C_j and κ a threshold value.

After laser curve detection, we must classify these laser curves. Usually, in multi-line laser generator, they use color-coded to classify the laser lines. However, this method is complex, computational time consuming and cannot be applied in this work at which the mono color multi-line laser used. Hence, we proposed an approach to classify the laser stripes by segmenting the image into different parts at which the laser lines deformed by the gap and flush on the object surface in industrial environment. Firstly, we cluster the detected curves into 2 sets deformed by the gap and flush. Each set contains \mathcal{S} curves, where \mathcal{S} equals to N_c , and N_c is the number of lines generated by the multi-laser lines generator. Then, scanning them in y ascending order, due to the system configuration, the i^{th} laser curve is decided to belong to the i^{th} laser stripe.

6.3 Proposed ground surface estimation

The objective of this section is to reconstruct the surface based on obtained 3D coordinate points.

First, laser points are mapped into 3D coordinate which is chosen as laser coordinate. By using the information in the calibration step, we can express the estimated curves function in the laser coordinate, which is chosen as world coordinate. The physical transformation part between laser projector coordinate and camera coordinate is the sum of the effects of some rotation ${}_{Laser}^{Cam}R$ and some translation ${}_{Laser}^{Cam}T$, which were exactly calculated in camera and laser calibration process. In homogeneous coordinate, we can combine these within single matrix as following

$${}_{Laser}^{Cam}W = \begin{bmatrix} {}_{Laser}^{Cam}R, & {}_{Laser}^{Cam}T \end{bmatrix} = \begin{bmatrix} {}_{Laser}^{Cam}r_1 & {}_{Laser}^{Cam}r_2 & {}_{Laser}^{Cam}r_3 & {}_{Laser}^{Cam}T \end{bmatrix} \quad (6-2)$$

The relationship between camera coordinate and laser coordinate can be illustrated by equation

$${}^{Cam}P = {}_{Laser}W^{Cam} {}^{Laser}P \quad (6-3)$$

Suppose that the detected laser point ${}^{IP}P$ is the image of real laser point ${}^{Laser}P$, ${}^{IP}M_{Cam}$ is camera intrinsic matrix, using camera pinhole model we have the following equation

$$W_{Cam} {}^{IP}M^{-1} {}^{IP}P = {}^{Cam}P \quad (6-4)$$

In laser coordinate, each laser plane is represented as follows:

$$\begin{aligned} & \begin{bmatrix} 0 & 0 & n_j & q_j \end{bmatrix} {}^{Laser}P = 0 \\ & \text{subject to } n_j^2 + q_j^2 \neq 0 \\ & \text{and } j = 0, \dots, N_c - 1 \end{aligned} \quad (6-5)$$

Let us denote $\Lambda_j = \begin{bmatrix} {}^{IP}M_{Cam}^{-1} {}^{IP}P & -{}_{Laser}r_1 & -{}_{Laser}r_2 & -{}_{Laser}r_3 \\ 0 & 0 & n_j & q_j \end{bmatrix}$

By solving equation (5), (6) and (7), the laser coordinate of laser point ${}^{Laser}P$ can be calculated as following

$${}^{Laser}P(j) = \Phi (\Lambda^T \Lambda)^{-1} \Lambda^T \begin{bmatrix} {}_{Laser}T_{Cam} \\ 0 \end{bmatrix} \quad (6-6)$$

with $\Phi = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Second, the surface is reconstructed using approach proposed in [19]. The interpolation problem is express as a linear algebra problem

$$\mu_p x = y \quad (6-7)$$

where the vector x has $nx \times ny$ length with nx is the number of points in the x direction, and ny is the number of grid nodes in the y direction. Thus μ_p has np rows, corresponding to each data point supplied by the user, and $nx \times ny$ columns.

At every point of the surface, interpolation algorithm will try to make the partial derivatives of the surface in neighboring cells to be equal. As the result, we obtain the linear equation

$$v_p x = 0 \quad (6-8)$$

where the derivatives are approximated using finite differences of the surface at neighboring points.

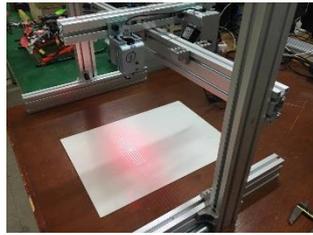
Finally, we obtain an optimization problem: find x to minimize

$$\|\mu_p x - y\|^2 + \tau \|v_p x\|^2 \quad (6-9)$$

where τ is a parameter that we can choose in order to control the relative grid stiffness.

6.4 Experiment Results

In this section, we verify the accuracy of the proposed vision system through real experiments using a camera and a five-line laser generator. The test platform was implemented in C/C++. The camera resolution is 2590×1942 and the focal length is $16mm$. The working distance is less than $1000mm$.



(a)



(b)

Figure 6-1 The plane surface angle measurement setup; (a) Experiment testbed, (b) monocular structured laser vision module.

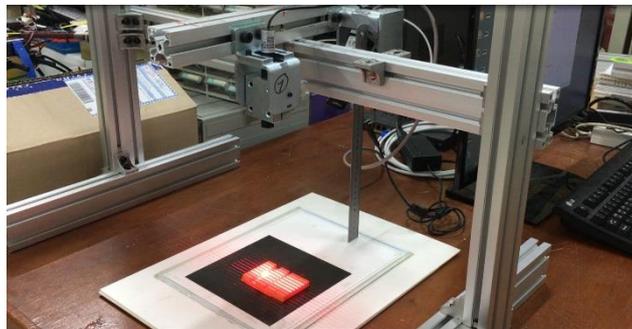


Figure 6-2 Real 3D reconstruction testbed.

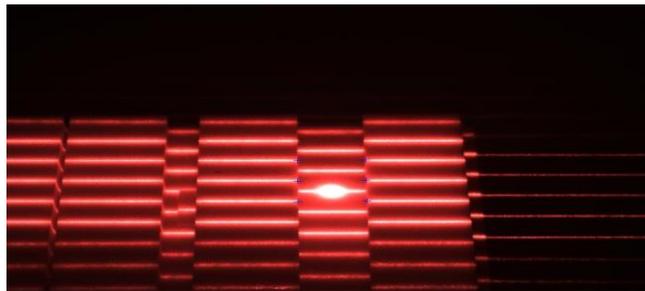


Figure 6-3 The complex surface captured by the camera.

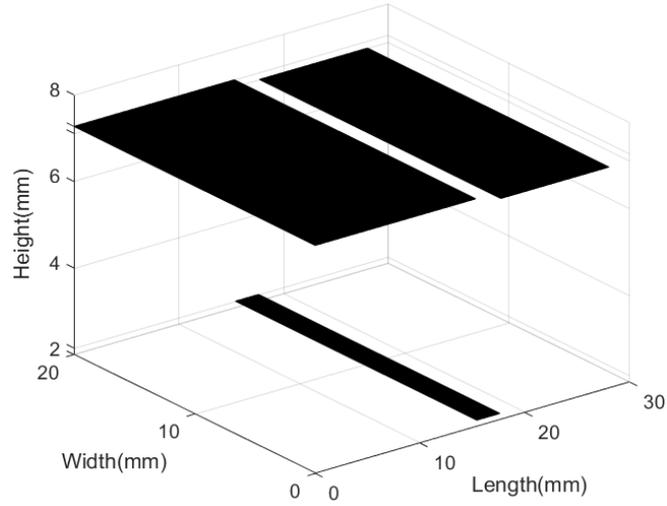


Figure 6-4 3D surface reconstruction.

Table 6-1 Angle between the flat surface and horizontal estimation

Ground truth angle (ϖ) in degrees					
$\varpi = 0.50^\circ$		$\varpi = 30.55^\circ$		$\varpi = 60.05^\circ$	
Estimated angles ($\hat{\varpi}$) and estimation error ($\tilde{\varpi}$) in degrees					
$\hat{\varpi}$	$\tilde{\varpi}$	$\hat{\varpi}$	$\tilde{\varpi}$	$\hat{\varpi}$	$\tilde{\varpi}$
0.457	0.043	30.474	0.076	59.919	0.131
0.546	0.046	30.471	0.079	60.104	0.054
0.550	0.050	30.633	0.083	60.117	0.067
0.457	0.043	30.636	0.086	60.171	0.121
0.450	0.050	30.640	0.090	59.965	0.085
0.561	0.061	30.643	0.093	59.952	0.098
0.564	0.064	30.463	0.087	59.888	0.162
0.567	0.067	30.471	0.079	59.954	0.096
0.550	0.050	30.624	0.074	59.951	0.099
0.563	0.063	30.627	0.077	60.088	0.038
Mean angle measurement error in degrees					
$\tilde{\varpi} = 0.0537$		$\tilde{\varpi} = 0.0824$		$\tilde{\varpi} = 0.0951$	
Total mean estimation error in degrees: $\tilde{\varpi} = 0.077$					

5.1 Performance w.r.t the changing of surface angle

This experiment is performed for different angle measurement to examine the proposed vision system angle estimation performance. In this experiment, the proposed vision system is

applied to measure the angle between a flat plane surface and the horizontal. The measurement results are compared with the ground truth angle obtained from inclinometer. The experiment setup is described in Fig. 6-1.

The flat plane surface is placed at various positions with different angles w.r.t. the horizontal. The proposed system projects the laser lines on the flat plane surface and estimates the angle between the plane and the horizontal $\hat{\varpi}$. The estimated angle $\hat{\varpi}$ is then compared with the ground truth angle ϖ . The estimation error $\tilde{\varpi}$ is computed by

$$\tilde{\varpi} = |\varpi - \hat{\varpi}| \quad (6-10)$$

Three different angles between the plane and the horizontal are measured and shown in Table 6-1. Each position is performed 10 measurements. As can be seen from Table 6-1, the proposed vision system gives a high accuracy. The total angle estimation error is $\tilde{\varpi} = 0.077$ where the maximum estimation error is $\tilde{\varpi} = 0.162$.

5.2 Performance w.r.t the changing of distance from surface to proposed vision system

This experiment evaluates the proposed vision system performance in distance measurement. In this experiment, a known-size complex surface with different height flat planes is reconstructed by using proposed vision system. The experiment setup is shown in Fig.6-2. Fig.6-3 and Fig.6-4 illustrate the image of the laser projected on the surface and surface 3D reconstruction, respectively. The average distance from the surface to the vision system measurement error is around 0.05 mm.

5.2 Performance w.r.t the changing of ground surface height

This experiment is performed for different gap and flush values on the surface. The measurement working distance is kept at around 220 mm. A known-size sample pattern, where the

sample drawing is presented in Fig.6-5, is place at various positions. The sample contains five flat planes F1, F2, F3, F4 and F5. Let denote the position at which we measure the gap and flush between F1 and F2, F2 and F3, F3 and F4 are S1, S2, S3, respectively. The proposed camera-laser module projects the laser lines on the measurement positions and measure the gap \widehat{G} and flush \widehat{F} . In order to examine the proposed camera-laser module measurement accuracy and precision, at each position, the gap and flush measurement test is performed independently 30 times.

The measurement results are compare with the ground truth values. The measurement error varies from 0.006mm to 0.05mm. The maximum of measurement error is 0.05mm with the maximum of mean of measurement error is 0.0201mm as can be seen in Fig.6-6

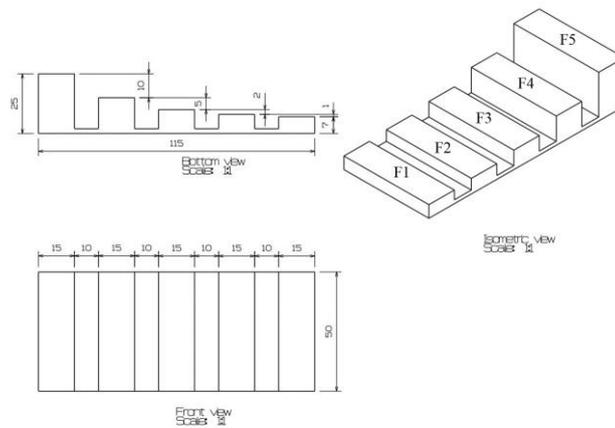
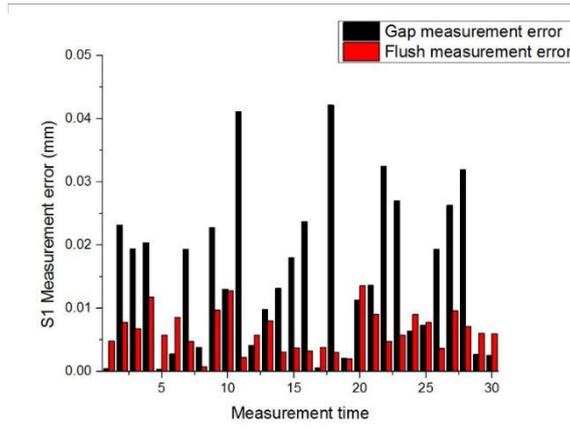
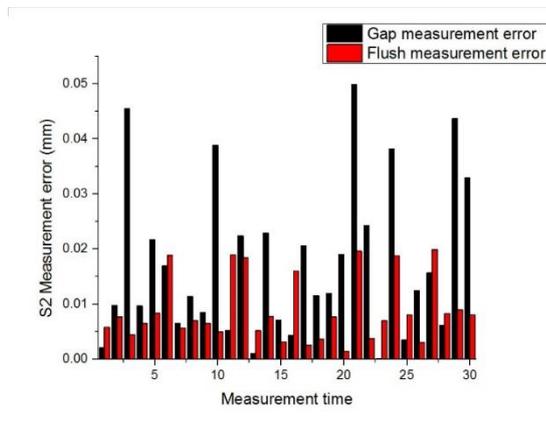


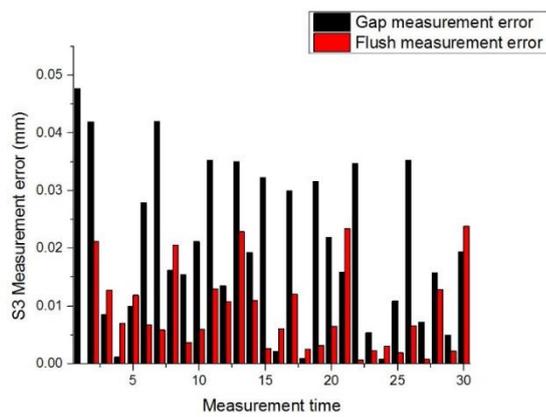
Figure 6-5 Sample pattern.



(a)



(b)



(c)

Figure 6-6 3D surface reconstruction measurement. (a) S1 measurement error; (b) S2 measurement error; (c) S3 measurement error.

6.5 Chapter summary

Inexpensive laser generator and a monocular camera vision system are proposed to estimate the orientation of and distance to static ground surface. We contribute image processing algorithms and 3D reconstruction for ground surface geometry estimation. Experiment results show that the proposed system can estimate the ground surface geometry with high accuracy. In many UAVs, a camera is already equipped for SLAM and other applications. The proposed surface geometry information estimation can be easily implemented in such cases by adding a single inexpensive multi-line laser.

Chapter 7

CONCLUSION

This thesis we contribute 3D measurement using monocular structured light vision. The 3D measurement system is calibrated by using a plane containing a rectangle. The proposed extrinsic calibration of camera and structured light can get reasonable calibration results which have been suitable robotic monocular structured light vision systems.

In 2D coordinate system, the monocular vision measures the target object position by dynamic fusion of color and shape. We build and update the statistical model of the object in different feature spaces. Then by fusing possible target positions in the feature spaces the final target position is validated. The experiment results demonstrate robust object position measurement performance under various scenarios, including perspective changes, drastic illumination changes, and occlusions.

The 3D measurement sensor using CMOS camera with double rate technology to capture the laser stripes generated from low-power generator reflex mirrors from the object surface to ensure real-time measurements. The laser stripes are digitized and map in to 3D coordinate system to perform 3D object measurement. The experimental results demonstrated that the measurement system can deal with complex surface with high accuracy, satisfying the demanding 3D measurement requirements in industrial applications.

The 3D measurement system is light, simple in structure, and cheap, all characteristics which make it suitable for a compact industrial robot. In many robot systems, a camera is already equipped for object grasping and other applications. The proposed 3D measurement system can be easily implemented in such cases by adding a single inexpensive source light.

PUBLICATIONS

A. International Journals

1. Thi-Trang Tran, Cheolkeun Ha, “Extrinsic Calibration of a Camera and Structured Multi-Line Light using a Rectangle,” *Int. J. Precis. Eng. Manuf.*, Volume 19, Issue 2, pp 195–202, 2018 (SCIE)
2. Thi-Trang Tran, Cheolkeun Ha, “Non-contact Gap and Flush Measurement using Monocular Structured Multi-line Light Vision for Vehicle Assembly,” *International Journal of Control, Automation and Systems*, 2018. (Accepted -SCIE)
3. Thi-Trang Tran, Cheolkeun Ha, “Self-tuning Proportional Double Derivative Like Neural Network Controller for a Quadrotor,” *International Journal of Aeronautical and Space Sciences*, 2018. (Selected for publication-SCIE)
4. Thi-Trang Tran, Cheolkeun Ha, “Dynamic Fusion of Color and Shape for Accurate Circular Object Tracking,” *Lecture Notes in Artificial Intelligence (LNAI)*, 2018. (Accepted -SCOPUS)
5. Thi-Trang Tran, Cheolkeun Ha, “Slippage Estimation Using Sensor Fusion,” *Intelligent Computing Theories and Application, Lecture Notes in Computer Science*, vol 9772, 2016. (SCOPUS)

B. International Conferences

6. Xuan Vinh Ha, Cheolkeun Ha and Trang Thi Tran, “An adaptive control for the under-actuated gripper in the object grasping tasks using the CCD camera and Laser-Based Slippage Estimation,” In *The 21st Int. Conf. on Mechatronics Technology*, Ho Chi Minh, Vietnam, 2017.
7. Thi-Trang Tran, Cheolkeun Ha, “Ground Surface Geometry Estimation using Monocular Structured Light Vision,” In *Asia-Pacific International Symposium on Aerospace Technology*, Seoul, Korea, 2017.

8. Thi-Trang Tran, Cheolkeun Ha, “Non-contact Gap and Flush Measurement using Monocular Structured Light Vision,” In *16th International Conference on Control, Automation and Systems*, HICO, Gyeongju, Korea, 2016.
9. Tran Thi Trang, Cheolkeun Ha, “An efficient approach for circular shape target recovery,” In *2015 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*, Busan, Korea, 2015.
10. Tran Thi Trang, Cheolkeun Ha, “Irregular Moving Object Detecting and Tracking Based on Color and Shape in Real-time System,” In *2013 International Conference on Computing, Management and Telecommunications (ComManTel)*, Ho Chi Minh, Vietnam, 2013.

REFERENCES

- [1] B. Culshaw, G. Pierce, and P. Jun, "Non-contact measurement of the mechanical properties of materials using an all-optical technique," *IEEE Sensors J.*, vol. 3, no. 1, pp. 62–70, Feb. 2003.
- [2] M.-T. Ha, H.-Y. Kim, C.-G. Kang, "A precision stopping measurement device to automatically detect position errors of an urban train at railway stations", *International Journal of Control, Automation and Systems*, vol. 15, Issue 2, pp. 848-856, 2017.
- [3] "Simultaneous Ultrasonic Measurement of Thickness and Speed of Sound in Elastic Plates using Coded Excitation Signals", *IEEE Transactions on Ultrasonics, Ferroelectrics, and frequency control*, vol. 14, no. 8, 2016.
- [4] A. Caarullo and M. Parvis, "An ultrasonic sensor for distance measurement in automotive applications," *IEEE Sensors J.*, vol. 1, no. 3, pp. 143–147, Oct. 2001.
- [5] Z.S. Lim, S.T. Kwon, M.G. Joo, "Multi-object identification for mobile robot using ultrasonic sensors", *International Journal of Control, Automation and Systems*, vol. 10, Issue 3, pp. 589-593, 2012.
- [6] Franc -a Joa ão G, Gazziro, Ma ´rio A, Ide Alessandro N, Saito, Jose ´ H. A 3D scanning system based on laser triangulation and variable field of view. In: *Conference on image processing*, IEEE, 2005.
- [7] V.-D. Hoang, K.-H. Jo, "Automatic calibration of camera and LRF based on morphological pattern and optimal angular back-projection error", *International Journal of Control, Automation and Systems*, vol. 13, Issue 6, pp. 1436-1445, 2015.
- [8] Richard Hartley and Andrew Zisserman, "Multiple View Geometry in computer Vision", Cambridge University Press, March 2004.
- [9] M.-H. Le, H.-H. Trinh, V.-D. Hoang, and K.-H. Jo, "Automated architectural reconstruction using reference planes under convex optimization," *International Journal of Control, Automation and Systems*, vol. 10, Issue 2, pp. 362-373, 2012.

- [10] Y.S. Suh, N.H.Q. Phuong, H.J. Kang, “Distance estimation using inertial sensor and vision”, *International Journal of Control, Automation and Systems*, vol. 11, no. 1, pp. 211-215, 2013.
- [11] J.K. Oh, S. Lee, C.-H. Lee, “Stereo vision based automation for a bin-picking solution”, *International Journal of Control, Automation and Systems*, vol. 11, no. 1, pp. 211-215, 2013.
- [12] Lowe, D.: Distinctive image features from scale-invariant keypoints, cascade filtering approach. *IJCV* 60(2004) 91 – 110.
- [13] H. Bay, T. Tuytelaars, and L. Van Gool. SURF: Speeded Up Robust Features. In *ECCV* (1), pages 404–417, 2006.
- [14] N. Dalai, and B. Triggs, “Histograms of oriented gradients for human detection,” *Computer Vision and Pattern Recognition* (vol. 1, pp. 886–893), June 2005.
- [15] G. Mori, S. Belongie, and J. Malik, “Efficient shape matching using shape contexts,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27 (2005): 1832–1837.
- [16] J. Matas, O. Chum, M. Urba, and T. Pajdla, “Robust wide baseline stereo from maximally stable extremal regions,” *Proceedings of the British Machine Vision Conference* (pp. 384–396), 2002.
- [17] A. C. Berg and J. Malik, “Geometric blur for template matching,” *IEEE Conference on Computer Vision and Pattern Recognition* (vol. 1, pp. 607–614), Kauai, Hawaii, 2001.
- [18] A. Bosch, A. Zisserman, and X. Muñoz, “Image classification using random forests and ferns,” *IEEE International Conference on Computer Vision*, Rio de Janeiro, October 2007.
- [19] Song, J. Y., Park, H. Y., Kim, H. J. and Jung, Y. W., “Development of Defect Inspection System for PDP ITO Patterned Glass,” *Int. J. Precis. Eng. Manuf.*, Vol. 7, No. 3, pp. 18-23, 2006.
- [20] Mahmoud, D., Khali, A. & Younes, M., “Optimization of the Longitudinal Resolution of a Structured Light Shape Reconstruction System, a DOE Approach,” *Int. J. Precis. Eng. Manuf.*, Vol. 16, No. 9, pp. 1935-1939, 2015.
- [21] Im, S., Lee, Y., Kim, J., “A Solution for Camera Occlusion Using a Repaired Pattern from a Projector,” *Int. J. Precis. Eng. Manuf.*, Vol. 17, No. 11, pp. 1443-1450, 2016.

- [22] D. Ortin, J. Montiel, A. Zisserman, "Automated multisensor polyhedral model acquisition," In Proc. IEEE International Conference on Robotics and Automation, pp. 1007–1012, 2003.
- [23] Baltzakis, H., Argyros, A., Trahanias, P., "Fusion of laser and visual data for robot motion planning and collision avoidance," *Machine Vision and Application*, Vol. 15, No. 2, pp. 92-100, 2003.
- [24] Y. Liu, R. Emery, "Using EM to learn 3d environment models with mobile robots," In Proc. 18th International Conference on Machine Learning, 2001.
- [25] Zhang, Z., "A Flexible New Technique for Camera Calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 11, pp. 1330-1334, 2000.
- [26] Hartley, R. and Zisserman, A., "Multiple View Geometry in Computer Vision," Cambridge University Press, pp. 310-324, 2003.
- [27] Anwar, H., Din, I., and Park, K., "Projector Calibration for 3D Scanning using Virtual Target Images," *Int. J. Precis. Eng. Manuf.*, Vol. 13, No. 1, pp. 125-131, 2012.
- [28] Kimura, M., Mochimaru, M., and Kanade, T., "Projector Calibration using Arbitrary Planes and Calibrated Camera," *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-2, 2007.
- [29] Martynov, I., Kamarainen, J.-K., and Lensu, L., "Projector Calibration By "Inverse Camera Calibration," in: *Scandinavian Conference on Image Analysis*, Heyden, A., Kahl, F., (Eds.), Springer, pp. 536-544, 2011.
- [30] Luo, H., Xu, J., Binh, N. H., Liu, S., Zhang, C., and Chen, K., "A Simple Calibration Procedure for Structured Light System," *Optics and Lasers in Engineering*, Vol. 57, pp. 6-12, 2014.
- [31] Chen, R., Xu, J., Chen, H., Su, J., Zhang, Z., and Chen, K., "Accurate Calibration Method for Camera and Projector in Fringe Patterns Measurement System," *Applied Optics*, Vol. 55, No. 16, pp. 4293-4300, 2016.

- [32] Fuqiang, Z., Guangjun, Z., "Complete calibration of a structured light stripe vision sensor through planar target of unknown orientations", *Image and Visual Computing*, Vol. 23, No. 1, pp. 59-67, 2005.
- [33] Sun, Q., Hou, Y., Tan, Q., & Li, G., "A Flexible Calibration Method Using the Planar Target with a Square Pattern for Line Structured Light Vision System", *PLoS ONE*, 9(9), e106911, 2014.
- [34] Tom B., Steven M., Richard G., "Design And Calibration of A Hybrid Computer Vision And Structured Light 3D Imaging System," 2011.
- [35] W.M.K.W.M. Khairrosfaizal, A.J. Nor'aini , "Eyes detection in facial images using circular Hough transform , " 5th Signal Processing & Its Applications (CSPA 2009) Kuala Lumpur. IEEE , pp.238–242, 2009.
- [36] M. Hassabalth, Shun Ido, "Eye Detection Using Intensity and Appearance Information," *MVA2009 IAPR Conference on Machine Vision Applications*, Yokohama, JAPAN, May 20-22 2009.
- [37] Ito Y., Ohyama W., Wakabayashi T., Kimura F., "Detection of eyes by circular Hough transform and histogram of gradient," *21st International Conference on Pattern Recognition (ICPR 2012)*; 11-15, Tsukuba. IEEE. pp. 1795-1798, 2012.
- [38] Da Fontoura Costa, L., Marcondes Cesar Jr. R "Shape analysis and classification: Theory and Practise," CRC Press, Boca Raton, FL, USA.
- [39] N. Barnes and A. Zelinsky, "Real-time radial symmetry for speed sign detection," *Intelligent Vehicles Symposium*, 14-17, pp. 566-57, 2004.
- [40] BenjaminHoferlin, KlausZimmermann, "Towards reliable traffic sign recognition," *Intelligent Vehicles Symposium*, pp. 324-329, 2009.
- [41] Mainzer, "Genetic algorithm for traffic sign detection," *International Conference Applied Electronics 2002*, Pilsen, Czech Republic. pp. 129–132, 2002.

- [42] Xinguo Yu, Hon Wai Leong, Changsheng Xu, Qi Tian, "Trajectory-based ball detection and tracking in broadcast soccer video," *IEEE Transactions on Multimedia*, Vol. 8, No. 6, 2006.
- [43] Zhang HY, Wu YD, Yang F, "Ball detection based on color information and Hough transform," *2009 International Conference on Artificial Intelligence and Computational Intelligence*, pp. 393-397, 2009.
- [44] J. Illingworth, J. Kittler, "A survey of the Hough transform," *Computer Vision, Graphics, and Image Processing* 44, pp. 87-116, 1988.
- [45] L. Xu, E.Oja, P.Kultanen, "A new curve detection method: Randomized Hough transform," *Pattern Recognition Letters* 11, pp. 331-338, 1990.
- [46] J. Illingworth, J. Kittler, "The adaptive Hough transform," *IEEE Transactions on Pattern Recognition and Machine Intelligence* 9, pp. 690-698, 1987.
- [47] R. K.K. Yip, P. K.S. Tam, D. N.K. Leung, "Modification of Hough transform for circles and ellipses detection using a 2-dimensional array," *Pattern Recognition* 25, pp. 1007-1022, 1992.
- [48] W. C. Y. Lam, S. Y. Yuen, "Efficient techniques for circle detection using hypothesis filtering and Hough transform," *IEEE Proc. Visual Image Signal Process* 143 (5), pp. 292-300, 1996.
- [49] Lu, W., Tan, J.L., "Detection of incomplete ellipse in images with strong noise by iterative randomized Hough transform (IRHT)" *Pattern Recognition* 41(4), pp. 1268-1279, 2008.
- [50] Yuen, S., Ma, C, "Genetic algorithm with competitive image labelling and least square," *Pattern Recognition* 33, pp. 1949-1966, 2000.
- [51] Yao, J., Kharna, N., Grogono, P, "Fast robust GA-based ellipse detection," *Proc. 17th Int. Conf. on Pattern Recognition ICPR-04, Cambridge, UK.* pp. 859-862, 2004.
- [52] Victor Ayala_Ramirez, Carlos H. Garcia-Capulin, Arturo Perez_Garcia, Raul E. Sanchez_Yanez, "Circle detection on the images using genetic algorithms," *Pattern Recognition Letters*, 2005.

- [53] E. Cuevas, F. Wario, D. Zaldivar, M. Pérez_Cisneros, "Circle detection on images using learning automata," *IET Computation Vis.* 6(2), pp. 121-132, 2012.
- [54] M. J. Swain, D. H. Ballard, "Color indexing," *Int. J. Computer Vision* 7(1), pp. 11-32, 1991.
- [55] A. R. Smith, "Color gamut transform pairs", *Computer Graphics* 12, pp. 12–19, 1978.
- [56] Zhang Y, Wang H, Liang Z, Tan M, Ye W and Lian B, "Existence probability map based circle detection method," *Computer Engineering and Application*, 2006.
- [57] S. Suzuki and K. Abe, "Topological structural analysis of digital binary images by border following," *Computer Vision, Graphics and Image Processing* 30, pp. 32–46, 1985.
- [58] Carsten Steger, "On the calculation of arbitrary moments of polygons," technical report FGBV-96-05, 1996.
- [59] Tran Thi Trang, CheolKeun Ha, "Irregular moving object detecting and tracking in real-time system," *International Conference on Management and Telecommunications (ComManTel 2013)*, Hochiminh city, Vietnam. pp. 415-419, 2013.
- [60] Gary R. Bradski, "Computer vision face tracking for use in a perceptual user interface," *Intel Technology Journal* 2(2), pp. 13-27, 1998.
- [61] Matthias Z, Joachim D, Heinrich N, "Entropy based camera control for visual object tracking," *IEEE ICIP 2002*, vol. 3, pp. 901-904, 2002.
- [62] Xie S, Lou J, Gong Z, Ding W, Zou H, Fu X, "Biomimetic control of pan-tilt-zoom camera for visual tracking based-on an autonomous helicopter," *IROS 2007. IEEE/RSJ.*, pp. 2138-2143, 2007.
- [63] Murad A H, Andrew D. Bagdanov, Jordi Gonazález and F. Xavier Roca, "Reactive object tracking with a single PTZ camera," *20th International Conference on Pattern Recognition*, pp. 1690-1693, 2010.
- [64] D. Douglas and T. Peucker, "Algorithms for the reduction of the number of points required for represent a digitized line or its caricature," *Canadian Cartographer* 10, pp. 112–122, 1973.

- [65] Zeng, X., Liu, Z., “A learning automaton based algorithm for optimization of continuous complex function,” *Inf. Sci.* 174, pp. 165–175, 2005.
- [66] H. Shinoda: Contact sensing, A state of the art. In *Japanese J. Robot. Soc. Jpn.*, vol. 20, no. 4, pp. 385–388, Japan (2002)
- [67] M. H. Lee and H. R. Nicholls: Tactile sensing for mechatronics-a state of the art survey. *Mechatronics*, vol. 9, pp. 1–31, (1999)
- [68] Dahiya, R.S., Metta, G., Valle, M., Sandini, G. Tactile Sensing—From Humans to Humanoids. *IEEE Trans. Robot.* 26, pp. 1–20, (2010)
- [69] Ferrier, N.J., Brockett, R.W.: Reconstructing the Shape of a Deformable Membrane from Image Data. *Int. J. Robot. Res.*, 19, pp. 795–816. 14, (2000)
- [70] Saga, S., Kajimoto, H., Tachi, S.: High-resolution tactile sensor using the deformation of a reflection image. *Sens. Rev.*, 27, 35–42. 15, (2007)
- [71] Johnson, M.K., Adelson, E.H.: Retrographic sensing for the measurement of surface texture and shape. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1070–1077, USA (2009)
- [72] Goro, O., Ashish, D., Norinao, W., Nobuhiko, M.: Vision Based Tactile Sensor Using Transparent Elastic Fingertip for Dexterous Handling. Chapter 7, *Mobile Robots: Perception and Navigation*, InTechOpen, (2007)
- [73] Yuji Ito, Youngwoo Kim, Chikara Nagai, Goro Obinata: Vision-based tactile sensing and shape estimation using a fluid-type touchpad. *IEEE Transaction on Automation*, Vol. 9, Issue 4, pp. 734-744, (2012)
- [74] Ito Y., Kim Y., Obinata G.: Contact Region Estimation Based on a Vision-Based Tactile Sensor Using a Deformable Touchpad. *Sensors J.* 14, pp. 5805-5822, (2014)
- [75] J. Ueda, A. Ikeda, T. Ogasawara: Grip-Force Control of an Elastic Object by Vision-Based Slip-Margin Feedback During the Incipient Slip. *IEEE Transactions on Robotics* , Vol.21 , Issue 6, pp. 1139-1147, (2005)

- [76] Tom, B., Steven, M., Richard, G: Design and calibration of a hybrid computer vision and structured light 3D imaging system. In: Proceeding of the 5th International Conference Automation, Robotics and Applications (ICARA), pp. 441-446. IEEE, Wellington (2011)
- [77] T. H. Wang, M. C. Lu, W. Y. Wang, C. Y. Tsai: Distance measurement Using Single Non-metric CCD Camera. In: Proceedings of the 7th WSEAS Int. Conf. on Signal Processing, Computational Geometry & Artificial Vision, Athens, Greece (2007).
- [78] <http://www.lmicorporation.com/277.html>
- [79] <http://www.nextsense.at/en/calipri/applications-solutions/gap-measurement-fold-evaluation/portable-gap-and-flush-measurement-system.php>
- [80] <https://www.autoevolution.com/news/ford-invests-100m-in-robots-with-special-vision-36089.html>
- [81] <https://www.baslerweb.com/en/>
- [82] Arbelaez, P., Maire, M., Fowlkes, C., and Malik, J., “Contour Detection and Hierarchical Image Segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 33, No. 5, pp. 898-916, 2011.
- [83] Forest, J., Salvi, J., Cabruja, E., Pous, C., “Laser stripe peak detector for 3d scanners. A FIR filter approach”, In: *Pattern Recognition, Proceedings of the 17th International Conference on ICPR 2004*, vol. 3, 2004.
- [84] Kumar, S., Tiwari, P. K., and Chaudhury, S., “An Optical Triangulation Method for Non-Contact Profile Measurement,” *Proc. of IEEE International Conference on Industrial Technology*, pp. 2878-2883, 2006.
- [85] Park, J. B., Lee, J. G., Lee, M. K., and Lee, E. S., “A Glass Thickness Measuring System using the Machine Vision Method,” *Int. J. Precis. Eng. Manuf.*, Vol. 12, No. 5, pp. 769-774, 2011.
- [86] Tran, TT. & Ha, C., “Extrinsic Calibration of a Camera and Structured Multi-Line Light using a Rectangle,” *Int. J. Precis. Eng. Manuf.*, Vol 19, Issue 2, pp. 195-202, 2018
- [87] Gupte, P. Mohandas, and J. Conrad, “A survey of quadrotor unmanned aerial vehicles,” In *Proc. IEEE Southeastcon*, Orlando, FL, USA, Mar. 2012, pp. 1–6.

- [88] Sharp, C., Shakernia, O., and Sastry, S., "A Vision System for Landing an Unmanned Aerial Vehicle," Proceedings of the 2001 IEEE International Conference on Robotics and Automation, Vol. 2, IEEE Publ., Piscataway, NJ, 2001, pp. 1720–1727.
- [89] Sinopoli, B., Micheli, M., Donato, G., and Koo, T., "Vision Based Navigation for an Unmanned Aerial Vehicle," Proceedings of the 2001 IEEE International Conference on Robotics and Auto, Vol. 2, IEEE Publ., Piscataway, NJ, 2001, pp. 1757–1764. Ferrier, N.J., Brockett, R.W., "Reconstructing The Shape of A Deformable Membrane from Image Data," Int. J. Robot. Res., 19, pp. 795–816. 14, 2000.
- [90] Saripalli, S., Montgomery, J., and Sukhatme, G., "Visually Guided Landing of an Unmanned Aerial Vehicle," IEEE Transactions on Robotics and Automation, Vol. 19, No. 3, 2003, pp. 371–380.
- [91] John, A., Lee, D., Lee, T., "Laser-based Guidance of a Quadrotor UAV for Precise Landing on an Inclined Surface," In Proc. Amer. Control Conf., Portland, OR, USA, Jun. 2014, pp. 1210–1215.
- [92] Kim, MS., Shin, HJ. & Park, YK., "Design Concept of High-Performance Flexible Tactile Sensors with A Robust Structure," Int. J. Precis. Eng. Manuf., Vol. 13, Issue 11, pp 1941-1947, 2012.
- [93] J. R. D'Errico. (2006, August 26). Surface Fitting Using Gridfit. Available: <http://www.mathworks.com/matlabcentral/fileexchange/8998-surface-fitting-using-gridfit>.