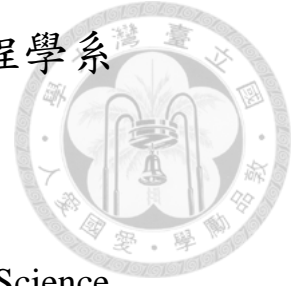國立臺灣大學電機資訊學院電機工程學系

碩士論文

Department of Electrical Engineering

College of Electrical Engineering and Computer Science

National Taiwan University

Master Thesis

利用色彩資訊及深度資訊之室內行動機器人

對行人偵測與追蹤

Pedestrian Detection and Tracking with Indoor Mobile

Robot Using both Color Information and Distance

Information

林俊榮

Jun-Rong Lin

指導教授:連豐力 博士

Advisor: Feng-Li Lian, Ph.D.

中華民國一百零三年七月

July 2014
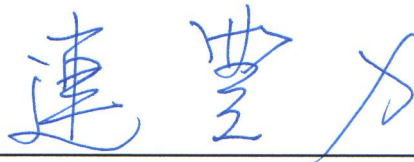
國立臺灣大學碩士學位論文
口試委員會審定書

利用色彩資訊及深度資訊之室內行動機器人對行人偵測與追蹤

Pedestrian Detection and Tracking with Indoor Mobile Robot Using both Color Information and Distance Information

本論文係林俊榮君（學號 R01921067）在國立臺灣大學電機工程學系完成之碩士學位論文，於民國一百零三年七月二十九日承下列考試委員審查通過及口試及格，特此證明
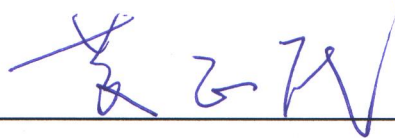
口試委員：

連豐力
（指導教授）

簡忠漢

李後燦

黃正民

系主任　　廖婉君

誌謝

即將要畢業，在研究所的兩年收穫良多。一路上的研究走來，首先要感謝的就是我的指導教授連豐力博士，在這嚴謹的研究態度以及看待事情的觀點使得我這兩年有很大的收穫。研究一個問題時，我終於懂得如何去定義一個問題、提出想法並且去解決這個問題。以及在這些複雜的研究問題上，如何用有效率的簡報讓大家了解，是我這兩年的成長。在口試時，感謝三位口試委員簡忠漢博士、李後燦博士及黃正民博士提供許多的建議，使得這本論文可以更加完整。

感謝 NCSLabers 的大家，由於你們的陪伴，才能讓我這麼順利的將研究完成。感謝志明學長，在研究遇到問題時，總是提供給我一些解決問題的想法。感謝意淳學姊，在我的研究上，總是會敘述一些我沒有想到的研究盲點。感謝俊兆學長，在研究上以及生活上，都受到你不少的幫忙。感謝飛紘學長，你每次說話的梗總是很好笑，研究上也會提供我閱讀論文的訣竅。感謝中易學長，你每次總會有出人意表的表現，而且在研究上甚至幫助我許多。感謝凱翔學長，每次都不言其煩的跟我討論研究的事情，並且製造很多實驗室的歡樂。感謝柏逸，在我遇到研究與生活瓶頸時，會與我一起討論。感謝家維，高等機器人學課程跟你討論讓我學習很多。感謝兆良，你生活上的率性，研究上獨特的想法都值得我去學習。感謝詠政，你在研究上的積極，以及生活上相處都是讓我值得效法。還有這一年相處過來的學弟們。感謝孝淇，在我的研究上會適時給我鼓勵，並且全力提供協助我。感謝登翔，在研究方面提供我一些想法，生活上的一些資訊也都會不吝的提供。感謝鍾潛，有時候在做實驗時還會麻煩到你，並且提供我做實驗時可以有不同的方式去表達。感謝博旭，你總是一個冷面笑匠，在研究中也是一位高深莫測的人。感謝偉哲，有些時候會跟你借 Matlab 書籍來參考，以及和你討論一些新想法。很開心可以在這間實驗室度過研究所生涯，少了你們其中一個，研究的生活就會變了調。

最後，感謝我的家人。爸爸、媽媽、姊姊以及哥哥。有你們陪著我、包容我，這研究的路才能夠如此順遂。未來的日子，我會繼續努力。在此，謹以此論文奉獻給予所有的有緣人，如果有提供給你們任何想法，這就是我的榮幸。

林俊榮 謹誌

中華民國一〇三年九月十九日

利用色彩資訊及深度資訊之室內行動機器人對行人偵測與追蹤

研究生:林俊榮   指導教授:連豐力 博士

國立臺灣大學 電機工程學系

摘要

生活之中,人機互動是一門重要的學問。如果機器人可以隨時隨地去跟人互動,那麼在生活上必定會便利不少。其中與人互動的情況中,追蹤行人就是其中一個必須研究的課題。機器人追蹤行人的問題,主要分成兩類:自我定位與地圖建構、行人偵測與追蹤。這些問題都是在做機器人追蹤行人時會碰到的項目。

以自我定位與地圖建構來說,要在未知路徑的情況下,得到行動機器人周邊及本身存在的資訊是一件困難的事。在本篇論文裡,主要解決行動機器人編碼器的硬體因素造成殘存誤差。利用粒子群最佳化演算法去找出更準確的機器人位置,然後利用這些資訊可以估測出機器人周遭環境地圖。由於雷射測距儀是一個相當精確的儀器,所以得到的距離資訊準確度是很高,在實驗部分會有驗證。在得到行動機器人的位置和周遭地圖之後,就可以分辨出靜態物地圖和動態物地圖。對行動機器人而言,是非常有用的資訊。然而,動態物的地圖推測會有一些問題,

這部分會在論文中詳細的描述到。

另外，在行人偵測部分，除了用雷射測距儀所得到的動態點之外，也可以用點的分群、大小以及資料可信度去判斷。另外，色彩資訊的添加可以當成是一個更強韌的條件。偵測邊緣形狀的霍式圓形轉換、色彩判定以及頭的所在位置都是偵測行人的準則。利用這些準則，幾乎可以更精確的去判斷行人。基於行動機器人追蹤，由於追蹤目標行人在生活環境中可能會有靜態障礙物擋住或者附近突然出現的行人，造成雷射測距儀的資料誤判。場景設定是在一般研究室或宿舍，通常人的穿著色彩分佈及紋理分佈會不盡相同，所以可以透過色彩資訊去加以判斷。並且使用距離資訊空間連續性去做一個強韌性判斷的準則。利用這些準則，就可以強韌地去判斷行人，並且去做目標行人的追蹤。在本篇論文中，最主要的就是用色彩資訊去解決在使用雷射測距儀時，行人偵測的判斷以及突然出現的人進而造成追蹤判斷的錯誤。

在實驗結果顯示出在這些環境當中，可以解決雷射測距儀資料連結的不足，並且可以做到追蹤目標行人。

關鍵字：雷射測距儀、全向攝影機、機器人自我定位、動態物偵測、行人偵測、行人追蹤。

Pedestrian Detection and Tracking with Indoor Mobile Robot Using both Color Information and Distance Information

Student: Jun-Rong Lin   Advisor: Dr. Feng-Li Lian

Department of Electrical Engineering
National Taiwan University

ABSTRACT

In daily life, that mobile robot communicates with pedestrian needs many tasks. The applications are used in guided vehicle, shopping cart, or office assistance. In this thesis, the tasks include self-localization, mapping, pedestrian detection, and target pedestrian tracking in unknown indoor environment.

To do self-localization and mapping, the accurate odometry of mobile robot is important. However, skidding and slipping can induce that the odometry is not equal to the real distance. In this thesis, particle swarm optimization (PSO) algorithm correct odometry in unknown indoor environment. The combination of the self-localization and the mapping is referred to as the simultaneous localization and mapping (SLAM) [39: Birk & Carpin 2006]. In the SLAM, once odometry of mobile robot is known, building a map is also a task which can be effectively solved at the

same time [39: Birk & Carpin 2006]. Afterward, moving object detection is based on the precise map.
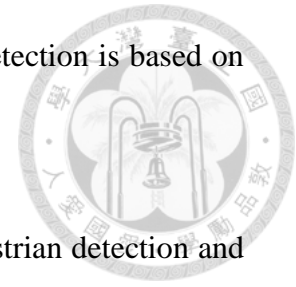
After the moving objects are detected, the next steps are pedestrian detection and target pedestrian tracking. In the pedestrian detection, the color image is regarded as an additional condition for the judgment based on the laser range finder (LRF) scan. In target pedestrian tracking, owing to pillars hindering or new pedestrians appearing, the data association may be error between two consecutive LRF scans. In this thesis, a method based on color distribution and color texture to track pedestrian in color images is proposed. The experiment demonstrates the target pedestrian in the new pedestrian appearing and the pillars hindering. Through the experiments, the performance of pedestrian detection and target pedestrian tracking is not good. However, the performance of pedestrian in color image is low owing to the resolution.

In the future, the detection and tracking moving object (DATMO) with LRF scan in Chapter 3 can combine the pedestrian detection and target pedestrian tracking with color image in this thesis.

Keywords: laser range finder, omnidirectional camera, robot self-localization, moving objects detection, pedestrian detection, target pedestrian tracking.

Contents

List of Figures

x

List of Tables

Chapter 1

Introduction

In this chapter, Section 1.1 states pedestrian tracking for mobile robot applications in daily life. In pedestrian tracking, two problems often occur in the sounding environment. The tasks include self-location, mapping, pedestrian detection, and target pedestrian tracking. The method of solutions states in Section 1.2. In this thesis, Section 1.3 states the contribution in this research field. Section 1.4 states the architecture in this thesis.

1.1 Motivation

The mobile robot applications in the surrounding environment are widely discussed such as automatic guided vehicle [1: Seifert & Kay 1995], shopping cart [2: Nishimura et al. 2007], or office assistance [3: Chen et al. 2011]. In the office assistance application, target pedestrian tracking is often hindered by pillar or is associated to other pedestrian. To track pedestrian, the research works include self-localization, mapping, pedestrian detection, and target pedestrian tracking. The primary objective is to construct the perception system using both the distance scan

and the color image to track the pedestrian for both LRF and omnidirectional camera mounting on mobile robot.
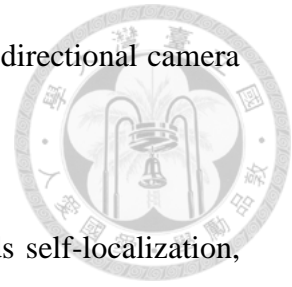
In the real-life, the mobile robot tracking the pedestrian needs self-localization, mapping, detection, and tracking. Self-localization and mapping are two of the fundamental capabilities for mobile robot [39: Birk & Carpin 2006]. Detection and tracking are also discussed in [24: Chang & Lian 2012], [27: Carballo et al. 2010], [35: Dalal & Triggs 2005]. However, the following two cases often occur in the sounding environment. There are many pillars in Figure 1.1 scene. One case is that the target pedestrian is hindered by pillar. The other case is that new pedestrian appears. The detail states in Section 1.2.

In this thesis, the objective can provide office assistance mobile robot in unknown indoor environment. The mobile robot in real-life can deal with the static pillar hindering the target pedestrian and other pedestrian suddenly appearing. After the target pedestrian hindered by pillars in the Figure 1.1 scene, the target pedestrian is not predictable with the LRF scan information. However, the color distribution is regarded as a condition for the judgment. In summary, the additional color image can provide additional information to detect the pedestrian and track the target pedestrian.

(a)   (b)   (c)
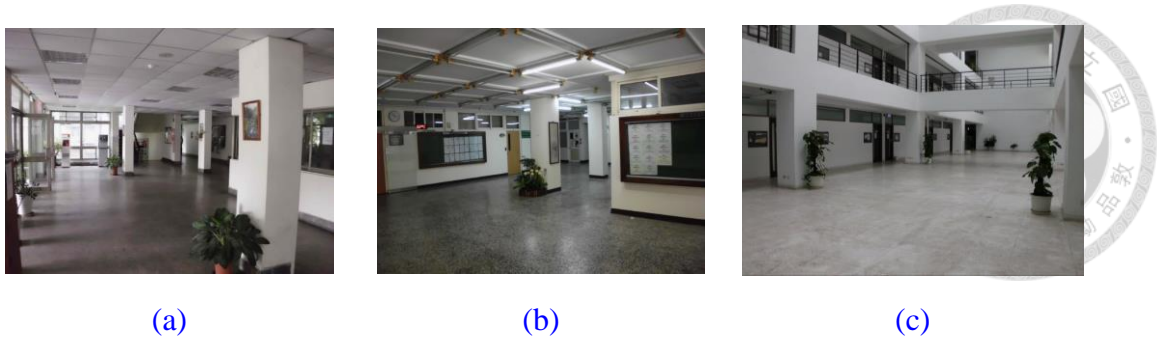
Figure 1.1 The Scene Shows the Pillars in the Indoor Environment.
(a) The 1st Men's Dorm in NTU
(b) Electrical Engineering Building No. 2 in NTU
(c) Ming-Da Building in NTU

1.2 Problem Formulation

That the mobile robot tracks the target pedestrian needs many tasks including localization, mapping, detection, and tracking. However, that the mobile robot tracks the target pedestrian is difficult in dynamic environment with pillars or other pedestrians. Figure 1.2 shows that multiple pedestrians make the mobile robot confuse with pillars. In Figure 1.2, there are pedestrian A1, pedestrian A2, pillar, robot, and LRF scan. Figure 1.2(a) shows the LRF scan at time t-2. The robot detects the pedestrian A1, pedestrian A2, and pillar. However, the robot only detects the pillar at time t-1 with LRF scan in Figure 1.2(b). At time t, two possible results appear in Figure 1.2(c) and Figure 1.2(d). However, the robot cannot distinguish between pedestrian A1 and pedestrian A2 from LRF scan. It is pillar hindering case. And Figure 1.3 shows that new pedestrian appears near the target pedestrian. In Figure 1.3,

there are pedestrian A1, pedestrian A2, pillar, robot, and LRF scan. In Figure 1.3(a), the robot only detects the pedestrian A1 with LRF scan. However, the pedestrian A1 and pedestrian A2 are detected with LRF scan in Figure 1.3(b). The robot cannot distinguish between pedestrian A1 and pedestrian A2 owing to the position. It is new pedestrian case. Two cases are discussed in this thesis.

In the mobile robotic field, the mobile robot localization and mapping is important. For robot position, skidding and slipping can induce mobile robot odometry is not equal to the real distance. Therefore, PSO algorithm [38: Li et al. 2011] corrects the mobile robot odometry through map construction. For map construction, the occupancy gird is mainly used in construct the map [39: Birk & Carpin 2006].

(a)

(b)

(c)

(d)
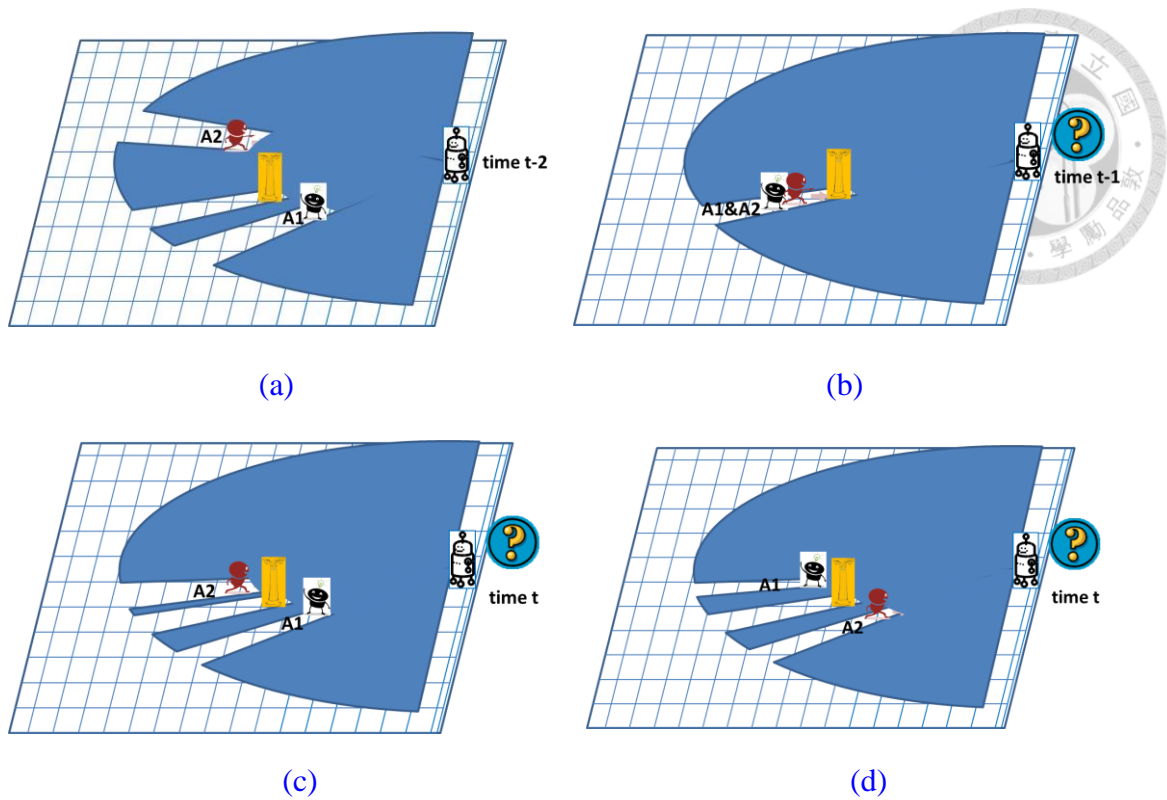
Figure 1.2 Dynamic Environment with Pillars Occlusion (with LRF Scan Grid Map in Specific Plane)

(a) The robot detects A1 and A2 candidate pedestrians and pillar at time t-2

(b) The robot only detects pillar at time t-1

(c) and (d) are two case making the robot confuse the candidate pedestrians with LRF sensor at time t.

(a)

(b)
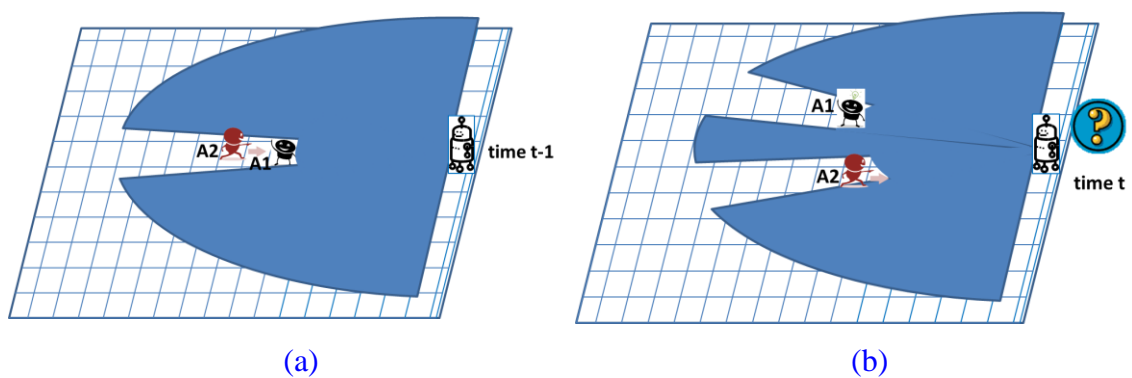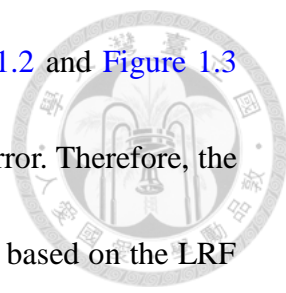
Figure 1.3 Dynamic Environment with New Pedestrian Appearing Nearby (with LRF Scan Grid Map in Specific Plane)

(a) The robot only detects A1 candidate pedestrian at time t-1.

(b) That A1 and A2 candidate pedestrians appear simultaneously makes the robot confuse at time t.

Pedestrian detection and target pedestrian tracking in Figure 1.2 and Figure 1.3

are difficult. The data association between two LRF scans may be error. Therefore, the

color image is regarded as an additional condition for the judgment based on the LRF

scan. In [22: Wolf & Sukhatme 2004], Wolf and Sukhatme propose static map and

dynamic map. The static map includes many dynamic obstacles owing to inverse

observation model. Since inverse observation model predicts that the state from

unknown to occupied is static object, the pedestrian detected with LRF scan by robot

may be regarded as static object. Therefore, the pedestrian detection with color image

is necessary. In this thesis, it is necessary to adopt the features of head to detect

pedestrian. Both Hough circle transform and color distribution are used in head

detection in each candidate pedestrian [48: Zhao et al. 2012]. Although the pedestrian

candidates are selected, the target pedestrian tracking is still a difficult issue because

of data association in unexpected position in Figure 1.2 and Figure 1.3. Using color

distribution and local binary map (LBP) algorithm is a powerful method to track

pedestrian in the dynamic environment with pillars or other pedestrians. The color

distribution means that the histogram is calculated in each color channel. And the LBP

algorithm calculates the relative neighbor value in each pixel [49: Rahimi et al. 2013].

With the target pedestrian, the mobile robot can continuously track the target

pedestrian.

1.3 Contribution

The thesis proposes a system structure includes localization, mapping, pedestrian detection, and tracking target pedestrian.

For the localization and the mapping, the initial position and the map construction are two problems. The PSO algorithm [38: Li et al. 2011] corrects the mobile robot odometry. In experimental result, the PSO algorithm compares to the ICP algorithm. For ICP algorithm [24: Chang & Lian 2012], the distance error is minimized. However, local optimal solution is a problem in scan matching. The PSO algorithm in this thesis can overcome the problem. In static map construction and dynamic map construction through inverse observation model in [22: Wolf & Sukhatme 2004], the color feature can robustly judge the moving pedestrians in previous unknown area.

For pedestrian detection and target pedestrian tracking in pillar hindering or new pedestrian appearing, the data association may be error [12: Ueda et al. 2011]. In color image, the Hough circle transformation, size, and color distribution are methods to judge pedestrian and track the same person. In this thesis, the pillar hindering case and the new pedestrian appearing case are solved in pedestrians with different color space and color texture.

The experimental results show in Chapter 5. The experimental results and

analysis shows the performance of the pedestrian detection and the target pedestrian tracking in unknown indoor environment. In the future works, the DATMO with LRF scan in Chapter 3 can combine the pedestrian detection and target pedestrian tracking with color image in this thesis.
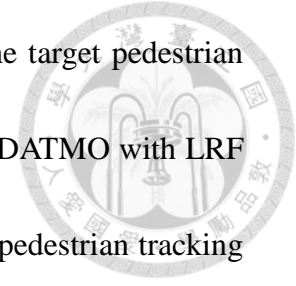
1.4 Organization of the Thesis

This thesis includes six chapters. The remainder of this thesis is organized as follows. Chapter 2 states the literature of past research. This chapter includes two sections: simultaneous localization and mapping (SLAM), and pedestrian detection and tracking. Chapter 3 states robot localization and map construction. The tasks have the LRF usage, the robot localization, and the map construction in specific environment. Chapter 4 states the omnidirectional camera structure, the pedestrian detection by LRF scan spatial continuity, image color feature, and image edge feature, and target tracking by image color histogram, image LBP, and LRF scan. Chapter 5 shows the experimental result and analysis. In addition, the PSO algorithm compares to the ICP algorithm in SLAM. Both conclusions and feature works are presented in Chapter 6.

Chapter 2

Literature Survey
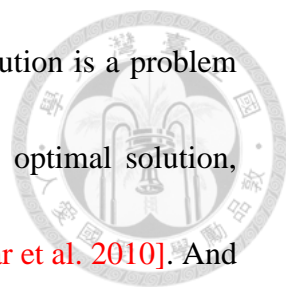
This chapter states the literature survey in the mobile robotic field. Section 2.1 states the simultaneous localization and mapping including self-localization and mapping. Self-localization and mapping are two of the fundamental capabilities for mobile robot [39: Birk & Carpin 2006]. In addition, pedestrian detection and target pedestrian tracking are researchable topics for the mobile robot. Section 2.2 states pedestrian detection and target pedestrian tracking. Figure 2.1 shows the SLAM categories in Section 2.1. And Figure 2.2 shows pedestrian detection and target pedestrian tracking categories in Section 2.2.

2.1 Simultaneous Localization and Mapping

SLAM is an important topic for a mobile robot in unknown indoor environment. Although many sensors can be selected, the laser range finder (LRF) often is used to SLAM [15: Wu et al. 2013], [16: Rusdinar et al. 2010]. LRF is a sensor commonly used owing to its accuracy in distance measurement [15: Wu et al. 2013].

In terms of methods, iterative close point (ICP) algorithm is commonly used [17:

Zhang 1994], [18: Lu & Milios 1994]. However, local optimal solution is a problem in scan matching. To overcome the problem relating to the local optimal solution, particle filter (PF) is proposed to correct the error pose [16: Rusdinar et al. 2010]. And extended Kalman filter (EKF) is used to decrease odometric error of the robot [20: Kang et al. 2010]. To overcome the problem relating to the outliers, random sample consensus (RANSAC) algorithm is used to filter outliers [19: Tong & Barfoot 2011].

To build the map, the occupancy grid map is used [21: Moravec & Elfes 1985], [39: Birk & Carpin 2006]. Since arbitrary data can be mapped, occupancy grid map is focused [55: Winner et al. 2012]. The occupancy grid map needs a resolution to discretize the environment [55: Winner et al. 2012]. Therefore, the occupancy grid map can be chosen depending on the requirements of the precision of the data [55: Winner et al. 2012]. The Bayesian probability grid map is used in [23: Thrun et al. 2005], [55: Winner et al. 2012]. The Bayesian probability grid map expresses the possibility of grid is occupied and there is a lot of merits of calculation [55: Winner et al. 2012].

In the dynamic environment, the occupancy grid map is difficultly built for the full of people. The inverse observation model is used to build static occupancy grid map [22: Wolf & Sukhatme 2004].

In this thesis, the PSO algorithm is proposed to overcome the local minimum

solution with ICP in [24: Chang & Lian 2012]. Figure 2.1 demonstrates simultaneous

localization and mapping categories.
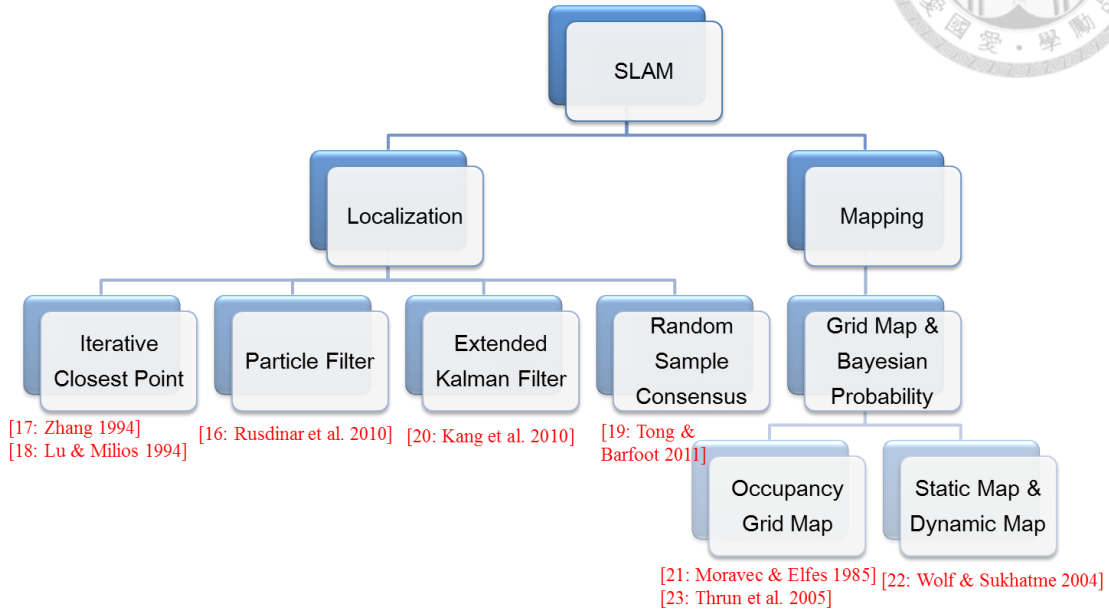
Figure 2.1 Simultaneous Localization and Mapping Categories

2.2 Pedestrian Detection and Tracking

Pedestrian detection and target pedestrian tracking play an important role in the

robotic field. Many sensors are implemented for pedestrian detection and target

pedestrian tracking. The sensors include laser-based sensor and vision-based sensor.

The laser-based sensor is used for distance measurement application [55: Winner et al.

2012]. And the vision-based sensor is used for color image application [55: Winner et

al. 2012].

For distance information such as LRF scan, many approaches are presented in

pedestrian detection and target pedestrian tracking. For pedestrian detection, inverse

observation model is used to differentiate between the dynamic objects and static

objects [22: Wolf & Sukhatme 2004]. However, the method detects the moving object.

In [26: Sung & Chung 2011] and [9: Chung et al. 2012], the clustering legs into a

pedestrian is presented. And using LRF scan in a two-layered arrangement to detect

features is presented in [27: Carballo et al. 2010]. For target pedestrian tracking, K

nearest neighbor (KNN) algorithm [3: Chen et al. 2011] and multiple hypothesis

tracking (MHT) algorithm [24: Chang & Lian 2012] are used with LRF scan in target

pedestrian tracking.

In the color image, background subtraction method acquires the moving objects

in static scene [29: Stauffer & Grimson 1999], [28: Lin & Huang 2011]. In [30: Lee et

al. 2003], background model updates based on Gaussian mixture model. In dynamic

scene, optical flow algorithm is applied [31: Enzweiler et al. 2008]. For pedestrian

detection, the image feature includes corner [34: Xu & Xu 2013], edge geometry [33:

Zhao et al. 2008], texture [32: Leithy et al. 2010], [41: Kun et al. 2012], and color

distribution [33: Zhao et al. 2008]. The image features are regarded as condition

judgments for pedestrian detection. What is more, Dalal and Triggs [35: Dalal &

Triggs 2005] present histograms of oriented gradients (HOG) feature vectors to detect

the pedestrian. Moreover, there are various approaches to track target pedestrian with

color image. For handle the occlusion case, Lin and Huang [28: Lin & Huang 2011]
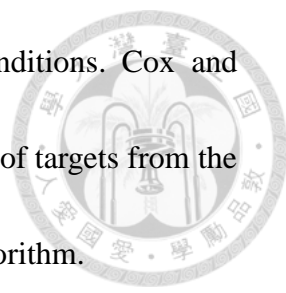
use either Kalman-filter or mean-shift algorithm in different conditions. Cox and

Hingorani [25: Cox & Hingorani 1996] enumerate multiple models of targets from the

latest three frames through multiple hypothesis tracking (MHT) algorithm.

For fusion of LRF scan and color image, a recognition method to track running

pedestrians is presented [12: Ueda et al. 2011]. In [37: Kristou et al. 2011], the target

pedestrian tracking uses the LRF scan. However, the pedestrian detection uses the

color image.

In this thesis, the pedestrian extraction uses LRF scan spatial continuity, image

color feature, and image edge feature. For tracking target pedestrian, color distribution

is regard as a conditional judgment. The mobile robot robustly detects pedestrian and

tracks target pedestrian. Figure 2.2 shows the pedestrian detection and target

pedestrian tracking categories.
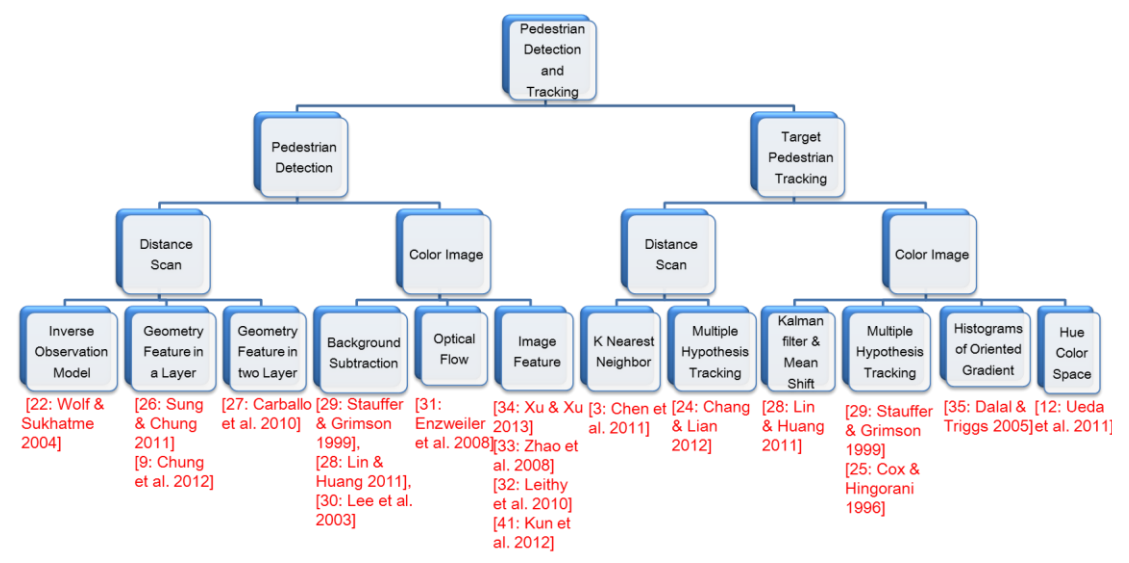
Figure 2.2 Pedestrian Detection and Target Pedestrian Tracking Categories

Chapter 3

Simultaneous Localization and Mapping

Self-localization and mapping are two of the fundamental capabilities for mobile robot [39: Birk & Carpin 2006]. The combination of the self-localization and the mapping is referred to as the simultaneous localization and mapping (SLAM) [39: Birk & Carpin 2006]. In unknown indoor environment, SLAM is a researchable task in this chapter. In this thesis, LRF is used owing to its accuracy in distance measurement [15: Wu et al. 2013]. The LRF operation principle and the LRF limitation are presented in Section 3.1. For self-localization, skidding and slipping can induce that the odometry is not equal to the real distance. To solve the problem, PSO algorithm [38: Li et al. 2011] is presented in Section 3.2. In the SLAM, once odemetry of mobile robot is known, building a map is also a task which can be effectively solved at the same time [39: Birk & Carpin 2006]. To build the map, the occupancy grid map is used in Section 3.3.

3.1 Laser Range Finder Usage and Limitation

LRF is a sensor measuring the distance. To use the LRF, the LRF operation principle and the LRF limitation should be known. In Section 3.1.1, the operation principle of LRF is presented. Section 3.1.2 states the limitation of LRF in specific scene.

3.1.1 Introduction of Laser Range Finder

For mobile robot, LRF is a common sensor. Compared with other sensors, LRF is a sensor commonly used owing to its accuracy in distance measurement. So the LRF is prevailing in the mobile robot.

The operation principle of LRF uses time of flight (ToF) to estimate the distance from specific angle [4: Okubo et al. 2009]. Figure 3.1 shows the operation principle of LRF. First, the laser emits an infrared beam and rotating mirror changes the beam's direction [4: Okubo et al. 2009]. Then the laser hits the surface of an object and is reflected [4: Okubo et al. 2009]. ToF is proportional to distance measurement. Since the infrared beam is rapid, the scan rate can achieve at least ten scans per second. From the direction of mirror, the phase of emitted can be estimated. Finally, the position of object is calculated.
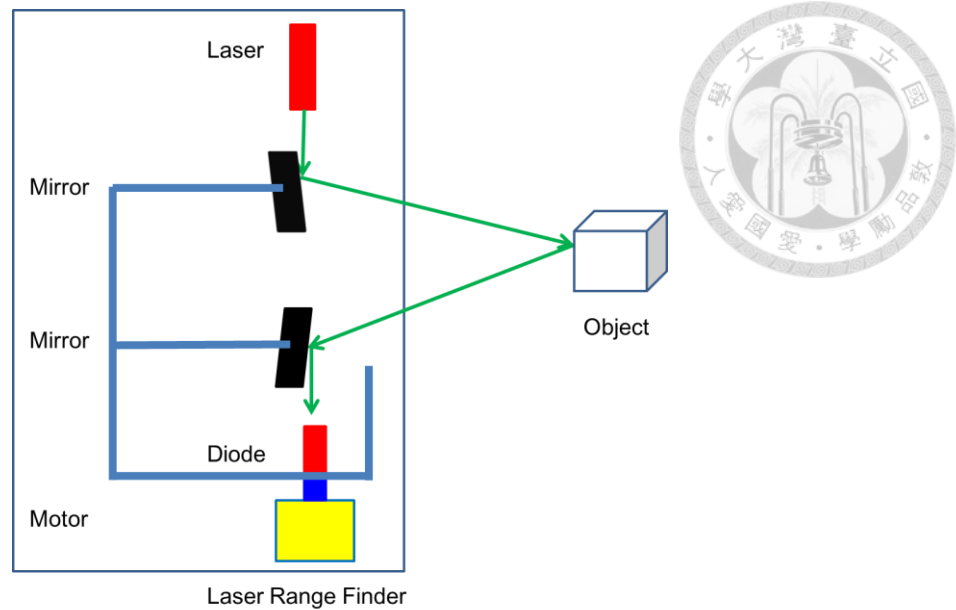
Figure 3.1 The Operation Principle of Laser Range Finder

3.1.2 The Limitation of Usage in the Glass Environment

LRF is a sensor commonly used owing to its accuracy in distance measurement [15: Wu et al. 2013]. However, the limitation of LRF is about environment texture. In the environment with glass, the light may refract the ray of light in the environment with glass like Figure 3.2(a). The incident light, laser beam, can be divided into diffusive reflection, specular reflection, and refraction, as shown in Figure 3.2(a). The diode only absorbs the diffusive reflection. Therefore, the missing data may occur. The real-scene is Ming-Da Building 2F having two glass window, as shown in Figure 3.2(b). In this scene, the data may miss.
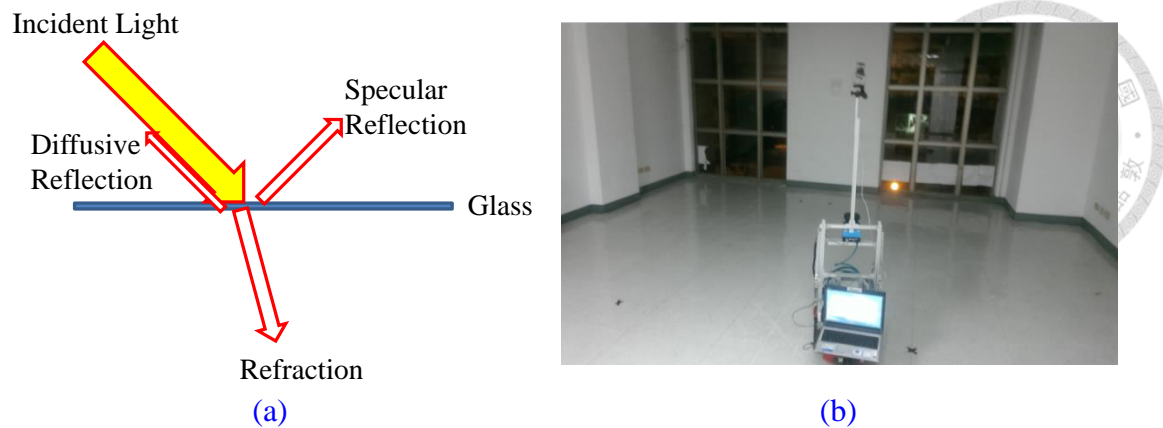
Figure 3.2: The Texture Influence in Glass Environment
(a) Theorem of the optical relation
(b) Scene in Ming-Da Building 2F

3.2 Robot Localization

Self-localization is an important task for mobile robot in unknown indoor environment. Although the mobile robot estimates the position with encoder, the scan matching technique can acquire precise position of the mobile robot. Generally, iterative closest point (ICP) algorithm [43: Besl & McKay 1992] is widely used to the scan matching. However, local optimal solution is a problem for ICP algorithm. The scan matching technique needs another algorithm.

PSO algorithm [44: Kennedy & Eberhart 1995] is a feasible method in scan matching technique. The PSO algorithm applies to minimize the distance energy function F(x) in the approximately global optimization problem [45: Eberhart & Shi 1998]. The particle swarm model sets the N particles in the D-dimensional problem space. The i-th particle owns the self-position x_i^t, self-velocity v_i^t, and distance

energy function $f_i^t(x)$ in the search domain at time t. Let each particle know the best

its position $pbest_i^t$ and the best position in N particles $gbest_i^t$ before time t.

Therefore, the position and the velocity of each particle with N particles are expressed

as follows:

$$x_i^t = (x_{i1}^t, x_{i2}^t, \dots, x_{iN}^t) \tag{3.1}$$

$$v_i^t = (v_{i1}^t, v_{i2}^t, \dots, v_{iN}^t) \tag{3.2}$$

The i-th particle is expressed as a point owning the position x_i^{t+1} and velocity

v_i^{t+1} at time t+1 according to the following equations:

$$v_i^{t+1} = \omega.v_i^t + c_1.rand().(pbest_i^t - x_i^t) + c_2.rand().(gbest_i^t - x_i^t) \tag{3.3}$$

$$x_i^{t+1} = x_i^t + v_i^t.\Delta t \tag{3.4}$$

where ω is an inertia weight, c_1 is a cognitive coefficient, c_2 is a social coefficient,

and rand() is a random probability in [0, 1]. x_i^{t+1} updates through Equation (3.4).

However, to avoid overshooting the global solution, v_i^t sets the threshold value v_{max}.

Figure 3.4 shows the illustration of particle motion and the process of particle motion.

In Figure 3.3(a), the plane represents the distance energy function. The curve

represents the pass path for each particle by time t. The circle represents current

position for each particle. The square presents the best position for each particle. And

the triangle represents best position for all particles. Each particle owns its position

and velocity. The velocity of each particle is determined by Equation (3.3) and the

position of each particle is determined by Equation (3.4). Every particle can affect
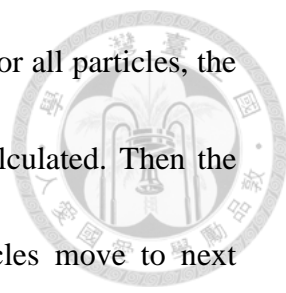
each other. Figure 3.4(b) shows the flow chart of particle motion. For all particles, the position and the velocity are set. The energy function $F(x)$ is calculated. Then the pbest and the gbest are picked in the particles. And the particles move to next position. Until the particles stay the same positions, the particles are global optimal positions. If the particles do not stay the same positions, the steps are iterative.

(a)

(b)
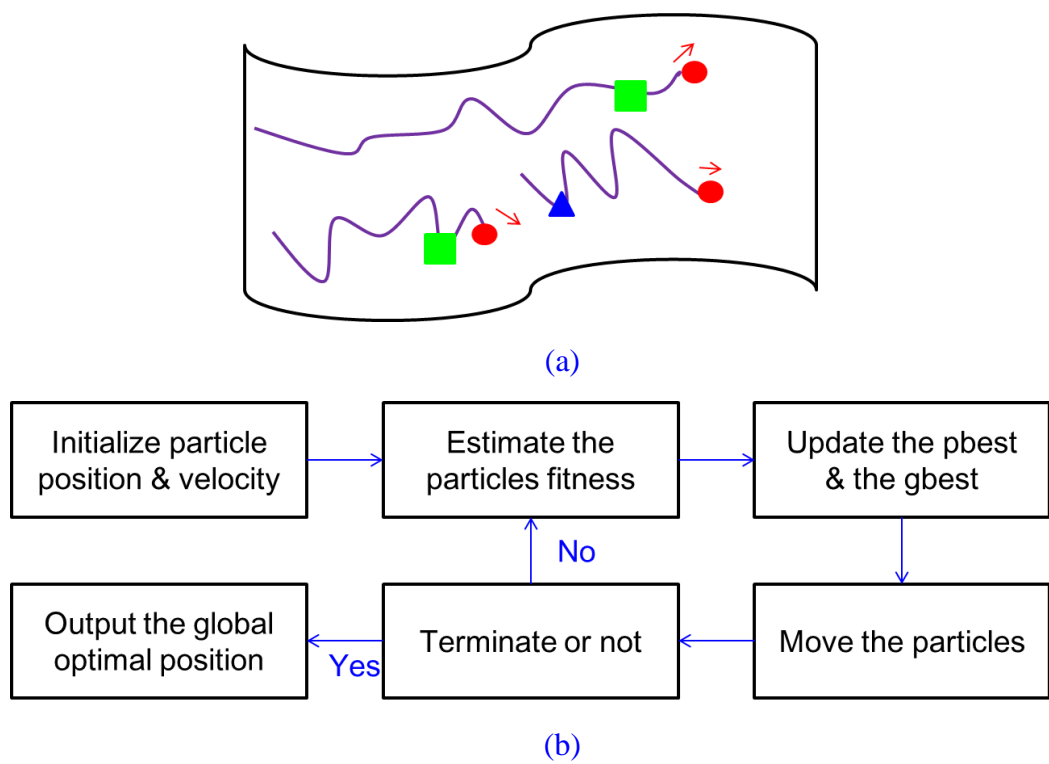
Figure 3.3 The Illustration and Process of Particle Selection
(a) The illustration of each particle moving
      i.   The line represents the pass path;
     ii.   The circle represents current position.
    iii.   The triangle represents best position in all particles.
    iv.   And the square is the best position in its particle.
(b) The particle selections of the PSO algorithm

Figure 3.4 shows scan matching points of two data sets results with the LRF. The scan matching points of two data sets include the red points and the blue points, as

shown in Figure 3.4. When the measurement error is not eliminated in time, it causes

the inaccurate map. Figure 3.4(a) and (c) show the LRF scan based on odometry in

Ming-Da 5F and Ming-Da 4F. Generally, PSO algorithm can reduce the measurement

error to enhance the map accuracy, as shown in Figure 3.4 (b) and (d). The other

experimental results in scan matching are shown in Section 5.3.

To acquire the more accurate robot position in populated environment, the

pedestrian need to be detected. Section 3.3 states dynamic map concept. And Section

4.3 states pedestrian detection based on dynamic map.
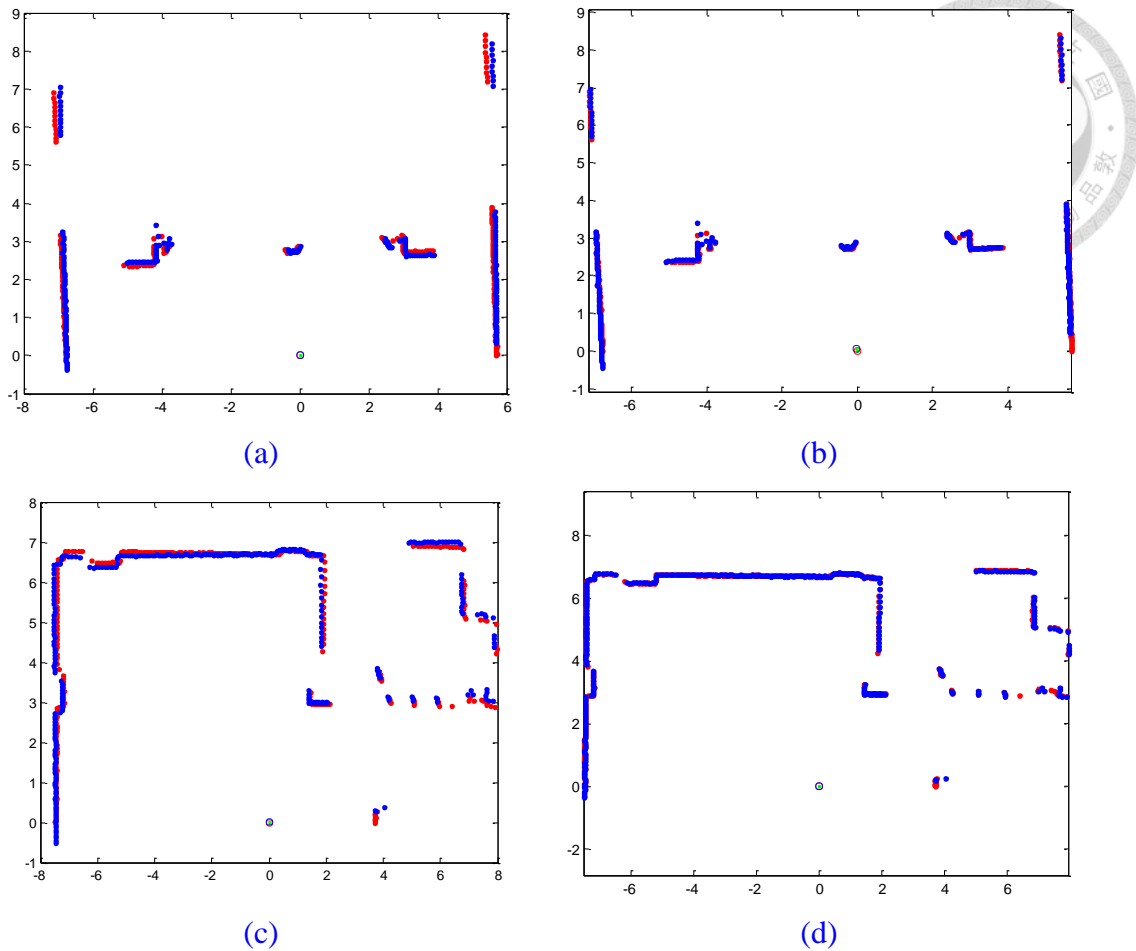
Figure 3.4 Scan Matching 2 Data Result (Target and Source are in the LRF scans)

(unit: meter)

(a) Encoder in Ming-Da Building 5F

(b) PSO algorithm process from (a)

(c) Encoder in Ming-Da Building 4F

(d) PSO algorithm process from (c)

3.3 Map Construction

In map construction, occupancy grid map is an important method. Section 3.3.1 uses

the occupancy grid map to construct the map. The inverse observation model is used

to build static occupancy grid map and dynamic occupancy grid map [22: Wolf &

Sukhatme 2004] in Section 3.3.2.

3.3.1 Grid Map Construction

To make mobile robot move arbitrarily, mapping is a necessary task. However, map construction is difficult for the mobile robot in dynamic environment. What is more, the inaccuracy measurement of LRF scan may cause the map false.

Occupancy grid map is a main method in map construction [23: Thrun et al. 2005]. The occupancy grid map needs a resolution to discretize the environment [55: Winner et al. 2012]. Therefore, the occupancy grid map can be chosen depending on the requirements of the precision of the data [55: Winner et al. 2012]. In occupancy grid map, three states include free, occupancy, and unknown. Bayesian probability update robustly applies to the occupancy grid map state as follows:

$$\text{p}(S^t|Z^t, S^{t-1}) = \alpha. \, \text{p}(S^t|Z^t). \, \text{p}(S^{t-1}|Z^{t-1}, S^{t-2}) \tag{3.5}$$

$$\text{Bel}(S^t) = \alpha. \, \text{p}(S^t|Z^{t-1}). \, \text{Bel}(S^{t-1}) \tag{3.6}$$

where the Bayesian probability $\text{p}(S^t|Z^t, S^{t-1})$ represents $\text{Bel}(S^t)$. S^t is map state in the specific position at time t, and Z^t is the measurement in the specific position at time t. α is a normalization coefficient.

By Equation (3.6), the occupancy grid map updates the state through iterative method. With this method, the grid occupancy probability only knows the previous occupancy grid map $\text{Bel}(S^{t-1})$ and the inverse observation probability $\text{p}(S^t|Z^{t-1})$.

3.3.2 Static Map and Dynamic Map

For the occupancy grid map, both the static map concept and the dynamic map concept are proposed [22: Wolf & Sukhatme 2004]. The dynamic map can be estimated from the following equation:

$$\frac{p(D^t|Z^1,...Z^t,S^{t-1})}{1-p(D^t|Z^1,...,Z^t,S^{t-1})} = \frac{p(D^t|Z^t,S^{t-1})}{1-p(D^t|Z^t,S^{t-1})} \cdot \frac{1-p(D)}{p(D)} \cdot \frac{p(D^{t-1})}{1-p(D^{t-1})} \quad (3.7)$$

where S^t is the state at time t, Z^t is the measurement at time t, and D^t is the dynamic state at time t.

However, the $p(D^t|Z^t, S^{t-1})$ needs the inverse observation model to update. Table 3.1 shows the occupancy probability p(S) in map construction with state. Three states include free, occupancy, and unknown. The threshold is set to 0.2 and 0.8.

Table 3.1 The Occupancy Probability of Each State S

State	Occupancy probability
Free	p(S) ≤ 0.2
Unknown	0.2 < p(S) < 0.8
Occupied	0.8 ≤ p(S)

With Table 3.1 result, the inverse observation model establishes in Table 3.2 [22: Wolf & Sukhatme 2004]. In [22: Wolf & Sukhatme 2004], Wolf and Sukhatme propose static map and dynamic map. The static map includes many dynamic

obstacles owing to inverse observation model. Since inverse observation model predicts that the state from unknown to occupied is static object, the pedestrian detected with LRF scan by robot may be regarded as static object. The probability is low in dynamic objects. It means that the objects are static. $p(D^t|Z^t, S^{t-1})$ is estimated as follows:

Table 3.2 Inverse Observation Model

S^{t-1}	Z^t	$p(D^t|Z^t, S^{t-1})$
Free	Free	Low
Unknown	Free	Low
Occupied	Free	Low
Free	Occupied	High
Unknown	**Occupied**	**Low**
Occupied	Occupied	Low

Figure 3.5 demonstrates an example of the inverse observation model analysis of two consecutive LRF scan. The state of pillar and the state of pedestrian are unknown for the robot at time t-1, while the pillar and the pedestrian are detected for robot at time t. Therefore, the pillar and the pedestrian are regarded as static objects from Table 3.2. In fact, the pillar should be regarded as a static object and the pedestrian should be regarded as a dynamic object. Table 3.2 is not obviously sufficient. In this scene, the moving object is only pedestrian. The pedestrian detection in Section 4.4 can solve the problem of inverse observation model in Table 3.2.
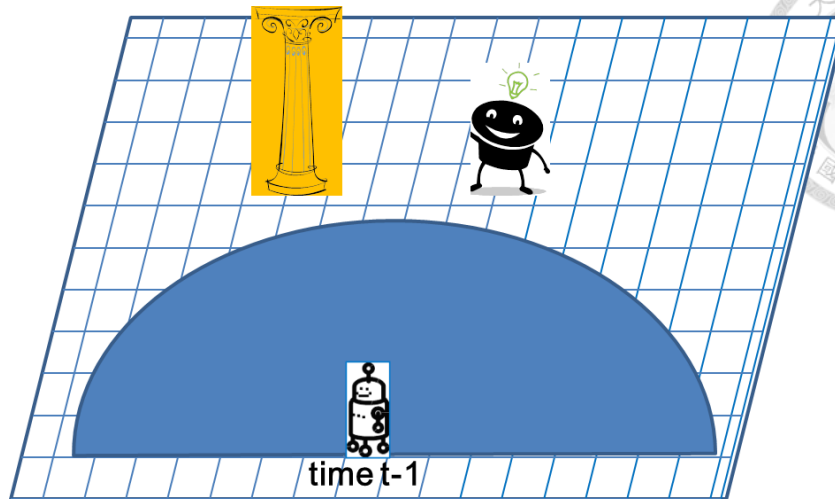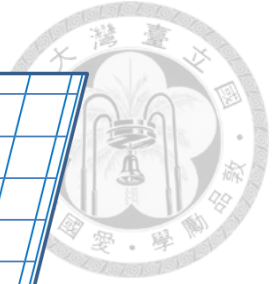
(a)
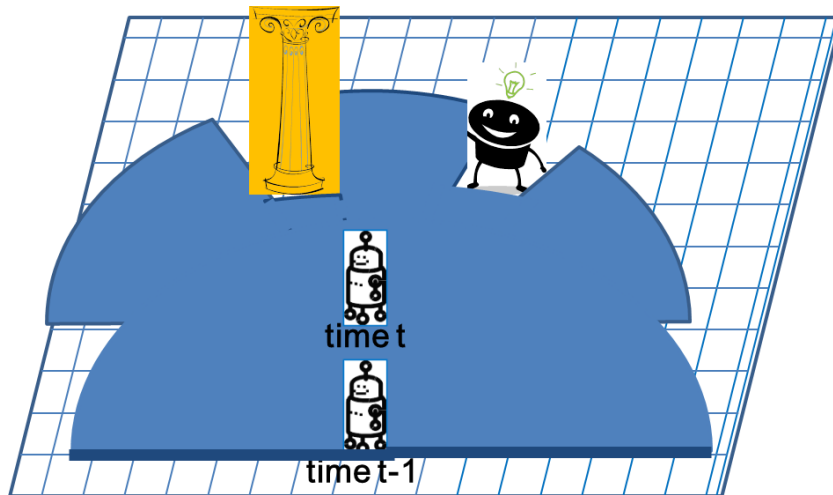
(b)

Figure 3.5 The Inverse Observation Model Problem (with LRF Scan Grid Map)

(a) The pillar and the pedestrian are unknown for the robot at time t-1.

(b) Either the pillar or the pedestrian is regarded as the static in inverse observation model in LRF scan at time t. The pedestrian detection judgment describes in Section 4.3.

Chapter 4

Pedestrian Detection and Target Pedestrian Tracking

For pedestrian detection and target pedestrian tracking, both LRF scan and color image are used in this chapter. Section 4.1 states the operation principle of omnidirectional camera and states the problem of the equipment. However, the combination of LRF and omnidirectional camera is difficult since the sensors are not calibrated. The calibration between LRF and omnidirectional camera can be divided into horizontal adjustment, translation, and rotation. In the calibration, the rotation calibration of the combination of LRF and omnidirectional camera is a researchable question. Section 4.2 states rotation calibration of the combination of LRF and omnidirectional camera. For pedestrian detection, the non-pedestrian needs to be filtered out. In this thesis, the methods with the LRF scan and the color image are presented in Section 4.3. LRF scan roughly judges pedestrian. Then the Hough circle transform and the color distribution are the judgment with the color image. With the above methods, the pedestrian detection can be implemented. In target pedestrian tracking, owing to pillars hindering or new pedestrians appearing, the data association

may be error between two consecutive LRF scans. The color distribution and the local

binary pattern (LBP) algorithm are used in the problems of Section 4.4.
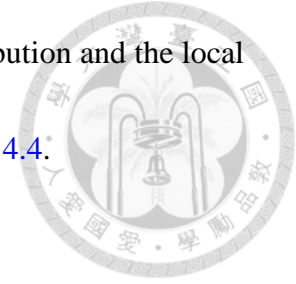
4.1 The Operation Principle of Omnidirectional Camera

In the pedestrian detection problem and the target pedestrian tracking problem of

Section 1.2, the LRF scan and the color image should be used. Owing to wide field of

view (FOV), the omnidirectional camera is necessarily used in the problems. Section

4.1.1 states the operation principle of omnidirectional camera. What is more, Section

4.1.2 presents the histogram equalization for low-light omnidirectional camera image.

4.1.1 Introduction of Omnidirectional Camera

The color image is widely used in the mobile robotic field. The color feature

plays an important role in pedestrian detection and target pedestrian tracking.

Therefore, the camera is often mounted on the mobile robot.

In this thesis, one of the problems is pillars hindering. The data association may

be error between two consecutive LRF scans. To search the target pedestrian, the

omnidirectional camera is necessarily used. The omnidirectional camera's field of

view is 360 degree. Therefore, the omnidirectional camera is regarded as an available

tool to track the target pedestrian in the pillars hindering problem.
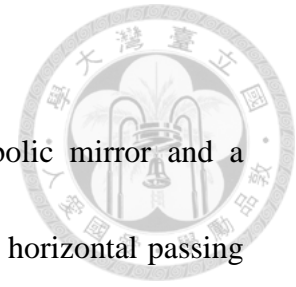
The structure of omnidirectional camera includes a hyperbolic mirror and a camera under the mirror like Figure 4.1 [5: Yagi et al. 2005]. The horizontal passing through the virtual center line (HPVCL) maintains the same height in projection [53: Yang & Lian 2012], [5: Yagi et al. 2005]. The operation principle makes a light flight to the upper center $(0, c)$. When the light touches the hyperbolic mirror, the light reflects to the other center $(0, -c)$. The image appears in the process.
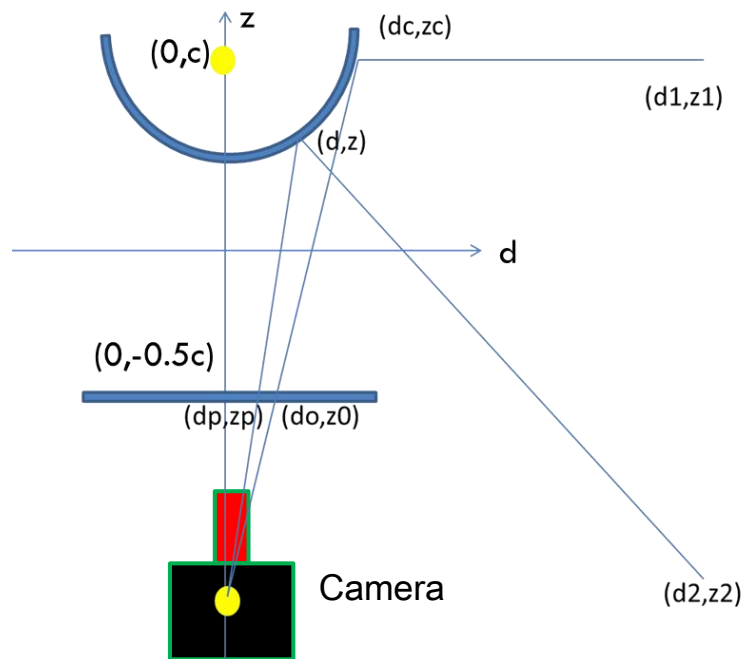
Figure 4.1 The Omnidirectional Camera Structure

4.1.2 The Lightness of Omnidirectional Camera

Although the omnidirectional camera owns many advantages, it still overcomes a low-light problem in Figure 4.2. Since the light does not directly flight to image plane,

the color image is dark. The low-light causes the image dark, as shown in Figure

4.3(a). However, the real scene from digital camera is bright, as shown in Figure

4.3(b). That the lighting sources in the beginning put on the floor seems unworkable

because of the unknown environment. The caution leads to two influences. One is the

edge threshold value sets small. As a result, the noise easily interfaces the results. The

other is each color channel distribution is dense. Therefore, using the color space to

tracking target pedestrian is more difficult. Two methods are present to improve the

influences. One uses the histogram equalization [6: Gonzalez & Woods 2008] stated

in the next paragraph. The other enhances light through the aperture. In this thesis, the

histogram equalization is used. With the process, the color image results obtain more

robust.

Figure 4.2 The Light Depends on the Active Lighting Source

| (a) | (b) |
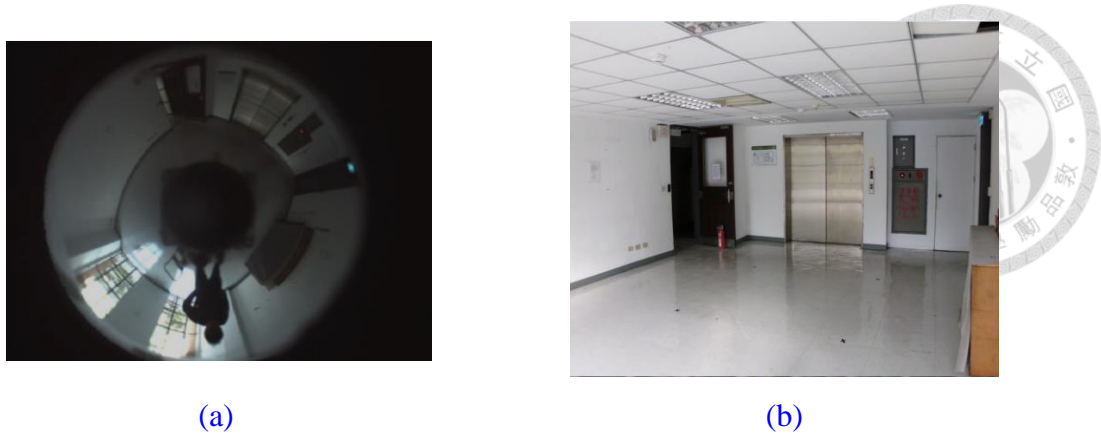
Figure 4.3 Display Image with Lightness Problem in Ming-Da Building 2F
(a) The omnidirectional image has the light problem
(b) The real-scene with digital camera

Histogram Equalization is a method making the intensity in image uniform [6: Gonzalez & Woods 2008]. The variables are shown in Figure 4.4. Let $P_r(r)$ be the probability of the intensity. Assume the output intensity s, and the definition of s is the following:

$$s = T(r) = \int_0^r P_r(w)dw \tag{4.1}$$

In this transform, the probability of s is the cumulative distribution function (CDF) of the input r. And that is proved in [6: Gonzalez & Woods 2008]. The definition of P_s is the following:

$$P_s = \begin{cases} 1 & 0 \le s \le 1 \\ 0 & otherwise \end{cases} \tag{4.2}$$

where the probability of P_s is a uniform function. Owing to digital signal, the intensity s_k of image from Equation (4.1) is a discontinuous function in the image process as follows:

$$s_k = \sum_{j=0}^{k} p_r(r_j) \quad\quad (4.3)$$

The distribution of intensity is sparse, as shown in Figure 4.3(c), while it is uniform. The edge extraction is more consistent than unprocessed image in dark image. The detail states in Section 5.2.
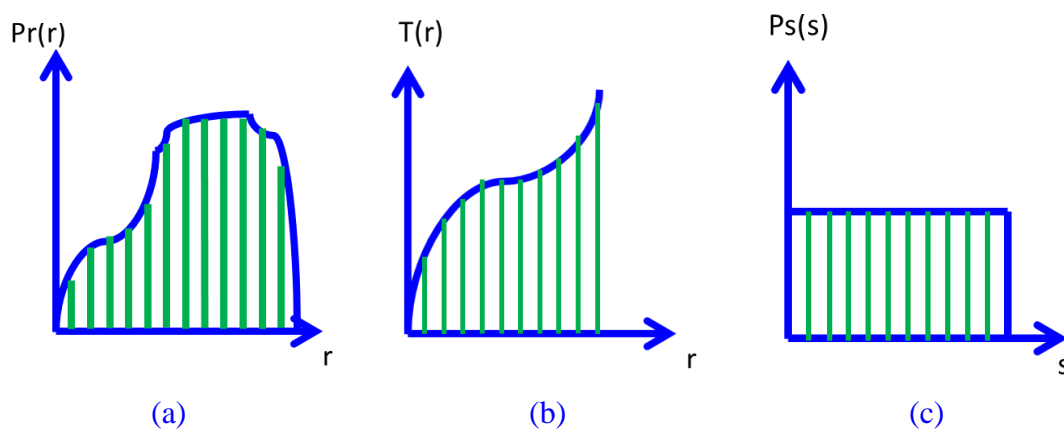
Figure 4.4 Each Variable in Histogram Equalization
(a) The probability of original intensity: $P_r(r)$
(b) The original intensity transforms output intensity: $T(r)$
(c) The probability of the output intensity: $P_s(s)$

4.2 Sensors Calibration

Before using the sensor, the calibration is an important task. In Section 4.2.1, the calibration is divided into horizontal adjustment, translation, and rotation. And the solutions are presented. Section 4.2.2, Section 4.2.3, and Section 4.2.4 are a series of the solutions for the rotation calibration.

4.2.1 The Description of Calibration

Combining the LRF and the omnidirectional camera can acquire the abundant information in signal process. Most of all, the data association is the most important problem in sensors calibration. In this calibration, the adjustment of six freedoms is divided into horizontal adjustment, translation, and rotation. The horizontal adjustment uses the gradienter to calibrate the inclination. The translation is to align the geometry center of the LRF and geometry center of the omnidirectional camera in different horizontal plane. The rotation problem is discussed in Section 4.2.2, 4.2.3, and 4.2.4. Figure 4.5 shows the calibration problem for horizontal adjustment, translation, and rotation. The gradienter is used to calibrate the horizontal adjustment. Furthermore, the vernier caliper is used to align the geometry center. For the rotation problem, the angle matching is a method. In this thesis, the break point and the angular point in LRF scan and vertical line in color image are regarded as feature and presented in Section 4.2.2 and Section 4.2.3.

(a)

(b)

(c)

(d)
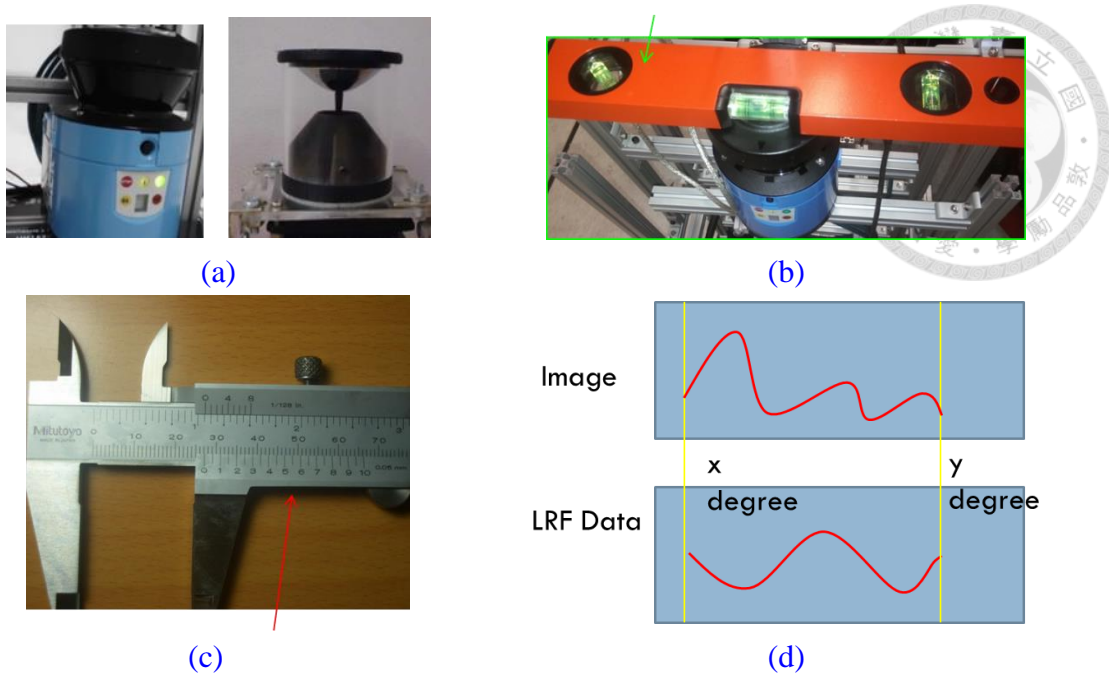
Figure 4.5 Calibration Problems: Horizontal Adjustment, Translation, and Rotation

(a) The laser range finder and omnidirectional camera

(b) The gradienter for the horizontal adjustment in different plane

(c) The vernier for the translation adjustment

(d) The plane rotation problem sketch

4.2.2 Break Point and Angular Point Detection

To do data association, using the feature of data is necessary. The indoor environment is full of the walls. The break point for LRF scan and the angular point for LRF scan are shown in Figure 4.6 [7: Jia et al. 2010]. Here, the break point is presented based on point-distance-based segmentation method [8: Rebai et al. 2009]. The distance between two continuous points in LRF scan is expressed as follows:

$$D(r_i, r_{i+1}) = \sqrt{r_{i+1}^2 + r_i^2 - 2.r_{i+1}^2.r_i^2.cos\Delta\alpha} \tag{4.4}$$

If the distance is more than the threshold value D_{th}, the two points are the break

points. Therefore the threshold value sets as follows:

$$D_{th} = C_0 + \frac{C_1 . min\{r_i, r_{i+1}\}}{cotg(\beta) . \left(\cos\left(\frac{\Delta\alpha}{2}\right) - \sin(\frac{\Delta\alpha}{2})\right)}$$ (4.5)

The parameters C_1, C_0, and β are presented in [8: Rebai et al. 2009]. Next, the

angular point is introduced in [7: Jia et al. 2010]. The start point links end point to be

a line. If the distance in point to the line is more than the threshold value δ, the

angular point appears. In Figure 4.6, the idea of feature detection in LRF scan is

presented. The 'V' presents the break point and the 'X' presents the angular point. For

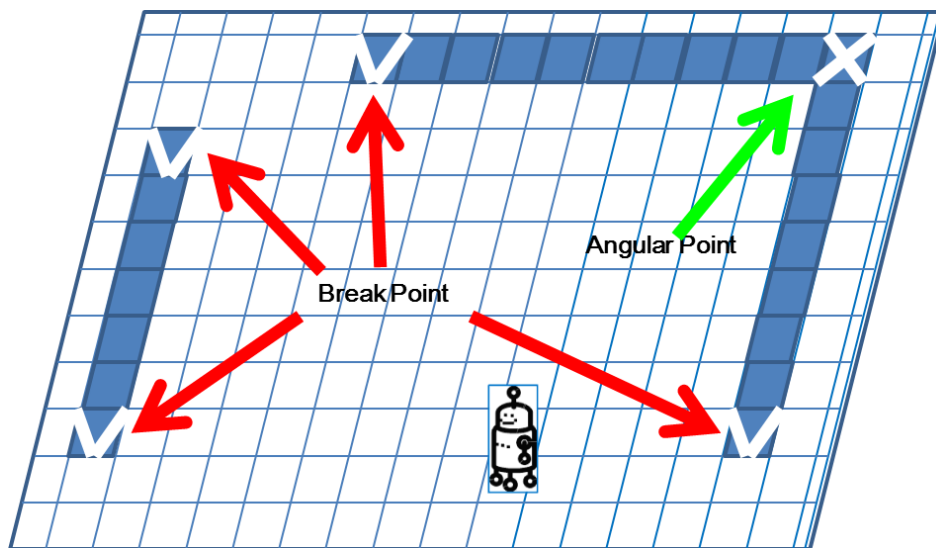the indoor scene, both the break point and the angular point may be corner.

Figure 4.6 The Break Points and Angular Points (With LRF Scan Grid Map)
In LRF scan, two consecutive points determinate the corner. Finding the break
point, the angular point owns max distance large a threshold value between two
break points link.

4.2.3 Vertical Line Detection

The vertical line for color image is also an important feature in unknown indoor environment. The image of omnidirectional camera has distortion. Because of both angular matching and image distortion, the panorama is necessary. To expand the panorama, the projection center search is first of all. To sum up, the process needs to find the projection center, expand the panorama, and detect the vertical line, as shown in Figure 4.13.

The flow chart of projection center search is shown in Figure 4.7. The projection center searching through the HPVCL is set in image center. First, the lower image is cut. The gray broad in image appears with RGB filter. Then, the region growing and the image filling are used to the image. The data x, y, and R relationship is shown as follows:

$$(x - c_1)^2 + (y - c_2)^2 = R^2 \tag{4.6}$$

where the coefficients c_1, and c_2 are unknown. Next, through least square method [10: Gander et al. 1994], the optimal projection center can be obtained. Finally, the iterative method is continuous until it converges.

With above method, the optimal projection center converges in the omnidirectional camera image.
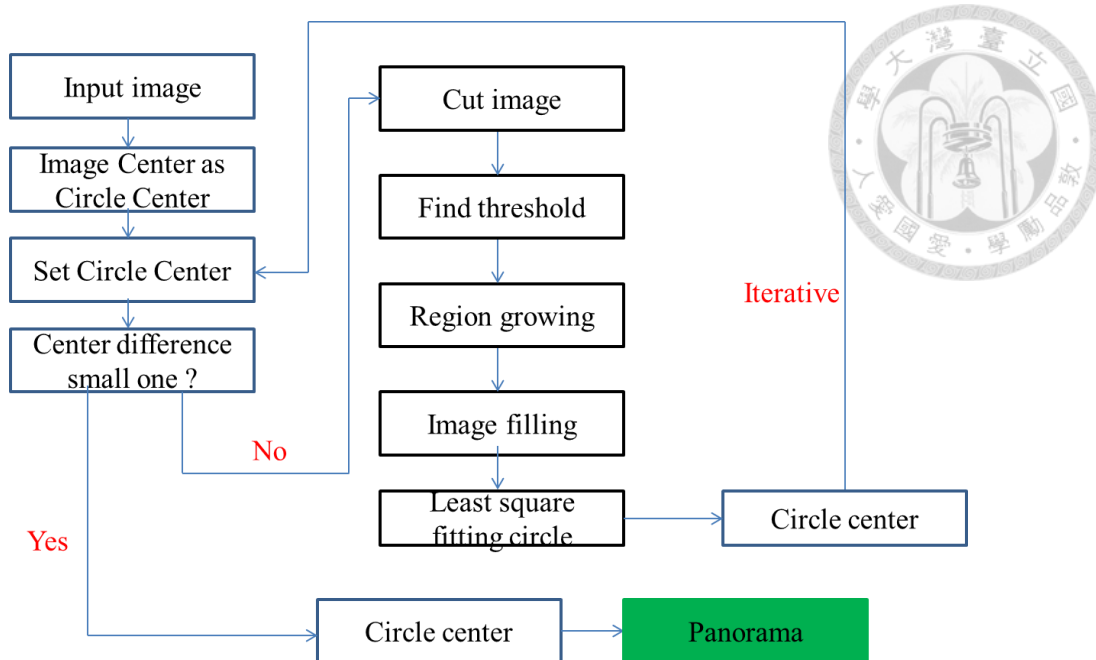
Figure 4.7 The Flow Chart of Projection Center Search

(a)

(c)

(b)

(d)

(e)
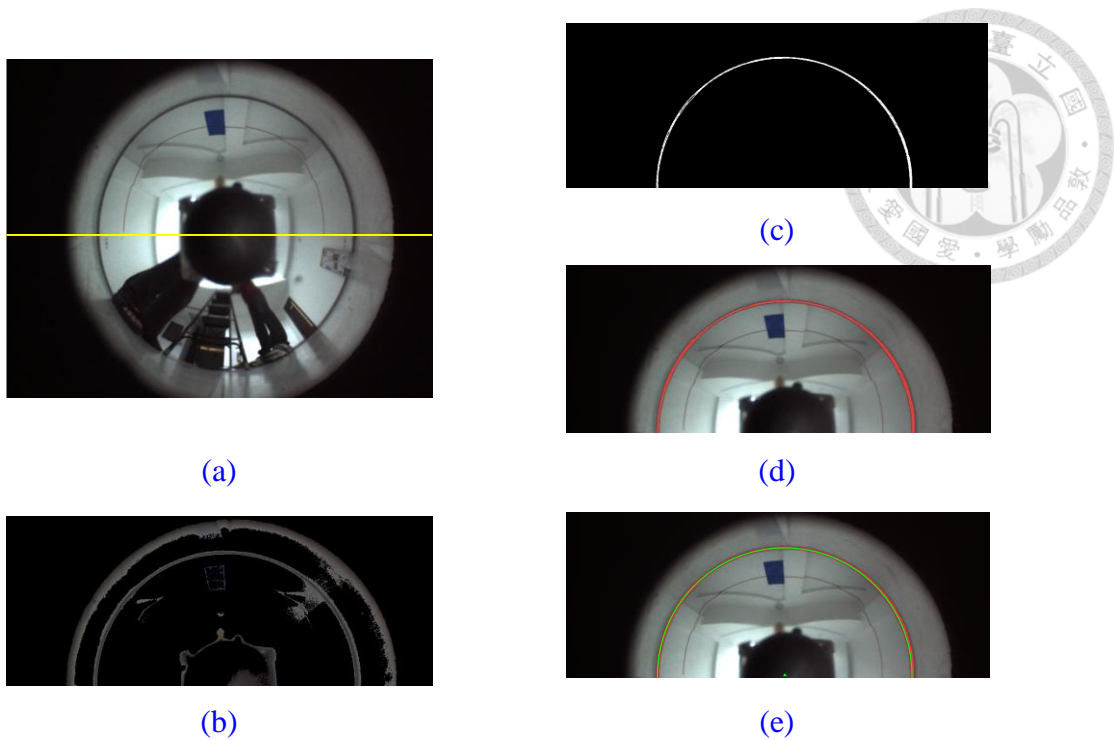
Figure 4.8 Results of the Projection Center Search Algorithm

(a) Cut image

(b) Set the threshold value

(c) Region growing

(d) Image fill the holes

(e) Least square fitting circle

The panorama remedies the distortion image. For [11: Grassi & Okamoto 2006], that the panorama image depends on radius and angle from projection center is shown in Figure 4.9. The vertical axis of panorama is radius and the horizontal axis of panorama is angle. The larger the radius from projection center is, and the less the distortion in omnidirectional image is. The idea of panorama is that the Cartesian coordinate converts the polar coordinate. Figure 4.10 demonstrates an example for panorama image in real scene. Panorama image is shown in Figure 4.10(b).

Figure 4.9 The Idea Process of Panorama Image

The vertical axis is radius and the horizontal axis is angle. The larger the radius is, and the less the distortion is.

(a)

(b)

Figure 4.10 Real Scene Image Process in Ming-Da Building 5F
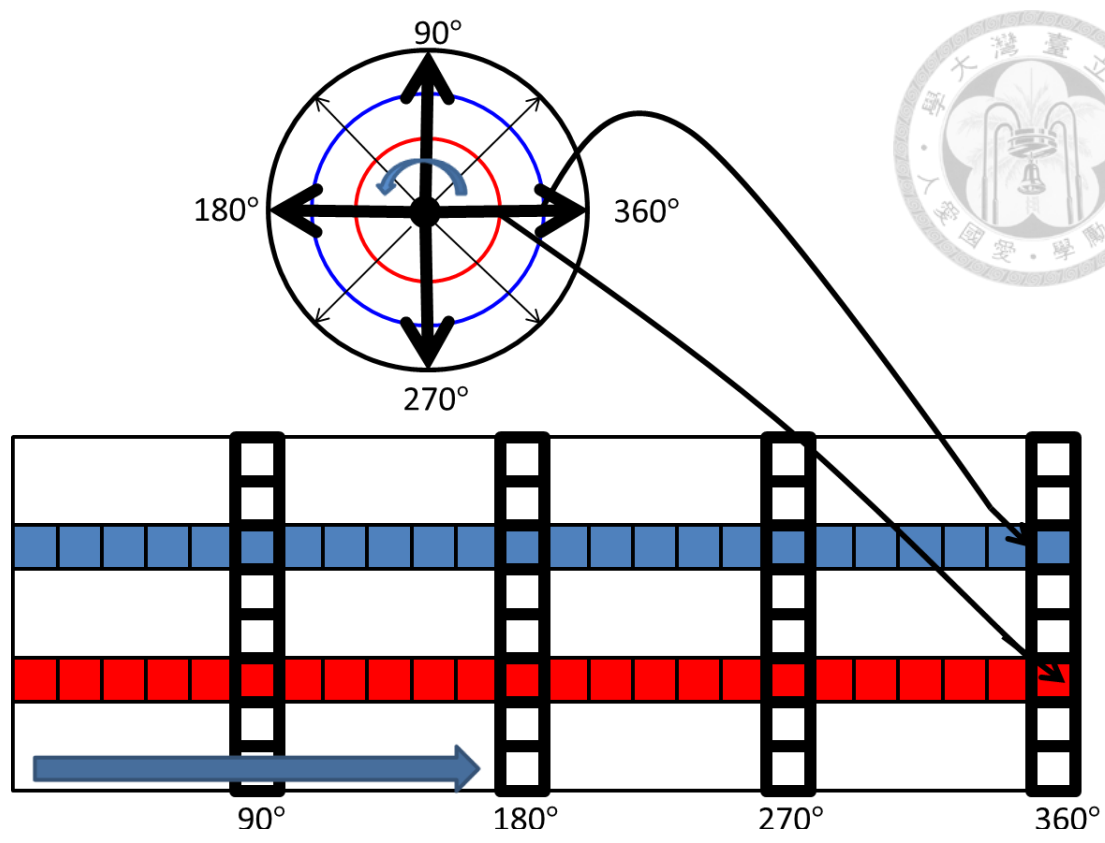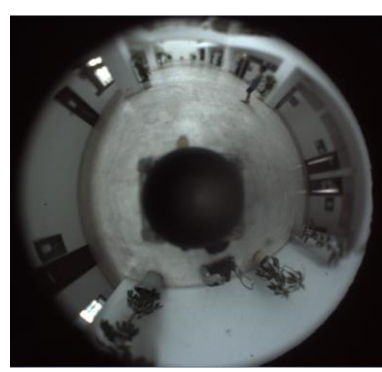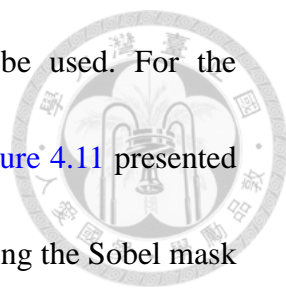
(a) Omnidirectional camera image

(b) Panorama image in (a) from the idea of Figure 4.9

For the vertical line detection, many edge detectors can be used. For the panorama image, the Sobel vertical edge detection as shown in Figure 4.11 presented in [6: Gonzalez & Woods 2008] seems to be a practical method. Using the Sobel mask in Figure 4.11 does convolution with the original image. If the intensity difference in vertical direction is more than threshold value, the pixel is considered as a vertical edge. The main problem is the lightness of color image stated in Section 4.1.2. If the gradient ∇f is more than threshold T, the vertical edge is detected. However, the threshold value in Sobel vertical edge detection can vary dramatically because of the low-light environment for omnidirectional camera. Figure 4.12 demonstrates an example of low-light environment for omnidirectional camera. The threshold is small, as shown in Figure 4.12(a), the edge can be detected. However, the threshold is large, as shown in Figure 4.12(b). The low-light problem causes the edge threshold often needs to change. Figure 4.14 presents the vertical line extraction results with histogram equalization of low-light environment for omnidirectional camera. Histogram equalization is used to Figure 4.14(c). The threshold values in Sobel vertical edge detection is set the same. As expected, Figure 4.12(d) appears more vertical line than Figure 4.12(b). Therefore, the vertical line detection includes inputting an image, finding projection center, expanding the panorama, using histogram equalization, using Sobel vertical mask, and using area filtering, as shown

in Figure 4.13.

-1	0	1
-2	0	2
-1	0	1

Figure 4.11 The Sobel Vertical Edge Detector

137	I41	115
137	I51	115
137	I61	115

200	I42	55
200	I52	55
200	I62	55

(a)   (b)

Figure 4.12 The Example of Two Image for Sobel Vertical Edge Detector

(a) The intensity of left column and right column is nearly

(b) The intensity of left column and right column is sparse

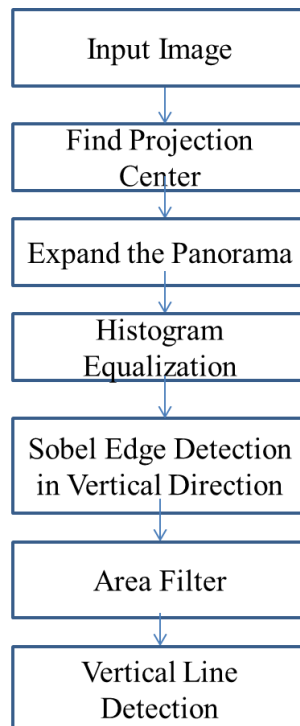Input Image
↓
Find Projection Center
↓
Expand the Panorama
↓
Histogram Equalization
↓
Sobel Edge Detection in Vertical Direction
↓
Area Filter
↓
Vertical Line Detection

Figure 4.13 The Process of Vertical Line Detection

(a)

(e)

(b)

(f)

(c)

(g)

(d)

(h)
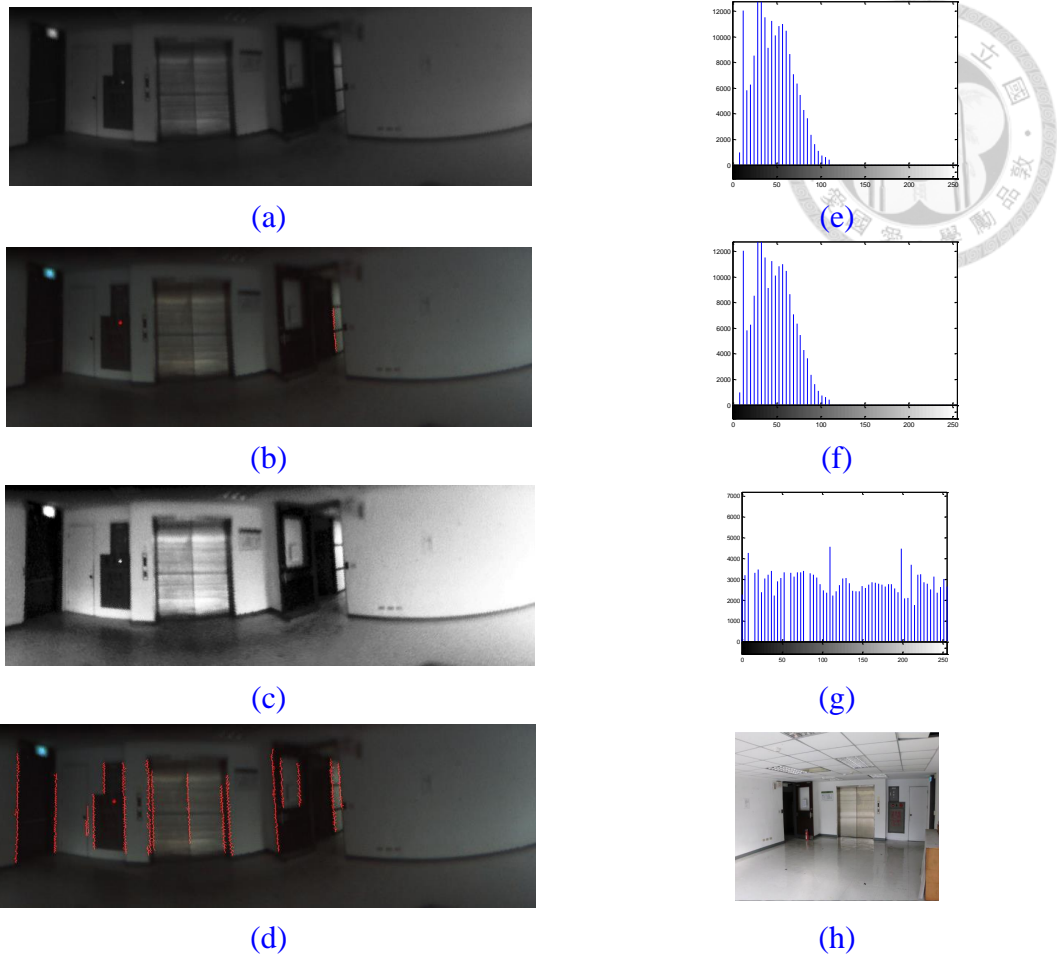
Figure 4.14 Vertical Edge Detector in the Same Threshold Value

(a) The original image converts to gray

(b) The edge detector result of (a)

(c) Using histogram equalization process

(d) The edge detector result of (c) and the real scene in digital camera

(e) The histogram from (a)

(f) The histogram from (b)

(g) The histogram from (c)

(h) The real-scene with digital camera in Ming-Da Building 2F

4.2.4 Data Association

In the sensors calibration, the data association is an important task. However, the

distance in LRF scan matches the pixel in color image is a problem. In [13: Bacca et

al. 2013], the corner point matches the vertical line like in Figure 4.15(a). What is more, the dashed plane means the LRF scan plane. In [12: Ueda et al. 2011] and [14: Scaramuzza et al. 2006], the corner of vertical line is estimated in omnidirectional image. With the above transform function, the rotation angle can be estimated at the polar coordinates. From vertical line, matching the corner point needs a polynomial function with order 4. In unknown indoor environment, the corner of LRF scan in specific plane and vertical line of color image in digital camera are obvious features. The results show in Section 5.2.2.

(a) (b)

Figure 4.15 The Data Association Conception Matching the Feature Points
(a) The dashed plane means the LRF scan plane. From vertical line, matching the corner point needs a polynomial function with order 4. The height convert pixel shows in Section 5.2.
(b) A door of Real-Scene for Vertical Line and Corner. In door, the corner of LRF scan in specific plane and vertical line of color image in digital camera are obvious features.
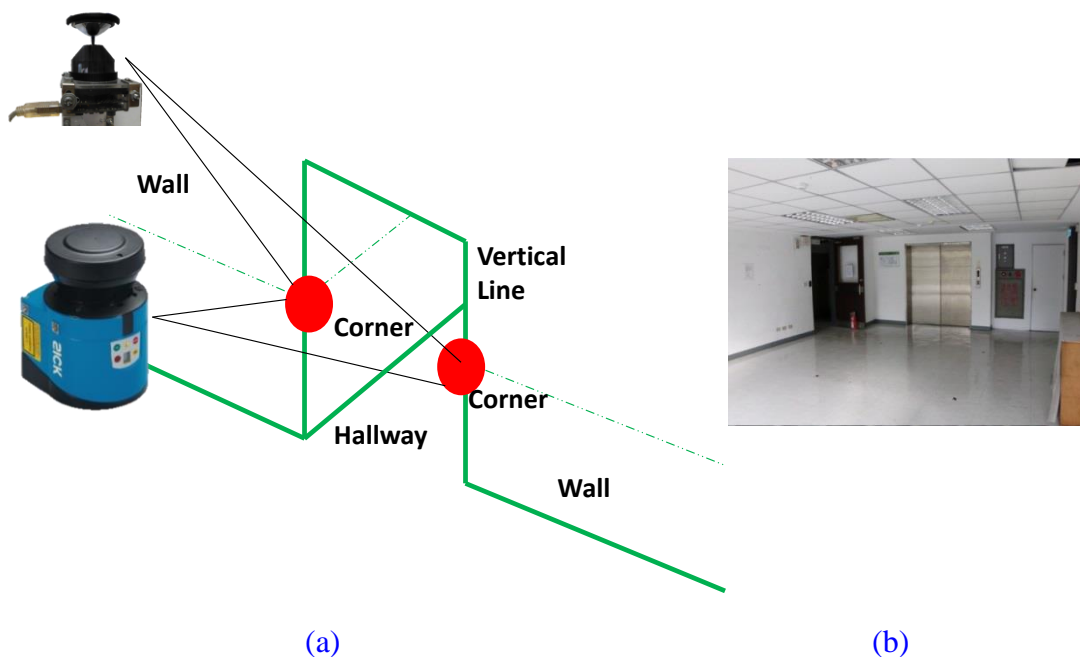
4.3 Pedestrian Detection

In pedestrian detection, both omnidirectional camera and LRF are used in the thesis. Section 4.3.1 presents a preprocessing for pedestrian detection with LRF scan. Both the left line and the right line of bounding box are from LRF scan. Section 4.3.2 states the lower line of bounding box extraction in pedestrian detection. However, the upper line of bounding box extraction for pedestrian needs to judge the pedestrian since the height of pedestrian is unknown. Section 4.3.3 provides judgment for pedestrian detection. Color distribution and Hough circle transform are used to estimate the head of pedestrian. Section 4.3.4 states the idea that the pedestrian extraction corrects the global static map.

4.3.1 Pedestrian Detection Preprocessing with Laser Range Finder

For pedestrian detection, inverse observation model is used in [3: Chen et al. 2011]. However, the inverse observation model with LRF scan does not detect the pedestrian whose state is from unknown to occupied. In summary, the moving object in LRF scan is sometimes regarded as static object in environment. Take for example, the dynamic map in Cartesian coordinate shows in Figure 4.16(a). However, the pedestrian in dynamic map is not obviously. Therefore, that the preprocessing for
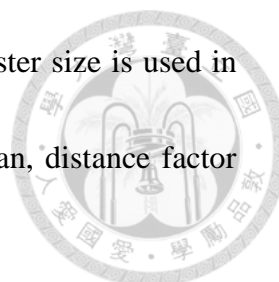
Figure 4.17(a) includes data segmentation, distance factor, and cluster size is used in the thesis. The following states the data segmentation for LRF scan, distance factor for LRF scan, and cluster size for LRF scan.

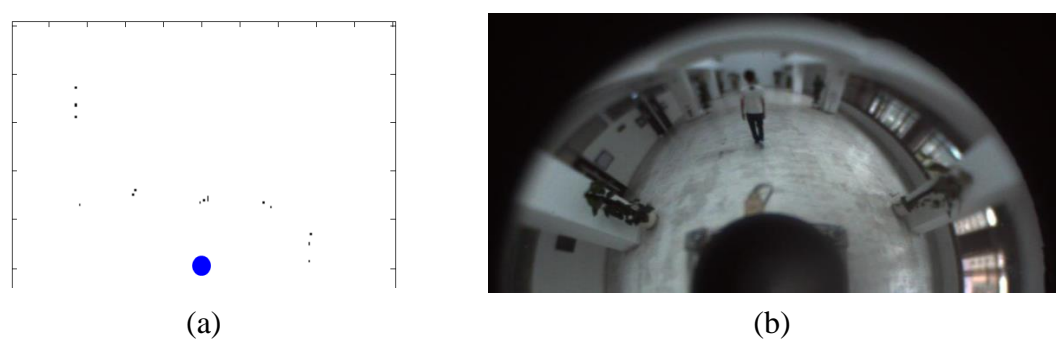(a)   (b)

Figure 4.16 The Moving Objects in Dynamic Map with LRF scan (a grid: 200cm)
(a) By inverse observation model, some moving objects points are noise points in Cartesian coordinate.
(b) The color image in omnidirectional camera with semi-image

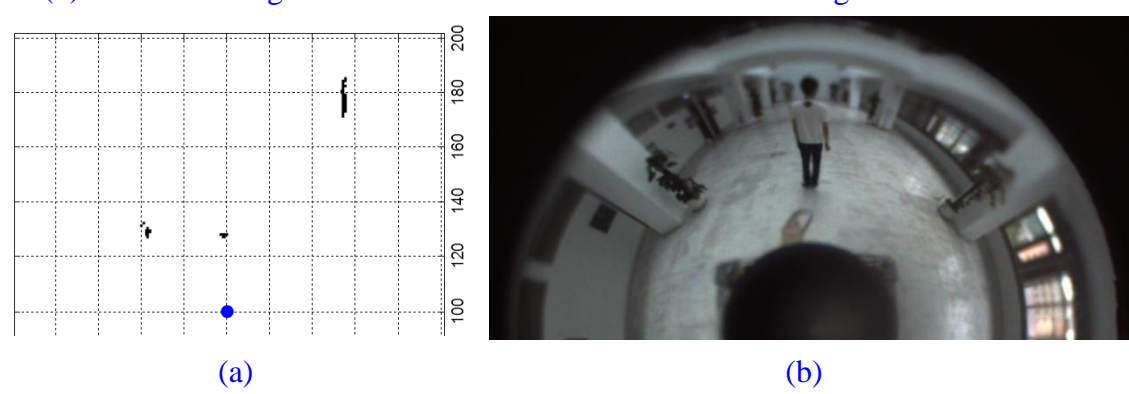(a)   (b)

Figure 4.17 The Measurement Points in LRF scan (unit : 10cm)
(a) The points in measurement map are error in Cartesian coordinate. The unreasonable pedestrian points should be eliminated. The LRF scan shows the LRF points needed to be solved.
(b) The color image in omnidirectional camera with semi-image

The idea of data segmentation is same as Section 3.2.2. For each cluster, if two consecutive points distance $D(r_i, r_{i+1})$ with Equation (3.4) in Section 3.2.2 is larger than D_{th} with Equation (3.5) in Section 3.2.2. Then the cluster divides into two
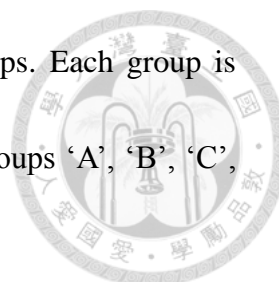
clusters. With the method, the points cluster around several groups. Each group is regarded as an object. The measurement point divides into four groups 'A', 'B', 'C', and 'D', as shown in Figure 4.18(a).

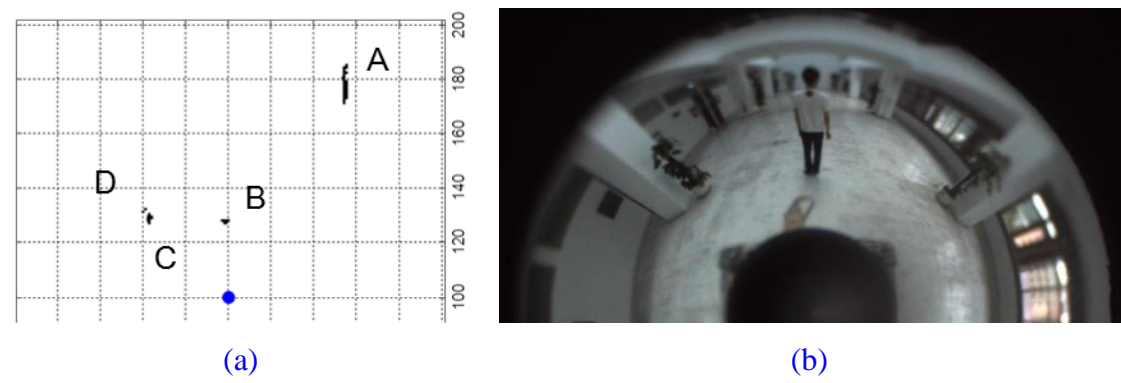(a)   (b)

Figure 4.18 The Data Segmentation Result in LRF scan (unit : 10cm)
(a) The data segmentation divides into four groups as the picture in Cartesian Coordinate. In picture, the scan divides into four groups A, B, C, and D.
(b) The color image in omnidirectional camera with semi-image

In general, the distance from the mobile robot affects the accuracy. Therefore, the threshold distance for measurement points is set. The distance from the mobile robot is far. Then, the confidence is low. In Figure 4.19(a), the distance of wall achieves eight meter. The far distance for measurement points is eliminated.
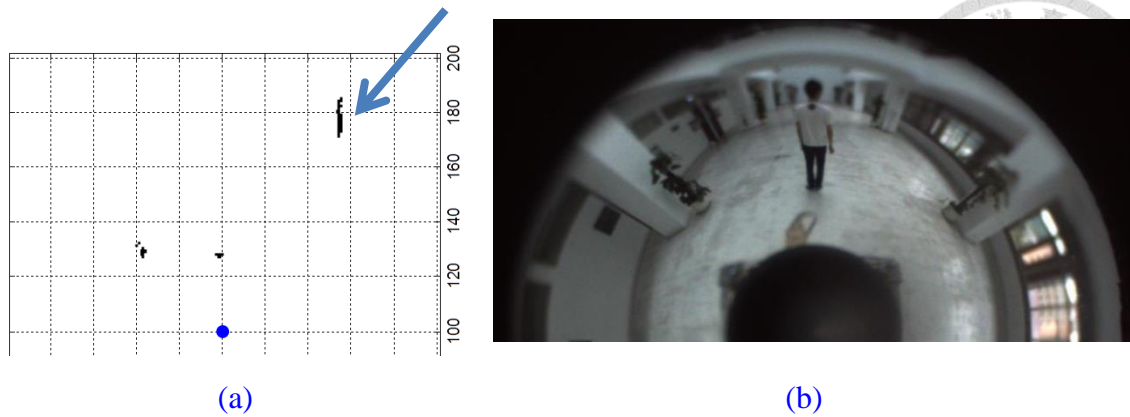
<center>(a)   (b)</center>

Figure 4.19 The Distance Factor in LRF Scan (unit : 10cm)

(a) The distance is far, and then the confidence is low. In this picture, the distance achieves eight meter. The error increases with distance. The LRF scan is in Cartesian Coordinate.

(b) The color image in omnidirectional camera with semi-image

Finally, the cluster size should be considered in the pedestrian detection. For the pedestrian data, the points in each cluster contain specific number. For example, the wall in Figure 4.20 is detected as a moving object. However, the number of measurement points obviously is non-pedestrian. Using the three judgments, the pedestrian is roughly extracted.
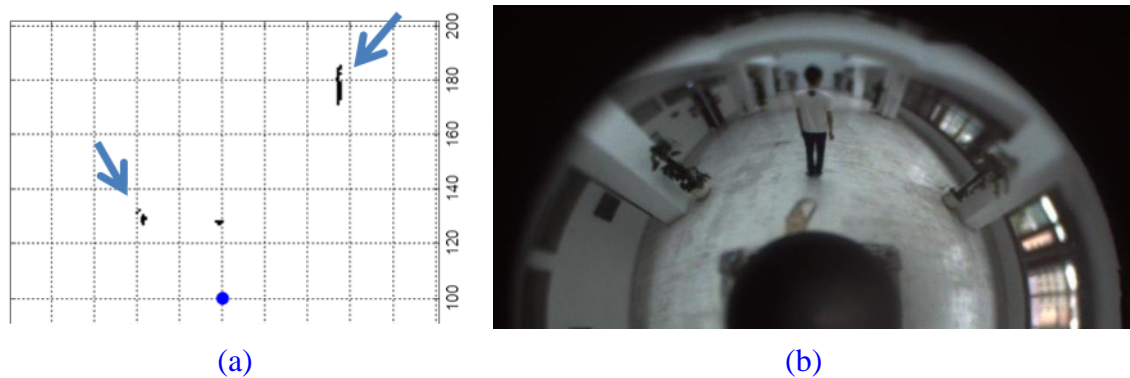

<center>(a)   (b)</center>

Figure 4.20 The Cluster Size of Elimination in LRF Scan (unit : 10cm)

(a) The pedestrian size owns specific size, and the unreasonable points are eliminated. The LRF scan is in Cartesian Coordinate.

(b) The color image in omnidirectional camera with semi-image.

4.3.2 Lower Line of Bounding Box Extraction

For the pedestrian detection, the LRF scan only the distance information seems to be insufficient. However, the pedestrian information includes not only the distance from mobile robot but also the color of pedestrian. Hence, the color image is regarded as an additional condition for the judgment based on the laser range finder (LRF) scan.

The omnidirectional camera is used to detect pedestrian since the field of view (FOV) is wide. For fusion of LRF and omnidirectional camera, the bounding box needs to be extracted. The left line of bounding box and the right line of bounding box are from the LRF scan. Therefore, the upper line of bounding box and the down line of bounding box states in Section 4.3.2 and Section 4.3.3.

Through the sensor calibration in Section 4.2, the sensors are directly used. To acquire lower line of bounding box, the operation principle of omnidirectional camera is used. Therefore, the omnidirectional camera mirror is discussed. Figure 4.21 shows the undistrtion image by a toolbox in [46: Bouguet 2013]. Hence, the omnidirectional camera mirror model can be estimated.

Figure 4.21 The Omnidirectional Camera Model
(a) The original image with digital camera
(b) The undistortion image with digital camera

From [5: Yagi et al. 2005], the omnidirectional mirror is a hyperbolic surface. To acquire the mirror equation, image processing is necessary. The hyperbolic surface region is selected from Figure 4.22(a), as shown in Figure 4.22(b). Canny edge detector [6: Gonzalez & Woods 2008] is used. Generally, the Canny edge detector is a powerful tool to deal with edge problems [6: Gonzalez & Woods 2008]. The result by the Canny edge detector is shown in Figure 4.22(c).

(a) (b) (c)

Figure 4.22 The Hyperbolic Surface Extraction
(a) The original image with digital camera
(b) Select the hyperbolic surface region in Figure 4.22(a)
(c) Using the Canny edge detector in Figure 4.22(b)

After acquiring the edge information, the hyperbolic surface equation is the next

problem. Least-square fitting (LSF) of Hyperbolae [47: O'Leary & Murray 2004] is minimum error method for curve fitting. In [47: O'Leary & Murray 2004], the process uses the single value decompositio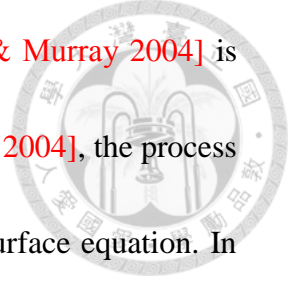n (SVD) to fit the hyperbolic surface equation. In Figure 4.23(b), the curve is the Least-Square Fitting curve from Figure 4.23(a).
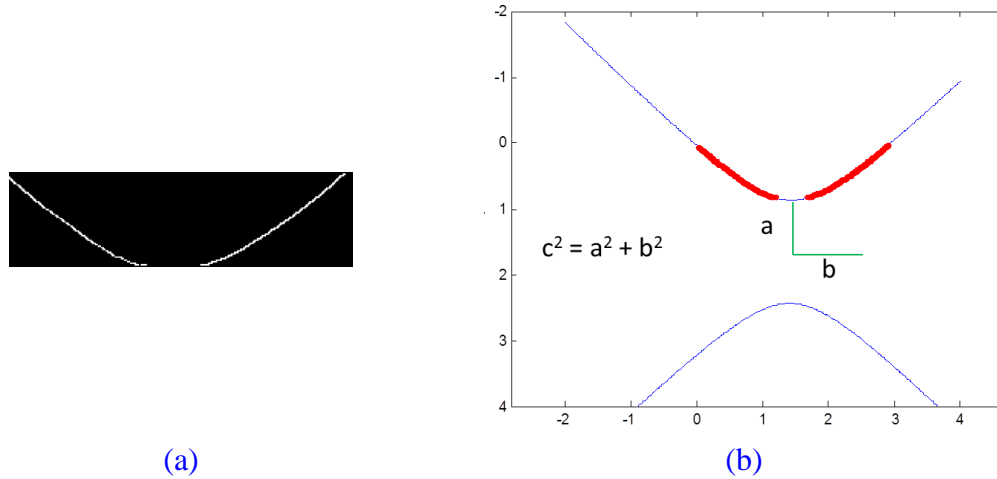

(a)   (b)

Figure 4.23 The Curve Fitting of SVD algorithm
(a) Using the Canny edge detector in Figure 4.22 (c)
(b) The curve is the Least-Square Fitting curve from Figure 4.23(a).

The above paragraph describes the curve fitting with SVD algorithm. Figure 4.24 shows the model of omnidirectional camera in [5: Yagi et al. 2005]. The parameters (d_1, z_1), (d_c, z_c), and (d_0, z_0) are HPVCL parameters. By optical relation of hyperbolic surface, it needs the parameter a, b, and c. Then, the light from (d_2, z_2) to $(0, c)$ on the hyperbolic surface is (d, z), and it turns to $(0, -c)$ on the image plane (d_p, z_p). From Equation (4.6) in Section 4.2.3, (d_0, z_0) is known. Therefore, (d_p, z_p) is obviously acquired by the optical relation of hyperbolic surface. The equations are from following Equation (4.7) to (4.12). The known parameters

49

are $d_1, d_2, z_1, z_2, d_c, z_c, d_0, z_0, z_p$, a, b, and c. The parameters d, z, and d_p is used.

Through the following equation, the result is shown in Figure 4.25. However, the

arrow represents the error measurement in LRF scan. The LRF scan converts the pixel

in omnidirectional camera image. The LRF mounts on 90 centimeter as same as the

wall Height. So the measurement in LRF scan may be some error. Figure 4.26(c)

shows the left line, the right line, and the lower line of bounding box.
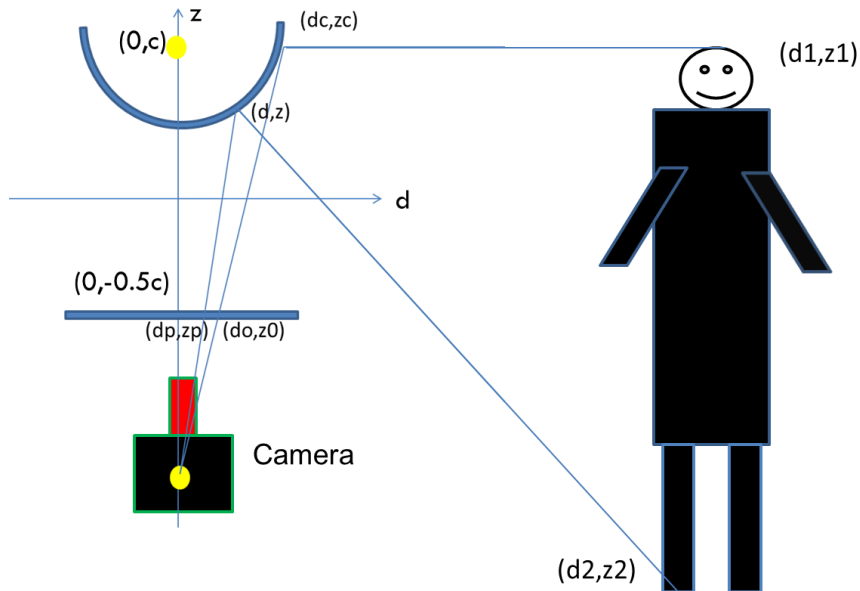

Figure 4.24 The Omnidirectional Camera Model with Pedestrian Estimation in Omnidirectional Camera Image Pixel from Projection Center

$$\text{dis} = \left(\frac{z_2 - z_1}{d_1} \right) \tag{4.7}$$

$$d_c = \sqrt{\frac{a^2}{-1 + \left(\frac{c}{b} \right)^2}} \tag{4.8}$$

$$d = \frac{-dis.z_1 - \sqrt{(dis.z_1)^2 - \left(dis^2 - \left(\frac{b}{a} \right)^2 \right) . (z_1{}^2 - b^2)}}{dis^2 - \left(\frac{b}{a} \right)^2} \tag{4.9}$$

$$z = z_1 + dis.d \tag{4.10}$$

$$d_0 = 0.25 d_c \tag{4.11}$$

$$z_0 = z_p = -0.5c \tag{4.12}$$

$$d_p = \frac{d.c}{2(c + z)} \tag{4.13}$$
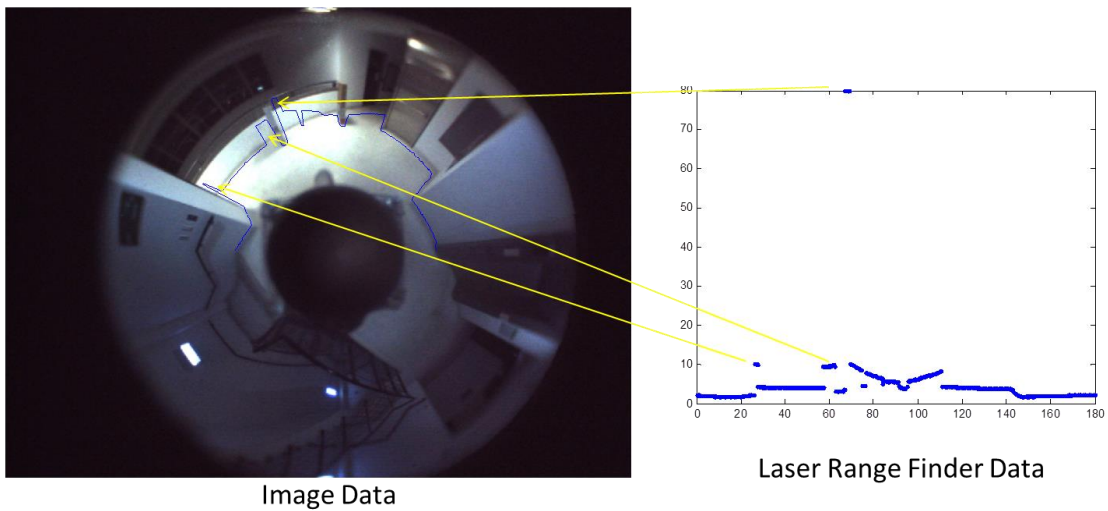

Image Data

Laser Range Finder Data

Figure 4.25 The Ground Bounding Box Result Demonstration (unit: pixel & meter) The LRF scan converts the pixel in omnidirectional camera image. The LRF mounts on 90 centimeter as same as the wall Height. So the measurement in LRF scan may be some error.

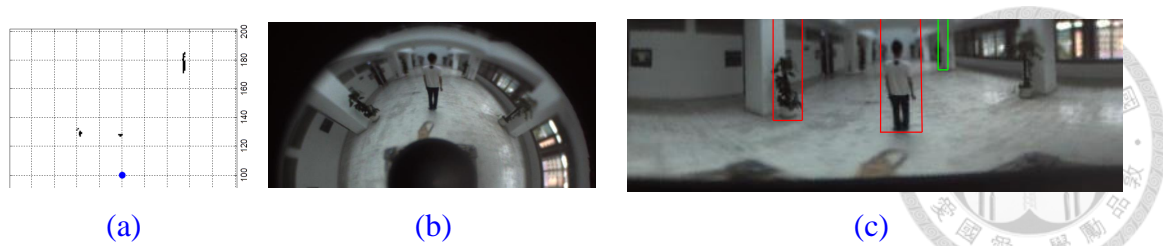<div align="center">(a)   (b)   (c)</div>

Figure 4.26 The Left Line, the Right Line, and the Lower Line of Bounding Box.

(a) The measurement points in LRF scan with some preprocessing

(b) Omnidirectional camera color image with semi image

(c) Left line, right line, and lower line of bounding box with panorama image

4.3.3 Upper Line of Bounding Box Extraction

For the bounding box, the upper line of bounding box extraction is a difficult task owing to unknown height of pedestrian. In the pedestrian detection, head detection is a practical method. In this thesis, Hough circle transform and color distribution are conditional judgments for the pedestrian detection [48: Zhao et al. 2012].

Hough circle transform is a circle with radius r and center (x_0, y_0) can be expressed as the following parametric equations:

$$x = x_0 + r_0.\cos\theta \tag{4.14}$$

$$y = y_0 + r_0.\sin\theta \tag{4.15}$$

where the angle θ is the full range 360 degree. Hough circle transform owns a voting mechanism from (x, y) coordinate to (x_0, y_0, r) coordinate [48: Zhao et al. 2012], [36: Ballard 1981]. Figure 4.27 illustrates the Hough circle transform procedure. First, any

point in (x, y) coordinate converts to a circular cone in (x_0, y_0, r_o) coordinate through Equation (4.14) and (4.15), as shown in Figure 4.27(a) and (b). Next, owing to adding other point, another circular cone in (x_0, y_0, r_o) appears, as shown in Figure 4.27(c) and (d). Finally, the voting number decides the circle. That the (x, y) coordinate converts to the (x_0, y_0, r_o) coordinate is two-dimension converts to three-dimension. So the radius usually is fixed in some range, as shown in Figure 4.27(e) and (f). For the pedestrian detection, circle radius in real-life is approximately from 15 cm to 20 cm. And head body ratio is 4 to 9.

For the color distribution [33: Zhao et al. 2008], [49: Rahimi et al. 2013], YCbCr color space extracts luminance in Y channel. The color space decides pedestrian model with both Cb channel and Cr channel. Equation (4.16) describes RGB color space converts to YCbCr color space [50: YCbCr from wiki 2014]. Equation (4.17) and (4.18), the color space threshold value in pedestrian detection is presented in [33: Zhao et al. 2008], [49: Rahimi et al. 2013], [48: Zhao et al. 2012].

For above conditional judgment, the pedestrian can be distinguished. Figure 4.28 shows a test of pedestrian detection. As expected, the Hough circle transform can detect the head of pedestrian, as shown Figure 4.28(d). In summary, the Hough circle transform and the color distribution are efficiently the methods for pedestrian detection.
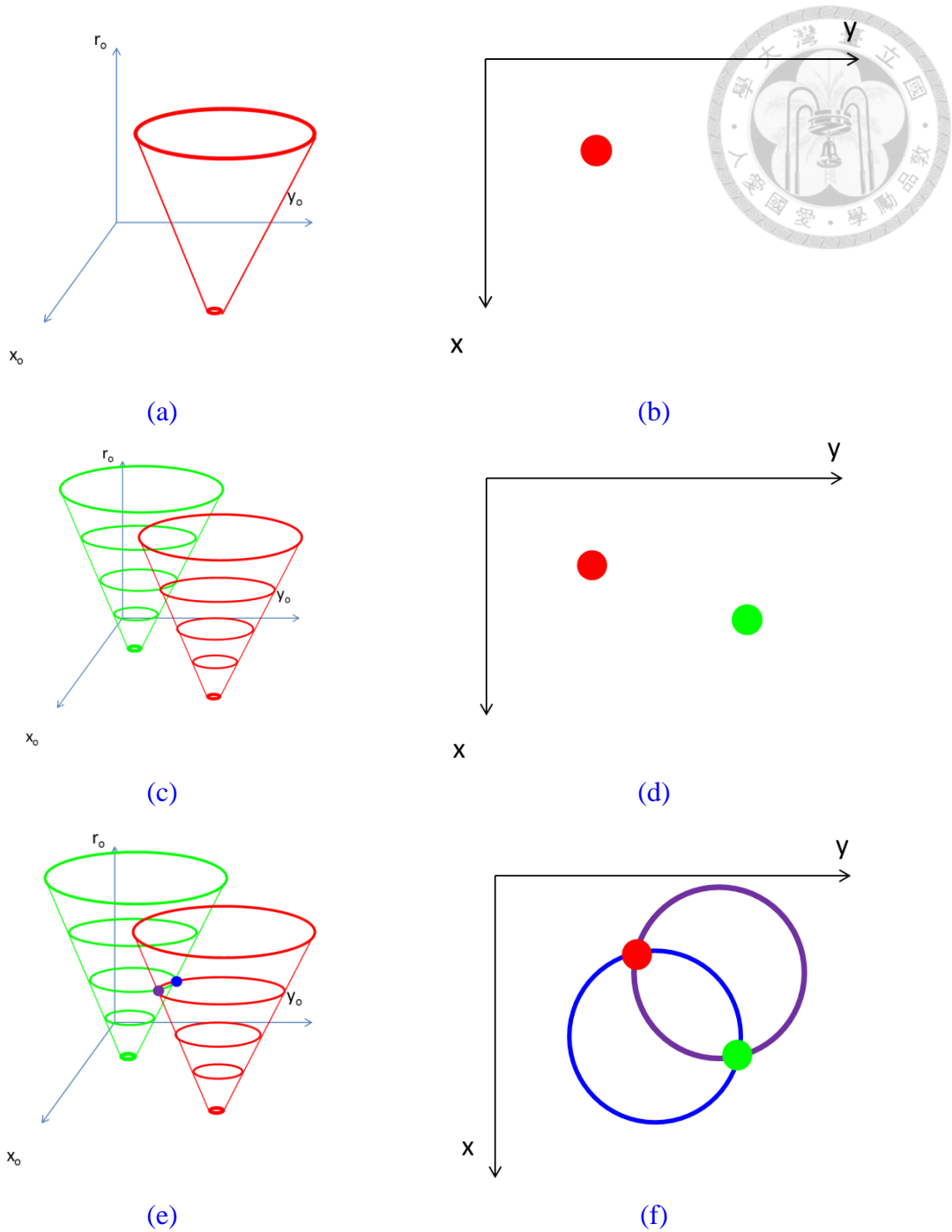
Figure 4.27 The Hough Circle Transform

(a) The (x_0, y_0, r) coordinate with Equation (4.14) and Equation(4.15)

(b) The (x, y) coordinate from (a) with Equation (4.14) and Equation(4.15)

(c) and (d) two points in (x, y) coordinate make a line intersection

(e) and (f) set the specific plane r value for Hough circle transform making the dimension reduction.

Then (a)~(f) is Hough circle transform operation process

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.533 & 24.966 \\ -37.797 & -74.203 & 112.000 \\ 112.000 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$ (4.16)

$$R_{Cb} = [111,170]$$ (4.17)

$$R_{Cr} = [115,195]$$ (4.18)

(a)

(b)   (c)   (d)
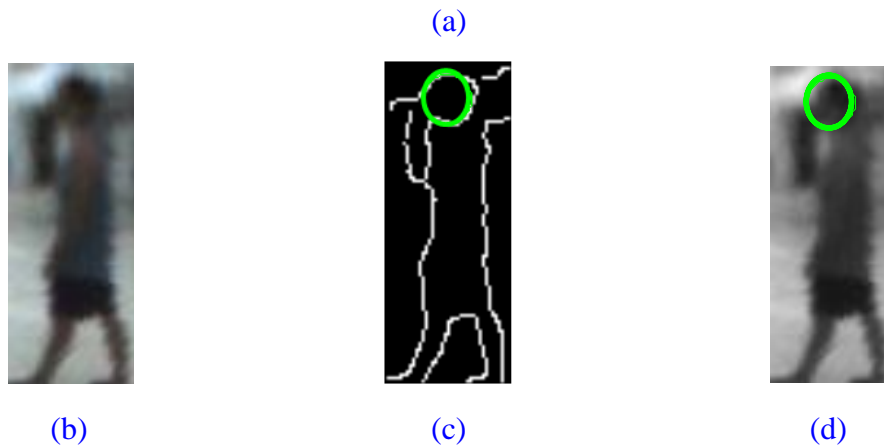
Figure 4.28 Test of Pedestrian Detection

(a) Original panorama image in Ming-Da 5F

(b) Original image in specific region

(c) Histogram equalization, Canny edge detection, and Hough circle transform

(d) Circle radius, head body ratio, and color distribution verification

4.3.4 Pedestrian Points Filtering in Static Map Construction

To obtain accurate the global static map, the pedestrian filtering is an important item in static global map construction. When the pedestrian points are filtered, the other points in scan are matched by PSO algorithm in Section 3.2. The static global

map is accurate owing to the pedestrian points filtering. The procedure is shown in

Figure 4.29. The pedestrians are detected with LRF scan at time t and t+1, as shown

in Figure 4.29(a). However, the pedestrians are filtered at time t and t+1, as shown in

Figure 4.29(b). The scan matching between two LRF scans, as shown in Figure

4.29(c). However, the thesis only states the idea. In the future work, pedestrian points
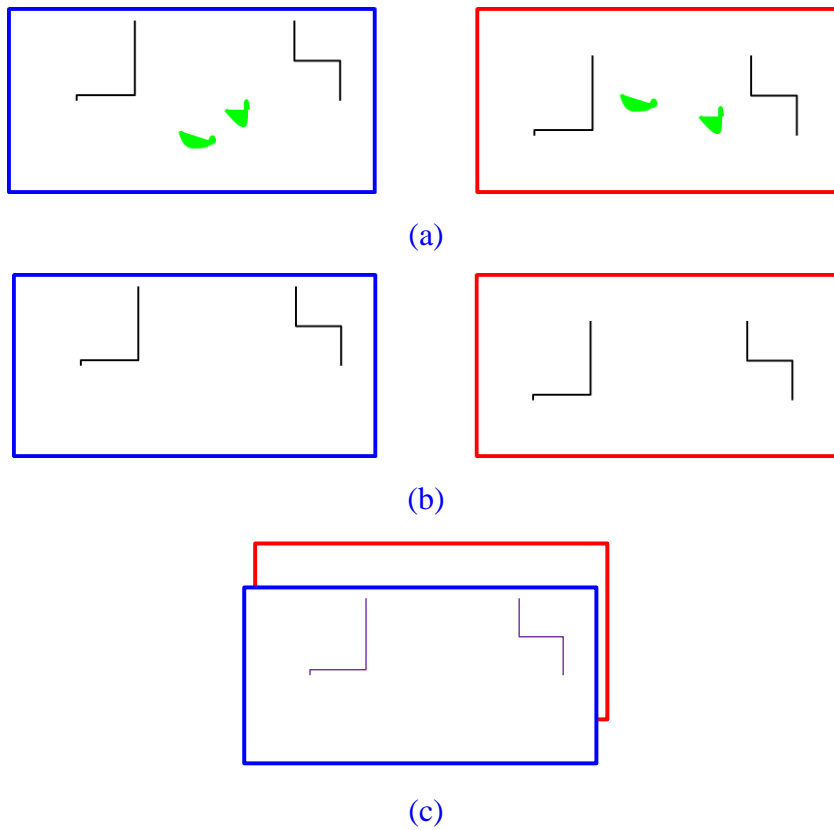
filtering in global static map may be a practical method.

(a)

(b)

(c)

Figure 4.29 Pedestrian Points Filtering Process
(a) Origin raw data in time t and t+1
(b) Pedestrian points filtering in time t and t+1
(c) Static global map in pedestrian points filtering in time t and t+1

4.4 Target Pedestrian Tracking

For target pedestrian tracking, the color image is important. However, the color image seems to solve the problems in Section 1.2. Section 4.4.1 and Section 4.4.2 describe the color distribution and the color texture. Section 4.4.3 provides a method Bhattacharyya distance for two histograms comparison. With the coefficients, Section 4.4.4 states the update method.

4.4.1 Color Distribution

Color distribution is a method to track target pedestrian. Color distribution can divide into three independent channels. In this thesis, YCbCr color space is presented in [50: YCbCr from wiki 2014]. The RGB color space converts the YCbCr color space with Equation (4.16) in Section 4.3.3. The YCbCr color space is widely used to digital video. Y is the luma component, Cb is the blue-difference chroma component, and Cr is the red-difference chroma component.

In the target pedestrian tracking, the color image is used. Generally, color spaces of two near images are similar. Therefore, the all channels of YCbCr color space are simultaneously adopted.

4.4.2 Local Binary Patterns

To track target pedestrian, the color distribution is obviously insufficient. For each pixel, neighbor relative intensity is also important. In [49: Rahimi et al. 2013], LBP is a texture operator. The LBP algorithm states the neighbor relative intensity for each pixel. Owing to the LBP algorithm, the target pedestrian tracking is robust against pedestrian deformation.

LBP algorithm is often used because of low computational complexity [49: Rahimi et al. 2013]. The LBP value represents relative intensity with neighbor. In an intensity of 3x3 cells, the center cell $I0$ compares with neighbor. If $I0$ is larger than the cell, the cell marks 0. Likely, if the $I0$ is smaller than the cell, the cell marks 1. Figure 4.30(b) is a result through the LBP operator in Figure 4.30(a). The LBP value is binary bit from $I1$ to $I8$, as shown in Figure 4.30(b). The LBP value is $(10100100)_2$. The decimal is 164 in Equation (4.19). Each pixel in image does the LBP operator. Therefore, the result is LBP image.

(a)   (b)

Figure 4.30 The Example for LBP Operator

(a) The intensity of 3*3 cell from pixel I0

(b) The LBP result from Figure 4.30 (a)

Decimal: $1 * 2^{(8-1)} + 1 * 2^{(6-1)} + 1 * 2^{(3-1)} = 164$ (4.19)

4.4.3 Bhattacharyya Distance

Bhattacharyya distance is used to judge the similarity of two histograms [51: Bhattacharyya distance from wiki 2014]. Equation (4.20) represents the Bhattacharyya distance coefficient (BC), N is the number of intensity, N_{total} is the total number of pixels in image, a_i is the number of previous color space for intensity i, and b_i is the number of current color space for intensity i. That the BC equals one means two histograms is same. That the BC is close to zero means two histograms are different. Figure 4.31 are examples for the BC.

$$BC = \sum_{i=1}^{N} \frac{\sqrt{a_i * b_i}}{N_{total}}$$ (4.20)
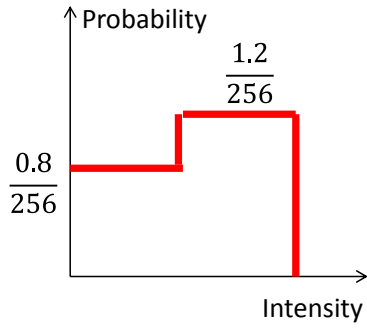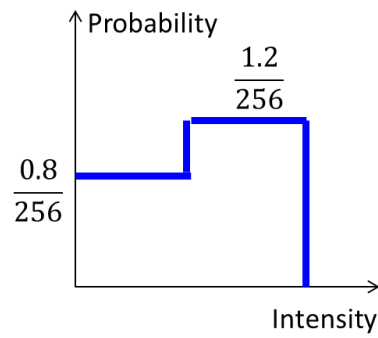
(a) previous color histogram   (b) current color histogram   (c) BC

$$BC = \sum_{i=0}^{127} \sqrt{\frac{0.8}{256} * \frac{0.8}{256}} + \sum_{i=128}^{255} \sqrt{\frac{1.2}{256} * \frac{1.2}{256}} = 1.000$$
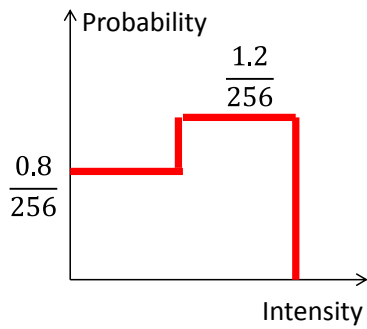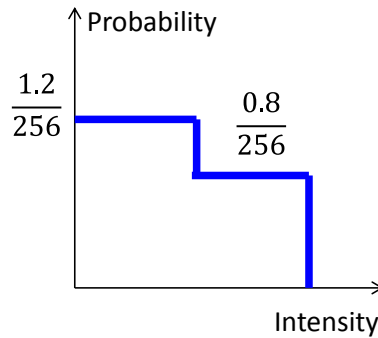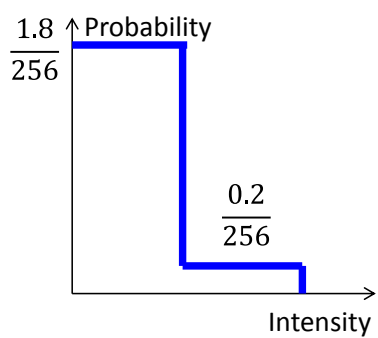
(d) previous color histogram   (e) current color histogram   (f) BC

$$BC = \sum_{i=0}^{127} \sqrt{\frac{0.8}{256} * \frac{1.2}{256}} + \sum_{i=128}^{255} \sqrt{\frac{1.2}{256} * \frac{0.8}{256}} = 0.9798$$
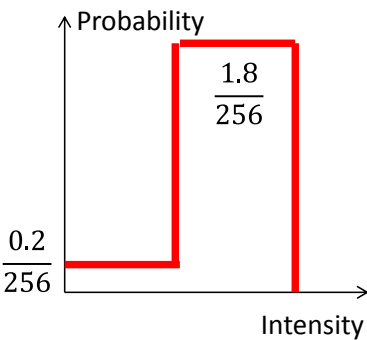
(g) previous color histogram   (h) current color histogram   (i) BC

$$BC = \sum_{i=0}^{127} \sqrt{\frac{0.2}{256} * \frac{1.8}{256}} + \sum_{i=128}^{255} \sqrt{\frac{1.8}{256} * \frac{0.2}{256}} = 0.600$$

Figure 4.31 The Bhattacharyya Distance Coefficient Illumination

(a)~(c) The two same histogram and Bhattacharyya distance coefficient

(d)~(f) The two different histogram and Bhattacharyya distance coefficient

(g)~(i) The more different the two histogram are, the closer to 0 the Bhattacharyya distance coefficient is

4.4.4 Database Update

For each candidate pedestrian data, it owns four Bhattacharyya distance coefficients including Y color intensity histogram, Cb color intensity histogram, Cr color intensity histogram, and LBP intensity histogram. If the four Bhattacharyya distance coefficients are larger than threshold values, the pedestrian is the target pedestrian. The new pedestrian replaces the original target pedestrian to the new target pedestrian.

For data association, spatial constraint needs to be considered. If the Euclidean distance between original pedestrian data and new pedestrian data is larger than a threshold value in a period time, the spatial constraint deletes the pedestrian data. Over the period time, the pedestrian location is not important.

With both the four Bhattacharyya distance coefficients and the spatial constraint, the target pedestrian tracking is complete.

Chapter 5

Experimental Results and

Analysis

Several experiments are undertaken in Ming-Da Building. Self-localization, mapping, pedestrian detection, and target pedestrian tracking are evaluated with the offline experimental systems. Section 5.1 states the equipment in the experiments. Section 5.2 states the sensors calibration between the LRF and the omnidirectional camera. What is more, Section 5.2 discusses the error of sensor measurement. In Self-localization and mapping, Section 5.3 states methods including PSO algorithm and occupancy grid map. The accuracy of mapping results by ICP algorithm and PSO algorithm is shown in Section 5.4. For pedestrian detection, Section 5.5 shows the performance and the process. Likely, Section 5.6 shows the performance and the process in target pedestrian tracking.

5.1 Hardware Platform

The following equipment is used in the experiment, as shown in Figure 5.1.

The Pioneer 3-DX in Figure 5.1(a) is a differential-drive mobile robot. The

LMS-100 is a LRF providing the distance scan in a plane. And the VS-C14U-80-ST is

an omnidirectional camera, and the field of view (FOV) of the omnidirectional

camera covers the 360 degree. Both the LRF and the omnidirectional camera are

mounted on the mobile robot, as shown in Figure 5.2. The VS-C14U-80-ST is

mounted on the mobile robot at 180 cm from ground. What is more, the LMS 100 is

mounted on the mobile robot at 90 cm from ground.

The URG-04LX-UG01 is also LRF used in these experiments as a ground truth

of the robotic position, as shown in Figure 5.1(d).

(a)   (b)

(c)   (d)

Figure 5.1: The Experimental Instrument
(a) The model number Pioneer 3DX from mobile robot
(b) The model number LMS-100 from LRF
(c) The model number VS-C14U-80-ST from omnidirectional camera
(d) The model number URG-04LX-UG01 from LRF
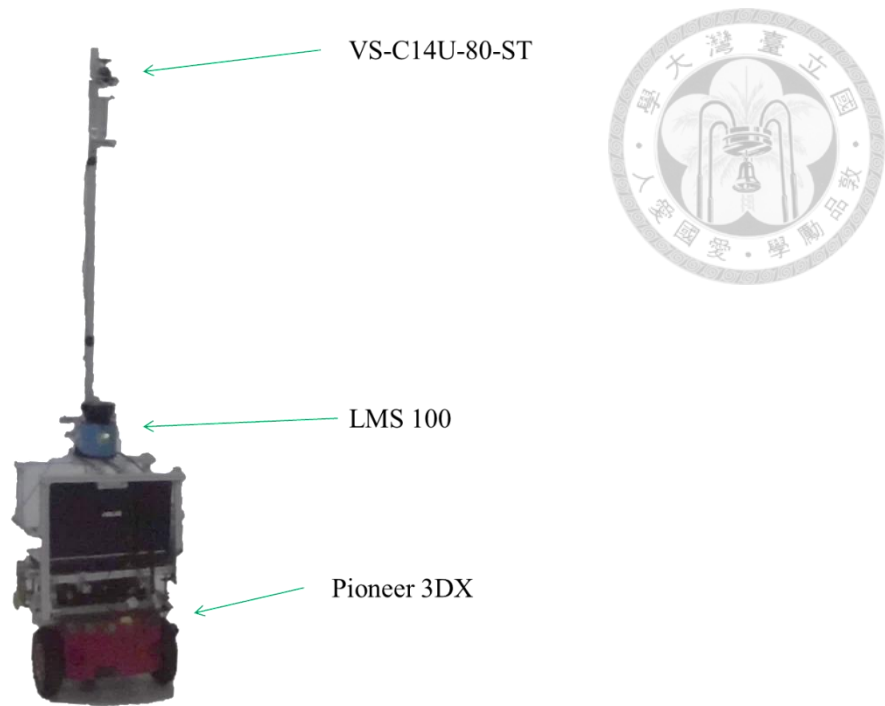
Figure 5.2 The Experimental Mobile Robot
The VS-C14U-80-ST is mounted on the mobile robot at 180 cm from ground.
And the LMS 100 is mounted on the mobile robot at 90 cm from ground.

5.2 Accuracy of Sensor Measurement

In Section 5.2.1, the LRF scan accuracy is tested. Section 5.2.2 analyzes

accuracy in both the panorama image of omnidirectional camera and rotation

calibration.

5.2.1 Laser Range Finder Accuracy

I. LMS-100

The experimental results probe into the accuracy of LRF before using the

equipment. The main tasks include two analyses. One is acceptance rate, and the other

is error. The details are presented as follows.

Acceptance rate means the sensor accepts the emissive light. Table 5.1 shows the total acceptance in each angle with 100 pieces of data in the environment with glass. The measurement points in the environment with glass, as shown in Figure 5.3(a).

In this experiment, Section 3.1.2 states the glass environment affecting the acceptance rate. Then, Figure 5.4 demonstrates the analysis of acceptance rate in the glass environment. Here, the x-axis is angle, and the y-axis is percentage. The glass environment may cause the acceptance rate reduction. To prove the thinking, the same style experiment is also implemented in network control systems laboratory (NCSLab) with non-glass environment, as shown in Figure 5.5. The measurement points in 60 cm plane from ground and the acceptance rate achieve 100% with LRF in each angle in Figure 5.6. In addition, the acceptance rate in glass environment is 95.6% in Table 5.1. In this condition, the glass environment affects the acceptance rate.

(a)

(b)

(c)
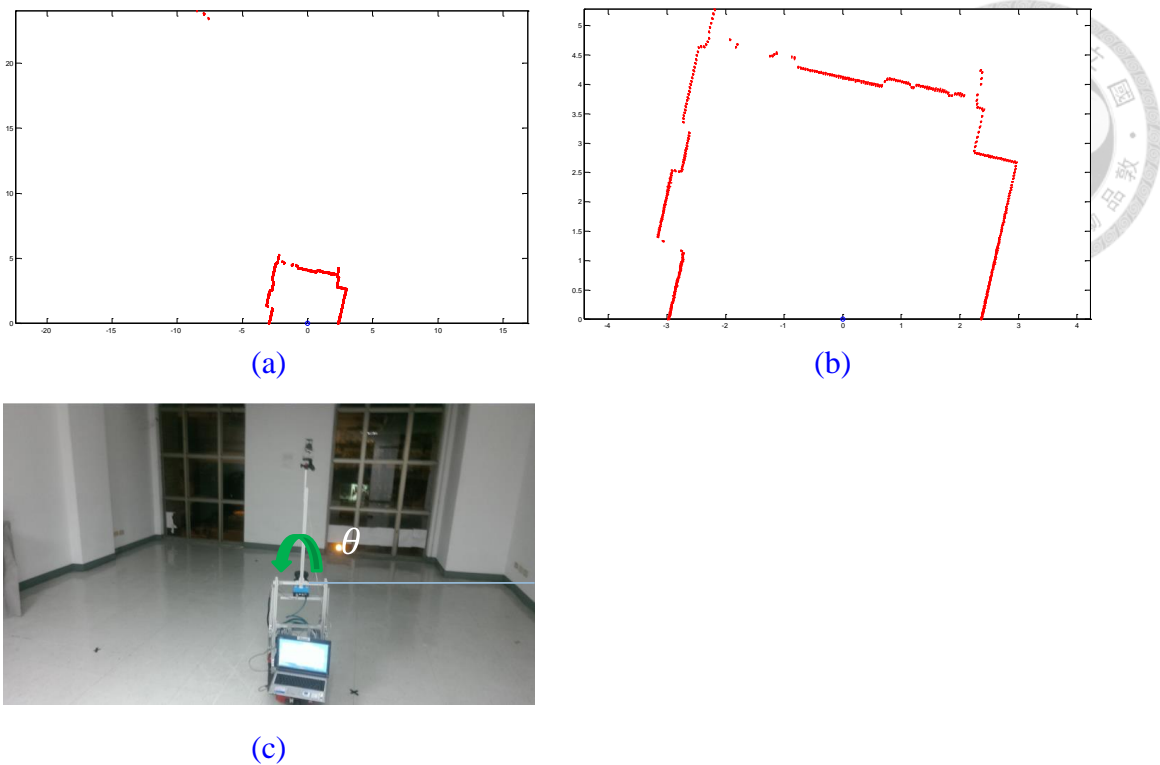
Figure 5.3 Experimental Environment with Glass in Ming-Da Building 2F (unit:m)

(a) Measurement points in 60 cm plane from ground, some distances exceed 20 m

(b) The measurement points contain less 20 m for LRF scan from (a).

(c) Real-scene in environment with digital camera

Table 5.1 Acceptance Rate in the Ming-Da Building 2F

Measurement Number (Total 100 data)	Total of Each Angle (361 angle)	Percentage
0	12	3.3%
1~99	4	1.1%
100	345	**95.6%**
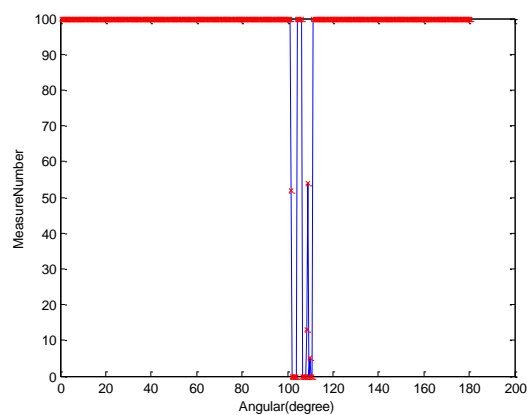
Figure 5.4 Plot the Acceptance Rate with Each Angle in the Table 5.1
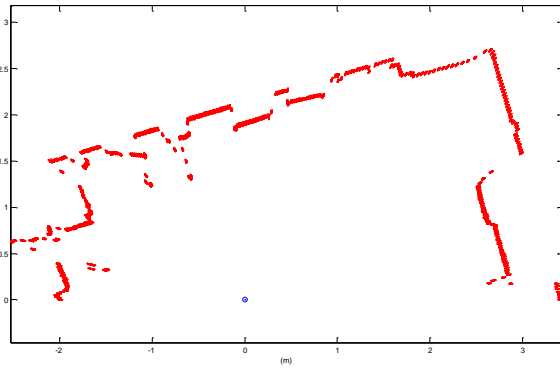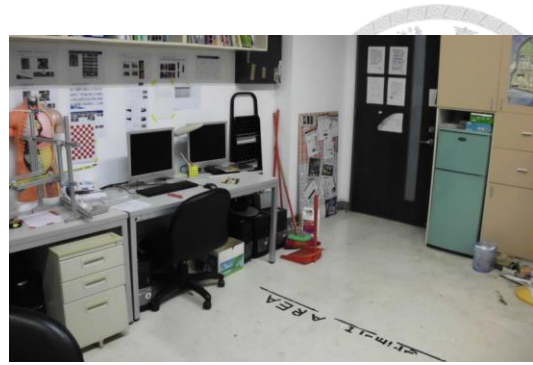
(a)   (b)

Figure 5.5 Experimental Environment in Ming-Da Building Lab 601 (NCSLab)

(a) Measurement points in 60 cm plane from ground and the acceptance rate
    achieve 100% with LRF in each angle (unit: meter)

(b) Real-scene with digital camera in non-glass environment

To judge the data performance, the mean and the error are presented. Equation

(5.1) and (5.2) show the mean and the error:

$$\text{mean(data)} = \frac{\sum all\ data}{item\ number} \tag{5.1}$$

$$\text{error} = \frac{\sum(data - mean(data))}{item\ number} \cdot \frac{100\%}{mean(data)} \tag{5.2}$$

The following results in Figure 5.6 indicate the error analysis from Equation (5.2)

with different distance. In length 4.5 meter with maximum error, the error achieves up

to 2%. For the data number from 100 pieces of data to 650 pieces of data, the error

does not obviously change.

67
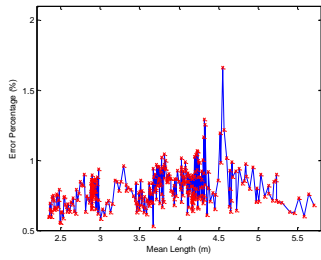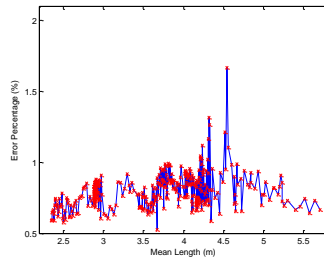
(a) 100 Sampling data   (b) 150 Sampling data   (c) 200 Sampling data
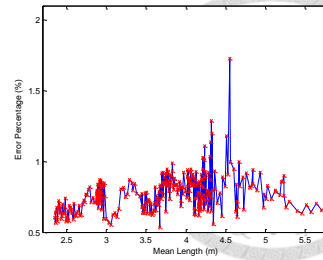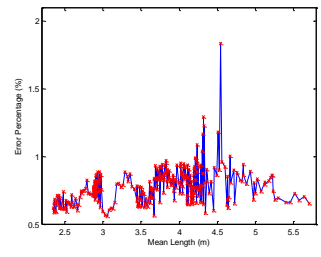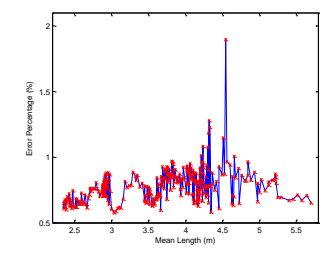
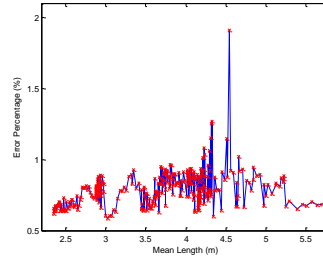(d) 250 Sampling data   (e) 300 Sampling data   (f) 350 Sampling data
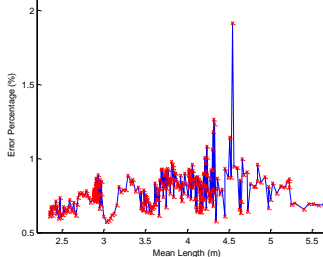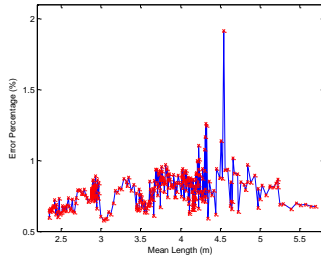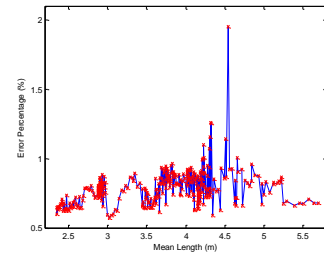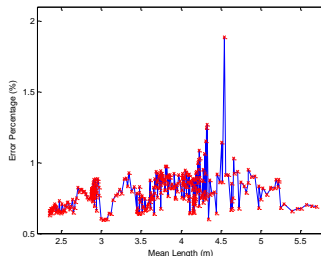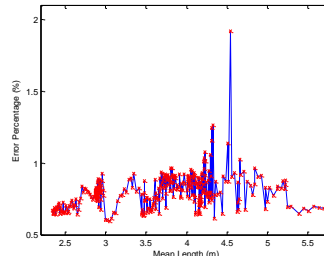
(g) 400 Sampling data   (h) 450 Sampling data   (i) 500 Sampling data

(j) 600 Sampling data   (k) 650 Sampling data

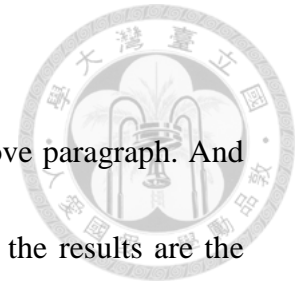Figure 5.6 Data Number of Sampling in Distance Length with Error Percentage

Analysis

x-axis: distance length

y-axis: error percentage

The maximum error is about 4.5 meter achieving up to 2%.

II. URG-04LX-UG01

Similarly, both acceptance rate and error are presented in above paragraph. And

the experiments are the same with above paragraph. In NCSLab, the results are the

following in Table 5.2, Figure 5.7, Figure 5.8, and Figure 5.9. The real-scene in

NCSLab is shown in Figure 5.7. However, the measurement unstable with different

value for corner, as shown in Figure 5.7(c) and (d). The acceptance rate is only 54.8%.

However, the acceptance rate of LMS-100 in NCSLab can achieve 100%. In the error

analysis, the error in different number data is shown in Figure 5.6. The error analysis

is for URG-04LX-UG01 from Equation (5.1) and (5.2). Although, the regular is not

found, there are two peaks close to 0.5m and 2m in these data. Compared with

LMS-100, the mean error is 0.56%. The mean error is smaller than LMS-100 0.81%.

But the maximum error in URG-04LX-UG01 achieves 40.02%. The maximum error

in LMS-100 is only 2.33%. In summary, the error of URG-04LX-UG01 has smaller

mean error than LMS-100. However, the maximum error of URG-04LX-UG01 is

larger than LMS-100. Table 5.3 states the details.
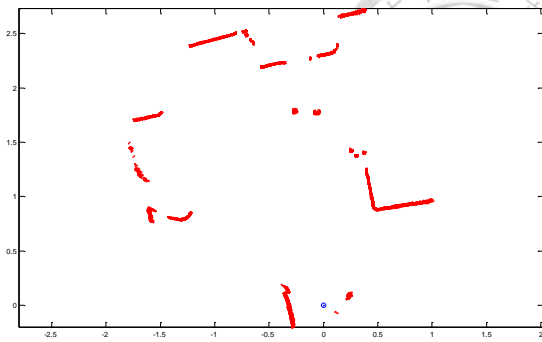
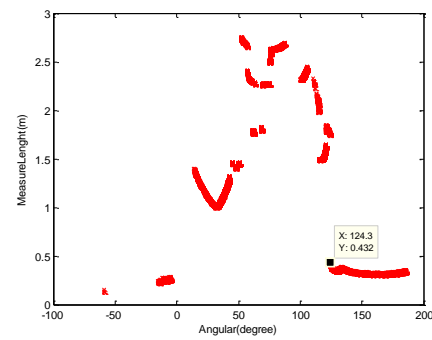| (a)	(b)
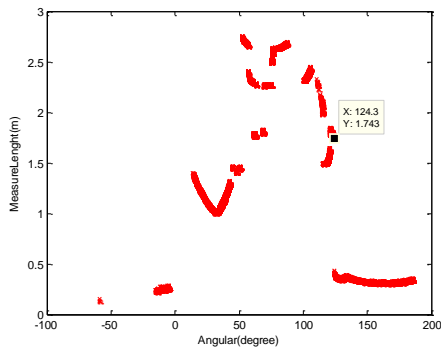
(c)	(d)

Figure 5.7 The Measurement Points of the LRF Scans

(a) Real-scene with digital camera

(b) Plot measurement points in NCSLab (unit: m)

(c) The left value of 124.3 degree: 1.743% error with 100 pieces of data

(d) The right value of 124.3 degree: 0.432% error with 100 pieces of data

(c) and (d) represent the measurement unstable with different value for corner

Table 5.2 Acceptance Rate in the NCSLab in 100 data

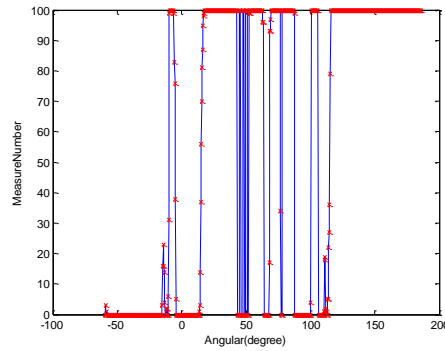| Measurement Number (Total 100 data)	Total of Each Angle (361 angle)	Percentage
0	253	37.1%
1~99	55	8.1%
100	374	**54.8%**

Figure 5.8 Plot the Acceptance Rate Result with Each Angle in Table 5.2
The x-axis is angular information, and the y-axis is measurement number. In this picture, the acceptance is about 50%.

(a) 100 data   (b) 150 data   (c) 200 data

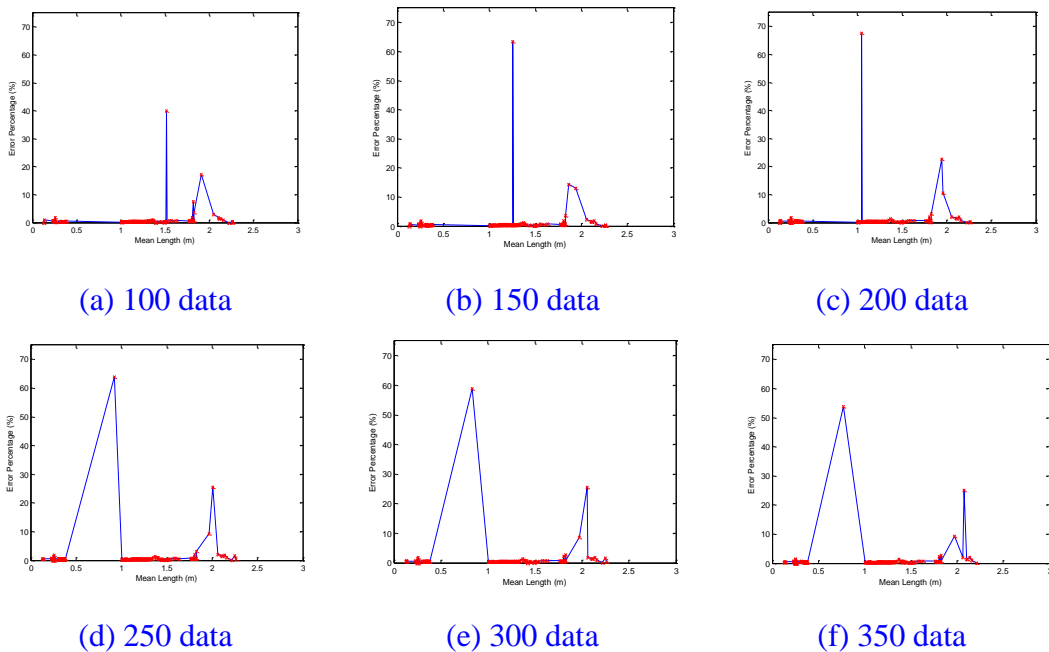(d) 250 data   (e) 300 data   (f) 350 data

Figure 5.9 Errors with Equation (5.1) and Equation (5.2)
x-axis: distance length
y-axis: error percentage
The regular of measurement error is not found. But, there are two peaks close to 0.5m and 2m in these data.

The LMS100 and the URG-04LX-UG01 have their advantages. Table 5.3 states the properties of LRFs. The items include measurement range, measurement angle,

resolution, measurement full ratio, maximum error, mean error, and scan time.

Table 5.3 Properties of Laser Range Finder

URG-04LX-UG01	Item	LMS100
0.02m ~ 4m	Measurement Range	**0.05 m ~ 20 m**
240 °	Measurement Angle	180 °
0.36 °	Resolution	0.5 °
54.8 %	Measurement Full Ratio	**95.6 %**
40.02 %	Maximum Error	**2.33 %**
0.56 %	Mean Error	0.81 %
100 msec	Scan Time	**20 msec**

5.2.2 Sensors Calibration

I.   Break Point and Angular Point Detection

The LRF uses the distance of two consecutive points to do segmentation. Then, the dynamic threshold D_{th} and δ are selected. The partitions divide into break points and angular points in Figure 5.11. In Figure 5.10(a), '∗' represents the break point and 'o' represents the angular point. Hence, the x-axis represents the number of angle. The 361 angle numbers cover the 180 degree, and the angle resolution is 0.5 degree, In addition, the LRF is mounted on the mobile robot at 90 cm from ground.
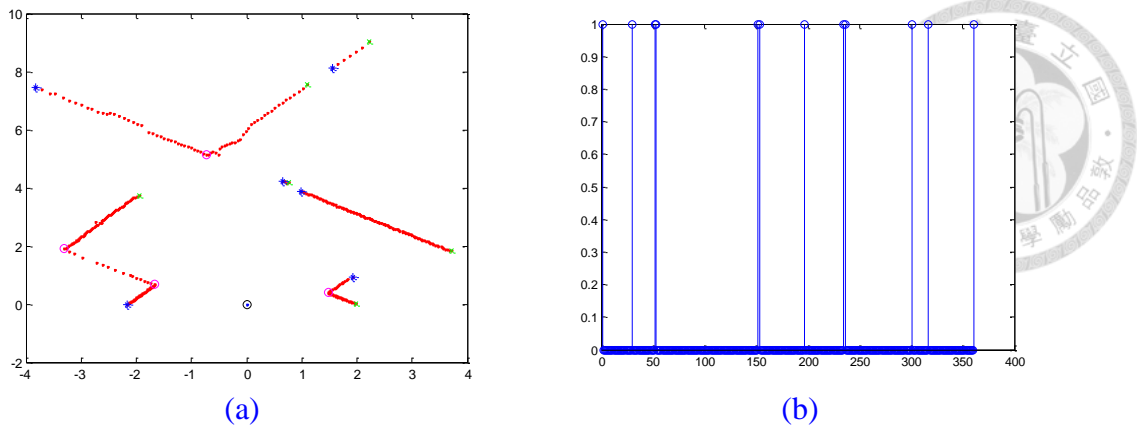
(a)   (b)

Figure 5.10 The Break Points and Angular Points with LRF Data in Ming-Da
Building 6F

(a) The break point is '∗' and the angular points 'o'

(b) The stem plot represents the break points and angular points with the number
of angle

x-axis: 0.5 degree is a number

y-axis: 1(true), 0(false)

Take for example in Figure 5.11, the '0°:14.5°(1.72)' means that 14.5° is angular

points from Section 4.2, 0° is angular point in Section 4.2, and 1.72 meter is the

maximum distance in the point to the line from start to end. Similarly, the

'26°:75.5°(3.55)' means that 26° and 75.5° are break points in Section 4.2, and 3.55

meter is the maximum distance in the point to the line from start to end.
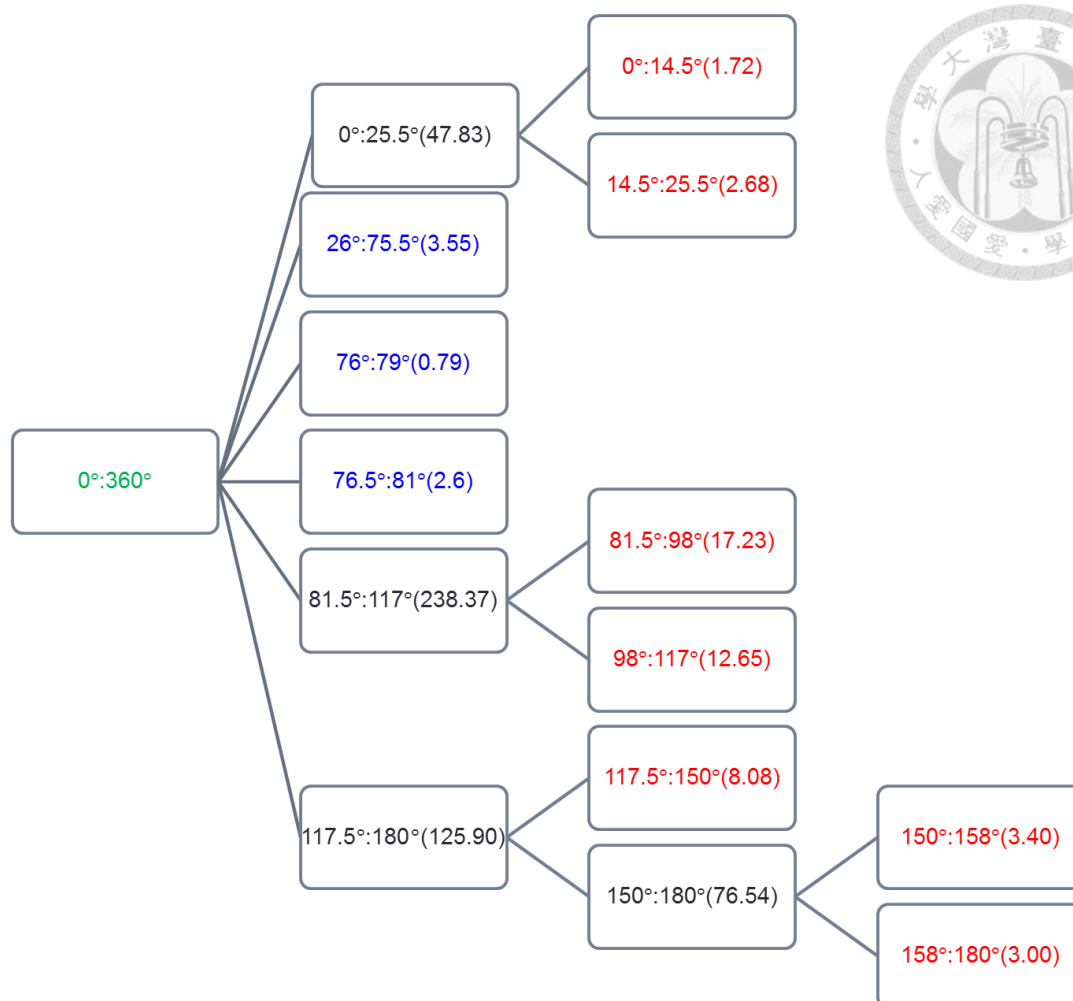
Figure 5.11 The classification represents the break point '∗' and
the angular point 'o'.

For example, 0°:14.5°(1.72) is a symbol. 0° and 14.5° are angular points, and 1.72 meter is the maximum distance in the point to the line from start to end. Similarly, 26°:75.5°(3.55) is a symbol. 26° and 75.5° are break points, and 3.55 meter is the maximum distance in the point to the line from start to end.

II. Vertical Line Detection

To do the panorama of omnidirectional camera, the projection center needs to be found. In the Ming-Da Building 4F, the floor owns the gray board considering the feature. The omnidirectional camera is put on the floor and turned over. The position of the omnidirectional camera in Ming-Da Building 4F is shown in Figure 5.12.

74

(a)   (b)

Figure 5.12 The Position of the Omnidirectional Camera in Ming-Da Building 4F
The omnidirectional camera is put on the floor and turned over.

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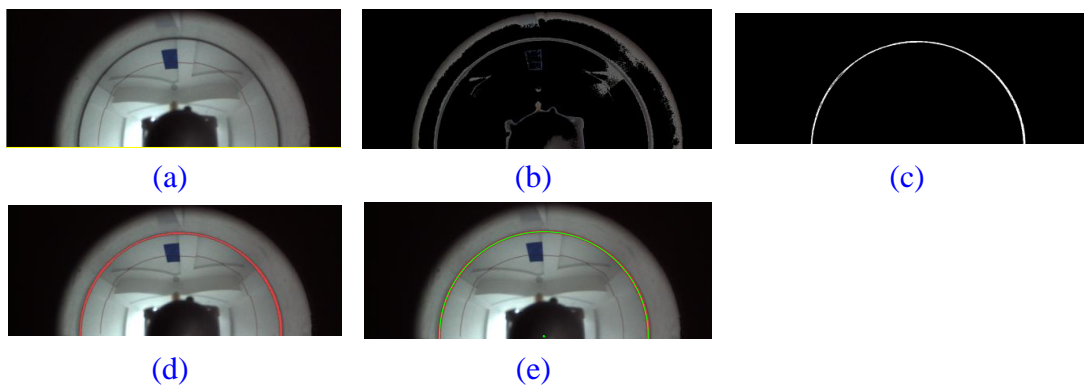
(d)   (e)

Figure 5.13 The Process in Finding the Projection Center in Ming-Da Building 4F
(a) Cut image in omnidirectional camera
(b) Set the threshold value in RGB color space from (a)
(c) Region growing in specific point from (b)
(d) Image fill the holes from (c)
(e) Least square fitting circle from points in (d)

And then, Figure 5.13 shows the procedure in finding the projection center through the HPVCL. First, Figure 5.13(a) is an input image. Then, the lower image is cut. Using the RGB filter, the gray color data is shown in Figure 5.13(b).Next, using the region growing, and the image is shown in Figure 5.13(c). Using the image filling, the image is shown in Figure 5.13(d). Finally, Table 5.4 shows the iterative results, the initial guess is the center of image (512,384). The second iterative result converges to

the projection center (530,394). Hence, the positive x-axis is right, and the positive

y-axis is down.

Table 5.4 Projection Center Search

	Initial result	First iterative	Second iterative	Final result
X (pixel)	512	530.147	530.1288	530
Y (pixel)	384	394.2445	394.2285	394

The resolution of panorama image is 350x721, as shown in Figure 5.14(b). The

resolution in Figure 5.14(a) is 1024x768. And the analysis of panorama needs to be

deeply probed. In Figure 5.15, the gray board is the feature extracted. The extracted

feature compares to the radius of projection center. First, using the RGB filter for

input image gets the image1. Then, using the Sobel horizontal edge detector in [6:

Gonzalez & Woods 2008] for the image1 gets image2. In image 2, picking the lower

line in two lines can get image3. Finally, using the opening algorithm in [6: Gonzalez

& Woods 2008] can get image4. The image4 includes the radius of projection center

names HPVCL.

(a)   (b)

Figure 5.14 The Panorama Image from Omnidirectional Camera
(a) The original image from omnidirectional camera
(b) The panorama image from (a)

(a)   (b)

(c)   (d)

(e)   (f)
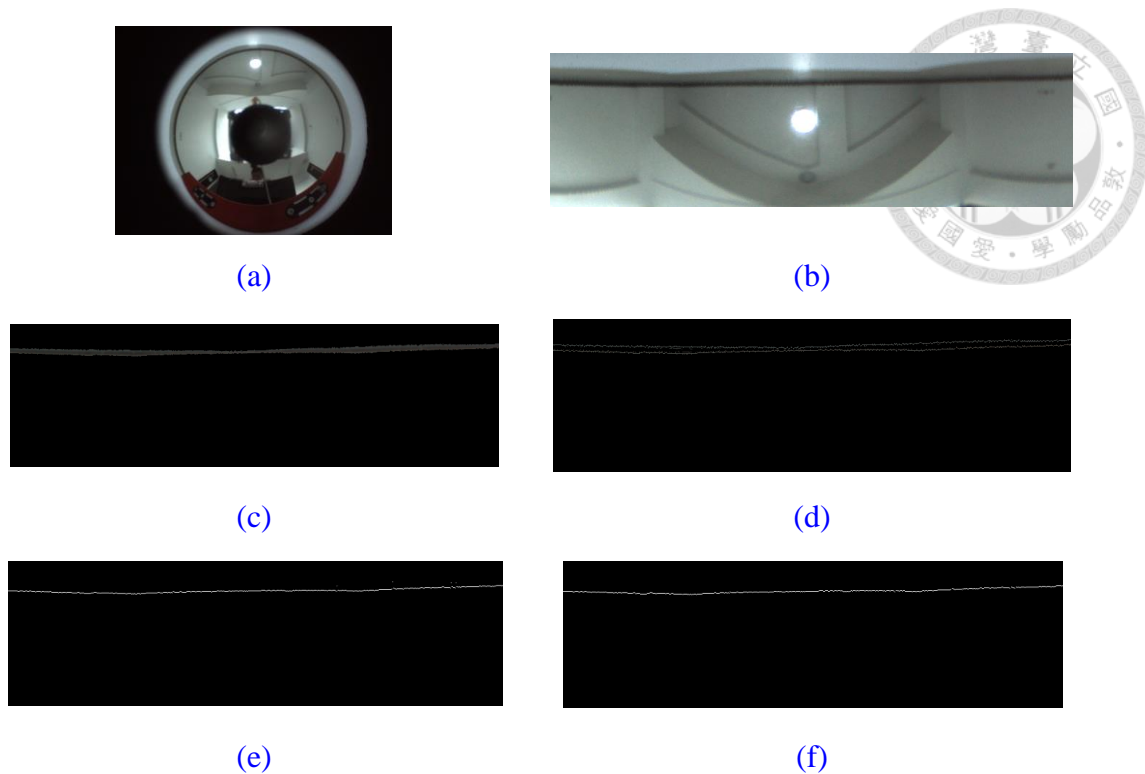
Figure 5.15 The SHL Extraction Method from Omnidirectional Camera in Ming-Da Building 4F

(a) The original image from omnidirectional camera

(b) The panorama image from (a)

(c) Using the RGB filter from (b)

(d) The Sobel horizontal edge detector in [6: Gonzalez & Woods 2008] from (c)

(e) Pick the down line in two line in (d)

(f) Using the opening algorithm in [6: Gonzalez & Woods 2008] from (e)

The accuracy of the HPVCL is shown in Figure 5.16 and the formulas are from Equation (5.3) and (5.4). In Figure 5.16, Current_Radius is the radius from projection center to gray board, and Pixel_Radius means the radius in Figure 5.13(e) is 308. The histograms are shown in Figure 5.16.
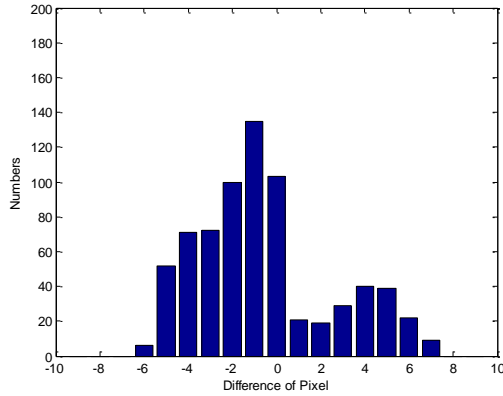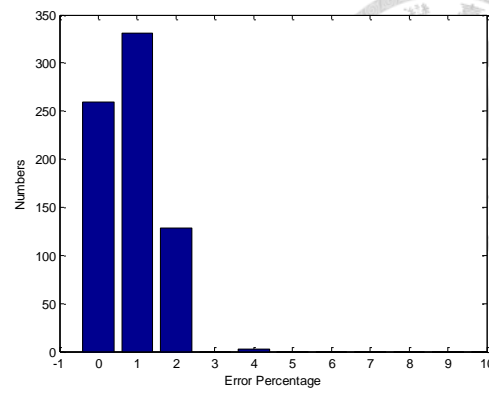
(a)   (b)

Figure 5.16 The analysis of Figure 5.15

(a) Difference of Pixel = Current_Radius - Pixel_Radius

   x-axis is the difference number, and y-axis is the number of pixels

(b) Error Pixel = abs(Difference of Pixel)/Pixel_Radius*100 %

   x-axis is the error percentage, and y-axis is the number of pixels

$$\text{Difference of Pixel} \; = \; \text{Current\_Radius} \; - \; \text{Pixel\_Radius} \tag{5.3}$$

$$\text{Error Pixel} = \left| \frac{\text{Difference of Pixel}}{\text{Pixel\_Radius}} \cdot 100\% \right| \tag{5.4}$$

The Sobel vertical edge detector results in Figure 5.17. And the histogram equalization improves the edge detection. Figure 5.17(c) is the original panorama image. However, Figure 5.17(g) is the vertical lines by Sobel vertical detector. As expected, Figure 5.17(g) appears more vertical line than Figure 5.17(c). The stem plot for vertical line is extracted, as shown in Figure 5.17(g). The data association needs to the specific pixel in vertical line of image, and the next paragraph states the detail.
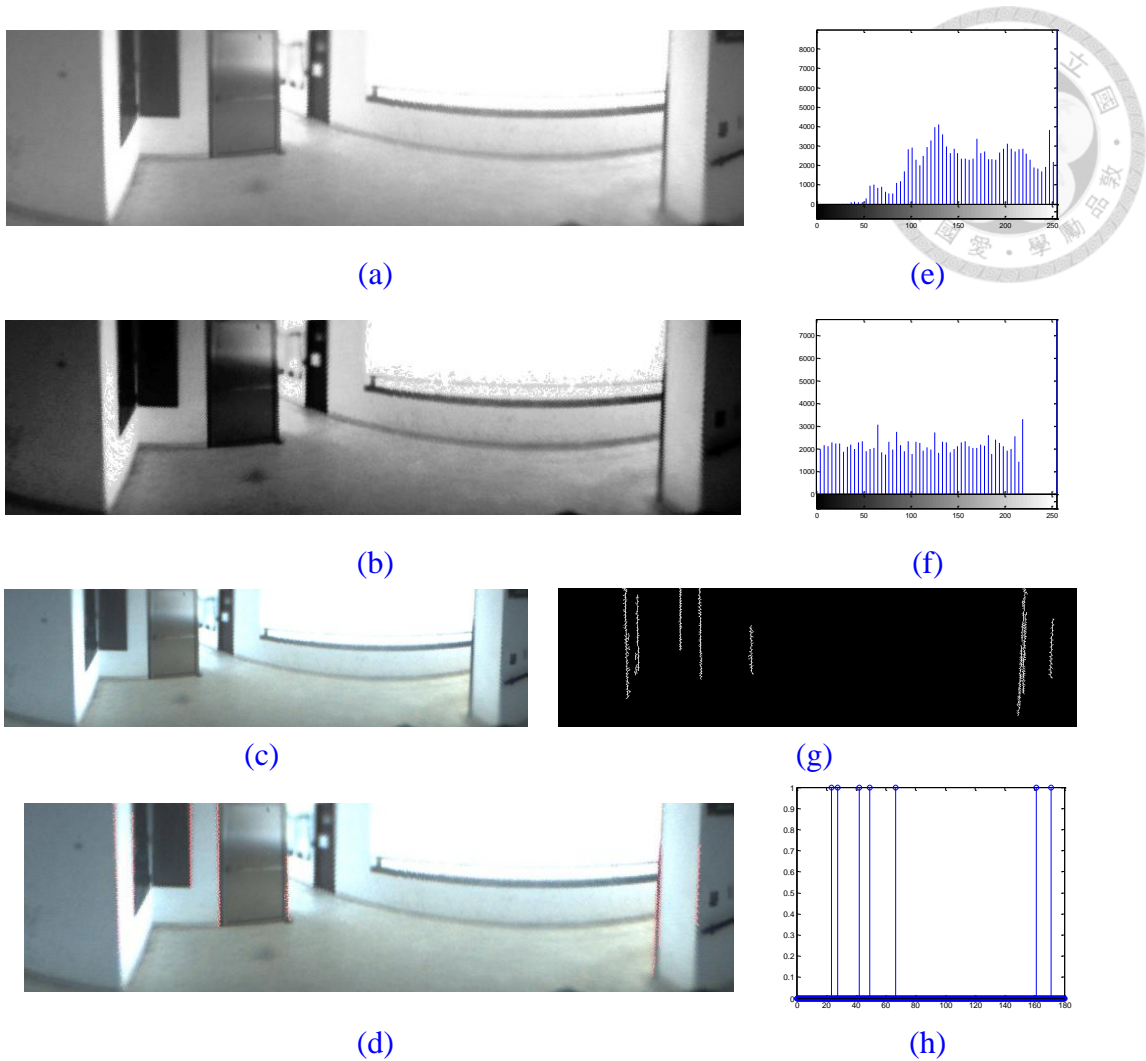
Figure 5.17 The Sobel Vertical Edge Detector for Ming-Da Building 6F

(a) Original image gray value from omnidirectional camera

(b) The histogram equalization process from (a)

(c) The panorama original image from 5.14(a)

(d) The vertical line angle from (c)

(e) The histogram from (a)

(f) The histogram from (b)

(g) The Sobel vertical line detection in (c)

(h) The stem plot for the real-scene in (d)

III. Angular Rotation Calibration

To pick the specific pixel matching the LRF scan point, the distance and pixel

relationship needs to be acquired. In Figure 5.18 and Table 5.5 with the distance and

the pixel relationship [12: Ueda et al. 2011], the LRF is mounted the same height of

the suitcase. The corner in LRF scan matches the suitcase corner pixel in color image.

The distance from LRF scan and pixel from color image relationship is shown in

Table 5.5. Therefore, the curve fitting of the distance and pixel is presented in Figure

5.19. The curve fits with a polynomial function with order four [14: Scaramuzza et al.

2006]. The function is $y = 0.0015x^4 - 0.0245x^3 - 1.0363x^2 + 22.4902x + 160.0462$.

The following six cases are shown in Figure 5.18. And the complete 20 cases are

shown in Appendix A.1.
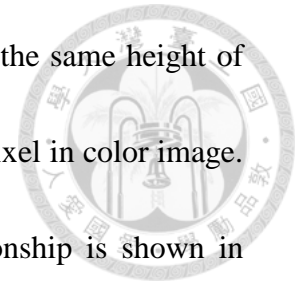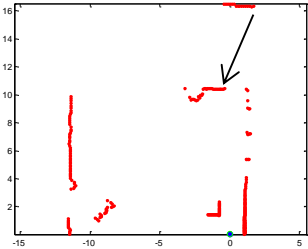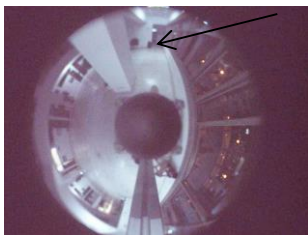
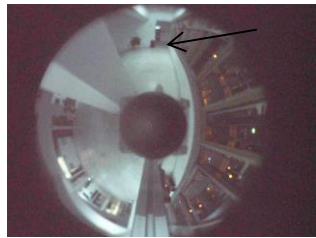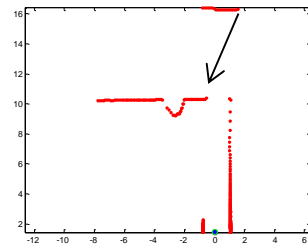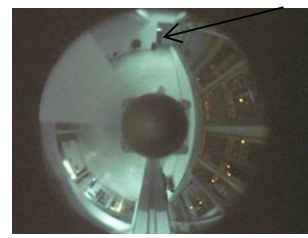**(a)** 10.46 m   **(b)** 10.32 m   **(c)** 10.23 m
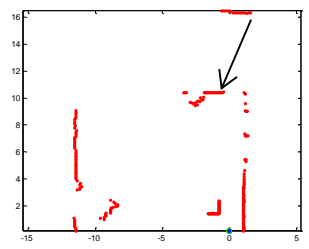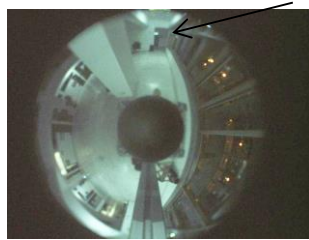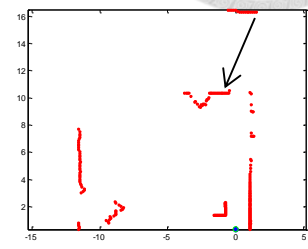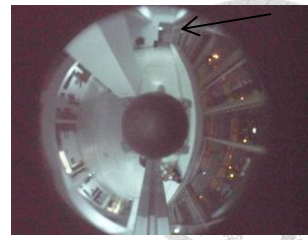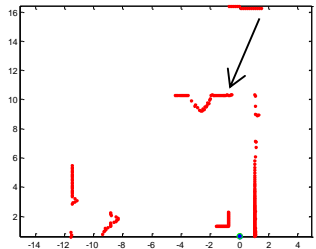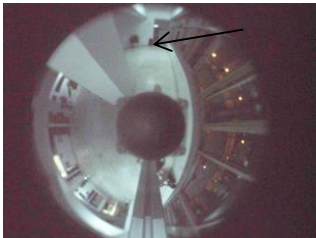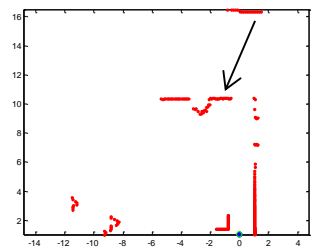
(d) 9.71 m   **(e)** 9.43 m   **(f)** 8.91 m

Figure 5.18 (a)~(f) The Distance Converts the Pixel (Remainder in Appendix A.1)

The LRF mounts on mobile robot, and the height is 60 cm.

The corner in LRF scan matches the suitcase corner pixel in color image.

Table 5.5 The Distance Data and The Pixel Data

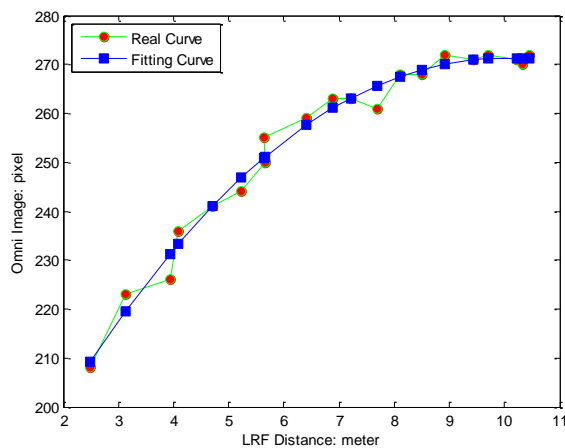| Distance (meter)	Pixel from Projection Center (Pixel)	Distance (meter)	Pixel from Projection Center (Pixel)
2.48	208	7.21	263
3.12	223	7.70	261
3.93	226	8.11	268
4.09	236	8.50	268
4.70	241	8.91	272
5.23	244	9.43	271
5.66	250	9.71	272
5.63	255	10.23	271
6.40	259	10.32	270
6.89	263	10.46	272

Figure 5.19 The Curve Fitting in Table 5.5

Using least square method to fit curve with a polynomial function with order 4

Quadratic function:

$$y = 0.0015x^4 - 0.0245x^3 - 1.0363x^2 + 22.4902x + 160.0462$$

With the relationship, the LRF scan and the color image are picked by manual

operation. The rotation data is shown in Table 5.6. And the rotation is 4.75°.

Table 5.6 The Rotation Angle with Each Feature

Item	Angle with Corner or Vertical Line			
Laser Range Finder	158.0°	117.5°	25.5°	14.5°
Omnidirectional Camera	156.5°	113.0°	19.0°	9.0°
Difference (4.75°)	2.5°	4.5°	6.5°	5.5°

5.3 Static and Dynamic Map

Section 5.3.1 shows the results in different algorithm including ICP and PSO in SLAM. And Section 5.3.2 displays the map construction results.

5.3.1 The Algorithms of Localization

When the robot moves, the localization is important for the robot. To correct the robotic position, the PSO algorithm is used to LRF scans. In the experiments, three cases, including straight path, circle path, and square path, are discussed. Figure 5.20 shows the experimental scene and the robotic path.
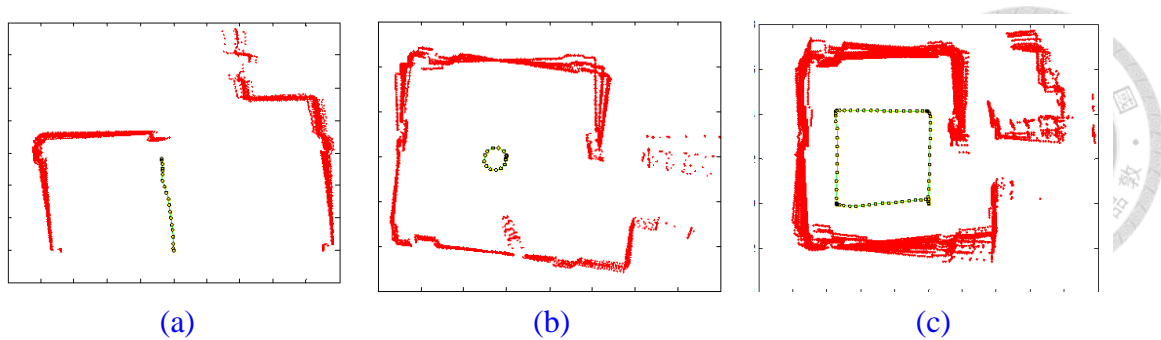
| (a) | (b) | (c) |

Figure 5.20 The Robotic Position with Encoder in Ming-Da Building 4F
with Different Path
(a) The robot walks straight path
(b) The robot walks circle path
(c) The robot walks square path

Therefore, Figure 5.21 shows the results with the odometry results, the ICP algorithm results, and the PSO algorithm results in three cases. In Figure 5.21, both the ICP algorithm and PSO algorithm possess good performance in the straight path case and the circle path. However, the wall divides two parts with the ICP algorithm in the square path, as shown in Figure 5.21(f). Then, the PSO algorithm result is shown in Figure 5.21(i). The possible reason is that the robot initial position causes the ICP to fall into the local minimum result. Then the PSO algorithm is an approximate global search. However, once the robot owns good initial position, the PSO algorithm affected by other particles causes the performance worse than the performance of ICP algorithm. In summary, the PSO algorithm owns near truth position in some cases. In Section 5.4, the location error is analyzed.

(a)　　　　　　　　　　(b)　　　　　　　　　　(c)

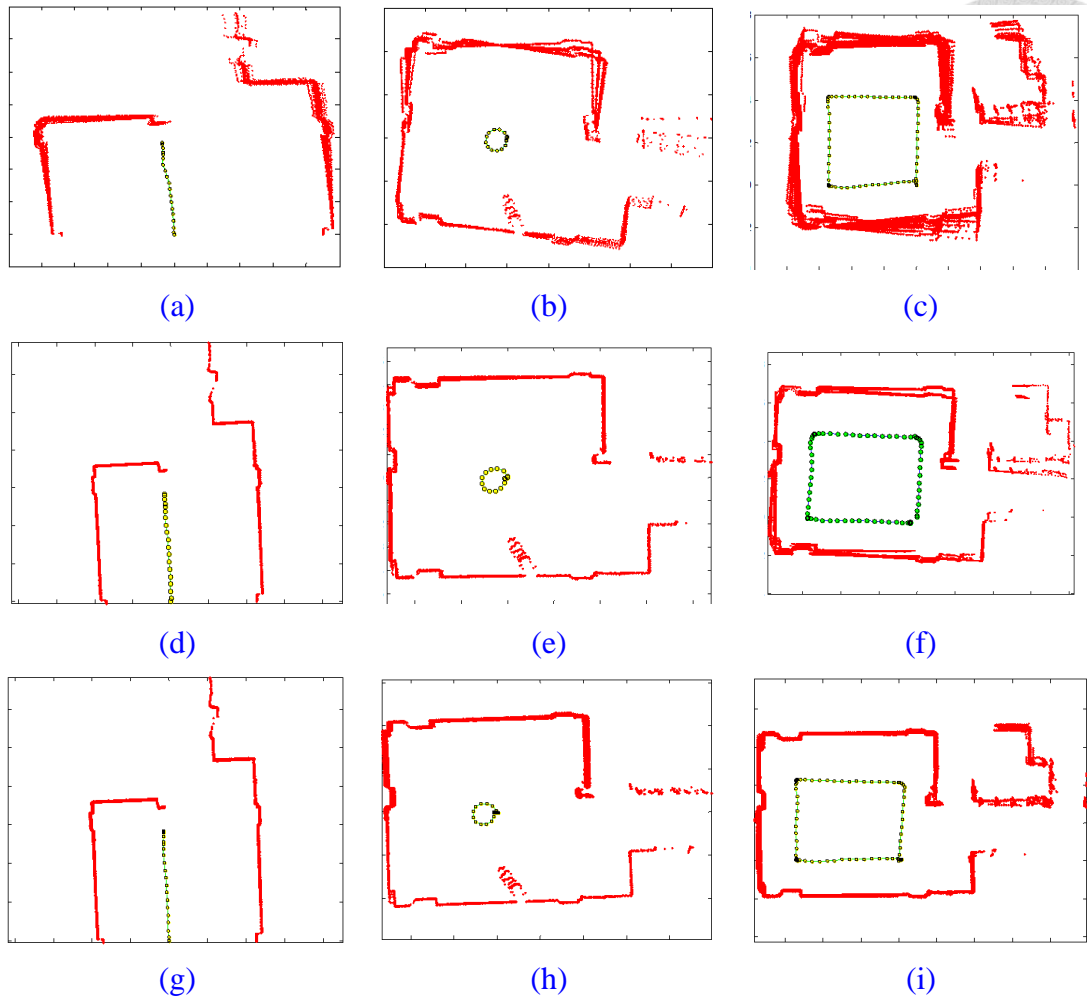(d)　　　　　　　　　　(e)　　　　　　　　　　(f)

(g)　　　　　　　　　　(h)　　　　　　　　　　(i)

Figure 5.21 The Localization with the Odometry, the ICP Algorithm, and the PSO
Algorithm in the Three Cases.

(a)~(c) The robot walks straight path, circle path, and square path with the
　　　odometry information.

(d)~(f) The robot walks straight path, circle path, and square path with the ICP
　　　algorithm. With the good initial position, the map is accuracy.

(g)~(i) The robot walks straight path, circle path, and square path with the PSO
　　　algorithm. The PSO algorithm can solve the local minimum problem.

5.3.2 The Map Construction

To make mobile robot move arbitrarily, mapping is a necessary task. The map

state updates according to the occupancy grid map robustly demonstrating the state

not interfered by noise. The map shows in Figure 5.22. The Figure 5.22(b), (d), and (f)

shows measurement points according to the raw data by PSO algorithm. Figure 5.22(a), (c), and (f) are the state of occupancy grid map from Figure 5.22(b), (d), and (f). Hence, the white is free, the black is occupied, and the gray is unknown. Figure 5.22 illustrates that the map construction is good in different conditions. First, the two walls mean the near measurement with two sides. Next, the one wall and the one empty mean the near measurement with one side and the far measurement with the other side. Finally, the two far walls mean the far measurement with two sides.
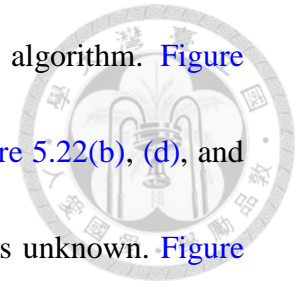
(a) Two Walls
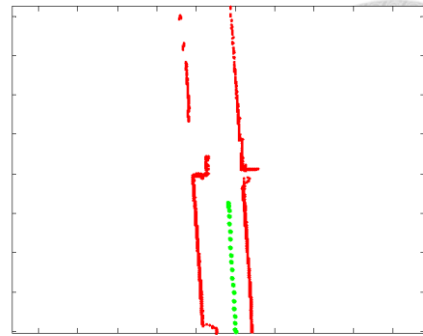
(b) Two Walls

(c) One Wall, One Empty
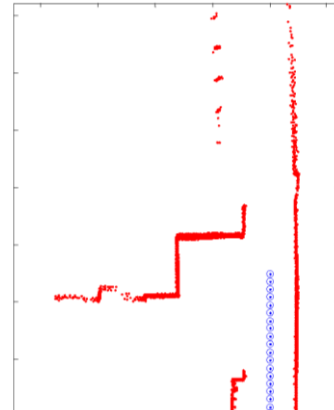
(d) One Wall, One Empty

(e) Two Far Walls
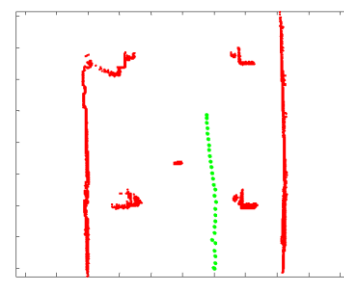
(f) Two Far Walls

Figure 5.22 The Map Construction by the Occupancy Grid Map and PSO Algorithm
(a), (c), and (e): The occupancy grid map.
White: free. Black: occupied. Gray: unknown.
(b), (d), and (f): Measurement points in a plane when the robot moves.

5.3.3 The Dynamic Map

For the local dynamic map, the inverse observation model acquires the moving objects in that time. However, the state from unknown to occupied causes false judgment in Section 3.3. The global map is shown in Figure 5.23(e). What is more,

the dynamic map is shown in Figure 5.23(a) and (c). But some noise measurements

exist owing to the inaccuracy position. Figure 5.24 shows the state from unknown to

occupancy in Ming-Da Building 5F. In Figure 5.24(e), whether the measurement

points are static objects or dynamic objects is not reasonable in the populated

environment. The dynamic maps are shown in Figure 5.24(a) and (c). In summary, the

state from unknown to occupancy needs to be judged by the color image. Therefore,

the pedestrian detection is solved the problem in Section 5.5.

(a)

(b)

(c)

(d)

(e)

(f)
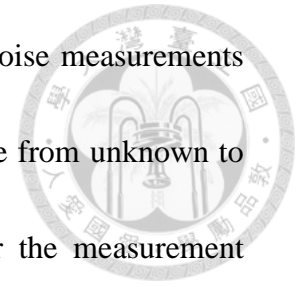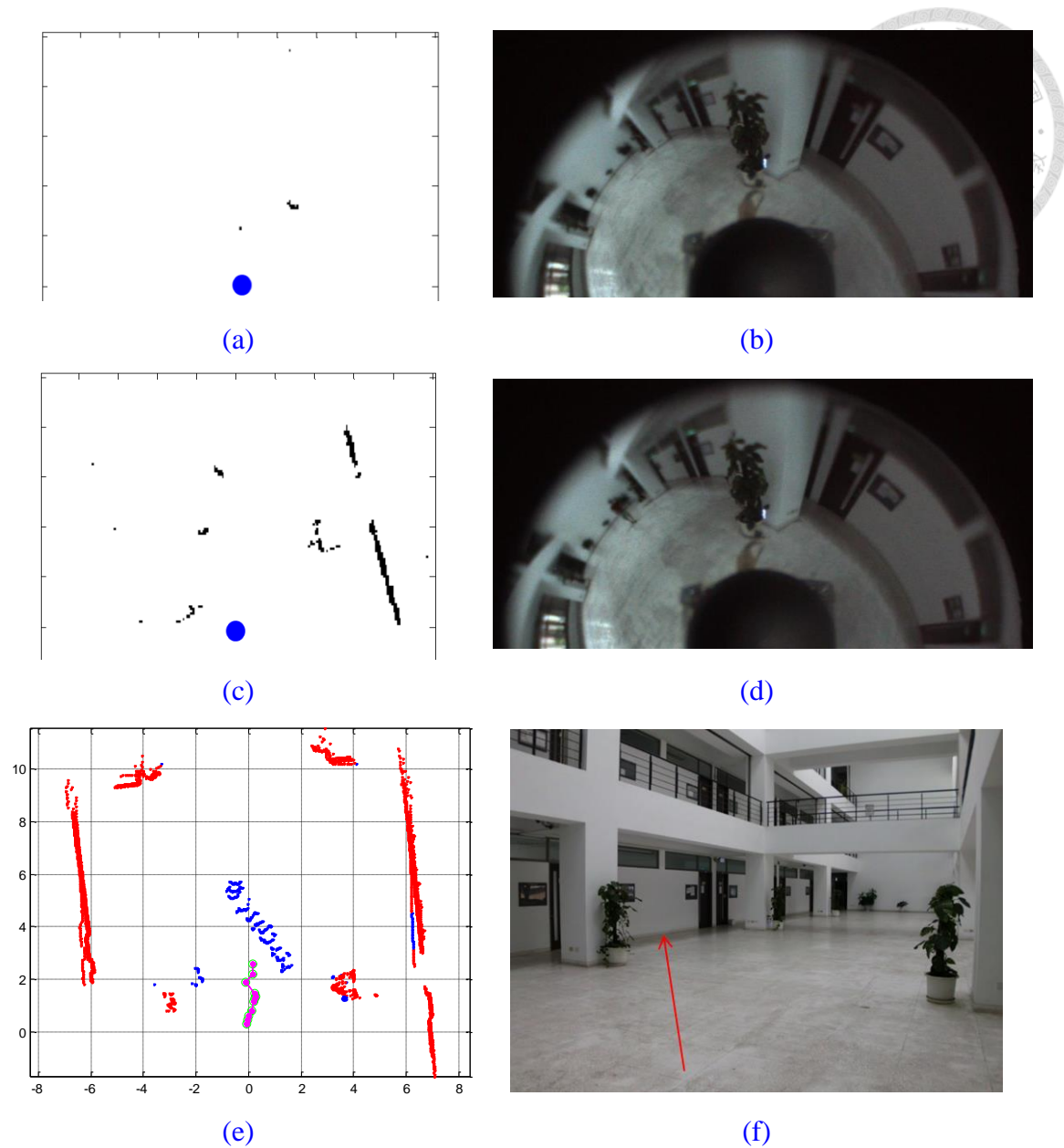
Figure 5.23 The State from Free to Occupancy in Pedestrian Walk

(a) and (c) the dynamic map are from free to occupancy, and the some noisy measurement exists. (a grid: two meter)

(b) and (d) are the omnidirectional image in (a) and (c)

(e) The global map from frame 13~23 (unit: meter)

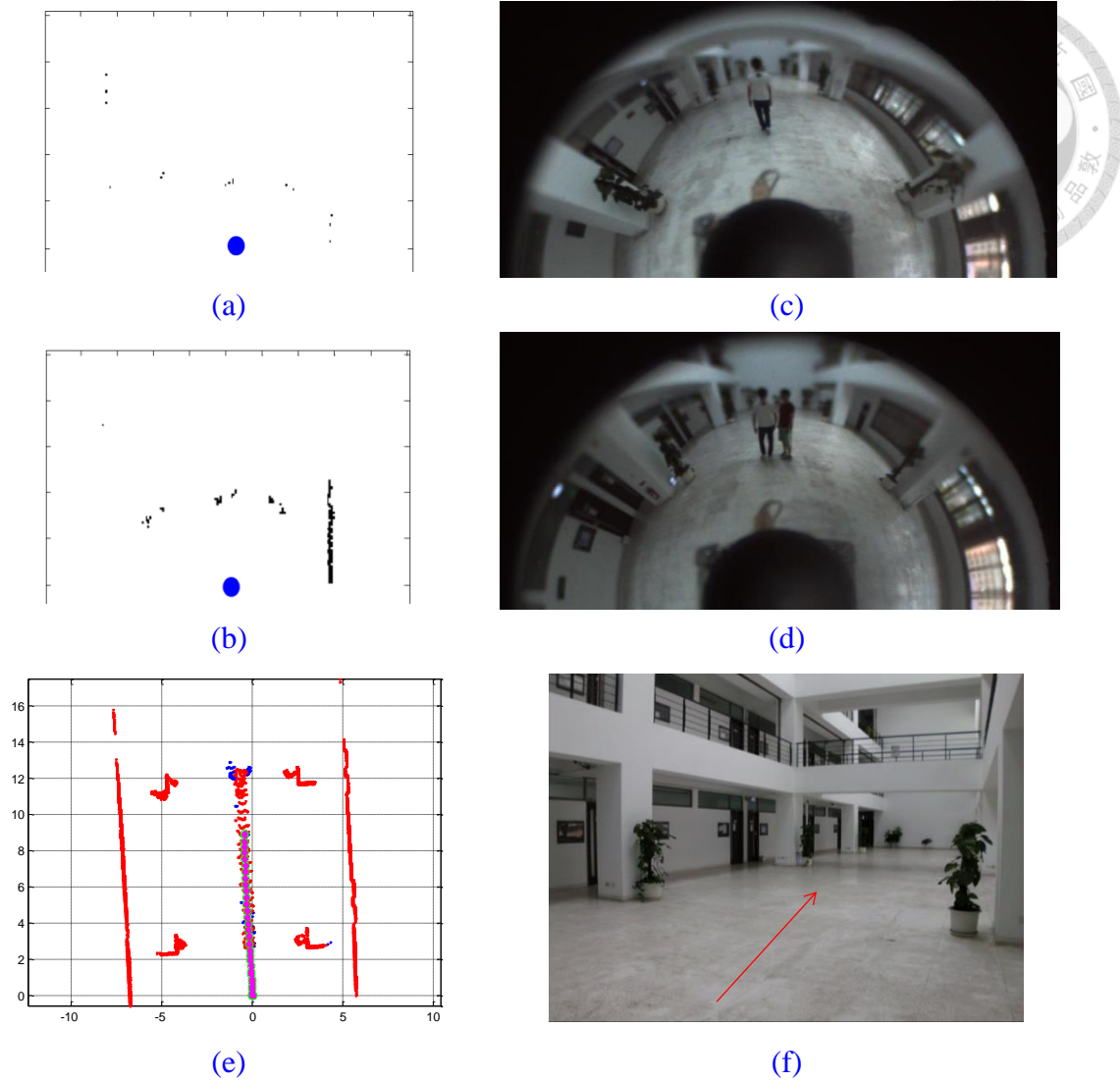(f) The Real-Scene with digital camera in Ming-Da Building 5F

Figure 5.24 The State from Unknown to Occupancy in Pedestrian Walk

(a) and (c) are the dynamic map from unknown to occupancy (a grid: two meter)

(b) and (d) are the omnidirectional image in (a) and (c)

(e) The global map from frame 1~35. However, that the pedestrian is considered as
   static objects is false. (unit: meter)

(f) The Real-Scene in Ming-Da Building 5F

5.4 Localization Accuracy

The stationary LRF URG-04LX-UG01 measures the robot position as a ground truth. The mobile robot owns the LRF LMS-100. The scene is in Ming-Da Building 4F, as shown in Figure 5.25.

LMS 100   URG 04-LX

(a)   (b)

Figure 5.25 The Experimental Scene in Ming-Da Building 4F
(a) The LRF LMS100 mounts on the Pioneer 3-DX for localization
(b) The URG 04-LX measurement as a ground truth

Since the measurement is a plane, it needs two linear independent data. In the analysis, the circle path and the square path are adopted. For ground truth robot points, the URG-04LX-UG01 measurement data center is the average of first and last point. The scheme is in Figure 5.26 [12: Ueda et al. 2011]. The error between the mapping result and the ground truth is defined as follows:

$$m_i = \sqrt{(x_i - x_{ground})^2 + (y_i - y_{ground})^2} \tag{5.5}$$

where x_i and y_i are i-th position, x_{ground} and y_{ground} are the ground truth from

Figure 5.26. Then, the standard derivation is defined as follows:

$$\sigma = \sqrt{\sum \frac{(m_i - \overline{m})}{N}} \tag{5.6}$$

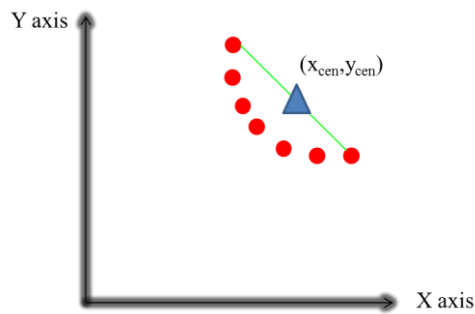where m_i is the i-th datum, $\overline{m}$ is the mean of all data, and N is the total number

of all data.

Figure 5.26 The Measurement Center of Robot Points for Ground Truth

I.   Circle Path

Figure 5.27(b) shows the circle path in Ming-Da Building 4F. Figure 5.27(a)

includes the odometry position, the ICP algorithm, the PSO algorithm, and the ground

truth position. The results of mean error and standard deviation (SD) error are shown

in Table 5.7. The position error shows in Figure 5.28 through Equation (5.5) and (5.6).
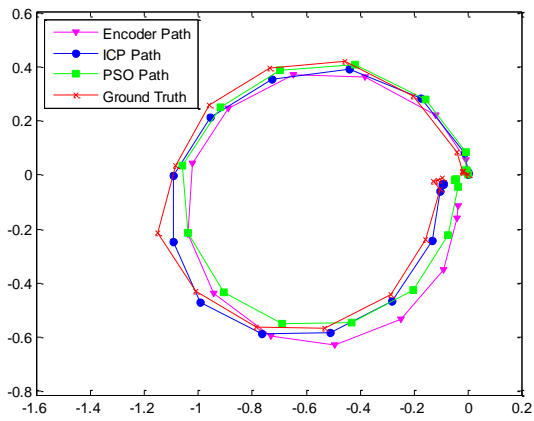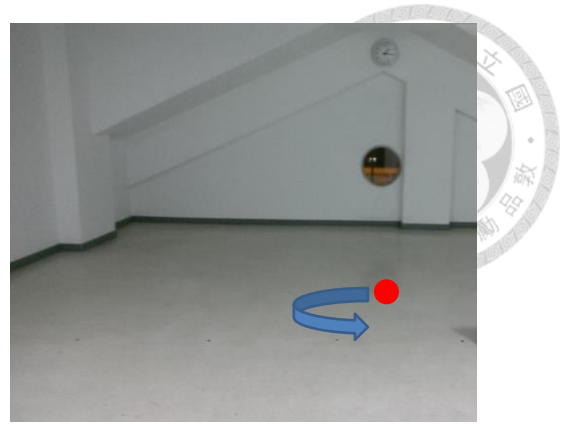
(a)	(b)

Figure 5.27 The Location of Each Results in Ming-Da Building 4F

(a) The odometry position, the ICP algorithm position, the PSO algorithm position, and the ground truth position. (unit: meter)

(b) Real scene with the path

Table 5.7 The Error of Each Algorithm with Ground Truth in Figure 5.26

Item	Encoder	ICP algorithm	PSO algorithm
Mean (meter)	0.0788	0.0288	0.0553
SD (meter)	0.0017	0.0003	0.0012

For Table 5.7, the mean error results of the ICP algorithm and the PSO algorithm are 0.0288 meter and 0.0533 meter. However, the mean error result of the encoder is 0.0788 meter. Owing to the robot position, the ICP algorithm owns the lower error than PSO algorithm. The mean error of the PSO algorithm is only 0.0533 meter.

Figure 5.28 demonstrates the distance error results. The x axis is the scan number, and the y axis is the distance error.
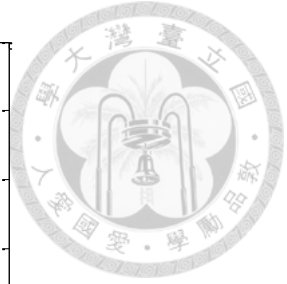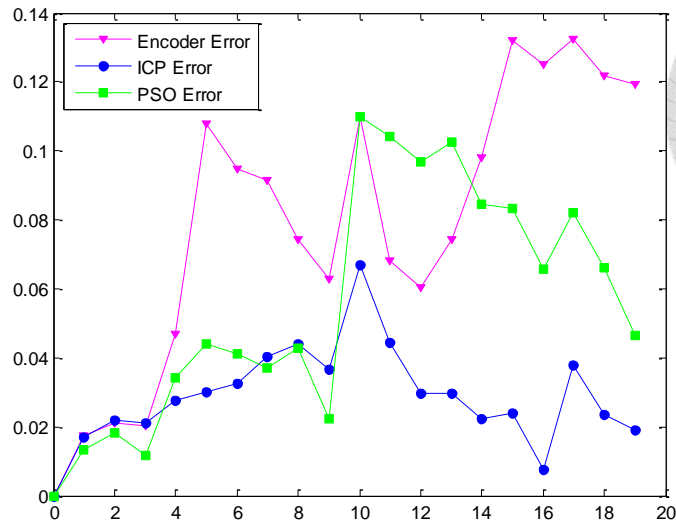
Figure 5.28 The Error Mean of each Scan Matching with Ground Truth Difference Owing to the robot position, the ICP algorithm owns the lower error than PSO algorithm.

x-axis: scan number (unit: meter)

y-axis: distance error

II.   Square Path

Figure 5.29(b) shows the square path in Ming-Da Building 4F. Figure 5.29(a) includes the odometry position, the ICP algorithm, the PSO algorithm, and the ground truth position. The results of mean error and standard deviation (SD) error are in Table 5.8. The position error shows in Figure 5.30 through Equation (5.5) and (5.6).
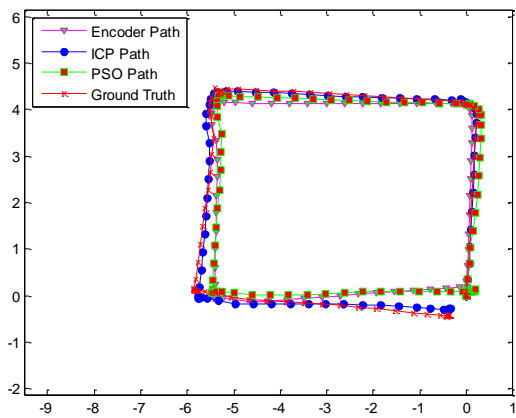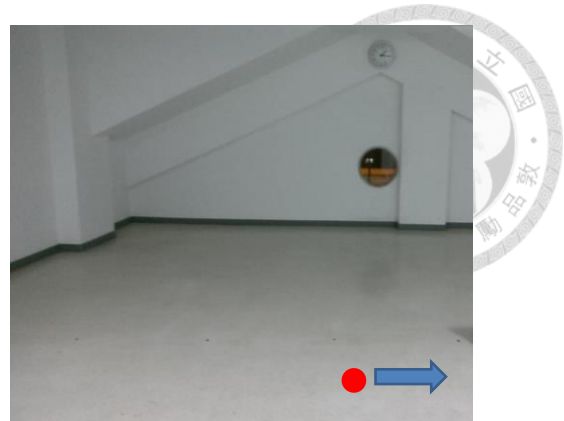
<div align="center">(a)   (b)</div>

Figure 5.29 The Location of Each Results in Ming-Da Building 4F

(a) The encoder position, the ICP algorithm position, the PSO algorithm position, and the ground truth position. (unit: meter)

(b) Real scene with the path

Table 5.8 The Error of Each Algorithm with Ground Truth in Figure 5.28

Item	Encoder	ICP algorithm	PSO algorithm
Mean (m)	0.3814	0.3652	0.1053
Standard (m)	0.0656	0.0718	0.0038

Mean error of the ICP algorithm is 0.3652 meter. However, mean error of the Encoder is 0.3814 meter. Owing to the robot position is not good. In circle path, the mean error difference of ICP algorithm and PSO algorithm is 0.0265 m. In square path, the mean error difference of ICP algorithm and PSO algorithm is 0.2599 meter. The ICP algorithm with excessive differences causes local optimal solution, as shown in Figure 5.29. So, the PSO algorithm can correct the local optimal solution owing to inaccurate robot position.
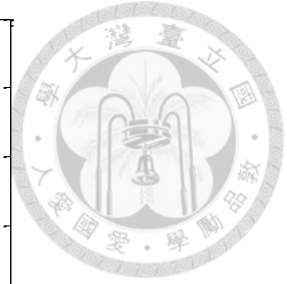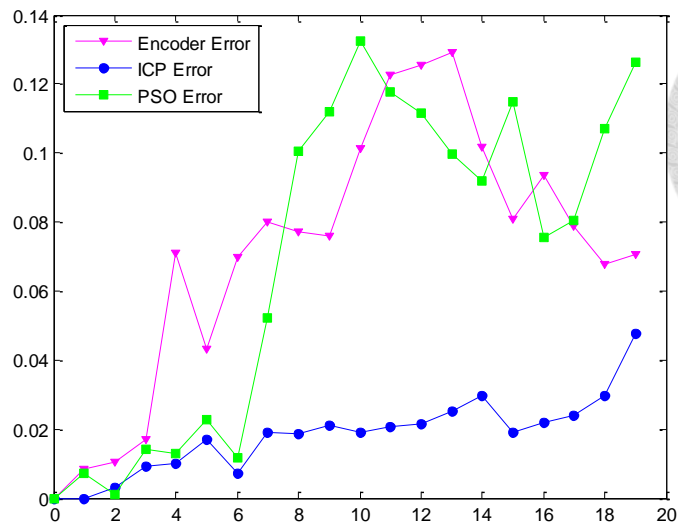
Figure 5.30 The Error of each Scan with Ground Truth

Owing to the robot position, the ICP algorithm owns the higher error than PSO

algorithm.

x-axis: scan number

y-axis: mean error (unit: meter)

5.5 Pedestrian Detection Performance

In pedestrian detection, bounding box selection is important. Section 5.5.1 states

lower line of bounding box. Then, the accuracy of the lower line of bounding box is

evaluated. Section 5.5.2 states the true or false for pedestrian detection in the color

image. Section 5.5.3 states the rate of accuracy in pedestrian detection.

5.5.1 Lower Line of Bounding Box

In color image, the lower line of bounding box needs the omnidirectional camera

model. In Figure 5.31(a), the omnidirectional camera in original image owns

96

distortion phenomenon. Therefore, the image in digital camera needs to be calibrated.

In [46: Bouguet 2013], the camera calibration toolbox uses 20 pieces of data with checkerboard to calibrate the image in Figure 5.31(a). Figure 5.31 shows the results of the image calibration. The 20 pieces of data with checkerboard are shown in Figure 5.31(b). Table 5.9 shows the camera intrinsic parameters. After calibrating the image in the digital camera, the omnidirectional camera in calibrating image is shown in Figure 5.31(d).

(a)

(b)

(c)

(d)
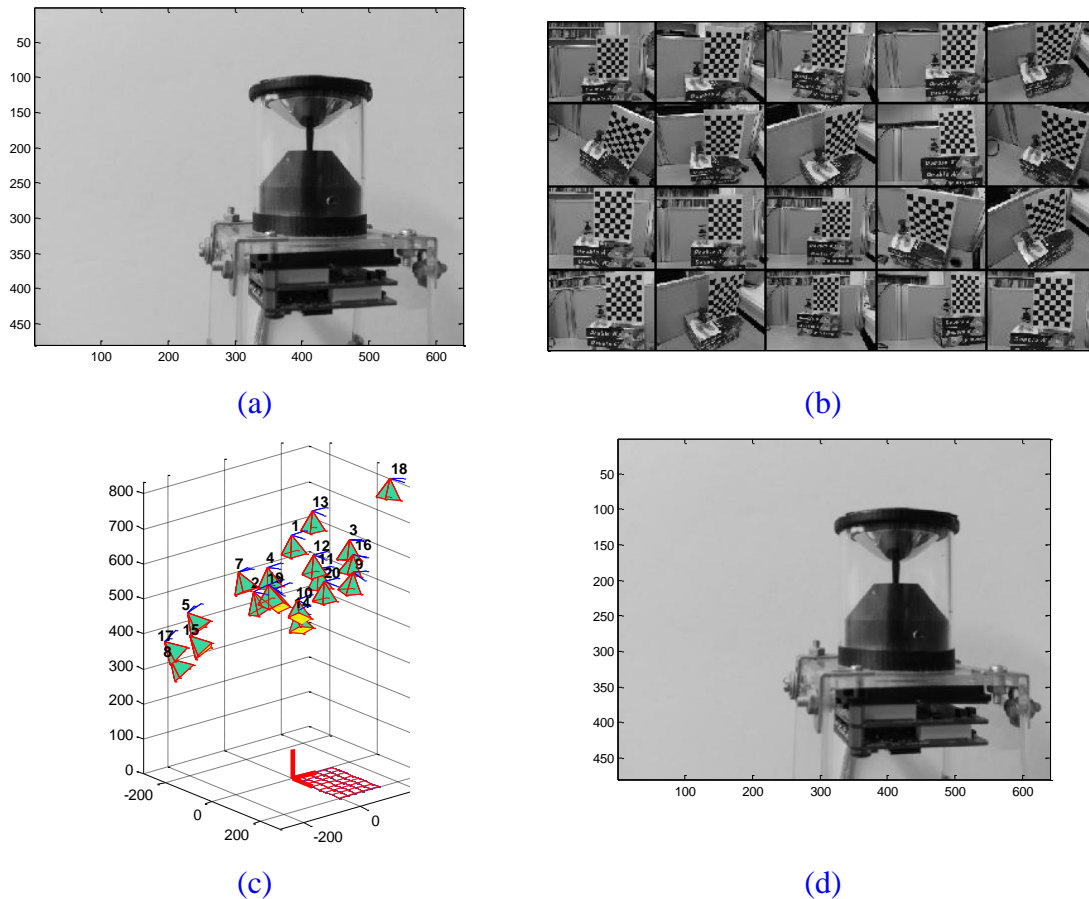
Figure 5.31 The Calibration of Omnidirectional Camera in Digital Camera Image

    (a) The omnidirectional camera in the original camera image

    (b) The different camera position with checkerboard image

    (c) The camera position on fixed checkerboard

    (d) The omnidirectional camera after calibrating camera image

Table 5.9 The Camera Intrinsic Parameter (uint: pixel)

Focal Length: fc	[688.99 688.99]
Principal point: cc	[319.50 239.50]
Skew: alpha_c	[0.00]
Distortion: kc	[0.00 0.00 0.00 0.00 0.00]

For the omnidirectional camera, the hyperbolic surface is used, as shown in

Figure 5.32. Then, the selected area in the image shows in Figure 5.32(a). Using

Canny edge detector, the edge demonstrates in Figure 5.32(b). Using SVD for least

square method, the parameters in hyperbolic surface are a, b, and c in Figure 5.32(c).

The hyperbolic surface equation and function shows as follows:

$$\frac{y^2}{a^2} - \frac{x^2}{b^2} = 1 \tag{5.7}$$

$$f(x, y) = -0.69x^2 + 0.02x.\,y + 0.72y^2 + 2.00x - 2.36y + 0.07 \tag{5.8}$$

Equation (5.7) is the hyperbolic surface equation $\frac{y^2}{a^2} - \frac{x^2}{b^2} = 1$. The function

$f(x, y) = -0.69x^2 + 0.02x.\,y + 0.72y^2 + 2.00x - 2.36y + 0.07$ is   shown   in

Equation (5.8). The parameters are obviously that a is 0.81 cm, b is 0.79 cm, and c is

1.13 cm.

(a)

(b)

$$-0.68873\ x^2 + -0.023617\ x\ y + \ldots + 0.069769 = 0$$

$$c^2 = a^2 + b^2$$
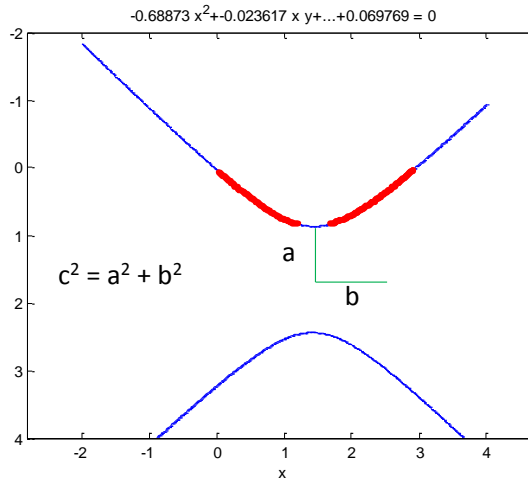
(c)

Figure 5.32 The Omnidirectional Camera Hyperbolic Surface

  (a) The selected area in the hyperbolic surface in calibration image

  (b) Canny edge detector for the image in (a)

  (c) The edge points in least square with SVD

With Equation (5.7) and (5.8), the omnidirectional camera model can be used.

From Equation (4.7) to (4.13), the lower line of bounding box can be extracted. Figure

5.33 shows the lower line of bounding box and the difference with the gray board.

Here, the horizontal axis is the percentage of error. What is more, the vertical axis is

the number of the each point in laser scan owning 361 pieces of data. Therefore, the

far the LRF distance is, the worse the error may be. Then, the error is under 10%. The

lower line of bounding box owns high accuracy. The difference between the lower

line of bounding box and gray board is shown in Table5.10. Then, Figure 5.34(a)

99

shows the mean error. Figure 5.34(b) shows the standard difference error. The mean

and the standard difference are from Equation (5.1) and (5.6). The lower line of

bounding box is shown in Figure 5.33(a) to (e). With the LRF scan in omnidirectional

camera, the pixel in omnidirectional camera can be acquired in semi-circle. Then the

error estimation is in the histogram. Figure 5.33 demonstrates the five pieces of data
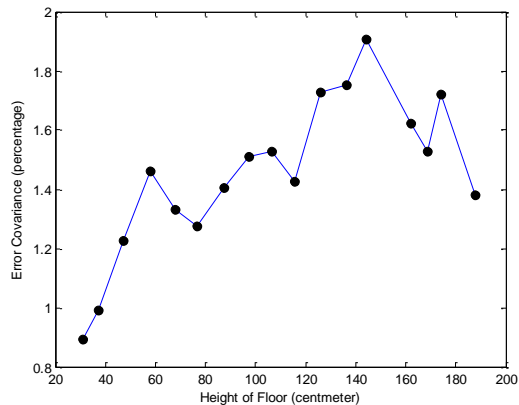
in Appendix A.2.

(a) 67.8 cm   (b) 58.0 cm   (c) 47.3 cm

(d) 37.4 cm   (e) 30.8 cm
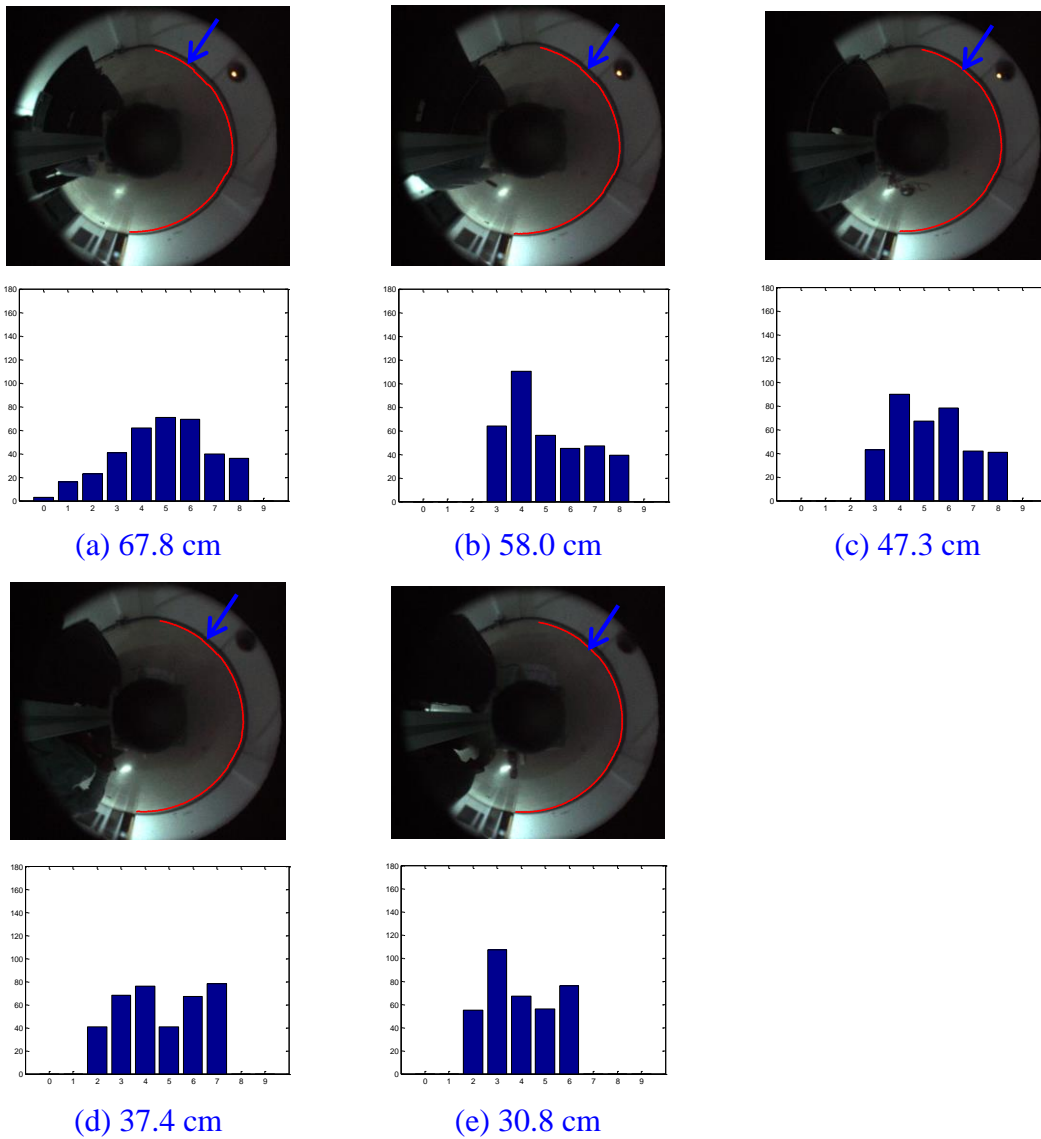
Figure 5.33 The Height Convert Pixel in Ming-Da Building 4F (Remainder in Appendix A.2)

Table 5.10 The Results in Figure A.2

Height(cm)	Mean	Standard Difference
30.8	1.85	0.89
37.4	2.35	0.99
47.3	2.80	1.23
58.0	2.72	1.46
67.8	3.21	1.33
76.6	3.27	1.28
87.2	3.36	1.41
97.4	3.73	1.51
106.3	3.93	1.53
115.6	3.96	1.43
125.8	3.42	1.73
136.4	4.20	1.75
144.2	4.90	1.91
162.2	5.05	1.62
168.8	5.30	1.53
174.2	4.82	1.72
187.9	3.98	1.38

(a)   (b)

Figure 5.34 The Mean and Standard Difference in Table 5.10

(a) Mean Error in Table 5.10.

(b) Standard Difference Error in Table 5.10
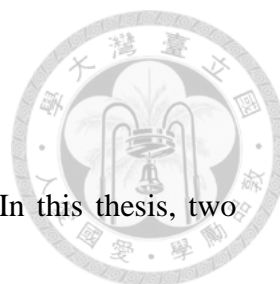
5.5.2 Pedestrian Map

In the scenes, the dynamic objects are only the pedestrians. In this thesis, two cases are discussed in Section 1.2. One is the target pedestrian with pillars environment. The other is the target pedestrian with other pedestrians abruptly appearing. Figure 5.35 shows the conditions with LRF scan in Ming-Da Building 5F. Hence, the blue points represent the pedestrian, the green points represent the mobile robot, and the red points represent the static objects.

(a)   (b)
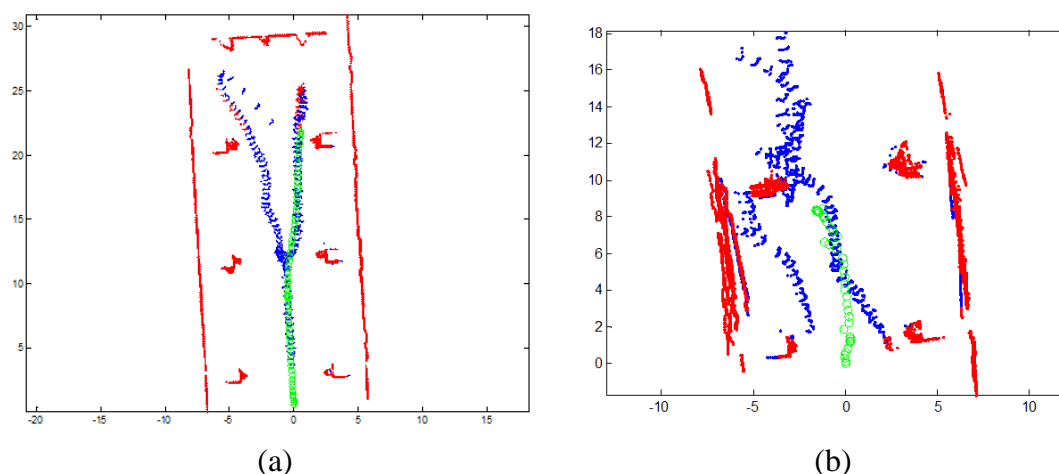
Figure 5.35 The Scene illumination with Laser Range Finder and Omnidirectional Camera (unit: meter)

(a) The target pedestrian with other pedestrians abruptly appearing
(b) The target pedestrian with pillars environment

To acquire accurate pedestrian map, the color image needs to be added in LRF scan. Figure 5.36 shows candidates in the color image. In Figure 5.36, the resolution with (a) is 350x721, (b) is 68x257, (c) is 44x257, (d) is 40x253, and (e) is 81x322. With the Hough circle transform method and color distribution method, the results are

pedestrian, as shown in Figure 5.37. Figure 5.37(a) depicts all candidates in the color image. However, Figure 5.37(b), (c), (d), and (e) show the Hough circle transform in each candidate. For the pedestrian detection, circle radius in real-life is approximately from 15 cm to 20 cm. And head body ratio is 4 to 9. Therefore, using the conditional judgment can acquire result, as shown in Figure 5.37(f).
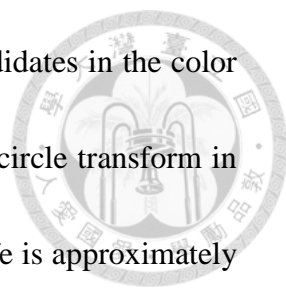
(a)
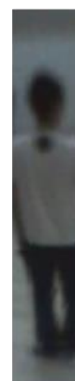
(b)   (c)   (d)   (e)

Figure 5.36 The Candidates in Dynamic Map
(a) The panorama image for down, left, right bound
(b)~(e) Each candidate in dynamic map in (a)
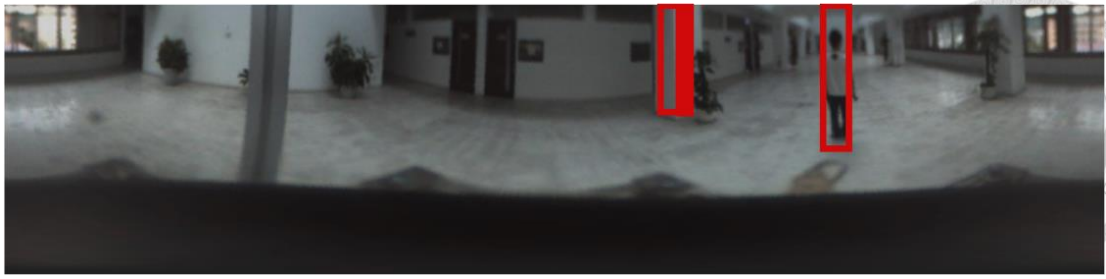
(a)

(b)   (c)   (d)   (e)

(f)

Figure 5.37 The Hough Circle Transform for Each Candidate

(a) The panorama image for lower line, left line, right line of bounding box

(b)~(e) Hough circle transform for each candidate in color image

(f) The pedestrian candidate for complete bounding box

104

5.5.3 The Detection Accuracy

In the pedestrian detection, receiver operating characteristic (ROC) is a method confirming the performance in statistic. In [54: Kay 1988], the basic definition contains true positive (TP), true negative (TN), false positive (FP), and false negative (FN). In Figure 5.38, P is a reasonable result in the condition, and N is an unreasonable result in the condition. Similarly, P' is a reasonable result in test outcome, and N' is an unreasonable result in test outcome.

Figure 5.38 The Illustration of Receiver Operating Characteristic
P is a reasonable result in the condition, and N is an unreasonable result in the condition. Similarly, P' is a reasonable result in test outcome, and N' is an unreasonable result in test outcome.

In [54: Kay 1988], accuracy (ACC) is a judgment for ROC analysis. The following cases are discussed through ACC as follows:

$$\text{ACC} = \frac{TP + TN}{P + N} \tag{5.9}$$

For the cases, the accuracy is as following:

I.   Dynamic Environment with New Pedestrian Appearing Nearby

For the new pedestrian appearing, Figure 5.35(a) shows the pedestrian and the robot path. And the detection result is shown in Table 5.11. The ACC of detection is 74%, and total frame are 126. Hence, the true frame is like Figure 5.37(e). And the false frame is like Figure 5.37(b), (c), and (d). In this case, the true detection is 110 and the false detection is 37.

Table 5.11 The ACC of detection with new pedestrian appearing nearby

ACC	Pedestrian Detection Rate
Case I	74 % (110:37)

II.   Dynamic Environment with Pillars Occlusion

For the pillars occlusion, Figure 5.35(b) shows the pedestrian and the robot path. And the detection results show in Table 5.12. The ACC of detection for Case II-1 is 58% and total frame are 148. And for Case II-2, the ACC of detection is 63% and total frame is 159. Hence, the true frame is like Figure 5.37(e). And the false frame is like Figure 5.37(b), (c), and (d). With Table 5.12, the pedestrian detection achieves 58% and 63%. In Case II-1, the true detection is 88 and the false detection is 35. In Case II-2, the true detection is 85 and the false detection is 50. The ACC is low. The caution may be the pillar occlusion without pedestrian detection. The caution causes the false increasing. And in next section, the target pedestrian tracking is discussed.
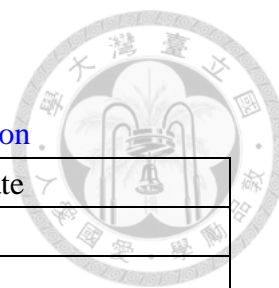
Table 5.12 The ACC of detection with Pillars Occlusion

ACC	Pedestrian Detection Rate
Case II-1	58 % (88:35)
Case II-2	63 % (85:50)

5.6 Target Pedestrian Tracking Performance

Section 5.6.1 shows results in target pedestrian tracking problem. Section 5.6.2

provides the rate of accuracy in target pedestrian tracking. The algorithms are based

on the color image. The target pedestrian tracking is discussed.

5.6.1 The Tracking Results and Accuracy

For target pedestrian tracking, laser range finder scan may cause data association

error in pillar occlusion or new pedestrian appearing. However, color image owns the

color information but does not own accuracy distance. For the color image, the color

distribution and local binary pattern (LBP) are used in target pedestrian tracking.

Section 5.6.2 discusses the accuracy.

The following two cases may cause data association error. First, Case I is that the

mobile robot follows the target pedestrian, as shown in Figure 5.35(a). When the new

pedestrian appears nearly, the laser data may match error. Then, Case II-1 is the

mobile robot follows the target pedestrian in Figure 5.35(b). However, the mobile

robot loses the target pedestrian owing to a pillar occlusion. Then both the target

pedestrian and the other pedestrian simultaneously appear. Finally, Case II-2 is also

the mobile robot follows the target pedestrian in Figure 5.35(b). The target pedestrian

and the other pedestrian are occluded by the same pillar. When the other pedestrian

walks for a long time, the target pedestrian walks in front the mobile robot. This

section discusses the problems.

The following are the results for target pedestrian tracking. Some true target

pedestrian tracking results are shown in Figure 5.39, Figure 5.40, and Figure 5.41.

(a) Frame 1   (b) Frame 16

(c) Frame 20   (d) Frame 23
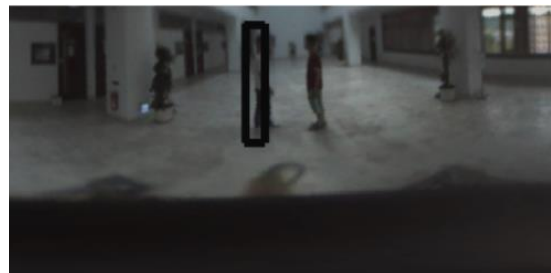
(e) Frame 24   (f) Frame 39
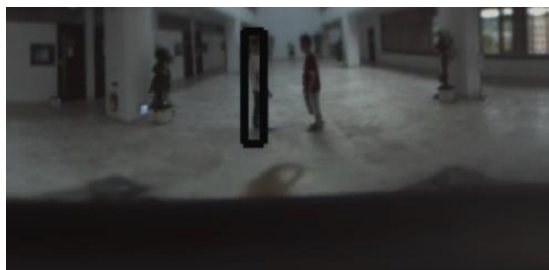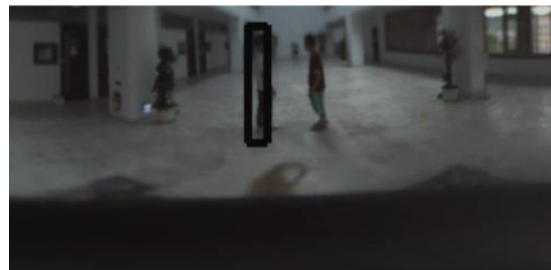
(g) Frame 43   (h) Frame 44

Figure 5.39 Tracking Target Results in Case I from Figure 5.34(a)

(a)~(h) For the target tracking results for true frame in Case I

5.6.2 The Tracking Results and Accuracy

Similarly, the accuracy of the target pedestrian tracking is shown in Table 5.13

and Table 5.14. The analysis of above cases is ACC in Equation (5.9).

I. Dynamic Environment with New Pedestrian Appearing Nearby

For the new pedestrian appearing, Figure 5.35(a) shows the pedestrian and the robot path. And the target pedestrian tracking result is shown in Table 5.13. The ACC of target pedestrian tracking is 73%, and total frame are 126. Hence, the true frames are like Figure 5.39(a) to (e). In this case, the true target pedestrian tracking is 84 and the false target pedestrian tracking is 31.

Table 5.13 The ACC of Tracking with New Pedestrian Appearing Nearby

ACC	Target Pedestrian Tracking Rate
Case I	73 % (84:31)

II. Dynamic Environment with Pillars Occlusion

For the pillars occlusion, Figure 5.35(b) shows the pedestrian and the robot path. And the target pedestrian tracking result is shown in Table 5.14. The ACC of target pedestrian tracking for Case II-1 is 48% and total frame are 148. The true frames are in Figure 5.40(a) to (c). And for Case II-2, the ACC of target pedestrian tracking is 53% and total frame is 159. Hence, the true frames are like Figure 5.41(a) and (b). In Case II-1, the true target pedestrian tracking is 29 and the false target pedestrian tracking is 32. In Case II-2, the true target pedestrian tracking is 41 and the false target pedestrian tracking is 36. In summary, the ACC is low. The caution may be the pillar occlusion without target pedestrian tracking. The caution causes the false

111

increasing.
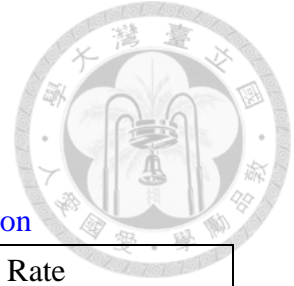
Table 5.14 The ACC of Tracking with Pillars Occlusion

ACC	Target Pedestrian Tracking Rate
Case II-1	48 % (29:32)
Case II-2	53 % (41:36)

Chapter 6
Conclusions and Future Works

In this chapter, Section 6.1 presents the conclusions in the research. And Section 6.2 states the future works in the future.

6.1 Conclusions

In this thesis, the main issue is pedestrian tracking with a mobile robot mounted on the laser range finder and the omnidirectional camera. The LRF operation principle and the LRF limitation are presented.

For simultaneous localization and mapping, the robot position and the map construction are two important tasks. The particle swarm optimization (PSO) is used to robot position inaccuracy. Therefore, the PSO algorithm can correct the odometry in local minimum solution. Once odometry of mobile robot is known, building a map is also a task which can be effectively solved at the same time [39: Birk & Carpin 2006]. For the occupancy grid map, the Bayesian rule in the occupancy grid map is according to the previous estimation and current measurement. Next, the inverse observation model divides into the static objects and dynamic objects based on the global accuracy map. With the information, the robot preliminary judges possible

pedestrians and avoids the obstacles.
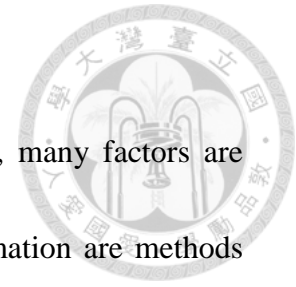
For the detection pedestrians and tracking target pedestrian, many factors are considered. The color distribution and the Hough circle transformation are methods for the pedestrian detection. By the experimental tests, the methods are reasonable. For the target pedestrian tracking, the color distribution and the local binary pattern (LBP) are used to the data association error in the real scene. Through the experiments, the idea is practical. However, the resolution of pedestrian in color image is low owing to the omnidirectional camera. In the same color image, the FOV is complementary to the resolution.

6.2 Future Works

For the future works, the vertical edge can make the accumulation error of robot position eliminate in simultaneous localization and mapping in [13: Bacca et al. 2013]. In [39: Birk & Carpin 2006], the multiple robots can communicate each other to construct the map increasing exploration rate. The pedestrian detection and the target pedestrian tracking can use more feature. Take for example, the shape of pedestrian [35: Dalal & Triggs 2005], the database construction in specific pedestrian [40: Wang et al. 2011], or combining the skeleton [28: Lin & Huang 2011] is a research direction. The 3D distance sensor such as Kinect, stereo camera, or 3D LRF is an available

operation for the indoor robot. To acquire the resolution in color image, the digital camera is used for pedestrian detection and target pedestrian tracking.
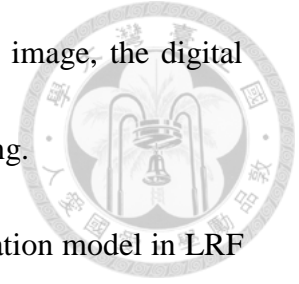
In summary, the moving object detection with inverse observation model in LRF can be a preprocessing for pedestrian detection and target pedestrian tracking. With the inverse observation model for LRF scan, the performance of the pedestrian detection and the target pedestrian tracking should increase. In this thesis, the methods provide an idea for office assistant robot. In the future, the office assistant robot is widely used.

References

[1: Seifert & Kay 1995]

R. W. Seifert and M.G. Kay, "Evaluation of AGV routing strategies using hierarchical simulation," in Proceedings of Winter Simulation Conference, Arlington, USA, pp. 850-856, Dec. 3-6, 1995.

[2: Nishimura et al. 2007]

S. Nishimura, H. Takemura, and H. Mizoguchi, "Development of attachable modules for robotizing daily items -person following shoppingcart robot-," in Proceedings of IEEE Robotics and Biomimetics, Sanya, China, pp. 1506-1511, Dec. 15-18, 2007.

[3: Chen et al. 2011]

C. L. Chen, C. C. Chou, and Feng-Li Lian, "Trajectory Planning for Human Host Tracking and Following of Slave Mobile Robot on Service-Related Tasks," in Proceedings of IEEE International Conference on Robotics and Biomimetics, Phuket, Thailand, pp. 2419-2420, Dec. 7-11, 2011.

[4: Okubo et al. 2009]

Y. Okubo, C.Ye, and J. Borenstein, "Characterization of the Hokuyo URG-04LX Laser Rangefinder for mobile Robot Obstacle Negotiation," in Proceedings of SPIE Conference Unmanned Robotic and Layered Systems, Orlando, USA, Apr. 30, 2009.

[5: Yagi et al. 2005]

Y. Yagi, K. Imai, K. Tsuji, and M. Yachida, "Iconic Memory-Based Omnidirectional Route Panorama Navigation," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27, No. 1, pp. 78-87, Jan. 2005.

[6: Gonzalez & Woods 2008]

R. C. Gonzalez and R. E. Woods, "Digital Image Processing," 3rd adapted ed., Editor: S. G. Miaou, Taiwan: Pearson, Jun. 2008.

[7: Jia et al. 2010]

S. Jia, H. Yang, X. Li, and W. Fu, "LRF-Based Data Processing Algorithm for Map Building of Mobile Robot," in Proceedings of IEEE International Conference on Information and Automation, Harbin, China, pp. 1924 - 1929, Jun. 20-23, 2010.

[8: Rebai et al. 2009]

K. Rebai, A. Benabderrahmane, O. Azouaoui, and N. Ouadah, "Moving Obstacles Detection and Tracking with Laser Range Finder," in Proceedings of International Conference on Advanced Robotics, pp. 1-6, Jun. 22-26 2009.

[9: Chung et al. 2012]

W. Chung, H. Kim, Y. Yoo, C. B. Moon, and J. Park, "The detection and following of human legs through inductive approaches for a mobile robot with a single laser range finder," IEEE Transactions Industrial Electronics, Vol. 59, No. 8, pp. 3156–3166, Aug. 2012.

[10: Gander et al. 1994]

W. Gander, G.H. Golub, and R. Strebel, "Least-Square Fitting of Circles and Ellipses," BIT Numerical Mathematics 34, No. 43, pp. 558-578, Dec. 1, 1994.

[11: Grassi & Okamoto 2006]

V. Grassi Jr., and J. Okamoto Jr., "Development of an omnidirectional vision system," Journal of the Brazilian Society of Mechanical Sciences and Engineering, Vol. 28, No.1, pp. 58-68, Jan.-Mar. 2006.

[12: Ueda et al. 2011]

H. Ueda, J. H. Lee, S. Okamoto, B. J. Yi, and S. Yuta, "People tracking method for a mobile robot with laser scanner and omni directional camera," in Proceedings of international Conference on Ubiquitous Robots and Ambient Intelligence, Incheon, Korea, pp. 503–507, Nov. 23-26, 2011.

[13: Bacca et al. 2013]

B. Bacca, X. Cufí, and J. Salví, "Vertical edge-based mapping using range-augmented omnidirectional vision sensor," IET Computer Vision, Vol. 7, No. 2, pp. 135-143, Apr. 2013.

[14: Scaramuzza et al. 2006]

D. Scaramuzza, A. Martinelli, and R. Siegwart, "A Flexible Technique for Accurate Omnidirectional Camera Calibration and Structure from Motion," in Proceedings of IEEE International Conference of Vision Systems, New York, USA, pp. 45-52, Jan. 5-7, 2006.

[15: Wu et al. 2013]

C. H. Wu, Y. H. Chen, Y. Y. Lee, and C. H. Tsai, "A Fast Genetic SLAM Approach for Mobile Robots," in Proceedings of 14th ACIS International Conference on Artificial Intelligence Software Engineering, Networking and Parallel/Distributed Computing, Honolulu, USA, pp. 563-568, Jul. 1-3, 2013.

[16: Rusdinar et al. 2010]

A. Rusdinar, J. Kim, and S. Kim "Error Pose Correction of Mobile Robot for SLAM Problem using Laser Range Finder Based on Particle Filter," in Proceedings of International Conference on Control, Automation and Systems, Gyeonggi-do, Korea, pp. 52–55, Oct. 27-30, 2010.

[17: Zhang 1994]

Z. Zhang, "Iterative Point Matching for Registration of Free-Form Curves and Surfaces," International Journal of Computer Vision, Vol. 13, No. 2, pp. 119-152,

1994.

[18: Lu & Milios 1994]

F. Lu and E.E. Milios, "Robot Pose Estimation in Unknown Environments by Matching 2D range scans," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Seattle, USA, pp. 935-938, Jun. 21-23, 1994.

[19: Tong & Barfoot 2011]

C. H. Tong and T. D. Barfoot, "Batch Heterogeneous Outlier Rejection for Feature-Poor SLAM," in Proceedings of IEEE International Conference on Robotics and Automation, Shanghai, China, pp. 2630-2637, May 9-13, 2011.

[20: Kang et al. 2010]

J. G. Kang, W. S. Choi, S. Y. An, and S. Y. Oh, "Augmented EKF based SLAM method for Improving the Accuracy of the Feature Map," in Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, Taipei, Taiwan, pp. 3725-3731, October 18-22, 2010.

[21: Moravec & Elfes 1985]

H. Moravec and A. Elfes, "High Resolution Maps from Wide Angle Sonar," in Proceedings of IEEE International Conference on Robotics and Automation, Missouri, USA, pp. 116-121, Mar. 1985.

[22: Wolf & Sukhatme 2004]

D. Wolf and G. Sukhatme. "Online Simultaneous Localization and Mapping in Dynamic Environments," in Proceedings of IEEE International Conference on Robotics and Automation, New Orleans, pp. 1301-1307, LA, USA, Apr. 26 – May 1, 2004.

[23: Thrun et al. 2005]

S. Thrun, W. Burgard, and D. Fox, "Probabilistic Robotics," Editor: R. Arkin, London: The MIT Press, 2005.

[24: Chang & Lian 2012]

F. M. Chang and F. L. Lian, "Polar Grid Based Robust Pedestrian Tracking with Indoor Mobile Robot using Multiple Hypothesis Tracking Algorithm," in Proceedings of SICE Annual Conference, pp. 1558-1563, Akita, Japan, Aug. 20-23, 2012.

[25: Cox & Hingorani 1996]

l. J. Cox and S. L. Hingorani, "An efficient Implementation of Reid's Multiple Hypothesis Tracking Algorithm and Its Evaluation for the Purpose of Visual Tracking," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.18, No.2, pp. 138-150, Feb. 1996

[26: Sung & Chung 2011]

Y. Sung and W. Chung, "Human Tracking of a Mobile Robot with an onboard LRF (Laser Range Finder) using Human Walking Motion Analysis," in Proceedings of International Conference on Ubiquitous Robots and Ambient Intelligence, Incheon, Korea, pp. 366-370, Nov. 23-26, 2011.

[27: Carballo et al. 2010]

A. Carballo, A. Ohya and S. Yuta "People Detection using Range and Intensity Data from Multi-Layered Laser Range Finders," in Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, Taipei, Taiwan, pp. 5849-5854, Oct. 18-22, 2010.

[28: Lin & Huang 2011]

D. T. Lin, and K. Y. Huang, "Collaborative Pedestrian Tracking and Data Fusion With Multiple Cameras," IEEE Transactions on Information Forensics and Security, Vol. 6, No. 4, pp.1432-1444, Dec. 2011.

[29: Stauffer & Grimson 1999]

C. Stauffer and W.E.L Grimson, "Adaptive background mixture models for real-time tracking," in Proceedings of IEEE Conference Computer Vision and Pattern Recognition, Fort Collins, USA, Vol. 2, pp. 246–252, Jun. 23-25, 1999.

[30: Lee et al. 2003]

D. S. Lee, J. J. Hull, and B. Erol, "A Bayesian Framework for Gaussian Mixture Background Modeling," in Proceedings of IEEE International Conference on Image Processing, Vol. 3, pp. 973-976, Sep. 14-17, 2003.

[31: Enzweiler et al. 2008]

M. Enzweiler, P. Kanter, and D. M. Gavrila, "Monocular Pedestrian Recognition Using Motion Parallax," in Proceedings of IEEE Intelligent Vehicles Symposium, Eindhoven, Holland, pp. 792-797, Jun. 4-6, 2008.

[32: Leithy et al. 2010]

A. Leithy, Mohamed N. Moustafa, and Ayman Wahba, "Cascade of Complementary Features for Fast and Accurate Pedestrian Detection," in Proceedings of the 4th Pacific-Rim Symposium on Image and Video Technology, Singapore, pp. 343-348, Nov. 14-17, 2010.

[33: Zhao et al. 2008]

M. Zhao, D. Sun, and H. He, "Hair-color Modeling and Head Detection," in Proceedings of the 7th World Congress on Intelligent Control and Automation, Chongqing, China, Jun. 25-27, 2008.

[34: Xu & Xu 2013]

F. Xu, and F. Xu, "Pedestrian Detection Based on Motion Compensation and HOG/SVM Classifier," in Proceedings of the Fifth International Conference on Intelligent Human-Machine Systems and Cybernetics, Hangzhou, China, pp.

334-337, Aug. 26-27, 2013.

[35: Dalal & Triggs 2005]

N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," in Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, Vol. 1, pp. 886-893, June 25, 2005.

[36: Ballard 1981]

D. Ballard. "Generalizing the Hough Transform to Detect Arbitrary Shapes," IEEE Transactions Transactions on Pattern Analysis and Machine Intelligence, Vol. 13, No. 2, pp. 111–122, 1981.

[37: Kristou et al. 2011]

M. Kristou, A. Ohya, and S. Yuta, "Target person identification and following based on omnidirectional camera and LRF data fusion," in Proceedings of IEEE International Symposium on Robot and Human Interactive Communication, Atlanta, GA, USA, pp. 419-424, July 31 - August 3, 2011.

[38: Li et al. 2011]

H. Li, T. Shen, and X. Huang, "Approximately Global Optimization for Robust Alignment of Generalized Shapes," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 33, No. 6, pp. 1116-1131, Jun. 2011.

[39: Birk & Carpin 2006]

A. Birk and S. Carpin "Merging Occupancy Grid Maps From Multiple Robots," in Proceedings of the IEEE Transactions, Vol. 94, No. 7, pp.1384-1397, July 2006.

[40: Wang et al. 2011]

M. Wang, H. Qiao, and B. Zhang, "A New Algorithm for Robust Pedestrian Tracking Based on Manifold Learning and Feature Selection," IEEE Transactions on Intelligent Transportation Systems, Vol. 12, No. 4, pp.1195-1208, Dec. 2011.

[41: Kun et al. 2012]

Z. Kun, S. Fengchi, Y. Jing, "An Autonomous Target-Tracking Algorithm Based on Visual Feature," in Proceedings of the 31st Chinese Control Conference, Hefei, China, pp. 4936-4941, July 25-27, 2012.

[42: Ishikawa et al. 2009]

T. Ishikawa, M. Kourogi, T. Okuma, and T. Kurata, "Economic and Synergistic Pedestrian Tracking System for Indoor Environments," in Proceedings of International Conference of Soft Computing and Pattern Recognition, Malacca, Malaysia, pp. 522-527, Dec. 4-7, 2009.

[43: Besl & McKay 1992]

P. J. Besl and N. D. McKay, "A method for registration of 3d shape," IEEE

Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, No. 2, pp. 239-256, Feb. 1992.

[44: Kennedy & Eberhart 1995]

J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in Proceedings of IEEE International Conference on Neural Networks, Perth, Australia, Vol. 4, pp. 1942-1948, Nov. 27 - Dec. 1, 1995.

[45: Eberhart & Shi 1998]

R. C. Eberhart and Y. Shi "Comparison between genetic algorithms and particle swarm optimization," in Proceedings of Conference on Evolutionary Programming, San Diego, USA, Vol. 1447, pp. 611-616, Mar. 25–27, 1998.

[46: Bouguet 2013]

J. Y. Bouguet.(2013, Dec. 2). In Camera Calibration Toolbox for Matlab. Retrieved Jul. 8 2014, from http://www.vision.caltech.edu/bouguetj/calib_doc/

[47: O'Leary & Murray 2004]

P. O'Leary, and P. Z. Murray, "Direct and Specific Least- Square Fitting of Hyperbolae and Ellipses," Journal of Electronic Imaging, Vol. 13, No. 3, pp. 492–503, Jul. 2004.

[48: Zhao et al. 2012]

M. Zhao, D. H. Sun, Y. Tang, and H. P. He1, "Head Detection Based on 21HT and Circle Existence Model," in Proceedings of the 10th World Congress on Intelligent Control and Automation, Beijing, China, pp. 4875-4880, Jul. 6-8, 2012.

[49: Rahimi et al. 2013]

S. Rahimi, A. Aghagolzadeh, and H. Seyedarabi, "Three camera-based human tracking using weighted color and cellular LBP histograms in a particle filter framework," in Proceedings of Iranian Conference on Electrical Engineering, Mashhad, pp. 1-6, May 14-16, 2013

[50: YCbCr from wiki 2014]

YCbCr. (2014, February 8). In Wikipedia. Retrieved July 8 2014, from http://en.wikipedia.org/wiki/YCbCr

[51: Bhattacharyya distance from wiki 2014]

Bhattacharyya Distance. (2014, May 31). In Wikipedia. Retrieved July 8 2014, from http://en.wikipedia.org/wiki/Bhattacharyya_distance

[52: Vu et al. 2007]

T. D. Vu, O. Aycard, and N. Appenrodt, "Online Localization and Mapping with Moving Object Tracking in Dynamic Outdoor Environments," in Proceedings of IEEE Intelligent Vehicles Symposium, Istanbul, Turkey, pp. 190-195, Jun. 13-15, 2007.

[53: Yang & Lian 2012]

J. Y. Yang and F. L. Lian, "Omnidirectional Vision-Based Robot Localization using Vertical Line Matching and Detection of Vanishing Point and Floor Region
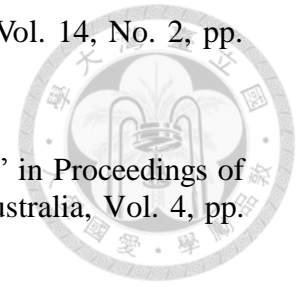
using Edge Orientation Information," 1st ed., Taipei, Taiwan, Editor: NTU, 2012, pp. 47–48.

[54: Kay 1988]

S. M. Kay, "Fundamentals of Statistical Signal Processing," Vol. 2: Detection Theory, US: Prentice Hall, 1998.

[55: Winner et al. 2012]

R. Grewe, M. Komar, A. Hohm, S. Lueke, and H. Winner "Evaluation Method and Results for the Accuracy of an Automotive Occupancy Grid," in Proceedings of IEEE International Conference on Vehicular Electronics and Safety, Turkey, Istanbul, pp. 19-24, July 24-27, 2012.
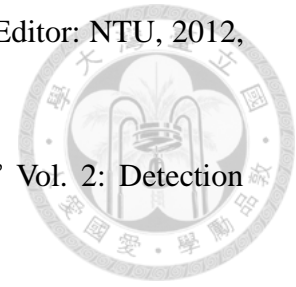
Appendix

Owing to the numerous data, appendix shows the complete data. Appendix A.1 presents distance convert pixel in 20 pieces of image data. Appendix A.2 provides the accuracy of lower line of bounding box extraction.

A.1 Distance Convert Pixel

To pick the specific pixel matching the LRF scan, the distance and pixel relationship needs to be acquired. In Figure A.1 and Table A.1 with the distance and the pixel relationship [12: Ueda et al. 2011], the LRF is mounted the same height 60 cm of the suitcase. In Figure A.1, the corner in LRF scan matches the suitcase corner pixel in color image. The distance from LRF scan and pixel from color image relationship is shown in Table A.1.

(a) 10.46 m   (b) 10.32 m   (c) 10.23 m
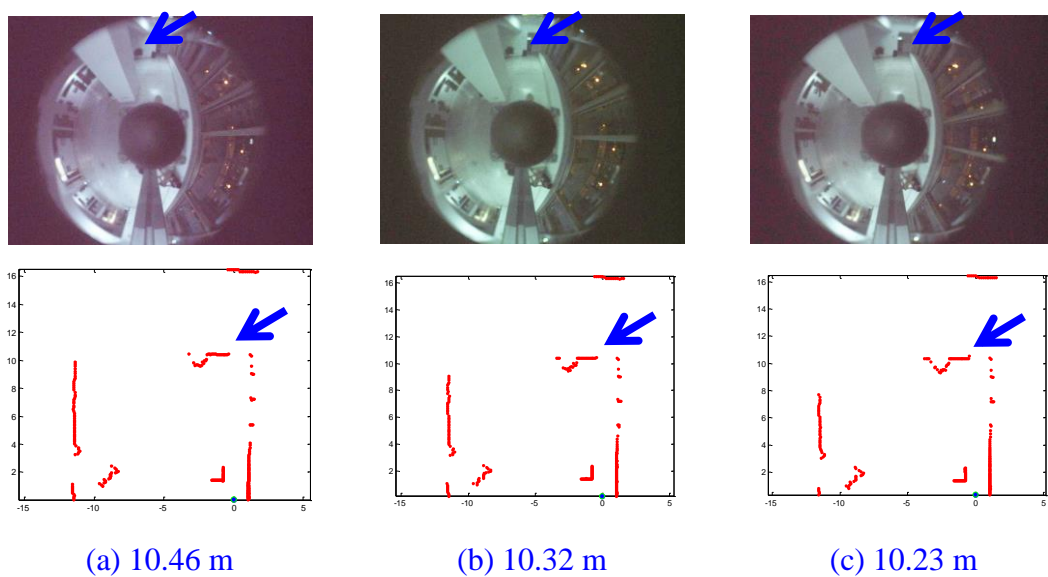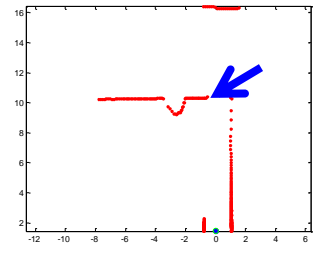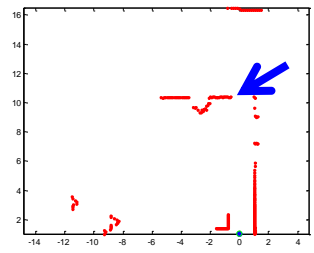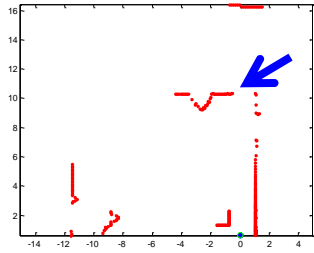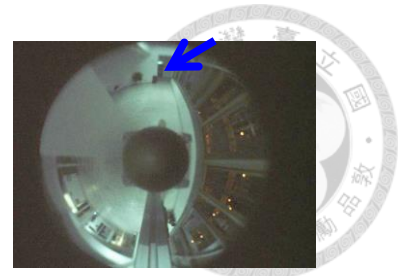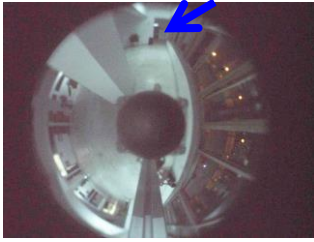
(d) 9.71 m   (e) 9.43 m   (f) 8.91 m
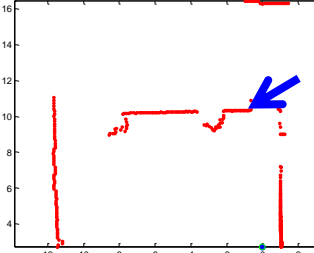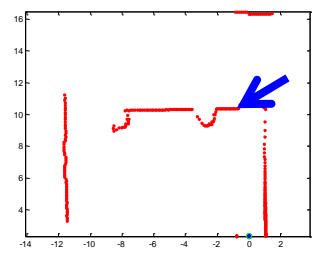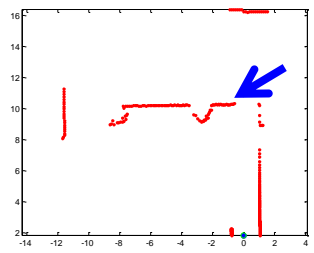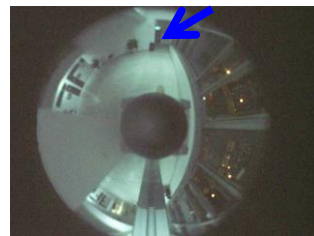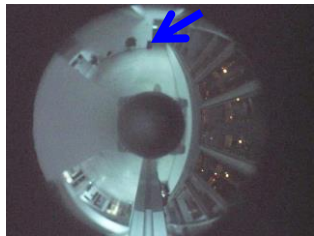
(g) 8.50 m   (h) 8.11 m   (i) 7.70 m

(j) 7.21 m   (k) 6.89 m   (l) 6.40 m

(m) 5.63 m   (n) 5.66 m   (o) 5.23 m

(p) 4.70 m   (q) 4.09 m   (r) 3.93 m
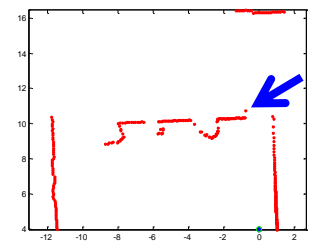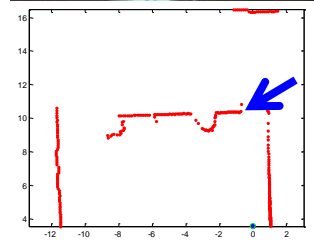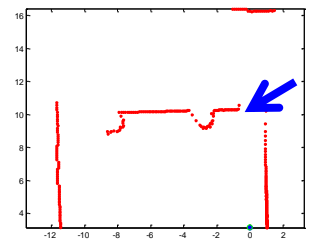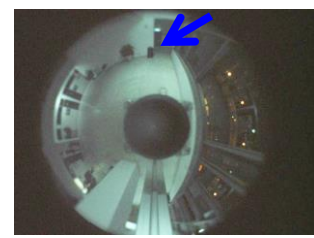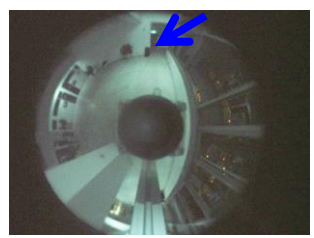
(s) 3.12 m   (t) 2.48 m
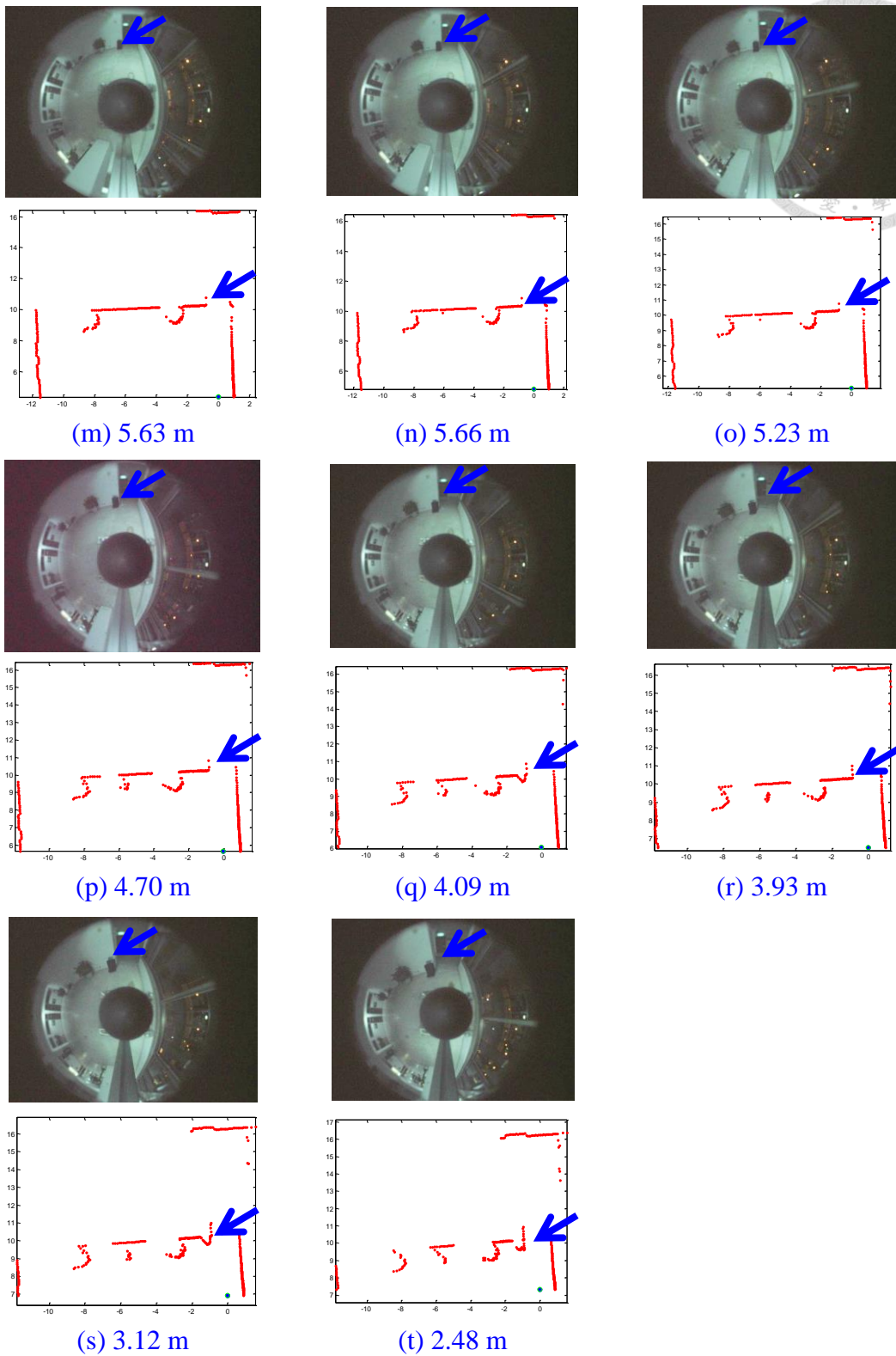
Figure A.1 (a)~(t) The Distance Converts the Pixel (From Figure 5.18)

The LRF mounts on mobile robot, and the height is 60 cm.

The corner in LRF scan matches the suitcase corner pixel in color image

Table A.1 The Distance Data and The Pixel Data in Figure A.1

Distance (m)	Pixel from Projection Center (Pixel)	Distance (m)	Pixel from Projection Center (Pixel)
2.48	208	7.21	263
3.12	223	7.70	261
3.93	226	8.11	268
4.09	236	8.50	268
4.70	241	8.91	272
5.23	244	9.43	271
5.66	250	9.71	272
5.63	255	10.23	271
6.40	259	10.32	270
6.89	263	10.46	272

A.2 Lower Line of Bounding Box Extraction

With Equation (5.7) and Equation (5.8), the lower line of bounding box extraction shows from omnidirectional camera model in Figure A.2. From Figure A.2, the far the LRF distance is, the worse the error may be. However, the error is under 10%. The lower line of bounding box extraction has high accuracy. Figure A.2 demonstrates the 20 pieces of data form Figure 5.33. The lower line of bounding box shows from Figure A.2(a) to (q). With the LRF scan in omnidirectional camera, the pixel in omnidirectional camera can be acquired in semi-circle. Then the error estimation is in the histogram. Hence, the x axis is error pixel percentage and the y axis is the error pixel number.
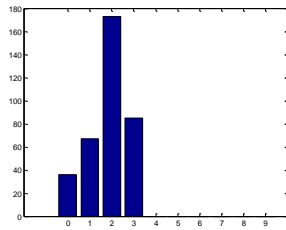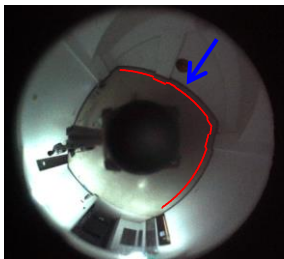
(a) 187.9 cm
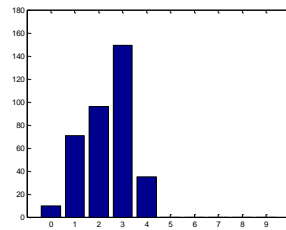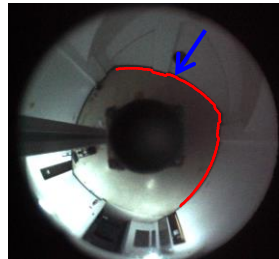
(b) 174.2 cm
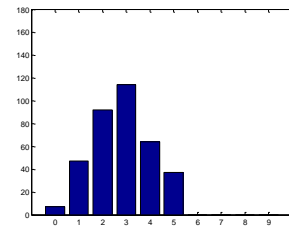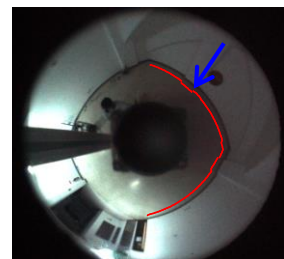
(c) 168.8 cm
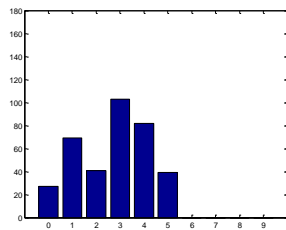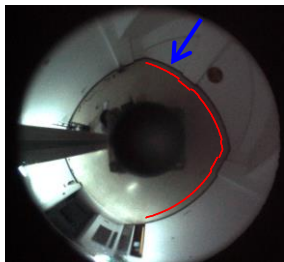
(d) 162.2 cm
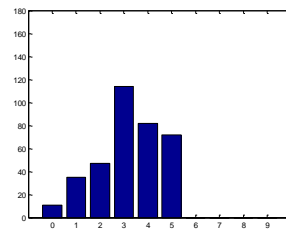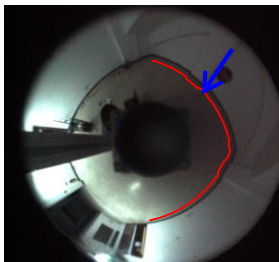
(e) 144.2 cm
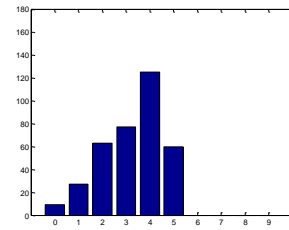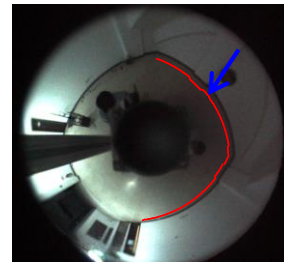
(f) 136.4 cm
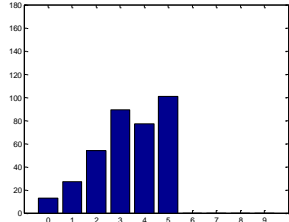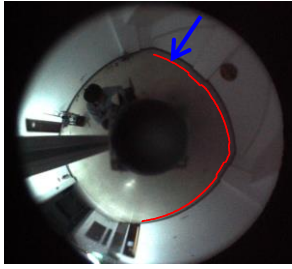
(g) 125.8 cm
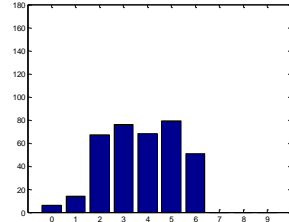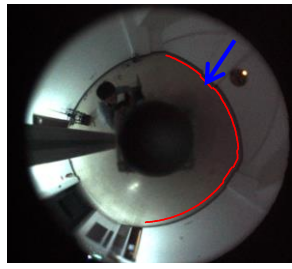
(h) 115.6 cm
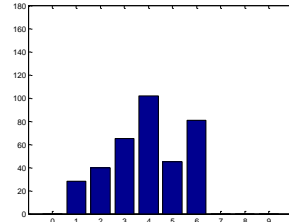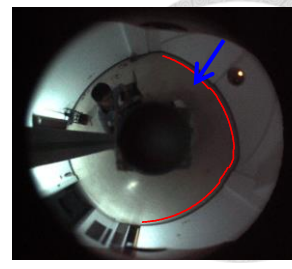
(i) 106.3 cm
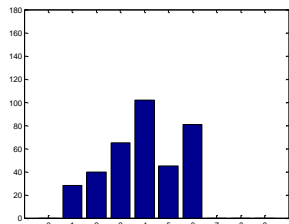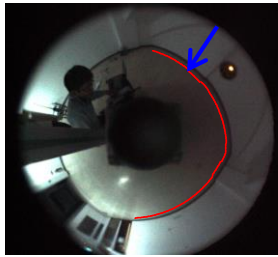
(j) 97.4 cm
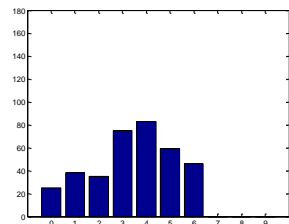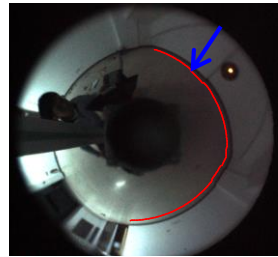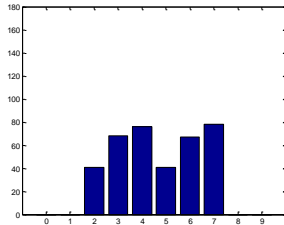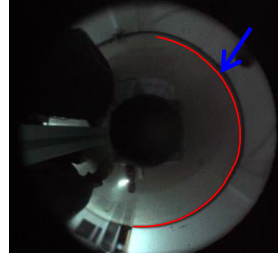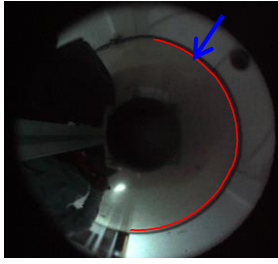
(k) 87.2 cm
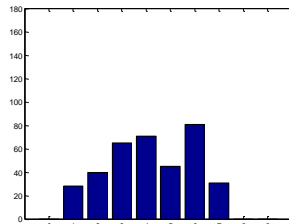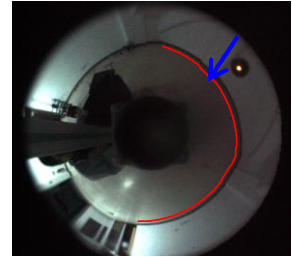
(l) 76.6 cm
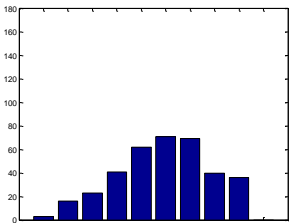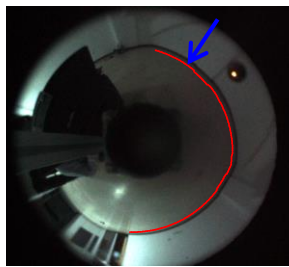
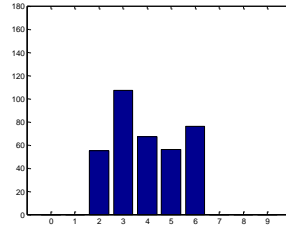(m) 67.8 cm
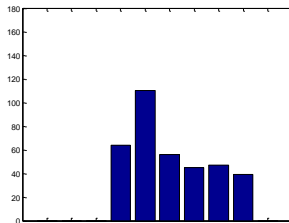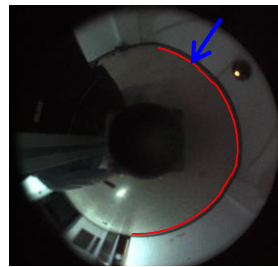
(n) 58.0 cm
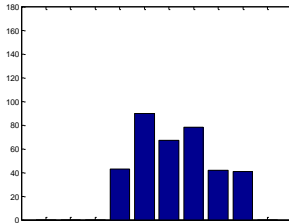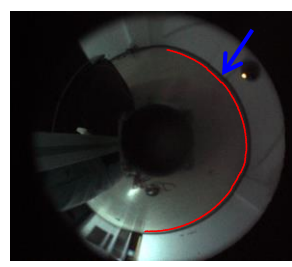
(o) 47.3 cm

(p) 37.4 cm   (q) 30.8 cm

Figure A.2 The Height Convert Pixel in Difference Height (From Figure 5.33)
The height convert pixels show from (a) to (q). With the LRF scan in
omnidirectional camera, the pixel in omnidirectional camera can be acquired. Then
the error estimation is in the histogram. The x-axis is error pixel percentage and the
y-axis is the error pixel number.

Table A.2 The Results in Figure A.2

Height(cm)	Mean	Standard Difference
30.8	1.85	0.89
37.4	2.35	0.99
47.3	2.80	1.23
58.0	2.72	1.46
67.8	3.21	1.33
76.6	3.27	1.28
87.2	3.36	1.41
97.4	3.73	1.51
106.3	3.93	1.53
115.6	3.96	1.43
125.8	3.42	1.73
136.4	4.20	1.75
144.2	4.90	1.91
162.2	5.05	1.62
168.8	5.30	1.53
174.2	4.82	1.72
187.9	3.98	1.38