Stony Brook University

Indoor Mobile Robot Localization using Laser-activated RFID

A Thesis Presented

by

Wenfei Liu

to

The Graduate School

in Partial fulfillment of the

Requirements

for the Degree of

Master of Science

in

Mechanical Engineering

Stony Brook University

December 2007

Stony Brook University

The Graduate School

Wenfei Liu

We, the thesis Committee for the above candidate for the

Master of Science degree,

hereby recommend acceptance of this thesis.

Yu Zhou, Ph.D., Advisor
Department of Mechanical Engineering, Stony Brook University

Jeffrey Ge, Ph.D., Chairperson
Department of Mechanical Engineering, Stony Brook University

Xing Wang, Ph.D., Third Member
Department of Electrical Engineering, Stony Brook University

This thesis is accepted by the Graduate School.

Lawrence Martin
Dean of the Graduate School

Abstract of the Thesis

Indoor Mobile Robot Localization using Laser-activated RFID

by

Wenfei Liu

Master of Science

in

Mechanical Engineering

Stony Brook University

2007

Localization is a fundamental problem in autonomous mobile robot navigation. It refers to the determination of the position and orientation of a mobile robot through the analysis of sensory data. By some authors the robot localization problem has been stated as the “most fundamental problem to providing robots truly autonomous capabilities" [1].

In this thesis, a novel localization technique for indoor mobile robot navigation using a collection of laser-activated RFID tags is presented. The laser-activated RFID tag is designed and used as the artificial landmark in the proposed localization system. They are distributed in the indoor environment. The robot localization is achieved through the combination of the stereo vision and RFID technologies and based on the principle of trilateration or triangulation. The localization system functions like an indoor GPS. The stereo vision provides the relative position between the robot and tags and the RFID tags

provide their absolute positions in the world coordinate. The robot can be uniquely located in the indoor environment by combining all the relative and absolute position information. Preliminary research shows that the proposed system is promising to provide a robust and accurate indoor localization method for mobile robots.

A new localization algorithm based on the trilateration and triangulation is presented in the second part. This algorithm can recover the robot position and orientation from a single image of identified landmarks taken by an onboard camera. The visual angle between two landmarks can be derived from their projections in the same image. The distances between the optical center and the landmarks can be calculated from the visual angles and the known landmark positions based on the law of cosine. The robot position can be determined using the principle of trilateration. The robot orientation is then computed from the robot position, landmark positions and their projections. Extensive simulation has been carried out. A comprehensive error analysis provides the insight on how to improve the localization accuracy.

Finally, a 2-axis laser scanner model is established to analyze the reflection relationship between the input and output laser beam direction caused by the rotating mirrors. The simplified model is calibrated and used in our localization system control the direction of laser beam.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First I would like to express my sincere gratitude to Professor Yu Zhou, my academic advisor, for his excellent guidance and constant support during these years. Without his instructions and ideas this thesis even can not be finished.

Many thanks to my lab-mate, Xiaoyu Cheng, and also my friends, Zhe Zhang, Hong Guo, and Xu Han for their cooperation. They make the journey of research more than interesting.

Last but not least, thanks to my wife, Ke Jin, for her strongest support through those years since we met.

Chapter 1 Introduction

1.1 Motivation and Background

Localization is a fundamental problem in autonomous mobile robot navigation. It refers to the determination of the position and orientation of a mobile robot through the analysis of sensory data. It is a key component in many successful autonomous robot systems. If a robot does not know where it is relative to the environment, it is difficult to decide what to do. The robot will most likely need to have at least some idea of where it is to be able to operate and act successfully. By some authors the robot localization problem has been stated as the “most fundamental problem to providing robots truly autonomous capabilities" [1].

Tremendous effort has been made to solve the problem of mobile robot localization, and a number of methods have been developed. Existing localization techniques for mobile robots can be classified into two categories: relative localization and absolute localization. The relative localization usually uses odometry, gyroscope and accelerometer. Since the position estimation is based on the previous position, the error in the estimation will accumulate and increase over time. On the contrary, the absolute localization supplies information about the location of the robot independent of previous location estimates. The location is not derived from integrating a sequence of

measurements, but directly from one measurement. It uses active beacon, landmark and map to derive the position information. This has the advantage that the error in the position does not grow unbounded, as is the case with relative position techniques. Obviously, the absolute localization has the advantages over the relative localization. However, two potential problems arise when using absolute localization: 1) how to get the measurement and 2) how to identify the landmarks.

The Global Positioning System (GPS) is one of the most successful tools used in the outdoor localization systems using absolute localization method [2-12]. By measuring the time-of-flight of the satellites' radio frequency signals which include information about the momentary locations of the satellites, a ground-based receiver can identify the satellites and compute its own position based on the principle of trilateration. With clock synchronization and error compensation, the accuracy of GPS-based positioning has been improving.

Due to the limitation of radio frequency signals, GPS can not be used in the indoor environment. Nonetheless, by recognizing the convenience brought by GPS in outdoor navigation, some researchers have made their effort to investigate GPS-like localization systems for mobile robots in indoor environments. These systems intend to achieve 3D indoor localization based on the active beacon without substantial online processing such as local map generation and matching. Many localization systems have been developed in the indoor environments [13-15]. However, their accuracy may be impaired by signal reflection and distortion due to the obstacles.

Vision sensors like cameras are also widely used in the localization system. They can provide much more useful geometric and radiometric information. Based on the

camera model, we can derive the distance measurement that we need. The landmarks in images are easy to be identified for the human eyes, but not for computers. Since the images are usually 2-D dimensional and information redundant, it requires a lot of image processing and computation work to extract useful information. Typical image processing work includes filtering, edge detection, and segmentation.

Recently, Radio Frequency Identification (RFID) becomes popular in both research and industry areas. It is a technology for automatically identifying objects by assigning an ID code to an electronic tag and retrieving the ID code using a wireless transceiver operating at radio frequencies. Some research has been conducted on RFID-based indoor mobile robot localization, where RFID tags are used as artificial landmarks [18-22]. A tag is easily identified by retrieving its ID code using an RFID reader. The tag ID can be used as an index to retrieve the accurate position of a tag from a database. However, they face the same problems of signal reflection and distortion.

Inspired by the some indoor localization system and RFID technologies, we are trying to combine the traditional stereo vision and the RFID technology to fully utilize their respective advantages. We propose a new indoor localization system for mobile robots using laser-activated RFID tags.

1.2 Objectives and Methodology

The objective of our research is to create a fast and accurate absolute localization technique for mobile robots moving in indoor environments. We intend to learn from the merits of the existing localization techniques and solve the problems with the existing

localization systems. The design of the proposed localization system is inspired by the GPS and RFID technologies.

The proposed localization system is artificial landmark-based, with a set of artificial landmarks pre-installed in an indoor environment with known absolute positions with respect to a global frame of reference, a set of onboard sensors to detect the artificial landmarks and an onboard computer to process the sensory data and localize a mobile robot. It is less expensive than a design which uses fixed sensors to track a mobile robot. Moreover, the distance between an artificial landmark and the mobile robot is measured along the line-of-sight. The general idea of the proposed localization technique is as following:

1) Artificial landmark design: A new type of RFID tag, the laser-activated RFID (LARFID) tag, is designed on the basis of an active RFID tag and used as the artificial landmark. Each LARFID tag has a unique ID, and after installation its absolute position in an indoor environment will be written into its own memory. The LARFID tag comes with a bright LED which makes it highly detectable from camera images. The LARFID tag will be activated by a laser beam, and then its ID and position information can be retrieved by an RFID reader. The LARFID tag will be deactivated when the laser beam is removed.

2) Landmark detection and localization: An LARFID tag is represented by its LED, and its position is defined by the position of the LED. Stereo vision is used to detect the LARFID tags. The position of an observed LARFID tag relative to the mobile robot can be calculated based on the perspective geometry. A laser beam will be shot from the mobile robot to activate the tag, and an onboard RFID reader will detect the activated tag.

The tag ID will be used to retrieve the stored absolute tag position via a follow-up inquiry by the RFID reader. Each time only one tag will be activated.

3) Mobile robot localization: The position and orientation of the mobile robot in the global frame of reference can be calculated based on the principle of trilateration or triangulation. The position of a LARFID tag relative to the mobile robot contains both the distance and bearing information. The absolute position of the tag is also known via RFID communication. The position and orientation of the mobile robot can be determined based on three or more tags.

4) Beyond the localization system, the algorithm part is also reviewed, which gives us an inspiration to combine the trilateration and triangulation. A new localization algorithm is presented to recover the position and orientation using a single image of identified landmarks.

1.3 Thesis Structure

Chapter 2 reviews current localization system and algorithm. Different kinds of localization system, GPS-based outdoor localization, GPS-liked indoor localization, and RFID-based indoor localization are classified. The different features of those systems give us an inspiration to create a new system to overcome current disadvantages. In the algorithm part, the traditional trilateration and triangulation algorithm is also provided. Although many improvements have been made on those algorithms, here we provide a novel localization algorithm to overcome their respective disadvantages.

In Chapter 3, a design scheme and prototype of our localization system is provided. The design scheme is illustrated at the beginning. Then a detailed description of

components and their functionality is elaborated. The stereo vision system, RFID tags, and laser scanner are three most important parts in our system. Both the hardware and software system have been established. Preliminary research shows that the proposed system is promising to provide a robust and accurate indoor localization method for mobile robots.

Chapter 4 presents a new localization algorithm based on the trilateration and triangulation. This algorithm can recover the robot position and orientation from a single image of identified landmarks taken by an onboard camera. Extensive simulation has been carried out. A comprehensive error analysis provides the insight on how to improve the localization accuracy.

Chapter 5 introduces the laser scanner modeling and calibration. It is a practical problem we solved in the construction of the localization system. A 2-axis laser scanner consists of two independent rotating mirrors. Geometric model is established to analyze the orientation of the final laser output beam caused by the two reflective mirrors. The simplified model is calibrated and used in our localization system control the direction of laser beam.

Chapter 6 presents the conclusions and future works of this research.

Chapter 2 Literature Review

2.1 Localization System

Many localization systems have been developed using different technology. GPS is widely used in the outdoor environment. Due to the nature of radio frequency signal, it can not be deployed in the indoor environment. Therefore, GPS-liked indoor system is developed to achieve 3D indoor localization without substantial online processing such as local map generation and matching. Recently, RFID technology is also used in the localization. A detailed review is provided in the following sections.

2.1.1 GPS –based Outdoor Localization

Global Positioning System (GPS) is a worldwide radio-navigation system developed by the US Department of Defense. It is formed from a constellation of 24 Earth-orbiting satellites (with 3 backup satellites) and their ground stations. These satellites are used by GPS receivers as reference points. The absolute position of each satellite is maintained and updated at the satellite, which eliminates the need of a central database. By measuring the time-of-flight of the satellites' radio frequency signals which include information about the momentary locations of the satellites, a ground-based receiver can identify the satellites and compute its own position based on the principle of

trilateration. With clock synchronization and error compensation, the accuracy of GPS-based positioning has been improving. Regular GPS provides an accuracy of <10m (after the removal of Selective Availability) [27]. Differential GPS, with pre-determined reference stations, can offer a positioning accuracy of 1-2m [27]. Recent technology based on phase measurement can achieve accuracy at centimeter or lower level, and real-time-kinematic (RTK) GPS can provide high positioning accuracy in motion [27].

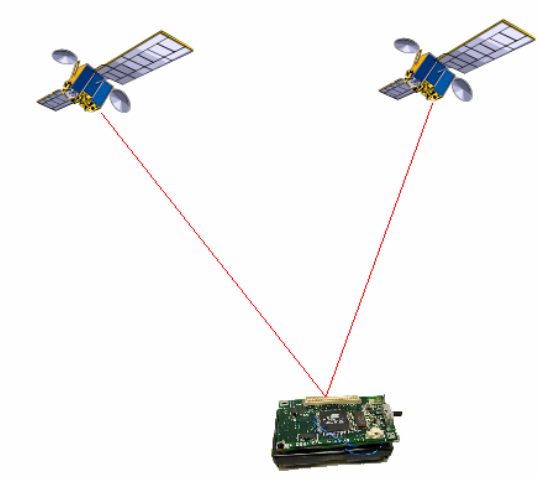

Fig. 2-1 GPS-based Outdoor Localization

GPS provides a convenient and powerful tool for outdoor navigation of mobile robots (Fig. 2-1). Li used a simple GPS for coarse estimation of robot position in outdoor environments [2]. Lenain et al. used RTK GPS for accurate path tracking of an agricultural robot in presence of slippage [3]. However, GPS receivers require an unobstructed view of the sky, so they are used only outdoors. They often have increased errors near trees or tall buildings due to signal blockage and multi-path interference [4, 5]. In other words, accurate distance measurement depends on “line-of-sight”. In addition, latency caused by data transmission, processing and communication with multiple

satellites increases the difficulty in dynamic applications [6]. A well-accepted solution is to fuse GPS data with odometry or inertial measurements to improve the outdoor positioning accuracy, which combines the long-term stability of GPS and real-time property of dead-reckoning [7-12].

It is many people believe that GPS will become the universal standard for outdoor applications [4，27]. However, an indoor equivalent to GPS is difficult to realize. Due to signal blockage and interference, none of the currently existing RF-based trilateration systems work reliably indoors. If the line-of-sight between stationary and onboard components can be maintained, RF-based solutions may work indoors as well.

2.1.2 GPS-like Indoor Localization

Recognizing the convenience brought by GPS in outdoor navigation, some researchers have made their effort to investigate GPS-like localization systems for mobile robots in indoor environments. These systems intend to achieve 3D indoor localization without substantial online processing such as local map generation and matching. They are in general active beacon-based. Lai and Wu developed a robot localization system consisting of one onboard laser range finder with pan/tilt and four reflectors fixed in the environment [13]. Based on the measurement of distances between the reflectors and laser range finder, the 3D position of the laser range finder can be determined using trilateration; and the 3D orientation of the mobile robot can be obtained by measuring three positions of the movable laser range finder. However, a robust method for distinguishing the reflectors is missing, which jeopardizes the validity of robot localization. Yi and Choi developed a robot localization system consisting of four fixed

ultrasonic generators and a pair of onboard ultrasonic receivers [14]. The distances are calculated based on the time-of-flight. To avoid crosstalk between the ultrasonic signals, the RF receivers and transmitter are added to the ultrasonic generators and receivers respectively, and the ultrasonic generators are activated successively using different RF channels. However, in less open indoor environments, inaccuracy may arise in distance measurement due to the reflection of ultrasonic signals. Prigge and How developed a robot localization system based on a low frequency magnetic field which is generated by a few electromagnetic beacons and can penetrate some line-of-sight obstacles [15]. The position of a sensor is calculated based on the amplitude measurement of the magnetic field vector which is the sum of superimposed beacon fields. Individual beacon fields are distinguished based on encoded time-varying magnetic fields. However, signal distortion may result from eddy field noise caused by time-varying magnetic fields, additive white noise in circuits, and existence of ferromagnetic materials in the environments. Active bat [37] (Fig. 2-2) is another localization system developed in the indoor environments based on ultrasound signals respectively. However, its accuracy may also be impaired by signal reflection and distortion due to the obstacles.

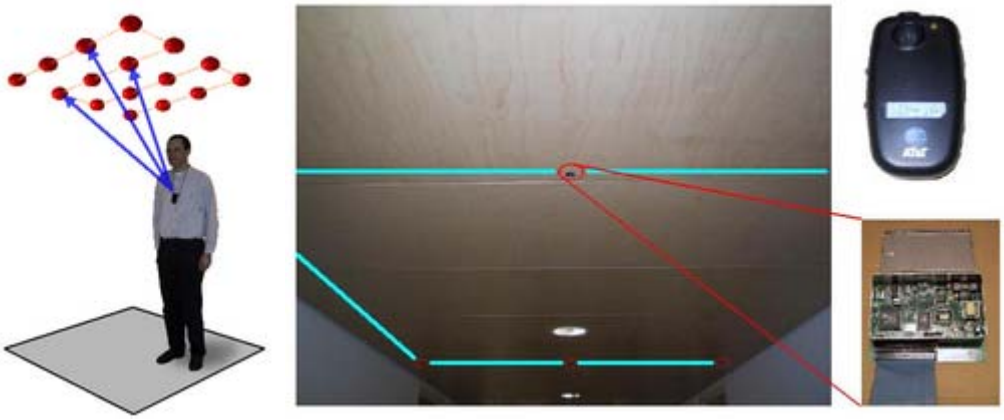

Fig. 2-2 Active Bat Indoor Localization System

2.1.3 RFID-based Indoor Localization

Radio Frequency Identification (RFID) is a technology for automatically identifying objects by assigning an ID code to an electronic tag and retrieving the ID code using a wireless transceiver operating at radio frequencies. RFID features non-contact and non-line-of-sight readability. The RFID reader can detect the existence of a tag within a certain range. RFID tags can be active, which have a power supply for sending their responses, or passive, which are powered by the RF energy transferred from the reader. In general, active tags are more reliable and more expensive, and can be read over a distance of tens of meters; while passive tags are less reliable and cheaper, and can be read over a distance from a few centimeters to a few meters [16].

Recently, some research has been conducted on RFID-based indoor mobile robot localization, where RFID tags are used as artificial landmarks. A tag is easily identified by retrieving its ID code using an RFID reader. The tag ID can be used as an index to retrieve the accurate position of a tag from a database. Alternatively, since many RFID tags have their own memory, the position information of each tag can be stored on the tag directly and read out by an RFID reader later, which eliminates the need of a central database. This function can be very helpful in dynamic environments where landmarks move and need to update their locations from time to time.

Existing RFID-based localization operates generally in two different ways. One way is to calculate the distance between a tag and a reader based on the time-of-flight of RF signals. Kim et al. developed an RFID system, including three orthogonal antennae, which determines the direction of a tag by comparing the signal strength in each direction and measures the tag-reader distance from phase shift [17]. Inaccuracy may arise due to

RF signal blockage and interference. Another way is based on the fact that a reader can detect a tag within a certain range. In the works of Tsukiyama [18,19] and Kulyukin et al. [20-22], passive RFID tags were used as landmarks to remind the mobile robot its position and help to decide next movement. The robot positioning accuracy depends on the RFID range and is coarse in general. Yamano et al. proposed an RFID-based mobile robot localization method using Support Vector Machine Learning [23]. The robot positioning accuracy depends on the learning resolution, and 20% ambiguity was reported for locating the robot at correct locations. Chae and Han proposed an active RFID tag-based localization method which divides the environment into small regions and determines the region where the robot lies by weighing each tag based on its distance to its region boundary [24]. The success rate of locating the robot in correct regions is reported as 91.5%. Hahnel et al. developed an RFID-based mobile robot localization method by first learning a probabilistic sensor model (Fig. 2-3), which describes the likelihood of detecting a tag given its position relative to the antenna, and then localize the robot using Monte Carlo localization based on that model [25]. However, the robot localization accuracy using only RFID tags is low, with a reported average error about 2m. Jia et al. used passive RFID tags as landmarks with known absolute positions in vision-based localization [26]. When the mobile robot moves into the range of a tag, the onboard reader detects the tag and the onboard camera localized the robot relative to the tag. When neighboring ranges overlap, ambiguity may arise due to the mismatch between the detected and observed tags. Otherwise, when they are totally separated, the mobile robot may not be able to localize itself in the transitional region between two neighboring

ranges. In a word, the inaccuracy and ambiguity induced by the RFID range are the major problems of the existing RFID-based indoor mobile robot localization.

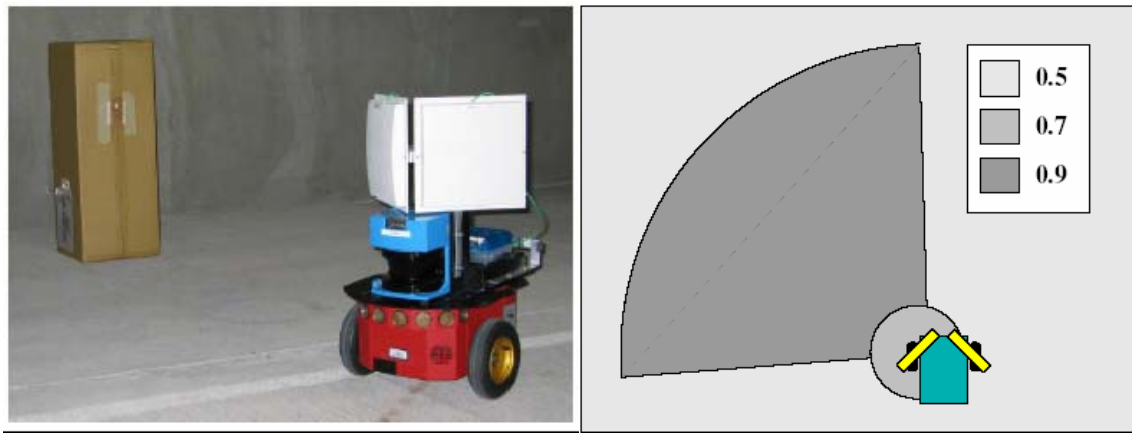

Fig. 2-3 RFID-based Indoor Localization System

2.2 Localization Algorithm

Trilateration and triangulation are two widely adopted localization methods used in different systems. Many improvements have been made to solve the problem efficiently and accurately. Trilateration is a traditional method to locate an object using distance measurement. However, its disadvantage is that it can not recover the orientation information. Triangulation, instead of using distance measurements, uses bearing measurements among references to locate an object. Here we review some works related to these two methods.

2.2.1 Trilateration Method

Trilateration is a method which determines the position of an object based on simultaneous distance measurements from three or more known references (Fig. 2-4). It

is equivalent to finding the intersection of three or more spheres. Manolakis provided a closed-form solution to the trilateration problem [30]. He also presented a systematic error analysis method, taking into account the distance measurement error [31]. Coope derived two alternative closed-form solutions for trilateration, one based on Gaussian elimination and the other based on orthogonal decomposition and transformation [32]. He also presented a nonlinear least-square optimization method to get the approximate solution from erroneous measurements. Recently, Thomas and Ros proposed a general formulation for the closed-form solution using the Cayley-Menger determinant defined by constructive geometric arguments [33]. Fang showed that the formulation of trilateration can be simplified by defining the problem in a frame attached to one of the references [34].

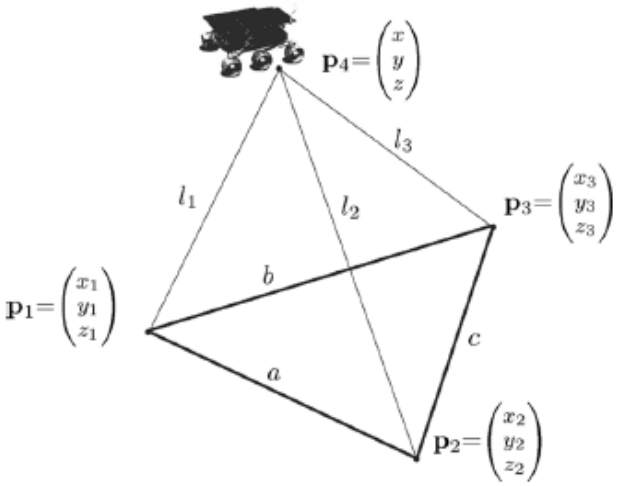

Fig. 2-4 Trilateration Problem

Many applications have been developed based on the principle of trilateration. The Global Positioning System (GPS) uses trilateration to locate a receiver based on the travel distances of radio frequency signals from the satellites [35]. It provides a powerful tool

for outdoor mobile robot localization and navigation [36]. In indoor environments, some localization systems have been developed, such as SpotOn [37], Active Badge [38], Active bat [39], Cricket [40], based on radio frequency [37], infrared [38], and ultrasound [39, 40] signals respectively. Recently, Zhou et al. have also introduced a new indoor localization method for mobile robots based on the laser-activated RFID landmarks [41].

2.2.2 Triangulation Method

Triangulation, instead of using distance measurements, uses bearing measurements among references to locate an object. The basic idea of triangulation is that in a plane containing the references and object, the object is located at the intersection of the circles each of which is determined by two references and the bearing between them (Fig. 2-5). Cohen compared four solution methods for triangulation, i.e. iterative search, geometric triangulation, geometric circle intersection and Newton-Raphson iterative method, and showed that geometric circle intersection is the most robust one among them [42]. Betke and Gurvits presented a position estimation algorithm using the complex numbers representation of the landmarks [43]. The main advantage of this algorithm is the linear time complexity with respect to the number of landmarks and the robustness to the noisy input. Shimshoni also presented an algebraic solution by applying several transformations to the linear system of equations which are defined by triangulation constraints [44]. He showed that these transformations indeed improved accuracy. Sutherland and Thompson discovered that the position error is influenced by both the input bearing error and the distribution of landmarks [45].

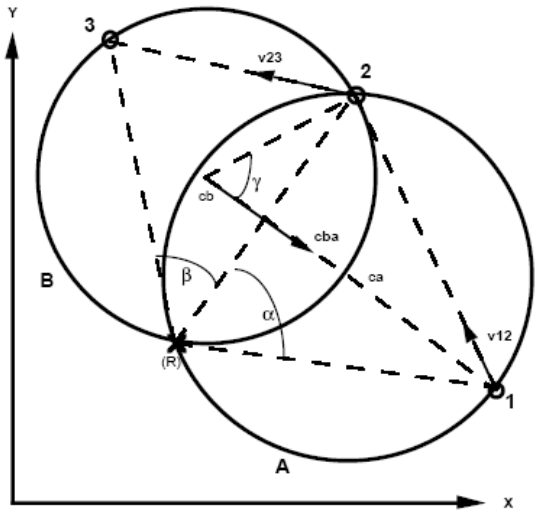

Fig. 2-5 Triangulation Problem

Vision sensors are widely used to measure the bearings. The structured features, such as doors and wall corners, are extracted from images. The 2D bearings of landmarks can be recovered from vertical edges. Muñoz and Gonzalez developed a 2D landmark-based triangulation algorithm in which the bearings are derived from a single image [46]. Other systems utilizing the bearing measurements for localization can be found in [43, 47-49].

Chapter 3 Laser-activated RFID-based Indoor Localization System

Inspired by the GPS and RFID technologies, we propose a new indoor localization system for mobile robots. The proposed localization system is artificial landmark-based, with a set of artificial landmarks pre-installed in an indoor environment with known absolute positions with respect to a global frame of reference, a set of onboard sensors to detect the artificial landmarks and an onboard computer to process the sensory data and localize a mobile robot. It is less expensive than a design which uses fixed sensors to track a mobile robot. Moreover, the distance between an artificial landmark and the mobile robot is measured along the line-of-sight.

3.1 Design Scheme

A novel type of RFID tag, the laser-activated RFID (LARFID) tag, is designed on the basis of the existing RFID technique, and used as the artificial landmark in the proposed indoor localization system for mobile robots. The LARFID tag extends an active RFID tag with a light-emitting diode (LED) and a laser-activated switch circuit. A function diagram of the LARFID-based indoor localization system for mobile robots is presented in Fig. 3-1.

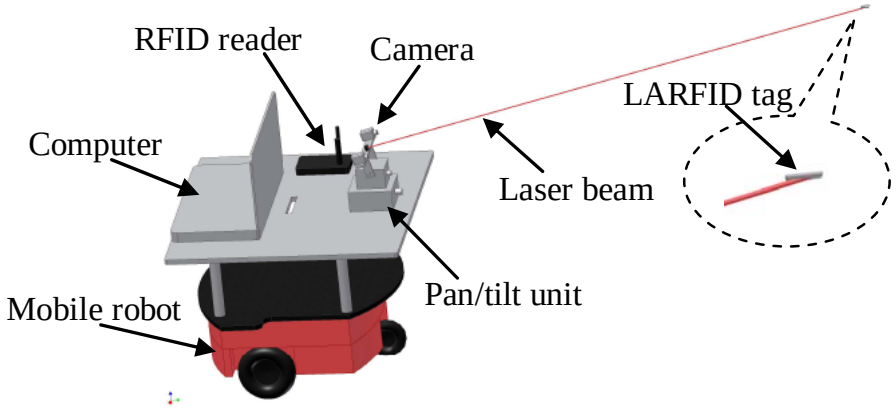

Fig. 3-1 Function of the LARFID-based Localization System

3.1.1 Active RFID Tag

An active read/write RFID tag is used as the base of the designed artificial landmark. The main reason for using the active RFID tag is because it has its own power supply which can be easily controlled to activate and deactivate the tag. Moreover, comparing with passive RFID tags, active tags work more stable and can be detected over a longer range. As a result, a relatively small number of artificial landmarks are needed for mobile robot localization.

The unique ID of an RFID tag is used to identify the associated artificial landmark. An RFID reader is used to receive the assigned tag ID. The tag ID can be used as the index to retrieve the absolute position of the associated artificial landmark from a database. However, instead of doing this, in our design the absolute position of an artificial landmark is written into the memory of the underlying RFID tag and read out by the RFID reader. With this strategy, there is no need to maintain a central database for all the artificial landmarks in an environment for the purpose of mobile robot localization.

More importantly, the stored position information can be updated, which makes the designed artificial landmark applicable to an indoor environment under change.

By identifying each artificial landmark using the unique ID of the underlying RFID tag, we intentionally avoid the shape-based and pattern-based artificial landmark designs. It may potentially reduce the cost of designing and fabricating a large number of uniquely-shaped or patterned artificial landmarks for applications in large indoor environments. As a result, the artificial landmarks in an environment have the same appearance.

3.1.2 LED for Landmark Detection

With the understanding that the line-of-sight plays an important role in accurately locating an artificial landmark relative to the mobile robot, we choose to use a line-of-sight-based method, instead of a time-of-flight-based method, for landmark localization. An onboard stereo vision unit, consisting of two cameras with fixed transformation between them, is used to detect the artificial landmarks and measure the position of an observed artificial landmark relative to the mobile robot.

In order to make landmark detection robust and avoid substantial online image processing, a bright LED is attached to each RFID tag and kept on constantly. In our design, the LED is regarded as the representative point feature of an artificial landmark, and the position of the artificial landmark is represented by the position of its LED. The apertures of the onboard cameras are intentionally minimized. As a result, with regular ambient lighting, the LEDs are shown as bright spots outstanding from the dark background in the images. To find the artificial landmarks from an image, we only need

to segment out the LED spots and find their centers, which does not require very complicated image processing procedure.

During each localization operation, a pair of images is taken from the stereo vision unit, one from each camera. The correspondence between the LED spots in the two images can be established by using a method such as epipolar line search [28] or combinatorial optimization [29]. In the work of this paper, we assume that a same set of artificial landmarks can be observed by both cameras, which can be realized in practice by adjusting the relative transformation between the two cameras and limiting the pan/tilt range of the stereo vision unit according to the distribution of the artificial landmarks in the environment. More complicated situations, e.g. unequal number of LED spots appearing in the two images, will be studied in our future work. With the known correspondence, the position of each observed landmark (LED) relative to the mobile robot can be calculated based on the perspective geometry.

3.1.3 Laser Activation for Localization

Given the positions of some landmarks relative to the mobile robot, we also need to know the absolute positions of those landmarks in a global frame of reference in order to localize the robot in the same frame. In our design, the absolute position of an artificial landmark is stored in the memory of the associated RFID tag, and can be retrieved by inquiry with the tag ID using an RFID reader installed on the mobile robot.

There is a potential problem with regular RFID tags. With regular RFID tags, the RFID reader detects all the functioning RFID tags within its reading range. Meanwhile, stereo vision recovers some artificial landmarks based on their LEDs. However, the

correspondence relationship between the artificial landmarks detected by the RFID reader and those observed by stereo vision is missing. Moreover, depending on how the artificial landmarks are distributed in the environment, at each moment some of them may be caught by the cameras, others may not. It may also be possible that the stereo vision unit detects some artificial landmarks which lie outside the range of the RFID reader and cannot be detected by the RFID reader.

To solve this problem, we add to each active RFID tag a photodiode-based switch circuit which stands between the battery and function part of the RFID tag (Fig. 1). In the proposed system, a laser pointer with pan/tilt capability will also be installed on the mobile robot. The modified RFID tag is called laser-activated RFID (LARFID) tag which will be activated by a laser beam sent from the mobile robot. Usually the switch circuit on an LARFID tag is off, and the tag is not powered and not functioning. When an LARFID tag is detected and its position relative to the mobile robot is calculated via stereo vision, the robot shoots a laser beam to the LARFID tag. The laser beam hits the photodiode and generates a voltage signal to open the switch circuit and turn on the power of the tag. Once the LARFID tag is activated, it will be detected by the onboard RFID reader. Then the RFID reader can retrieve the absolute position of the tag, which is stored in the tag memory, by a follow-up inquiry with the tag ID. After this operation, the mobile robot turns off the laser beam. Once the laser beam is removed from the photodiode, the switch circuit will be closed, and the tag will be deactivated. Intentionally, only one laser beam is used, and each time only one LARFID tag is activated by the laser beam. The mobile robot decides which tag to target, and the RFID reader knows which

tag it detects. In this way, the correspondence between an observed LARFID tag and its ID is established successfully.

To avoid using an LARFID tag which is observed by stereo vision but lies outside the range of the RFID reader, we set a threshold value for valid distance measurement based on the capability of the RFID reader, tags and laser beam and the accuracy of distance measurement. Only those tags at a distance from the mobile robot shorter than the threshold are considered valid for robot localization.

3.1.4 Laser Activation for Robust Landmark Detection

Besides retrieving the absolute position of an LARFID tag, the laser activation operation provides an easy way to check and correct the landmark matching and reconstruction results from stereo vision. In general, if the position of an LARFID tag relative to the mobile robot is reconstructed based on the correct correspondence, then when the laser beam hits the targeted tag, it will be activated and detected by the RFID reader. Otherwise, no tag physically exists at the targeted location, and nothing will be detected by the RFID reader, which means a mismatch in stereo vision. Then the landmark matching in stereo vision can be partially redone with only those mismatches. As a result, landmark detection and localization becomes more robust. An alternative method can be also used when only a small number of tags are observed. In fact, with N tags caught by two cameras, there are totally N^2 possible tag locations in space. By shooting laser beam to every possible location, the robot can eventually figure out the N valid tag locations.

3.2 System Prototype

A prototype system has been built for the feasibility study of the LARFID-based indoor localization technique for mobile robots. The system consists of a mobile subsystem (Fig. 3-2) and a stationary subsystem (Fig. 3-3).

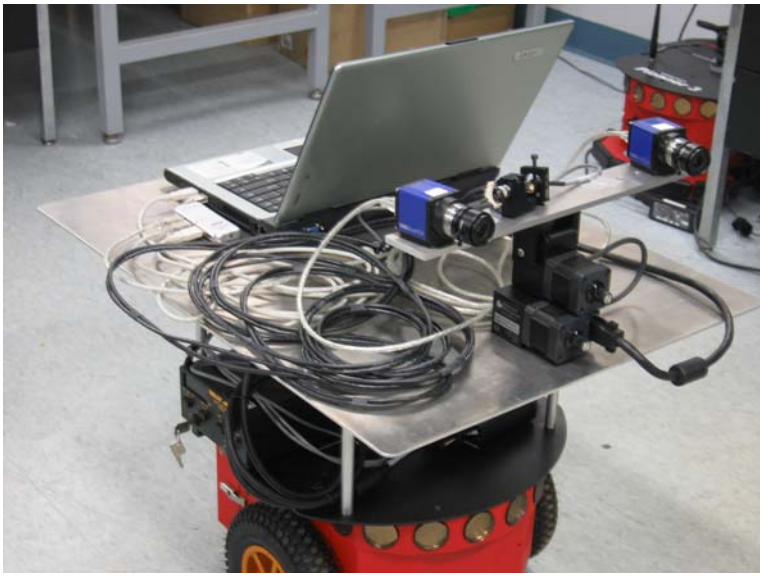

Fig. 3-2: Mobile Subsystem

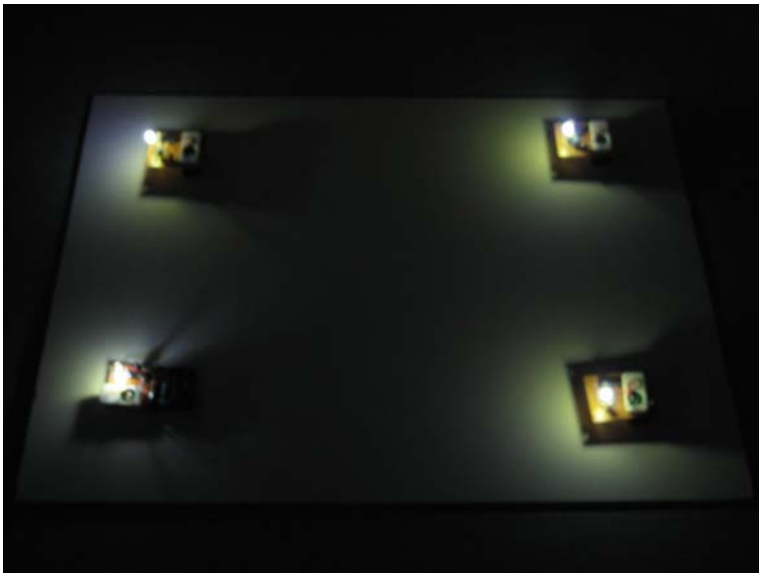

Fig. 3-3: Stationary Subsystem

3.2.1 System Components

The mobile subsystem includes a mobile robot, onboard notebook, stereo vision system (two cameras), pan-tilt, laser generator and scanner, data acquisition board (DAQ), and RFID reader. The stationary subsystem is actually a world coordinate defined by the modified RFID tags. The detailed view of the main components is presented in the following diagram (Fig. 3-4), which also illustrates the relationship between those components.

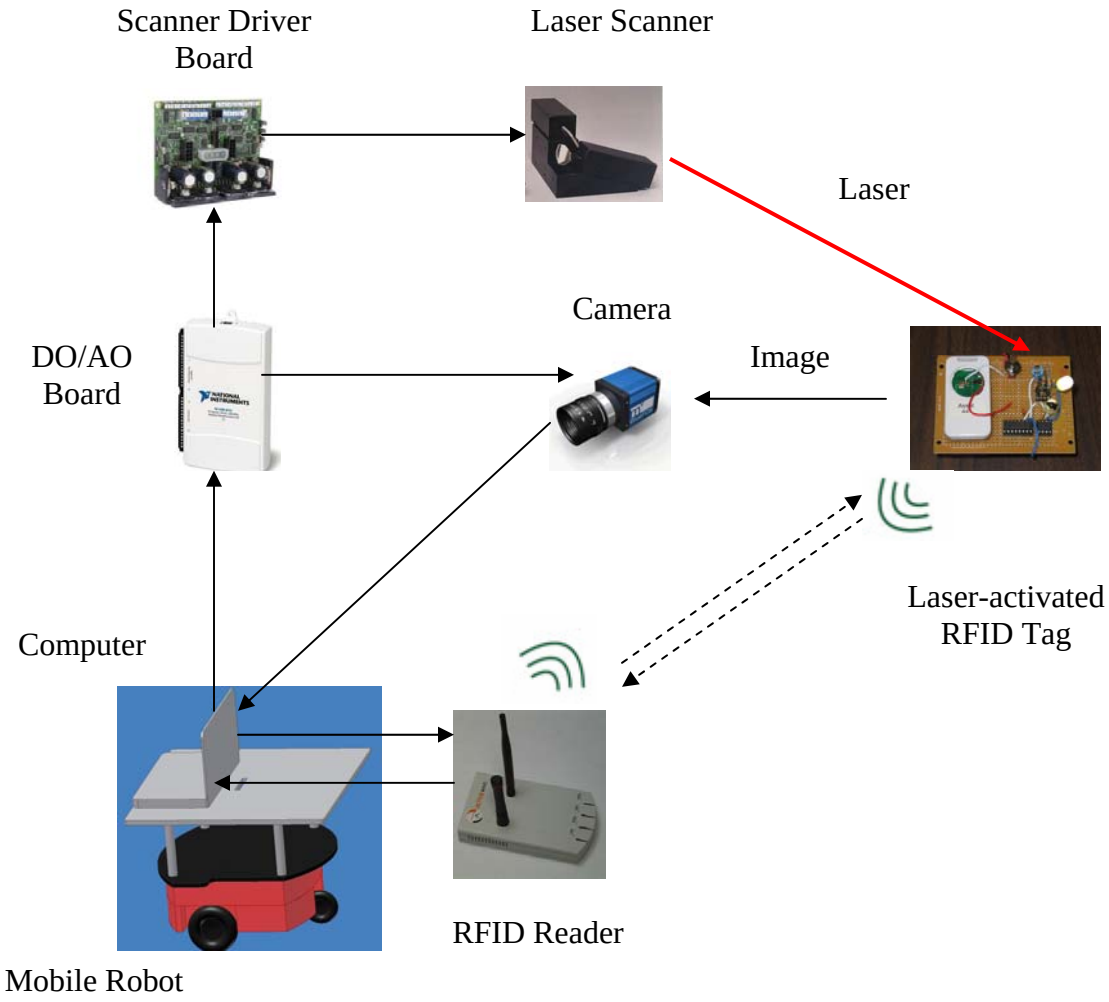

Fig. 3-4 LARFID-based Localization System Components

In our experimental system, an Activmedia Pioneer3-DX mobile robot is used as the sensor carrier (Fig. 3-2). Pioneer3-DX is an agile and versatile intelligent mobile robotic platform updated to carry loads more robustly and to traverse sills more surely with high-performance current management to provide power when it's needed. In the feasibility study, we did not use the original onboard sensors, such as encoders, gyroscope and sonar ring, with the purpose of testing the proposed LARFID-based localization technique.

A laptop computer is mounted on the top of the mobile robot (Fig. 3-2), with Intel Celeron M 1.5GHz CPU and 512MB RAM. All the above-mentioned onboard devices are controlled by the computer, and the computations for stereo vision, laser scanning and localization are implemented by the computer.

For detecting the LARFID tags and measuring their relative positions with respect to the mobile robot, a stereo vision unit, consisting of two Matrix Vision BlueFox USB cameras with Kowa 12mm lenses, is installed on the mobile robot (Fig. 3-2). This stereo vision set is supported by a Directed Perception PTU-D46-17 controllable pan-tilt unit which provides the stereo vision unit the landmark search capability.

Melles Griot Laser Diode Driver is used to generate the laser beam. The size and intensity of laser can be manually adjusted according to the experiment’s needs. Since the laser beam is used to activate the RFID tags, its direction has to be controlled accurately. The Cambridge Technology Galvanometer with two controllable mirrors (Fig. 3-4) can accurately reflect the laser beam to the designed direction. A corresponding Scanner Driver Board is also required.

The National Instrument NI USB-6211 data acquisition board (DAQ) is used to generate the digital and analogy signal. The reason we need to use digital and analogy signal is due to the cameras and laser scanner. To capture the simultaneous images by two cameras, a rising edge pulse signal is used to trigger both cameras. Also, the two analogy input signals are required by the laser scanner server board to control the mirror rotating angles.

For identifying the tags and retrieving their absolute positions, an ActiveWave standard RFID reader is installed under the aluminum frame (Fig. 3-2). The RFID reader is designed for fast and easy system integration without losing performance. The RFID reader consists of a real time processor, operating system, virtual portable memory, and transmitter/receiver unit in one small self-contained module that is easily installed in the ceiling or in any other convenient location. This RFID reader has a range of 85m for reading and 30m for writing. The corresponding active tags have the ability of long read/write ranger and large data storage capacity. Further more, to satisfy our system needs, the active tags are extended with a light-emitting diode (LED) and a laser-activated switch circuit.

By combining all those components together, we create a new indoor localization system. First of all, the LED illuminated tags are captured by the cameras simultaneously. We can adjust the orientation of robot and pan-tilt to make sure enough tags are captured. After simple image processing, the relative position of tags in the camera reference is derived. Using the transformation between the camera and laser scanner, the position in the scanner reference is derived and the corresponding laser beam is controlled to activate the RFID tags. Now only the signal from activated tag can be received by the RFID

reader. Therefore, a pair of identified absolute position and relative position is matched. Based on the trilateration, we can successfully recover the robot position in the world coordinate. The more detailed technical part about the stereo vision, modified tags, and the laser scanner is presented in the following sections.

3.2.2 Stereo Vision Unit Calibration

Stereo vision is a process of reconstructing the object depth from the fusion of two images captured simultaneously. Usually it requires two cameras. The basic principle of stereo vision is based on the geometric relationship between cameras and object, which is illustrated in the Fig 3-5.

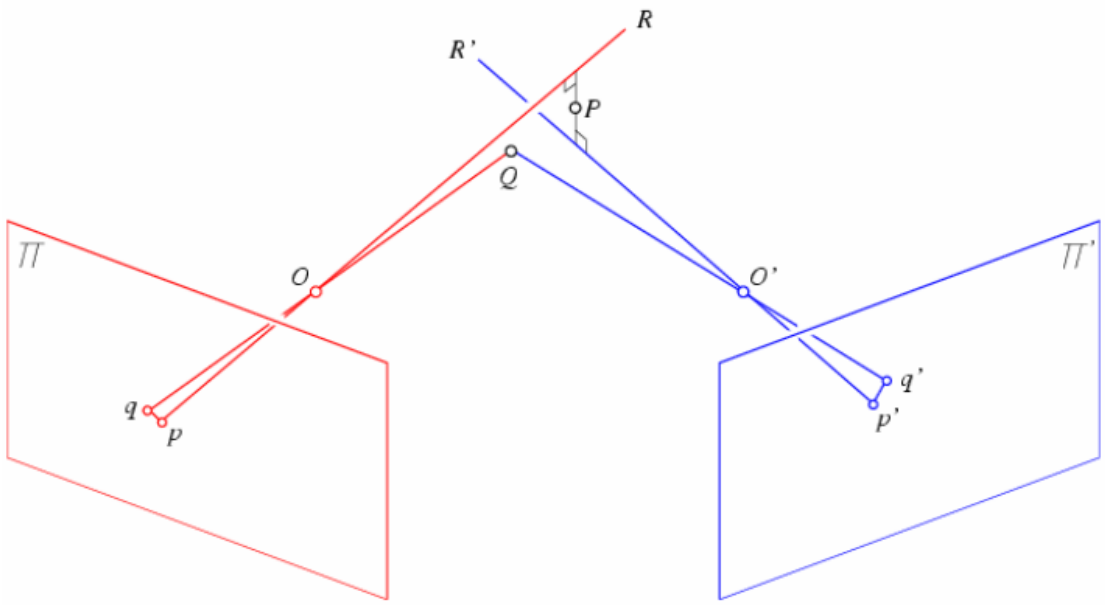

Fig 3-5 Stereo Vision Model

Several kinds of methods are available to reconstruct the object. One way is to find the mid-point of line segment perpendicular to the ray R and R'. Alternatively, we can reconstruct a scene point using a purely algebraic approach:

$$\begin{cases} pMP = 0 \\ p'M'P = 0 \end{cases} \Leftrightarrow \begin{pmatrix} [p_{\times}]M \\ [p'_{\times}]M' \end{pmatrix} P = 0 \qquad 3.1$$

where M and M′ are calibration matrix for two cameras. This is an over constrained system of four independent linear equations in the coordinates of P that is easily solved using the linear least square techniques.

No matter what method we use, we have to rely on the camera intrinsic parameters and the transformation matrix between two cameras. This is the purpose of our stereo vision calibration.

The perspective camera model is used to calibrate the camera parameters. Also, the distortion is considered according to the "Plumb Bob" model [51],

$$x_d = (1 + k_1 r^2 + k_2 r^4 + k_5 r^6)\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 2k_3 xy + k_4(r^2 + 2x^2) \\ k_3(r^2 + 2y^2) + 2k_4 xy \end{pmatrix}, \qquad 3.2$$

where $r^2=x^2+y^2$, (x,y) is the normalized image projection, and k_1 through k_5 are the distortion coefficients.

A camera calibration toolbox for Matlab is available at "http://www.vision.caltech.edu/bouguetj/calib_ doc/". A printed black/white checkerboard pattern was used (Fig. 3-6), with different poses and at different distances from 1m to 2m, to find out the intrinsic parameters of the cameras, such as focal length, image center and distortion coefficients, and the transformation between the two cameras. The calibration results of the intrinsic parameters are listed in Table 3-I. Only k_1 and k_2 are listed in the table because k_3, k_4, and k_5 are ignorable small.

TABLE 3-I
CALIBRATED INTRINSIC CAMERA PARAMETERS

Parameter	Left Camera	Right Camera
Focal length (pixel)	(1629.3, 1632.5)	(1634.9,1637.5)
Image center (pixel)	(317.68, 227.17)	(327.35, 237.29)
Distortion coefficients	-0.37199, 0.038818…	-0.38701, 0.23067…

Considering a point P in the 3D space, its two coordinate vectors $\mathbf{X_R}$ and $\mathbf{X_L}$ in the left and right camera reference frames respectively are related to each other through the rigid motion transformation $\mathbf{X_R = R * X_L + T.}$ The transformation matrix between these two cameras can be derived from the images of chessboard taken by two cameras simultaneously.

$$ {}^{left}T_{right} = \begin{pmatrix} 0.9659 & 0.0063 & 0.2589 & -290.9523 \\ -0.0056 & 1.0000 & -0.0033 & 2.6767 \\ -0.2589 & 0.0017 & 0.9659 & 19.3497 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \qquad 3.3 $$

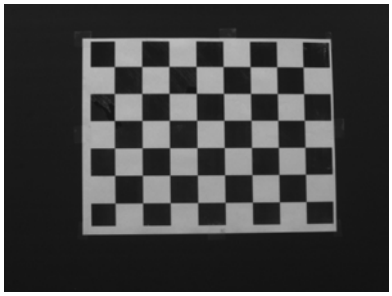

Fig. 3-6 Camera Calibration Pattern

3.2.3 Laser Scanner Calibration

The laser scanner consists of two rotating mirrors (Fig. 3-7). Each of the mirrors is controlled by an independent analogy signal. There is a linear relationship between the voltage of input analogy signal and degree of mirror rotating. With different rotating angles, the laser beam points to one orientation. The modeling of the laser scanner is established in the Chapter 5. Here we just use the derived simplified results from Chapter 5. In the virtual mirror coordinate frame that we mentioned in Chapter 5, the orientation of the output laser beam is expressed by two rotating angles, θ_1 and θ_2:

$$P_{out,0} = [-d \cdot \tan 2\theta_1,\ 0,\ 0]',$$

$$n_{out} = [-\tan 2\theta_1,\ \cos 2\theta_2,\ \sin 2\theta_2]'. \qquad 3.4$$

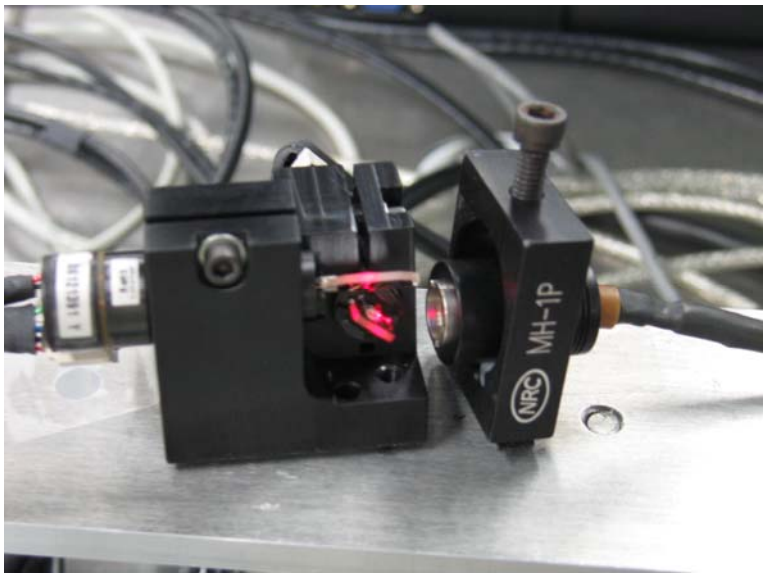

Fig. 3-7 Laser Diode and Laser Scanner

On the other hand, we can reconstruct the relative position of the tag Pc in the camera coordinate using stereo vision. However, this position is only related to the camera coordinate, whose origin, the optical center, can not be physically measured. As

we can see in the Fig 3-8, if we want to shoot the laser beam onto the tag, we have to find the transformation ${}^{m}_{c}T$ between the camera frame and laser scanner frame. Here arises the laser scanner calibration problem.

Assuming ${}^{m}_{c}T = [t_1; t_2; t_3; t_4]$, where $t_4=[0, 0, 0, 1]$, we have:

$$\frac{t_1 \cdot Pc+d \cdot \tan 2\theta_1}{t_2 \cdot Pc} = \frac{-\tan 2\theta_1}{\cos 2\theta_2}$$

$$\frac{t_1 \cdot Pc+d \cdot \tan 2\theta_1}{t_3 \cdot Pc} = \frac{-\tan 2\theta_1}{\sin 2\theta_2}. \qquad 3.5$$

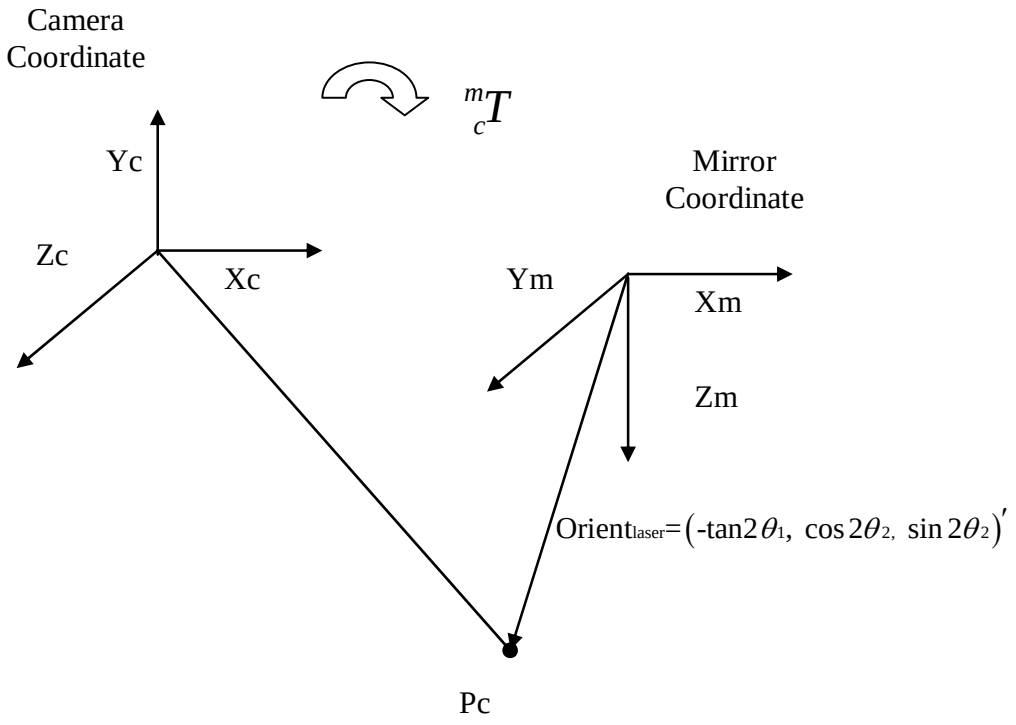

Fig. 3-8 Transformation between Camera and Scanner Frames

Since we are trying to find the transformation matrix ${}^{m}_{c}T$ by calibration, we know the predetermined position Pc and corresponding orientation.

Let us re-organize previous equations, we have:

$$t_1 \cdot Pc \cdot \cos 2\theta_2 + t_2 \cdot Pc \cdot \tan 2\theta_1 = -d \cdot \tan 2\theta_1 \cdot \cos 2\theta_2 ,$$

$$t_1 \cdot Pc \cdot \sin 2\theta_2 + t_3 \cdot Pc \cdot \tan 2\theta_1 = -d \cdot \tan 2\theta_1 \cdot \sin 2\theta_2 . \quad 3.6$$

But those equations are still not ready for use. Let us take the transpose on both sides, which only transpose the vectors and changes the sequence of vector multiple. We will have:

$$Pc^T \cdot \cos 2\theta_2 \cdot t_1^T + Pc^T \cdot \tan 2\theta_1 \cdot t_2^T + 0 \cdot t_3^T = -d \cdot \tan 2\theta_1 \cdot \cos 2\theta_2 ,$$

$$Pc^T \cdot \sin 2\theta_2 \cdot t_1^T + 0 \cdot t_2^T + Pc^T \cdot \tan 2\theta_1 \cdot t_3^T = -d \cdot \tan 2\theta_1 \cdot \sin 2\theta_2 . \quad 3.7$$

Convert it to the matrix form: $UT = D$, 3.8

where $U=\begin{pmatrix} Pc^T \cdot \cos 2\theta_2 & Pc^T \cdot \tan 2\theta_1 & 0 \\ Pc^T \cdot \sin 2\theta_2 & 0 & Pc^T \cdot \tan 2\theta_1 \end{pmatrix}$, $D=\begin{pmatrix} -d \cdot \tan 2\theta_1 \cdot \cos 2\theta_2 \\ -d \cdot \tan 2\theta_1 \cdot \sin 2\theta_2 \end{pmatrix}$ and $T=\begin{pmatrix} t_1^T \\ t_2^T \\ t_3^T \end{pmatrix}$.

This is a typical linear least-square problem. U is a 2by12 matrix with known variables, D is known 2 elements vector, and T is 12by1 vector which includes 12 unknown variables. Obviously, we need at least 6 points to find a unique solution. With enough n points, we will have 2n equations. U_{2n} is a 2nby12 matrix and D_{2n} is known 2n elements vector. According to the linear least square method, the solution is

$$T = \left(U_{2n}{}^T U_{2n}\right)^{-1} \cdot U_{2n}{}^T \cdot D_{2n} . \quad 3.9$$

Reorganize the T to the transformation matrix ${}^m_c T$. According to the equation 3.5, we can derive the rotating angles of scanners in the system conversely,

$$P_M = {}^m_c T \cdot P_C .$$

$$\theta_1 = -\frac{1}{2} \cdot \tan^{-1}\left(\frac{P_{M,1}}{\sqrt{P_{M,2}{}^2 + P_{M,3}{}^2} + d}\right),$$

$$\theta_2 = \frac{1}{2} \cdot \tan^{-1}\left(\frac{P_{M,3}}{P_{M,2}}\right) . \quad 3.10$$

3.2.4 Modified RFID Tags

A novel type of RFID tag, the laser-activated RFID (LARFID) tag, is designed on the basis of the existing RFID technique, and used as the artificial landmark in the proposed indoor localization system for mobile robots. The LARFID tag extends an active RFID tag with a light-emitting diode (LED) and a laser-activated switch circuit (Fig. 3-9). Fig. 3-10 is the prototype of the revised RFID tag.

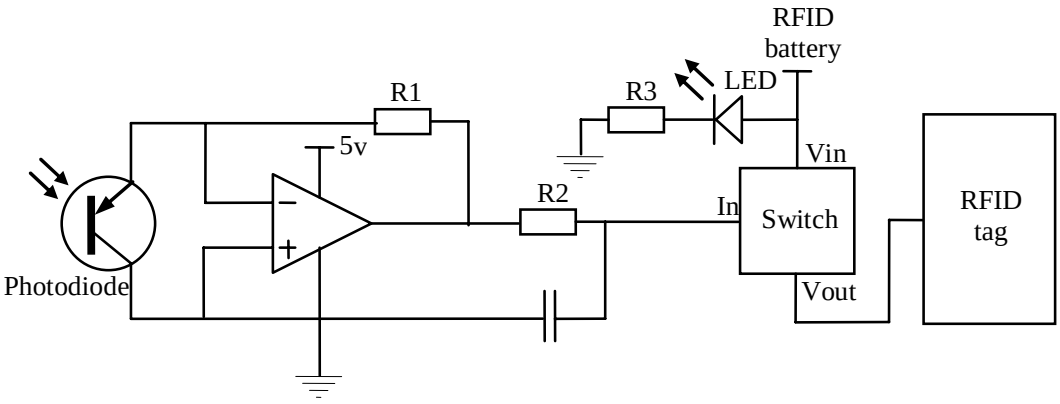

Fig. 3-9 Block diagram of the Laser-activated RFID Tag

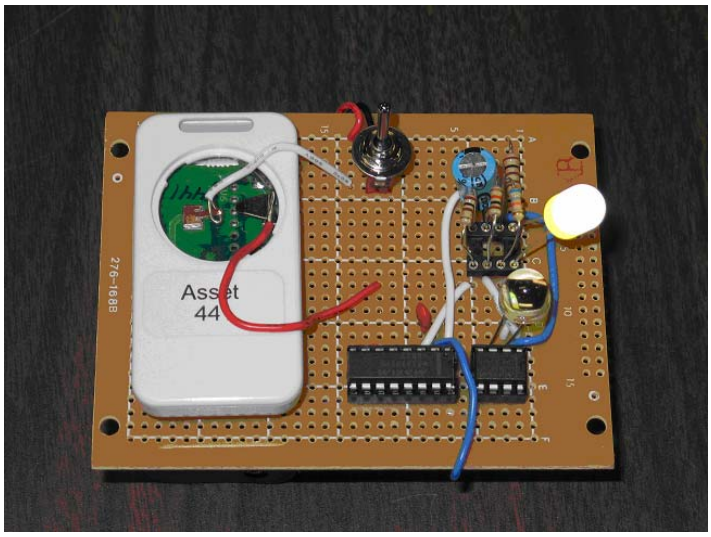

Fig. 3-10 Prototype LARFID tag

3.3 Software Design

Autonomous controlling software using C/C++ is developed to control and test our localization system (Fig. 3-11). Four code libraries for hardware are embedded in the software: robot library, camera capturing library, data acquisition board (DAQ) library and RFID library.

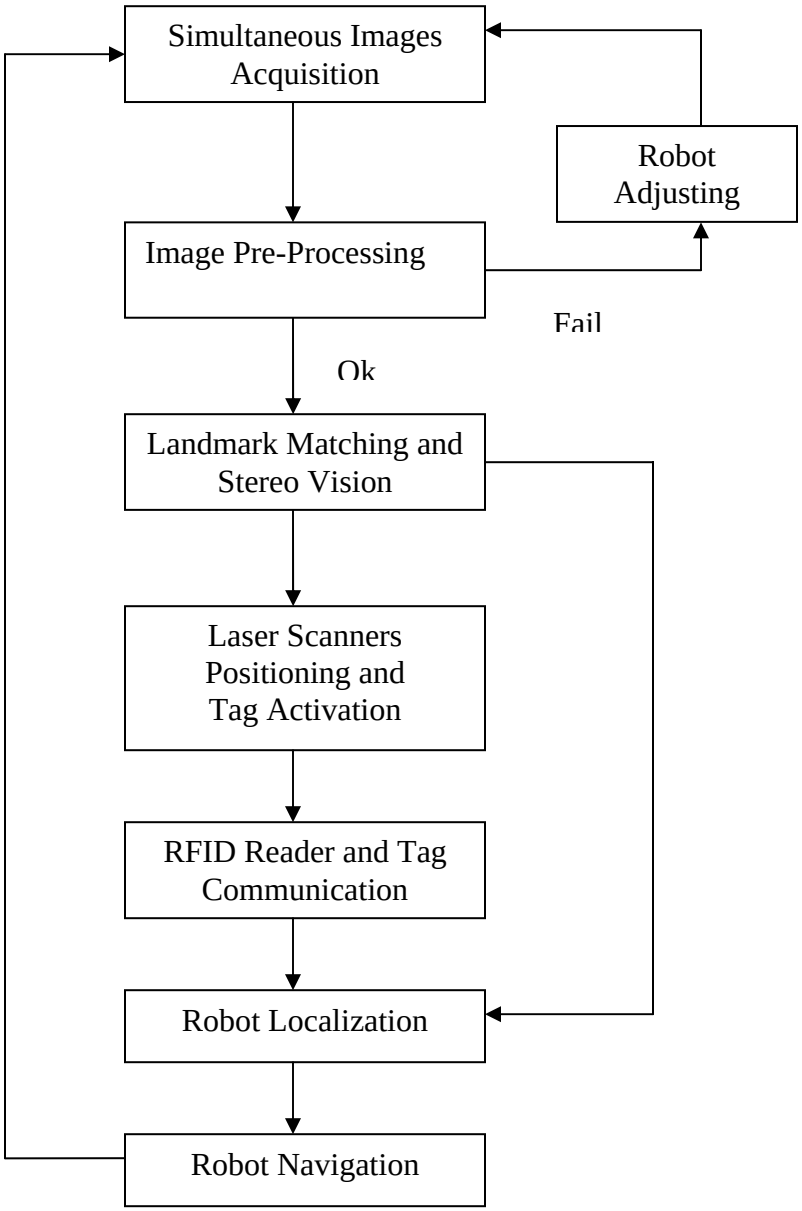

Fig. 3-11 Software Flow Chart

3.4 Experiments

In the previous section, we have mentioned the way to calibrate transformation between the camera frame and laser scanner frame. The transformation matrix is

$$ {}^{m}_{c}T = \begin{Bmatrix} -0.9886 & 0.0081 & -0.1506 & 142.3 \\ -0.1586 & -0.0049 & 0.9873 & 54.7 \\ 0.0081 & 1.0000 & -0.0021 & -0.0023 \\ 0 & 0 & 0 & 1 \end{Bmatrix}. \qquad 3.11 $$

With this transformation matrix ${}^{m}_{c}T$, we can convert any 3-D position derived by stereo vision into laser scanner reference frame in real time. Therefore, current system can autonomously capture and process the images, direct the laser to the tags and activate them, and navigate the robot moving forward. But duo to the recent arrival of RFID reader application programming interface (API), we haven't been able to integrate this library into our system. We can only manually receive the date from RFID tags. This is one of our most important future works to complete this project.

The accuracy of the whole system lies on the on accuracy of stereo vision. The localization accuracy of the proposed system has been tested extensively. Given the absolute positions of the tags in a global frame of reference, the localization accuracy for the mobile robot depends on the positioning accuracy of the tags relative to the robot. The positions of the tags relative to the mobile robot are obtained via stereo vision.

Relative distance measurement accuracy was tested. We fixed two LARFID tags on the ceiling with a certain distance which was carefully measured by hand. We grabbed two images simultaneously from the two cameras, corrected the distortion, segmented the bright spots and reconstructed the 3D positions of those two tags with respect to a camera frame. The distance between the two reconstructed tags was calculated and compared

with the known physical distance. The difference between them gives the error in relative distance measurement. This procedure was repeated many times with various distances between the tags and various distances between the tags and the cameras. The root-mean-square error was taken as the error measure (Fig. 3-13). The results show that the measurement error generally increases with respect to the distance under measurement. The underlying reason is that the resolution of the cameras decreases as the distance increases. We notice that the distance measurement of the experimental localization system has an accuracy of <20mm from a distance of about 7m, which, as we believe, is sufficient for most of the indoor mobile robot applications.

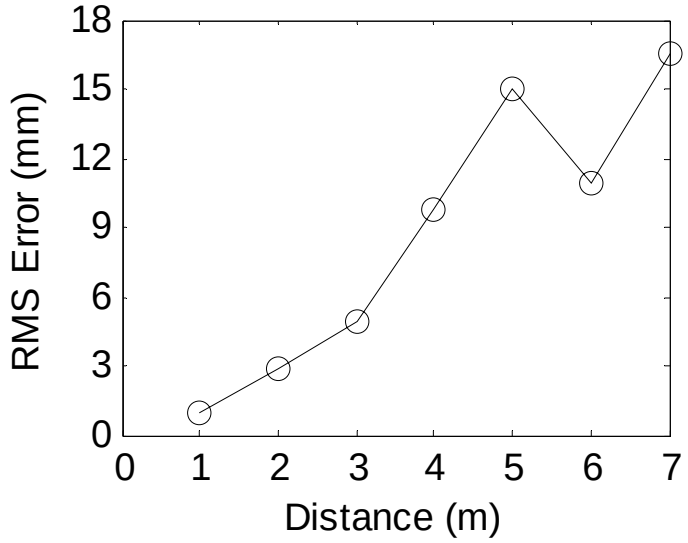

Fig. 3-12 Variation of root-mean-square (RMS) error in relative distance measurement with respect to the distance between the cameras and landmarks

3.5 Summary

This chapter introduces a new artificial landmark-based indoor localization system for mobile robots. The proposed localization technique is inspired by the existing GPS and RFID-based localization techniques. In the proposed system, a new type of RFID tag,

the laser-activated RFID tag, is designed and used as the artificial landmark. The LARFID tags function like the indoor equivalent of the GPS satellites. Stereo vision and LARFID are combined, together with trilateration or triangulation, to localize a mobile robot in an indoor environment. Feasibility study shows that the proposed system is promising to provide a robust and accurate indoor localization for mobile robots.

Some future research need to be conducted to complete the proposed system. The most important part is to integrate the RFID library into our controlling system, which makes this prototype fully autonomous. The design of the LARFID tag needs to be finely tuned to make it more compact. At the same time, a simultaneous localization and mapping algorithm is under consideration with the intention to automate the pre-installation process of the LARFID tags in an unexplored indoor environment and reduce the cost of installation and maintenance. Also under consideration is the multi-robot operation with the LARFID-based localization system.

Chapter 4 Robot Self-Localization based on a Single Image of Identified Landmarks

The previous chapter presents a new indoor localization system for mobile robot. The traditional trilateration method is used in the system. It can provide an accurate position based on the distance measurement using stereo vision. Nonetheless, the orientation is not easy to derive from the above information. Based on the idea of trilateration and triangulation, here we presented a new localization algorithm to discover the position and orientation of the robot. Further more, it only requires one single image. Before we introduce the algorithm, let us take a quick a look at the traditional localization algorithm.

Trilateration determines the position of an object based on simultaneous distance measurements from three or more known references. It is equivalent to finding the intersection of three or more spheres. Radio frequency signals and stereo vision sensors are usually used to measure the distance. This method is simple and straightforward. However, the disadvantage of this method is that it can not determine the orientation which is quite important for the localization of mobile robot.

Triangulation, instead of using distance measurements, uses bearing measurements among references to locate an object. The basic idea of triangulation is that in a plane containing the references and object, the object is located at the intersection of the circles each of which is determined by two references and the bearing between them. The structured features, such as doors and wall corners, are extracted from images. The 2D bearings of landmarks can be recovered from vertical edges.

So far, the research on triangulation has been focused on 2D situation, mainly because of the complicated mathematical formulation of the 3D triangulation. As a result, existing triangulation algorithms are constraint to 2D robot workspaces such as a flat-floor indoor environment. Meanwhile, though trilateration can successfully provide 3D robot positions, the orientation information usually can not be directly derived.

Combining trilateration with bearing estimation, we propose a novel method to recover the 3D position and orientation of a mobile robot from a single image of landmarks taken by an onboard camera. We assume that the involved landmarks are identified, and their positions are known from a map or a database. The robot localization consists of four steps:

1) In the linear camera perspective projection model, the visual angle between two landmarks is derived from their projections in the same image. With the distances between the optical center and the landmarks, the robot position is calculated using trilateration.
2) Distortion is corrected using “Plumb Bob” model [51]. The relative position of landmarks in the camera reference frame can be derived visual angles and the landmark positions.

3) The robot orientation is computed from the robot position, landmark positions and their projections.

4.1 Localization Algorithm

In this section, our localization algorithm is presented. The visual angle between two landmarks can be derived from their projections in the same image based on the perspective geometry. The distances between the optical center and the landmarks are computed from the constraint equations, as well as the relative position of landmarks in the camera frame. The robot position is determined using the principle of trilateration. A distortion rectified camera model is considered to improve the accuracy and estimate the relative position of landmarks in the camera reference frame. The robot orientation is then computed based on the camera model and landmark positions.

4.1.1 Robot Position Estimation

Our derivation is based on the pinhole camera model (Fig. 4-1). Here **P** denotes the optical center, xyz denotes the camera frame, UV denotes the image plane, and **C** denotes the image center of UV. In principle, the optical axis **PC** is perpendicular to the plane UV, and the length of **PC** is the focal length f. In addition, $\mathbf{L}_i$ where $i = (1,2,3)$, denotes the ith landmark with a known position defined in the world frame XYZ, and $\mathbf{l}_i$ denotes the projection of $\mathbf{L}_i$ in the image plane UV.

The camera model follows the principle of perspective projection. A detailed description of camera parameters and estimation algorithms can be found in a computer

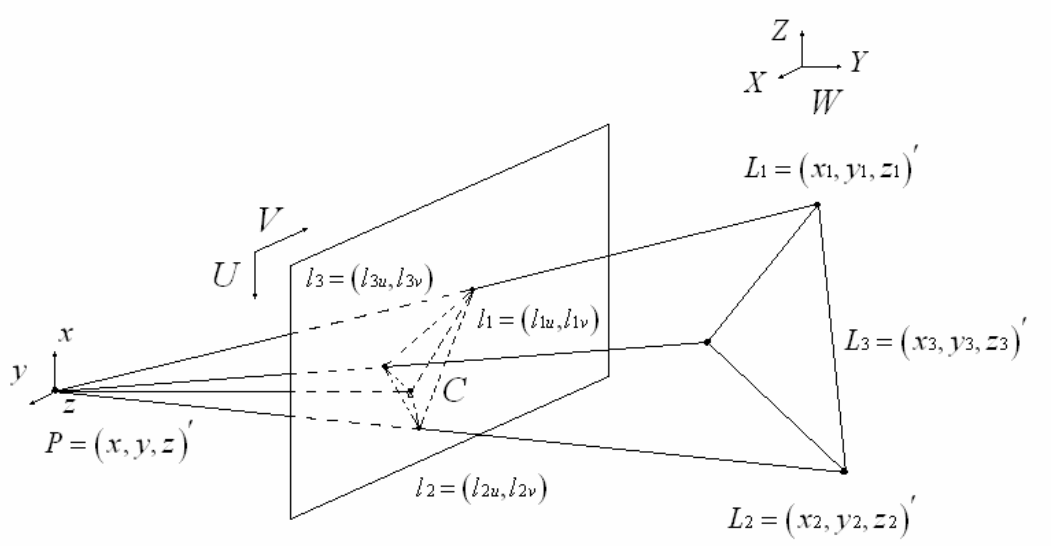

Fig. 4-1 Pinhole Camera Model

vision book such as [50]. A standard calibration toolbox can also be found at "http://www.vision.caltech.edu/bouguetj/calib_doc/".

Assuming that the camera is fixed on the mobile robot, robot positioning is equivalent to finding the position of the optical center $\mathbf{P} = (x, y, z)'$. Besides, the optical axis $\mathbf{PC}$ (Fig. 4-1) is chosen to represent the robot orientation in our method. We assume that the image center $\mathbf{C}$ and focal length f are known (in practice they can be obtained from camera calibration), and

$$\|\mathbf{PC}\| = f,\ \mathbf{C} = (c_u, c_v)',$$

$$\mathbf{l_1} = (l_{1u}, l_{1v})',\ \mathbf{l_2} = (l_{2u}, l_{2v})',\ \mathbf{l_3} = (l_{3u}, l_{3v})',$$

$$\mathbf{L_1} = (x_1, y_1, z_1)',\ \mathbf{L_2} = (x_2, y_2, z_2)',\ \mathbf{L_3} = (x_3, y_3, z_3)'.$$

In the image frame, all the lengths are measured in the unit of pixel. Since $\mathbf{PC}$ is perpendicular to the image plane UV, we have

$$\begin{aligned} \|\mathbf{Pl_i}\| &= \sqrt{f^2 + (l_{iu} - c_u)^2 + (l_{iv} - c_v)^2} \\ \|\mathbf{l_i l_j}\| &= \sqrt{(l_{iu} - l_{ju})^2 + (l_{iv} - l_{jv})^2} \end{aligned}, \tag{4.1}$$

where $i = (1,2,3)$, $j = (1,2,3)$, and $i \neq j$.

Applying the law of cosine to the triangle $\mathbf{l_i P l_j}$, we obtain

$$\angle \mathbf{l_i P l_j} = \cos^{-1}\left(\frac{\|\mathbf{Pl_i}\|^2 + \|\mathbf{Pl_j}\|^2 - \|\mathbf{l_i l_j}\|^2}{2\|\mathbf{Pl_i}\|\|\mathbf{Pl_j}\|} \right). \tag{4.2}$$

Moreover, in the pinhole camera model, the visual angle between landmarks i and j, $\angle \mathbf{L_i P L_j}$, is the same angle between their projections, $\angle \mathbf{l_i P l_j}$. For the convenience of expression, we define $\phi_{12} = \angle \mathbf{L_1 P L_2} = \angle \mathbf{l_1 P l_2}$ as the visual angle between landmarks 1 and 2. Visual angles ϕ_{13} and ϕ_{23} are defined similarly.

Applying the law of cosine to the triangles $\mathbf{L_1 P L_2}$, $\mathbf{L_1 P L_3}$ and $\mathbf{L_2 P L_3}$ respectively, we have

$$\begin{aligned} \|\mathbf{PL_1}\|^2 + \|\mathbf{PL_2}\|^2 - 2\cos\phi_{12}\|\mathbf{PL_1}\|\|\mathbf{PL_2}\| - \|\mathbf{L_1 L_2}\|^2 = 0 \\ \|\mathbf{PL_1}\|^2 + \|\mathbf{PL_3}\|^2 - 2\cos\phi_{13}\|\mathbf{PL_1}\|\|\mathbf{PL_3}\| - \|\mathbf{L_1 L_3}\|^2 = 0 \\ \|\mathbf{PL_2}\|^2 + \|\mathbf{PL_3}\|^2 - 2\cos\phi_{23}\|\mathbf{PL_2}\|\|\mathbf{PL_3}\| - \|\mathbf{L_2 L_3}\|^2 = 0 \end{aligned}, \tag{4.3}$$

where $\|\mathbf{L_i L_j}\|$ denotes the known distance between landmarks i and j, $\|\mathbf{L_i L_j}\| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}$, and $\|\mathbf{PL_i}\|$ denotes the unknown distance between the optical center and landmark i.

Newton's method is used to compute $\|\mathbf{PL_i}\|$ in Equ.4.3. Once the distances $\|\mathbf{PL_i}\|$ are obtained, the optical center can be located by solving a trilateration problem,

$$\begin{aligned}(x-x_1)^2+(y-y_1)^2+(z-z_1)^2&=\|\mathbf{PL_1}\|^2\\(x-x_2)^2+(y-y_2)^2+(z-z_2)^2&=\|\mathbf{PL_2}\|^2\\(x-x_3)^2+(y-y_3)^2+(z-z_3)^2&=\|\mathbf{PL_3}\|^2\end{aligned}. \qquad 4.4$$

Both closed form and iterative form solutions are available [30-33].

To sum up, the distances between the optical center and the landmarks are calculated from the visual angles and landmark positions using the law of cosine, and the robot position is estimated by trilateration. During this procedure, only the landmark positions and projections are used as input. No direct distance measurement is required.

A further inspection on Equ.4.3 and Equ.4.4 reveals the geometric meaning of this problem. Substituting Equ.4 into Equ.3, we can formulate a system of equations

$$\begin{aligned}&\cos^2\phi_{ij}\left[(x-x_i)^2+(y-y_i)^2+(z-z_i)^2\right]\left[(x-x_j)^2+(y-y_j)^2+(z-z_j)^2\right]\\&=\left[x^2+y^2+z^2-x(x_i+x_j)-y(y_i+y_j)-z(z_i+z_j)+x_ix_j+y_iy_j+z_iz_j\right]^2\end{aligned}. \qquad 4.5$$

In fact, if we generalize a circle in the 2D triangulation, determined by two landmarks and the mobile robot, into the 3D space, it becomes a spindle torus which is generated by revolving the circle around the chord connecting the two landmarks. The surface defined by Equ.4.5 is in fact a part of the spindle torus, generated by revolving the arc L_1PL_2 about the chord L_1L_2 (Fig. 4-2). Like the intersection of circles in the 2D space, the mobile robot localization problem in the 3D space becomes to find the intersection of three tori.

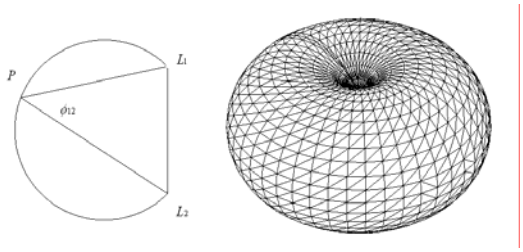

Fig. 4-2 Spindle Torus

4.1.2 Rectified Relative Position Estimation

This previous linear perspective model is quite straightforward and easy to understand. However, in the practice, the camera lens has always some distortion. The "Plumb Bob" model [51] is introduced briefly here.

Assuming P to be a point with a coordinate vector $X^c = [X_c, Y_c, Z_c]$ in the camera coordinate system. Let x_{n1} be the normalized projection on the image plane (Fig. 4-3). Then, in a linear model:

$$x_n = \begin{bmatrix} X_c / Z_c \\ Y_c / Z_c \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}. \tag{4.6}$$

If we assume $r^2 = x^2 + y^2$, after including the lens distortion the new normalized coordinate x_d is defined as follows:

$$x_d = \begin{bmatrix} x_d(1) \\ x_d(2) \end{bmatrix} = (1 + kc(1)r^2 + kc(2)r^4 + kc(5)r^6)x_n + dx \tag{4.7}$$

where dx is the tangential distortion defined as:

$$dx = \begin{bmatrix} 2kc(3)xy + kc(4)(r^2 + 2x^2) \\ kc(3)(r^2 + 2y^2) + 2kc(4)xy \end{bmatrix} \tag{4.8}$$

In equations (4.7) and (4.8), kc is a vector that contains both radial and tangential distortion coefficients. Once distortion is applied, the pixel coordinates $[u, v]$ of the projection on the image plane is:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = A \begin{bmatrix} x_d(1) \\ x_d(2) \\ 1 \end{bmatrix} \tag{4.9}$$

where A is the camera intrinsic parameter matrix as shown

$$A = \begin{bmatrix} f_u & \alpha & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}. \qquad 4.10$$

All the camera intrinsic parameter matrix and radial and tangential distortion coefficients can be calibrated using the toolbox provided above.

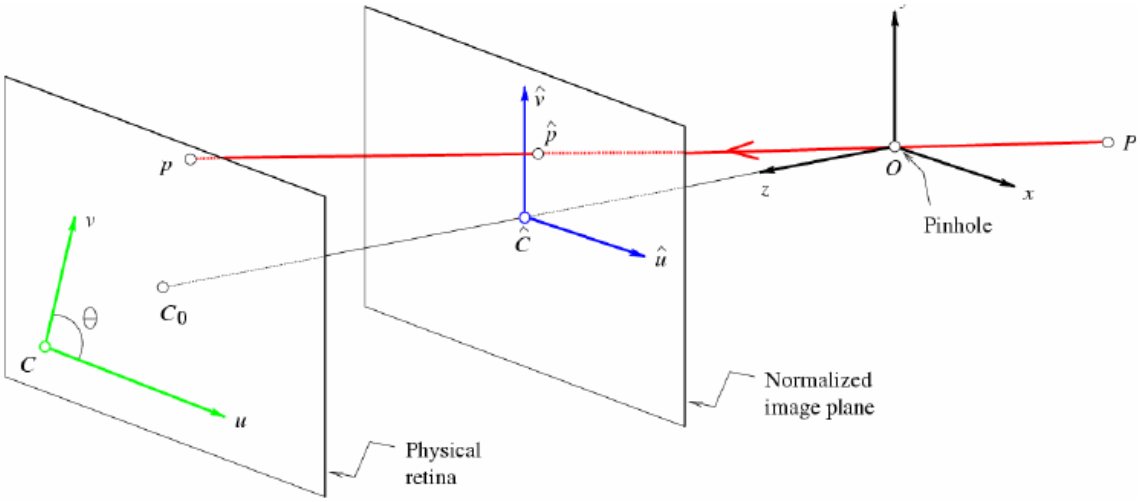

Fig. 4-3 Normalized Image Plane

From the initial image projection $\mathbf{l}_1 = (l_{1u}, l_{1v})'$, the normalized projection x_{n1} can be derived inversely according to the equation (4.7), (4.8) and (4.9). Now the direction vector of $\overrightarrow{PL_1}$ can be expressed as $\overrightarrow{n_{n1}} = [x_{n1}(1), x_{n1}(2), 1]'$. Adding one proportional parameter λ_1, we have

$$\overrightarrow{PL_1} = \lambda_1 \cdot [x_{n1}(1), x_{n1}(2), 1]'. \qquad 4.11$$

Since $\left\| \overrightarrow{PL_1} - \overrightarrow{PL_2} \right\| = \left\| \overrightarrow{L_1L_2} \right\|$, we have equation

$$\left\| \overrightarrow{PL_1} \right\|^2 + \left\| \overrightarrow{PL_2} \right\|^2 - 2\overrightarrow{PL_2} \bullet \overrightarrow{PL_2} = \left\| \overrightarrow{L_1L_2} \right\|^2,$$

$$\lambda_1^2 \left\|\overrightarrow{n_{n1}}\right\|^2 + \lambda_2^2 \left\|\overrightarrow{n_{n2}}\right\|^2 - 2\lambda_1\lambda_2 \overrightarrow{n_{n1}} \bullet \overrightarrow{n_{n2}} = \left\|\overrightarrow{L_1L_2}\right\|^2 . \quad 4.12$$

This equation is actually consistent with equation 4.3. With another two equations, we can find the proper λ_i for those vectors using Newton's method.

Now the relative position of landmarks in the camera reference frame can be expressed as $\overrightarrow{PL_i} = \lambda_i \cdot [x_{ni}(1), x_{ni}(2), 1]'$. To verify our method, we will compare our method with stereo vision in the experiment part.

To recover the position of camera optical center in the world frame, we have to apply the trilateration method discussed in the previous section.

4.1.3 Orientation Estimation

In a 3D environment, an orientation can be represented by a directional vector. Since the camera is fixed on the mobile robot, finding the robot orientation is equivalent to finding the camera orientation. Therefore, the optical axis $\overrightarrow{\mathbf{PC}}$ (Fig. 4-1) is chosen to represent the robot orientation.

In the world frame XYZ, the lines $\overrightarrow{\mathbf{PC}}$ and $\overrightarrow{\mathbf{PL_i}}$ can be expressed as following

$$\text{line } \mathbf{PC}: \frac{x - x_p}{p_x} = \frac{y - y_p}{p_y} = \frac{z - z_p}{p_z}, \quad 4.13$$

$$\text{line } \mathbf{PL_i}: \frac{x - x_p}{x_i - x_p} = \frac{y - y_p}{y_i - y_p} = \frac{z - z_p}{z_i - z_p}, \quad 4.14$$

where $(x_p, y_p, z_p)'$ is the estimated robot position $\mathbf{P}$ (the position of the camera optical center), $(x_i, y_i, z_i)'$ is the global position of landmark $\mathbf{L_i}$, and $(p_x, p_y, p_z)'$ is the directional vector of $\overrightarrow{\mathbf{PC}}$ which needs to be determined.

Corresponding to Equ.4.13 and Equ.4.14, the angle between $\overrightarrow{\mathbf{PC}}$ and $\overrightarrow{\mathbf{PL_i}}$ satisfies the following relationship

$$\phi_{ic} = \cos^{-1} \frac{p_x(x_i - x_p) + p_y(y_i - y_p) + p_z(z_i - z_p)}{\sqrt{p_x^2 + p_y^2 + p_z^2}\sqrt{(x_i - x_p)^2 + (y_i - y_p)^2 + (z_i - z_p)^2}}. \quad 4.15$$

By normalizing the directional vector $\mathbf{PC}$, we have $p_x^2 + p_y^2 + p_z^2 = 1$, and Equ.4.8 can be rewritten as

$$\phi_{ic} = \cos^{-1} \frac{p_x(x_i - x_p) + p_y(y_i - y_p) + p_z(z_i - z_p)}{\sqrt{(x_i - x_p)^2 + (y_i - y_p)^2 + (z_i - z_p)^2}}. \quad 4.16$$

On the other hand, we can calculate ϕ_{ic} using the method mentioned in previous section. For the linear perspective model, we have

$$\phi_{ic} = \cos^{-1}\left(\frac{\|\mathbf{Pl_i}\|^2 + \|\mathbf{PC}\|^2 - \|\mathbf{l_iC}\|^2}{2\|\mathbf{Pl_i}\|\|\mathbf{PC}\|}\right) = \cos^{-1}\frac{f}{\|\mathbf{Pl_i}\|}, \quad 4.17$$

where we have used the facts that $\|\mathbf{Pl_i}\| = \sqrt{f^2 + (l_{iu} - c_u)^2 + (l_{iv} - c_v)^2}$, $\|\mathbf{PC}\| = f$, and $\|\mathbf{l_iC}\| = \sqrt{(l_{iu} - c_u)^2 + (l_{iv} - c_v)^2}$. The quantities f, l_{iu}, l_{iv}, c_u and c_v can be obtained directly from the image.

According to Equ.4.16 and Equ.4.17, we have

$$p_x(x_i - x_p) + p_y(y_i - y_p) + p_z(z_i - z_p) = \frac{f\sqrt{(x_i - x_p)^2 + (y_i - y_p)^2 + (z_i - z_p)^2}}{\sqrt{f^2 + (l_{iu} - c_u)^2 + (l_{iv} - c_v)^2}}. \quad 4.18$$

The system of equations represented by Equ.4.18 with $i = (1,2,3)$ can be written in the following matrix form

$$\overrightarrow{\mathbf{PC}} = \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} = \begin{pmatrix} x_1 - x_p & y_1 - y_p & z_1 - z_p \\ x_2 - x_p & y_2 - y_p & z_2 - z_p \\ x_3 - x_p & y_3 - y_p & z_3 - z_p \end{pmatrix}^{-1} \begin{pmatrix} \dfrac{f\sqrt{(x_1 - x_p)^2 + (y_1 - y_p)^2 + (z_1 - z_p)^2}}{\sqrt{f^2 + (l_{1u} - c_u)^2 + (l_{1v} - c_v)^2}} \\ \dfrac{f\sqrt{(x_2 - x_p)^2 + (y_2 - y_p)^2 + (z_2 - z_p)^2}}{\sqrt{f^2 + (l_{2u} - c_u)^2 + (l_{2v} - c_v)^2}} \\ \dfrac{f\sqrt{(x_3 - x_p)^2 + (y_3 - y_p)^2 + (z_3 - z_p)^2}}{\sqrt{f^2 + (l_{3u} - c_u)^2 + (l_{3v} - c_v)^2}} \end{pmatrix}, \quad 4.19$$

which gives the direction of the optical axis $\overrightarrow{\mathbf{PC}}$ (equivalent to the robot orientation as stated at the beginning of first section).

With the nonlinear distortion model, we have principle point $\mathbf{C} = (u_0, v_0)'$. Similarly, the directional vector of $\overrightarrow{\mathbf{PC}}$ in the camera frame is $\overrightarrow{n_c}$ = [$x_c(1)$, $x_c(2)$, 1]′ and

$$\phi_{ic} = \cos^{-1}\left(\frac{\overrightarrow{n_c}' \bullet \overrightarrow{n_{n1}}}{\|\overrightarrow{n_c}\|\|\overrightarrow{n_{n1}}\|}\right) = \cos^{-1}\frac{x_{n1}(1)x_c(1) + x_{n1}(2)x_c(2) + 1}{\sqrt{x_{n1}(1)^2 + x_{n1}(2)^2 + 1}\sqrt{x_c(1)^2 + x_c(2)^2 + 1}}. \quad 4.20$$

Therefore,

$$p_x(x_i - x_p) + p_y(y_i - y_p) + p_z(z_i - z_p) = \frac{\overrightarrow{n_c}' \bullet \overrightarrow{n_{n1}}}{\|\overrightarrow{n_c}\|\|\overrightarrow{n_{n1}}\|}. \quad 4.21$$

$$\overrightarrow{\mathbf{PC}} = \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} = \begin{pmatrix} x_1 - x_p & y_1 - y_p & z_1 - z_p \\ x_2 - x_p & y_2 - y_p & z_2 - z_p \\ x_3 - x_p & y_3 - y_p & z_3 - z_p \end{pmatrix}^{-1} \begin{pmatrix} \dfrac{\overrightarrow{n_c}' \bullet \overrightarrow{n_{n1}}}{\|\overrightarrow{n_c}\|\|\overrightarrow{n_{n1}}\|} \\ \dfrac{\overrightarrow{n_c}' \bullet \overrightarrow{n_{n2}}}{\|\overrightarrow{n_c}\|\|\overrightarrow{n_{n2}}\|} \\ \dfrac{\overrightarrow{n_c}' \bullet \overrightarrow{n_{n3}}}{\|\overrightarrow{n_c}\|\|\overrightarrow{n_{n3}}\|} \end{pmatrix}. \quad 4.22$$

4.2 Simulation Analysis

This section carries out an error analysis of our localization method, based on the simulation with erroneous inputs, which provides the insight on how to improve the localization accuracy and robustness using our method.

4.2.1 Simulation Setting

The global positions and image projections of the landmarks are the basic inputs to our algorithm. In our simulation, we assume that they may have uncorrelated zero-mean Gaussian noises. The statistics of the output errors in the robot position and orientation estimation is obtained from extensive numerical computation.

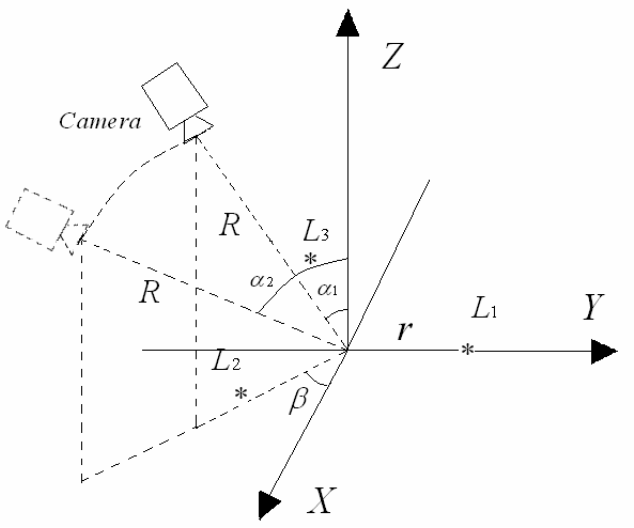

Fig. 4-4 Simulation Setting

In the simulation, we put three landmarks (star symbols) at vertices of an equilateral triangle inscribed in a circle with radius r (Fig. 4-4). r is also referred as the baseline. The circle is laid on the XY plane of the world frame and centered at the origin. Then the positions of the landmarks can be expressed as

$$\mathbf{L}_1 = (0, r, 0)', \mathbf{L}_2 = \left(r\sqrt{3}/2, -r/2, 0\right)', \mathbf{L}_3 = \left(-r\sqrt{3}/2, -r/2, 0\right)'.$$

We let the camera move on a spherical surface with a radius of R and centered at the origin of the world frame. We also keep the optical axis of the camera pointing to the origin of the world frame. A tilt angle α is defined as the angle between the Z axis of the world frame and the camera optical axis. A pan angle β is defined as the angle between the X axis of the world frame and the XY projection of the optical axis. With fixed R and α, the camera trajectory is a circle. By varying R, α and β, we can have an understanding of how these factors affect the localization accuracy.

In the following, we will examine the effect of six parameters, i.e. standard deviation σ_l of landmark position error, standard deviation σ_p of projection segmentation error, baseline r, sphere radius R, tilt angle α and pan angle β, on the robot localization error.

4.2.2 Landmark Position Error

We have pointed out two kinds of errors above, the landmark position error and the projection segmentation error. Position Dilution of Precision (PDOP) is used here as a performance index to evaluate the effect of the landmark position error. PDOP is a GPS term used in geometrics engineering to describe the geometric strength of satellite configuration. It is a unit less measure that reflects positioning accuracy, defined as

$$PDOP = \frac{\sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}}{\sigma_l}, \qquad 4.23$$

where σ_x^2, σ_y^2, and σ_z^2 are the standard deviations of the robot position estimation error in x, y and z direction respectively, and σ_l is the standard deviation of the landmark position error. To evaluate the accuracy of this method in orientation estimation, we extend the concept of PDOP into Orientation Dilution of Precision (ODOP), with a unit of degree per millimeter,

$$ODOP = \frac{\sigma_a}{\sigma_l}, \qquad 4.24$$

where σ_a is the standard deviation of the robot orientation estimation error which is defined by the angle between the actual robot orientation and the error-corrupted one.

Simulation has been carried out to examine the effect of the landmark position error. The Gaussian noises are added into the landmark positions, $\mathbf{L}_i = \mathbf{L}_{i0} + \boldsymbol{\delta}\mathbf{L}_i$, where $\boldsymbol{\delta}\mathbf{L}_i$ denotes the Gaussian random error with a standard deviation σ_l, and $\mathbf{L}_{i0}$ denotes the actual value of the landmark positions. With different α ranging from 5° to 80°, we can evaluate both PDOP and ODOP at different configurations. Forty camera positions, decided by β, are sampled on the circle defined by each R and α.

Fig. 4-5 is generated by setting σ_l =1mm, r=1000mm and R=1500mm and calculating the PDOP and ODOP at various camera positions on the sphere. It turns out that the PDOP (and ODOP) at different camera positions on a circle are close to each other. Thus, a mean value of these DOP is presented for each circle along with the standard deviation (Fig.4-5). The other curves in Fig.4-5 are generated with the same parameters except σ_l =3mm. Comparing these two curves, we notice they have almost the same PDOP and ODOP. In fact, we have also seen similar results with other σ_l values. Therefore, we conclude that the standard deviations of the robot position

estimation error and orientation estimation error are linearly proportional to that of the landmark position error. The accuracy of the robot localization can be improved with accurately mapped landmark positions.

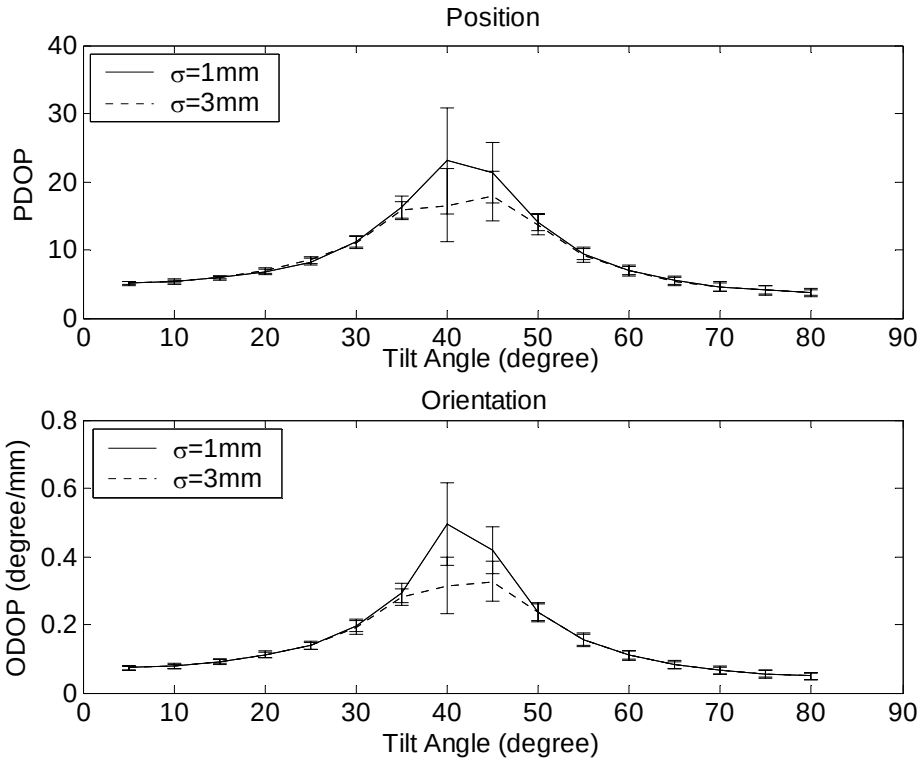

Fig. 4-5 PDOP and ODOP with different standard deviations of landmark position error

Another simulation has been carried out to examine the effect of different R values while having fixed ratio r/R. We have already known the PDOP and ODOP results generated with r/R=1000/1500 and σ_l =1mm in Fig.4-5. Another example is generated from r=2000mm, R=3000mm and σ_l =1mm. The comparison between these curves in Fig. 4-6 shows that they have the same results of PDOP. However, the value of ODOP with r/R=1000/1500 is twice as big as that with r/R=2000/3000. In fact, we have obtained the similar results from different R but with the same r/R. Clearly, with the fixed ratio, the

position error is constant while the orientation error is inversely proportional to the radius R within small angles. A rough explanation can be derived by differentiating Equ.4.17

$$\Delta\phi_{ic} = \Delta\left\|\mathbf{Pl_i}\right\|\cos\phi_{ic} / \left\|\mathbf{Pl_i}\right\|\sin\phi_{ic} . \qquad 4.25$$

Since ratio r/R is fixed, $\left\|\mathbf{Pl_i}\right\|$ and ϕ_{ic} doesn't change, $\Delta\phi_{ic}$ is proportional to the $\Delta\left\|\mathbf{Pl_i}\right\|$. Moreover, according to the camera model in Fig. 4-1, we can roughly estimate:

$$\Delta\left\|\mathbf{Pl_i}\right\| \approx \frac{\Delta\left\|\mathbf{L_i}\right\|}{R} f . \qquad 4.26$$

By combining Equ.4.25 and Equ.4.26, we can say that visual angle errors are approximately inversely proportional to R with small angles. According to the linear transformation between the visual angles and robot orientation in Equ.4.19, R will have the same effect on the final orientation error.

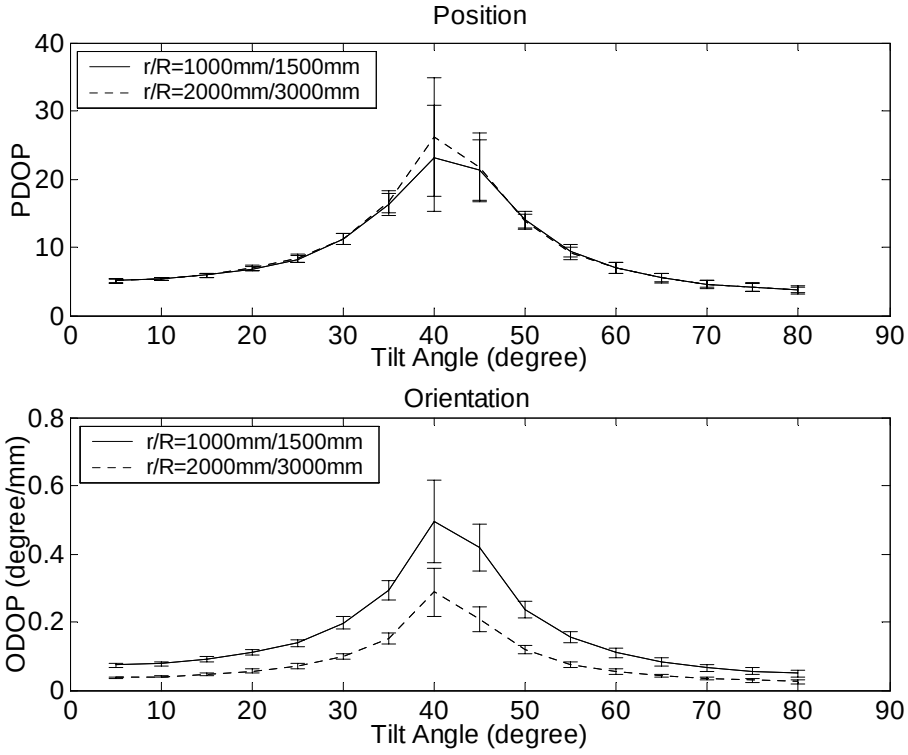

Fig. 4-6 PDOP and ODOP with the same ratio r/R

The effect of variable r/R is more complicated. A further discussion about the trend of PDOP and ODOP with different r/R is provided in later section.

4.2.3 Segmentation Error of Image Projections

Segmenting landmark projections from the image may introduce some random segmentation error. In the simulation, the Gaussian noises are added to the positions of the landmark projections in the image, $\mathbf{l}_i = \mathbf{l}_{i0} + \delta\mathbf{l}_i$, where $\delta\mathbf{l}_i$ is the Gaussian random error with standard deviation σ_p, and $\mathbf{l}_{i0}$ is the actual positions of the landmark projections in the image.

Since the standard deviation of the projection error is in the unit of pixel, the PDOP is no longer available in this situation. A new measure which we call Pseudo-Position Dilution of Precision (PPDOP) is defined in the unit of millimeter per pixel

$$PPDOP = \frac{\sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}}{\sigma_p}. \qquad 4.27$$

Similarly, Pseudo-Orientation Dilution of Precision (PODOP) is defined in the unit of degree per pixel,

$$PODOP = \frac{\sigma_a}{\sigma_p}. \qquad 4.28$$

Fig. 4-7 is generated by setting r=1000mm, R=1500mm, and σ_p =1 pixel. Fig.4-7 and Fig.4-5 show the different results with respective measures.

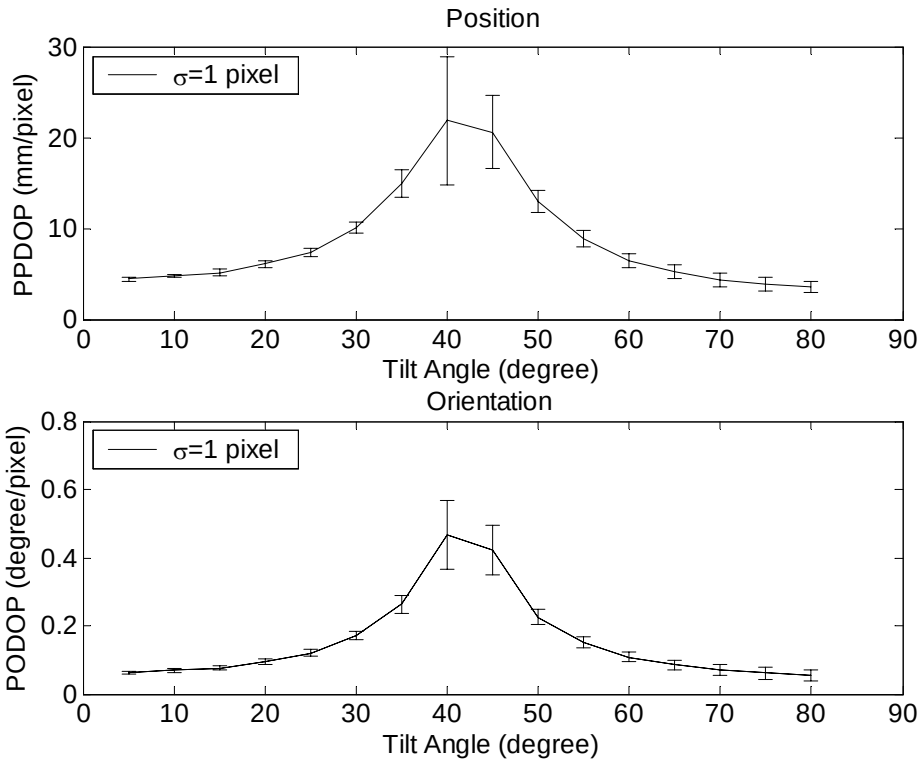

Fig. 4-7 PPDOP and PODOP with image projection error

These two results are actually unified if we take Effective Pixel Size (EPS) into account. EPS is defined as a measure of the length that one pixel of the image occupies in a plane that is perpendicular to the optical axis. The unit of EPS is millimeter per pixel. It is estimated from the pinhole camera model. The segmentation error of the landmark projection can be converted to the unit of millimeter using $\sigma_p \approx \sigma_l / EPS$, and we have

$$\begin{aligned} PDOP &\approx PPDOP / EPS \\ ODOP &\approx PODOP / EPS \end{aligned} . \qquad 4.29$$

The EPS at R=1500mm is around 0.95 pixel/mm. A conversion between Fig.4-5 and Fig.4-7 actually matches the relationship defined in Equ.4.29.

Fig.4-8 is generated with the same r/R while assuming r =2000mm, R=3000mm. The resulting PPDOP is twice as big as that in Fig.7 while the PODOP is almost the same.

The reason of doubled PPDOP is that EPS has also been doubled. For the PODOP, the effect of EPS is balanced out by that of R which has been explained in previous section. Therefore, we can conclude that at the fixed ratio r/R, the variation in distance only affects the robot localization accuracy due to the change in EPS.

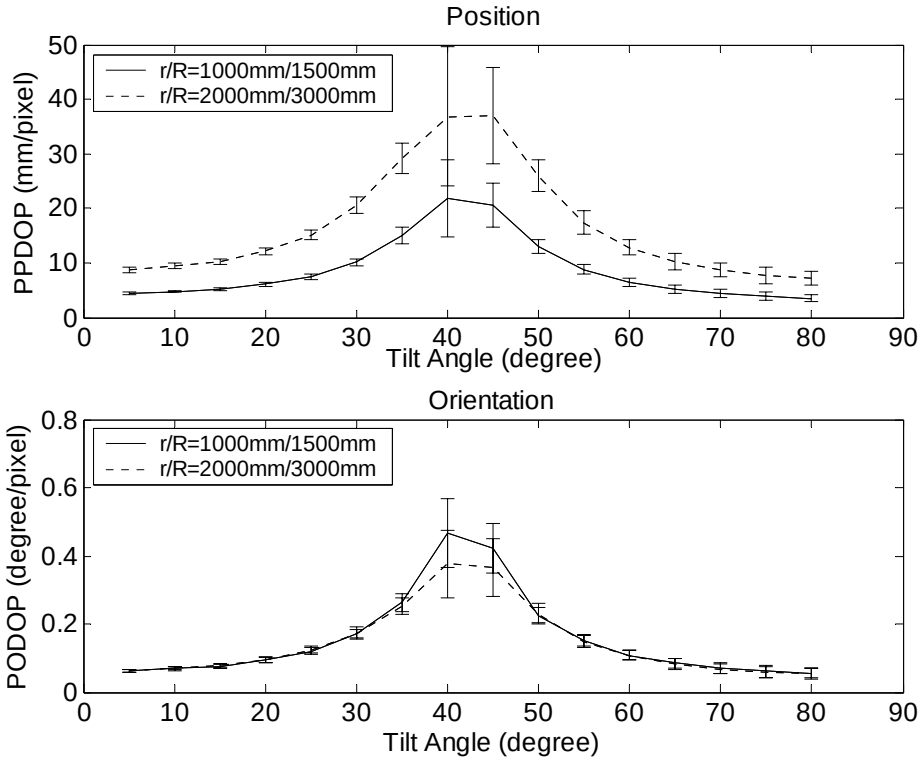

Fig. 4-8 PPDOP and PODOP with the same ratio r/R

4.2.4 Ratio r/h

Since all the figures above choose the same ratio r/R=1000/1500, they present the same trend of PDOP and ODOP. We have also examined the effect of different r/R values. Fig.4-9 includes the results of r/R=1000/1000, r/R=1000/1500 and r/R=1000/2600 respectively. Since the landmark position error and the projection segmentation error can be unified, only PDOP and ODOP are considered in these settings.

A comparison among these three curves shows that a smaller r/R results in a larger error in both robot position and orientation estimation. Therefore, to achieve high localization accuracy, a big r/R is desirable.

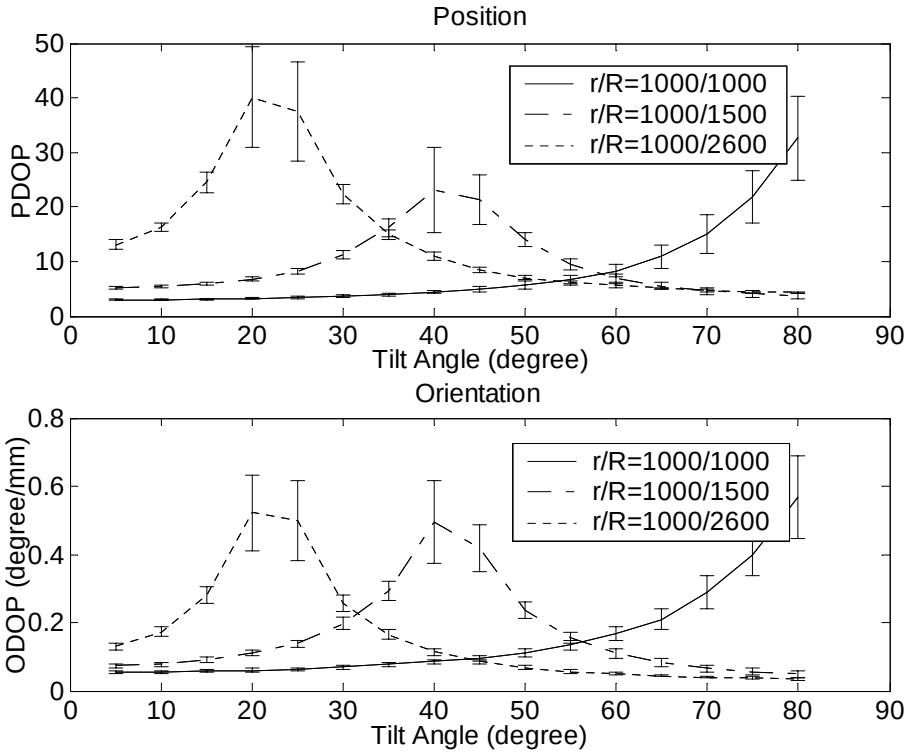

Fig. 4-9 PDOP and ODOP with different r/R

One may notice that there are obvious peaks of PDOP and ODOP at some tilt angle in each figure. The simulation results show that the peak usually occurs when the camera projection on the XY plane is located at or close to the circum-circle of the equilateral triangle, which means the tilt angle $\alpha = \sin^{-1}(r/R)$. The PDOP and ODOP can be extremely big around the peak. To achieve accurate robot localization, these camera poses should be avoided in the practice. An analytic explanation of this phenomenon is not available yet. It will be investigated in our future work.

4.3 Experiment

After this simulation was implemented and the result was published, we later on find a way to calculate the relative position of landmarks in the camera reference frame. Therefore, you don't see a separate analysis of the relative position accuracy. At that time, we mainly focus on the solution and accuracy of trilateration. The final position of camera in the world frame is the important result that navigates the robot moving forward. The relative positions of landmarks are just intermediate results.

If you think carefully, the influence of our method is already embedded in the simulation results. This simulation results give us the way how to improve the accuracy and distribute the landmarks. The result of dilution of position is also consistent with others' conclusion [30].

In the real experiment, the situation is more complicated than the simulation. The calibrated camera parameters have error, the lens distortion has to be considered, and the position and segmentation error is unknown. Most importantly, both the relative position of landmarks and the real position of camera can not be measured. What can we do to verify our method?

First, we will take measure the relative error of our method. Although we don't the real relative position of the landmarks, we can know the trajectory length if we translate those landmarks. If we keep the camera fixed during the landmark movement, the recovered distance of the two pair of positions should be equal to the trajectory length. The motion platform in the mechatronics lab can provide um accuracy movement which is good enough for our experiment. Two variables, baseline and distance, are tested in our experiment. At each set of variables, the landmarks are moving around as long as they are

inside the view of the camera and totally, we made 15 groups of data. The final results are presented in table 4.1.

Table 4.1 Relative Position Error

Baseline(mm) / Distance(m)	150mm		300mm	
	Mean	STD	Mean	STD
1.2m	0.0048	0.013	N/A	
2.4m	0.0033	0.029	0.0041	0.0088

The distance and baseline measurement are roughly estimated. They are not as accurate as that in the simulation. The mean and standard devastation are statistics of difference between the measured distance and trajectory length. The standard devastation is an important indicator of the accuracy. Both the value of mean and standard devastation is expressed in percentage. For example, at 1m distance and 150mm baseline, 100mm movement may cause a relative error with 0.48mm mean and 1.3mm standard devastation.

From the table, obviously, we can see that the long distance will cause big error when the baseline is fixed. On the other hand, the bigger baseline will eliminate the influence of longer distance. A good choice in the real situation is to keep the baseline and distance proportional to reduce the error caused by the longer distance.

The second experiment we did is to compare our method with stereo vision method. Although we know that even stereo vision can not give the true relative position of

landmarks, we assume it is “accurate”. We use two simultaneously captured images to recover the relative position. Also, we recover the same landmarks using one of these two images. Ideally, their values should be the same. But actually they are not. Still we use the mean and standard deviation to estimate the difference between those relative positions. The final results are presented in table 4.2.

Table 4.2 Relative Position Comparison between Stereo Vision and Our Method

Baseline(mm) / Distance(m)		150mm		300mm	
		Mean(mm)	STD(mm)	Mean(mm)	STD(mm)
1.2m	X	0.044	0.543	N/A	
	Y	-0.228	0.336		
	Z	-0.242	4.252		
2.4m	X	-0.267	0.607	0.165	0.749
	Y	0.164	0.292	0.030	0.664
	Z	-1.981	6.576	-0.871	5.930

The setting of experiment is similar as the previous one, except that we record the difference in the X, Y, and Z direction in millimeter. Clearly, the STD in the Z direction is much bigger than those in the X and Y directions. It is easy to understand if we take a look at the camera frame. In the camera frame, the coordinate value in the Z direction has bigger value compared with those in the X and Y directions. Therefore, the standard

deviation will be bigger. If we put the value into percentage, their values should be comparable.

Similarly, the shorter distance and longer baseline will reduce the difference between the results of our method and those of stereo vision. Most importantly, the standard deviations of those differences are acceptable, which means our method is comparable with stereo vision method.

As we presented annotation in the previous section, the final robot position is the result that interests us. To recover the absolute position, we will inevitably introduce the trilateration error discussed in the simulation. We use the relative position above derive from stereo vision and our method to recover the absolute position respectively. The mean and standard deviation of the difference between absolute positions are listed in the table 4.3.

Table 4.3 Absolute Position Comparison between Stereo Vision and Our Method

Baseline(mm) / Distance(m)		150mm		300mm	
		Mean(mm)	STD(mm)	Mean(mm)	STD(mm)
1.2m	X	21.02	18.49	N/A	
	Y	9.07	8.71		
	Z	2.42	10.66		
2.4m	X	22.99	33.54	2.34	50.24
	Y	13.79	23.85	-14.40	26.71
	Z	-30.30	23.84	24.95	15.93

As we can see, the absolute position statistics are much bigger than relative positions. This phenomenon is caused by the intrinsic error amplification of trilateration method. It can be simply illustrated using following diagram. As we know, the geometric meaning of trilateration is finding the intersection of circles. In the Fig 4-10, the distance between centers of circles is small compared with the radius, small change in radius will cause big change in the intersection position, which is the exactly the case of our experiment: in the view of camera, the baseline is always smaller than the distance between camera and landmarks. Also, the amplified error is no longer restricted in the Z direction. We have pointed out the way to improve the accuracy of trilateration by increasing baseline and reduce the distance.

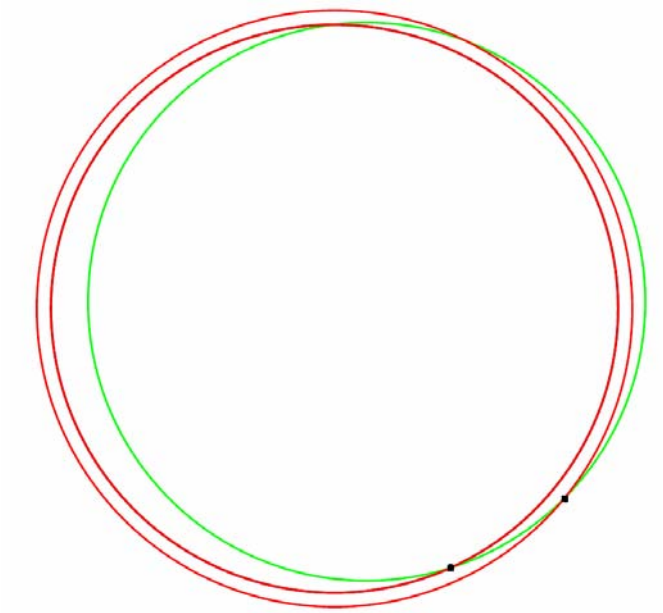

Fig 4-10 Intersection change with radius change

Finally, we are going to verify our derived orientation computation. From the equation 4.22, we know that the orientation accuracy largely relies on the accuracy of trilateration result. Using the different trilateration results derived from stereo vision and

our method respectively, we can compare the influence of trilateration error to the orientation. Since the orientation is a directional vector, we use the included angle to record the orientation difference. For the same position, ideally the orientation vector should be parallel. The comparison results are listed in the table 4.4. All the values are expressed in the unit of degrees.

Table 4.4 Orientation Comparison between Stereo Vision and Our Method

Baseline(mm) / Distance(m)	120mm		240mm	
	Mean(degree)	STD(degree)	Mean(degree)	STD(degree)
1.2m	0.0045	0.0035	N/A	
2.4m	0.0022	0.0032	6.7860e-004	7.6674e-004

As you can see, the mean and standard deviation value are quite small, which means that the trilateration error has quite limited influence to the orientation computation. Besides that, we can also see that with the same distance-to-baseline ratio, the orientation error will be reduced by the longer distance. This conclusion is consistent with the simulation.

When we translate the camera using the motion platform, the orientation will keep the same, i.e. be parallel. This provide us another way testify the accuracy of our orientation computation. This experiment is more meaningful because it will give us a dynamic view of orientation accuracy when the robot is moving. Using our method, we

estimation its orientation before and after translation movement. Table 4.5 gives a full record of its difference with difference parameter settings.

Table 4.5 Orientation Comparison in Translation Movement

Baseline(mm) / Distance(m)	120mm		240mm	
	Mean(degree)	STD(degree)	Mean(degree)	STD(degree)
1.2m	0.0496	0.0068	N/A	
2.4m	0.0252	0.0055	0.0432	0.0079

The orientation error caused by translation movement is within our acceptable range. And it is also constant with the simulation result.

4.4 Summary

This chapter introduces a novel effective localization algorithm for mobile robots based on one single image of a few identified landmarks taken by an onboard camera. The visual angle between two landmarks can be derived from their projections in the same image. The lens distortion is also considered in the model. The relative position of the landmarks can be calculated from the visual angles and the known landmark positions. The robot position can then be determined using the principle of trilateration. Finally, the robot orientation is computed from the robot position, landmark positions and their projections.

Simulation study has been conducted to evaluate the performance of our method with erroneous measurements. To achieve accurate robot localization, accurate landmark position estimation and projection segmentation are desired. Also, big baseline-distance ratio should be considered. Moreover, the camera poses associated with PDOP/ODOP peaks should be avoided.

In the experiment part, we verified the relative position and absolute error which is consistent with the simulation result. The comparison between the stereo vision and our method show that these two methods are comparable. Finally, the orientation experiment result shows that it can provide high accuracy data.

Chapter 5 Modeling of Two-axis Optical Laser Scanner

In the chapter 3, we have encountered the problem of converting camera frame into laser scanner frame. To control the output direction of laser beam, we have to find the relationship between the output orientation and mirror angles. The complexity of this problem lies on the laser scanner model. There are two mirrors in the 2-axis optical scanner, each of which can rotate about its revolute axis. The two mirrors rotate independently from each other, and their revolute axes are perpendicular to each other. Whenever the mirror rotates, the reflected laser beam will change its direction. Besides, the transformation between the mirror coordinates adds much more complication into the modeling. A local frame of reference can be attached to each mirror and rotate together with the mirror (Fig. 5-1).

5.1 Reference Frames of the Two-mirror System

To define the frames of reference in the 2-mirror system, we at first define the x axis of a mirror-fixed frame to be coincident with the revolute axis of the mirror. In this way, we have x_1 and x_2 for the two mirror frames respectively. Next we find the mutual

perpendicular, L, between x_1 and x_2. The intersection point between L and x_1 defines the origin of frame 1, O_1, and the intersection point between L and x_2 defines the origin of frame 2, O_2. Then the z axis of a mirror frame is defined as being perpendicular to the mirror plane and passing through the origin, i.e. z_1 for frame 1 and z_2 for frame 2. The y axis of a mirror frame can then be obtained by the rule of right hand. As a result, the xy plane coincides with the mirror plane, and the z axis points to the normal direction of the mirror plane. As shown in Fig.1, $x_1y_1z_1$ denotes frame 1, and $x_2y_2z_2$ denotes frame 2.

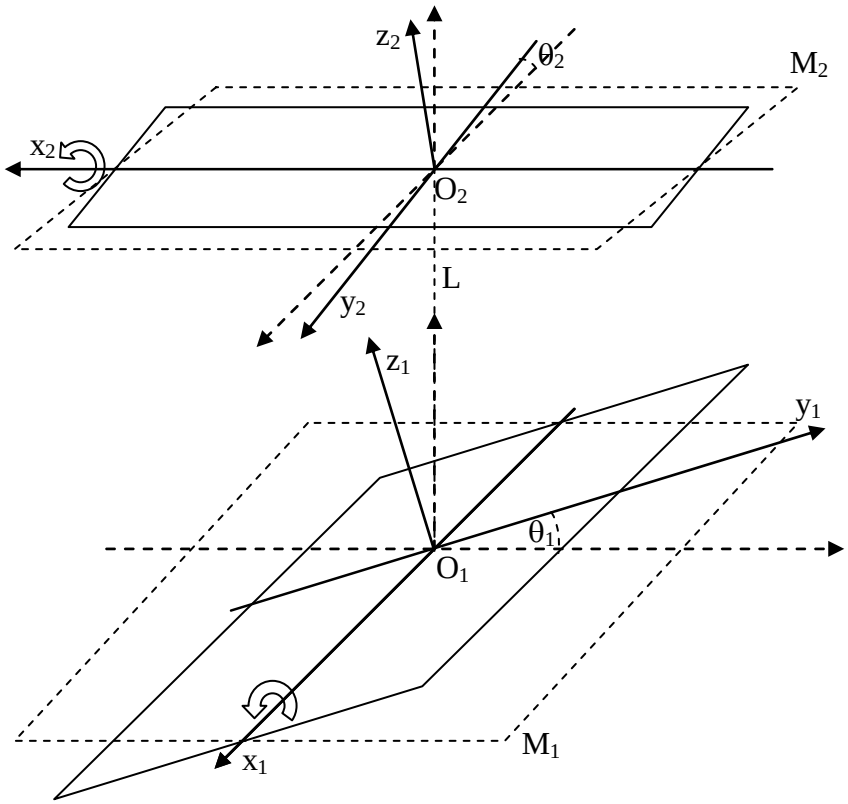

Fig. 5-1 Frames of reference in the two-mirror system

Without loss of generality, we assume that when the mirrors are at their neutral positions ($\theta_1=\theta_2=0$), the two mirror planes are physically parallel to each other (Fig. 5-2).

In this special case, it is clear that z1 coincides with z2. Moreover, we know that x1 is perpendicular to x2. Therefore, in this special case, the transformation between x1y1z1 and x2y2z2 can be obtained as.

$$ {}^{1}_{2}\mathbf{T} = \mathbf{R}_z(-\frac{\pi}{2})\mathbf{T}_z(a) = \mathbf{T}_z(a)\mathbf{R}_z(-\frac{\pi}{2}) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & a \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad 5.1 $$

where a is the distance between the two mirror planes, and

$$ {}^{2}_{1}\mathbf{T} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -a \\ 0 & 0 & 0 & 1 \end{pmatrix}. \qquad 5.2 $$

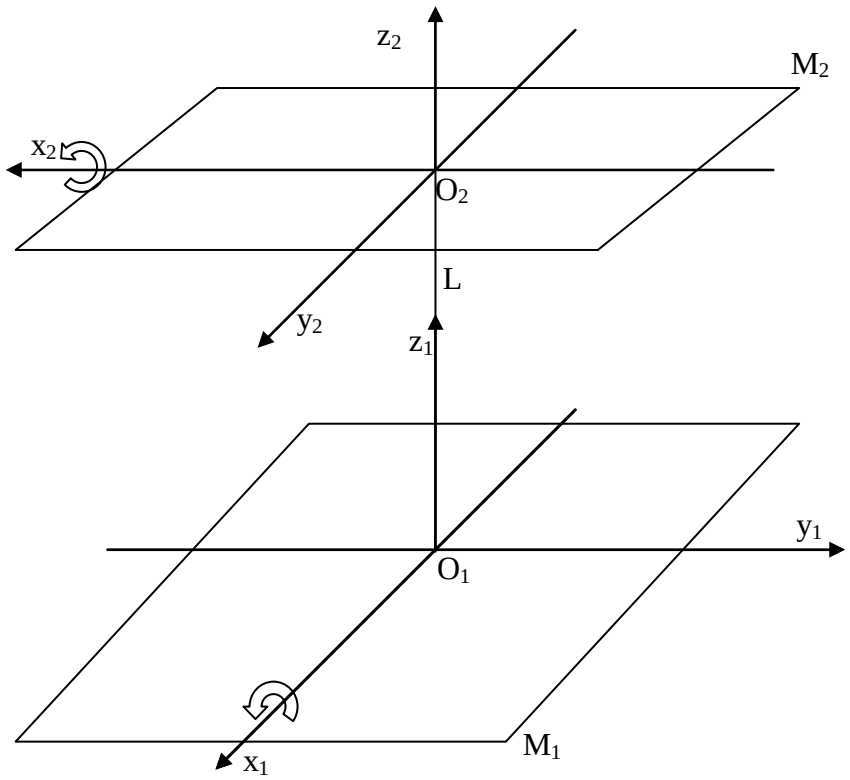

Fig. 5-2 Transformation between the two neutral mirror frames

5.2 Reflection inside a Mirror Frame

We initiate the discussion from the simplest case. Assume that a mirror-fixed frame of reference xyz has been defined following the rules described in previous section. We also assume that the incident ray hits the mirror at the origin of the frame, O. Then the reflected ray shoots out from O. Let $\mathbf{n}_{in}$ denote the directional vector of the incident ray and $\mathbf{n}_{out}$ the directional vector of the reflected ray. They are both normalized. Then $\mathbf{n}_{in}$ and $\mathbf{n}_{out}$ are symmetric with respect to z.

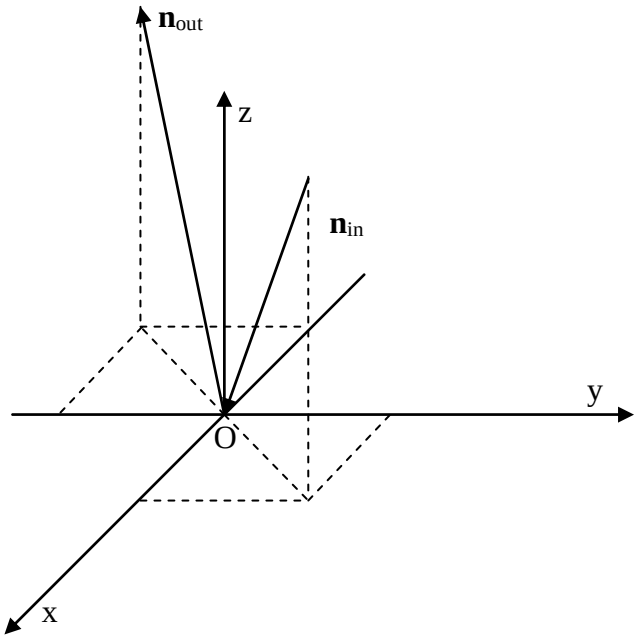

Fig. 5-3 Reflection at the origin of a mirror frame

To study the relationship between $\mathbf{n}_{in}$ and $\mathbf{n}_{out}$, we notice from Fig. 5-3 that $n_{out,z}=n_{in,z}$, $n_{out,x}= -n_{in,x}$, and $n_{out,y}= -n_{in,y}$, which means that if $\mathbf{n}_{in}=[a, b, c]^T$, $\mathbf{n}_{out}=[-a, -b, c]^T$. This relationship can also be written in the matrix form as

$$\mathbf{n}_{out} = \mathbf{T} \bullet \mathbf{n}_{in}, \qquad 5.3$$

where $\mathbf{T} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. The incident and reflected lines are defined respectively as

$$\begin{aligned} \mathbf{p}_{in} &= \lambda_{in}\mathbf{n}_{in} \\ \mathbf{p}_{out} &= \lambda_{out}\mathbf{n}_{out} \end{aligned}, \qquad 5.4$$

where λ_{in} and λ_{out} are linear parameters.

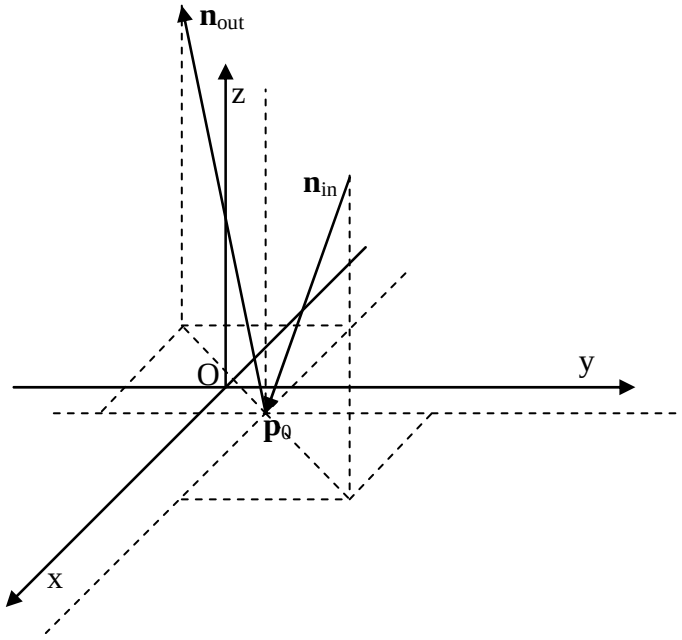

Fig. 5-4 Reflection inside a mirror frame

In general, when the incident ray hits a point on xy plane other than the origin (Fig. 5-4), the equations for the incident and reflected lines become.

$$\begin{aligned} \mathbf{p}_{in} &= \mathbf{p}_{in,0} + \lambda_{in}\mathbf{n}_{in} \\ \mathbf{p}_{out} &= \mathbf{p}_{out,0} + \lambda_{out}\mathbf{n}_{out} \end{aligned}, \qquad 5.5$$

where $\mathbf{p}_{in,0}$ denotes the position vector of an arbitrary point on the incident line, and $\mathbf{p}_{out,0}$ denotes the position vector of an arbitrary point on the reflected line. If we choose $\mathbf{p}_{in,0}$

and $\mathbf{p}_{out,0}$ to be both coincident with the incident point, $\mathbf{p}_0$, where the incident ray hits the xy plane, Equ.5.5 becomes

$$\begin{aligned} \mathbf{p}_{in} &= \mathbf{p}_0 + \lambda_{in}\mathbf{n}_{in} \\ \mathbf{p}_{out} &= \mathbf{p}_0 + \lambda_{out}\mathbf{n}_{out} \end{aligned} . \qquad 5.6$$

5.3 Reflection by a Rotating Mirror

Assume that the laser pointer is physically fixed in space. As the mirror rotates about its x axis, the mirror frame will also rotate (Fig. 5-5). As a result, the incident line will be redefined within the current frame, though physically the incident line does not move. Correspondingly, the reflected line will change.

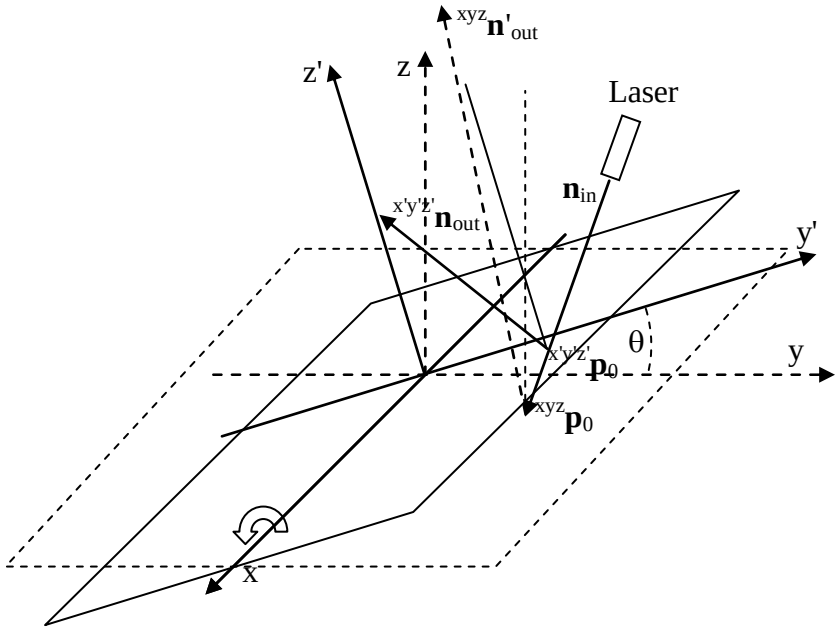

Fig. 5-5 Reflection by a rotating mirror

We begin with defining the physically fixed incident line in xyz which corresponds to the neutral mirror position, i.e. θ=0. According to Equ.5, the incident line in xyz is

$$^{xyz}\mathbf{p}_{in} = {}^{xyz}\mathbf{p}_{in,0} + {}^{xyz}\lambda_{in}\,{}^{xyz}\mathbf{n}_{in}, \qquad 5.7$$

where ${}^{xyz}\mathbf{p}_{in,0}$ denotes an arbitrary point on the incident line defined in xyz, ${}^{xyz}\mathbf{n}_{in}$ is $\mathbf{n}_{in}$ defined in xyz, and ${}^{xyz}\lambda_{in}$ is a linear parameter defined in xyz.

If we rotate the mirror about its x axis by an angle of θ, the frame xyz becomes x'y'z', i.e.

$$ {}^{xyz}_{x'y'z'}\mathbf{R} = \mathbf{R}_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}, \tag{5.8} $$

where x' coincides with x. $\mathbf{n}_{in}$ is defined in x'y'z' as

$$ {}^{x'y'z'}\mathbf{n}_{in} = {}^{x'y'z'}_{xyz}\mathbf{R}\,{}^{xyz}\mathbf{n}_{in} = {}^{xyz}_{x'y'z'}\mathbf{R}^{-1}\,{}^{xyz}\mathbf{n}_{in} = \mathbf{R}_x(-\theta)\,{}^{xyz}\mathbf{n}_{in}. \tag{5.9} $$

Similarly, $\mathbf{p}_{in,0}$ is defined in x'y'z' as

$$ {}^{x'y'z'}\mathbf{p}_{in,0} = {}^{x'y'z'}_{xyz}\mathbf{R}\,{}^{xyz}\mathbf{p}_{in,0} = \mathbf{R}_x(-\theta)\,{}^{xyz}\mathbf{p}_{in,0}. \tag{5.10} $$

Then, in x'y'z', the incident line can be written as (by applying Equ.5.8 to Equ.5.7)

$$ {}^{x'y'z'}\mathbf{p}_{in} = {}^{x'y'z'}\mathbf{p}_{in,0} + {}^{xyz}\lambda_{in}\,{}^{x'y'z'}\mathbf{n}_{in} = \mathbf{R}_x(-\theta)({}^{xyz}\mathbf{p}_{in,0} + {}^{xyz}\lambda_{in}\,{}^{xyz}\mathbf{n}_{in}). \tag{5.11} $$

Since the incident point in x'y'z', ${}^{x'y'z'}\mathbf{p}_0$, should be in the x'y' plane, to calculate it, we let

$$ 0 = {}^{x'y'z'}p_{in,0,3} + \lambda\,{}^{x'y'z'}n_{in,3}, \tag{5.12} $$

and obtain

$$ \lambda = -\frac{{}^{x'y'z'}p_{in,0,3}}{{}^{x'y'z'}n_{in,3}} = -\frac{-{}^{xyz}p_{in,0,2}\sin\theta + {}^{xyz}p_{in,0,3}\cos\theta}{-{}^{xyz}n_{in,2}\sin\theta + {}^{xyz}n_{in,3}\cos\theta}. \tag{5.13} $$

Substituting Equ.5.13 into Equ.5.11, we obtain

$$ {}^{x'y'z'}\mathbf{p}_0 = {}^{x'y'z'}\mathbf{p}_{in,0} + \lambda\,{}^{x'y'z'}\mathbf{n}_{in} = \mathbf{R}_x(-\theta)\left({}^{xyz}\mathbf{p}_{in,0} - \frac{-{}^{xyz}p_{in,0,2}\sin\theta + {}^{xyz}p_{in,0,3}\cos\theta}{-{}^{xyz}n_{in,2}\sin\theta + {}^{xyz}n_{in,3}\cos\theta}\,{}^{xyz}\mathbf{n}_{in}\right). \tag{5.14} $$

Then the incident line in x'y'z' can be rewritten as

$$ {}^{x'y'z'}\mathbf{p}_{in} = {}^{x'y'z'}\mathbf{p}_0 + {}^{x'y'z'}\lambda_{in}\,{}^{x'y'z'}\mathbf{n}_{in} = \mathbf{R}_x(-\theta)\left({}^{xyz}\mathbf{p}_{in,0} - \frac{-{}^{xyz}p_{in,0,2}\sin\theta + {}^{xyz}p_{in,0,3}\cos\theta}{-{}^{xyz}n_{in,2}\sin\theta + {}^{xyz}n_{in,3}\cos\theta}\,{}^{xyz}\mathbf{n}_{in} + {}^{x'y'z'}\lambda_{in}\,{}^{xyz}\mathbf{n}_{in}\right) $$

, 5.15

where ${}^{x'y'z'}\lambda_{in}$ is a new linear parameter defined in x'y'z'. The corresponding reflected line in x'y'z' can be written as

$$ \begin{aligned} {}^{x'y'z'}\mathbf{p}_{out} &= {}^{x'y'z'}\mathbf{p}_0 + {}^{x'y'z'}\lambda_{out}\,{}^{x'y'z'}\mathbf{n}_{out} \\ &= \mathbf{R}_x(-\theta)\left({}^{xyz}\mathbf{p}_{in,0} - \frac{-{}^{xyz}p_{in,0,2}\sin\theta + {}^{xyz}p_{in,0,3}\cos\theta}{-{}^{xyz}n_{in,2}\sin\theta + {}^{xyz}n_{in,3}\cos\theta}\,{}^{xyz}\mathbf{n}_{in}\right) + {}^{x'y'z'}\lambda_{out}\mathbf{T}\mathbf{R}_x(-\theta)\,{}^{xyz}\mathbf{n}_{in} \end{aligned} , \quad 5.16 $$

where ${}^{x'y'z'}\lambda_{out}$ is a linear parameter defined in x'y'z', and

$$ {}^{x'y'z'}\mathbf{p}_0 = \mathbf{R}_x(-\theta)\left({}^{xyz}\mathbf{p}_{in,0} - \frac{-{}^{xyz}p_{in,0,2}\sin\theta + {}^{xyz}p_{in,0,3}\cos\theta}{-{}^{xyz}n_{in,2}\sin\theta + {}^{xyz}n_{in,3}\cos\theta}\,{}^{xyz}\mathbf{n}_{in}\right), \quad 5.17 $$

$$ {}^{x'y'z'}\mathbf{n}_{out} = \mathbf{T}\mathbf{R}_x(-\theta)\,{}^{xyz}\mathbf{n}_{in}. \quad 5.18 $$

Here, the relationship between the incident and reflected lines, defined in Equ.5.3, is used.

For further discussion, we transform the reflected line into xyz, and obtain

$$ {}^{xyz}\mathbf{p}_{out,0} = {}^{xyz}_{x'y'z'}\mathbf{R}\,{}^{x'y'z'}\mathbf{p}_0 = {}^{xyz}\mathbf{p}_{in,0} - \frac{-{}^{xyz}p_{in,0,2}\sin\theta + {}^{xyz}p_{in,0,3}\cos\theta}{-{}^{xyz}n_{in,2}\sin\theta + {}^{xyz}n_{in,3}\cos\theta}\,{}^{xyz}\mathbf{n}_{in}. \quad 5.19 $$

$$ {}^{xyz}\mathbf{n}_{out} = {}^{xyz}_{x'y'z'}\mathbf{R}\,{}^{x'y'z'}\mathbf{n}_{out} = \mathbf{R}_x(\theta)\mathbf{T}\mathbf{R}_x(-\theta)\,{}^{xyz}\mathbf{n}_{in}. \quad 5.20 $$

$$ \begin{aligned} {}^{xyz}\mathbf{p}_{out} &= {}^{xyz}\mathbf{p}_{out,0} + {}^{x'y'z'}\lambda_{out}\,{}^{xyz}\mathbf{n}_{out} = {}^{xyz}_{x'y'z'}\mathbf{R}\left({}^{x'y'z'}\mathbf{p}_0 + {}^{x'y'z'}\lambda_{out}\,{}^{x'y'z'}\mathbf{n}_{out}\right) \\ &= \left({}^{xyz}\mathbf{p}_{in,0} - \frac{-{}^{xyz}p_{in,0,2}\sin\theta + {}^{xyz}p_{in,0,3}\cos\theta}{-{}^{xyz}n_{in,2}\sin\theta + {}^{xyz}n_{in,3}\cos\theta}\,{}^{xyz}\mathbf{n}_{in}\right) + {}^{x'y'z'}\lambda_{out}\mathbf{R}_x(\theta)\mathbf{T}\mathbf{R}_x(-\theta)\,{}^{xyz}\mathbf{n}_{in} \end{aligned} . \quad 5.21 $$

5.4 Reflection by Two Mirrors

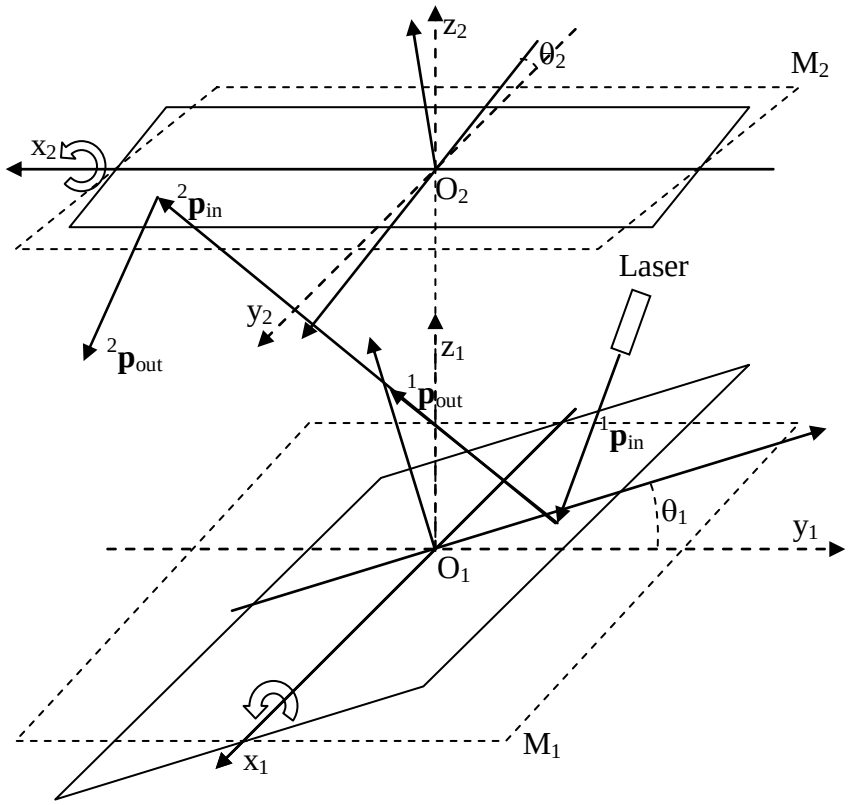

Fig. 5-6 Reflection by the two mirrors

Based on Equ.5.21, the reflected line by M_1 defined in $x_1y_1z_1$ is

$$ {}^1\mathbf{p}_{out} = ({}^1\mathbf{p}_{in,0} - \frac{-{}^1p_{in,0,2}\sin\theta_1 + {}^1p_{in,0,3}\cos\theta_1}{-{}^1n_{in,2}\sin\theta_1 + {}^1n_{in,3}\cos\theta_1}{}^1\mathbf{n}_{in}) + {}^{1'}\lambda_{out}\mathbf{R}_x(\theta_1)\mathbf{T}\mathbf{R}_x(-\theta_1){}^1\mathbf{n}_{in}. \quad 5.22 $$

To M_2, ${}^1\mathbf{p}_{out}$ is an incident line which can be transformed into $x_2y_2z_2$ as

$$ {}^2\mathbf{p}_{in} = {}^2_1\mathbf{T}{}^1\mathbf{p}_{out} = {}^2_1\mathbf{T}({}^1\mathbf{p}_{in,0} - \frac{-{}^1p_{in,0,2}\sin\theta_1 + {}^1p_{in,0,3}\cos\theta_1}{-{}^1n_{in,2}\sin\theta_1 + {}^1n_{in,3}\cos\theta_1}{}^1\mathbf{n}_{in}) + {}^{1'}\lambda_{out}\,{}^2_1\mathbf{T}\mathbf{R}_x(\theta_1)\mathbf{T}\mathbf{R}_x(-\theta_1){}^1\mathbf{n}_{in}. $$

5.23

Defining

$$ {}^{2}\mathbf{p}_{in,0} = {}^{2}_{1}\mathbf{T}({}^{1}\mathbf{p}_{in,0} - \frac{-{}^{1}p_{in,0,2}\sin\theta_1 + {}^{1}p_{in,0,3}\cos\theta_1}{-{}^{1}n_{in,2}\sin\theta_1 + {}^{1}n_{in,3}\cos\theta_1}{}^{1}\mathbf{n}_{in}), \quad 5.24 $$

$$ {}^{2}\mathbf{n}_{in} = {}^{2}_{1}\mathbf{T}\mathbf{R}_x(\theta_1)\mathbf{T}\mathbf{R}_x(-\theta_1){}^{1}\mathbf{n}_{in}, \quad 5.25 $$

Based on Equ.5.21, we can obtain the reflected line by M_2 defined in $x_2y_2z_2$ as

$$ {}^{2}\mathbf{p}_{out} = ({}^{2}\mathbf{p}_{in,0} - \frac{-{}^{2}p_{in,0,2}\sin\theta_2 + {}^{2}p_{in,0,3}\cos\theta_2}{-{}^{2}n_{in,2}\sin\theta_2 + {}^{2}n_{in,3}\cos\theta_2}{}^{2}\mathbf{n}_{in}) + {}^{2'}\lambda_{out}\mathbf{R}_x(\theta_2)\mathbf{T}\mathbf{R}_x(-\theta_2){}^{2}\mathbf{n}_{in}. \quad 5.26 $$

In the practice, the situation may not be exactly the same as we discussed. We have to adjust the mirror frame and reflection equation accordingly. If when $\theta_1=\theta_2=0$, the two mirror planes are not parallel to each other, we can find a parallel configuration for the 2-mirror system at first. If $\theta_1=\theta_{1,0}$ and $\theta_2=\theta_{2,0}$ at this parallel configuration, the θ_1 and θ_2 in Equ.5.22 through Equ.5.26 should be replaced with θ_1-$\theta_{1,0}$ and θ_2-$\theta_{2,0}$ respectively.

In addition, if the revolute axes of the two mirrors are not perfectly perpendicular to each other, Equ.5.1 should be adjusted as

$$ {}^{1}_{2}\mathbf{T} = \mathbf{R}_z(-\frac{\pi}{2} + \alpha_0)\mathbf{T}_z(a) = \mathbf{T}_z(a)\mathbf{R}_z(-\frac{\pi}{2} + \alpha_0), \quad 5.27 $$

where $\pi/2+\alpha_0$ is the actual angle from x_1 to x_2 about z_1.

5.5 Simplified Scanner Model Derivation

The scanner we are using has some features which can help us simplify the model. Assume the scanners are centered, at their zero position. The laser beam comes in from right, is reflected by the first mirror (We call it scanner X) up to the second mirror (scanner Y), and then by mirror Y to the target (Fig 5-7). Usually the beam enters parallel to the base of the mount and perpendicular to the axis of the X scanner, and leaves parallel to the base of the mount and perpendicular to the axis of the Y scanner. That is,

the beam changes direction by 90 degrees and height by the distance between centers of the X and Y scanners.

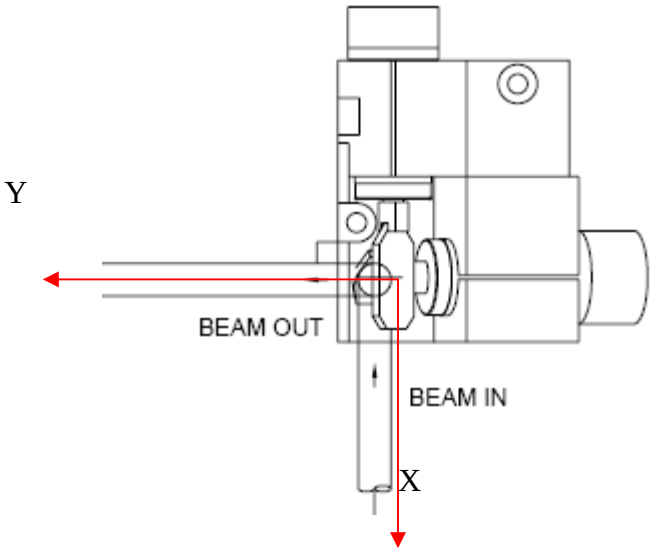

Fig. 5-7 Simplified Scanner Model (1)

A virtual mirror coordinate is set by choosing zero-centered beam output as Y-axis and rotating axis of Y scanner as X axis (Fig. 5-7). All the following vectors are measured in this coordinate system.

Assume that distance between centers of the X and Y scanner is d, and the tilted angle of scanner X is α (Fig. 5-8).

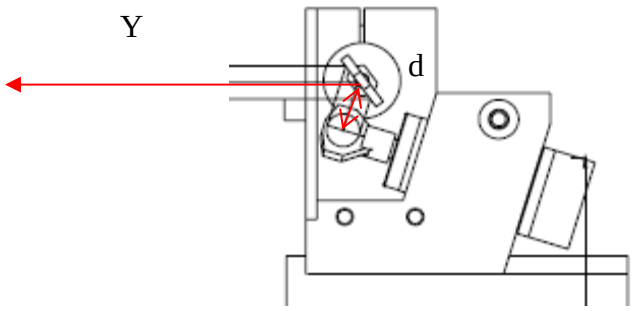

Fig. 5-8 Simplified Scanner Model (2)

Let us first examine the path of the zero-center beam (Fig. 5-8). In the virtual mirror coordinate, the initial direction of laser beam after reflected by scanner X is $n_{in} = [0, -d \cdot \sin\alpha, -d \cdot \cos\alpha]'$, and the reflection point is the origin $P_{in,0} = [0, 0, 0]'$. As we have discussed in the section 4.6, we have to take consideration of the initial position of the scanner Y. Compared with the virtual mirror coordinate, the initial angle of mirror Y is $\theta_{2,0} = -(90 + \alpha)/2$. Substitute these parameters into equ.5-21, we will have $P_{out,0} = [0, 0, 0]'$ and the directional vector

$$\begin{aligned} n_{out} &= R_x(\theta_{2,0}) \cdot T \cdot R_x(-\theta_{2,0}) \cdot [0, -d \cdot \sin\alpha, -d \cdot \cos\alpha]' \\ &= [0, d, 0]' \end{aligned}, \qquad 5.28$$

where ${}^{xyz}_{x'y'z'}\mathbf{R} = \mathbf{R}_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$, and $\mathbf{T} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

This result is consistent with the reality that the final laser output is along with the Y axis.

Now let us further analyze the influence of two rotating mirrors. According to the Fig. 5-8, when the scanner rotates θ_1 degrees, the direction of beam after reflected by scanner X is $n_{in} = [-d \cdot \tan 2\theta_1, -d \cdot \sin\alpha, -d \cdot \cos\alpha]'$, and now the reflection point is the $P_{in,0} = [-d \cdot \tan 2\theta_1, 0, 0]'$. If the scanner Y rotates θ_2, the actual angle we should use in the computation is $\theta_2 + \theta_{2,0}$.

According to the equ.5-21, we have

$$P_{out,0} = [-d \cdot \tan 2\theta_1, 0, 0]' \qquad 5.29$$

$$\begin{aligned} n_{out} &= R_x(\theta_2 + \theta_{2,0}) \cdot T \cdot R_x(-\theta_2 - \theta_{2,0}) \cdot [-d \cdot \tan 2\theta_1, -d \cdot \sin\alpha, -d \cdot \cos\alpha]' \\ &= [-d \cdot \tan 2\theta_1, d \cdot \cos 2\theta_2, d \cdot \sin 2\theta_2]' \\ &= [-\tan 2\theta_1, \cos 2\theta_2, \sin 2\theta_2]' \end{aligned} \qquad 5.30$$

5.6 Summary

This chapter gives a full description of the 2-axis laser scanner modeling, starting from reflection inside a mirror frame, reflection by a rotating mirror to the reflection by two rotating mirrors. The orientation of the output beam in the original reference frame can be derived using the input beam and rotating angle. The most important result we can get from one rotating mirror is

$$
\begin{aligned}
{}^{xyz}\mathbf{p}_{out} &= {}^{xyz}\mathbf{p}_{out,0} + {}^{x'y'z'}\lambda_{out}\,{}^{xyz}\mathbf{n}_{out} = {}^{xyz}_{x'y'z'}\mathbf{R}({}^{x'y'z'}\mathbf{p}_0 + {}^{x'y'z'}\lambda_{out}\,{}^{x'y'z'}\mathbf{n}_{out}) \\
&= ({}^{xyz}\mathbf{p}_{in,0} - \frac{-{}^{xyz}p_{in,0,2}\sin\theta + {}^{xyz}p_{in,0,3}\cos\theta}{-{}^{xyz}n_{in,2}\sin\theta + {}^{xyz}n_{in,3}\cos\theta}\,{}^{xyz}\mathbf{n}_{in}) + {}^{x'y'z'}\lambda_{out}\mathbf{R}_x(\theta)\mathbf{T}\mathbf{R}_x(-\theta)\,{}^{xyz}\mathbf{n}_{in}
\end{aligned}
$$

The practical situation is also consideration. This model can be used in a general purpose. Considering the scanner feature, a simplified model is used in the real computation. The final direction of laser beam can be estimated by the two rotating angles θ_1 and θ_2, with $P_{out,0} = [-d\cdot\tan 2\theta_1,\ 0,\ 0]'$ and $n_{out} = [-\tan 2\theta_1,\ \cos 2\theta_2,\ \sin 2\theta_2]'$.

Chapter 6 Conclusions and Future Works

6.1 Summary of Contributions

The primary contributions of this work are introducing a new artificial landmark-based indoor localization system for mobile robots and presenting a new localization algorithm based on a single image with identified landmarks.

The proposed localization technique is inspired by the existing GPS and RFID-based localization techniques. In the proposed system, a new type of RFID tag, the laser-activated RFID tag, is designed and used as the artificial landmark. The LARFID tags function like the indoor equivalent of the GPS satellites. Stereo vision and LARFID are combined, together with trilateration or triangulation, to localize a mobile robot in an indoor environment. Feasibility study shows that the proposed system is promising to provide a robust and accurate indoor localization for mobile robots.

A novel effective localization algorithm for mobile robots is discovered based on one single image of a few identified landmarks. The visual angle between two landmarks can be derived from their projections in the same image. The lens distortion is also considered in the model. The relative position of the landmarks can be calculated from the visual angles and the known landmark positions. The robot position can then be determined using the principle of trilateration. Finally, the robot orientation is computed

from the robot position, landmark positions and their projections.

Finally, a practical problem of modeling a 2-axis laser scanner is solved. This model illustrates the reflection inside a mirror frame, reflection by a rotating mirror and the reflection by two rotating mirrors. The orientation of the output beam in the original reference frame can be derived using the input beam and rotating angle. The practical situation is also consideration. Considering the scanner feature, a simplified model is used in the real computation.

6.2 Future Work

Some future research need to be conducted to complete the proposed system. The most important part is to integrate the RFID library into our controlling system, which makes this prototype fully autonomous. At the same time, a simultaneous localization and mapping algorithm is under consideration with the intention to automate the pre-installation process of the LARFID tags in an unexplored indoor environment. Besides, a full calibration based on the complete laser scanner model will further increase the activation accuracy.

As we can see in the simulation part of our new algorithm, the error can be extremely big around the peak. An analytic explanation of this phenomenon is not available yet. It will be investigated in our future work.

References

[1] I. J. Cox, Blanche: Position estimation for an autonomous robot vehicle, Autonomous Mobile Robots: Control, Planning, and Architecture (Vol. 2), IEEE Computer Society Press, Los Alamitos, CA (1991), 285-292.

[2] S. G. Li, “Localization along routes, based upon iconic and Global Positioning System information in large-scale outdoor environments”, Advanced Robotics, vol. 15, pp. 749-762, 2001.

[3] R. Lenain, B. Thuilot, C. Cariou, and P. Martinet P. “High accuracy path tracking for vehicles in presence of sliding: Application to farm vehicle automatic guidance for agricultural tasks”, Autonomous Robots, vol. 21, pp. 79-97, 2006.

[4] J. Borenstein, H. R. Everett, L. Feng, and D. Wehe, “Mobile robot positioning: sensors and techniques”, Journal of Robotic Systems, vol. 14, pp. 231-249, 1997.

[5] A. El-Rabbany, Introduction to GPS: The Global Positioning System. Norwood, MA: Artech House, 2002.

[6] D. Bouvet, and G. Garcia, “GPS latency identification by Kalman filtering”, Robotica, vol. 18, pp. 475-485, 2000.

[7] A. Georgiev, and P. K. Allen, “Localization methods for a mobile robot in urban environments”, IEEE Transactions on Robotics and Automation, vol. 20, pp. 851-864, 2004.

[8] O. Yukumoto, Y. Matsuo, and N. Noguchi, “Robotization of agricultural vehicles (Part 1) - Component technologies and navigation systems”, JARQ-Japan Agricultural Research Quarterly, vol. 34, pp. 99-105, 2000.

[9] S. Panzieri, F. Pascucci, and G. Ulivi, “An outdoor navigation system using GPS and inertial platform”, IEEE-ASME Transactions on Mechatronics, vol. 7, pp. 134-142, 2002.

[10] J. L. Martinez, R. Molina-Mesa, A. Mandow, and C. A. Rodriguez-Serrano, “Continuous localization via wide-area Differential Global Positioning System for outdoor navigation of mobile robots”, Integrated Computer-Aided Engineering, vol. 11, pp. 1-13, 2004.

[11] K. Ohno, T. Tsubouchi, B. Shigematsu, and S. I. Yuta, “Differential GPS and odometry-based outdoor navigation of a mobile robot”, Advanced Robotics, vol. 18, pp. 611-635, 2004.

[12] C. J. Wu, T. L. Chien, T. L. Lee, and L. C. Lai, “Navigation of a mobile robot in outdoor environments”, Journal of the Chinese Institute of Engineers, vol. 28, pp. 915-924, 2005.

[13] L. C. Lai, C. J. Wu, “An active system for three-dimensional localization of mobile robots”, Journal of Intelligent & Robotic Systems, vol. 44, pp. 123-137, 2005.

[14] S. Y. Yi, B. W. Choi, “Autonomous navigation of indoor mobile robots using a global ultrasonic system”, Robotica, vol. 22, pp. 369-374, 2004.

[15] E. A. Prigge, J. P. How, “Signal architecture for a distributed magnetic local positioning system”, IEEE Sensors Journal, vol. 4, pp. 864-873, 2004.

[16] K. Finkenzeller, RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification, 2nd Edition. West Sussex, England: John Wiley & Sons, 2003.

[17] M. Kim, T. Takeuchi, N. Y. Chong, "A 3-axis Orthogonal Antenna for Indoor Localization", First International Workshop on Networked Sensing Systems, pp. 59-62, 2004.

[18] T. Tsukiyama, "Global navigation system with RFID tags", Proceedings of SPIE, vol. 4573, pp. 256-264, Mobile Robots XVI, 2002.

[19] T. Tsukiyama, "World map based on RFID tags for indoor mobile robots", Proceedings of SPIE, vol 6006, pp. 412-419, Intelligent Robots and Computer Vision XXIII: Algorithms, Techniques, and Active Vision, 2005.

[20] V. Kulyukin, C. Gharpure, J. Nicholson, S. Pavithran, "RFID in robot-assisted indoor navigation for the visually impaired", Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1979-1984, 2004.

[21] V. Kulyukin, C. Gharpure, J. Nicholson, "RoboCart: toward robot-assisted navigation of grocery stores by the visually impaired", Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 2845-2850, 2005.

[22] V. Kulyukin, C. Gharpure, J. Nicholson, G. Osborne, "Robot-assisted wayfinder for the visually impaired in structured indoor environments", Autonomous Robots, vol 21, pp. 29-41, 2006.

[23] K. Yamano, K. Tanaka, M. Hirayama, E. Kondo, Y. Kmuro, M. Matsumoto, "Self-localization of mobile robots with RFID system by using support vector

machine", Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 3756-3761, 2004.

[24] H. Chae, K. Han, "Combination of RFID and vision for mobile robot localization", Proceedings of the 2005 International Conference on Intelligent Sensors, Sensor Networks and Information Processing Conference, pp. 75-80, 2005.

[25] D. Hahnel, W. Burgard, D. Fox, K. Fishkin, M. Philipose, "Mapping and localization with RFID technology", Proceedings of the 2004 IEEE International Conference on Robotics and Automation, pp. 1015-1020, 2004.

[26] S. Jia, W. Lin, K. Wang, K. Takase, "Network distributed multi-functional robotic system supporting the elderly and disabled people", Journal of Intelligent and Robotic Systems, vo. 45, pp. 53-76, 2006.

[27] V. Ashkenazi, D. Park, and M. Dumville, "Robot positioning and the global navigation satellite system", *Industrial Robot*, vol. 27, pp. 419-426, 2000.

[28] M. Trucco, and A. Verri, Introductory Techniques for 3-D Computer Vision, New Jersey: Prentice Hall, 1998.

[29] A. Jain, Y. Zhou, T. Mustufa, C. Burdette, G. S. Chirikjian, and G. Fichtinger, "Matching and Reconstruction of Brachytherapy Seeds Using the Hungarian Algorithm (MARSHAL) ", Medical Physics, vol. 32, pp. 3475-3492, 2005.

[30] D. E. Manolakis, "Efficient solution and performance analysis of 3-D position estimation by trilateration," IEEE Trans. Aerosp. Electron. Syst., vol. 32, pp. 1239–1248, Apr. 1996.

[31] D. E. Manolakis and M. E. Cox, “Effect in range difference position estimation due to stations’ position errors,” IEEE Trans. Aerosp. Electron. Syst., vol. 34, pp. 329–334, Jan. 1998.

[32] I. D. Coope, “Reliable computation of the points of intersection of n spheres in R^n ,” Australian, New Zealand Ind. Appl. Math. J., pt. C, vol.42, pp. 461–477, 2000.

[33] Federico Thomas and Lluís Ros, “Revisiting trilateration for robot localization,” IEEE Trans. Robot. Automat., vol. 21, pp. 93–101, Feb. 2005.

[34] B. T. Fang, “Trilateration and extension to global positioning system navigation,” J. Guidance, Contr., Dynam., vol. 9, no. 6, pp. 715–717, 1986.

[35] A. El-Rabbany, Introduction to GPS: The Global Positioning System. Norwood, MA: Artech House, 2002.

[36] V. Ashkenazi, D. Park, and M. Dumville, “Robot positioning and the global navigation satellite system,” Industrial Robot, vol. 27, pp. 419-426, 2000.

[37] Jeffrey Hightower, Gaetano Borriello and Roy Want, “SpotON: An Indoor 3D Location Sensing Technology Based on RF Signal Strength,” UW CSE Technical Report, 2000.

[38] R. Want, A. Hopper, V. Falco, J. Gibbons. “The Active Badge Location System,” ACM Transactions on Information Systems, Vol. 40, No. 1, pp. 91-102, January 1992

[39] Andy Ward, Alan Jones, and Andy Hopper, “A New Location Technique for the Active Office,” IEEE Personal Communications, Vol. 4, No. 5, October 1997, pp. 42-47.

[40] N. Priyantha, A. Miu, H. Balakrishnan, and S. Teller, “The Cricket Compass for Context-Aware Mobile Applications,” In Proc. 7th ACM MOBICOM Conf., pages 1–14, Rome, Italy, July 2001.

[41] Yu Zhou, Wenfei Liu and Peisen Huang, “Laser-activated RFID-based indoor localization system for mobile robots (To be published),” 2007 IEEE International Conference on Robotics and Automation.

[42] Charles Cohen and Frank V. Koss, “A Comprehensive Study of Three Object Triangulation,” in Proc. SPIE Vol. 1831, p. 95-106, Mobile Robots VII, May 1993.

[43] M. Betke and K. Gurvits, “Mobile robot localization using landmarks,” IEEE Trans. Robot. Automat., vol. 13, pp. 51–263, Apr. 1997.

[44] Ilan Shimshoni, “On mobile robot localization from landmark bearings,” IEEE Trans. Robot. Automat., vol. 18, pp. 971–976, Dec. 2002.

[45] K. T. Sutherland and W. B. Thompson, “Inexact navigation,” in Proc.1993 IEEE Int. Conf. Robot. Automat., Atlanta, GA, May 1993, vol. 1, pp. 1–7.

[46] A. J. Muñoz and J. Gonzalez, “Two-dimensional landmark-based position estimation from a single image,” in Proc. IEEE Int. Conf. Robotics and Automation, 1998, pp. 3709–3714.

[47] H. Ishiguro, K. Kato, and M. Barth, “Identifying and localizing robots with omnidirectional vision sensors,” in Panoramic Vision: Sensors, Theory, and Application, R. Benosman and S. B. Kang, Eds. New York: Springer-Verlag, 2001, pp. 376–391.

[48] Stephen Se, D. Lowe, and J. Little, “Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks,” Int. Journal of Robotics Research, 21(8):735–758, 2002.

[49] A.J. Davison and D.W. Murray, “Mobile robot localization using active vision,” In European Conference on Computer Vision, 1998.

[50] David A. Forsyth and Jean Ponce, “Computer Vision: A Modern Approach,” Prentice Hall, 2002, pp. 20–54.

[51] D. C. Brown, “Decentering distortion of lenses”, *Photometric Engineering*, vol. 32, pp. 444-462, 1966.