ABSTRACT

Title of dissertation: Pointing, Acquisition, and Tracking Systems for Free-Space Optical Communication Links:

Tzung-Hsien Ho
Doctor of Philosophy, 2007

Dissertation directed by: Professor Christopher C. Davis
Department of Electrical and Computer Engineering

Pointing, acquisition, and tracking (PAT) systems have been widely applied in many applications, from short-range (e.g. human motion tracking) to long-haul (e.g. missile guidance) systems. This dissertation extends the PAT system into new territory: free space optical (FSO) communication system alignment, the most important missing ingredient for practical deployment.

Exploring embedded geometric invariances intrinsic to the rigidity of actuators and sensors is a key design feature. Once the configuration of the actuator and sensor is determined, the geometric invariance is fixed, which can therefore be calibrated in advance. This calibrated invariance further serves as a transformation for converting the sensor measurement to actuator action.

The challenge of the FSO alignment problem lies in how to point to a 3D target by only using a 2D sensor. Two solutions are proposed: the first one exploits the invariance, known as the linear homography, embedded in the FSO applications which involve long link length between transceivers or have planar trajectories. The second one employs either an additional 2D or 1D sensor, which results in invariances

known as the trifocal tensor and radial trifocal tensor, respectively. Since these invariances have been developed upon an assumption that the measurements from sensors are free from noise, including the uncertainty resulting from aberrations, a robust calibrate algorithm is required to retrieve the optimal invariance from noisy measurements.

The first solution is sufficient for most of the PAT systems used for FSO alignment since a long link length constraint is generally the case. Although PAT systems are normally categorized into coarse and fine subsystems to deal with different requirements, they are proven to be governed by a linear homography. Robust calibration algorithms have been developed during this work and further verified by simulations. Two prototype systems have been developed: one serves as a fine pointing subsystem, which consists of a beam steerer and an angular resolver; while the other serves as a coarse pointing subsystem, which consists of a rotary gimbal and a camera. The average pointing errors in both prototypes were less than 170 and 700 μrads, respectively.

PAT systems based on the second solution are capable of pointing to any target within the intersected field-of-view from both sensors because two sensors provide stereo vision to determine the depth of the target, the missing information that cannot be determined by a 2D sensor. They are only required when short-distance FSO communication links must be established. Two simulations were conducted to show the robustness of the calibration procedures and the pointing accuracy with respect to random noise.

Pointing, Acquisition, and Tracking Systems for Free-Space Optical Communication Links

by

Tzung-Hsien Ho

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2007

Advisory Committee:
Professor Christopher C. Davis, Chair/Advisor
Professor Thomas Antonsen
Professor Gilmer L. Blakenship
Professor Robert W. Gammon
Professor Julius Goldhar

Dedication

To my parents

for their continuous guidance and dedication

To my fiancee, Jen-Wei

for her love and support

Acknowledgements

My sincere gratitude goes to my advisor, Professor Christopher C. Davis, for his continuous guidance, patience, and support throughout my graduate studies at the University of Maryland. His liberal attitude gave me the freedom to explore the topics that I am interested in. His wisdom and knowledge led me through many hardships in the course of my research. My appreciation also goes to Professor Stuart Milner for continuously investigating new ideas with me as well as collaborating with me in numerous publications.

I am fortunate enough to work with the best colleagues in the world throughout my years of graduate studies. I am particularly thankful to Dr. Igor Smolyaninov, Professor Robert Gammon, Professor William Levine, and Dr. Vildana Hodzic for their advice to both my career and research. I am very grateful to Dr. Sugianto Trisno and John Rzasa, for the fruitful discussions and collaborations with them; and most importantly, for their friendship. I also thank Dr. Heba Yuksel, Dr. Quirino Balzano, Aniket Desai, Jaime Llorca, Clint Edwards, Yuju Hung, and Felice Venin for the technical and non-technical discussions we have had.

Finally, I want to thank my fiancee, Jen-Wei, for her love and support during my many years of study. I truly appreciate her outstanding patience for bearing our continuous separation. I could not have done it without her. I also want to express my gratitude to my parents for their continuous guidance and dedication throughout my entire educational path. They are the inspiration for me to succeed.

Table of Contents

List of Figures

List of Algorithms

List of Abbreviations

2D	Two-Dimensional
3D	Three-Dimensional
AMS	Algebraic Minimization Scheme
AOP	Axis of Propagation
AR	Angular Resolver
BER	Bit Error Rate
BS	Beam Steerer
CCD	Coupled Charge Detector
CPAT	Coarse Pointing, Acquisition, and Tracking
CPC	Compound Parabolic Concentrator
DB,db	Decibel
ECEF	Earth Centered Earth Fixed
EDFA	Erbium Doped Fiber Amplifier
FCC	Federal Communications Commission
FOV,fov	field of view
FPAT	Fine Pointing, Acquisition, and Tracking
FPM	Focal Plane Motion
FSO	Free Space Optical Communication
FSM	Fast Steering Mirror
GMS	Geometric Minimization Scheme
GPS	Global Positioning System
He	Helium
Ne	Neon
IMU	Inertial Measurement Unit
ISS	Inertial Sensing System
LANs	Local Area Networks
LED	Light Emitting Diode
LEP	Lateral Effect Photodiode
LOS	Line-of-Sight
LS	Least Squares
LVDT	Linear Variable Displacement Transducer
MFD	Mode Fiber Diameter
PRBS	Pseudo Random Binary Sequence
PSD	Position Sensing Detector
QD	Quadrant Detector
RF	Radio Frequency
RTT	Radial Trifocal Tensor
SNR	Signal to Noise Ratio
SVD	Singular Value Decomposition
TLS	Total Least Squares
TOF	Time of Flight
TOI	Target of Interest
TT	Trifocal Tensor
WiMAX	Worldwide Interoperability for Microwave Access

Chapter 1

Introduction

Pointing, acquisition and tracking (PAT) systems have been widely applied in numerous industrial, military, and commercial applications. Typical examples include robotic arm positioning [4], laser ranging [37], satellite ranging, virtual reality user interaction [55], missile guidance [38] and directional communication systems alignment [26]. This work explores the geometric invariance of the configuration and further applies it to solve the alignment problem in free-space optical (FSO) communications. In general, a PAT system is capable of acquiring remote targets, aiming the local lasers toward the acquired targets, and further locking the laser on the acquired targets despite disturbances induced by the target's motion, atmospheric turbulence, or random jitter.

1.1 Overview of Free Space Optical Communications

With the deployment of third generation (3G) communication systems, the search for next generation communication systems, commonly referred to as fourth generation (4G) systems, has begun, and its goal is to provide ubiquitous connectivity and seamlessly integrated operations among different scenarios ranging from short-range, high-mobility cellular systems, to long-range, high data-rate directional systems. It is commonly agreed that 4G systems will not be based on a single access

technique but will encompass a number of different complementary access technologies [51].

Free space optical communication systems have emerged as a viable technology in this next generation indoor and outdoor broadband revolution. A free space optical (FSO) link is commonly described as a direct line-of-sight (LOS), point-to-point, long-haul, optical communications link. It utilizes a laser or light emitting diode (LED) to transmit signals through the atmosphere. Because of light's great capacity to carry information, FSO systems can provide services ranging from short-distance, broadcasting networks connecting high-mobility terminals (known as diffusive systems), to long-haul, point-to-point links bridging the fiber backbone and local network, and even between the ground station and a spaceship orbiting Mars (known as LOS systems) [1].

Diffusive systems provide omnidirectional data access by utilizing a wide line-of-sight (WLOS) transceiver or a diffuse reflection from obstacles, where the transceiver alignment is redundant since no direct line-of-sight (LOS) path is required from the transmitter to the receiver. Several systems have been proposed including WLOS systems composed of quadratic transceivers by Trisno et.al. [67] and compound parabolic concentrators (CPC) by Carruther et.al. [11], and a diffuse-reflection system composed of white LEDs and power line by Kaverhrad et.al.[3].

LOS systems are capable of supporting long-distance communication because the small beam divergence of the optical signal prevents the transmitted power from being dissipated. A small beam divergence offers essentially very secure channels with low probability of interception and simultaneously enhances the signal-to-

noise ratio (SNR) of the received signals which therefore improves the bit-error rate (BER). Recently, a 80 Gbps FSO system over a few kilometers was demonstrated by a joint project between the Applied Physics Laboratory and AOptix Technologies [60].

The drawback of a LOS system is the requirement of LOS paths between transceivers. The establishment and maintenance of the LOS paths are subject to a PAT system whose performance depends on a seamless collaboration among: (1) optical/electrical, (2) closed-loop control, and (3) geometric mapping subsystems, as shown in figure 1.1.

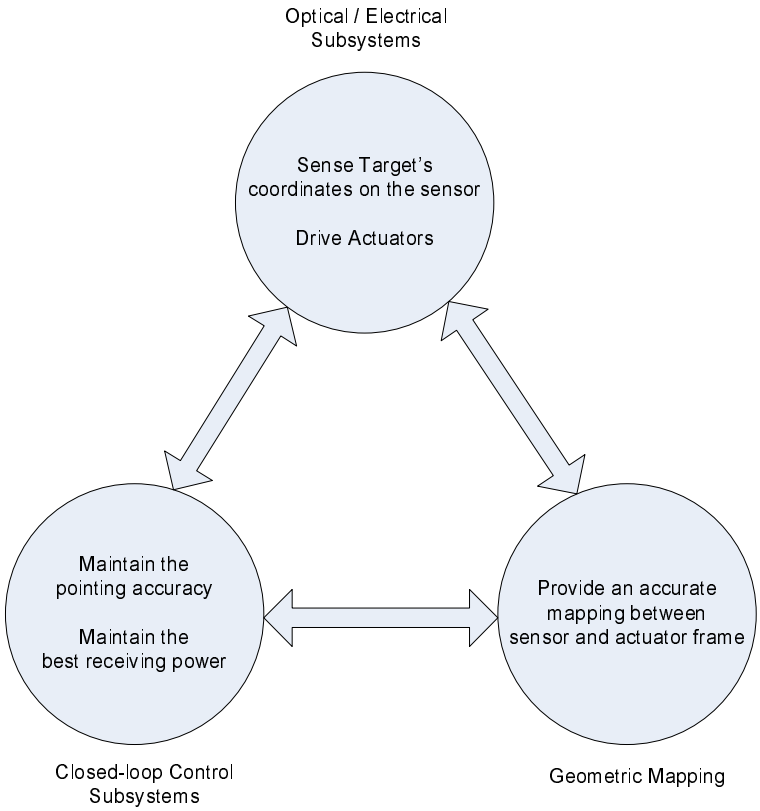

Figure 1.1: Tasks of the three essential subsystems including (1)optical/electrical, (2)closed-loop control, and (3)geometric mapping in a PAT system

The objective of this research is to investigate and develop a systematic PAT design method for LOS systems. Among many of the FSO applications, we put the emphasis on the PAT system used in LOS transceivers for solving the last-mile problem since it is one of the toughest challenges for such 4G communication systems.

1.1.1 Last Mile Problem and Its Solutions

Bandwidth is a scarce and exhaustible resource and can therefore be over-utilized. When the bandwidth use is above capacity, congestion occurs and networks can no longer run efficiently. The current solution is to form individual cells, also known as local area networks (LANs), for frequency reuse. However, the infrastructure of the connection between the LANs to the fiber backbone or among the LANs is still inadequate because of the significant mismatch between available backbones (Gb/s to Tb/s) and current broadband wireless networks (Kb/s to Mb/s). This mismatch problem is also known as the last-mile problem. It has been estimated that over 90% of U.S. businesses fall less than a mile short of a fiber backbone [2].

The IEEE developed standard IEEE 802.16 [15] in October 2001 known as the Worldwide Interoperability for Microwave Access (WiMAX) standard for solving the last-mile problem. The original 802.16a specifies WiMAX at 10 to 66 GHz, and the later version 802.16-2004 added support for 2 to 11 GHz. Ironically, from 10 to 66 GHz, only 24.125 ± 0.125 and 61.25 ± 0.25 GHz are for industrial, scientific, and medical (ISM) use and $59 - 64$ GHz is for unlicensed devices. All other bands

are licensed and are regulated by the Federal Communications Commission (FCC). However, the two license-free bands at 24 GHz and 61 GHz are strongly attenuated by water vapor and oxygen, respectively, which implies large powers to support long-haul communications [12]. Also, the long wavelength of radio-frequency (RF) waves leads to a large diffraction angle, which further results in poor reception since the received power varies with the inverse square of the diffraction angle.

In comparison to WiMAX systems, FSO technology offers the potential of broadband communication capacity over unlicensed optical wavelengths along with its narrow beam divergence and the immunity from being absorbed by water and oxygen. We believe that by equipping it with PAT capability, the FSO system can be a competitive alternative for solving the last-mile problem.

1.1.2 Last Mile Networks

Compared to individual FSO links among backbones and LANs, an FSO-based last-mile network promises better quality of service (QoS) and higher availability through its capability of autonomous reconfigurability to deal with changing atmospheric and traffic conditions in dynamic environments. The new topology, defined as the physical link configuration of the network, is generated from an optimization process subject to multi-objective cost functions regarding several network parameters, such as throughput, packet loss, or power consumption. In FSO communication systems, these network parameters must be retrieved from the PAT system.

Therefore, an ideal PAT system, capable of supporting a last mile network, must have (1) a wide field-of-view (around 2π steradians) to probe the potential targets, (2) a high alignment resolution (around 1 milli-rads) to establish the LOS connectivity, and (3) a short settling time of the tracking loop to compensate for undesired motion. However, these requirements involve tradeoffs, which cannot be satisfied simultaneously within a single design. Lambert [41] and Gibson[21] have both suggested an integrated PAT structure based on two systems:

1. Fine angular PAT (FPAT) system: Building and maintaining an FSO link are the main tasks of an FPAT system, which needs a high pointing resolution and a short settling time. The two criterions can be satisfied much easier by narrowing down the range of uncertainty, defined as the angular aperture of a cone, with its apex at the center of transmitter or receiver of the host, covering the area where the target may reside, through a CPAT system.

2. Coarse angular PAT (CPAT) system: Probing potential targets and reducing the range of uncertainty are the responsibilities of an CPAT system. The former depends on a sensor with a wide view and the latter depends on estimating the pointing vector between the target and the actuator.

A finite-state diagram, plotted in figure 1.2, describes a possible integration between the coarse and fine pointing systems.

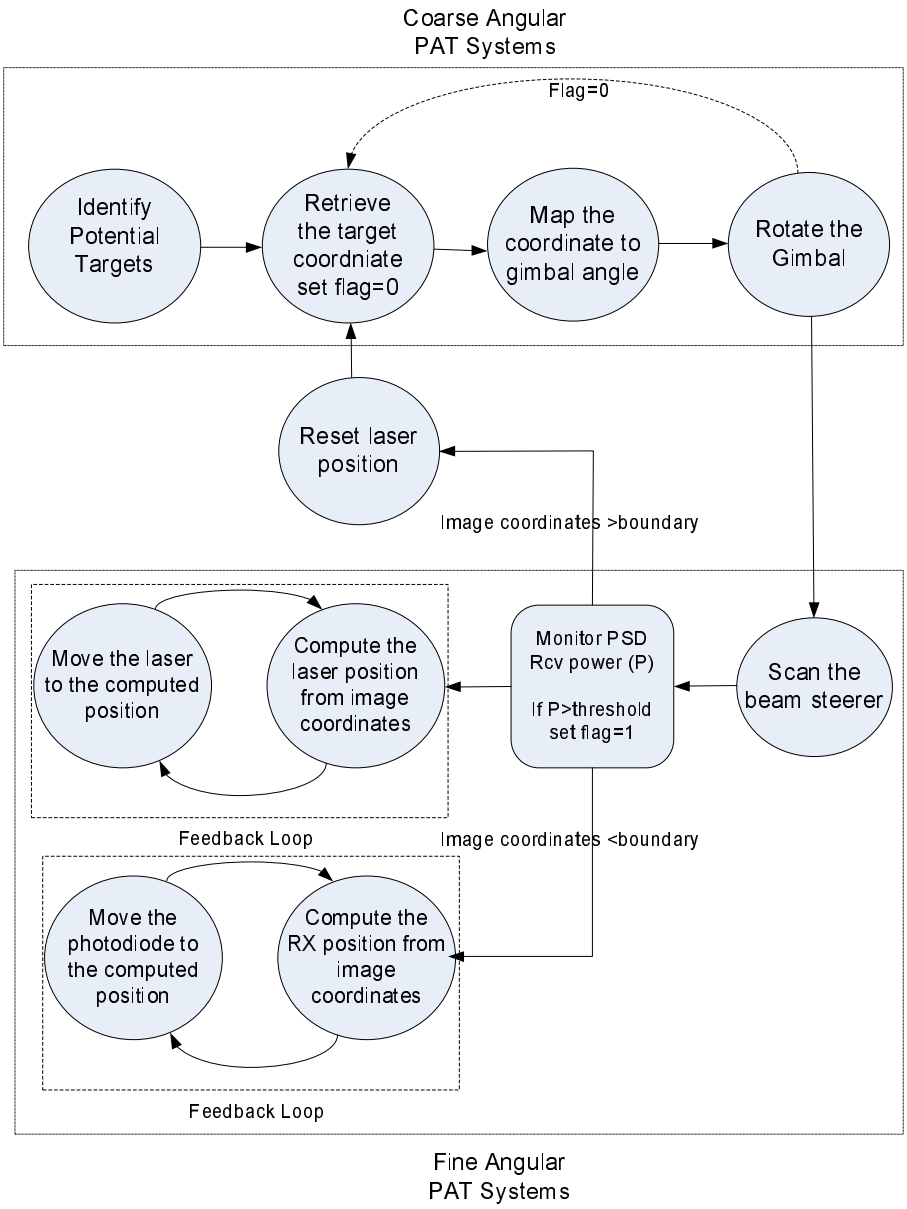

Figure 1.2: Finite state diagram of a PAT system: the upper is the CPAT system and the bottom is the FPAT system.

1.2 Overview of Pointing, Acquisition, and Tracking Systems

1.2.1 Previous Work

Pointing, acquisition, and tracking systems have been successfully implemented in many applications ranging from short-distance cases such as human motion tracking to long distance applications such as missile guidance systems. Different applications may adapt different principles of operation into the PAT system design. Rolland et.al [71] reviewed these techniques and further classified them into seven categories including time of flight (TOF), spatial scan, inertial sensing, mechanical linkages, phase-difference sensing, direct-field sensing, and hybrid methods.

Among these techniques, the spatial scan method, which is based on analyzing the incoming light ray to determine the orientation of a target, is the best match to the capabilities of an FSO system. The sensor of the spatial scan method is usually a combination of a front-end optical system and a position sensing diode (PSD), including coupled charge detectors (CCD), quadrant detectors (QD), and lateral effect detectors (LEP). The CCD-based sensor can simultaneously measure the incident angles for multiple rays, whereas the QD-based and LEP-based sensor can only measure the angle of one ray.

1.2.2 Fine Angular Pointing, Acquisition, and Tracking Systems

The goal of the FPAT system is to complete the link, which implies that the alignment procedure must take the received power into consideration. Also, the FSO system, incorporated with the FPAT system, must be compact enough to be

carried by the actuator of the CPAT system. These two conditions make the spatial scan method the best candidate, because this method (1) determines the orientation of the targets from the same light ray that carries the information bits and (2) can be easily incorporated into a traditional transceiver.

1.2.2.1 Enhanced FSO Transceivers

An FSO transceiver consists of a transmitter and a receiver to achieve duplex transmission. The data in the transmitter is first modulated onto an optical carrier, typically a laser, then the laser beam is collimated through an optical system, and finally transmitted as an optical field into the atmospheric channel. In order to comply with the PAT requirement, beam steering capability must be incorporated into the design, which converts a simple transmitter into a beam steerer. Gibson[21] categorized the fine laser beam steering systems into (1) mechanical and (2) non-mechanical. Mechanical Beam steerers have advantages in their large steering range and inexpensive design. Non-mechanical beam steerer are useful for eliminating potentially bulky mechanical components and can have a high pointing accuracy.

At the receiver, the arriving optical field is first collected through an optical front-end and projected onto a photodiode for signal detection. For a high-speed FSO application, the power collected from the front-end optics may not be focused onto the photodiode because of pointing errors resulting from turbulence or misalignment. A better strategy is to utilize a PSD to first determine the location of the focused spot and then move the photodiode to optimize the received power using

feedback control. An FSO receiver capable of estimating azimuthal and elevation angles is defined as an angular resolver (AR).[1]

The combination of the beam steerer and AR enhances the FSO transceiver with fundamental pointing and tracking capability. If the beam steerer and AR are combined such that their optical axes are identical, the resulting transceiver is denoted as mono-static as in figure 1.3 (a); otherwise it is denoted as bi-static as in figure 1.3 (b).

Generally, mono-static transceivers suffer from strong interference resulting from strong energy coupled from self-reflection between the forward and backward links. Most mono-static transceivers require additional power-isolation devices (e.g. a polarizing beam splitter) to prevent this effect, called narcissus.

1.2.2.2 Introduction to Transceiver Alignment

An FSO link is established if the optical axes of the local/remote beam steerer and the remote/local AR are aligned to the vector connecting between the local/remote beam steerer and remote/local AR, respectively. Since aligning a vector to the other vector in general takes 2 rotations (one in azimuth and the other in elevation), it requires 4 rotations to complete a link and 8 rotations to develop a duplex channel (2 from each beam steerer and AR). In general, the image position in the local AR is capable of providing only the rotation angles for the local AR

[1] A new generation of image sensor with a photoreceiver capability was fabricated in Japan by Kagawa et.al. [34], which can select the best photoreceiver based on the image position without additional motion.

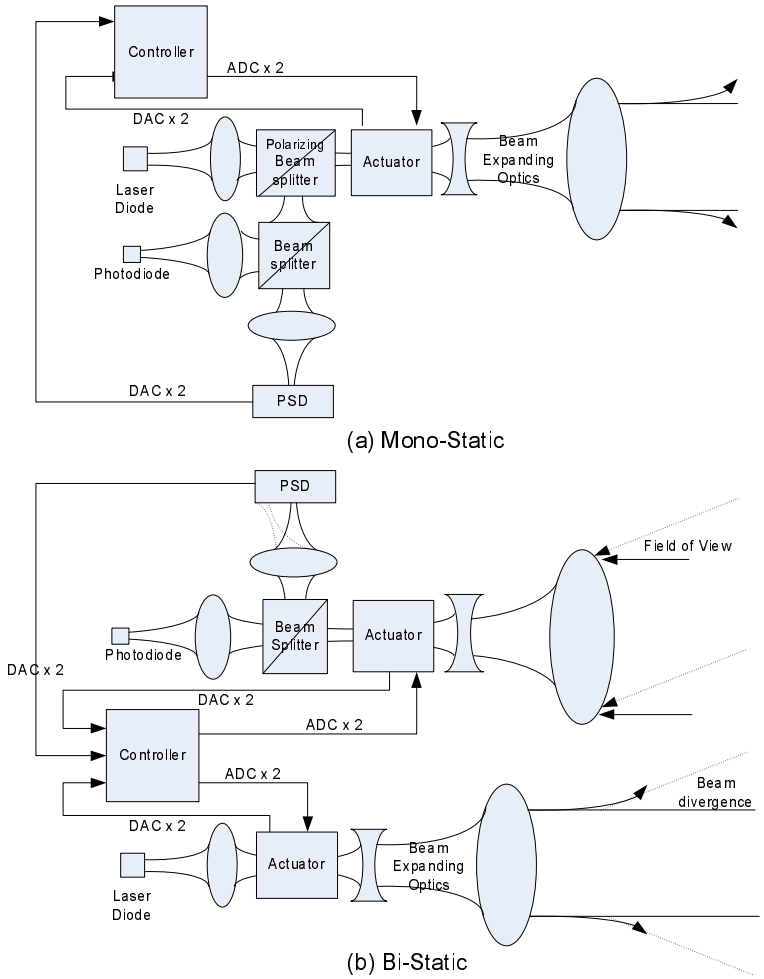

Figure 1.3: Schematic of an FSO transceiver following (a) a mono-static design (b) a bi-static design

which optimizes the received power but not the rotation angles which leads the local beam steerer to the remote AR. A theoretical proof is given in section 2.2. Since each link must be aligned individually, we therefore define this alignment problem as the single alignment problem. The details are depicted in figure 1.4 (a).

If the transceivers are mono-static, and since the optical axes for the local beam steerer and AR are identical, the alignment takes only 4 rotations. Most importantly, the image position in the local AR is sufficient to determine the rotation angles for both the AR and beam steerer, which implies that once either one of the two links is built, the other link can be automatically aligned. Such an alignment problem is defined as a coupled alignment problem because the two links are geometrically related. The details are shown in figure 1.4 (b).

In this work, we propose a scenario where the alignment can still be treated as a coupled alignment problem even though the transceivers are not mono-static. In this scenario, once either one of the two links is built, the other link can be formed since the two links are related by a linear mapping, which can be calibrated in advance. Such a scenario takes place if the following inequality is satisfied (shown in figure 1.4 (c)):

$$\|L\| > \frac{\|T\| \sin\left(\theta_{bdiv} + \theta_{TL}\right)}{\sin \theta_{bdiv}} \tag{1.1}$$

where $\|L\|$ is the distance from the local AR to the remote beam steerer, $\|T\|$ is the displacement between the local beam steerer and AR, θ_{div} is the beam divergence of the beam steerer, and $\theta_{TL} = \arccos \frac{TL}{\|T\|\|L\|}$ is the angle between the vector T and L.

For example, let us consider a duplex communication channel formed by a

pair of bi-static FSO transceivers. Normally, the displacement between the AR and beam steerer of a bi-static transceiver is about 10 cm and the beam divergence is about 1 mrad. Then, a single alignment problem for the bi-static transceivers can be converted into a coupled alignment problem if the distance from the local AR to the remote beam steerer is more than 100 m. Since most of the link lengths of FSO systems are greater than 100 m, it is safe to assume that most FSO transceivers fall into this particular scenario.

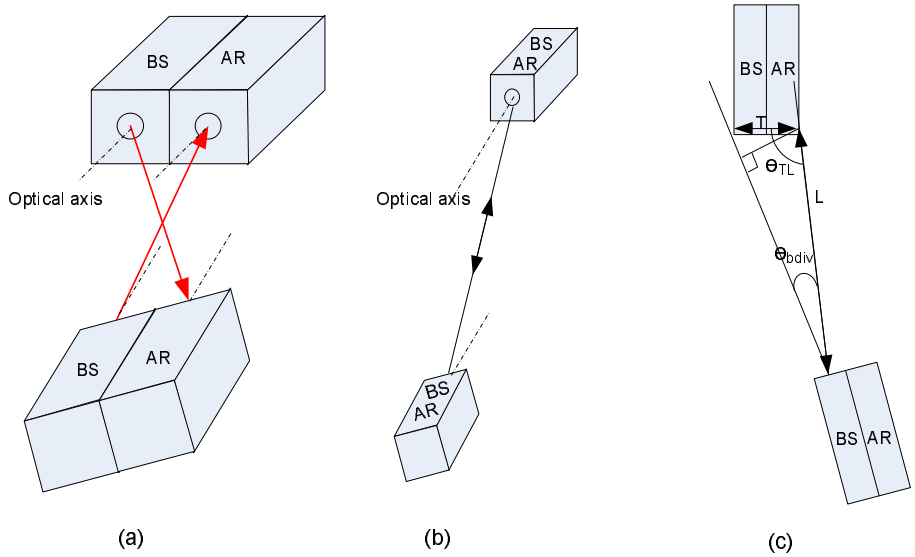

Figure 1.4: Different alignment problems for a pair of FSO transceivers: (a) Single alignment (between two bi-static transceivers with a short link length), (b) Coupled alignment (between two mono-static transceivers), and (c) Coupled alignment (between two bi-static transceivers with a long link length). Note that BS uses as an abbreviation for "beam steerer".

Therefore, we can assume that most FSO transceiver alignments are coupled alignment problems that can always be solved in three steps:

1. Apply a scanning process and point the local beam steerer to the remote AR by trial-and-error.

2. Compute the rotation angles for the remote AR and beam steerer according to the focused spot on the remote AR. Point the remote beam steerer back to the local AR.

3. Compute the rotation angles for the local AR and beam steerer according to the focused spot in the local AR. Point the local beam steerer back to the remote AR.

1.2.2.3 Limitations and Challenges

In the coupled alignment problem, since the pointing angle is computed by transforming the detected image position on the AR with a pre-calibrated mapping, the pointing resolution is directly bounded by the accuracy of the pre-calibrated mapping and AR measurements. The former is determined by the robustness of the algorithm and the latter is affected by (1) quality of the image point, (2) sensitivity of the PSD, (3) range of uncertainty of the AR, and (4) linearity of the AR.

We summarize the limitations of the pointing accuracy below:

1. Robustness of algorithms: The robustness is a performance indicator of an algorithm in the presence of measurement noise. If the algorithm is not robust, the estimated mapping can be far from the true mapping because of measurement errors.

2. Quality of the image spot: If no turbulence is present, the planar wavefront of the incident field produces a tight spot on the PSD through a front-end optical system and the size of the spot defines the image quality. The aberrations of the front-end optical system affect the image position in two ways: (1) the distortion displaces the image position in the radial direction, and (2) other aberrations produce an asymmetric image spot that shifts the centroid of the spot. The theoretical analysis is in section 2.1.2.

3. Sensitivity of the PSD: The sensitivity is a measure of the weakest signal a PSD is able to detect, which is limited by the noise originating in the internal impedance of the detector and the associated electronic circuitry. Given the same received power, a PSD with a higher sensitivity has a higher signal-to-noise ratio (SNR), which usually implies a better resolution. A sensitivity analysis with respect to the QD and LEP is provided in section 3.3.2.3.

4. Range of uncertainty of the AR: The range of uncertainty of the AR is normally equal to its field-of-view and can be approximated as $\theta_{rou} \sim \frac{D}{f}$, where D is the width of the PSD and f is the focal length of the front-end optical system. Given the same sensor resolution, the pointing resolution is inversely proportional to the range of uncertainty.

5. Linearity of the AR: The linearity of the AR is controlled by both the optical system and PSD, including the linear relations embedded (1) between the incident angle and the focused image position and (2) between the focused image position and the digitized position.

Besides the pointing resolution, the other important consideration is the settling time, usually written as $n\%$ of settling time, which is a measure of the time to stabilize an actuator within $\pm n\%$ of a given position [19]. The settling time must be short enough to prevent the target from disappearing from the field-of-view of the AR or the region covered by the beam steerer. The limitation of the settling time is formulated in section 2.3. The feedback and feedforward control systems which are used to achieve this requirement are introduced in sections 2.4 and 2.5.

1.2.3 Coarse Angular Pointing, Acquisition, and Tracking Systems

The goal of the CPAT system is to (1) probe potential targets for network optimization and (2) narrow down the range of uncertainty for the FPAT system. Unlike the FPAT system, which must be incorporated into the FSO transceiver, the design of the CPAT system allows external hardware, which provides more flexibility.

Among the 7 methodologies proposed by Rolland et.al. [71], to the author's knowledge, only the spatial scan method [21] and the hybrid method, a combination of the time-of-flight and the inertial sensing methods [16, 57], have been realized as CPAT systems for FSO transceivers, which may due to the long link length between FSO transceivers.

1.2.3.1 Camera Based Systems

The most popular CPAT systems for commercial FSO transceivers are based on the spatial scan method, including the DT series from Canobeam, FlightLite

series from LightPointe, and TeraScope series from MRV, because of its simplicity and reliability. These transceivers are equipped with a pre-calibrated telescope parallel to the FSO transceiver for coarse alignment. By moving the transceiver until the target overlaps the pre-determined position in the telescope, the coarse alignment is completed. Although the telescope provides a narrow fov, it is still a spatial scanning sensor.

Our efforts focus on improving the previous CPAT system in three areas, namely: (1) a wider field-of-view ($\sim 2\pi$ steradians), (2) controllability over multiple transceivers, and (3) capability of solving the single alignment problem by using multiple sensors.

The current CPAT system is based on a combination of cameras and mechanical rotary gimbals, also known as the camera-based CPAT system. Compared to the narrow field-of-view telescope, off-the-shelf perspective or fisheye cameras are capable of acquiring more potential targets. The main idea is to first explore the geometric invariance among the cameras and the gimbals, calibrate it in advance, and then apply it to the pointing process.

1.2.3.2 GPS and ISS Hybrid Systems

Recently, with the improvement of inertial sensing systems, several papers [16, 57] have investigated a CPAT system, belonging to the hybrid method, consisting of global positioning systems (GPS) and inertial sensing systems (ISS).

GPS sensors determine a target's 3D position by measuring the propagation

time of pulsed signals from the target to at least three references. Two GPS sensors installed in both transceivers provide an accurate measurement of the pointing vector in the frame of the GPS sensors, also known as the earth-centered earth-fixed (ECEF) frame [64]. The pointing vector is in the ECEF frame, which requires a mapping from the ECEF frame to the beam steerer frame. An ISS system consists of three accelerometers and three gyroscopes providing 6 dimensional information for updating any changes in the mapping. The pointing resolution is mostly determined by the accuracy of the estimated mapping from the ISS system.

1.3 Organization

The remainder of this dissertation is organized as follows: Chapter 2 reviews in depth the theoretical analysis and the characterization of an FPAT system. Chapter 3 describes the design, construction, and characterization of a prototype FPAT system based on focal plane motion (FPM). Simulations and experimental investigation are applied to evaluate the beam divergence and pointing resolution. Chapter 4 provides the theoretical framework for the CPAT systems, with an emphasis on the estimation error characterization and wide-angle projective geometry. Chapter 5 introduces robust algorithms to develop prototype CPAT systems based on single or coupled cameras for different FSO scenarios. This dissertation concludes with chapter 6, which summarizes the contributions of this work and suggests several possible directions for future improvement.

Chapter 2

Theoretical Studies of Fine Angular Pointing, Acquisition, and Tracking System

This section provides a theoretical background for a fine angular PAT system design. It covers five important topics, which include:

1. Paraxial Imaging and Aberration theory: The paraxial approximation characterizes a ray transformation by taking the zero, first, and second order deformation of the wavefront into account, including piston, x-tilt, y-tilt, and defocus in the Zernike polynomial [10]. It provides a linear relationship between the image position and incident angles. It has been shown that the paraxial approximation characterizes the performance of the optical system better when the optical paths of the system are closer to the optical axis. Besides, since this approximation matches with human visual perception, known as the perspective view since the Renaissance [30], most optical imaging systems, used to reproduce the human view, can be modelled by this approximation with high accuracy.

   The paraxial approximation characterizes the first three orders of the wavefront. The rest of the components of the wavefront, defined as the wavefront aberrations, result in the difference between the paraxial and true image positions, known as the ray aberrations. The design of an optical imaging system

is to minimize the aberrations in order to reproduce the perspective images seen in human vision.

In this section, the main aberrations, known as Seidel aberrations, are characterized to provide a general understanding. Notice that the compensation of the aberrations is not the focus of this dissertation since the aberrations are relatively small and vary with respect to the environment and setup (e.g. ambient light or shadows). Instead, the ray aberrations are treated as additive noise, which affects the estimation accuracy of the geometric parameters and our goal is to investigate estimation algorithms that are still robust in the presence of noise.

2. Homography Mapping: This section distinguishes our FPAT system from others. Most FPAT systems operate on an ideal assumption that both the the beam steerer and AR have parallel optical axes, where the mapping between the local AR and beam steerer is only an Euclidean transformation [23], including only the scaling factors in the X and Y direction and a centroid translation. This assumption is only true through a deliberate pre-alignment.

   In this section, we prove that the mapping between the local AR and beam steerer is a linear homography as long as the FSO transceivers satisfy the inequality in 1.1. Adapting this homography concept into the FPAT design can improve the pointing resolution and reduce the design complexity.

3. Image Motion Analysis: The LOS paths between FSO transceivers may be blocked by the motion of the remote or local transceivers. These motions can

be divided only into rotation and translation because the transceivers are rigid bodies. This analysis provides solutions to two questions: (1) What kind of motion most likely results in loss of LOS path? and (2) How quickly should the LOS path be re-established before the induced motion breaks it again?

4. Feedback Control: Linear Feedback controllers provide their desired output by feeding errors back to the actuator. The dynamics of most actuators can be characterized as a linear model with different states [14]. The states are often considered as kinematic parameters (e.g. position, velocity and acceleration) in the modelling process. Those states affecting the control input must be updated by referring to the measurements and the actuator model. The estimator gain is a weighting parameter, which estimates the state parameters from the current measurement. The major task of the feedback can be divided in three steps:

    - Apply a control input to the actuator and measure the difference between the desired and measured sensor output.
    - Update the state parameters with the product of the difference and the estimator gain.
    - Update the control input with the product of state parameters and the control gain.

    In general, overwhelming control and estimator gains produce instability whereas a small feedback gain results in a slow response. The best control gain can be

selected using optimal control theory and the best estimator gain is normally determined through a Kalman estimation process, which provides the estimated state parameters maximizing the likelihood function if the noises are Gaussian. Most importantly, a Kalman estimation process updates the states recursively, which minimizes the computational load of the microprocessor.

5. Feed-Forward Control Theorem: Feedback control usually assumes the noise in the model or the measurements is white, which is defined as a sequence whose autocorrelation is a delta function. For colored noise, a linear dynamic model can be created to whiten the colored noise. By augmenting this new linear model into the original dynamics, a feed-forward controller is created.

   The name feed-forward is given because the computed control input, delivered to the actuator, contains a component that can cancel the output produced by the colored noise. For example, building sway results in random disturbances in FSO alignment; however, these disturbances are not purely random but can be modelled by a second-order spring system. The feed-forward controller can better reject such disturbances. Applying feedforward control into a PAT system was proposed by Skormin et.al. [58].

2.1 Optical Imaging Theory

2.1.1 Paraxial Imaging

In general, ray tracing in an optical system depends on solving Snell's law or using Fermat's principle [10], whose solution depends on perfect knowledge of each element of the system and the overall configuration. Normally, this information is unavailable for most commercial optical systems. Also, the nonlinearity of the solution results in one-to-many correspondences between the incident ray vector and the image position, which implies a spot on the image may result from rays with different incident angles.

By introducing the paraxial approximation, Snell's law and propagation law describing the optical ray path can be approximated by [13]

$$\begin{aligned} r' &\sim \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} \begin{pmatrix} r \\ \theta \end{pmatrix} && \text{; Propagation Law} \\ \theta' &\sim \begin{pmatrix} 1 & 0 \\ \frac{h(n_2-n_1)}{Rn_1} & \frac{n_2}{n_1} \end{pmatrix} \begin{pmatrix} r \\ \theta \end{pmatrix} && \text{; Snell's Law} \end{aligned} \tag{2.1}$$

where $(r, \theta)^T, (r', \theta')^T$ are the position and angle of the ray before and after propagation, d is the propagation distance, (n_1, n_2) are the refractive indices of the two media, h is the incident height, and R is the radius of the conic interface.

These equations imply that the nonlinear ray propagation process can be simplified as the product of the 2 ray transfer matrices. Without loss of generality, let the ray transfer matrix of an imaging system be $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, which can be decomposed

as

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & h_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \begin{pmatrix} 1 & h_1 \\ 0 & 1 \end{pmatrix}$$

where $h_1 = \frac{d-1}{c}$ represents the first principal plane, where the ray passing through the first focal point F_1 travels parallel to the axis, and $h_2 = \frac{a-1}{c}$ represents the second principal plane where the ray parallel to the optical axis passes through the second focal point F_2. F_1, F_2 are shown in figure 2.1

If an object is placed far from the imaging system, according to the paraxial approximation, its image is produced at F_2, the intersection between the optical axis and the current paraxial marginal rays. This focusing property helps to solve $d_2 = -a/c$ and the relation between the incident rays and projected distances can be determined as:

$$\begin{pmatrix} r_i \\ \theta_i \end{pmatrix} = \begin{pmatrix} 1 & d_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & d_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} r_o \\ \theta_o \end{pmatrix} \tag{2.2}$$

$$= \begin{pmatrix} 0 & b - \frac{ad}{c} \\ c & d \end{pmatrix} \begin{pmatrix} r_o + d_1\theta_o \\ \theta_o \end{pmatrix}, \tag{2.3}$$

where (r_o, θ_o) is the radius and angle at the object plane, and (r_i, θ_i) is the radius and angle at the image plane.

Generally, the media of the image plane and the object plane are identical, which provides a constraint that the determinant of the ray transfer matrix is 1. Applying this constraint, the previous equation can be cast as

$$\begin{pmatrix} r_i \\ \theta_i \end{pmatrix} = \begin{pmatrix} 0 & -\frac{1}{c} \\ c & d \end{pmatrix} \begin{pmatrix} r_o + d_1\theta_o \\ \theta_o \end{pmatrix}. \tag{2.4}$$

By defining the focal length $f = -\frac{1}{c}$, the linear model between the projected image position and the incident angle can be drawn as:

$$r_i = f\theta_o \tag{2.5}$$

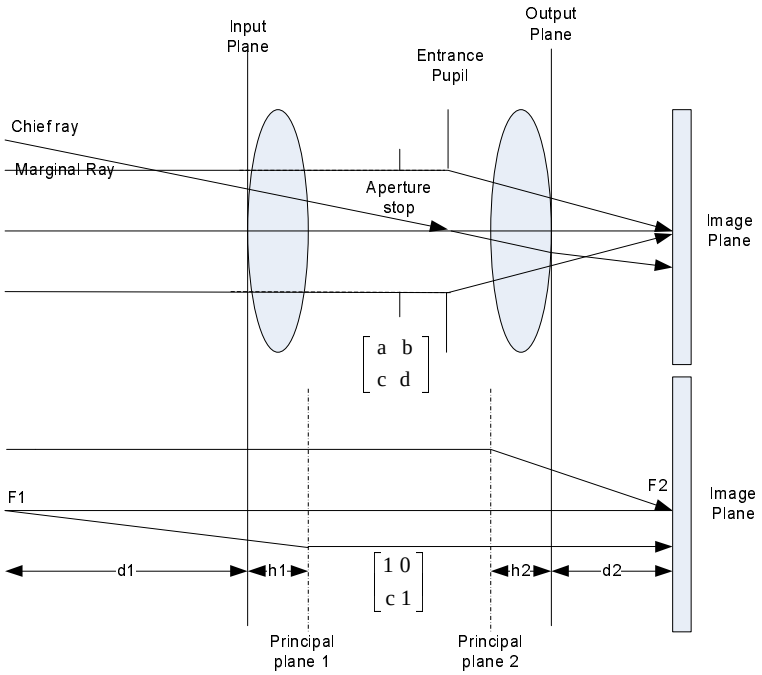

Figure 2.1: Imaging model geometry

This result provides a linear and one-to-one mapping between the incident angles and the projected image position and the mapping contains only one unknown. This implies that by estimating f from limited measurements, all the incident angles and image position pairs are uniquely determined.

2.1.2 Aberration Theory

The aberrations, resulting from high-order residues of the wavefront, produce the difference between the true and paraxial image position. The residues of the

wavefront and the displacement between the true and paraxial image position are known as the wavefront and ray aberrations, respectively. In general, aberrations can be categorized into monochromatic and chromatic and only the former exists in FPAT system since the operating laser is a monochromatic source.

Let V be the optical path length, defined as $\int n(s)dr(s)$, where $r(s)$ is the ray path. According to Maxwell equations, if the two rays have an identical origin but experience two optical paths lengths V_1, V_2 in the same uniform medium, the difference between the two ray vectors can be cast as

$$dr = \frac{1}{n}(\frac{\delta(V_1 - V_2)}{\delta x}, \frac{\delta(V_1 - V_2)}{\delta y}, \frac{\delta(V_1 - V_2)}{\delta z}) \tag{2.6}$$

where n is the refractive index of the medium.

The third order primary aberrations, known as the Seidel aberrations, can be characterized by two paraxial rays, the paraxial chief ray and the axial ray [59]. Figure 2.2 shows the optical path difference between the paraxial chief ray (from O to P_0) and one of the oblique rays (from O_1 to P). Note that O, O_0, O_1 are on the same wavefront. The wavefront aberrations (up to third order) can be formulated as

$$\begin{aligned} W &= n(O_0Q - O_1Q) = \triangle n(O_1Q) \\ &\sim \triangle n(OP_0 - QP_0) \end{aligned}$$

where $\triangle$ represents the increment of the optical path.

Let the coordinates of O and Q be $(0, \bar{y}, \bar{z})$ and (x, y_1, z), respectively. The phase difference W is

$$W = \triangle n(((l - \bar{z})^2 + (\eta - \bar{y})^2)^{\frac{1}{2}} - ((l - z)^2 + (\eta - y_1)^2 + x^2)^{\frac{1}{2}}) \tag{2.7}$$

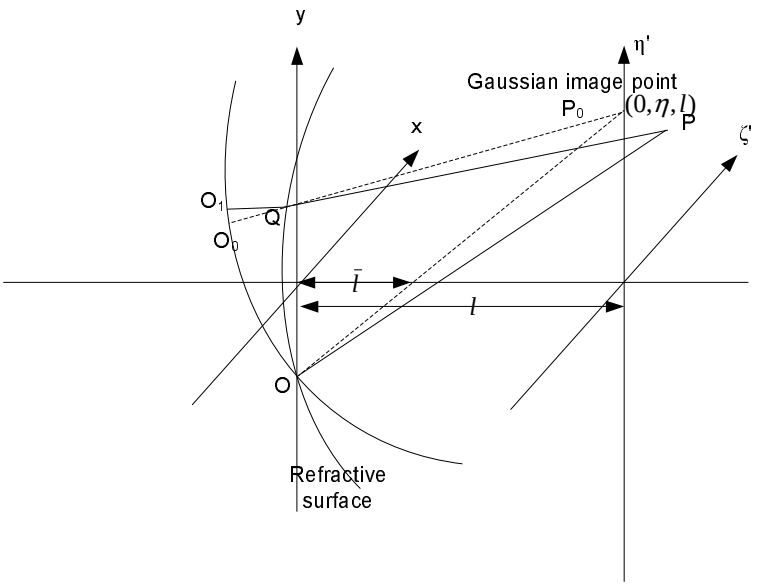

Figure 2.2: Optical path difference in the approximation of the primary aberrations

After plugging into the spherical equation of the refractive surface along with the Lagrange invariant ($\triangle n(\frac{\eta y}{l}) = 0$) and refractive invariant ($\triangle n(c - \frac{1}{l}) = 0$), the constant and second order terms of W are eliminated. $\bar{y}$ is the paraxial chief ray and η is its corresponding image, which can be determined by $\frac{\eta}{l} = \bar{y}(\frac{1}{l} - \frac{1}{\bar{l}})$. By replacing η, y_1 with their paraxial equivalents, the Seidel aberrations passing through a single surface can be derived as

$$\begin{aligned} W &= \triangle n(\tfrac{1}{8} - \tfrac{c}{4l^2} + \tfrac{1}{8l^3})(x^2+y^2)^2 \\ &\quad + \triangle n(\tfrac{c^3}{2} - \tfrac{c}{2l\bar{l}} - \tfrac{c}{2l^2} + \tfrac{1}{2l^2\bar{l}})\bar{y}y(x^2+y^2) \\ &\quad + \triangle n(\tfrac{c^2}{2} - \tfrac{c}{l\bar{l}} + \tfrac{1}{2l\bar{l}^2})\bar{y}^2y^2 \\ &\quad + \triangle n(\tfrac{c^3}{4} - \tfrac{c}{4\bar{l}^2} - \tfrac{c}{4l^2} + \tfrac{1}{4l\bar{l}^2})\bar{y}^2(x^2+y^2) \\ &\quad + \triangle n(\tfrac{c^3}{2} - \tfrac{c}{2l\bar{l}} - \tfrac{c}{2l^2} + \tfrac{1}{2\bar{l}^3})\bar{y}^3y \end{aligned}$$

After summarizing the aberrations from all the surfaces and introducing the maximum image height (η_{max}) and maximum height at the exit pupil (h_p) from the

paraxial chief ray and axial ray, respectively, the Seidel wavefront aberrations are cast as [74]

$$\begin{aligned}
W &= \frac{1}{8}S_I\frac{(x_p^2+y_p^2)^2}{h_p^4}+\frac{1}{2}S_{II}\frac{y_p(x_p^2+y_p^2)}{h_p^3}\frac{\eta}{\eta_{max}}+\frac{1}{2}S_{III}\frac{y_p^2}{h_p^2}\frac{\eta^2}{\eta_{max}^2} \\
&\quad +\frac{1}{4}(S_{III}+S_{IV})\frac{x_p^2+y_p^2}{h_p^2}\frac{\eta^2}{\eta_{max}^2}+\frac{1}{2}S_V\frac{y_p}{h_p}\frac{\eta^3}{\eta_{max}^3} \\
S_I &= -\sum A^2 h\triangle(\frac{u}{n}) \\
S_{II} &= -\sum \bar{A}Ah\triangle(\frac{u}{n}) \\
S_{III} &= -\sum \bar{A}^2 h\triangle(\frac{u}{n}) \\
S_{IV} &= -\sum H^2 c\triangle(\frac{1}{n}) \\
S_V &= -\sum \frac{\bar{A}}{A}h\triangle(\frac{u}{n})+\frac{\bar{A}}{A}H^2 c\triangle(\frac{1}{n})
\end{aligned}$$

where (x_p, y_p) represent the location of the ray at the exit pupil, (η) is the image height, $A = nh(c - 1/l)$ represents the paraxial marginal angle times the refractive index of the medium, $\bar{A} = n\bar{h}(c-1/\bar{l})$ represents the incident angle of the chief ray times the refractive index of the medium, and H represents the Lagrange invariant, which can be expressed as $n(\bar{h}u - h\bar{u})$.

The ray vector resulting from the wavefront aberration can be formulated as equation 2.6. If the propagation distance from exit pupil to the image plane is R, the corresponding image shifts, defined as the ray aberrations, are summarized below:

Spherical aberration:

$$\begin{pmatrix} \delta\zeta & \delta\eta \end{pmatrix} = \begin{pmatrix} -\frac{R}{n}S_I\frac{r_p^2}{2h_p^4}x & -\frac{R}{n}S_I\frac{r_p^2}{2h_p^4}y \end{pmatrix}$$

Coma:

$$\begin{pmatrix} \delta\zeta & \delta\eta \end{pmatrix} = \begin{pmatrix} -\frac{R}{n}S_{II}\frac{2x_py_p}{h_p^3}\frac{\eta}{\eta_{max}} & -\frac{R}{n}S_{II}\frac{x^2+3y^2}{h_p^3}\frac{\eta}{\eta_{max}} \end{pmatrix}$$

Astigmatism and coma:

$$\begin{pmatrix} \delta\zeta & \delta\eta \end{pmatrix} = \begin{pmatrix} \frac{R}{2n}(S_{III}+S_{IV})\frac{x}{h_p^2}\frac{\eta^2}{\eta_{max}^2} & -\frac{R}{n}S_{III}\frac{y}{h_p^2}\frac{\eta^2}{\eta_{max}^2}+\frac{R}{2n}(S_{III}+S_{IV})\frac{y}{h_p^2}\frac{\eta^2}{\eta_{max}^2} \end{pmatrix}$$

Distortion:

$$\begin{pmatrix} \delta\zeta & \delta\eta \end{pmatrix} = \begin{pmatrix} 0 & -\frac{R}{2n}S_V\frac{\eta^3}{\eta_{max}^3} \end{pmatrix}$$

We plot the Seidel ray aberrations, obtained from different optical systems, in figure 2.3. It can be seen clearly from the plot that the aberrations, except for spherical aberrations, produce a non-symmetrical beam shape, which result in the shift of the centroid.

2.2 Homography Mapping

Most alignment problems can be categorized as coupled alignment problems since the link distance satisfies the inequality in 1.1. The displacement between the beam steerer and AR of the PAT system is negligible in the coupled alignment problem, which implies only an unknown rotation matrix between the beam steerer and AR. Let the imaging system satisfy the paraxial approximation. Any ray vector expressed in the frame of AR and beam steerer is related by

$$v_{BS} = Rv_{AR} \tag{2.8}$$

where $R = R_\alpha R_\beta R_\gamma$ and $R_\alpha, R_\beta, R_\gamma$ represent rotation matrices with respect to yaw, pitch, and roll angles. Note that R is an identity matrix in the mono-static transceiver.

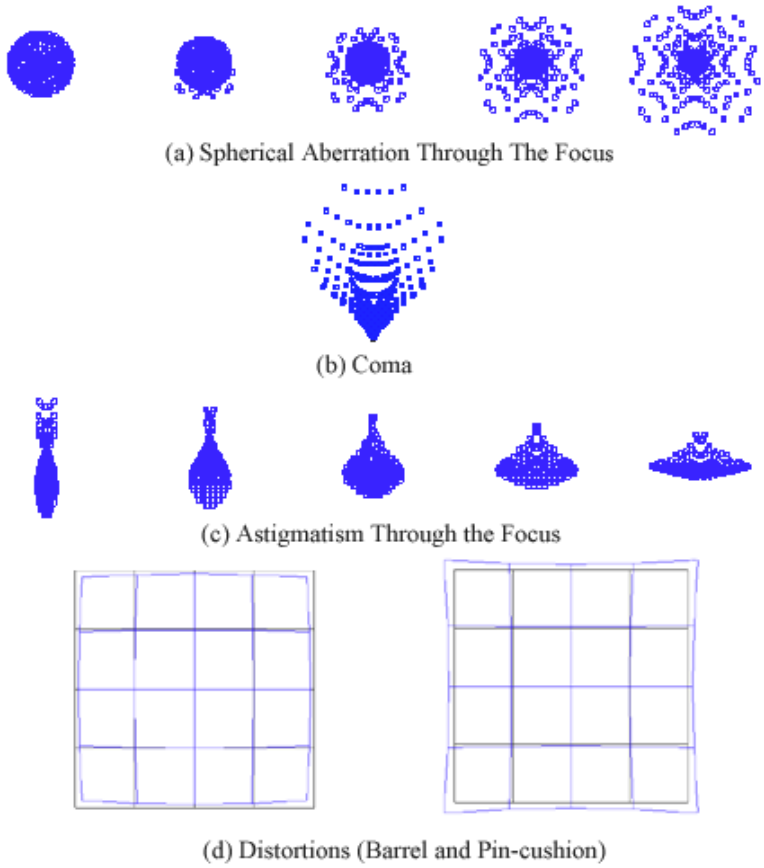

Figure 2.3: Spot diagrams resulting from the Seidel aberrations (a) spherical aberrations (b) coma (c) astigmatism (d) distortions (barrel and pin-cushion). Except for the spherical aberrations, any other aberrations result in an asymmetric spot diagram and a centroid shift.

The transmitted ray vector in the beam steerer results from moving a laser diode, preferred to be a pig-tailed fiber laser for its light weight, on a plane parallel to the focal plane of the collimator. If the fiber is close to the optical axis, the resulting transmitted vector can be approximated as

$$v_{BS} = \begin{pmatrix} \frac{1}{f_{coll}} & 0 \\ 0 & \frac{1}{f_{coll}} \end{pmatrix} \left(\begin{pmatrix} d_{LA}^x \\ d_{LA}^y \end{pmatrix} - \begin{pmatrix} o_x \\ o_y \end{pmatrix} \right)$$

$$\begin{pmatrix} d_{LA}^x \\ d_{LA}^y \\ 1 \end{pmatrix} = \begin{pmatrix} f_{coll} & 0 & o_x \\ 0 & f_{coll} & o_y \\ 0 & 0 & 1 \end{pmatrix} v_{BS}$$

where f_{coll} is the focal length of the collimator, (d_{LA}^x, d_{LA}^y) are the fiber coordinates, and (o_x, o_y) is the coordinates of the intersection between the optical axis and the motion plane.

The motion of the fiber may be provided by a two-axis moving platform whose moving axes are not necessarily orthogonal. Also, the sensors used to measure the fiber position may have different axial gains and a different zero position. Taking these into consideration, the mapping between the sensor output and the transmitted ray vector can be generalized to be an affine transformation matrix M_{LA}, a 3×3 upper triangular matrix, instead of the Euclidean transformation matrix used in the last equation.

The PSD used to measure the received ray vector also suffers from similar problems, such as different axial gains and zero positions. Analogously, the image readout from the AR and the received ray vector can also be related by an affine transformation matrix. Combined with equation 2.8, the readouts from the AR and

the laser position sensors can be described by a 3×3 linear homography:

$$\begin{pmatrix} d^x_{LA} \\ d^y_{LA} \\ 1 \end{pmatrix} = \lambda M_{LA} R M^{-1}_{PSD} \begin{pmatrix} d^x_{PSD} \\ d^y_{PSD} \\ 1 \end{pmatrix} \tag{2.9}$$

where (M_{LA}, M_{PSD}) are affine transformation matrices with form $\begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & 1 \end{pmatrix}$ and λ is an unknown scaling factor.

$M_{LA} R M^{-1}_{PSD}$ is a 3×3 matrix and can be calibrated in advance. The calibration procedure and algorithm are listed in section 3.3.3 and algorithm 5.1, respectively.

With a pre-calibrated homography and a readout from AR indicating the target location, the desired output from the laser position sensor is

$$\begin{aligned} d^x_{LA} &= \frac{H^1 \begin{pmatrix} d^x_{PSD} & d^y_{PSD} & 1 \end{pmatrix}^T}{H^3 \begin{pmatrix} d^x_{PSD} & d^y_{PSD} & 1 \end{pmatrix}^T} \\ d^y_{LA} &= \frac{H^2 \begin{pmatrix} d^x_{PSD} & d^y_{PSD} & 1 \end{pmatrix}^T}{H^3 \begin{pmatrix} d^x_{PSD} & d^y_{PSD} & 1 \end{pmatrix}^T} \end{aligned} \tag{2.10}$$

where H^i is the i^{th} row of H and $H = M_{LA} R M^{-1}_{PSD}$.

As for received power optimization, since there is generally no displacement between the photodiode and the AR, the mapping between the AR and the best photodiode position shares the same relationship, a linear homography.

One then concludes that once the incoming ray vector is revealed from the AR output, the positions of the laser diode and the photodiode can be found from

equation 2.10. The linear homography required in equation 2.10 can be calibrated in advance by algorithm 5.1.

2.3 Image Motion Analysis

Define a coordinate system with origin at the local transceiver and let the remote transceiver be at (X, Y, Z). The remote and local transceivers have translation velocities of t^r, t^l and angular velocity of ω^r, ω^l, respectively. The relative velocity with respect to the local host can be cast as

$$\frac{dP}{dt} = (t^r - t^l) + \omega^r \times R - \omega^l \times P \tag{2.11}$$

where $P = (X, Y, Z)$ represents the vector from the local to the remote transceiver and R represents the distance from the object centroid to the remote rotation center.

Assuming the imaging system has a focal length of f, the location of the remote transceiver in the local image frame can be approximated as

$$p = f\frac{P}{Z} + O$$

where $O : (o_x, o_y)$ is the origin shift between the camera and image frame.

The image velocity can be computed by taking the derivative of p with respect to t. By putting the terms related to the link length (Z) together, the image motion can be cast as follows

$$\begin{aligned} \frac{dp}{dt} &= \begin{pmatrix} \frac{f((t_x^r - t_x^l) + R_z\omega_y^r - R_y\omega_z^r) - p_x((t_z^r - t_z^l) + R_y\omega_x^r - R_x\omega_y^r)}{Z} \\ \frac{f((t_y^r - t_y^l) + R_x\omega_z^r - R_z\omega_x^r) - p_y((t_z^r - t_z^l) + R_y\omega_x^r - R_x\omega_y^r)}{Z} \end{pmatrix} \\ &+ \begin{pmatrix} \omega_x^l \frac{p_x p_y}{f} - \omega_y^l(f + \frac{p_x^2}{f}) + \omega_z^l y_p \\ \omega_x^l(f + \frac{p_y^2}{f}) - \omega_y^l \frac{p_x p_y}{f} - \omega_z^l y_p \end{pmatrix} \end{aligned} \tag{2.12}$$

where (p_x, p_y) are the image coordinates.

2.3.1 Pointing Induced Tracking Errors

The link length between two FSO systems is normally long, which implies that the image motion in the local AR is only affected by the local angular velocity since the first term of equation 2.12 is negligible. In other words, the motion from the target is immeasurable in the local AR, which is known as the pointing induced tracking errors (PITE) problem, commonly described by FSO engineers.

LOS paths are kept once the transceiver stays within the beam divergence of the other transceiver and a tracking loop must be applied to the link maintenance because of the limited beam divergence. Since the AR in each transceiver only provides half of the motion, the LOS paths can only be maintained by cooperation between the transceiver pair.

2.3.2 Minimum Settling Time

Generally, once an image spot is imaged on the PSD, the image's coordinates are mapped to a desired output from the laser position sensor in the beam steerer. However, because of the errors coupled from the measurement, calibration and target motion, the computed position may be off from the true position. As long as the offset is within the region created by the beam divergence of the beam steerer, the link can still be established.

The back-projection of this region in the PSD frame is the tolerance region,

which serves as a bound for the aberrations and the image motion caused by mobility. In other words, the control loop in the beam steerer must settle the laser diode to the computed steering angles before the image spot drifts out of the tolerance region. It can be seen that a large tolerance region is beneficial to the link alignment but may simultaneously damage the quality of the link because a large tolerance region implies a large beam divergence.

Assuming the tolerance region is a small circle, the minimum settling time is defined as

$$T_{set} = \frac{r_{tol}}{\left\| \frac{dp}{dt} \right\|}$$

where r_{tol} is the radius of the tolerance region.

2.4 Feedback Controller

Once the positions of the laser/photodiode are determined from the image position at the AR, actuators must be applied to move the laser/photodiode toward the computed location fast and accurately, which relies on a feedback controller.

This section provides the concepts of the digital feedback control, including the optimal control theory and the Kalman estimator, an optimal estimator for Gaussian noise [20]. Throughout this section, the digitized actuator model is assumed to be linear and accepts control input (e.g. voltage or current) from an external port. Sensors, also suffering from additive random noises, are attached to the actuator to

probe the current status (e.g. displacement, velocity). The overall model is cast as

$$\begin{aligned} x(k+1) &= \Phi x(k) + \Gamma u(k) \\ y(k) &= Cx(k) \end{aligned} \tag{2.13}$$

where k is the time constant, x is the state of the actuator, u is the control input, y is the sensor output, Φ is the state transition matrix for the evolution of x, Γ is the gain of the control input, and C is the response of the sensor, known as the observation matrix.

2.4.1 Optimal Control

Optimal control is attractive for its ability to select the desired pole locations for the single input single output (SISO) system and handle the multi-input multi-output (MIMO) system.

The optimal control is to select the control input that minimizes the control power as well as the difference between the desired output and the measurements. If the desired output is 0, the cost function ζ is best cast as

$$\begin{aligned} \zeta &= \min_u \frac{1}{2} \sum_{k=0}^{N} x^T(k) Q_1 x(k) + u^T(k) Q_2 u(k) \\ \text{subject to} \quad & x(k+1) = \Phi x(k) + \Gamma u(k) \end{aligned} \tag{2.14}$$

where k is the time index, (Q_1, Q_2) are symmetrical and nonnegative definite weighting matrices and $Q_1 = C^T C$.

Using the Lagrange method, we can derive the Euler-Lagrange equations,

whose state transition matrix is known as the control Hamiltonian matrix.

$$\begin{pmatrix} x(k+1) \\ \lambda(k+1) \end{pmatrix} = \begin{pmatrix} \Phi + \Gamma Q_2^{-1}\Gamma^T\Phi^{-T}Q_1 & -\Gamma Q_2^{-1}\Gamma^T\Phi^{-T} \\ -\Phi^{-T}Q_1 & \Phi^{-T} \end{pmatrix} \begin{pmatrix} x(k) \\ \lambda(k) \end{pmatrix} \tag{2.15}$$

where $\lambda(k)$ is the Lagrange coefficient at time k.

By imposing an innovative constraint, $\lambda(k) = S(k)x(k)$, to the last equation, the Riccati equation describing the evolution of $S(k)$ and the optimal control gain $K(k)$ relating $u(k)$ and $x(k)$ can both be derived as follows.

$$\begin{aligned} S(k) &= \Phi^T(S(k+1) - S(k+1)\Gamma R^{-1}\Gamma^T S(k+1))\Phi + Q_1 \\ K(k) &= (Q_2 + \Gamma^T S(k+1)\Gamma)^{-1}\Gamma^T S(k+1)\Phi \end{aligned} \tag{2.16}$$

These equations show that the optimal gain can only be solved backwards in time since the appropriate boundary condition exists only after reaching the steady state.

If the controller runs for a long period, instead of computing the time-varying gain, the steady state $S(\infty)$ can be solved directly from the quadratic algebraic Riccati equation, and the positive definite root is the desired solution for $S(\infty)$. The algebraic Riccati equation is listed below

$$S(\infty) = \Phi^T(S(\infty) - S(\infty)\Gamma R^{-1}\Gamma^T S(k+1))\Phi + Q_1$$

A better algorithm to solve S_∞ is through eigendecomposition of the control Hamiltonian matrix. Since the eigenvalues are reciprocal, the state transition matrix

in the Euler-Lagrange equation can be decomposed as

$$\begin{aligned} H_c &= W H_c^* W^{-1} \\ &= \begin{pmatrix} X_I & X_0 \\ \Lambda_I & \Lambda_0 \end{pmatrix} \begin{pmatrix} E^{-1} & 0 \\ 0 & E \end{pmatrix} \begin{pmatrix} X_I & X_0 \\ \Lambda_I & \Lambda_0 \end{pmatrix}^{-1} \end{aligned}$$

where $H_c = \begin{pmatrix} \Phi + \Gamma Q_2^{-1} \Gamma^T \Phi^{-T} Q_1 & -\Gamma Q_2^{-1} \Gamma^T \Phi^{-T} \\ -\Phi^{-T} Q_1 & \Phi^{-T} \end{pmatrix}$, E^{-1} is a diagonal matrix containing the stable eigenvalues ($z < 1$), $(X_I^T, \Lambda_I^T)^T$ are the eigenvectors associated with the stable eigenvalues, and $(X_0^T, \Lambda_0^T)^T$ are the eigenvectors associated with the unstable eigenvalues.

Therefore, a new state $(x^*(k), \lambda^*(k))$ is obtained by multiplying the inverse of the eigenvectors (W) to the old state, and equation 2.15 becomes

$$\begin{pmatrix} x^*(k+1) \\ \lambda^*(k+1) \end{pmatrix} = H_c^* \begin{pmatrix} x^*(k) \\ \lambda^*(k) \end{pmatrix}$$

The convergence of $\lambda^*(k)$ concludes that the initial condition of $\lambda^*(0) = 0$. After converting this result to the original state, the $S(\infty)$ and optimal gain in steady-state can be derived as

$$\begin{aligned} S(\infty) &= \Lambda_I X_I^{-1} \\ K(\infty) &= (Q_2 + \Gamma^T S(\infty) \Gamma)^{-1} \Gamma^T S(\infty) \Phi \end{aligned} \tag{2.17}$$

2.4.2 Kalman Estimator

From equation 2.17, the optimum control input is the product of the optimum gain constant and the current state information. The former is computed from

optimal control theory and the latter must be from an optimal estimator for two reasons: (1) sensor capabilities are limited and insufficient for probing all the states for most systems and (2) the measurements are inevitably affected by noise (e.g. circuit thermal noises). The Kalman filtering process provides an optimal estimator for all the states if the measurement noises are additive Gaussian. Most importantly, it provides a recursive updating process that is crucial for real-time implementation.

A noise source is further added into equation 2.13 to evaluate the best estimator, as

$$\begin{aligned} x(k+1) &= \Phi x(k) + \Gamma u(k) + \Gamma_1 w(k) \\ y(k) &= Hx(k) + v(k) \end{aligned} \tag{2.18}$$

where $\Gamma_1 w(k)$ is the state disturbance resulting from a Gaussian distribution $N(0, R_w)$ and $v(k)$ is the measurement error with a Gaussian distribution $N(0, R_v)$.

The two noise sources result in the prediction and estimation stages of the Kalman filtering process. The former is to estimate the best current state, defined as the prior state, based on the previous states and the latter is to update the prior state to estimate the best posterior state from the measurement. Three notations are first defined before introducing the recursive algorithm.

$$\begin{cases} x(k), & \text{the true state at time } k; \\ \hat{x}_{k|k-1}, & \text{the prior state at time } k; \\ \hat{x}_{k|k}, & \text{the posterior state at time } k. \end{cases}$$

2.4.2.1 Prediction

Since $w(k)$ is independent of $x(k)$, the best prior state can be estimated as

$$\hat{x}_{k+1|k} = \Phi\hat{x}_{k|k} + \Gamma u(k) \tag{2.19}$$

Comparing the best estimator to the state space model, the error covariance matrix, defined as $M(k+1)$, is equal to

$$\begin{aligned} M(k+1) &= E((x(k+1) - \hat{x}_{k+1|k})(x(k+1) - \hat{x}_{k+1|k})^T) && (2.20) \\ &= \Phi\Sigma_{x_{k|k}}\Phi^T + \Gamma_1 R_w \Gamma_1^T && (2.21) \end{aligned}$$

2.4.2.2 Estimation

Given a linear model $Y = HX + V$, where V is independent to X, the linear least squares estimator of X and the error covariance matrix are as follows [54]. A detailed explanation can be found in the appendix.

$$\begin{aligned} \hat{X} &= E(X|Y) && (2.22) \\ &= \mu_X + \Sigma_{XY}\Sigma_Y^{-1}(Y - \mu_Y) \\ \Sigma_{\hat{X}} &= \Sigma_X - \Sigma_X H^T \Sigma_Y^{-1} H \Sigma_X && (2.23) \end{aligned}$$

where Σ_{XY} is the covariance between X and Y, Σ_Y is the variance of Y, μ_X is the mean of X and μ_Y is the mean of Y.

Given all the previous measurements from $y(0)$ to $y(k)$, recorded as y_0^k, this optimal posterior state can be formulated as

$$\begin{aligned} \hat{x}_{k|k} &= E(x|y_0^k) \\ &= E(x|y_0^{k-1} \mid y(k)|y_0^{k-1}) && (2.24) \end{aligned}$$

Since $(w(k), v(k))$ are Gaussian processes and $v(k)$ is independent of y_0^{k-1}, $P(x(k)|y_0^{k-1}, y(k)|y_0^{k-1})$ can be verified to be a jointly Gaussian distribution and expressed as

$$P\left(\begin{array}{cc} x(k)|y_0^{k-1} & y(k)|y_0^{k-1} \end{array} \right) = N(\left(\begin{array}{c} \hat{x}_{k|k-1} \\ H\hat{x}_{k|k-1} \end{array} \right), \left(\begin{array}{cc} M(k) & M(k)H^T \\ HM(k) & HM(k)H^T + R_v \end{array} \right))$$

where $N(a, b)$ is a Gaussian distribution with mean a and covariance matrix b.

According to equations 2.22 and 2.24, the optimal posterior state and its error covariance matrix are

$$\begin{aligned} x_{k|k} &= \hat{x}_{k|k-1} + M(k)H^T(HM(k)H^T + R_v)^{-1}(y - H\hat{x}_{k|k-1}) \\ &= \hat{x}_{k|k-1} + L(k)(y - H\hat{x}_{k|k-1}) \qquad (2.25) \\ \Sigma_{x_{k|k}} &= M(k) - M(k)H^T(HM(k)H^T + R_v)^{-1}HM(k) \qquad (2.26) \end{aligned}$$

$L(k)$ is the gain of the estimator, which updates the prediction states with the measurement.

Equations 2.21 and 2.26 form a recursive process for updating the error covariance matrices; while equations 2.19 and 2.25 provide the values of the prior and posterior states.

2.4.2.3 Steady State Gain

Since equations 2.21 and 2.26 have a structure analogous to equation 2.16, the estimator Hamiltonian matrix can be derived by comparing the coefficients as

$$H_e = \left(\begin{array}{cc} \Phi^T + H^T R_v^{-1} H \Phi^{-1} \Gamma_1 R_w \Gamma_1^T & -H^T R_v^{-1} H \Phi^{-1} \\ -\Phi^{-1}\Gamma_1 R_w \Gamma_1^T & \Phi^{-1} \end{array} \right) \qquad (2.27)$$

Using the eigen-decomposition algorithm described in optimal gain, the steady state Kalman filter gain is

$$L_\infty = M_\infty H^T (H M_\infty H^T + R_v)^{-1} \tag{2.28}$$

where $M_\infty = \Lambda_I X_I^{-1}$ and $(X_I^T, \Lambda_I^T)^T$ are the eigenvectors of H_e associated to the stable eigenvalues.

Note that, the selection of Γ_1 must assure that the estimator tracks all the states, which provides the feedback mechanism to correct these states from measurements. The states without feedback will drift away from reality because of the imperfect model of the actuator or disturbances in use.

2.5 Feed-Forward Control

Feed-forward control is commonly used if the state-space model for the dynamics of the motion are known or can be estimated. The two most common forms are disturbance rejecter and reference follower. The former is to better reject colored noise and the latter is to better drive the actuator according to a pre-determined moving trajectory. By incorporating the dynamics, the control input then contains the components capable of neutralizing the dynamics induced from disturbances and reference commands.

2.5.1 Disturbance Rejecter

Assuming that the actuator has a linear model as equation 2.18, the colored noise model of the actuator and sensor can be formulated as follows:

$$\begin{pmatrix} \rho_d(k+1) \\ \rho_s(k+1) \end{pmatrix} = \begin{pmatrix} \Phi_d & 0 \\ 0 & \Phi_s \end{pmatrix} \begin{pmatrix} \rho_d(k) \\ \rho_s(k) \end{pmatrix}$$

$$\begin{pmatrix} w(k) \\ v(k) \end{pmatrix} = \begin{pmatrix} H_d & 0 \\ 0 & H_s \end{pmatrix} \begin{pmatrix} \rho_d(k) \\ \rho_s(k) \end{pmatrix} + \begin{pmatrix} w_r(k) \\ v_r(k) \end{pmatrix}$$

where (ρ_d, ρ_s) are the states, (Φ_d, Φ_s) are the state transition matrices, (H_d, H_s) are the observation matrices of the noise, and $(w_r(k),\ v_r(k))$ are the residual noise in the actuator and sensor, respectively,

The augmented model can then be represented as

$$\begin{aligned} x_d(k+1) &= \Phi_w x_d(k) + \Gamma_w u(k) \\ y(k) &= H_w x_d(k) \end{aligned} \tag{2.29}$$

where $x_d(k) = \begin{pmatrix} x(k)^T & \rho_d(k)^T & \rho_s(k)^T \end{pmatrix}^T$, $\Phi_w = \begin{pmatrix} \Phi & \Gamma H_d & 0 \\ 0 & \Phi_d & 0 \\ 0 & 0 & \Phi_s \end{pmatrix}$,

$H_w = \begin{pmatrix} H & 0 & H_s \end{pmatrix}$, and $\Gamma_w = \begin{pmatrix} \Gamma^T & 0 & 0 \end{pmatrix}^T$.

The augmented model is only applied to the optimal estimator gain computation because the focus of the feedforward control is to estimate the noise states and incorporate those into the control input. Since the noise states do not alter the control input u, the optimal control gain is identical to the one computed from the

original state-space model and the control input is

$$u = -Kx(k) - H_d\rho_d(k) + \bar{N}r(k)$$

where $r(k)$ is the reference signal, $\bar{N} = (K(I_{m\times m}-\Phi)^{-1}\Gamma+I_{n\times n})(H(I_{m\times m}-\Phi)^{-1}\Gamma)^{-1}$ is the steady-state gain, and (m, n) are the number of the actuator's states and observations.

An estimator, including model prediction and measurement refinement, are applied to retrieve the true states, which can be formulated as

$$\begin{aligned} \hat{x}_d(k+1) &= \Phi_w\hat{x}_d(k) + \Gamma_w u(k) + L(y(k) - H_w\hat{x}_d(k)) + \Gamma_w\bar{N}r(k) \\ &= (\Phi_w - LH_w - \Gamma_w K_a)\hat{x}_d(k) + Ly(k) + \Gamma_w\bar{N}r(k) \end{aligned} \tag{2.30}$$

where $\hat{x}_d(k)$ represents the estimation of $\begin{pmatrix} x(k)^T & \rho_d(k)^T & \rho_s(k)^T \end{pmatrix}^T$ and $K_a = \begin{pmatrix} K & H_d & 0 \end{pmatrix}$.

By combining equations 2.29 and 2.30, the overall model can be expressed as

$$\begin{aligned} \begin{pmatrix} x(k+1) \\ \hat{x}_d(k+1) \end{pmatrix} &= \begin{pmatrix} \Phi & -\Gamma K_a \\ LH & \Phi_w - LH_w - \Gamma_w K_a \end{pmatrix} \begin{pmatrix} x(k) \\ \hat{x}_d(k) \end{pmatrix} + \Gamma_a w_r(k) + L_a v_r(k) + \bar{N}_a r(k) \\ e(k) &= \begin{pmatrix} H & -H_w \end{pmatrix} \begin{pmatrix} x(k) \\ \hat{x}_d(k) \end{pmatrix} \end{aligned} \tag{2.31}$$

where $\Gamma_a = \begin{pmatrix} \Gamma^T & 0 \end{pmatrix}^T$, $L_a = \begin{pmatrix} 0 & L^T \end{pmatrix}^T$, and $\bar{N}_a = \begin{pmatrix} \Gamma^T & \Gamma_a^T \end{pmatrix}^T \bar{N}$.

The frequency responses of the residue with respect to the disturbance at the

actuator and the reference signal can be expressed as

$$\frac{e(z)}{w_r(z)} = \begin{pmatrix} H & -H_w \end{pmatrix} (zI - \begin{pmatrix} \Phi & -\Gamma K_a \\ LH & \Phi_w - LH_w - \Gamma_w K_a \end{pmatrix})^{-1} \Gamma_a$$

$$\frac{e(z)}{r(z)} = \begin{pmatrix} H & -H_w \end{pmatrix} (zI - \begin{pmatrix} \Phi & -\Gamma K_a \\ LH & \Phi_w - LH_w - \Gamma_w K_a \end{pmatrix})^{-1} \bar{N}_a$$

A detailed model plotted in Simulink is shown in figure 2.4.

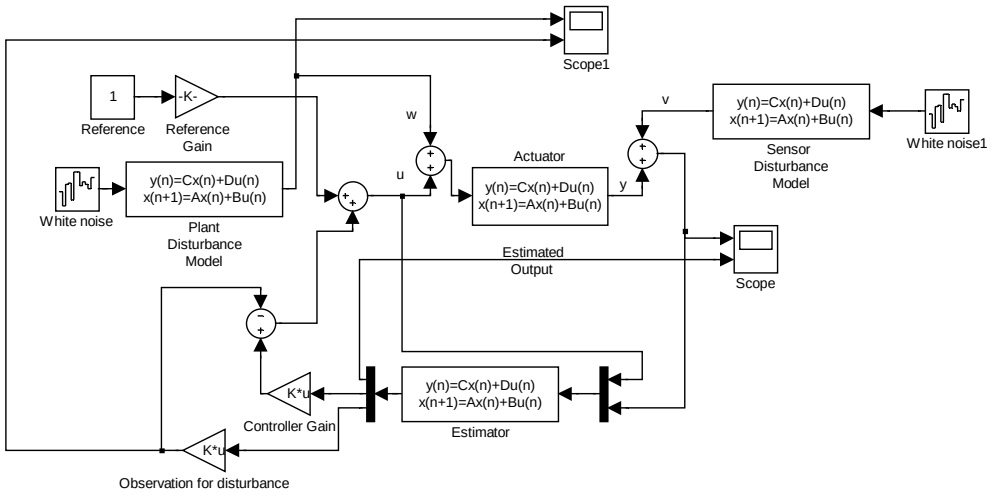

Figure 2.4: Block diagram of a disturbance rejecter

2.5.2 Reference Follower

Generally, a linear controller drives the output to zero when the system is stabilized. In order to follow a reference signal, a set of new state parameters are defined whose output represents the difference between the measurement and the reference signal. Thus, the change of the reference signal can be regarded as the disturbance in this new system model.

Let the dynamics of the reference signal model be

$$\begin{aligned} \rho(k+1) &= \Phi_r \rho(k) \\ r(k) &= H_r \rho(k) \end{aligned}$$

By augmenting the reference model, the model can be written as

$$\begin{aligned} x_r(k+1) &= \Phi_{wr} x_r(k) + \Gamma_{wr} u(k) \\ e(k) &= H_{wr} x_r(k) \end{aligned} \tag{2.32}$$

where $x_r(k) = \begin{pmatrix} x(k)^T & \rho(k)^T \end{pmatrix}^T$, $\Phi_{wr} = \begin{pmatrix} \Phi & \Gamma H_r \\ 0 & 1 \end{pmatrix}$, $H_{wr} = \begin{pmatrix} H & 0 \end{pmatrix}$, and $\Gamma_{wr} = \begin{pmatrix} \Gamma^T & 0 \end{pmatrix}^T$.

This form is analogous to the state space model for the disturbance rejector. Therefore, the estimator design is similar to equation 2.30. The difference is that the current estimator uses errors between the reference and output signals to update the state.

$$\hat{x}_r(k+1) = (\Phi_{wr} - L H_{wr}) \hat{x}_r(k) + \Gamma_{wr} u(k) + L(y(k) - r(k)) \tag{2.33}$$

where $\hat{x}_r(k)$ represents the estimations of $\begin{pmatrix} x(k)^T & \rho(k)^T \end{pmatrix}^T$.

Similar to the disturbance rejecter, the control signal is the summation of both the actuator and augmented states, formulated as $u(k) = -K_{ra} \hat{x}_r(k)$, where $K_{ra} = \begin{pmatrix} K & H_r \end{pmatrix}^T$. Therefore, the overall system model, including the actuator

and the estimator, can be written as

$$\begin{pmatrix} x(k+1) \\ \hat{x}_r(k+1) \end{pmatrix} = \begin{pmatrix} \Phi & -\Gamma_{wr}K_{ra} \\ LH & \Phi_{wr} - LH_{wr} - \Gamma_{wr}K_{ra} \end{pmatrix} \begin{pmatrix} x(k) \\ \hat{x}_r(k) \end{pmatrix} - L_{ra}r(k) + \Gamma_{ra}w(k)$$

$$\hat{e}(k) = \begin{pmatrix} 0 & H_{wr} \end{pmatrix} \begin{pmatrix} x(k) \\ \hat{x}_r(k) \end{pmatrix} \tag{2.34}$$

where $L_{ra} = \begin{pmatrix} 0 & L^T \end{pmatrix}^T$, $\Gamma_{ra} = \begin{pmatrix} \Gamma^T & 0 \end{pmatrix}^T$, and $w(k)$ is the disturbance in the actuator.

The frequency responses of the residue with respect to the disturbance at the plant and the reference signal are

$$\frac{\hat{e}(z)}{w(z)} = \begin{pmatrix} 0 & H_{wr} \end{pmatrix} (zI - \begin{pmatrix} \Phi & -\Gamma K_{ra} \\ LH & \Phi_{wr} - LH_{wr} - \Gamma_{wr}K_{ra} \end{pmatrix})^{-1}\Gamma_{ra}$$

$$\frac{\hat{e}(z)}{r(z)} = -\begin{pmatrix} 0 & H_{wr} \end{pmatrix} (zI - \begin{pmatrix} \Phi & -\Gamma K_{ra} \\ LH & \Phi_{wr} - LH_{wr} - \Gamma_{wr}K_{ra} \end{pmatrix})^{-1}L_{ra}$$

Figure 2.5 shows the design diagram using Simulink.

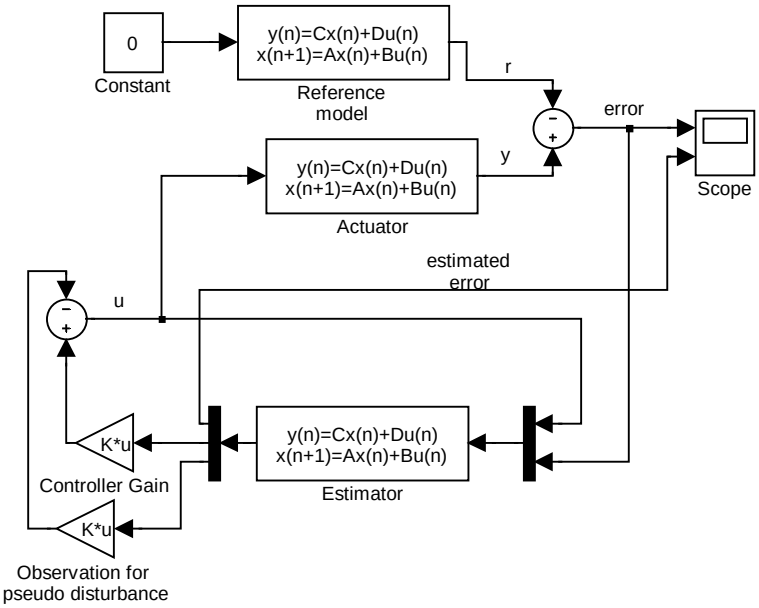

Figure 2.5: Block diagram of a voltage follower

Chapter 3

Design and Analysis of Fine Angular Pointing, Acquisition, and Tracking Systems

Research in fine angular pointing and tracking (FPAT) systems has blossomed since the 1980s after the appearance of satellite communication. Although several companies have brought commercial systems for long-haul point-to-point links to the market since the 1990s, such as Canon, Terabeam, MRV, fSONA and LightPointe, not until recent years have the new generation of FSO transceivers been equipped with a FPAT system, such as the DS series from Canon and Compass (compact airborne stabilization system) from NovaSol. Currently, many research groups are still working on improving such systems to have: (1) a faster tracking rate, (2) a higher accuracy, (3) a larger tracking range, and (4) a low cost.

Gibson [21] categorized fine laser beam steering systems into mechanical and non-mechanical. Their pros and cons are discussed in the previous chapter. Mechanical systems have been the favorite in this field and several FPAT systems were developed by different research groups, including fast-steering mirrors (FSM) based systems by UCLA [5], the Canada Space Agency [39], and NovaSol, and Risley prism based systems by Ball Aerospace and Technologies Corp [56]. As for the non-mechanical systems, Nikulin et.al [50] implemented a beam steerer with acoustic-optical Bragg cells. Arnon et.al [53] proposed using gratings to steer the

beam. McManamon [47] demonstrated a beam steering device with a liquid crystal.

A FPAT system is usually integrated with traditional FSO transceivers to achieve reliable connectivity. Its objective is to (1) complete the alignment to establish a link, (2) reject the disturbance induced by unknown noise sources, and (3) track the angular motion from the transceiver pair. The FPAT system is normally triggered after two FSO transceivers have been roughly pointed with a CPAT system.

An FPAT system consists of three subsystems, which are

1. Beam Steerer: points the beam precisely according to a set of computed azimuthal and elevation angles provided by the AR.

2. Angular Resolver (AR): resolves the incident angle of the incoming beam and optimizes the received power. The AR in the FPAT system mainly consists of a small field-of-view imaging system (reflective or refractive), a PSD (QD or LEP), and a two-axis actuator.

3. Internal Controller: provides a feedback control signal, which rejects the disturbance resulting from unknown noise sources; and a feedforward control signal, which cancels the angular change induced by the jitter or follows a pre-determined scanning pattern.

3.1 A PAT System using Focal Plane Motion

The linear relationship between the incident angle and the projected image position on the focal plane is revealed in section 2.1.1. According to the reciprocal of

the optical rays, once an optical source is placed at any image spot on the focal plane, the transmitted ray will produce the outgoing angle analogous to the corresponding incident angle. Moving the optical source on the focal plane is the beam steering mechanism used in the proposed FPAT system, also known as the focal plane motion (FPM) based system.

Some questions wait to be answered before implementing the system. For example, if the moving plane is the focal plane, the transmitted beam is too collimated to be used in real FSO applications. How can we retrieve a desired beam divergence without affecting the beam steering capability? Some preliminary simulations and experiments for this new FPAT system using a focal plane motion are reported to demonstrate its capability. Compared to the system based on the fast-steering mirrors (FSM), the proposed system is very low-cost, which is an important incentive for popularizing FSO products.

The major difference between the FPM-based and the FSM-based system is the mechanism to determine the incident angle. An FPM-based system determines the incident angle from the position of the focused spot using the paraxial approximation; whereas an FSM-based system physically rotates the two-axis mirror until the beam re-focused at a pre-determined position and then the incident angle is measured by the sensor of the FSM after the mirror is stabilized. To summarize, the FPM-based system is capable of providing faster angular information to the beam steerer with slight errors, and the FSM-based system provides accurate angular information but at a slower rate.

It occurs to us that an FPM-based system can outperform the FSM-based

system if the aberrations can be minimized to a level that does not affect the pointing accuracy of the beam steerer, since a correct incident angle can lead the beam steerer to point at the target much faster. A detailed investigation of the aberrations in the optical system is listed below.

Besides affecting the pointing accuracy, aberrations also result in the photodiode being slightly displaced from the optimal receiving position, which may degrade the link quality. This problem can be easily solved by a non-imaging system, such as an compound parabolic concentrator (CPC). A study is described in the following section to show that a CPC can expand the effective receiving area of the photodiode, which effectively mitigates the position offset from the aberrations.

Figure 3.1 shows a diagram of the FPM-based FPAT system and figure 3.2 describes its function flow chart. Notice that the current implementation applies two PSDs to measure the incoming angle of the other transceiver and the motions of the local beam steerer. The PSD used at the beam steerer can be freely replaced by other compact sensors, such as strain gauge sensors or linear variable displacement transducers (LVDT).

Each part of the FPM-based system has been either simulated or implemented for further verification. This is mainly divided into three sections: (1) beam steerer, (2) angular resolver, and (3) controller design.

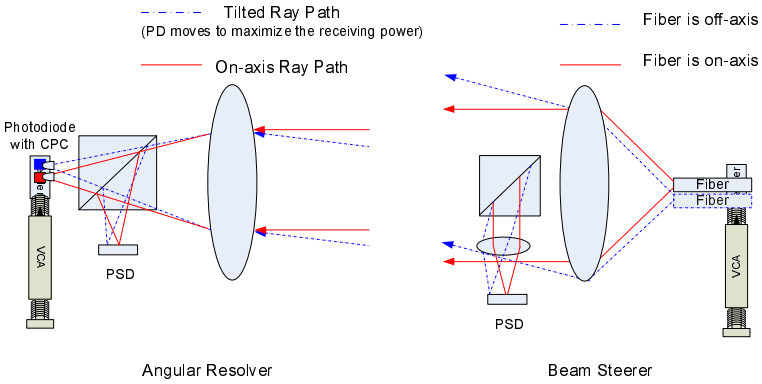

Figure 3.1: Proposed fine angular PAT system diagram: angular resolver is on the left and beam steerer is on the right. The PSD in the beam steerer measures the position of the fiber, which can be replaced by other compact sensors.

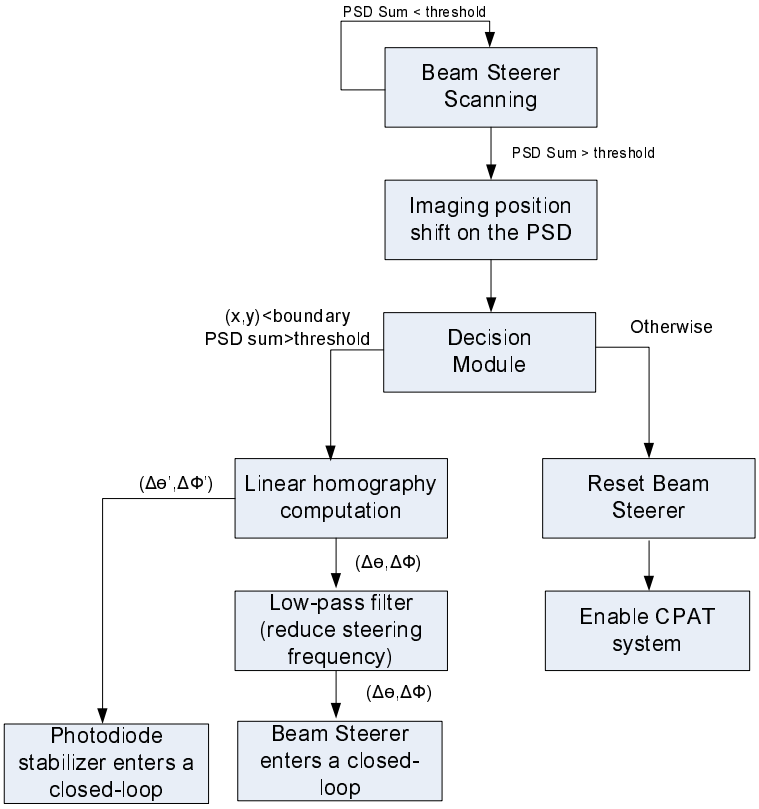

Figure 3.2: Flow chart for the proposed FPAT system

3.2 Beam Steerer

The current moving mechanism is provided by a two-axis moving stage, which moves a fiber tip near the focal plane of a collimator for beam steering. The objective of the beam steerer is to: (1) move the laser to the pre-determined position quickly and accurately, and (2) provide an accurate outgoing angle and a sufficiently large beam divergence for the transmitted beam after passing the optical system. The former depends on the actuators and the control loop and the latter depends on the optical system used for the collimation. Figure 3.3 shows a beam steerer in the laboratory.

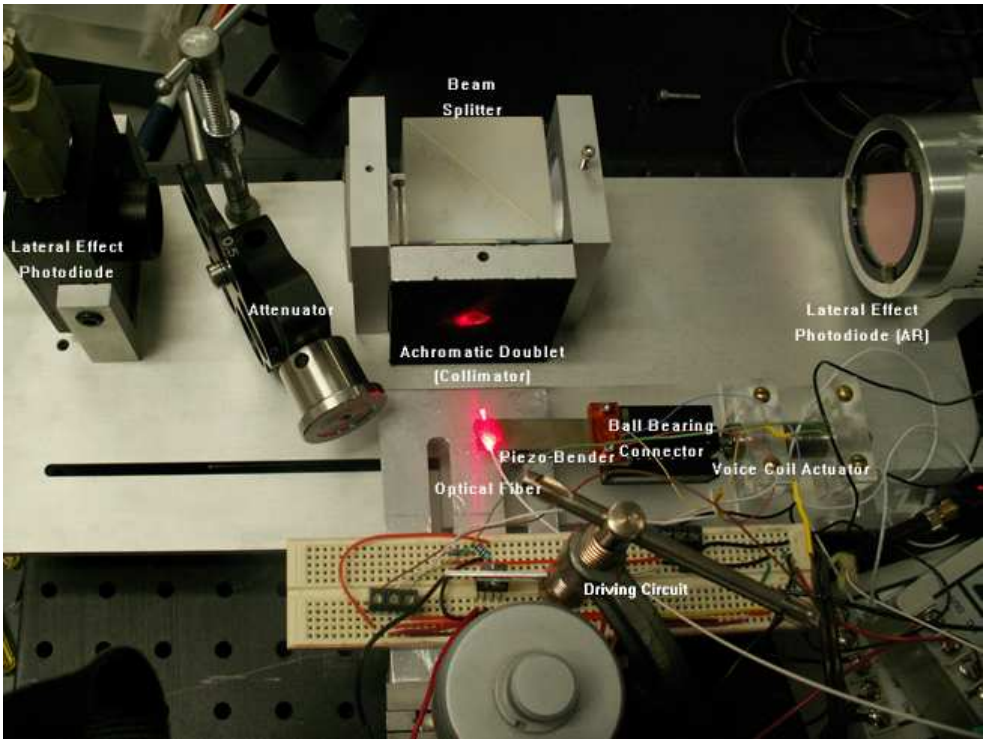

Figure 3.3: A beam steerer consisting of a two-axis moving platform, an achromatic lens, and a fiber.

The current specifications of the beam steerer are selected as: a beam divergence of 2 mrad, a steering range of ±25 mrad (±1.5°), and a steering resolution of 1 mrad.

3.2.1 Optical Systems

Given a steering range of ±25 mrad and a ±1 mm actuator, the collimator optical system must have a focal length of 40 mm according to the imaging equation 2.5 derived from the paraxial approximation and offer a good image quality for beam steering. A suitable lens is the achromatic lens (026-0190) from Optosigma. Combined with a single mode fiber (Thorlabs SM600) whose field diameter is 4.5 μm at 660 nm as well as a two-axis moving platform, a beam steerer is produced. This section evaluates if the optical system provides the desired beam divergences as well as accurate steering angles.

3.2.1.1 Beam Divergence

The guided field in a step-index monomode fiber is in the LP_{01} mode, whose field distribution in the core and cladding can be modelled by Bessel and modified Bessel functions, respectively [13]. For most monomode fibers, the guided wave at the fiber tip can be accurately approximated by a Gaussian beam with an infinite radius of curvature and a waist size of [9]

$$W = (0.65 + 1.619V^{-1.5} + 2.879V^{-6})a$$

where a is the fiber radius, V is the normalized frequency, and $2W$ is normally defined as the mode fiber diameter (MFD).

The propagation of the Gaussian beam through an optical paraxial ABCD systems is usually characterized by

$$q(z) = \frac{Aq(0) + B}{Cq(0) + D}$$

where $q(z) = \frac{1}{R(z)} - \frac{i\lambda}{\pi w(z)^2}$ is the beam parameter at distance z, $R(z)$ is the radius of curvature, $w(z)$ is the beam waist, and λ is the wavelength.

If the fiber is placed at a distance d_1 in front of a collimator with a ABCD ray transfer matrix, the new minimum beam waist can be cast as

$$w_o = \frac{w_f}{\sqrt{(D + Cd_1)^2 - C^2 q_f^2}}$$

where $q_f = \frac{i\lambda}{\pi w_f}$ is the beam parameter at the fiber tip and w_f is half of the mode fiber diameter.

Generally, q_f is negligible and the new minimum beam waist can therefore be simplified as $w_o = \frac{w_f}{D+Cd_1}$. Since the beam divergence and minimum beam waist are related by $\theta = \frac{\lambda}{\pi w_o}$, a linear relation between the beam divergence and d_1 can be derived as

$$\theta_o = \frac{\theta_f}{D}(1 + \frac{Cd_1}{D}) \tag{3.1}$$

where θ_f is the beam divergence at the fiber output.

Because C is the inverse of the focal length of the collimator, equation 3.1 reveals two important pieces of information for the beam steerer design:

1. The beam divergence resulting from a collimator increases linearly with the

distance from the fiber to the front focal plane for an on-axis gaussian beam.

2. Choosing a collimator with a long focal length can reduce the variation of the beam divergence induced by the motion.

Beam Divergence Simulation

This simulation was conducted to answer two questions associated with the current beam steerer consisting of an achromatic lens (Optosigma 26-0190) with 40 mm focal length and a single mode fiber (Thorlabs SM600) with a mode field diameter of 4.5 μm at 660 nm: (1) where should the fiber be placed to produce the beam divergence according to the specification? and (2) how does the beam divergence vary as the fiber moves?

Since the fiber tip may be off-axis when applying the focal plane motion, instead of taking the on-axis approximation as shown in the previous derivation, a simulation was conducted using the CodeV optical simulator from Optical Research Associates, where the beam divergence of an off-axis Gaussian beam was evaluated by a general algorithm proposed by Arnaud et.al [6].

The answer to the first question can be found by placing the fiber on the optical axis within a ± 1 mm span near the front focal plane, 39 mm in front of the first element of the collimator. The results, plotted in figure 3.4, show a linear relation between the fiber's displacement and the beam divergence, as predicted in the derivation. From the plot, it can be seen that the best location to place the fiber is between 38.6 to 39.4 mm in order to achieve the 2 mrad specification.

Since the beam divergence varies linearly with the distance from the fiber to

the front focal plane, the resulting beam divergences suffer from smaller impacts from the fiber motion if the moving plane has a distance farther away from the front focal plane. The current selection is to place the fiber at $z = 38.6$ mm, 0.4 mm away from the front focal plane.

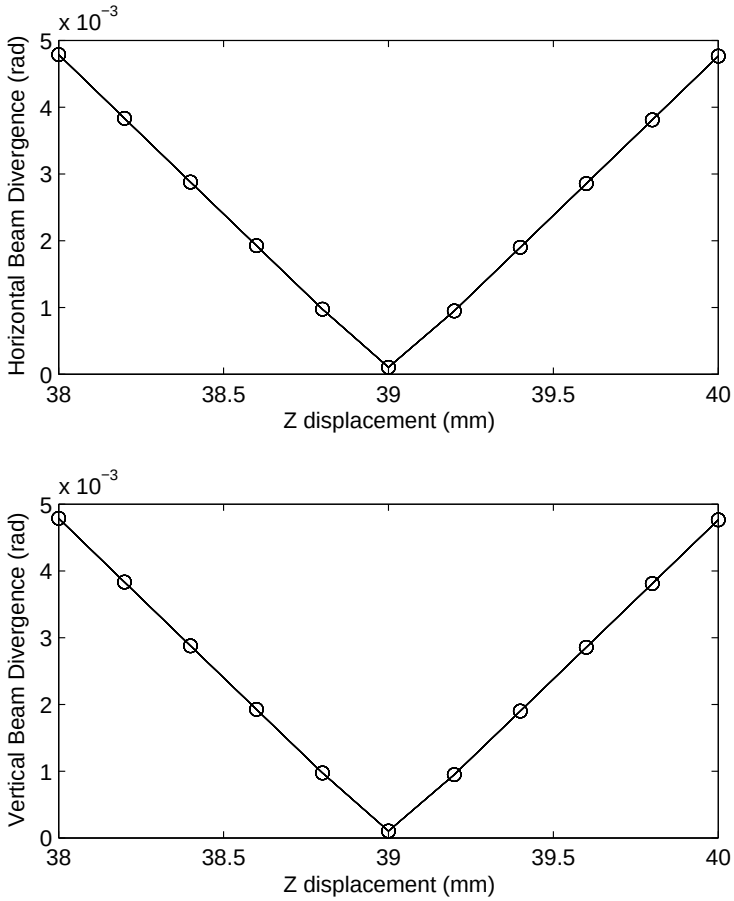

Figure 3.4: Horizontal and vertical beam divergences on a ± 1 mm span on both Y and Z axes at $x = 0$ mm.

The variation of the beam divergence was evaluated by moving the fiber within a ± 1 mm span on the XY plane at $z = 38.6$ mm and the results as shown in figure 3.5. The results indicate that the beam divergence of the fiber is 20% smaller at the

corner than at the centroid, which must be taken into consideration in the FPAT system design.

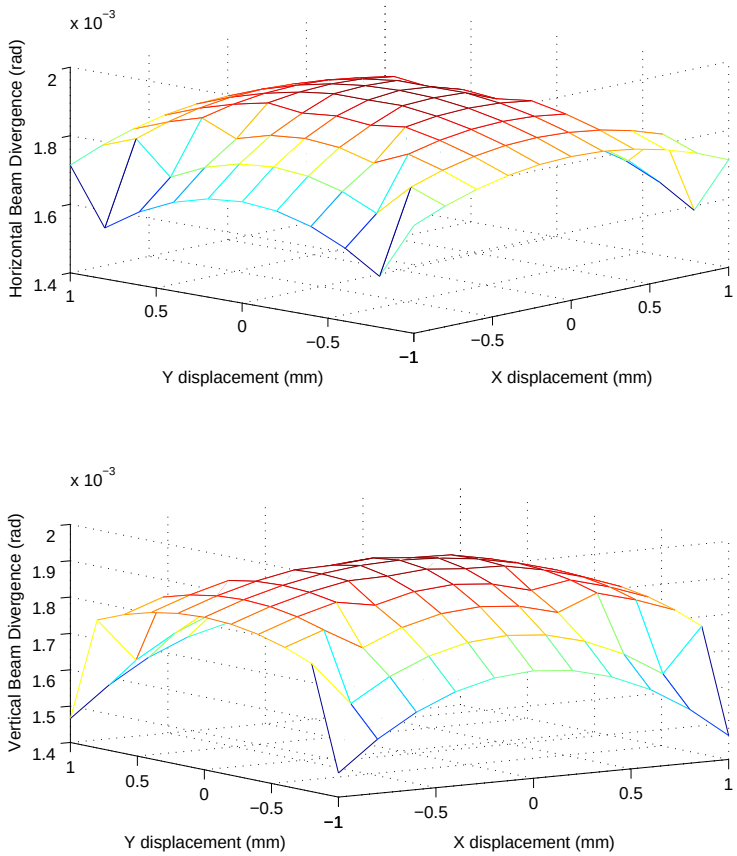

Figure 3.5: The variation of the horizontal and vertical beam divergences from a planar motion at $z = 38.6$ mm

3.2.1.2 Steering Angle

According to the paraxial approximation in section 2.1.1 and the reciprocal of the optical rays, the relation between the location of the fiber (x_f, y_f) on the focal plane and the steering angle (α, β) in azimuthal and elevation can be described by

a scaling matrix M as

$$\begin{pmatrix} \alpha \\ \\ \beta \end{pmatrix} = M \begin{pmatrix} x_f \\ \\ y_f \end{pmatrix} \tag{3.2}$$

This simulation investigated the validity of this relation when the fiber motion is not on the focal plane. The experimental setup was identical to the previous simulation and the fiber was placed at $z = 38.6$ mm and moved on a ± 1 mm span. Real ray tracing techniques were applied and the angles of the chief ray at different fiber locations were recorded and further compared with the computed angles according to equation 3.2.

The residual angular errors, defined as the difference between the simulated and computed beam divergence using equation 3.2, are shown in figure 3.6. The maximum residual error is $(0.575, 0.575)$ μrads and the average residual error is $(0.144, 0.144)$ μrads in both azimuthal and elevation directions. The results imply that the steering angle is linearly associated with the position of the fiber, even though it was not on the focal plane.

3.2.2 Actuators

Two different actuators, a piezo bender (PI-PL140) and a voice coil actuator (H2W-NCM02), were selected for building the two-axis moving platform because of the unbalanced load from each axis (one actuator carries only a fiber and the other actuator carries the load of the fiber and the first actuator) and their displacement ranges are $\pm$1mm and $\pm$1.9 mm, respectively.

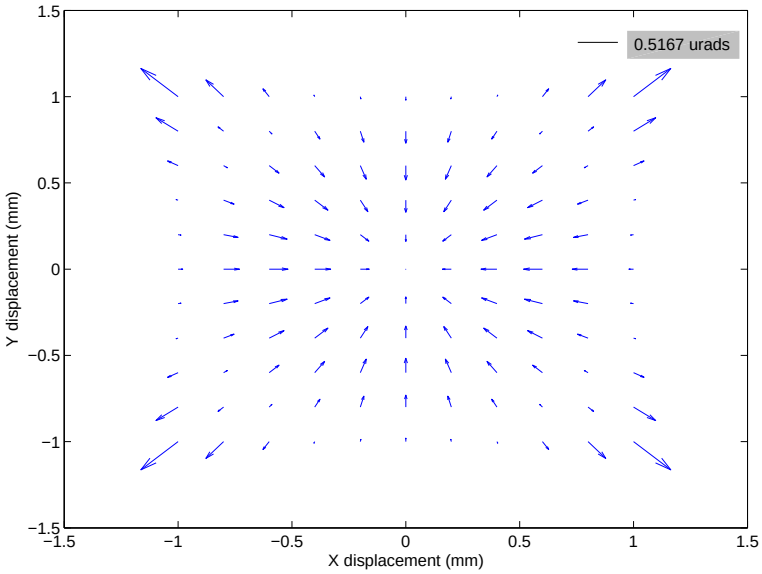

Figure 3.6: Residual angular errors by using the paraxial approximation model at different fiber positions. The length of the arrow represents the magnitude of the angular error.

3.2.2.1 Piezo Actuators

The piezoelectric effect was first observed by Jacques and Pierre Curie in 1880 [29]. They discovered that pressure applied to a quartz crystal can create an electrical charge in the crystal. In addition, they also verified that an electrical field applied to the crystal would lead to a strain of the material, defined as the geometrical expression of deformation caused by the action of stress on a physical body. It was also known as converse piezoelectricity. In general, the deformation relation can be characterized as [61]

$$D_i = d_{iJ}T_J + \epsilon_{ij}^T E_j \tag{3.3}$$

$$S_I = s_{IJ}^E T_J + d_{jI}E_j \tag{3.4}$$

where D represents the electric displacement, T represents the stress, S represents the strain, s^E represents the elastic compliance when subjected to a constant electric field, d represents the strain coefficient (m/V), ϵ^T represents permittivity measured at a constant stress, and E represents electric field. I, J : 1 to 6 and i, j : 1 to 3. The number $1 \rightarrow 6$ represent six axes. $1 \rightarrow 3$ are X, Y, Z axes and $4 \rightarrow 6$ are rotation axes with respect to X, Y, Z.

A piezo-bender, also known as a piezo-bimorph, consists of two attached beams of piezo material in d_{31} mode. The d_{31} mode piezo-material contracts or expands in the X-axis when the electric field is applied in the polarization axis Z if its polarization direction is in the Z axis. Combining two d_{31} piezo beams in the opposite or same polarization directions results in two types of piezo-benders, serial or parallel and both types have similar dynamic models.

Let us assume a serial piezo bender consisting of two symmetric d_{31} beams with opposite polarization directions and each beam has a thickness h, as shown in figure 3.7. The resulting strain and electric displacement can be modelled as

$$\begin{aligned} D_3^U &= -d_{31}T_1^U - \epsilon_{33}^T E_3 \\ S_1^U &= s_{11}^E T_1^U - d_{31}E_3 \\ D_3^L &= d_{31}T_1^L + \epsilon_{33}^T E_3 \\ S_1^L &= s_{11}^E T_1^L + d_{31}E_3 \end{aligned} \tag{3.5}$$

where (D^U, D^L) are the displacements, (T^U, T^L) are the stresses, (S^U, S^L) are the strains, and $(-E_3, E_3)$ are the applied electrical fields in the upper and lower beams.

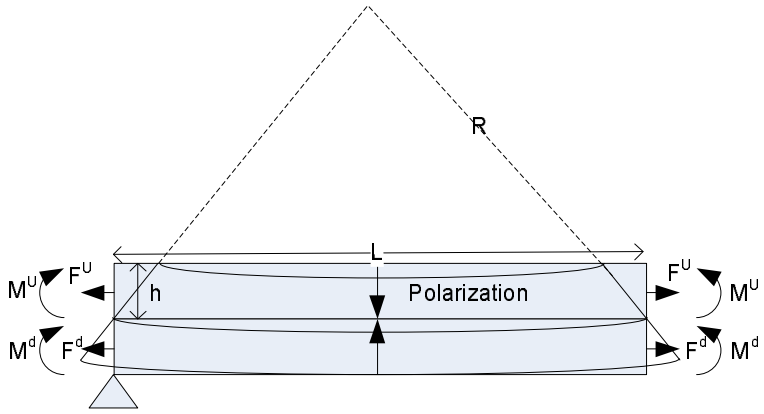

Figure 3.7: A serial piezo bender structure

The stress can be expressed as the combination of the force and bending moment as

$$T = \frac{F}{wh} + \frac{hM}{2I}$$

where (w, h) is the width and thickness of a beam, (F, M) is the force and bending moment, and $I = \frac{wh^3}{12}$ is the moment of inertia of the rectangular cross-section.

Since the two layers are identical and attached adhesively, four equilibria can be retrieved as

$$\begin{aligned}
\frac{F^U s_{11}^E}{wh} + \frac{hM^U s_{11}^E}{2I} - d_{31}E3 &= \frac{F^L s_{11}^E}{wh} - \frac{hM^L s_{11}^E}{2I} + d_{31}E3 \text{ ; strain equilibrium} \\
M^U + M^L &= \frac{h}{2}(F^U - F^L) \text{ ; moment equilibrium} \\
F^U &= -F^L \text{ ; force equilibrium} \\
M^U &= M^L \text{ ; bending curvature equilibrium}
\end{aligned} \tag{3.6}$$

And the force and bending moment can be solved as:

$$\begin{aligned}
F^U &= -F^L = \frac{d_{31}E_3 wh}{4s_{11}^E} \\
M^U &= M^L = \frac{d_{31}E_3 wh^2}{8s_{11}^E}
\end{aligned}$$

The strains in the upper and lower beams result from force, bending moment, and the E-field, equal to

$$\begin{aligned}
S^U &= \frac{d_{31}E_3}{4} - \frac{3d_{31}E_3}{2h}(z - \frac{h}{2}) - d_{31}E_3 \\
S^L &= -\frac{d_{31}E_3}{4} - \frac{3d_{31}E_3}{2h}(z - \frac{h}{2}) + d_{31}E_3
\end{aligned}$$

Plugging the last two equations back to equation 3.5, the stress T is

$$T^U = (\frac{d_{31}E_3}{4} - \frac{3d_{31}E_3}{2h}(z - \frac{h}{2}))/s_{11}^E \tag{3.7}$$

$$T^L = (-\frac{d_{31}E_3}{4} - \frac{3d_{31}E_3}{2h}(z - \frac{h}{2}))/s_{11}^E \tag{3.8}$$

Since the energy density of an infinitesimally small element is $u = \frac{1}{2}(ST + DE)$, plugging the stress and strain into this equation and integrating over the three axes, the total energy of the two layer bender with no external force is expressed as

$$U = LwhE_3^2(\epsilon_{33}^T - \frac{d_{31}^2}{4s_{11}^E}) \tag{3.9}$$

If an additional force F_1 is applied to the bender at distance x away from the free end, an additional moment is created, which results in an additional stress of

$$T_1 = \frac{3F_1(L-x)z}{2wh^3}$$

Taking the additional stress to compute the energy density and integrating over three axes, the total energy is equal to

$$U = \frac{s_{11}^E F_1^2 L^3}{4wh^3} + \frac{3F_1 E_3 d_{31} L^2}{4h} + LwhE_3^2(\epsilon_{33}^T - \frac{d_{31}^2}{4s_{11}^E}) \tag{3.10}$$

Since $E_3 = \frac{-V}{2h}$, according to the Castigliano's second theorem [28], the displacement (δ) is obtained by taking derivatives of the total energy with respect to the related parameters as

$$\delta = \frac{s_{11}^E L^3}{2wh^3} F_1 - \frac{3d_{31}L^2}{8h^2} V$$

If the additional force working toward the piezo-bender is applied by an external force (F_{ext}), the dynamics of the piezo-bender can be expressed by Newton's law as

$$F_{ext} = F_1 + m\ddot{\delta} + b\dot{\delta}$$

Thus, the bender model can be formulated as a second-order system

$$m\ddot{\delta} + b\dot{\delta} + \frac{2wh^3}{3s_{11}^E L^3}\delta = F_{ext} - \frac{d_{31}wh}{2s_{11}^E L} V \tag{3.11}$$

The previous equation does not take hysteresis into consideration. Several methods are proposed to model this nonlinear behavior, which are:

1. Boun-Wen hysteresis model [76]: This model describes the hysteresis as a differential equation, where

$$\dot{z} = \alpha\dot{x} - \beta|\dot{x}|z|z|^{n-1} - \gamma\dot{x}|z|^n$$

where z represents the state variable, x represents the excitation, α controls the input force amplitude, β, γ control the shape of the hysteresis loop, and n represents the smoothness of the transition from the elastic to plastic response. The advantage of the Boun-Wen model is that the state variable z can be regarded as an additional strain and incorporated into equation 3.11 [42].

2. Preisach model [46]: The Preisach model is based on two important observations of the hysteretic actuator.

   (a) The wipe-out property describes that the output from a hysteretic actuator is determined not only by the current input but also by the alternating series of dominant input extrema.

   (b) The congruency property suggests that all minor hysteresis loops corresponding to back-and-forth variation of inputs between the same two consecutive extremum values are congruent.

   The Preisach model contains a set of known Preisach operators and an unknown Preisach function. The model is

$$f(t) = \int \int_{\alpha>\beta} \mu(\alpha, \beta) \hat{\gamma_{\alpha\beta}} d\alpha d\beta \tag{3.12}$$

   where $\hat{\gamma_{\alpha\beta}}$, known as the hysteresis operators, can be interpreted as a two-position unit relay operator with an on switch at α and an off switch at β and $\mu(\alpha, \beta)$, known as the Preisach function, includes the parameters of the Preisach model.

3. Prandtl-Ishlinskii (PI) Model [40]: The PI model behaves similarly to the Preisach model. Instead of using Preisach operators, a set of PI operators, similar to regular backlash operators, are applied to describe the hysteresis loop. The PI model has an analytical solution for the hysteresis compensator compared to the numerical solution in the Preisach model.

Incorporating hysteresis into the bender's model promises a more stable control loop but increases the implementation cost as well as the complexity. Normally, the hysteresis affects about $7 \sim 15\%$ of the motion range, which can simply be treated as actuator noise in feedback control.

3.2.2.2 Piezo-Bender Model Identification

The identification of the current piezo-bender (PI-PL140) utilizes a pseudo random binary sequence (PRBS) to probe its open-loop response. The PRBS is a set of binary sequences, whose spectrum is close to white within a pre-determined bandwidth. Since the current piezo-bender has a natural frequency at 160 Hz (unload), the bandwidth of the PRBS is set to 200 Hz.

As shown in the previous section, a two-layer piezo-bender can be described by a linear model, known as the ARX model, which can be written as

$$A(z)y(n) = B(z)u(n - nk) + e(n)$$

where $y(n)$ is the output at time n, $u(n)$ is the input at time n, $e(n)$ is the estimation error at time n, nk is the delay between input and output, and $A(z) = 1 + \sum_{k=1}^{na} a_k z^{-k}$, $B(z) = \sum_{k=1}^{nb} b_k z^{-k}$ are two linear discrete filters whose order are

na and nb, respectively.

Our goal was to estimate all the coefficients from a_1 to a_{na}, b_1 to b_{nb}, and nk to characterize the piezo-bender. Considering the estimation results and the physical model of the piezo-bender derived above, an ARX(3,1,1) model was selected and expressed below

$$M_{PB} = \frac{0.2365z^{-1}}{1 - 1.929z^{-1} + 1.604z^{-2} - 0.4161z^{-3}} \tag{3.13}$$

After identification, a synthesized output was generated by the model coefficients and the input samples and the residue between the real and synthesized output was computed. The residue mainly resulted from imperfect modelling caused by the hysteresis. The standard deviation of the residual signal for the piezo-bender was 0.34 V for a ±3 V output, about 11% of the entire range.

3.2.2.3 Voice Coil Actuators

The voice coil actuator is driven by the Lorentz force resulting from the field current and magnetic field.

The input voltage equals the sum of potential on the resistance, inverse potential coupled from the moving speed, and the potential stored in the inductance,

$$V_{in} = I_f R + K_m^c \dot{x_L} + L\dot{I_f} \tag{3.14}$$

where I_f represents the current in the VCA, L represents the equivalent inductance, and K_m^c represents the coupling coefficient.

The acceleration results from the magnetic force produced by the field current,

which equals

$$F = L \times BI_f = K_m I_f = m_L \ddot{x_L} + \alpha \dot{x_L} + (k_1 + k_2) x_L \tag{3.15}$$

where B : magnetic field, L : the vector of the current path, α : friction coefficient, k_1, k_2 : the spring coefficients, and K_m represents the gain between the current and the induced torque.

$$x_{n+1} = Ax_n + Bu_n \tag{3.16}$$

Combining the two equations, the frequency response between V_{in} and x_L can be written as

$$\frac{x_L}{V_{in}} = \frac{K_m}{s^3 L m_L + s^2(\alpha L + R m_L) + s(K_m K_m^c + R\alpha + L(k_1 + k_2)) + R(k_1 + k_2)}$$

Also, its state-space model can be written as

$$\dot{\mathbf{x}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\frac{R(k_1+k_2)}{Lm_L} & -\frac{K_m K_m^c + R\alpha + L(k_1+k_2)}{Lm_L} & -\frac{Rm_L + L\alpha}{Lm_L} \end{pmatrix} \mathbf{x} + \begin{pmatrix} 0 \\ 0 \\ \frac{K_m}{Lm_L} \end{pmatrix} V_{in} \tag{3.17}$$

$$Y = \begin{pmatrix} M & 0 & 0 \end{pmatrix} \mathbf{x} \tag{3.18}$$

where $\mathbf{x} = \begin{pmatrix} x & \dot{x} & \ddot{x} \end{pmatrix}^T$, M : magnification of the lens set.

If L is relatively small compared to other coefficients, the resulting state-space can be simplified to a second-order model as

$$\dot{\mathbf{x}} = \begin{pmatrix} 0 & 1 \\ -\frac{k_1+k_2}{m_L} & -\frac{R\alpha + K_m K_m^c}{Rm_L} \end{pmatrix} \mathbf{x} + \begin{pmatrix} 0 \\ \frac{K_m}{Rm_L} \end{pmatrix} V_{in} \tag{3.19}$$

where $\mathbf{x} = \begin{pmatrix} x & \dot{x} \end{pmatrix}^T$

3.2.2.4 VCA Model Identification

The H2W-NCM02 VCA has a maximum stroke length of 3.8 mm, a maximum force of 6.6 N, and a continuous force of 2.2 N. The load on the VCA is around 32 grams, including a piezo-bender, a ball bearing stage, and the VCA moving weight. Two identical springs, each with a spring constant of 7.8 lbs/in (1366 N/m), are attached on each side of the shaft to conserve kinetic energy. The natural frequency can be computed as

$$f = \frac{1}{2\pi}\sqrt{\frac{k_1 + k_2}{m_L}} = 46.5$$

Considering the estimated results along with the physical model derived above, an ARX (2,1,2) was selected and expressed as follows:

$$M_{VCA} = \frac{0.1066z^{-2}}{1 - 1.826z^{-1} + 0.9043z^{-2}} \tag{3.20}$$

The identified model has a natural frequency of 45.9 Hz, close to the natural frequency (46.5 Hz) computed from the specifications. The residual noise from the difference between the synthesized signal and the measurement has a standard deviation of 0.22 V for a ±5 V output, about 4% of the entire range, which results from the static friction of the ball bearing.

3.3 Angular Resolver

The angular resolver (AR), consisting of an optical imaging system and a position sensing diode (PSD), is mainly to resolve the incident angle of the beam from the other FSO transceiver in order to optimize the received power. According

to the derivation in section 2.2, the location of the beam transformed by a pre-calibrated linear homography serves as the reference to move the beam steerer.

The objectives of the angular resolver are to: (1) project the incident angle to the image position using the paraxial approximation, (2) digitize and report the image position accurately, and (3) move the photodiode toward the pre-determined position fast and accurately.

The current specification of the AR is the following: an angular resolution of 1 mrad, a half field-of-view of 30 mrad (2°), chosen to be slightly larger than the steering range of the beam steerer, and a receiving aperture greater than the lateral correlation length ($\sim \sqrt{\lambda L}$), mitigating jitter from turbulence, where λ is the wavelength and L is the link length [80].

3.3.1 Optical Systems

To meet described specifications, an achromatic lens from Optosigma (26-1460) is used, which has a 25 mm radius aperture, a focal length of 100 mm, and a 10 mm LEP. The resulting AR has a semi field-of-view of 2.85°, an aperture of 25 mm greater than the correlation length (17.8 mm) when applied to a 500 m link distance. This section is dedicated to evaluating the validity of the paraxial approximation on the current AR, equivalent to studying the aberration errors. As described in section 2.1.2, two different errors are produced by the aberrations: the chief ray displacement, resulting from the distortion, and the centroid shift, caused by other aberrations.

One may ask why we do not evaluate the Seidel aberrations directly. Our explanation is that the magnitude of each of the Seidel aberrations only represents their individual largest displacement from the paraxial position, which does not provide the centroid information. For example, if an optical system suffers from strong aberrations but the resulting spot diagram is close to symmetric, its performance can be superior than another optical system with smaller but asymmetrical aberrations.

In addition, a combined design including a photodiode and a compound parabolic concentrator, used to optimize the received power, is proposed and analyzed, which compensates for a small position shift of the photodiode caused by aberrations.

3.3.1.1 Chief Ray Displacement

This simulation evaluated the magnitude of the distortion from the residual errors, defined as the displacement between the true positions of the chief ray and the image positions computed from the paraxial model. The incident angles move from 0° to 2.8° with a 0.4° spacing in both X and Y axes.

Besides, the additional errors resulting from the defocus of the image plane were also evaluated by placing the image plane at 0, 1 and 5 mm away from the best focus position to estimate the additional residual errors. The results are summarized at Table 3.1.

The results indicate that the maximum residual error is only 0.8 μm on a 10

Sensor Location (mm)	Maximum residual error (um)		Average residual error (um)	
	X	Y	X	Y
0	0.8132	0.8132	0.1805	0.1805
1	0.8179	0.8180	0.1815	0.1815
5	0.8369	0.8371	0.1858	0.1858

Table 3.1: Residual errors for PSD at different distances away from the best focus

mm sensing area. Also, the additional error resulting from the defocus of the image plane is around 0.01 μm and increases with the defocus length. The equivalent angular error was around 8 μrads, which was negligible in this prototype.

3.3.1.2 Image Centroid Shift

In general, the sensor reports the centroid position of the beam instead of the position of the chief ray. As described in section 2.1.2, the Seidel aberrations produce irregular spot diagrams and further result in centroid shift. By tracing 316 rays in the current optical system, the centroid position can be estimated as well as the difference between that and the position of the chief ray. Three angles, 0°, 1.42°, and 2.85° were investigated as well as the additional errors introduced by defocus. The upper part of figure 3.8 shows the displacement between the chief ray and the estimated centroid and the bottom part shows the radius of the spot diagram.

These results indicate that the projection error of the current AR is mainly from the centroid shift instead of the chief ray displacement. The magnitude of the error is around 0.1 mm on the PSD, corresponding to 100 μrad for an incident angle

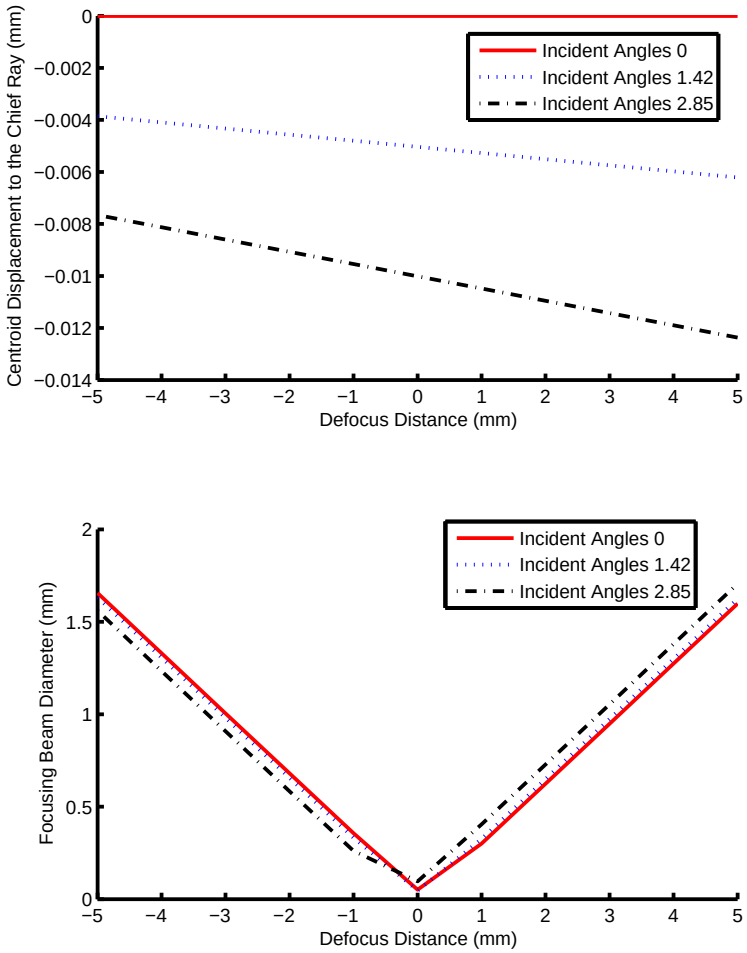

Figure 3.8: The upper plot shows the displacement between the centroid of the beam and the location of the chief ray, and the bottom plot shows the radius of the spot diagram.

of 2.85°. Its magnitude decreases with smaller incident angles.

3.3.1.3 Compound Parabolic Concentrator

As described in the beginning of the chapter, a compound parabolic concentrator (CPC) can be incorporated into the photodiode to solve for the slight offset from the power optimization operation because it expands the effective receiving aperture of the photodiode. In addition, a larger effective receiving aperture can mitigate the angular fluctuations induced by turbulence [66, 73].

A CPC consisting of two off-axis parabolic mirrors has a larger entrance pupil and a smaller exit pupil. Once the incident ray enters the entrance pupil with an angle smaller than the acceptance angle of the CPC, it reflects back to the exit pupil. The detailed geometric description is given by Winston and Welford book [75]. A CPC structure can be uniquely determined by its maximum acceptance angle, inversely proportional to the f-number of the focusing lens, and the receiving aperture, equivalent to the radius of the photodiode in the design.

$$\begin{aligned} \sin\theta_{ma} &= \tfrac{r}{R} \\ L &= (R+r)\cot\theta_{ma} \end{aligned} \tag{3.21}$$

where θ_{ms} represents the maximum acceptance angle, (R, r) represents the aperture of the entrance and exit pupil of the CPC respectively, and L represents the total length of the CPC.

The f-number of the current achromatic lens corresponds to a maximum incident angle of 14°. The exit aperture was set to be 100 μm, equivalent to the radius of the photodiode. The entrance aperture of the CPC was therefore 387 μm, which

is much larger than the shift resulting from the aberrations (around 18 to 70 μm). Since the effective aperture was increased by 4 times, the current AR can receive the entire power from misaligned rays up to $\pm$3.9 mrad. The design is shown in figure 3.9.

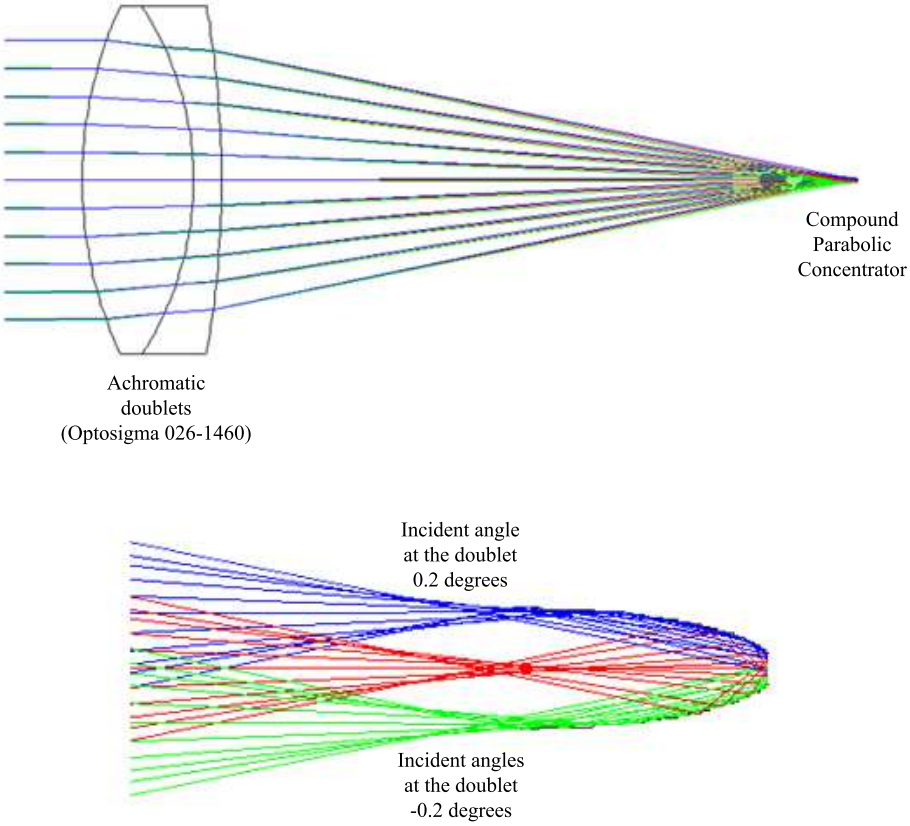

Figure 3.9: Compound parabolic concentrator integrated design: (Top:) A combination of an imaging system and a compound parabolic concentrator (Bottom:) A zoomed in view of the compound parabolic concentrator and the acceptance angles

3.3.2 Sensors

The incident rays from the remote transceiver produce a spot on the focal plane and the position is digitized by a position sensing diode. Three different position sensing diodes, (1) quadrant detectors (QD), (2) lateral effect photodiodes (LEP), and (3) charge coupled detectors (CCD), are available off-the-shelf. The first two diodes give an output which corresponds to the centroid of the total light quantity distribution on the active area, while the CCD samples the light quantity distribution at each pixel and outputs them sequentially. The QD and LEP are preferred in the FPAT system for their fast output data rate, whereas the CCD is often applied in the CPAT system for its multi-target acquisition capability.

3.3.2.1 Operating Principle

The QD utilizes four photodiodes positioned symmetrically around the center of the detector and separated by a narrow gap. If the focused beam is circularly symmetric, position information can be estimated by normalizing the intensity difference among the four photodiodes with the sum of the photodiodes.

In order to determine the displacement of the beam, all four photodiodes have to be illuminated by the focused beam. In other words, the size of the beam decides the measurable range (figure 3.10 (a)) and the beam must be expanded on purpose through defocusing to measure a larger displacement.

The displacement of the beam centroid on the QD in X and Y axes can be

formulated by

$$\begin{aligned} \frac{\sigma^x_{QD}}{d} &= k_{QD}\frac{((P_a+P_b)-(P_c+P_d))}{P_a+P_b+P_c+P_d} \\ \frac{\sigma^y_{QD}}{d} &= k_{QD}\frac{(P_a+P_d)-(P_b+P_c)}{P_a+P_b+P_c+P_d} \end{aligned} \tag{3.22}$$

where d represents the lateral span of the beam at the QD and k_{qd} represents the scaling factor converting the normalized voltage to the normalized distance.

The LEP consists of a single large-area photodiode with a uniform resistive sheet on both its cathode and anode. Two ohmic contacts are positioned at the opposite edges of each sheet, and the contact pairs are oriented perpendicularly to each other [78]. The photocurrent divided between the contacts is inversely proportional to the distance from the beam to each contact, which determines the position of the beam.

The centroid is estimated by normalizing the difference of the photocurrents, collected from the two contacts at each sheet, to their sum.

$$\begin{aligned} \frac{\sigma^x_{LEP}}{d} &= \frac{k_{LEP}(P_a-P_c)}{P_a+P_c} \\ \frac{\sigma^y_{LEP}}{d} &= \frac{k_{LEP}(P_d-P_b)}{P_d+P_b} \end{aligned} \tag{3.23}$$

where d represents the lateral span of the LEP, and k_{LEP} represents the scaling factor converting the normalized voltage to the normalized distance.

3.3.2.2 Signal-To-Noise Ratio

It is well known that in order to achieve a bit-error-rate of 10^{-9} in a communication system, the SNR ratio of the photodiode must be at least 21.5 db [13]. How about the signal-to-noise ratio of the PSD?

If the output of each photocurrent is a Gaussian variable whose variance results

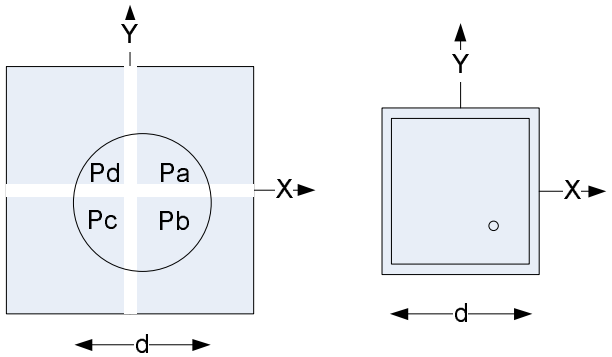

Figure 3.10: (a) Quadrant detectors (b) Lateral effect photodiodes

from additive noise, a correct output can be estimated by taking the average of multiple outputs because the distribution provides an unbiased estimator. However, the output of the PSD is the quotient between the difference of the photocurrents and the sum of the photocurrents. The average is a biased estimator to this quotient in a low signal-to-noise ratio scenario, which is shown below.

Let the output of each photocurrent be a Gaussian variable and the difference and sum of the photocurrents be X and Y, which are joint Gaussian.

Theorem 3.3.1. *Given two jointly Gaussian variables, X and Y, their joint distribution can be expressed as*

$$f(X,Y) = N(\begin{pmatrix} m_x \\ m_y \end{pmatrix}, \begin{pmatrix} n_x^2 & rn_xn_y \\ rn_xn_y & n_y^2 \end{pmatrix})$$

where r represents the cross correlation coefficients between X and Y.

For the quotient, $Z = X/Y$, its distribution can be derived as

$$\begin{aligned} f(Z=z) &= \frac{n_xn_y\sqrt{1-r^2}/\pi}{n_y^2(z-rn_x/n_y)^2+n_x^2(1-r^2)}\exp\frac{-C}{2(1-r^2)} \\ &+\frac{B/A}{\sqrt{2\pi A}n_xn_y}erf(\frac{B}{\sqrt{2(1-r^2)A}})\exp\frac{-m_y^2(z-\frac{m_x}{m_y})^2}{2(n_x^2-2zrn_xn_y+n_y^2z^2)} \end{aligned} \tag{3.24}$$

where

$$
\begin{aligned}
A &= \frac{n_y^2(z-\frac{n_x}{n_y}r)^2+n_x^2(1-r^2)}{n_x^2n_y^2},\\
B &= \frac{(m_xn_y^2-n_xn_ym_yr)(z-\frac{n_x}{n_y}r)+n_x^2m_y(1-r^2)}{n_x^2n_y^2},\\
C &= \frac{(n_ym_x-n_xm_yr)^2+n_x^2m_y^2(1-r^2)}{n_x^2n_y^2}\\
erf(x) &= \frac{2}{\sqrt{\pi}}\int_0^x exp^{-t^2}dt
\end{aligned}
$$

The previous equation is a sum of two terms and the first term is negligible in normal operation. The second term provides a maximum likelihood position equal to $\frac{m_x}{m_y}$ at two operating scenarios: (1) the displacement is close to the centroid and (2) a sufficient high SNR ratio, which can be derived from the probability function given in equation 3.24.

This probability density function of $\frac{X}{Y}$ was verified by a Monte-Carlo simulation and also compared to the Gaussian $N(\frac{m_x}{m_y},\frac{n_x^2}{m_y^2})$ distribution, one usually used to model the PSD output. The results are shown in figure 3.11, where the SNR of the upper and lower plots are 141 and 10 dB.

The weighting function in the picture refers to the $\frac{B/A}{\sqrt{2\pi A}n_xn_y}$ in the second term of equation 3.24, which causes the mean shift in a low SNR ratio. The results from the Monte-Carlo simulation show that the proposed probability distribution is closer to reality and it converges to the commonly used Gaussian with an increment of the SNR.

The bias of the average estimator is further investigated by using Monte-Carlo methods with $20,000$ samples to reduce the variance. Figure 3.12 shows the difference between the mean position and the position of $\frac{m_x}{m_y}$ at different SNRs. The

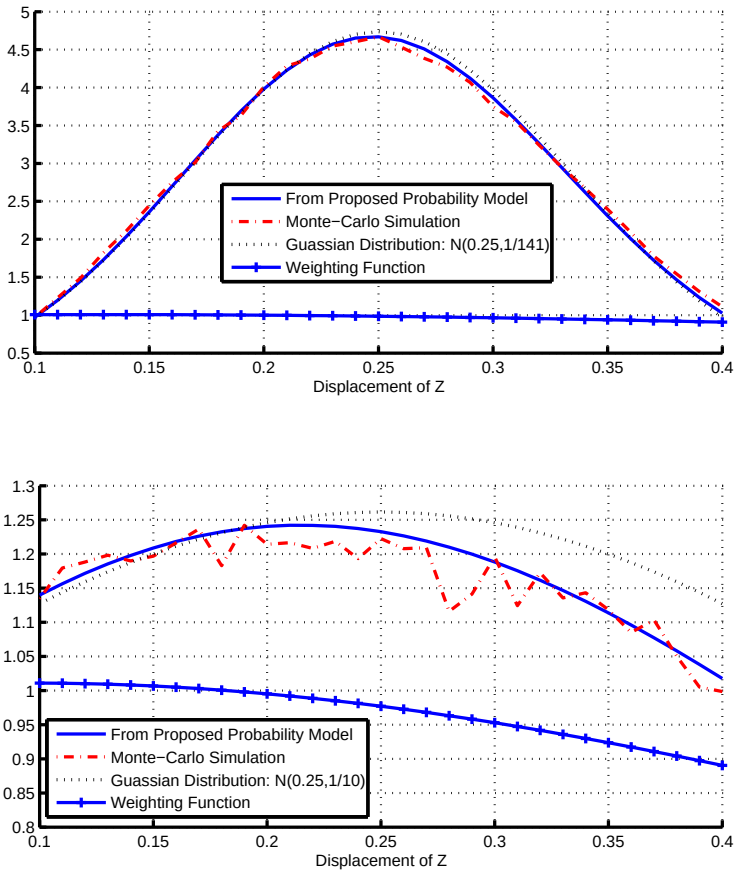

Figure 3.11: Maximum likelihood position shift under SNR (a) 141, and (b) 10 dB

results indicate that in order to have an unbiased estimation $< 0.5\%$ over the whole region, the minimum SNR must be above 24 dB.

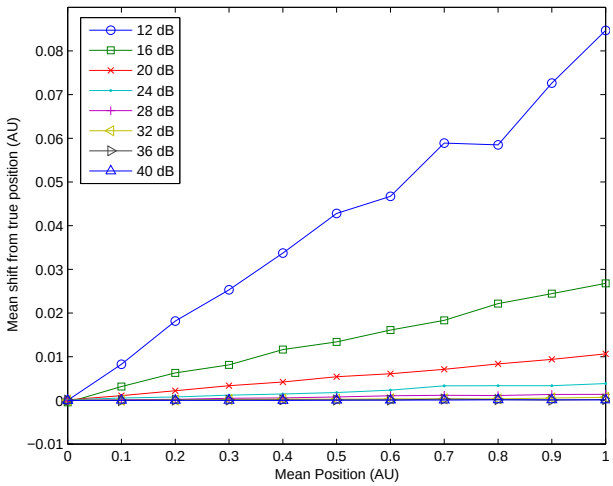

Figure 3.12: Mean shift from the normalization process for different SNRs

3.3.2.3 Comparison between QD and LEP

The comparisons are mainly in three aspects:

- Resolution: Assuming the SNR is sufficient for both QD and LEP, the standard deviation of the output is inversely proportional to the SNR of each photocurrent.

  In general, the noise sources of the photocurrent are from the preamplifier noise, such as shot noise and thermal (Johnson) noise [13] and the amplifier noise. In general, Johnson noise is regarded as the baseline noise, which bounds the resolution of the PSD, and can be represented as

$$V_N = \sqrt{\frac{4kT\triangle f}{R}}$$

where R represents the equivalent resistance of the sensor and k represents Boltzmann's constant.

The effective resistances of the LEP and QD are determined by the inter-electrode resistance of $10K\Omega$ and the feedback resistance of $10M\Omega$, respectively. Since the baseline noise is inversely proportional to the maximum resolution, it implies that QD has about 40 times better resolution than the LEP.

- Response Time: The response time of the PSD is determined by (1) the transit time of the photocarriers in the depletion region, (2) the diffusion time of the photocarriers generated outside the depletion region, and (3) the RC time constant used to guide the photocurrent to the associated circuit [36]. Generally, the RC time constant is dominant for both QD and LEP.

  The RC time constant of the pre-amplification circuitry is determined by the poles and zeros of the feedback loop used to stabilize the operation. This constant is inversely proportional to the operating bandwidth, formulated as $f = \frac{1}{2\pi RC}$, where R, C are the dominant resistance and capacitance of the diode. Normally, a stray capacitance across the feedback resistance is applied to compensate for the phase lag created by the internal capacitance.

  In a QD, since the small stray capacitance ($1 \sim 2$ pF) is sufficient to compensate for the internal capacitance (< 20 pF/quadrant), a feedback resistance of $10M\Omega$ can provide an operating bandwidth of 10 KHz. In a LEP, since the inter-electrode resistance ($10K\Omega$) and internal capacitance (100 pF) are

dominant, the operating bandwidth can be designed to be above 150 KHz.

- Linearity: The linearity of the QD is simulated by moving a uniformly illuminated circle toward the edge of the QD. The estimated displacement from the QD is computed from equation 3.22 and compared with the real beam displacement. The results are plotted in figure 3.13.

  It can be seen that the linearity of the QD between the estimated displacement and real displacement only exist within 20% of the beam radius. In other words, if the displacement of the beam is larger than that, the output of the QD does not correctly represent the image position.

  As for the LEP, the linearity of the whole sensor region is determined by the inter-resistance between the image spot on two electrodes. The linearity error is about 0.3% according to the specifications.

In conclusion, the QD is superior to the LEP in the resolution but worse in response time and linearity. In the FPM-based FPAT system, correct angular information on the AR is important since it determines the pointing angle of the beam steerer. Therefore, the FPM-based system must choose the LEP over the QD despite its larger noise, whereas the FSM-based system can select QD over LEP since the angular information is provided by the mirror sensors instead of the PSD.

3.3.2.4 Sensor Noise Measurement

This experiment characterized the sensor noise in an indoor environment. A pigtailed He-Ne laser was placed in front of a collimator and the parallel beam was

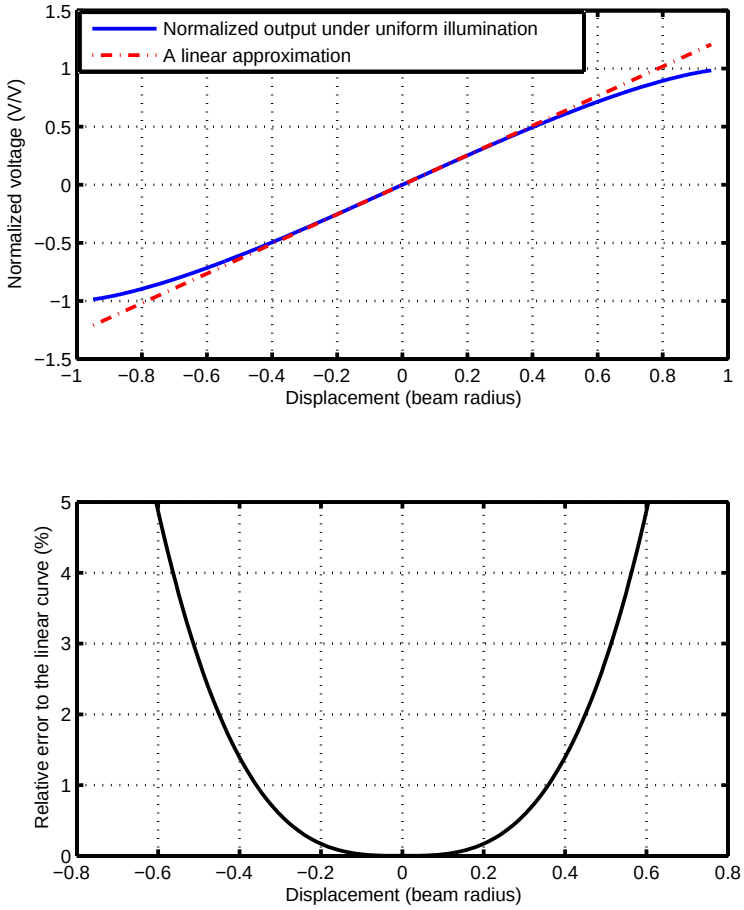

Figure 3.13: (a) The beam displacement with respect to the normalized output voltage (b) The error between the normalized output and its linear approximation

focused onto the PSD through a doublet. An attenuator and an interference filter were placed in front of the PSD to prevent saturation and background noise. The received laser power was 43.37 μW, which produced a photocurrent of 17.35 μA.

According to the specifications, the preamplifier noise is 1.3 $\frac{pA}{\sqrt{Hz}}$ and the response time is 5.66 μs, equivalent to 177 KHz, which results in the total noise of 0.54 nA and a sensitivity of -58.7 dbm for a responsivity at 0.4. The current amplifier has a current/voltage ratio of 64000, which amplifies the preamplifier noise to 34.6 μV for a unity noise figure.

Compared to the current measurement, the standard deviation of the signal fluctuation is 2.4 mV, which may result from (1) preamplifier noise, (2) amplification noise, and (3) background noise. Since the background noise is negligible because of the interference filter, it implies that the amplification noise has a noise figure around 18.42 db and the resulting sensitivity is -40.28 dbm. These fluctuations in the measurements contribute 12 μrad of angular uncertainty.

3.3.3 Calibration

As shown in section 2.2, the spot location of the AR and the transmitted ray vector is related by a linear homography. The calibration of the homography is the crucial step in the FPAT system since the beam steering process is an open-loop, which only relies on the spot position and the pre-calibrated homography without any feedback control. If the homography is erroneous, the beam steerer may never point to the target.

The goal of the calibration is to find the homography between the beam steerer and AR equivalent to $M_{LA}RM_{PSD}^{-1}$ in equation 2.9, where (M_{LA}, M_{PSD}) are affine transformation matrices with a form of $\begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & 1 \end{pmatrix}$ and R is the rotation matrix between the beam steerer and AR.

The homography calibration algorithm is listed in algorithm 5.1, which requires at least four corresponding vector pairs, from the fiber position as well as the focused image position. The main challenge is that the four transmitted rays from the fiber and the incident rays to the AR must be parallel. Figure 3.14 shows one alignment scheme which applies a corner cube reflector (CCR) to ensure parallellism between the transmitted and reflected ray. By moving the fiber to several locations, the corresponding AR outputs can be measured and the homography can be further estimated.

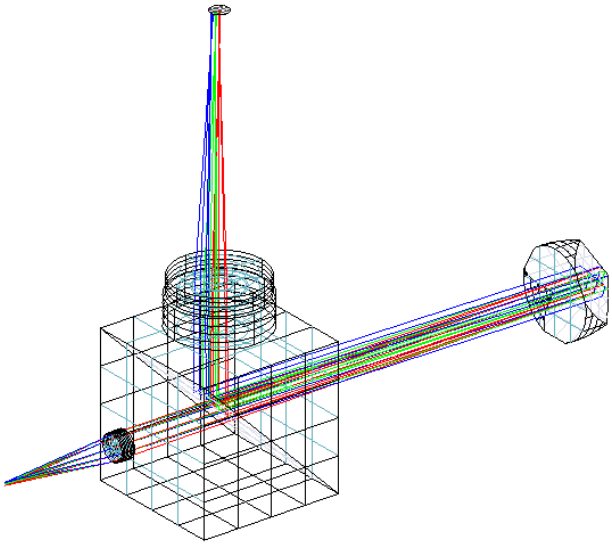

Figure 3.14: Simulation setup for the homography calibration

This scheme was implemented in CodeV, where the distance from the corner cube reflector to the beam splitter was 10cm and the fiber was moved in a 1×1 mm^2 area, at $z = 38.6$ mm. The optical systems of the collimator and AR are Optosigma 026-0190 and 026-1460, identical to the previous selections. Figure 3.15 shows the angular residual errors by dividing the difference between the estimated beam steerer angles and the real beam steerer angles by the focal length of the collimator. The maximum and average residual angular results are presented in table 3.2.

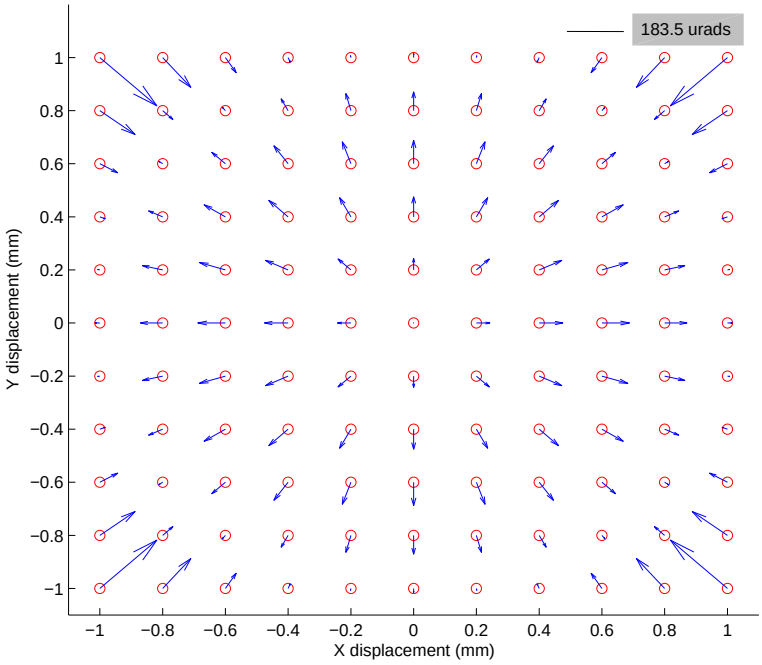

Figure 3.15: Residual errors obtained from the PAT optical system

A lab experiment was also conducted to verify the calibration accuracy and the setup was shown in figure 3.17. The PSD was selected to be the LEP from Ontrak. The distance from the beam splitter to the PSD was 10 cm and the fiber

moved in a $1 \times 1\ mm^2$ area close to the focal plane. The output from the PSD was measured by taking an average over 1500 samples to reduce the uncertainty. The residual errors are plotted in figure 3.16 and the maximum and average errors are listed in table 3.2.

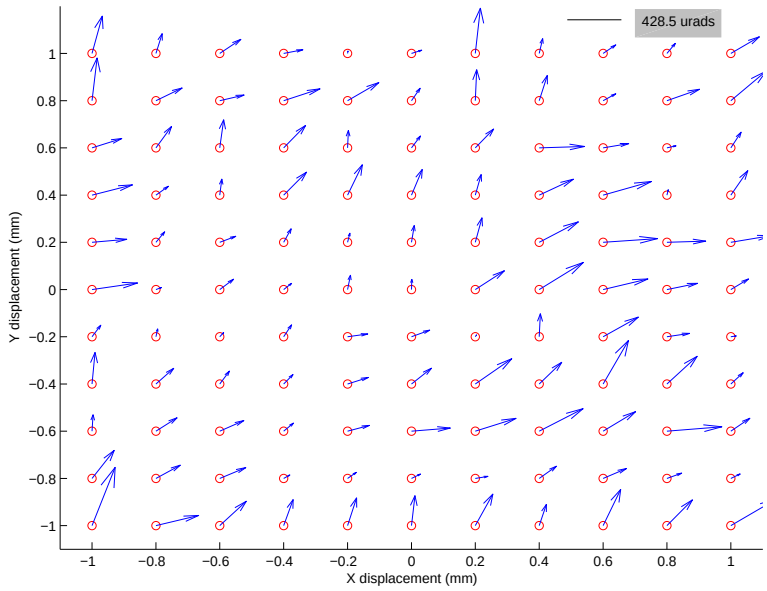

Figure 3.16: Residual errors from the homography experiment

The results show that the current homography is capable of providing an accuracy of 600 μrad to the beam steerer, which is within the range of the beam divergence (2 mrad). Compared to the simulation, the additional residual errors are about 3 times larger, which may result from the ambient light noise coupled from the transmitter since the combination of beam steerer and AR is mono-static.

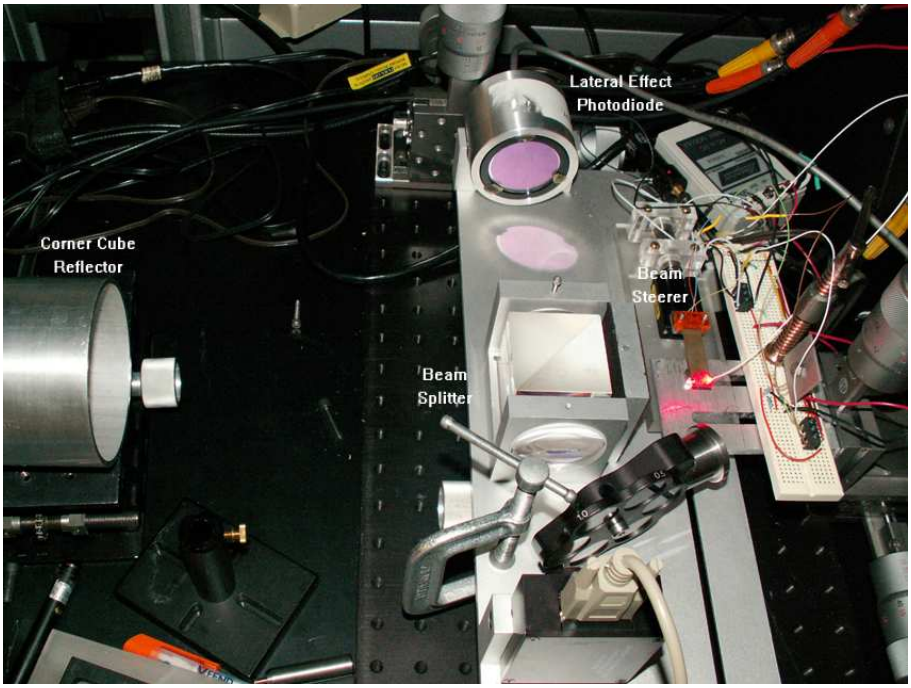

Figure 3.17: Experimental setup for the homography calibration

	Maximum Residual Angular Error(μrad)		Average Residual Angular Error(μrad)	
	X	Y	X	Y
Experiment	428.5	613.8	147.4	159.6
Simulation	183.5	183.5	44.47	44.47

Table 3.2: Experimental homography mapping errors

3.4 Control Systems

The PSD is responsible for providing reference positions to the moving stages in the beam steerer and AR to complete and optimize the FSO link. Given a reference position, the two-axis moving stage must settle fast enough before the image drifts off from the tolerance region, as described in section 2.3.2, which relies on a control system.

3.4.1 Axial Decouple

Should the two actuators be designed as either two separate SISO systems or one MIMO system? This question is entirely determined by the motion coupling between the two axes of the moving stage, which is verified by using the measurements obtained from the actuator model identification.

When the VCA in the X-axis was driven by a ± 1 V PRBS sequence, a span of ± 5 V in the X axis and a coupled signal in the Y-axis with a magnitude 0.25 V were produced. When the piezo-bender in Y-axis was driven by the same signal, it produced a span ± 3.4 V in the Y axis and the coupled signal in the X-axis of 0.2 volts. Therefore, the motion coupling between two axes are 2% and 3.68%, respectively, which is even smaller than the modelling error. Thus, we can conclude that the two axial motions are uncoupled.

3.4.2 Step Reference Follower

A step reference follower must be designed to move the fiber to a reference position computed from the AR. According to sections 2.4 and 2.5, a step reference follower can adapt either the disturbance rejecter or the reference follower structure [20]. An evaluation is provided in this section using the two identified models of the actuators.

3.4.2.1 Optimal Control

The controller gain can be determined from the optimal control theory in section 2.4.1 and the locus of the gain is identical to both the disturbance rejecter and reference follower since the augmented states do not alter the control input. The only question is to select a meaningful Q_1 and Q_2 for the cost function. Assuming the system model is

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) \end{aligned}$$

After Q_1 is selected to be C^TC, this minimizes the difference between the output signal to the reference signal and the optimal locus of poles can be computed by varying Q_2. We also assume that the settling time of the system must be less than 15 ms, which implies that all the desired poles in the unit circle must be within $Z < 0.7356$ if the controller is operated at 1 kHz. Thus, Q_2 of the VCA and piezo-bender are selected to be 0.2042 and 0.4677, respectively. The locus of the poles and the settling time criterion are both plotted in figure 3.18.

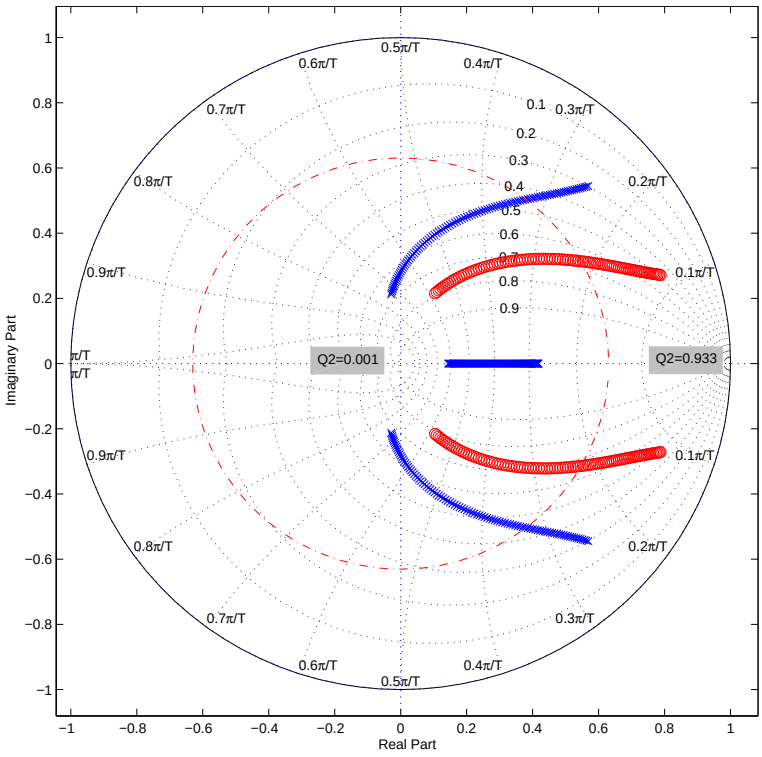

Figure 3.18: The locus of poles for the optimal controller: the cross signs are the locii of the VCA, the circular signs are the locii of the piezo-bender, and the dashed circle is the bound to achieve a desired settling time.

3.4.2.2 Kalman Estimator

Incorporating a constant reference signal into the disturbance rejecter and reference follower structures are quite different. In the former, a steady-state gain ($\bar{N}$) for the reference signal must first be computed from equation 2.31 and then a constant disturbance model is augmented into the system model to reject the difference between the computed and the real steady-state gain resulting from an imperfect model. In the latter, the constant reference model is directly augmented into the system model as shown in equation 2.32.

The key to applying Kalman estimators is to select the following parameters: the error covariance matrices of the two actuators and two-axis sensor and the coupling constant of the actuator noise that determines the impact of the actuator noise on the actuator states.

The error covariance matrices of the VCA and piezo-bender were estimated from the residual errors in the identification experiments (sections 3.2.2.4 and 3.2.2.2) since the main actuator noises were mainly from imperfect models. The two-axis sensor error is dominanted by the coupled motion from other axis, around 0.2 V. The coupling constant for both actuators was selected to be an identity matrix, which implies that the noise affects both the original and augmented states.

3.4.2.3 Step Reference Follower Simulations

Two controllers, based on disturbance rejecter and reference follower structures, were implemented by combining their individual optimal controller and Kalman

estimator. The frequency response of the actuator output with respect to the reference signal and the disturbance signals are plotted in figure 3.19.

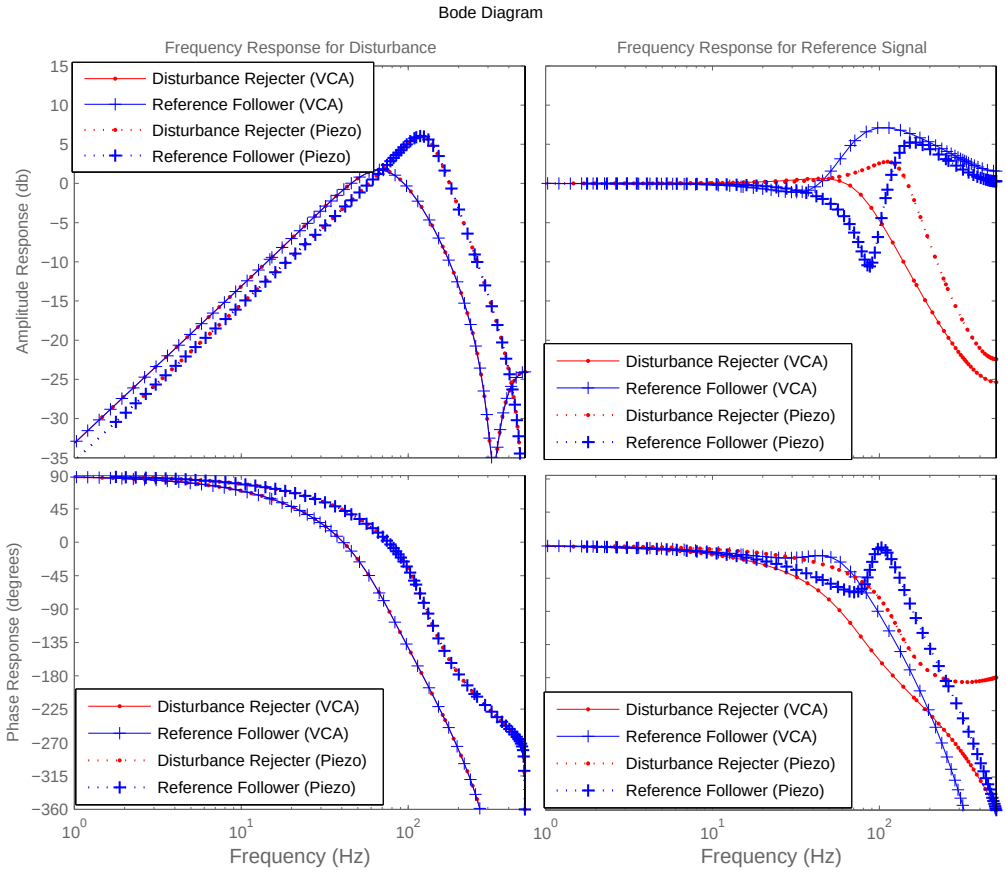

Figure 3.19: The simulated frequency response of the transfer function: (Left) between the reference signal and the resulting output and (Right) between the noise signal and the resulting output.

The results indicate that both structures provide identical disturbance rejection capability but the disturbance rejecter structure has a much smoother response for the constant reference signal. This advantage is from the complexity we paid

to compute the steady-state reference gain, which greatly reduces the magnitude of the constant disturbance because the close estimation of the actuator models leads to that of the real steady-state reference gain.

Besides, the results also indicate the capability of the piezo-bender and VCA. The VCA has a slower response compared to the piezo-bender because of a heavier load, which results in a 10 Hz bandwidth deficiency. However, since the VCA model is better estimated, a smaller Kalman gain is required to update the state parameters, which yields 5 db gain over high-frequency disturbance noise.

3.4.2.4 Step Reference Follower Experiment

The previous results showed the superiority of the disturbance rejecter structure and therefore it was selected to be implemented in the real FPAT system. The experimental setup is identical to the calibration experiment, shown in figure 3.17, without the corner cube reflector. Figure 3.20 shows that the 5% settling time of the VCA and piezo-bender are around 20 and 18 ms. It implies that the pointing errors of the VCA and piezo-bender are both smaller than 50 μrad after the fiber settles.

A white noise sequence, with 50 Hz 3-dB bandwidth, is applied to probe the noise rejection capability of the controller. The spectrum of the residual noise, defined as the difference between the output and reference signal, with/without applying the step reference controller is shown in figure 3.21.

The results show that the 3 db noise rejection bandwidth of the VCA is greater

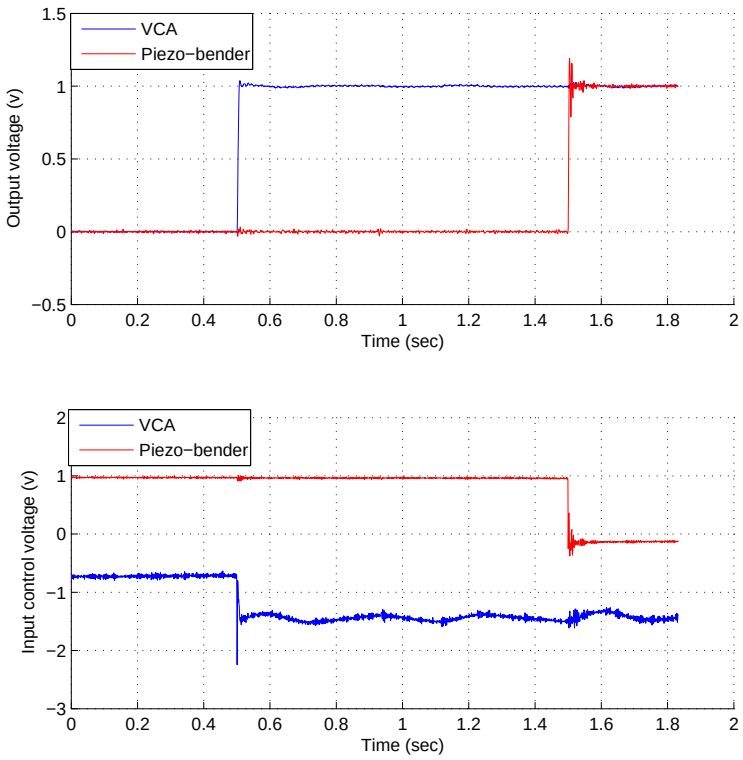

Figure 3.20: Step response of VCA and piezo-bender from the experiment.

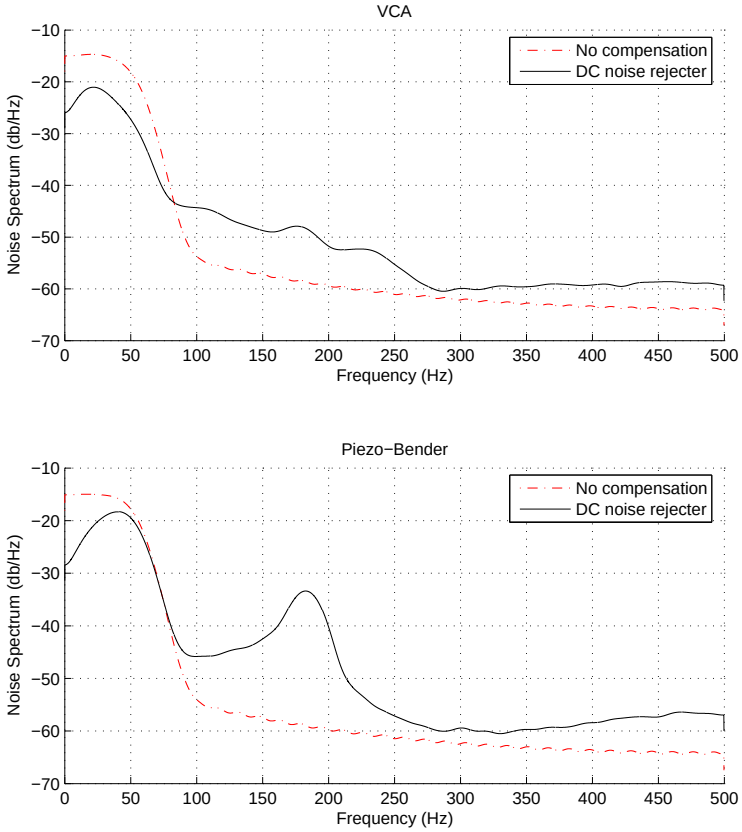

Figure 3.21: Residual noise spectral comparisons with/without the controller: (Top) VCA and (Bottom) Piezo-bender.

than 50 Hz and that of the piezo-bender is around 50 Hz. Compared to the simulation, the piezo-bender performs much worse than expectations, which requires further investigation.

3.4.3 Notch Controller

The previous design assumes a constant disturbance to model the difference between the estimated and real steady state reference output and treats other disturbances as white, which is not true in general. The residual error of the actuator clearly indicates that the disturbance errors resulting from the imperfect model normally produce large magnitude errors near the actuator's natural frequency, which motives us to test a notch controller design.

The idea is to convert the desired notch filter into a state-space form and augment it to the previous step reference follower model. The augmented model creates additional zeros at the desired frequencies resulting in a deep rejection band.

3.4.3.1 Notch Controller Simulation

Figure 3.22 shows the frequency response of the actuator output with respect to the reference and disturbance signals. The notch is assigned at the natural frequencies of the VCA and piezo-bender, 45 Hz and 112.5 Hz, respectively.

Compared to the disturbance rejecter, the notch controller produces a deep rejection band at the desired frequency with about 3 dB cost at the low frequencies and also provides an identical output response with respect to the reference signal.

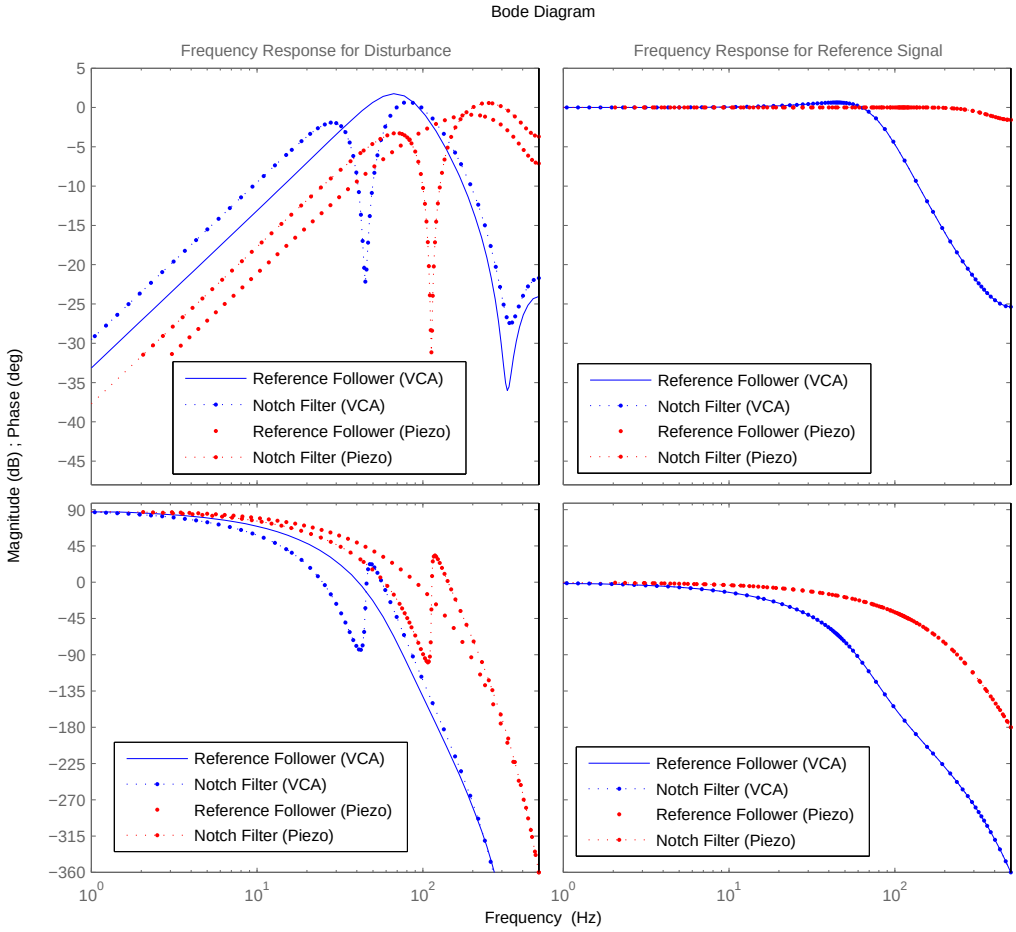

Figure 3.22: Simulated frequency response of the actuator output with respect to the reference and disturbance signals of a notch controller

3.4.3.2 Notch Controller Experiment

The notch filter was also implemented on the current two-axis moving platform. Except for the augmented states, all other coefficients are identical to the step reference follower.

A pseudo random binary disturbance with an input voltage ±0.3 V was applied to probe the noise rejection performance of the step reference follower and the notch controller. Figure 3.23 shows the residual noise spectrums of different controllers.

The results show that the notch controller effectively mitigates the induced peak disturbance at 45 Hz for the VCA and 120 Hz for the piezo-bender but imposes a penalty at other frequencies. The average residual noise amplitude falls 61% in the VCA and 11% in the piezo-bender. The inferior performance of the piezo-bender may result from the small drift of the natural frequency, which makes the notch unable to compensate for the largest noisy components. The frequency drift comes from the nonlinearity of the hysteresis, which drives the natural frequency higher when the input amplitude is smaller [33].

3.5 Link Budget Computation

The previous specifications of beam steerers and AR are combined to specify a enhanced FSO transceiver, operating at 660 nm, with 100 Mb/s transmission speed, and a link length around 500 m. An ideal photodiode (FDS10) from Thorlabs operates at wavelength from 200 to 1100 nm and has a noise equivalent power (NEP) of $5 \times 10^{-14}\ W/\sqrt{Hz}$. The path loss of the link is computed in table 3.3.

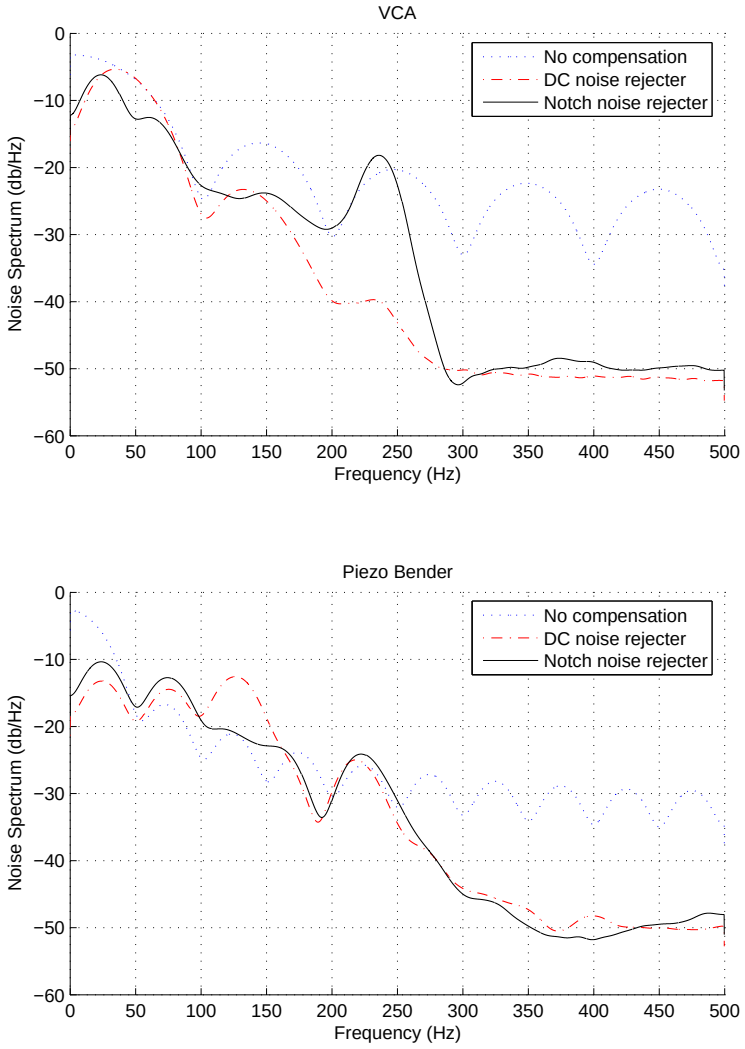

Figure 3.23: Residual noise spectral comparisons among three scenarios: no controller, step reference follower, and notch controller: (Top) VCA and (Bottom) Piezo-bender

System Components	Descriptions	Power
Fiber loss	32% efficiency	−4.9 dB
Collimator loss	Due to a small aperture and 90 transmission	−1.32 dB
Double Pane Window	15% reflectivity	−0.71 dB
Atmospheric Transmission	Assuming 10 km visibility	−1.7 dB
FSO loss	500 m, 2 mrad, and 50 mm receiving diameter	−26.02 dB
Beam Splitter	50% transmission	−3 dB
Receiving Lens	90% transmission	−0.46 dB
Interference Filter	50% transmission	−3 dB

Table 3.3: Link Path Loss Analysis

The total path loss from the remote transmitter to the local receiver is 41.11 dB. The communication photodiode, operating at 100 Mbps, has a maximum sensitivity around −63.01 dBm, and the PSD, operating at 15 Kbps, provides a maximum sensitivity around −58.7 dBm. Also, the communication system requires a 21.5 dB gain to achieve the BER performance of 10^{-9} and the current LEP requires a 24 dB gain to achieve an estimation error $< 0.5\%$. By providing a 10 dB link margin to both the communication and acquisition links, the required laser powers for each system are 9.12 mW and 43.75 mW.

3.6 Summary

This chapter introduced a novel FPAT prototype based on moving the fiber near the focal plane, also known as the focal plane motion, which can be an economic

solution to solve the coupled alignment problem between two FSO transceivers. This prototype includes a beam steerer and an angular resolver to transmit and receive optical signals. The main idea is to calibrate the mapping, shown to be a linear homography, between the AR and the fiber position in the beam steerer. Then, once an incident ray from the target is imaged on the local AR, the corresponding fiber position is determined, which guarantees accurate pointing from the beam steerer to the target.

The steering resolution, range, and beam divergence of the beam steerer are designed to be 1 mrad, $\pm 1.43^\circ$ (50 mrad), and 2 mrad, respectively. The simulations indicate that the specifications can all be met by the current selections, including the achromatic lens (026-0190) from Optosigma and the optical fiber (SM-600) from Thorlabs. More importantly, the simulation shows that the focal plane motion has almost no effect on the steering angle and varies only by 20% of the beam divergence.

The angular resolver, consisting of an optical system and a PSD, determines the pointing resolution of the system, since the AR contributes most of the errors to the estimated homography: If the signal-to-noise ratio of the PSD is sufficient, the aberrations of the optical system are the dominant errors; otherwise, the error brought by a biased position estimation from the PSD is dominant. The homography is crucial because it is like an open-loop operation, which cannot be corrected in the operation of the FPAT system.

A calibration procedure to retrieve the homography was introduced and the performance was evaluated by experiment. The results indicated that the average and maximum pointing error by using the estimated homography as the mapping is

about 217 μrad and and 730 μrad, respectively, which satisfies the accuracy specification. Also, once the fiber position is computed, a closed-loop controller is capable of settling the motion within 20 ms and maintaining the pointing error within 50 μrad, which were also confirmed by experiment.

Chapter 4

Theoretical Studies in Coarse Angular Pointing, Acquisition, and Tracking Systems

This section provides a theoretical background for a coarse angular PAT (CPAT) system design based on a camera based methodology, also known as the spatial scan method described by Rolland [71]. It covers five important topics including:

1. Chromatic Aberrations: To recognize the targets of interest (TOI), the CPAT system must rely on some unique features of the targets and one of them is color. Therefore, if the CPAT utilizes imaging systems to recognize the target, the receiving optics do not only suffer from the monochromatic aberrations but also from the chromatic aberrations.

2. Perspective Imaging Theorem and Homogeneous Coordinates: The perspective imaging model is identical to the paraxial imaging model, in which the aberration is negligible. This model, also known as the pin-hole model in the computer vision community [23, 68], has been widely applied to the camera imaging system because the imaging lens is always designed to have minimum aberrations.

   A camera consists of an imaging lens and a position sensing diode, which utilizes a coupled charge detector (CCD) instead of QD or LEP. Compared to the AR of the FPAT system, the camera based system is capable of monitor-

ing individual light spots imaged on each pixel, which provides considerable freedom for selecting targets of interest (TOI). However, because of imperfect manufacturing, the optical axis of the imaging lens and that of the CCD may not be the same [77]. In addition, the pixel array on the CCD has a different spacing in X and Y axes in order to create a better visualization for users. The ratio between the spacing is called the aspect ratio, which requires it to be calibrated to properly retrieve the embedded geometry.

3. Forward and Backward Error Propagation: In the CPAT system, all the parameters used for the pointing procedure must be retrieved in advance through a calibration procedure. The measurement errors from the procedure are coupled into the estimated parameters, which result in further alignment errors in the pointing process. This section describes a methodology to compute the estimation errors and the pointing errors from random measurement errors, which can serve as an indicator to measure the robustness of the calibration procedure.

4. Two-Axis Rotary Gimbal Model: A motorized platform is selected as the actuator to carry the FSO transceiver because of its large pointing range and fast acquisition speed. Since the link distance in most FSO applications is long, a rotary-based motorized system is a better candidate compared to the translation-based system. This section develops a geometric model to map the transceiver ray vector to the azimuthal and elevation angles of the two-axis rotary gimbal.

5. Radial Trifocal Tensors: A trifocal tensor is capable of encoding the geometric relation embedded among three different frames. Many researchers apply this technique to reconstruct a 3D structure of the objects from three perspective cameras [23]. We adapt this technique to design a three-dimensional CPAT system and the details are revealed in section 5.2.

   We also discover that the trifocal tensor is actually over-constrained for the problem since it provides four dimensional geometric information (2 from each perspective camera) to estimate a three-dimensional position. Since most of the wide-angle cameras suffer from only distortion, which only appears in the radial direction, 1 dimensional geometric information is still preserved. By replacing one of the two perspective cameras with a wide-angle camera, the resulting system is still sufficient to uniquely determine a 3D position but with a much wider fov. Instead of the trifocal tensor, the geometric relationship among the three components are encoded by a radial trifocal tensor (RTT). This section is dedicated to introducing the geometric properties of the RTT and the detailed CPAT design is revealed in section 5.4.

Throughout this chapter, tensor notation is applied and summarized below. Image points and lines are represented by homogenous column and row 3-vectors, respectively. For example, given a 3D point $x = (x^1, x^2, x^3)^T$ and a line vector $l = (l_1, l_2, l_3)$, their tensor notations are x^i and l_i. The ij-th entry of a matrix A is denoted by a^i_j, index i being the contravariant (row) index and j being the covariant (column) index. The equation $x' = Ax$ is equivalent to $x'^i = \sum_j a^i_j x^j$,

which is written as $x'^i = a^i_j x^j$ in repeated subscript tensor notation.

4.1 Chromatic Aberrations

In addition to achromatic aberrations, the camera suffers from chromatic aberration, which results from the wavelength dispersion of lens material. The transverse wavefront aberration can be analyzed by using two parallel rays. The wavefront aberration with respect to the wavelength variation is the following

$$\delta_\lambda W = \sum_{k=1}^{a} -i_k h_k \left(\frac{\delta n_k}{n_k} - \frac{\delta n_{k-1}}{n_{k-1}}\right) \tag{4.1}$$

where a represents the number of elements in the optical system, h_k is the height of the second ray at surface k, δn_k represents the small refractive index change from the wavelength variation, n_k is the refractive index after refracting at surface k, and i_k is the incident angle at surface k.

The overall chromatic transversal shift is equal to

$$\delta_\lambda \eta = \frac{-R_a}{n_a h_a} \sum_{k=1}^{a} i_k h_k \left(\frac{\delta n_k}{n_k} - \frac{\delta n_{k-1}}{n_{k-1}}\right) \tag{4.2}$$

where R_a represents the distance from the exit pupil to the image plane.

4.2 Perspective Imaging Theorem and Homogeneous Coordinates

A camera system consists of an imaging lens and a CCD array, which is similar to the components of the angular resolver in the previous chapter. Compared to the angular resolver, a camera has the capability to detect multiple targets simultaneously because each pixel in the CCD array operates independently.

Since the camera is a pixel-based system, its resolution is limited by (1) the aberrations of the lens and (2) the spacing of pixels of the CCD array. Typically, the spacing of the pixel works as the upper bound of the resolution and the imaging lens is designed to suppress the aberrations in order to approach that bound.

4.2.1 Homogeneous Coordinates

If the aberrations of the lens are negligible, the incident ray and focused position can be modelled by the perspective imaging model, in which both incident angles (ϕ, θ) are assumed to be preserved. The details of this imaging model are plotted in figure 4.1.

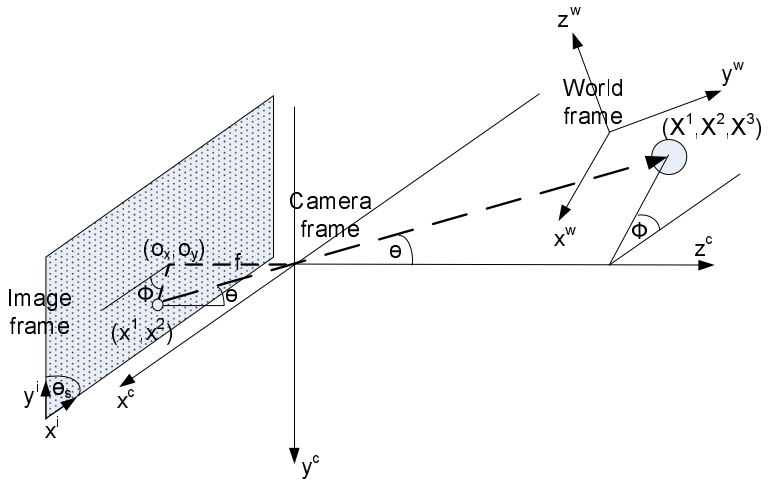

Figure 4.1: Perspective camera imaging model: the bold-dashed lines indicate that any image point on the image plane actually represents a ray vector connecting from the object to the center of the camera, which is identical to a human's perspective.

In the camera frame, the image can be formulated as

$$\begin{aligned} \frac{X_c^1}{X_c^3} &= f\tan\theta\cos\phi \\ \frac{X_c^2}{X_c^3} &= f\tan\theta\sin\phi \end{aligned} \tag{4.3}$$

This image is further sampled by the CCD array for processing. The sampling process includes (1) a non-uniform sampling in X and Y axes, (2) an axial skewness (θ_s) between the two axes, and (3) an origin shift, shown as:

$$\begin{aligned} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} &= \begin{pmatrix} \frac{1}{d_x} & 0 \\ 0 & \frac{1}{d_y} \end{pmatrix} \begin{pmatrix} \sin\theta_s & \cos\theta_s \\ 0 & 1 \end{pmatrix} \begin{pmatrix} f & 0 \\ 0 & f \end{pmatrix} \begin{pmatrix} \frac{X_c^1}{X_c^3} \\ \frac{X_c^2}{X_c^3} \end{pmatrix} + \begin{pmatrix} o_x \\ o_y \end{pmatrix} \\ &= \begin{pmatrix} \frac{f\sin\theta_s}{d_x} & \frac{f\cos\theta_s}{d_x} \\ 0 & \frac{f}{d_y} \end{pmatrix} \begin{pmatrix} \frac{X_c^1}{X_c^3} \\ \frac{X_c^2}{X_c^3} \end{pmatrix} + \begin{pmatrix} o_x \\ o_y \end{pmatrix} \end{aligned}$$

where (x^1, x^2) represent the image position, (X_c^1, X_c^2, X_c^3) represent the 3D target position on the camera frame, f represents the focal length of the optical system, (d_x, d_y) represent the spacing of the CCD array of the X and Y axes respectively, and (o_x, o_y) represent the origin shift, also known as the principal point.

By introducing the homogeneous coordinates, where a 1 is inserted into the image frame, all these parameters can be incorporated as a linear mapping converting the camera frame into the image frame as the following equation:

$$\begin{pmatrix} x^1 \\ x^2 \\ 1 \end{pmatrix} = \begin{pmatrix} f' & f's & o_x \\ 0 & f'\alpha & o_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{X_c^1}{X_c^3} \\ \frac{X_c^2}{X_c^3} \\ 1 \end{pmatrix} \tag{4.4}$$

where $\alpha = \frac{d_x}{\sin\theta_s d_y}$ is defined as the aspect ratio of the CCD, $s = \cot\theta_s$ is defined as the skewness between two axes, and $f' = \frac{f\sin\theta_s}{dx}$ is defined as the effective focal length of the camera.

This linear 3×3 matrix used to map the camera frame to the image frame is defined as the intrinsic matrix, which contains 5 unknown coefficients[68], $(f^{'}, \alpha, s, o_x, o_y)$.

Physically, the homogeneous expression implies that an image point in the perspective camera is not merely a 2D planar point (x, y) but a line $(x, y, 1)^T$ which connects the camera center and the 3D object. This point-of-view matches with the perspective graphical painting concept as described by Wikipedia [30], which results in the notation of "perspective camera". This line is composed of two vectors: a line on the image plane and a line from the image plane to the camera origin, which are shown in bold-dashed in figure 4.1.

4.2.2 World Coordinates Transformation

In order to generalize the previous imaging model, we introduce an arbitrary frame system, defined as the world frame. Since the target and the camera are rigid bodies, the transformation between the world and the camera frame can be formulated by a rotation matrix and a translation vector, where :

$$\begin{pmatrix} x^1 \\ x^2 \\ 1 \end{pmatrix} = \frac{1}{\lambda} M \begin{pmatrix} R & T \end{pmatrix} \begin{pmatrix} X_w^1 \\ X_w^2 \\ X_w^3 \\ 1 \end{pmatrix} \tag{4.5}$$

where $\lambda = R_j^3 X_w^j + T^3$ represents the depth, defined as the projected distance on the optical axis from the target to the camera, M represents the intrinsic matrix, and $\begin{pmatrix} R & T \end{pmatrix}$ is denoted as extrinsic matrix (R represents a 3×3 rotation matrix and T represents the translation vector between the world and camera frame).

For a single camera, λ is normally treated as an unknown scaling factor since it cannot be calibrated from the image, known as the depth ambiguity. Therefore, equation 4.5 implies that the two position vectors, (1) transforming from the world frame to the image frame and (2) the measured image point, should be parallel. This parallelism can be easily checked by taking the cross product as

$$\begin{pmatrix} x^1 \\ x^2 \\ 1 \end{pmatrix}_{\times} M \begin{pmatrix} R & T \end{pmatrix} \begin{pmatrix} X_w^1 \\ X_w^2 \\ X_w^3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{4.6}$$

where $(a)_{\times}$ defines a skew-symmetric matrix as $\begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}$ representing the cross product operator of $a : (a_1, a_2, a_3)^T$, which is related to the cross product according to $a \times b = (a)_{\times} b$

Given each position vector pair, this equation provides two independent constraints for solving the linear mapping. Since the rotation matrix and the translation vector contain 3 unknowns each, this linear mapping contains 11 unknown parameters. Therefore, all the coefficients can be solved by 6 3D points and their corresponding image points.

4.3 Estimation Error Propagation

In both the CPAT and FPAT systems, a calibration procedure is required for an open-loop mapping, which transforms the sensor frame to the actuator frame. Most of the calibration procedures are developed based on the validity of the paraxial/perspective imaging model and the aberrations are treated as additive random noise, except for distortion.

The randomness results from the uncontrollability of the achromatic aberrations since their magnitude depends on the incident position of the ray at the entrance pupil and cannot be determined by observation [10]. If one takes the chromatic aberration in section 4.1 into consideration, the imaging position is further randomized because it also depends on the incident position. Therefore, the calibration algorithm has to be robust with respect to the random errors and a theoretical evaluation must be developed to evaluate the algorithm. The following sections are dedicated to developing a methodology to (1) evaluate the robustness of the algorithm under measurement noise, (2) model the impact of the measurement noises on the estimated parameters and the further pointing error. The camera calibration process in section 4.2 serves as an example to describe the methodology.

4.3.1 Error Functions

Error functions are important in the parametric estimation process because they serve as an indicator to verify whether the optimum has been achieved and also as a mean to update the optimum iteratively. As for the camera calibration

process, the most intuitive error function should be the Euclidean distances, calculated between the measured image points x and the re-projected points $M(R|T)X_w$. However, the Euclidean-based error function used in the camera calibration can be highly nonlinear and its optimizer cannot be estimated directly. In general, the optimization of the Euclidean distances must be estimated through iterative algorithms, such as the Newton method or Levenberg-Marquardt algorithm . These algorithms need to have a sufficiently good initial guess in order to converge to the correct optimizer. Such guesses usually rely on other error functions whose minimizer has a closed-form solution and provides a sub-optimal estimation. For example, the parallelism between the measured points x and the computed points $M(R|T)X_w$ is selected in the camera calibration application.

4.3.1.1 Algebraic Distance

Equation 4.6 measures the parallelism between the measured and computed points by taking the norm of the cross product between them. This function is known as the algebraic distance and its minimizer provides a sub-optimal estimation of the mapping parameters:

$$d_{alg}(x, M(R|T)X_w) = \left\| \begin{pmatrix} 0^T & -X_w^T & x^2 X_w^T \\ X_w^T & 0 & -x^1 X_w^T \end{pmatrix} h \right\|^2 \tag{4.7}$$

where h represents the vectorization of $M(R|T)$, a 12×1 vector.

Let the right matrix be A. By stacking multiple position vector pairs (x, X_w), the mapping parameters can be estimated in the total least squares sense (TLS) as algorithm A.1. The position vectors must be normalized to 0 mean and an unit

length in advance [22] to increase numerical stability.

4.3.1.2 Mahalanobis and Euclidean Distance

The minimizer of the algebraic distance implies that all the position vector pairs are best parallelized without optimizing the re-projection error. If the covariance matrix of the noise is known, the Mahalanobis distance is commonly selected to evaluate the re-projection error and is listed below. If the additive noise is produced by a stationary Gaussian process, the minimizer of the Mahalanobis distance is also the maximum likelihood estimator.

$$d_{mah}(x, M(R|T)X_w) = \left\|(x - M(R|T)X_w)^T \Sigma_x^{-1} (x - M(R|T)X_w)\right\|^2 \tag{4.8}$$

where Σ_x is the covariance matrix of x.

However, the covariance matrix is not easily estimated in reality. An inferior selection is to evaluate the re-projection error in Euclidean distance. If the covariance matrix of the noises is isotropic and independent, minimizing the re-projection error in Mahalanobis distance is identical to minimizing it in Euclidean distance.

4.3.2 Forward Error Propagation

If the estimated mapping parameter P, equal to $M(R|T)$ in our case, is a random vector with mean $\bar{P}$ and variance Σ_P, the image position $f(P)$ computed from the parameter P is a random vector, where f is the function mapping P to X.

Theorem 4.3.1. *[23] Let P be a random vector in $\mathbb{R}^M$ with mean $\bar{v}$ and covariance matrix Σ_p and $f : \mathbb{R}^M \to \mathbb{R}^N$ is differentiable in a neighborhood of $\bar{P}$. Then, up to a*

first-order approximation, $f(P)$ is a random variable with mean $f(\bar{P})$ and covariance $J\Sigma_P J^T$, where J is the Jacobian matrix of f, evaluated at $\bar{P}$.

Proof. Assuming X is the image of P, the first order approximation gives

$$X = f(\bar{P}) + J(P - \bar{P}) \tag{4.9}$$

Therefore, X has a mean at $f(\bar{P})$ and a variance equal to $J\Sigma_P J^T$. □

4.3.3 Backward Error Propagation

Theorem 4.3.2. *[23] Let $f : \mathbb{R}^M \to \mathbb{R}^N$ be a differentiable mapping and let J be its Jacobian matrix evaluated at a point $\bar{P}$. Suppose that J has rank M and such that f is one-to-one in a neighborhood of $\bar{P}$. Let S_M be the image of f in $\mathbb{R}^N$ and let a mapping $\eta : \mathbb{R}^N \to S_M$. Let $f^{-1} \circ \eta : \mathbb{R}^N \to \mathbb{R}^M$ be the mapping that maps a measurement X to the set of parameters corresponding to the ML estimate $\bar{X}$. Let X be a random variable in $\mathbb{R}^N$ with mean $\bar{X} = f(\bar{P})$ and covariance matrix Σ_x. Then, $\hat{P} = f^{-1} \circ \eta(X)$ is a random variable with mean $\bar{P}$ and covariance matrix $J^T \Sigma_x^{-1} J$.*

Proof. If the measured image points X is a Gaussian random vector with a mean $\bar{X}$ and a covariance matrix Σ_x, the minimizer $\hat{X} = f(\hat{P})$ of the re-projection error in Mahalanobis distance is also the one maximizing the likelihood function $P(X|\hat{P})$ defined as

$$P(X = X|\hat{P}) = \frac{1}{2\pi^{N/2} det(\Sigma_x)} \exp^{\frac{1}{2}(X - f(\hat{P}))^T \Sigma_x^{-1} (X - f(\hat{P}))}$$

According to the first-order approximation where $\hat{X} = f(\hat{P}) = f(\bar{P}) + J(\hat{P} - \bar{P})$, the optimization problem can be formulated as

$$\min_{\hat{p}} \left\| X - \hat{X} \right\|_{\Sigma_x} = (X - \bar{X} - J(\hat{P} - \bar{P}))^T \Sigma_x^{-1} (X - \bar{X} - J(\hat{P} - \bar{P})) \tag{4.10}$$

Taking the derivative of $\hat{P} - \bar{P}$ leads to the maximum likelihood estimator as

$$\hat{P} - \bar{P} = (J^T \Sigma_x^{-1} J)^{-1} J^T \Sigma_x^{-1} (X - \bar{X}) \tag{4.11}$$

Therefore, the estimator of the parameters $\hat{P}$ is unbiased and its covariance matrix is $(J^T \Sigma_x^{-1} J)^{-1}$ if the measured image points have a covariance matrix Σ_x. □

The previous theorem derives the optimizer of an unconstrained optimization problem. Some parameter estimation problems may be constrained by additional criterions. For example, although there are 12 parameters in the camera calibration procedure, equation 4.5 can only estimate 11 parameters from the position vector pairs because of the depth ambiguity. It leaves us room to impose additional constraints to solve the problem easier, such as the norm is 1 in the TLS solution. The error performance of a constrained optimization problem is analyzed below.

Theorem 4.3.3. *[23] Let $f : \mathbb{R}^M \to \mathbb{R}^N$ be a differentiable mapping taking a parameter vector $\bar{P}$ to a measurement vector $\bar{X}$. Let S_p be a smooth manifold of dimension d embedded in $\mathbb{R}^M$ passing through point $\bar{P}$, and such that the map f is one-to-one on the manifold S_p in a neighborhood of $\bar{P}$, mapping S_p locally to a manifold $f(S_p) \subset \mathbb{R}^N$. The function f has a local inverse, denoted f^{-1}, restricted to the surface $f(S_p)$ in a neighborhood of $\bar{X}$ and let $\eta : \mathbb{R}^N \to f(S_p)$ be the mapping that takes a point in $\mathbb{R}^N$ to the closet point on $f(S_p)$ with respect to the Mahalanobis*

norm. Via $f^{-1} \circ \eta$, the probability distribution on $\mathbb{R}^N$ with covariance matrix Σ_x induces a probability distribution on $\mathbb{R}^M$ with a covariance matrix of first-order equal to

$$\Sigma_{\hat{P}} = Z(Z^T J^T \Sigma_x^{-1} J Z)^{-1} Z^T$$

where Z is a $M \times d$ matrix whose column vectors span the tangent space to S_p at $\bar{P}$.

Proof. A constraint, such as $C(P) = N$, bounds the minimizer to be on a submanifold $S_p \in R^d$. A map $g : R^d \to S_p$ is defined to map an open neighborhood U in R^d to an open set S_p which includes $\bar{P}$. Therefore, $g \circ f$ is one-to-one on the neighborhood U. If the Jacobian of U is Z, the Jacobian of $g \circ f$ is equal to ZJ. According to the previous theorem, the covariance matrix of a point in U is equal to $(Z^T J^T \Sigma_x^{-1} J Z)^{-1}$. By forwarding the error using the first theorem, the resulting covariance matrix of $\hat{P}$ can be shown as

$$\Sigma_{\hat{P}} = Z(Z^T J^T \Sigma_x^{-1} J Z)^{-1} Z^T \tag{4.12}$$

□

4.3.4 Camera Calibration Error Performance

The error performance of the camera parameter estimation due to the measurement error can be characterized by theorem 4.3.3. If we apply the estimated mapping to estimate another 3D point, the estimation error is defined by theorem 4.3.1. This section derives a theoretical error performance of the camera calibration application.

4.3.4.1 Calibration Error

Let us consider the most general case of the camera calibration, where the position vectors in world and image frames are both perturbed by additive noise, and simplify the notion of the linear mapping notion $M(R|T)$ as H. In this scenario, the parametric space includes: a linear mapping (H) and a set of estimated 3D points $\hat{X_w}$. The dimension of the parametric space is $3n + 12$, where n is the number of corresponding vector pairs. Also, because of the depth ambiguity, the scaling of the estimated homography is unconstrained. Therefore, the number of the essential parameters is actually $3n + 11$.

The dimension of the measurement space is $5n$ including $2n$ measurements from the camera points and $3n$ measurements from the 3D points, where n is the number of corresponding vectors. We define the following notion:

$$\begin{cases} X, & \text{Measured position vector pairs;} \\ \hat{X}, & \text{Estimated position vector pairs;} \\ \bar{X}, & \text{True position vector pairs;} \\ \left\| \hat{X} - \bar{X} \right\|, & \text{Estimation error;} \\ \left\| X - \hat{X} \right\|, & \text{Residual error.} \end{cases}$$

The maximum likelihood minimum occurs on the tangent space of the function f[49], whose dimension is equal to the number of essential parameters. A geometric plot is shown in figure 4.2.

Combining equations 4.11 and 4.12, the MLE with imposed constraints can

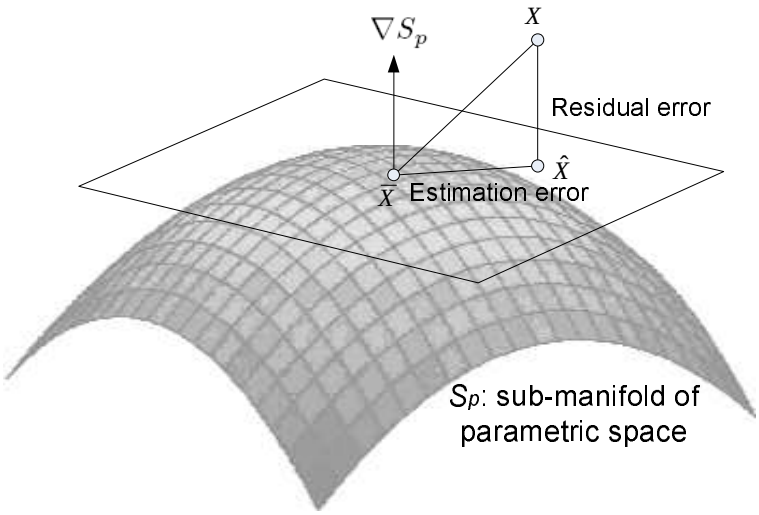

Figure 4.2: Geometry of the errors in measurement space: $\hat{X}$ is the estimated point which has the minimum distance to the measured point X

be formulated as

$$\hat{P} = JZ(Z^T J^T \Sigma_x^{-1} JZ)^{-1} Z^T J^T (X - \bar{X}) + \bar{P}$$

where h represents the vectorized H, Σ_x represents the covariance matrix of the measurements, $J = \begin{pmatrix} 0 & I \\ J_h & J_{X_w} \end{pmatrix}$, and $(J_h, J_{X_w}) = (\frac{\partial f}{\partial h}|_{h=\bar{h}, X_w=\bar{X_w}}, \frac{\partial f}{\partial X_w}|_{h=\bar{h}, X_w=\bar{X_w}})$.

The estimated position vector pairs are derived by plugging the MLE back into equation 4.9 as

$$\hat{X} = \bar{X} + JZ(Z^T J^T \Sigma_x^{-1} JZ)^{-1} Z^T J^T (X - \bar{X}) \tag{4.13}$$

Finally, the estimation error can be formulated as

$$\left\| \hat{X} - \bar{X} \right\|^2 = JZ(Z^T J^T \Sigma_x^{-1} JZ)^{-1} Z^T J^T \tag{4.14}$$

Z represents a mapping from the parametric space to the essential parameters subspace. Since J is rank-deficient, without loss of generality, Z can be chosen to

be the singular subspace of J corresponding to its non-zero singular values, which is equivalent to taking the pseudo-inverse as the following [49]:

$$\left\|\hat{X} - \bar{X}\right\|^2 = J(J^T\Sigma_x^{-1}J)^\dagger J^T \tag{4.15}$$

where $(.)^\dagger$ is the pseudo-inverse operator.

This equation shows the covariance matrix of the estimated world and image points. The mean of the trace of this matrix represents the root-mean-square (RMS) estimation error. In reality, the measurement noise covariance matrix is better modelled as $\Sigma_x = \text{diag}(\Sigma_{X_w}, \Sigma_{x_c})$. The RMS estimation errors of the x_c and X_w can be derived from equation 4.15 as

$$\begin{aligned} \|\hat{x_c} - \bar{x_c}\|_{RMS} &= \frac{trace\left(\begin{pmatrix} J_h & J_{X_w} \end{pmatrix}\begin{pmatrix} \Sigma_{\hat{h}} & -\Sigma_a W V^{-1} \\ -V^{-T}W^T\Sigma_a^T & \Sigma_{\hat{X_w}} \end{pmatrix}\begin{pmatrix} J_h^T \\ J_{X_w}^T \end{pmatrix}\right)^{1/2}}{\sqrt{N_{x_c}}} \\ \left\|\hat{X_w} - \bar{X_w}\right\|_{RMS} &= \frac{trace\left(V^{-T}W^T\Sigma_a W V^{-1} + V^{-1}\right)^{1/2}}{\sqrt{N_{X_w}}} \end{aligned} \tag{4.16}$$

where (N_{X_w}, N_{x_c}) are the number of measurements, $\Sigma_a = (U - WV^{-1}W^T)^{-1}$, $\Sigma_{\hat{h}} = U^{-1}(I + WV^{-T}W^T\Sigma_a^T)$ represents the covariance matrix of estimated h, $U = J_h^T\Sigma_{x_c}^{-1}J_h$, $V = \Sigma_{X_w}^{-1} + J_{X_w}^T\Sigma_{x_c}^{-1}J_{X_w}$, and $W = J_h^T\Sigma_{x_c}^{-1}J_{X_w}$

Consider a special case where the measurement noise of X_w and x_c are independently and identically distributed (i.i.d) with a variance σ^2. If the SVD of J is $U_J S_J V_J^T$, the RMS estimation error is then equal to

$$\left\|\hat{X} - \bar{X}\right\|_{RMS} = \frac{\sigma}{\sqrt{N}} trace\left(U_J \begin{pmatrix} I_{d\times d} & 0 \\ 0 & 0 \end{pmatrix} U_J^T\right)^{1/2} = \sigma\sqrt{\frac{d}{N}} \tag{4.17}$$

where N is the number of measurements and d represents the number of essential parameters (equal to $3n + 11$ in the camera calibration).

Given such an i.i.d noise, it can also be shown that if the estimator from the algorithm is optimal, the Pythagorean equality is reached among the estimation error, the residual error, and the measurement error [23].

$$\left\|X - \bar{X}\right\|_{RMS}^2 = \left\|X - \hat{X}\right\|_{RMS}^2 + \left\|\hat{X} - \bar{X}\right\|_{RMS}^2 \tag{4.18}$$

4.3.4.2 Point Transfer Error

The point transfer error results in an uncertainty of the homography H whose magnitude is equal to $\Sigma_{\hat{h}}$ in the previous equation.

If a set of new 3D points are measured and further transferred to estimate the image points using the estimated H, the transfer error results from the estimation error from the calibration and the measurement error of the new 3D points. Since the new 3D points and the estimated H are uncorrelated, the uncertainty of the transferred image points have a covariance matrix as

$$\Sigma_{x_c'} = J_{h'} \Sigma_{\hat{h}} J_{h'}^T + J_{X_w'} \Sigma_{X_w'} J_{X_w'}^T \tag{4.19}$$

where $(J_{h'}, J_{X_w'}) = (\frac{\partial f}{\partial h}|_{h=\bar{h}, X_w = X_w'}, \frac{\partial f}{\partial X_w}|_{h=\bar{h}, X_w = X_w'})$ and $X_w^{'}$ represents the set of new 3D points.

4.4 Two-axis Gimbal Model

Gibson compared the mechanical gimbal with some non-mechanical techniques in his surveys [21], where he pointed out the former has the advantage in its large

steering range and the disadvantages in its physical size, weight and power consumption. We prefer a two-axis rotary mechanical gimbal for three reasons: (1) no other economic alternative is capable of providing a steering range up to 4π steradians, (2) the recent progress in miniature motors, such as brushless pancake motors, greatly reduced the weight and physical size, and (3) the transceiver's ray vector and the 3D point illuminated by the ray in a rotary gimbal follow the perspective projection.

Let the optical axis of the FSO transmitter be the *axis of propagation* (AOP). If it is designed to have a minimum displacement with respect to the pivoting point of the gimbal, the AOP in the transceiver and gimbal frame can be simply related by

$$\begin{pmatrix} x'^1 \\ x'^2 \\ 1 \end{pmatrix} = \lambda_a R_a \begin{pmatrix} x_g^1 \\ x_g^2 \\ x_g^3 \end{pmatrix}$$

where R_a represents the mapping from the gimbal frame to transceiver, $\begin{pmatrix} x_g^1 & x_g^2 & x_g^3 \end{pmatrix}^T$ represents the AOP in the gimbal frame and λ_a is a scaling factor for normalization.

A 3D point aimed at by the transceiver is expressed in the gimbal frame as

$$\begin{pmatrix} x_g^1 \\ x_g^2 \\ x_g^3 \end{pmatrix} = \lambda_g (R_g | T_g) \begin{pmatrix} X^1 \\ X^2 \\ X^3 \\ 1 \end{pmatrix}$$

Combining the two equations above, the vector of the AOP and its target 3D

point can be expressed as

$$\begin{pmatrix} x'^1 \\ x'^2 \\ 1 \end{pmatrix} = \lambda_3(R_a R_g | R_a T_g) \begin{pmatrix} X^1 \\ X^2 \\ X^3 \\ 1 \end{pmatrix} = \lambda_3 P' \begin{pmatrix} X^1 \\ X^2 \\ X^3 \\ 1 \end{pmatrix} \tag{4.20}$$

If we select the gimbal frame as the one shown in figure 4.3, the rotation matrix that maps from the gimbal to the transceiver frame can be modelled as:

$$R_a = \begin{pmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ \sin\phi & 0 & \cos\phi \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \tag{4.21}$$

where ϕ : the azimuthal angle and θ : the elevation angle.

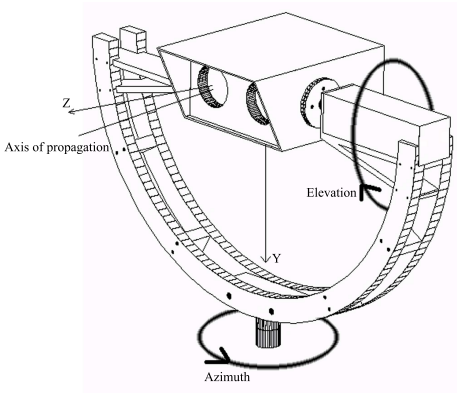

Figure 4.3: Two-axis rotary gimbal diagram

4.5 Radial Trifocal Tensor

4.5.1 Geometry

The radial trifocal tensor (RTT) was first introduced by Thirthala et.al [63] to calibrate wide-angle cameras. We proposed to apply that for expanding the field-of-view of a 3D CPAT system [27]. Before introducing it, we must understand the imaging model of a wide-angle camera. Compared to a perspective camera, since the distortion lies symmetrically in the radial direction from the centroid of the image plane, a wide-angle camera only preserves ϕ but not θ. Figure 4.4 shows the concept, and the imaging model can be formulated as

$$\begin{pmatrix} x_w^1 - o_x \\ x_w^2 - o_y \end{pmatrix} = \begin{pmatrix} x'^1 \\ x'^2 \end{pmatrix} = \lambda \begin{pmatrix} 1 & s \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} R^1 & T^1 \\ R^2 & T^2 \end{pmatrix} X_w \tag{4.22}$$

where (x_w^1, x_w^2) represents the pixel coordinates on the image plane.

Physically, an image point in the wide-angle camera only represents a line perpendicular to the optical axis, which contains only half of the information compared to the perspective camera and such a comparison can be made by observing figures 4.1 and 4.4. Combining equations 4.4, 4.20, and 4.22, the geometry among the cameras and the transceiver can be formulated as

$$\begin{pmatrix} P_{3\times4} & x & 0 & 0 \\ P'_{3\times4} & 0 & x' & 0 \\ P''_{2\times4} & 0 & 0 & x'' \end{pmatrix} \begin{pmatrix} X_w \\ -\lambda_1 \\ -\lambda_2 \\ -\lambda_3 \end{pmatrix} = 0_{8\times1} \tag{4.23}$$

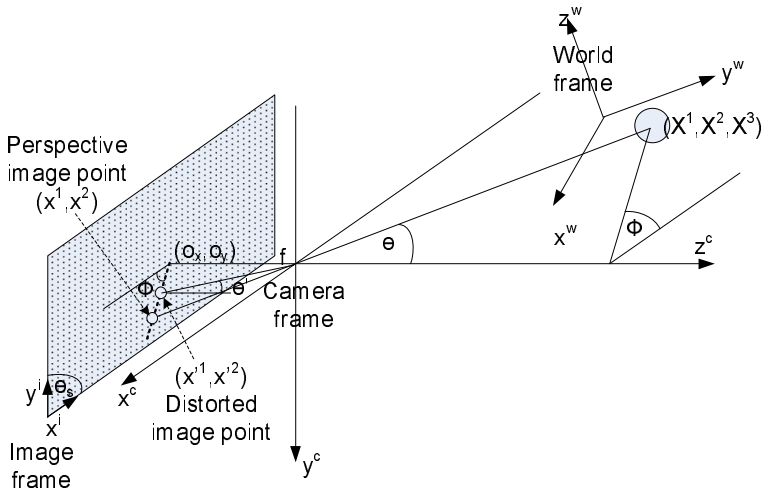

Figure 4.4: Wide-angle camera imaging model: (x^1, x^2) represents the perspective imaging point and (x'^1, x'^2) represents the distorted image point. The bold-dashed line shows the geometric information revealed by an image point.

where $x"$ is the image position in the wide-angle camera subtracting the principal point.

Because of the non-zero null vector, the minor matrix can be further expanded as

$$-\frac{1}{2}x^i x'^j x"^k \epsilon_{jqu}\epsilon_{ilm}\eta_{kr} \det \begin{pmatrix} P_l \\ P_m \\ P'_q \\ P"_r \end{pmatrix} = 0_u \tag{4.24}$$

where $\epsilon_{abc} = \begin{cases} 0, & \text{a,b, and c are not distinct;} \\ 1, & \text{if abc is an even permutation of 123;} \\ -1, & \text{if abc is an odd permutation of 123.} \end{cases}$,

$$\eta_{ab} = \begin{cases} 0, & \text{a and b are identical;} \\ 1, & \text{if ab is 12;} \\ -1, & \text{if ab is 21.} \end{cases} .$$

The radial trifocal tensor (T_{ik}^{q}) is further defined as

$$T_{ik}^{q} = -\frac{1}{2}\epsilon_{ilm}\eta_{kr} \det \begin{pmatrix} P_l \\ P_m \\ P_q^{'} \\ P_r^{"} \end{pmatrix} \tag{4.25}$$

The RTT defined in equation 4.25 is a $3 \times 3 \times 2$ matrix consisting of 18 elements but only 17 degrees of freedom because of an unknown scaling factor. Although there are three linear constraints provided by each triplet of the correspondences for solving the radial trifocal tensor, only two of them are independent. Therefore, 9 triplets are needed to determine an RTT from the previous equation.

4.5.2 Internal Constraints

According to equation 4.25, a RTT depends on the projective matrices ($P, P^{'}, P^{"}$) of each device. Since P and $P^{'}$ each have 11 degrees of freedom (dof) and $P^{"}$ has 7 dof, the number of overall dof is 29. We have the freedom to select an arbitrary world frame, which eliminates 15 dof and leaves only 14 dof. A RTT is represented by 17 parameters, which is over-parameterized. Therefore, 3 embedded constraints, known as the internal constraints, should be imposed on the RTT estimation.

Without loss of generality, we first select the world frame to be identical to the frame of the perspective camera ($P = (I_{3\times3}|0_{3\times1})$) to eliminate 11 dof. Under such

a selection, the origin of the perspective camera can be represented by the fourth column of the two respective projective matrices $(P_4^{'}, P_4^{"})$ in the transceiver frame and the wide-angle camera frame, also known as the epipole and distorted-epipole $(e^{'}, e^{"})$. Equation 4.25 can be expanded as:

$$\begin{aligned} T_{ik}^{q} &= P_i^{'q} P_4^{"\sim k} - P_4^{'q} P_i^{"\sim k} \\ &= P_i^{'q} L_k^{"} - e^{'q} P_i^{"\sim k} \end{aligned} \tag{4.26}$$

where $P_i^{"\sim k}$ denotes omission of the k^{th} element of $P_i^{"}$ and $L^{"}$ is the line vector passing through the distorted-epipole $e^{"}$.

Definition 4.5.1. If a 3D object is within the field-of-view of the perspective camera, (1) its scattered light ray projected to the perspective image frame (x), (2) the transceiver's laser ray pointing at it $(x^{'})$, and (3) the epipole $(e^{'})$, form a plane, known as the epipolar plane. This coplanar property is denoted as the epipolar constraint.

Definition 4.5.2. The plane consisting of the optical axis of the wide-angle camera and the origin of the perspective camera is denoted as the *axial plane*, whose plane equation depends on the setup of the two cameras.

Theorem 4.5.1. *If a 3D object is neither on the axial plane nor on the line connecting the gimbal's pivoting point and the perspective camera's origin, the matrix* $T_{ik}^{q} x^{i}$ *is rank-2 and its left null vector is the normal vector of the epipolar plane between the perspective camera and the transceiver. Figure 4.5 (a) demonstrates this concept.*

Proof. According to equation 4.26,

$$\begin{aligned} T_{ik}^{q} x^{i} &= x^{i} P_{i}^{'q} L^{"} - e^{'} P_{i}^{"\sim k} x^{i} \\ &= x^{'q} L_{k}^{"} - e^{'q} l_{k}^{"} \end{aligned}$$

where $l^{"}$ is a line vector in 1D passing through the distorted image point ($x^{"}$).

Since $T_{ik}^{q} x^{i}$ is a 3×2 matrix, its maximum rank is 2. Also, because the object is neither on the axial plane nor on the translation line between the gimbal and the perspective camera, the resulting image points $x^{'}$ and $l^{"}$ are not parallel to $e^{'}$ and $L^{"}$, respectively. Let the singular vectors of $x'qL_{k}^{"}$ be a_1 and b_1 and the values are σ_1. Without loss of generality, we can expand $e^{'}$ and $x^{"}$ as $\alpha_1 a_1 + \beta_1 a_1^{\perp}$ and $\alpha_2 b_1 + \beta_2 b_1^{\perp}$. Equation 4.26 can be shown as a rank 2 matrix and can be expressed as

$$T_{ik}^{q} x^{i} = \begin{pmatrix} a_1 & a_1^{\perp} & 0 \end{pmatrix} \mathrm{diag}(\sigma_1 + \alpha_1\alpha_2, \beta_1\beta_2, 0) \begin{pmatrix} b_1 & b_1^{\perp} & 0 \end{pmatrix}^{T}$$

Since its left null vector ($u^{'}$) is a zero vector, according to equation 4.26, the left null vector $u^{'}$ must be orthogonal to both $e^{'}$ and $x^{'}$, which is a normal vector of the epipolar plane. □

Theorem 4.5.2. *If a 3D object is located on either the axial plane or the line between the gimbal and the perspective camera, $T_{ik}^{q} x^{i}$ is a rank-1 matrix and the two left null vectors are orthogonal to the epipole and the right null vector is the distorted-epipole. This concept is shown in figure 4.5.*

Proof. Given the two conditions, equation 4.26 can be rewritten as

$$\begin{aligned} T_{ik}^{q} x^{i} &= (x^{'q} - e^{'q}) L_{k}^{"} \quad ; \text{On the axial plane} \\ &= e^{'q} (L_{k}^{"} - l_{k}^{"}) \quad ; \text{On the connected line} \end{aligned}$$

Therefore, the left two null vectors are orthogonal to $e^{'}$ and the right null vector is orthogonal to $L^{"}$, which is the distorted-epipole. □

Corollary 4.5.3. *Given $T^q_{ik}x^i$ is a rank-1 matrix, any determinant of the 2×2 minor is 0. If the perspective camera and the transceiver have a depth difference, the resulting locus of such x are three degenerate conics and the intersections are a single point and a line, which are the epipole and the distorted-epipole expressed in the frame of the perspective camera; otherwise, the intersections are two lines. This provides the three internal constraints for the radial trifocal tensor.*

4.5.3 Radial Trifocal Tensor and Projective Matrices

The previous section indicated that the world frame can be arbitrary selected, which implies that the projective matrices estimation is not unique for a given RTT. It can be proven by multiplying a 4×4 matrix P_M after $(P, P^{'}, P^{"})$ and P_M^{-1} before X_w in equation 4.23.

We again first select the world frame to be the perspective image frame and $P = \begin{pmatrix} \text{diag}(1,1,1) & 0_{3\times 1} \end{pmatrix}$. The only requirement to determine that the projective matrix of the transceiver frame is the epipolar constraint in definition 4.5.1. If the world frame is the same as the perspective image frame, this constraint can be formulated as

$$x^{'}(e^{'})_{\times}P^{'}_{1\sim 3}x = 0 \tag{4.27}$$

where $(e^{'})_{\times}P^{'}_{1\sim 3}$ is denoted as the fundamental matrix.

This equation implies that if the first three columns of the projective matrix

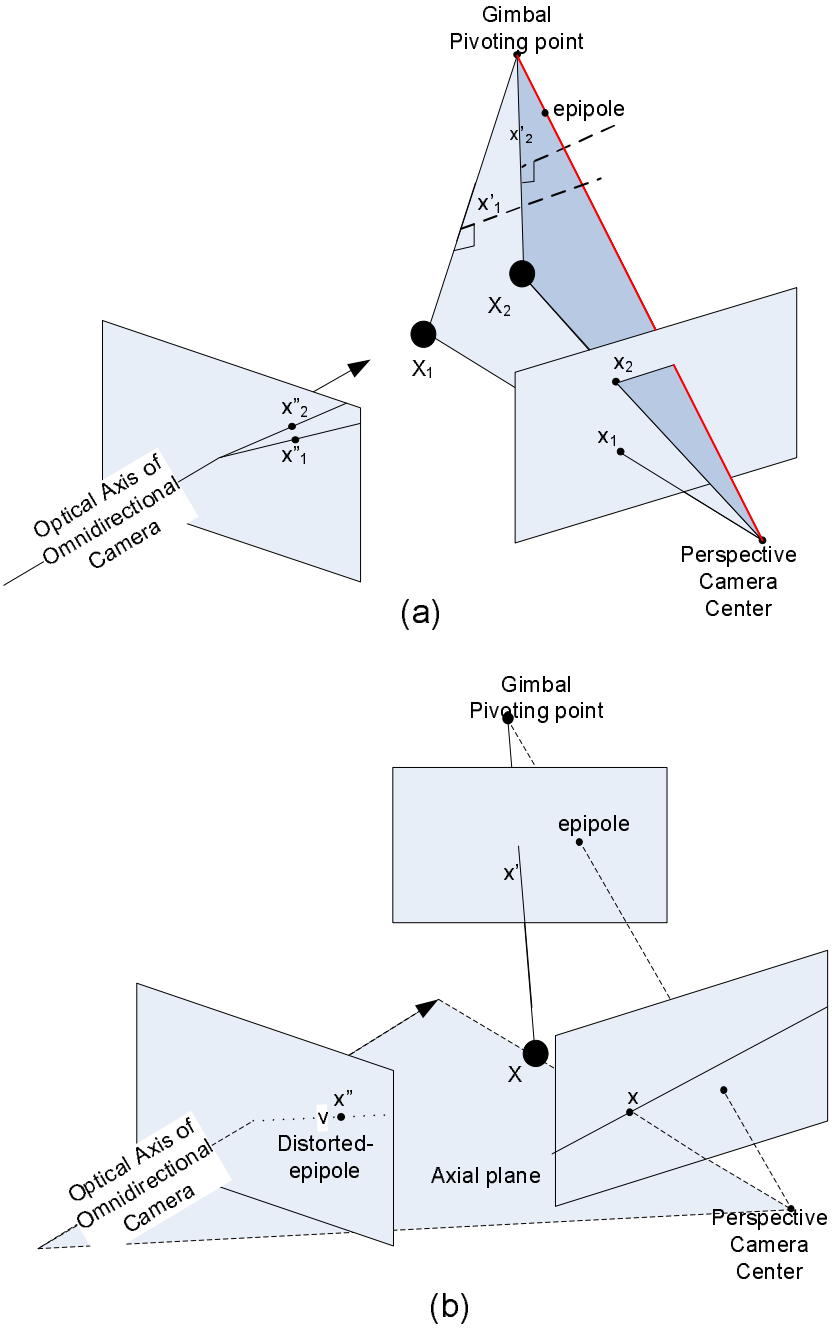

Figure 4.5: (a) The epipolar constraint between the actuator and the perspective camera and (b) the axial planar constraint among the three components

have the following form, the epipolar constraint is satisfied:

$$\hat{P}'_{1\sim3} = P'_{1\sim3} + ev^T$$

where v is any 3×1 vector.

It was shown that $e^{'}$ and $L^{"}$ can be estimated directly from the RTT. After normalizing them to a unit norm and multiplying by $L^{"}$ on both sides of equation 4.26, a qualified projective matrix to satisfy the epipolar constraint between the perspective image and the transceiver frame is obtained as

$$\begin{pmatrix} T^q_{ik} L^{"k} & e^{'} \end{pmatrix}$$

By replacing the selected $P^{'}$ back into equation 4.26, we can obtain the projective matrix of the wide-angle image frame as

$$P^{"} = \begin{pmatrix} P^{"}_1 & P^{"}_2 & P^{"}_3 & e^{"} \end{pmatrix}$$

where $P^{"}_i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} (L^{"} L^{"T} - \mathrm{diag}(1,1)) T_i e^{'}$

This section provided the necessary geometrical constraints for the RTT. In most applications, the RTT must be estimated in advance through a calibration process, such as combining equation 4.24 with least squares procedures. However, because of the measurement noise and the lens aberrations, the estimation is usually erroneous. A sophisticated estimation procedure which incorporates the internal constraints can greatly improve the estimation accuracy. The details are further revealed in section 5.4.

Chapter 5

Design and Analysis of Coarse Angular Pointing, Acquisition, and Tracking Systems

A coarse pointing acquisition and tracking system (CPAT) is a sub-module used to narrow down the range of uncertainty for the further FPAT system. Therefore, its pointing accuracy is determined by the FOV of the FPAT system. Generally, the CPAT system must have (1) a wide sensing FOV to acquire multiple targets simultaneously, (2) a large steering range to move transceivers, and (3) an autonomous methodology to acquire the target and steer the transceiver toward the selected target.

The earliest FSO CPAT system can be dated back to the Gemini 7 space mission in 1965 [52], where the CPAT system was operated by users with a sighting telescope and mechanical steering stages. In 1992, NASA started a series of ground-to-orbit FSO experiments such as the Galieo and GOLD programs [41, 31], whose CPAT systems relied on the precise ephemeris and altitude of the satellite. For most of the commercial ground-to-ground FSO systems such as Canobeam, MRV, and Lightpointe, their CPAT systems are almost identical to the one used in Gemini 7. The slow progress of the CPAT part of the FSO systems lies in that traditional FSO links are mostly static and therefore the alignment does not have to be updated dynamically.

Breakthroughs in sensors and actuators led to several successful demonstrations of mobile FSO applications, which further drove the demand of autonomous CPAT systems. Saw et.al [57] and Epple et.al [16] proposed a GPS aided inertial sensing system (GPS-ISS) hybrid CPAT system. The MUSIC and ESPRIT algorithms introduced by Manolakis [44] have been widely applied to RF radar systems for missile tracking, which is another alternative of the CPAT system. Gibson [21] suggested a camera-based CPAT system with steering mirrors, a similar idea to the FSM-based FPAT system. Most of the camera-based CPAT systems still follow the telescope concept, described in section 1.2.3.1, where the alignment is completed if the image of the target overlaps with a pre-calibrated pixel. However, this kind of system does not only reduce the acquisition speed but also complicates the target identification process since the entire image varies because of the camera or mirror motion.

Besides aligning mobile targets, CPAT systems can also serve as catalyst to upgrade the FSO communication systems from point-to-point links to a fully connected network [62]. Compared to multiple PtoP links connecting randomly with each other, a connected network often select the best links from all the possible routes to optimize some designated cost functions, such as the packet delay time or throughput rate. In a directional FSO network, the wide FOV of the CPAT system is the only access to the information about possible routes.

Our work emphasizes camera-based CPAT systems consisting of single or multiple cameras for target identification and a two-axis rotary gimbal for target pointing. Compared to the traditional ones, the camera is static with a large field-of-view,

close to 2π steradians. A static camera can easily detect mobile targets by subtracting the static background, which is beneficial to mobile FSO systems. Also, since the rotation angles for the gimbal are determined once the image is retrieved, it greatly increases the acquisition speed. Figure 5.1 shows a camera-based CPAT system.

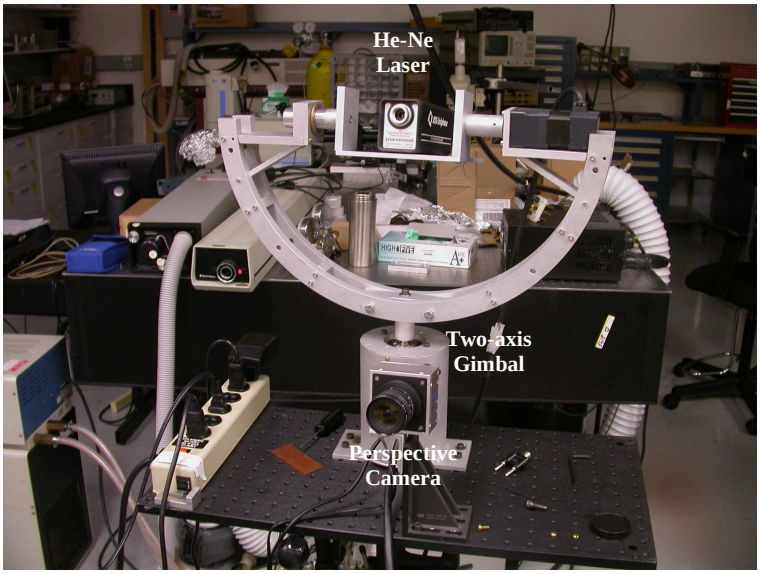

Figure 5.1: A camera-based coarse PAT system

The design of camera-based CPAT systems can be divided into two parts:

- Target identification, where the features or patterns of the target must be differentiated from a complicated background quickly and accurately. In FSO applications, target identification is relatively simple because (1) the background is mostly clean and (2) users can add unique features to the transceivers before the real deployment.

- Target pointing, where the pixel information acquired from the camera must be converted to assist the transceiver equipped with an FPAT system to point

to the target correctly. The challenge is to find the correct transformation since the sensing cameras are mostly installed outside the transceiver, which results in non-reciprocal receiving and transmitting rays between the remote and the local transceiver. This chapter is dedicated to solving this problem.

We have discussed the camera imaging model in section 4.2 and revealed that the camera preserves two dimensional spatial information (azimuthal and elevation angles), of the targets but not their distance. The distance information plays an important role in the CPAT system only if the following inequality is not satisfied:

$$\|L\| > \frac{\|T\| \sin\left(\theta_{F_{fov}} + \theta_{X2cam}\right)}{\sin \theta_{F_{fov}}} \tag{5.1}$$

where $\|L\|$ represents the link length (equal to the depth of the target), $\|T\|$ represents the translation distance between the camera and the transceiver, $\theta_{F_{fov}}$ represents the field-of-view of the FPAT systems and $\theta_{X2cam} = \arccos \frac{XT}{\|X\|\|T\|}$ is the angle between the vector from the transceiver to the target (X) and the vector from the transceiver to the camera.

This inequality is almost identical to the inequality 1.1. It implies that when the link length is far enough, the alignment of the CPAT system is a coupled alignment problem and can be solved by one camera as the FPAT system; otherwise, it is a singled alignment problem and therefore needs two cameras to determine the depth.

For example, if the FPAT system has a field-of-view 2° and the translation between the camera and the transceiver is 1 meter, only when the target has a non-planar motion and the motion takes place within 28.7 meters from the host is the

distance information of the target required.

This example shows that the distance information is not necessary in most of the FSO applications since the link length L is normally greater than the requirement or the motion of the target is on a plane. The CPAT system used in these two scenarios is denoted as a planar coarse pointing system. The specialty of such a system is that only one camera is required in the design. The details are discussed in the following section.

5.1 Planar Coarse Pointing Systems

In many of the FSO applications, the target stays a long distance away from the host or tends to move on a planar surface, where a single camera is sufficient to point to the target correctly. Figure 5.2 shows some of these applications. Compared to the general CPAT system, the mapping of the planar one is simpler and the mapping parameters are fewer, which usually leads to a better mapping estimation.

5.1.1 Coarse Pointing Systems For Planar Motion

5.1.1.1 Geometry

Let the normal vector of the plane be N and the plane equation be $N^T X = d$, where d represents the distance from the point X to the origin of the world frame.

Without loss of generality, the world frame is selected to be the transceiver frame, where ($x^{'} = X$). Combining the gimbal model in equation 4.21 and the perspective imaging model of the camera in equation 4.5, the geometry between the

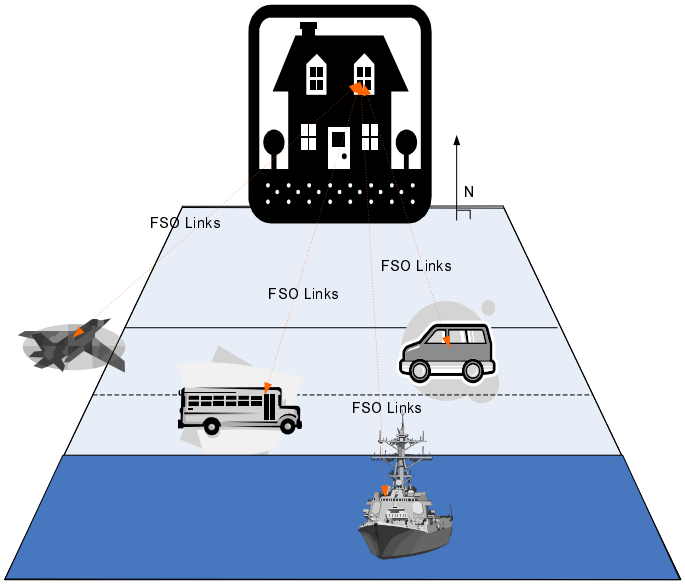

Figure 5.2: FSO applications for planar CPAT systems

vectors in the transceiver frame and the perspective image frame can therefore be formulated as

$$\begin{pmatrix} x^1 \\ x^2 \\ 1 \end{pmatrix} = \lambda M(R_{t2c} + \frac{1}{d}T_{t2c}N^T) \begin{pmatrix} x'^1 \\ x'^2 \\ 1 \end{pmatrix} \tag{5.2}$$

where M represents the intrinsic matrix and $(R_{t2c}|T_{t2c})$ represents the transformation from the transceiver to camera frame.

The previous equation shows that the ray vectors transmitting from the transceiver and received by the camera are related by a linear homography, which can be calibrated in advance.

5.1.1.2 Calibration/Pointing Algorithm

The calibration algorithm, depending on the cost function, is divided into two steps (1) algebraic minimization (AMS), which optimizes the parallelism between the position vectors and (2) geometric minimization (GMS), which optimizes the Mahalanobis distances.

Algorithm 5.1 Homography Estimation Algorithm

Require: Rotate the two-axis gimbal with 4 sets of different angles and record the laser ray vectors using equation 4.21 and corresponding image point in the camera.

Ensure: The 3D points illuminated by the laser must be coplanar.

1: Normalize the vector sets from each frame to a mean position at the origin and an average distance $\sqrt{2}$ with respect to the mean position.

2: Solve for the vectorized homography using the two vector sets according to algorithm A.1

$$\min_{h} \left\| \begin{pmatrix} 0^T & -x' & x^2 x' \\ x' & 0^T & -x^1 x' \end{pmatrix} h \right\| \quad \text{subject to } \|h\| = 1$$

where h is the vectorized $M(R_{t2c} + \frac{1}{d} T_{t2c} N^T)$.

3: Apply the estimated homography to compute the corresponding normalized camera vectors $\hat{x}$ and minimize the Mahalanobis distances between x and $\hat{x}$ with the Levenberg-Marquardt algorithm.

4: Denormalize the estimated homography to retrieve the optimal one.

Algorithm 5.2 Point Transfer Algorithm

1: Select the target image from the camera and use the coordinates to compute the corresponding laser vector.

2: Apply equation 4.21 to retrieve the rotation angles from the computed laser vector.

3: Update the homography according to the rotation angles after the pointing.

$$H^{i+1} = H^i (R^i)^T$$

where H^i represents the homography before i^{th} rotation and R^i is the rotation matrix applied at the i^{th} rotation.

5.1.1.3 Performance Evaluation

Three simulations were conducted to test the proposed algorithm. The first one is to show the estimation improvement from more corresponding vectors, the second one is to verify the robustness of the algorithm subject to different amount of measurement noise, and the last one demonstrates the expected pointing errors by using the proposed algorithm. In the current CPAT system, the corresponding vector pairs are provided by the two-axis gimbal and the camera. The resolution of the gimbal, $\pm 0.0036^\circ$, is limited by the optical encoders of the motor, which is negligible compared to the aberration error resulting from the lens. Therefore, in the first two simulations, the vectors obtained from the gimbal are assumed to have no additive random errors, where the number of the essential parameters is 8 including only the homography. According to equations 4.17 and 4.18, the theoretical estimation and residual RMS errors are

$$\begin{aligned} \left\| \hat{X} - \bar{X} \right\|_{RMS} &= \sigma \sqrt{\tfrac{4}{n}} \\ \left\| X - \hat{X} \right\|_{RMS} &= \sigma \sqrt{\tfrac{N-4}{n}} \end{aligned}$$

where n is the number of corresponding vectors.

The geometric setup was identical in the three simulations as follows: the camera and the gimbal were placed 1 m apart and the AOP of the transceiver and the optical axis of the camera were made parallel. The plane was placed 10 m away from the transceiver and perpendicular to the AOP. The camera was set to have a semi-fov of 10° and a resolution of 640×480 pixels.

Simulation 5.1: Error Reduction Simulation

The first task was to investigate the estimation performance when more correspond-

ing pairs were used. In the simulation, the number of correspondences increased from 4 to 29 with a spacing of 5. The image points were perturbed by an isotropic Gaussian noise with mean 0 and variance 1. For each number of correspondences, 100 tests were executed and the RMS of estimation and residual errors were computed. The results are shown in figure 5.3

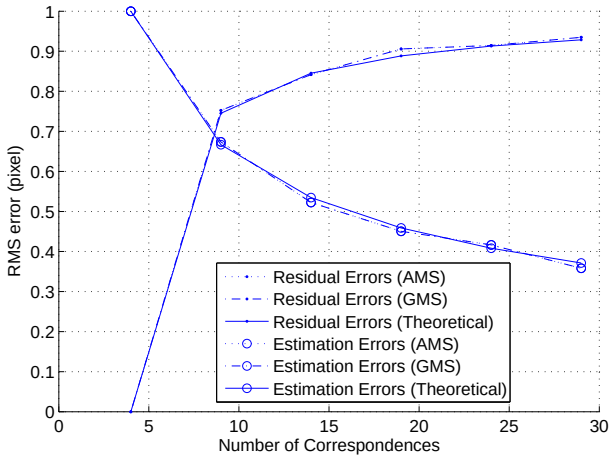

Figure 5.3: RMS Error performance with respect to different number of corresponding vector pairs. The reported RMS error is an average from 100 different independent trials.

The results show two important facts: (1) the estimation errors are reduced linearly when more corresponding vector pairs are applied in the algorithm, which implies the estimated homography is closer to the true homography. This improvement results from the independent measurements reducing the estimation uncertainty. (2) The estimation accuracy is almost identical between the AMS and combined (AMS+GMS) algorithms, which indicates that the linear AMS algorithm is sufficient to the current system. Implementing the AMS algorithm in a microproces-

sor is much cheaper than the combined algorithm.

Simulation 5.2: Robustness Simulation

The next simulation was to investigate the estimation performance with increasing noise. The number of corresponding vector pairs was limited to 14. The standard deviation of Gaussian noises used in the camera varied from 1 to 5. For each noise magnitude, 100 tests were applied to find the RMS of the estimation and residual errors, which are plotted in figure 5.4.

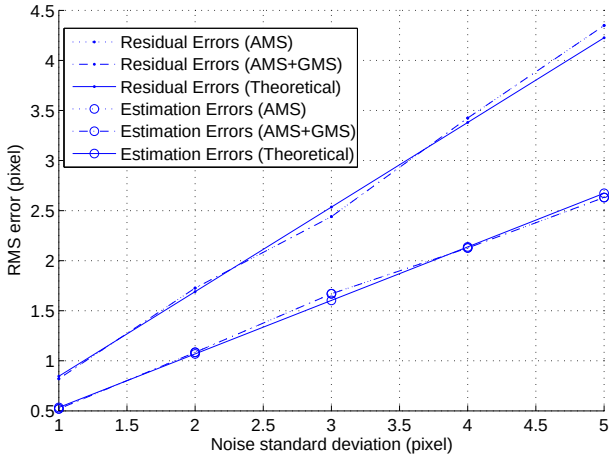

Figure 5.4: RMS Error performance with respect to different amounts of noise. The reported RMS error is an average from 100 different independent trials.

The results show the RMS estimation and the residual error increase linearly with the noise magnitude, which indicates the robustness of both algorithms with respect to the noise. Also, even with an increase of the noise, the AMS and the combined algorithms still perform almost identically.

Simulation 5.3: Pointing Error Simulation

As mentioned in section 4.3.4, the accuracy of the transferred points depends on

Position	Theoretical RMS Errors (mrad)		Simulated RMS Errors (mrad)	
	X	Y	X	Y
(640,1)	0.6023	0.7829	0.6326	0.8092
(1,1)	0.6618	0.8632	0.6646	0.8627
(640,480)	0.6562	0.8412	0.6854	0.8212
(1,480)	0.6317	0.8174	0.6228	0.8390
(320,240)	0.5285	0.6962	0.5341	0.7261

Table 5.1: Theoretical and Simulated Pointing RMS Errors. The reported pointing error is an average from 100 different independent trials.

the position of the new selected points and the covariance matrices of the estimated homography and the measurement noise formulated in equation 4.19. This accuracy determines the pointing resolution of the planar CPAT system.

A Monte-Carlo simulation was conducted to investigate the theoretical model of the pointing resolution. In this simulation, we assumed the measurements from both the gimbal and the camera were perturbed by Gaussian noises: $N(0, 1)$ to the camera and $N(0, 0.0036)$ to the gimbal.

We selected five points in the image frame including four corners and the centroid to analyze their transferred accuracy. The selected points were perturbed by 1 pixel as well. 100 trials were applied to find the RMS pointing errors. The results are displayed in table 5.1.

The results between the Monte-Carlo simulation and the theoretical model are closely matched, which is another indicator that the proposed algorithm produces an optimal estimation. The average pointing errors are around $0.6 \sim 0.8$ mrad. Notice that the measurement noise from the new selected image point is responsible

for most of the pointing error compared to the noise introduced by the estimated homography.

Experiment 5.1: Planar CPAT Systems

A CPAT experiment was conducted in the laboratory for verifying the pointing performance. The planar CPAT system includes a two-axis stepping-motor gimbal from Oriental-Motor Inc., a camera from Imperx (IPX-VGA210), and a CCTV lens from Pentax. Both stepping motors, AS46AA-H50 and AS66AA-H50, are operated in a closed-loop with the angular information fed from an optical encoder with a resolution of 0.0072°. The resulting camera has a horizontal FOV of 31.8°, a CCD resolution of 640×480, and a frame rate of 210 fps.

The CPAT system setup is shown in figure 5.1. The separation between the camera and the gimbal was 0.7 meter and the planar wall was 7 meters in front of the camera. 15 corresponding vector pairs from the gimbal and the camera were uniformly selected in the scene for homography estimation. Another 12 points were further selected for pointing verification. The camera scene and the selected points (with cross and start marks) are plotted in figure 5.5.

For the homography estimation experiment, the average residual errors between the estimated points and the measurements in X and Y axis of the gimbal were around 0.4156 and 0.4311 mrad, respectively. This amount of resolution was even sufficient for the FPAT system. As for the pointing experiment, the error was observed by measuring the distance between the selected image point and the laser spot in the camera. The average pointing errors were about 0.5824 and 0.5766 pixels, corresponding to 0.506 and 0.666 mrad.

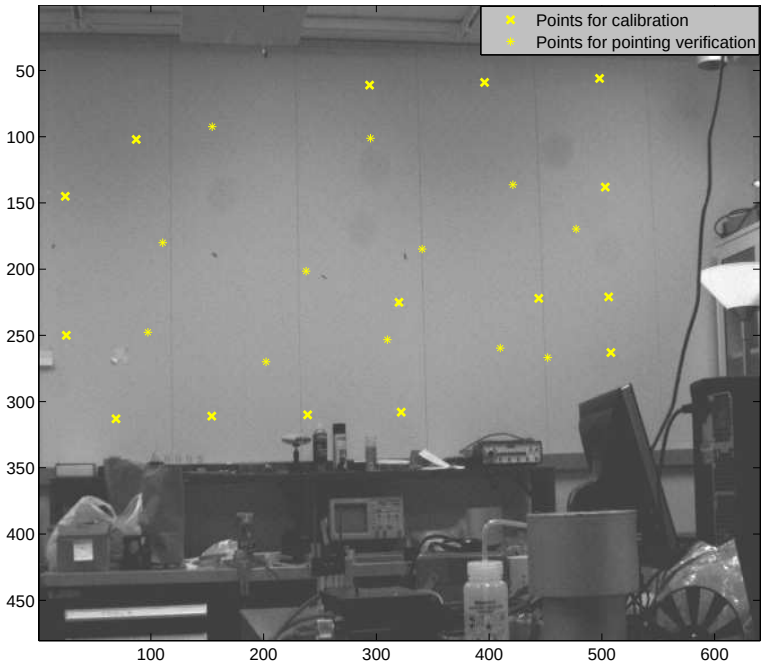

Figure 5.5: Experimental scene of homography estimation

5.1.2 Planar Coarse Pointing Systems For Long Link Distances

The previous results are applicable when the motion of the target is limited on a plane. If the link length is long enough, since the translation between the transceiver and the camera can be neglected, equation 5.2 can be cast as

$$\begin{pmatrix} x^1 \\ x^2 \\ 1 \end{pmatrix} = \lambda M R_{t2c} \begin{pmatrix} x'^1 \\ x'^2 \\ 1 \end{pmatrix} \tag{5.3}$$

The equation shows that the mapping between two frames for targets with a long link distance is equal to the product of the intrinsic and rotation matrix from the transceiver to the camera. If a plane with a large distance exists and the scattering of the laser on that plane is still visible by the camera, algorithm 5.1

is still applicable. If not, we introduce other algorithms to estimate the mapping MR_{t2c}.

5.1.2.1 Intrinsic Matrix Estimation

The intrinsic matrix must be known to transform the image frame to the camera frame in order to estimate the rotation matrix. The concept of the intrinsic matrix was discussed in section 4.2 and many algorithms have been proposed for it. These algorithms are mainly divided into two categories: on-line and off-line. On-line algorithms rely on multiple views and scene points to estimate the absolute conic [23], which is invariant under projective transformations. The absolute conic is equal to the $(MM^T)^{-1}$, where M is the intrinsic matrix. Triggs [65] proposed to estimate the absolute dual quadric from different views and the image of the this quadric is the absolute conic. The quadric is a surface in $\mathbb{P}^3$ defined by the equation $X^TQX = 0$, where Q is a symmetric 4×4 matrix. Note that each view contains five quadratic equations to solve for the absolute quadric.

Compared to the on-line algorithm, off-line algorithms produce more accurate and stable solutions since the applied constraints are linear. Off-line algorithms utilize a calibration object in a pre-determined frame. Tsai [69] first introduced a calibration algorithm where a cube with checkerboard pattern served as the calibration object. Zhang [81] proposed an improved algorithm which only requires a planar checkerboard pattern. It utilized the additional constraints imposed from the planar view and three views were required to fully calibrate the intrinsic matrix.

Because of its simplify and robustness, this algorithm has been widely accepted and a camera calibration toolbox based on such algorithm can be downloaded from Caltech's vision laboratory [32].

5.1.2.2 Rotation Matrix Estimation from Homography

Let the intrinsic matrix be pre-calibrated and then the problem lies in retrieving the rotation matrix. Ma et.al. [43] introduced a decomposition algorithm based on the homography with the form as ($H = \lambda(R + \frac{1}{d}TN^T)$). This algorithm derives the eigenvectors of the rotation matrix by its length-preservation property. The details of the algorithm were presented in algorithm 5.3 and the related proofs are shown in [26].

5.1.2.3 Rotation Matrix Estimation Through Essential Matrix

The other algorithm to retrieve the rotation matrix is through the essential matrix generated from the coplanar constraint discussed in section 4.5.3. Since the intrinsic matrix is assumed to be known, the mapping from the camera frame to the transceiver frame is then $(R_{c2t}|T_{c2t})$. Thus, the coplanar constraint can be formulated as

$$x^{'}(T_{c2t})_{\times}R_{c2t}x_c = 0 \tag{5.4}$$

where $x_c = M^{-1}x$ represents the incident ray vector expressed in the camera frame.

The matrix $(T_{c2t})_{\times}R_{c2t}$ is known as the essential matrix, which is a 3×3 matrix. Since it consists of only a rotation matrix and a translation vector, the

dof of the essential matrix are only 5. Therefore, 3 internal constraints have to be imposed when estimating the essential matrix from equation 5.4. The three internal constraints are: (1) the determinant is 0 and (2) two out of the three singular values of the essential matrix are equivalent. Wang et.al [72] proposed an essential matrix estimation algorithm by enforcing the equivalence of the two singular values in the optimization process. The corresponding pairs used in this algorithm are the ray vectors in the transceiver frame and the image positions in the camera frame, which are computed by converting the the measured position in the image frame with the intrinsic matrix.

We also experimented with the fundamental matrix estimation algorithm proposed by Hartley [23]. The fundamental matrix is defined in equation 4.27. We first estimated the fundamental matrix ($\hat{F}$) by applying the image points and the transceiver vectors as the corresponding pairs. Then, the essential matrix was computed as $\hat{E} = \hat{F}M$ since the image positions in the camera and image frame are related by an intrinsic matrix (M). Because of the measurement noise, the two singular values of $\hat{E}$ may not be equivalent. The closest essential matrix (E) in the Frobenius norm is

$$E = U\text{diag}(\frac{\sigma_1 + \sigma_2}{2}, \frac{\sigma_1 + \sigma_2}{2}, 0)V^T$$

where the SVD of $\hat{E}$ is $U\text{diag}(\sigma_1, \sigma_2, 0)V^T$.

Recall that $E = (T)_\times R$. Since $(T)_\times$ is skew-symmetric, any skew-symmetric matrix can be decomposed as $U_1\text{diag}(1, -1, 0)WU_1^T$, where U_1 is an unitary matrix.

If the SVD of the essential matrix is USV^T, the following equations can be derived:

$$\begin{aligned} E &= U\text{diag}(1,-1,0)V^T \\ (T)_\times R &= U_1\text{diag}(1,-1,0)WU_1^T R \end{aligned} \tag{5.5}$$

where $W = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ is an unitary matrix and $\text{diag}(1,-1,0) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

Since the singular values are equivalent, we can conclude that $U_1 = U$ and $V = R^TUW$ or $-R^TUW$. Then, R and T can be best cast as:

$$\begin{aligned} R &= UWV^T \text{ or } UW^TV^T \\ T &= U\text{diag}(1,-1,0)WU^T \text{ or } U\text{diag}(1,-1,0)W^TU^T \end{aligned} \tag{5.6}$$

Because of the repeated singular vectors, the estimated rotation matrix and translation vector have two solutions each, resulting in four different sets of solutions. Since the intersection of the two rays must lie in front of both the camera and the transceiver, this positive depth constraint helps us to select the correct solution among the four candidates. Equation 5.31 can be used to solve the intersection.

Simulation 5.4: Rotation Matrix Estimation Error

A Monte-Carlo simulation was conducted to verify the RMS estimation error of the rotation matrix among three different algorithms, which included (1) Ma's homography decomposition algorithm (denoted as H-based), (2) Hartley's algorithm (denoted as F-based) based on the fundamental matrix, and (3) Wang's algorithm based on the essential matrix (denoted as E-based).

In the simulation, the camera was set to have a translation $(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ away from the transceiver and there was no relative rotation between the transceiver and

Algorithm 5.3 Algorithm to Estimate the Rotation Matrix from Homography

Require: An estimated homography $\hat{H} = \lambda(R_{t2c} + \frac{1}{d}T_{t2c}N^T)$.

1: Decompose $\hat{H}$ using SVD and normalize $\hat{H}$ to have its second singular value 1. Let the SVD of the normalized $\hat{H}$ be $Udiag(\sigma_1, 1, \sigma_3)V^T$, where $\sigma_1 > \sigma_3$

2: Compute the vector

$$l_1 = \alpha_1 V_1 + \alpha_2 V_3$$

where V_i is the i^{th} column of V and $(\alpha_1, \alpha_3) = (\pm\sqrt{\frac{1-\sigma_3^2}{\sigma_1^2-\sigma_3^2}}, \pm\sqrt{\frac{\sigma_1^2-1}{\sigma_1^2-\sigma_3^2}})$

3: Let the cross product of V_2 and l_1 be l_3 and that of the $\hat{H}V_2$ and $\hat{H}l_1$ be L.The estimated rotation matrix is

$$\hat{R}_{t2c} = \begin{pmatrix} \hat{H}l_1 & \hat{H}V_2 & L \end{pmatrix} * \begin{pmatrix} l_1 & V_2 & l_3 \end{pmatrix}^T$$

4: The N and $\frac{1}{d}T_{t2c}$ are l_3 and $(\hat{H} - \hat{R}_{t2c})l_3$, respectively.

5: Estimate the intersection of the rays by solving equation 5.31. Among the four pairs of $\hat{R}_{t2c}$ and $\frac{1}{d}T_{t2c}$, the one providing the positive depth is the solution.

camera frames. The accuracy of the estimated rotation matrix was evaluated by the yaw, pitch, and roll angles, corresponding to the three rotations with respect to the X, Y and Z axes.

$$R = \begin{pmatrix} \cos\theta_{roll} & -\sin\theta_{roll} & 0 \\ \sin\theta_{roll} & \cos\theta_{roll} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta_{pitch} & 0 & -\sin\theta_{pitch} \\ 0 & 1 & 0 \\ \sin\theta_{pitch} & 0 & \cos\theta_{pitch} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{yaw} & -\sin\theta_{yaw} \\ 0 & \sin\theta_{yaw} & \cos\theta_{yaw} \end{pmatrix}$$

15 3D calibration points, used to estimate the fundamental and essential matrices, were randomly selected within the field-of-view of the camera and their depths were between 5 and 10 meters away from the camera. The same amount of calibration points were located on a plane with a normal vector $(0, \frac{1}{2}, \frac{1}{2})^T$ and a distance 10 meters away from the camera. The intrinsic matrix was assumed to be known perfectly and the image points were transformed from the image frame to the camera frame before the calibration. The image points in the camera were perturbed by a Guassian noise with 0 mean and a standard deviation varying from 1 to 5 pixels and

Algorithm 5.4 Algorithm to Estimate Rotation Matrix Through Essential Matrix

Require: Collect at least 8 corresponding vector pairs by rotating the gimbal.

Require: The measured image positions must be transformed to the camera frame by multiplying the inverse of the intrinsic matrix.

Ensure: Notice that the 3D bright spots illuminated by the transceiver must not be coplanar.

Essential Matrix Estimation

1: Estimate an initial essential matrix from equation 5.4:

$$\min_{e} \left\| f(x, x^{'})e \right\| \qquad \text{subject to } \|e\| = 1$$

where e is a vectorized essential matrix.

2: Impose the zero determinant constraint by finding α and β from e:

$$\begin{pmatrix} e^{1\sim3} & e^{4\sim6} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = e^{7\sim9}$$

3: Minimize the algebraic norm again with an imposed constraint

$$\min_{e} \left\| f(x, x^{'})e \right\| \quad \text{subject to } \|e\| = 1 \text{ and } \sigma_1 - \sigma_2 = 0$$

where (σ_1, σ_2) represent the two singular values of the estimated essential matrix.

4: Minimize the Mahalanobis distance :

$$\min_{e^{1\sim6}, \alpha, \beta} d(\hat{x}^{'}, x^{'})$$

Rotation Matrix Estimation

Require: An estimated essential matrix: $\hat{E} = \lambda(T_{t2c})_{\times} R_{t2c}$.

5: Solve for the R_{t2c} and T_{t2c} from equation 5.6.

6: Estimate the intersection of the rays by solving equation 5.31. Among the four combinations of $\hat{R}_{t2c}$ and T_{t2c}, the one providing the positive depth is the solution.

the ray vectors from the gimbal were perturbed by another Gaussian noise with 0 mean and a fixed standard deviation of 63 μrad. For each amount of the noise, 100 trials were executed and the RMS angular error of the estimated rotation matrices are shown in figure 5.6

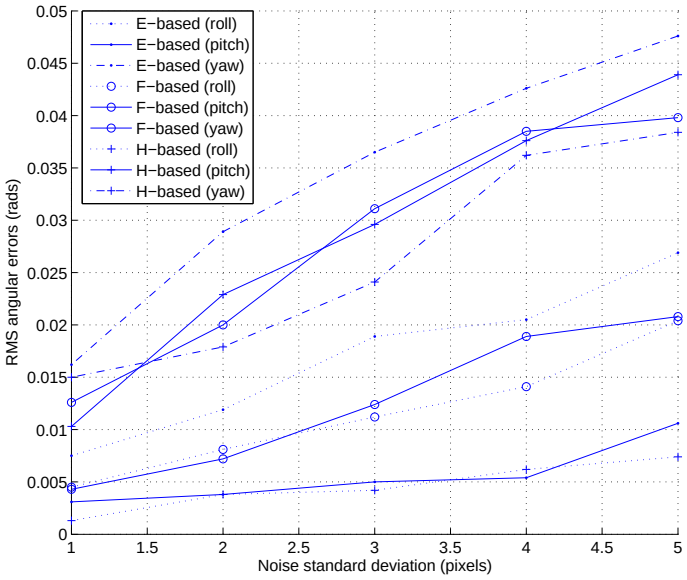

Figure 5.6: RMS angular error for the estimated rotation matrix

The results indicate that the angular errors could increase from 5 to 50 mrad with respect to the magnitude of the noise for all three algorithms. The inferior performance of estimating the rotation matrix arises from the fact that the linear homography, $M(R_{t2c} + \frac{1}{d}T_{t2c}N^T)$, embedded in the current configuration, is not the desired invariance, (MR_{t2c}). The numerical procedure therefore cannot provide a meaningful cost function for optimization, which leads to a suboptimal estimation. In the next section, we propose to apply additional optical systems to assist the calibration. The invariance in the resulting configuration is directly desired one, (MR_{t2c}).

We are interested in a general calibration algorithm without using additional optical systems. The current investigation focus on exploring the approximated linear constraints for the essential matrix which is closely related to the rotation matrix. In addition, a first-order theoretical analysis based on matrix perturbation theorem is under investigation, which is expected to help us characterizing the performance of the algorithms and understanding the error due to different geometrical scenarios.

5.1.2.4 Optical System Assisted Calibration

If the displacement between the camera and the transceiver can be minimized ($\|T_{t2c}\| \rightarrow 0$) or the distance of the calibration source can be extend to infinity ($d \rightarrow \infty$), the current homography can be converted into the desired homography. This transformation can be assisted by additional optical systems.

The first candidate is a beamsplitter, which produces a pseudo camera overlapped with the transceiver, shown in figure 5.7 (a). Since this estimated homography is the desired one, the accuracy can be expected to be better than 1 mrad. However, the field-of-view and steering range of the resulting CPAT system, limited by the size of the beam splitter, is often very small.

The other candidate is to create pseudo sources at infinity through the help of an imaging system and align the camera with the sources. This idea was adapted by the calibration process described in section 3.3.3, where the combination of a collimator and a corner cube reflector (CCR) projected the laser diode of the trans-

mitter as the pseudo source to infinity and projected onto the local AR. However, this process is not suitable for the camera/transceiver alignment because the CCD array which is easily saturated by the laser intensity from the transmitter and the coherency of the light may produce spurious images. Therefore, instead a CCR, a combination between several fiber tips and an imaging system, shown in figure 5.7 (b), can also project the source to infinity. The detailed procedure is listed below.

Algorithm 5.5 Optical System Assisted Calibration

1: Estimate the intrinsic matrix of the camera using off-line algorithms, such as the toolbox from Caltech [32].
2: Mark the principal point and the horizontal line passing through the principal point in the image plane.
3: Place a fiber at the focal plane of the imaging system.
4: Rotate the gimbal until the fiber receives the maximum power.
5: Rotate the gimbal in azimuthal direction in a small angle (fewer than the field-of-view of the imaging system).
6: Place another fiber next to the previous fiber which maximizes the received power.
7: Calibrate the camera such that the principal point overlaps with the first fiber and the horizontal line passes through the image of the two fiber tips.

This camera alignment process produces an uncertainty of around tens of μrad, which can be formulated as

$$\theta_{cal} \sim \frac{r_{fib}}{f}$$

where r_{fib} is the fiber radius and f is the focal length of the imaging system used in the calibration.

The remaining uncertainty is in the intrinsic matrix calibration process whose accuracy is usually less than 1 mrad, depending on the quality of the camera lens. Therefore, the resulting system has the potential to serve directly as an FPAT system.

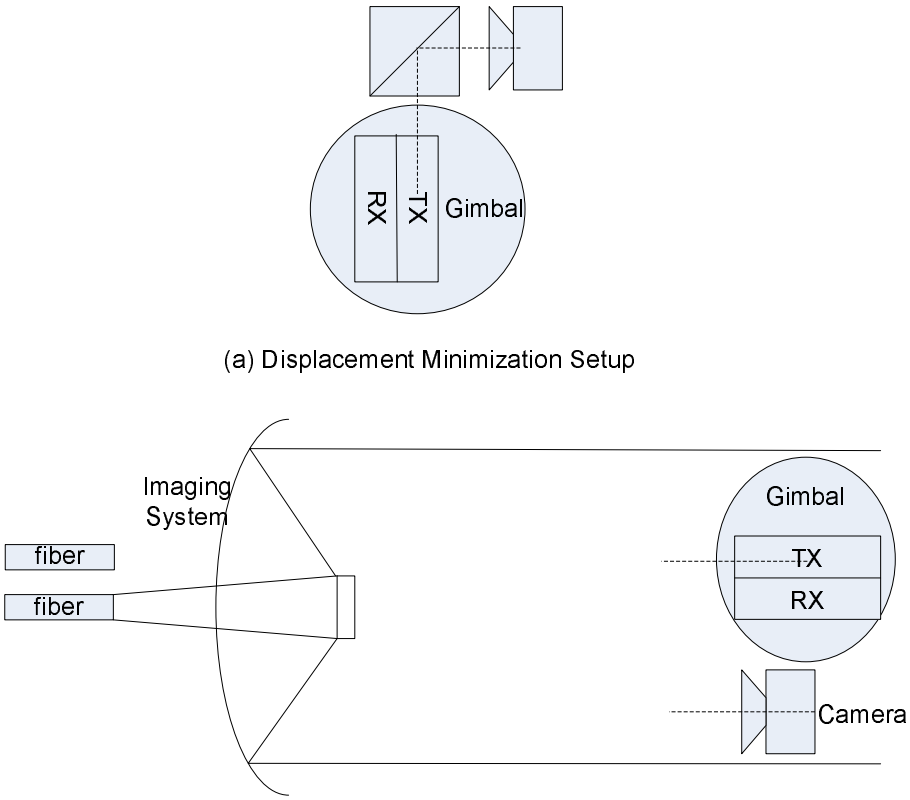

Figure 5.7: The configuration of the optical system assisted calibration

5.2 General Coarse Pointing Systems

A planar CPAT system is sufficient for most FSO applications. However, for the scenarios where the inequality of equation 5.1 fails, the target distance must be taken into account to provide an accurate pointing performance. Since a human's stereo vision can perceive an object's depth, the same concept can be implemented with two cameras. Instead of a linear homography mapping between a single camera and a transceiver, the transformation among the transceiver frame and the two different camera frames is governed by the trifocal tensor (TT).

5.2.1 Geometry

Consider a CPAT system consisting of two perspective cameras and a two-axis gimbal with a transceiver mounted on the pivoting point. The transformations from the world frame to the perspective image frame and transceiver frame are modelled by equations 4.5 and 4.20, respectively. By rewriting them in matrix form, the geometrical constraint can be summarized as

$$\begin{pmatrix} P & x & 0 & 0 \\ P' & 0 & x' & 0 \\ P" & 0 & 0 & x" \end{pmatrix} \begin{pmatrix} X_w \\ -\lambda \\ -\lambda' \\ -\lambda" \end{pmatrix} = 0_{9\times 1} \tag{5.7}$$

where $P, P"$ represent the perspective projection matrices of the two cameras, P' represents the transformation matrix of the two-axes gimbal, and X_w represents the coordinates of the target in the world frame.

The existence of the null vector shows that the maximum rank of the left matrix is 6. Since the matrix is 9×7, it implies any 7×7 minor matrices have a zero determinant. We choose a minor matrix as follows:

$$\det \begin{pmatrix} P & x & 0 & 0 \\ P'^j & 0 & x'^j & 0 \\ P'^l & 0 & x'^l & 0 \\ P"^k & 0 & 0 & x"^k \\ P"^m & 0 & 0 & x"^m \end{pmatrix} = 0$$

By defining the trifocal tensor as:

$$T_i^{qr} = \frac{1}{2}\epsilon_{ilm} \det \begin{pmatrix} P^l \\ P^m \\ P'^q \\ P"^r \end{pmatrix} \tag{5.8}$$

The geometric constraint can therefore be formulated by expanding the determinant

$$x^i x'^j x"^k \epsilon_{jqu} \epsilon_{krv} T_i^{qr} = 0_{uv} \tag{5.9}$$

where $\epsilon_{abc} = \begin{cases} 0, & \text{a,b, and c are not distinct;} \\ 1, & \text{if abc is an even permutation of 123; ,} \\ -1, & \text{if abc is an odd permutation of 123.} \end{cases}$

and $\epsilon_{abc} x^a$ is equivalent to the cross product operator $(x^a)_\times$ in equation 4.5.

5.2.2 Calibration Stage

Equation 5.8 shows that a trifocal tensor is a $3 \times 3 \times 3$ matrix consisting of 27 elements but 26 degrees of freedom because of an unknown scaling factor. According to the geometric constraint of 5.9, since $\epsilon_{abc}x^a$ is skew-symmetric and rank-deficient, each corresponding triplet provides only four independent equations to solve for the TT. Therefore, 7 triplets are required.

However, a qualified TT has only 18 dof, which can be computed as: The mapping from the world frame to each component's frame has 11 parameters which result in total 33 unknowns for two cameras and a transceiver. However, the world frame can be arbitrary selected, which reduces the unknowns by 15 and the number of the essential parameters is therefore only 18. There are 8 internal constraints which must be imposed to estimate the TT. These internal constraints can be explained by the degenerate conics concept in section 4.5.2. Note that by selecting the first perspective image frame to be the world frame, the TT can be cast as

$$T_i^{qr} = P_i^{'q} e^{"r} - e^{'q} P_i^{"r} \quad , i = (1, 2, 3) \tag{5.10}$$

where P_i is the i^{th} column of a projective matrix and $(e^{'}, e^{"}) = (P_4^{'}, P_4^{"})$ represents the image of the first perspective camera's origin expressed in the transceiver and the second perspective image frame, also known as the epipoles.

The algorithm to estimate the trifocal tensor includes three steps summarized briefly below. The details can be referred to in chapter 16 of Hartley's book [23].

Simulation 5.5: Robustness Simulation

The robustness of the trifocal tensor estimation algorithm is investigated in this

Algorithm 5.6 Trifocal Tensor Estimation Algorithm

Require: Rotate the two-axis gimbal to at least 7 positions and record the laser vectors with equation 4.21 and the corresponding image coordinates in the two perspective cameras. Notice that these points must not be coplanar.

AMS algorithm

1: Normalize the measured coordinates to a mean position at the origin and an average distance $\sqrt{2}$ with respect to the mean position for each component.
2: Use the norm of equation 5.9 as the cost function and solve for its minimizer as the initial TT.
3: Estimate a pair of $(P_4^{'}, P_4^{"})$ from the initial TT and solve for the previous minimization problem again with the additional linear constraint described in equation 5.10. The new minimizer is a qualified TT, which satisfies the 8 internal constraints.

GMS algorithm

4: Estimate the two projective matrices mapping the world frame to the the transceiver and the second perspective image frame from the qualified TT.
5: Triangulate three vectors to determine the coordinates of the points in 3D and then re-project the 3D points back to each frame.
6: Find the minimizer of the Mahalanobis distance between the computed points and the measured points with a numerical iterative method, such as Levenberg-Marquardt algorithm .
7: De-normalize the minimizer to retrieve the best-estimate TT.

simulation. The simulation scenario consisted of two perspective cameras and a two-axes gimbal. Both of the perspective cameras had a semi-FOV 15° and the setup is plotted in figure 5.8. 24 3D points were uniformly selected within the inner spherical region covered by the field-of-view of all three components. Additive Gaussian noise with 0 mean and standard deviation from 0.5 to 3 pixels was added to the triplets. For each different noise magnitude, 100 experiments were made and the average residual and estimation errors after step 2 (after imposing internal constraints) and 3 (after the Levenberg-Marquardt optimization process) are presented in figure 5.9.

The results show that the residual and the estimation errors both increase linearly with the magnitude of the measurement error and can be predicted precisely by the first-order error propagation in the current scenario. Using the AMS algorithm

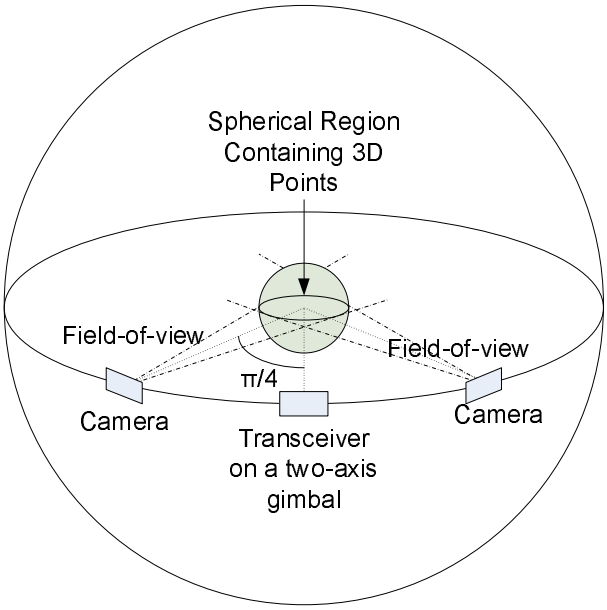

Figure 5.8: Robustness simulation setup: the radius of the outer sphere is $10\sqrt{2}$ and the radius of the inner sphere is $10\sqrt{2}\sin$ (semi-FOV)

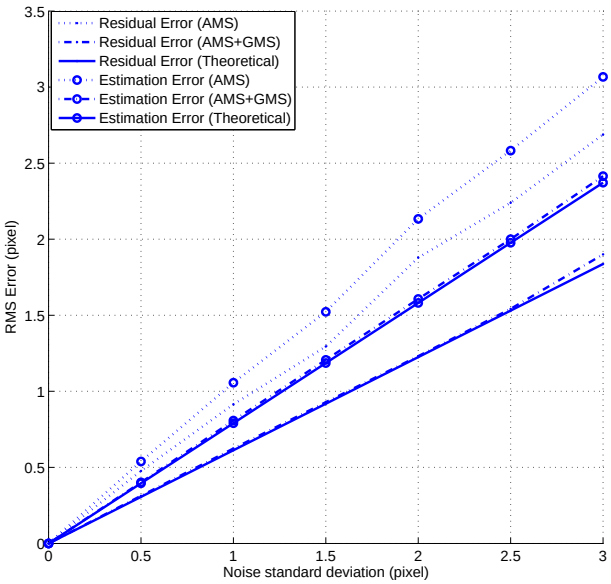

Figure 5.9: RMS error performance of the trifocal tensor estimation

alone produces about 30% higher estimation errors and 40% higher residual errors compared to the combined algorithm (AMS+GMS), which indicates the necessity of the iterative algorithm in TT estimation.

By changing the positions of the two perspective cameras on the outer sphere, we further investigated the effects for other geometric setups. The results indicated that a better AMS estimation can be achieved by simultaneously increasing the field-of-view of the camera and the sphere containing the 3D points.

5.2.3 Pointing Stage

After the trifocal tensor is retrieved, if the target appears in both of the perspective cameras, the corresponding rotation of the transceiver can be computed by rearranging equation 5.9 as

$$x^{j'} = x^i T_i^{jk} l_k^{"} \tag{5.11}$$

where x represents the image position at the first perspective camera, $l^{"}$ represents a line passing through the corresponding image position at the second perspective camera, and $x^{j'}$ represents the corresponding transceiver ray vector.

This simple equation leads to a challenging question:

What is the best way to select the best $l^{"}$?

Finding the line $l^{"}$ passing through the point $x^{"}$ is equivalent to looking for a vector v satisfying $v^{"T} x^{"} = 0$. This is an under-constrained problem and has infinite solutions. Among these solutions, the only line to avoid is the one intersected by the epipolar plane produced by x, $x^{"}$, and $e^{"}$ and the second camera's image plane.

This line results in the degeneracy of the equation 5.9: $x^i T_i^{jk} l_k^{"} = 0^j$ and can be computed as:

$$\epsilon_{rsk} e^{"s} (T^T)_i^{kj} e^{'j} x^i$$

Thus, the best selection of $l^{"}$ is obviously the line perpendicular to it and also passing through $x^{"}$.

The image pair $(x, x^{"})$ is selected by the user, which inevitably introduces measurement errors. This error affects the estimation of the line $l^{"}$ and further deteriorates the pointing accuracy. Hartley proposed a sophisticated algorithm to refine the selection by using the epipolar constraint introduced by the selected image pair $(x, x^{"})$ and the epipole $e^{"}$.

His idea was that if the selected image pair were the best, the line intersected by the epipolar plane formed by the best image pair and the epipole should have a minimum distance to the corresponding measured image point. By parameterizing the best image point, he turned the distance minimization problem into a single-variable and six-order polynomial function. The details are revealed in chapter 12 of his book [23].

Given the same setup as in simulation 5.5, the error improvement resulting from the point refinement is shown in figure 5.10.

5.3 Wide Field-Of-View Coarse Pointing Systems

A laser pointing/acquisition system requires a sensor to provide angular information with a high degree of accuracy for its autonomous pointing capability. A

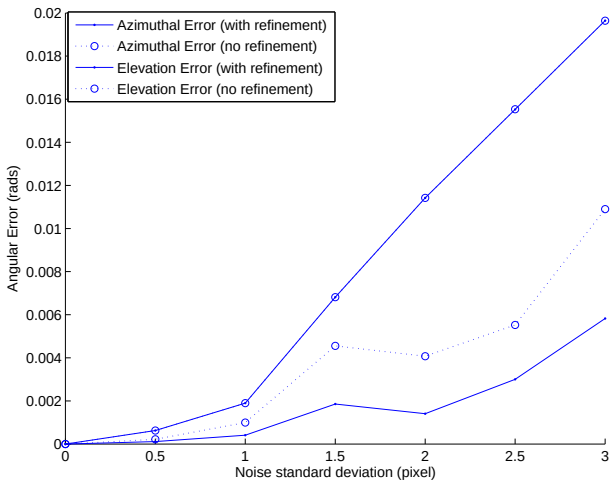

Figure 5.10: Improvement of the pointing error from the point refinement algorithm

perspective camera is an ideal solution because the lens is optimized as a first-order imaging device. Unfortunately, such a system suffers from a small FOV ($15° - 40°$). To increase the viewing angle for the imaging system will inevitably bring in higher order aberrations, especially barrel distortions[10]. Since the distortion lies symmetrically in the radial direction from the centroid of the image plane, it can be pre-calibrated and corrected [24].

Wide field-of-view cameras are generally categorized into two types:

1. Catadioptric camera: The term "catadioptric" refers to an imaging system involving both lenses and mirrors in the design. A catadioptric camera typically includes a convex mirror, which reflects the rays from a wide field-of-view into a perspective imaging system. Because of the simple structure of the mirror, the imaging model can be perfectly analyzed. Baker et.al. [7] presented a fixed viewpoint constraint, where all the light rays must pass through a single 3D point before imaging on the CCD array. Only the distortions produced by

the surfaces satisfying such a constraint can be geometrically corrected by the perspective image. The qualified surfaces include planar, elliptical, parabolic, and hyperbolic shapes, but only the last two are practical in producing a wider field-of-view. Geyer et.al. [18] proposed the catadioptric geometry, where they showed the catadioptric projection model was equivalent to a spherical projection, which was controlled by two cata-projection parameters: the distance from the sphere center to the equivalent viewpoint and to the image plane, respectively. Most importantly, they discovered that the image of a 3D line is a conic section in the catadioptric image and proposed to utilize these conic sections to estimate the intrinsic matrix and the two cata-projection parameters simultaneously. However, because of the limited fov and length of the line, the resulting conic sections only partially appear in the image, which may lead to an unstable estimation. Ying et.al. [79] showed that the catadioptric image of a 3D sphere also resulted in a conic section. They further proposed a calibration algorithm based on 3D spherical objects instead of straight lines. The image produced by spherical objects contains a full conic section, which greatly improves the estimation stability. The imaging model of the catadioptric camera is plotted in figure 5.11.

2. Dioptric camera: The distortion of the catadioptric camera can be fully modelled because the reflecting surface used to capture the light rays from a wide field-of-view is simple. As for the dioptric cameras, the light rays are captured by multiple refractive lenses instead of a simple reflecting surface, which

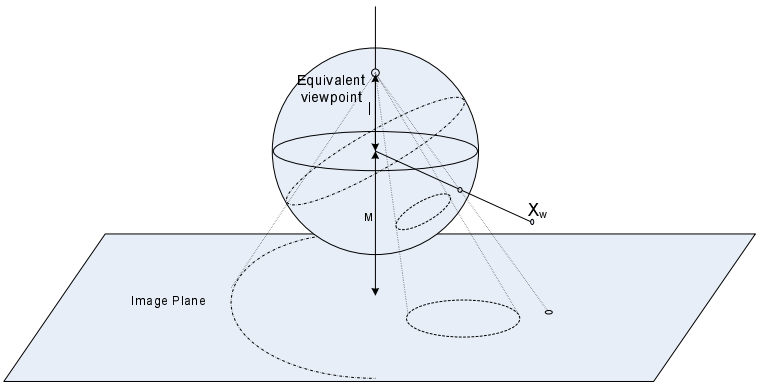

Figure 5.11: Catadioptric imaging model: This figure shows the catadioptric image of a 3D point, line, and a sphere. The point X_w is first projected as a point on a unit sphere through the sphere center and this point is further imaged through an equivalent viewpoint onto the image plane. l and M represent the distance from the unit sphere centroid to the equivalent viewpoint and to the image plane, respectively.

eliminates the good optical properties, such as the spherical projection, in central catadioptric cameras. We therefore consider a general distortion model by assuming the distortion lies symmetrically in the radial direction from the centroid of the image plane and the details are plotted in figure 4.4.

The wavefront aberration theorem[10] suggested that the distortion is only affected by the odd polynomials of the paraxial image radius r. This model is denoted as the radius-radius model and can be cast as

$$r^{'} = r + \sum_{k=1}^{n} a_k r^{2k+1} \tag{5.12}$$

where r and $r^{'}$ are the paraxial and dioptric image radii, respectively.

Since the paraxial image position is in general unknown, applying the pre-

vious equation to estimate the distortion usually leads to solving high-order equations, which is impractical. Several researchers proposed alternative distortion models which depend on the current image position instead of the paraxial image position. These models can be categorized into two types:

- Division model: Fitzgibbon[17] introduced a model to describe the relation between the dioptric image radius ($r^{'}$) and its paraxial radius (r), where

$$r = \frac{r^{'}}{1 + p_1 r^{'2}}$$

  Micusik et.al [48] and Thirthala et.al [63] further generated this model as

$$r = \frac{r^{'}}{1 + \sum_{k=1}^{n} p_k r^{'2k}} \tag{5.13}$$

- Theta-radius model: Incident angle (θ) v.s. Image radius ($r^{'}$): Margaret [45] suggested to adapt the Cartography projections to wide-angle lenses, where the distortion relationship is described by the incident angle (θ) and the dioptric radius ($r^{'}$) including:

$$\begin{cases} r^{'} = 2f \tan\frac{\theta}{2}, & \text{stereo-graphic;} \\ r^{'} = f\theta, & \text{equidistance;} \\ r^{'} = f \sin\theta, & \text{sine law;} \\ r^{'} = 2f \sin\frac{\theta}{2}, & \text{equi-solid angle.} \end{cases}$$

  Among the models above, she further suggested the stereo-graphic projection because of its area preservation property. Geyer [18] further showed that a catadioptric camera which uses a parabolic mirror as a reflector

follows this projection model. In addition, Ying [79] also showed that this model is identical to Fitzgibbon's division model [17]. Bakstein et.al [8] suggested to use a hybrid model given as

$$r^{'} = p_1 \sin p_2\theta + p_3 \tan p_4\theta$$

Finally, Kannala et.al [35] proposed to use a general model

$$r^{'} = \sum_{k=0}^{n} p_k\theta^{2k+1} \tag{5.14}$$

In order to make a comparison, 17 different wide-angle lens models in CodeV with a semi field-of-view larger than 80 degrees were selected to verify the correctness of the distortion model. Three models including the (1) Radius-Radius model, as described in equation 5.12, (2) Division Model, as described in equation 5.13, and (3) Theta-Radius model, as described in equation 5.14 were tested and the number of the parameters for each model was 3. 350 random rays with incident angles distributed within 0 to 70 degrees were applied and their corresponding image positions were recorded. The model parameters were first estimated and the incident angles were further computed from the estimated parameters and the image points. The results are shown in figure 5.12.

From the results, the theta-radius model outperforms the other two models and the estimation error is on the scale of milli-degrees for all 17 different wide-angle lenses. The division and the radius-radius models contribute similar amounts of error on the scale of 0.1 degrees for all the lenses. Although

the theta-radius model produces the best estimation results, incorporating it into the algorithm usually leads to solving a highly nonlinear equation. Among the three models, since the division model can be easily incorporated into the algorithm by using the denominator as the third component of the homogeneous coordinates and its error performance is acceptable for the CPAT system, it is selected as the model in the following algorithm.

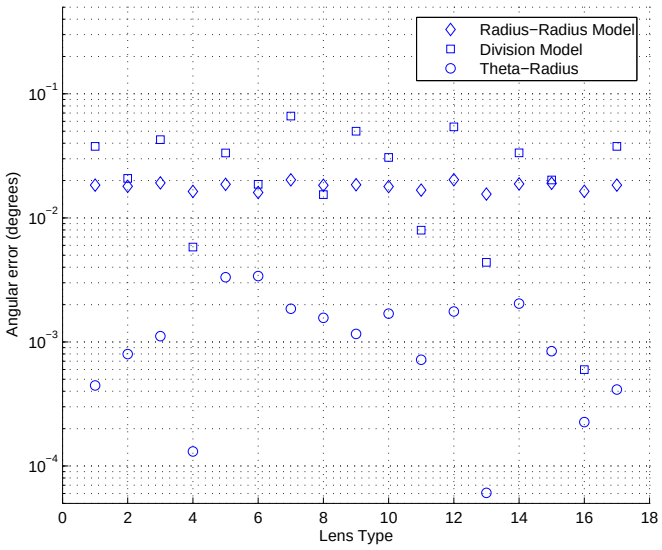

Figure 5.12: Estimated angular errors of 17 different wide angle lenses using different dioptric distortion models. The number of distortion parameters is 3 for each model.

Compared to the catadioptric camera, the dioptric camera has several advantages: (1) no dead-spot in the middle of the image, (2) a more compact design, and (3) a better aligned optical axis of the lens. Figure 5.13 shows a catadioptric camera, a dioptric camera, and their wide-angle images.

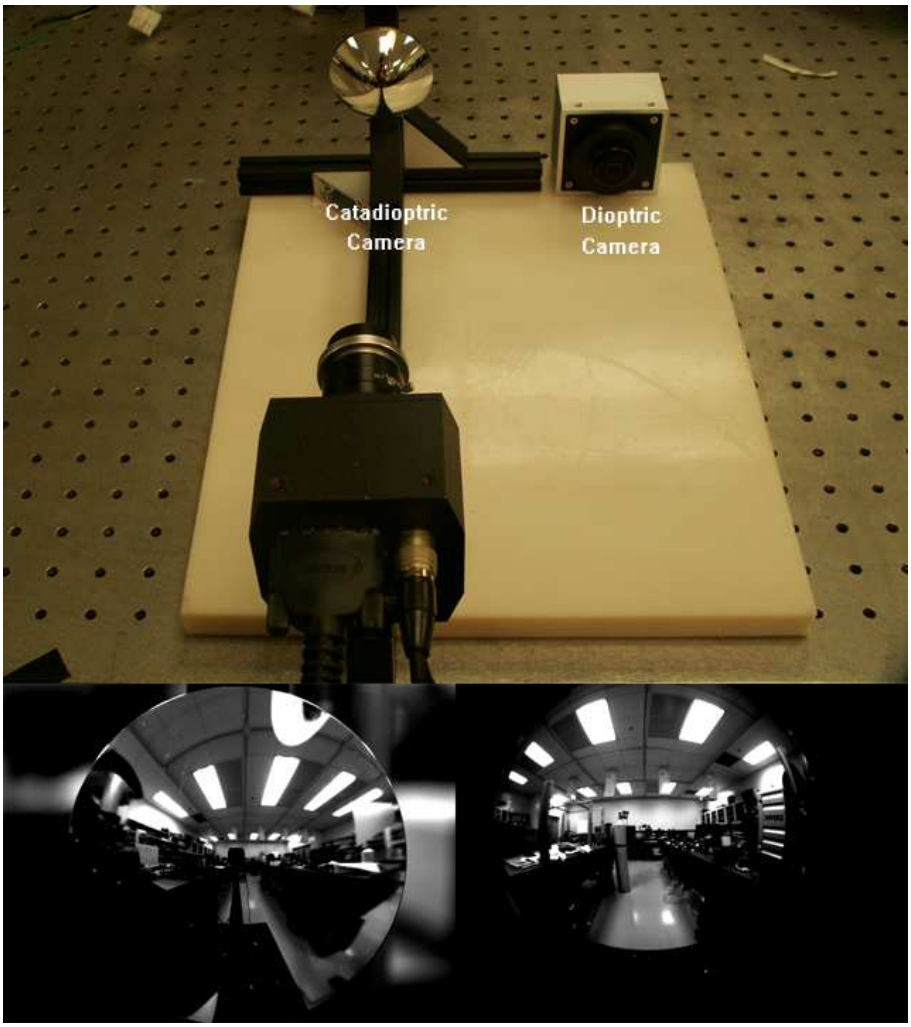

Figure 5.13: The top picture shows a catadioptric camera (left) and a dioptric camera (right). The bottom left and right images are taken by the catadioptric and dioptric camera, respectively, at the same location. The image indicates that the selected catadioptric camera has a wider FOV than the selected dioptric camera.

5.3.1 Distortion Correction Algorithm

In this section, an algorithm to estimate the intrinsic matrix and the distortion parameters is presented for the dioptric camera. This algorithm is mainly divided into two parts: (1) Intrinsic matrix estimation and (2) parametric distortion model estimation. As shown in figure 4.4, the planar angle (θ) is preserved in both the dioptric and catadioptric cameras. Because of the unknown distortion parameters, the paraxial image point is located on the line between the dioptric image point and the distortion center. The distortion center is assumed to be the principal point since the optical axis of the wide-angle lens is straight.

Our idea is to utilize multiple dioptric image points since the intersection of these 2D lines reveals the information of the principal point. After retrieving the principal point, the skewness and the aspect ratio can further be estimated by imposing the orthogonal constraint from the scene. Then, the distortion model can be characterized from the coordinates of the selected scene point.

5.3.1.1 Intrinsic matrix estimation

The mapping from a 3D point to its corresponding dioptric image point, according to equation 4.22 is

$$\begin{pmatrix} x_w^1 \\ \\ x_w^2 \end{pmatrix} = \lambda \begin{pmatrix} 1 & s \\ \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} P^1 \\ \\ P^2 \end{pmatrix} X_w + \begin{pmatrix} o_x \\ \\ o_y \end{pmatrix} \tag{5.15}$$

where P^n represents the n^{th} row of the transformation matrix $(R|T)$, (s, α) represent the skewness and aspect ratio of the camera, $X_w : (X_w^1, X_w^2, X_w^3, 1)^T$, and (x_o, y_o)

represent the principal point of the camera.

We further eliminate λ to obtain a linear equation

$$\alpha P^2 x^1 X_w + ((P^1 + sP^2)o_y - \alpha P^2 o_x)X_w - (P^1 + sP^2)x^2 X_w = 0 \tag{5.16}$$

This equation is a quadratic equation with 10 unknown parameters (4 for αP^2, 4 for $P^1 + sP^2$, and o_x, o_y). If we assume $(P^1 + sP^2)o_y - \alpha P^2 o_x$ as 4 new parameters, we can obtain a linear equation containing 12 parameters. Using 11 correspondences, the 12 parameters can be solved using algorithm A.1.

Let the 12 parameters be a_{1-12}, where a_{1-4} represent αP^2, a_{5-8} represent $-(P^1 + sP^2)$, and a_{9-12} represent $(P^1 + sP^2)o_y - \alpha P^2 o_x$. The principal point (o_x, o_y) can be estimated by solving

$$\begin{pmatrix} a_1 & a_5 \\ a_2 & a_6 \\ a_3 & a_7 \\ a_4 & a_8 \end{pmatrix} \begin{pmatrix} -o_x \\ -o_y \end{pmatrix} = \begin{pmatrix} a_9 \\ a_{10} \\ a_{11} \\ a_{12} \end{pmatrix} \tag{5.17}$$

After estimating the principal point, the skewness and the aspect ratio can also be retrieved by using the orthogonality of the transformation matrix. Recall that P^1, P^2 represent the first two rows of the transformation matrix. Therefore, the first three elements of the vectors are orthogonal and their individual norm is equivalent. The two constraints can be cast as

$$\begin{aligned} \frac{1+s^2}{\alpha^2} &= \frac{\sum_{i=5}^{7} a_i^2}{\sum_{i=1}^{3} a_i^2} \\ \frac{s}{\alpha} &= \frac{a_1 a_5 + a_2 a_6 + a_3 a_7}{\sum_{i=1}^{3} a_i^2} \end{aligned} \tag{5.18}$$

After recovering the intrinsic matrix, the extrinsic matrix can also be partially estimated (except for T^3) by

$$\begin{pmatrix} R^1 & T^1 \\ R^2 & T^2 \end{pmatrix} = \begin{pmatrix} 1 & s \\ 0 & \alpha \end{pmatrix}^{-1} \begin{pmatrix} -a^{5-8} \\ a^{1-4} \end{pmatrix} \qquad (5.19)$$
$$R^3 = (R^1)_\times R^2$$

where R^i represents the i^{th} row of the rotation matrix, T^i represents the i^{th} element of the translation vector, and $(.)_\times$ represents the cross product operator.

5.3.1.2 Incorporating the Distortion Model

We selected our distortion model to be the division model with 3 parameters as shown in equation 5.13. By incorporating the division model into the third component in the camera frame, the equation 5.15 can be expanded as

$$\begin{pmatrix} 1 & -\frac{s}{\alpha} & \frac{so_y}{\alpha} - o_x \\ 0 & \frac{1}{\alpha} & -\frac{o_y}{\alpha} \\ 0 & 0 & f(.) \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ 1 \end{pmatrix} = \lambda \begin{pmatrix} R^1 & T^1 \\ R^2 & T^2 \\ R^3 & T^3 \end{pmatrix} X_w \qquad (5.20)$$

where $f(.) = p_1 + p_2 d^2 + p_3 d^4$ represents the division model and
$d = \sqrt{(x^1 - o_x - \frac{s(x^2 - o_y)}{\alpha})^2 + (\frac{x^2 - o_y}{\alpha})^2}$.

Note that there are 4 unknowns in equation 5.20, which includes T^3 representing the object's depth and P_{1-3} representing the distortion parameters of the division model. By separating the unknown parameters and cancelling the unknown scaling factor λ, two constraints to estimate the distortion parameters can be cast

as

$$\begin{pmatrix} P^1X_w & 0 \\ 0 & P^2X_w \end{pmatrix} \begin{pmatrix} 1 & d^2 & d^4 & \frac{s(x^2-o_y)-\alpha(x^1-o_x)}{(R^1X_w+T^1)\alpha} \\ 1 & d^2 & d^4 & \frac{-(x^2-o_y)}{(R^2X_w+T^2)\alpha} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ T^3 \end{pmatrix} = \begin{pmatrix} \frac{\alpha(x^1-o_x)-s(x^2-o_y)}{\alpha} \\ \frac{x^2-o_y}{\alpha} \end{pmatrix} R^3X_w \tag{5.21}$$

Algorithm 5.7 Distortion Calibration Algorithm

Require: Take at least one image with an L-shape checkerboard and select at least 11 corner points from the checkerboard image in each picture.

1: Find a minimizer (composed of 12 parameters) regarding of the norm of equation 5.16 and estimate the principal point, skewness and aspect ratio from equations 5.17 and 5.18.
2: Retrieve the extrinsic matrix (except T_3) from equation 5.19.
3: Compute the estimated distorted image radius.
4: Find the minimizer regarding of the norm of equation 5.20. This minimizer consists of T_3 and the distortion parameters.

5.3.2 Performance Evaluation

The advantage of our algorithm is its capability to incorporate different images to retrieve a better estimation, which helps to mitigate the measurement uncertainty. Also, since we assume a general distortion model, this algorithm can also be applied to the catadioptric camera. The simulation results regarding the algorithm performance were presented in our work [25].

Figure 5.15 shows a corrected image taken by a dioptric camera whose imaging lens is a fisheye lens (ORIFL 190-3) from Omnitech. This lens has semi FOV up to 95 degrees on 1/3 inch CCD array. Because the algorithm requires a $3D$

calibration object, an L-shape checkerboard was assembled for the experiment. In this experiment, 18 control points were selected from a single image, which can be seen in the figure 5.14. Figure 5.15 was corrected based on the estimated parameters. The corrected image captured a field-of-view of 70°. The correction performance can be evaluated by observing the straight lines in the calibrated image since the straight lines are invariant in the perspective camera.

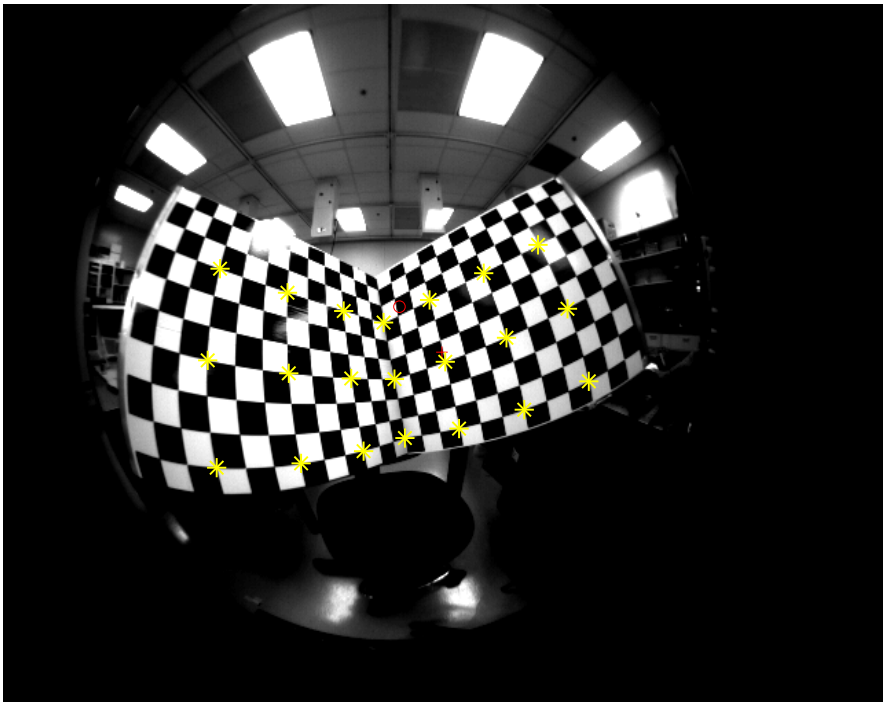

Figure 5.14: Distorted image taken from an Omnitech fisheye lens, Notice the plus sign is the center of the CCD, the circle sign is the principal center estimated from the proposed algorithm, and the 18 star signs indicates the selected control points for our algorithm.

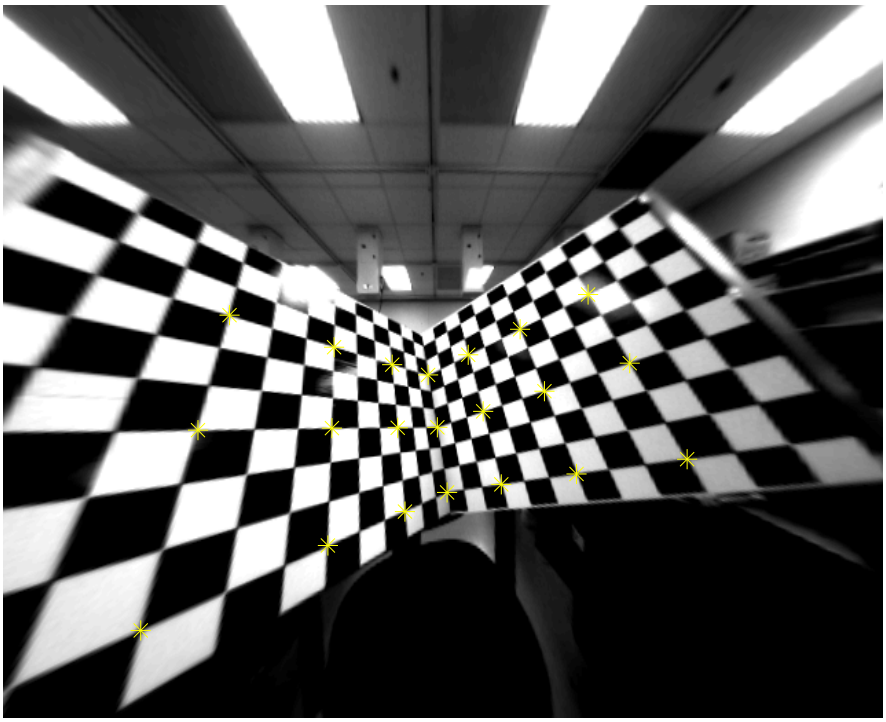

Figure 5.15: Distorted scene corrected by the proposed algorithm: the 18 star signs indicate the selected points corrected by our algorithm. The corrected scene has a semi FOV up to 70 degrees.

We also compared our results with the characteristic curve (incident angle v.s. projection radius) posted on their website[1].The parameters we extracted from their curve using the division model were $[165.40, -1.9e^{-3}, -2.37e^{-8}]$, versus our estimation distortion parameters, which were $[167.45, -2.5e^{-3}, -8.167e^{-9}]$. Since the distortion model can only be estimated after the intrinsic matrix is known, a close match in the distortion parameters represents an accurate estimation of the intrinsic matrix.

Also, we selected 190 angles uniformly distributed within ± 70 degrees to compute the difference of the angles while using two different sets of coefficients. The angular error histogram is plotted in figure 5.16, which shows the maximum angular error in the test was smaller than $0.9°$ and most of the errors are fewer than $0.1°$.

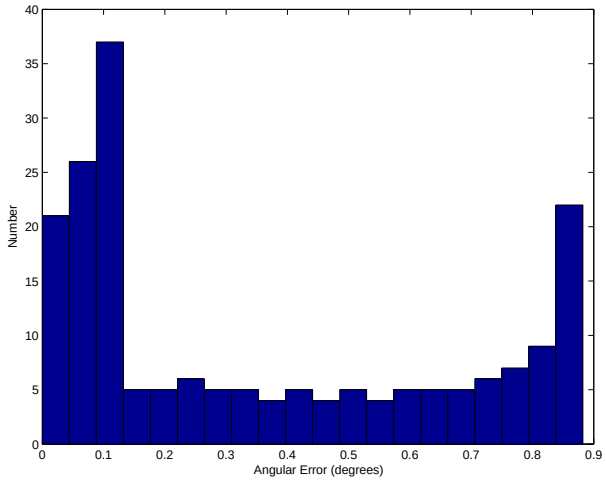

Figure 5.16: Angular error between the estimated coefficients and real coefficients from the datasheet of the OmniTech fisheye lens

[1]http://www.omnitech.com/fisheye.htm

5.3.3 Enhanced Planar Coarse Pointing Systems

The previous results suggest that the disadvantage brought by the wide-angle camera, namely distortions, can be effectively compensated for by a correction algorithm. Since the distortion is static, the calibration of the catadioptric/dioptric camera must only be performed once before applying it in the CPAT system. Once the calibration successfully corrects the distorted image position of the target into its perspective position, the corrected catadioptric/dioptric camera is equivalent to a perspective camera.

Therefore, the mapping between the frame of a singly corrected wide-angle camera and the transceiver remains a linear homography as described in section 5.1 and the CPAT system design algorithm introduced in algorithm 5.1. Furthermore, the mapping among the two corrected wide-angle cameras' frame and the transceiver remains to be the trifocal tensor as described in 5.2 and the CPAT system design algorithm shown in algorithm 5.6. The selection criterion between using the single or stereo cameras follows inequality 5.1.

5.4 Wide Field-Of-View and Three Dimensional Pointing Systems

Although a wide-fov CPAT system was proposed in the previous chapter and it provides a satisfactory performance, we are still interested in improving its pointing accuracy by giving up the distortion model. As shown in figure 5.12, the distortion cannot be completely compensated by the division model. The uncompensated error further affects the estimation of the homography or the trifocal tensor.

We therefore introduce another wide-fov CPAT system assisted by an additional perspective camera to achieve three-dimensional high-accuracy pointing. The theoretical basis of such a combination among a perspective camera, a wide-angle camera, and a two-axis gimbal has been discussed previously in section 4.5 and we have shown their geometric relation is encoded by a radial trifocal tensor (RTT). Compared to traditional stereo systems, the resulting one has a much wider 3D deterministic region where the coordinates of 3D points can be uniquely resolved. It is shown in figure 5.17.

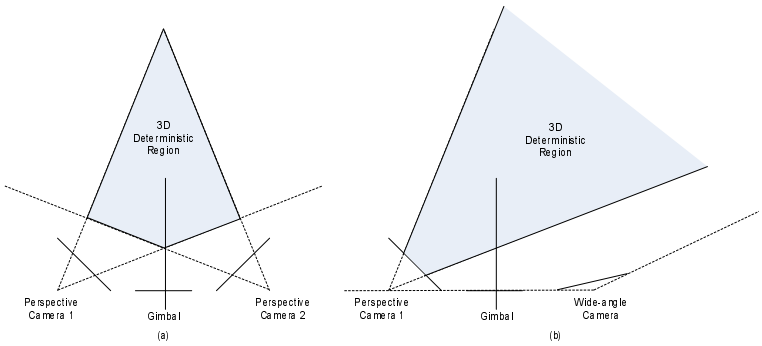

Figure 5.17: The region in which 3D coordinates can be uniquely determined (a) in a regular stereo vision system (b) in the proposed system

In addition, if the perspective camera can be rotated by another rotary gimbal, the field-of-view of the resulting system can achieve full hemisphere: (1) a target of interest is first selected from the wide-angle camera, (2) lock the target within the fov of the perspective camera by rotating the gimbal, and (3) pointing the transceiver toward the target according to the computed azimuthal and elevation angles.

5.4.1 Calibration Stage

The geometry of the RTT was introduced in section 4.5. This section focuses on a robust RTT estimation algorithm.

5.4.1.1 Algebraic Minimization Scheme (AMS)

Equation 4.24 provides three linear constraints to estimate the RTT cast as follows but only two out of three are independent. Notice that $x^{"}$ represents the pixel coordinate on the image plane subtracting the principal point estimated from equation 5.17.

$$T_{ik}^{q} x^{i} x^{'j} x^{"k} \epsilon_{jqu} = 0_u \tag{5.22}$$

By vectorizing T_{ik}^{q}, an algebraic cost function can be formulated as the following and we are interested in solving for the RTT which minimizes this function:

$$min_{vec(T)} \left\| f(x, x^{'}, x^{"}) vec(T) \right\| \tag{5.23}$$

$$\text{subject to } \|vec(T)\| = 1$$

where $f(x, x^{'}, x^{"})$ is a 3×18 matrix, $vec(.)$: matrix vectorization operator, (i.e. $vec(T) = (T_{11}^{1} T_{12}^{1} T_{11}^{2} T_{12}^{2} ...)^{T}$).

However, because of measurement noise and aberration errors, this RTT generally does not have the internal constraints imposed. The internal constraints that result from the intersections of the three degenerate conics are a point and a line in the perspective camera's frame, as described in section 4.5.2. The point and the line result from the epipole and the distorted-epipole, respectively. Therefore, imposing

the three constraints is equivalent to improving the estimation of a better epipole and distorted-epipole pair.

As for the epipole estimation, if we select x which is not on the locus of the axial plane or the line between the transceiver and the perspective camera, the $T^q_{ik}x^i$ is rank-2 and its left null vector represents the normal vector of the epipolar plane according to theorem 4.5.1. The vector which is orthogonal to all these normal vectors is the epipole, which is shown in figure 4.5. By picking at least two x, the epipole ($e^{'}$) can be estimated as

$$(e^{'})^T \begin{pmatrix} n_1 & n_2 & \cdots & n_k \end{pmatrix} = 0 \tag{5.24}$$

where n_j represents the null vector of $T^q_{ik}x^i$ resulting from j^{th} different x.

The distorted-epipole estimation relies on the constraint that the determinant of the minors of $T_i x^i$ is 0. For each constraint, the locus of x is a degenerate conic. It can be shown that at least one of the conic sections is rank-2 as long as the two cameras and the transceiver do not overlap. The constraints are summarized below:

$$\begin{aligned} x^T(T^1_{i1}T^2_{i2} + T^2_{i2}T^1_{i1} - T^2_{i1}T^1_{i2} - T^1_{i2}T^2_{i1})x &= 0 \\ x^T(T^2_{i1}T^3_{i2} + T^3_{i2}T^2_{i1} - T^3_{i1}T^2_{i2} - T^2_{i2}T^3_{i1})x &= 0 \\ x^T(T^1_{i1}T^3_{i2} + T^3_{i2}T^1_{i1} - T^3_{i1}T^1_{i2} - T^1_{i2}T^3_{i1})x &= 0 \end{aligned} \tag{5.25}$$

Theorem 5.4.1. *Let x be the null vector of the rank-2 conic section. Then, x is on the line intersected by the axial plane and the perspective image plane.*

Proof. Using the homogeneous coordinate, the degenerate conic is composed of two lines and the null vector is the intersection of these two lines.

Let x be the intersection of the two lines of the first conic section in equation 5.25. If x, the solution of the equation 5.25, is on the line intersected by the axial plane and the perspective image plane, this line must be one of the two lines of the first degenerate conic, which therefore must pass through the intersection. One concludes that the null vector must belong to the line intersected by the axial plane and the image plane. □

In most scenarios, the three conics are all rank-2. One method to find the distorted epipole is to plug the null vector n estimated from each degenerate conic back into $T_i n^i$. Since the distorted-epipole represents their common null vector, a TLS estimation of the distorted-epipole can be obtained by finding the null vector of the cascading $T_i n^i$.

$$\begin{pmatrix} T_i n_1^i \\ T_i n_2^i \\ T_i n_3^i \end{pmatrix} e^{"} = 0_{3\times 1} \tag{5.26}$$

where n_i is the null vector of i^{th} degenerate conics.

Another method is to directly estimate the $e^{"}$ from the three degenerate conics. By rewriting the rank-1 constraint, we can obtain the following rank-1 matrix

$$\begin{pmatrix} T_{1k}^1 e^{"k} & T_{2k}^1 e^{"k} & T_{3k}^1 e^{"k} \\ T_{1k}^2 e^{"k} & T_{2k}^2 e^{"k} & T_{3k}^2 e^{"k} \\ T_{1k}^3 e^{"k} & T_{2k}^3 e^{"k} & T_{3k}^3 e^{"k} \end{pmatrix} \tag{5.27}$$

Since the determinant of all the 2×2 minors is 0, 9 constraints can be cast and one of them is equal to

$$e^{"}(T_{1k}^1 T_{2k}^2 - T_{1k}^2 T_{1k}^2)e^{"} = 0$$

Since each of the constraints is scale-invariant and $e^{"}$ is only a vector, without loss of generality, we can scale $e^{"}$ to be $(x, 1)$, where x is an unknown scalar. If $e^{"}$ satisfies all nine constraints, the square-sum of the constraints should be minimized. Therefore, the distorted-epipole estimation problem turns into an optimization problem and the cost function is

$$\min_x \sum_{i=1}^{9} a_i^2 x^4 + 4a_i b_i x^3 + (2a_i c_i + 4b_i^2)x^2 + 4b_i c_i x + c_i^2$$

where $\begin{pmatrix} a_i & b_i \\ b_i & c_i \end{pmatrix}$ represents i^{th} 2×2 minors ($i = 1 \sim 9$).

The derivative of the cost function contains three roots. By plugging all the roots back into the cost function, the distorted-epipole can be expressed as $(x, 1)$, where x is the global minimizer of the cost function.

Figure 5.18 shows the relative error of the two different distorted-epipole estimation algorithms versus the measurement noise. The estimation $\hat{e}^{"}$ from both methods is first normalized to be $(\hat{e}^{"1}, 1)$ and the relative error is defined as $(\hat{e}^{"1} - \bar{e}^{"1})/\bar{e}^{"1}$, where $\bar{e}^{"1}$ represents the true value of the first element of the distorted-epipole.

The results indicated that the 9-constraints method performs slightly better than the conic-based method. Note that if the distorted-epipole is a line close to $y = 0$, the 9-constraints method must be modified to estimate the second element instead of the first one.

Equation 4.25 shows that $T_{ik}^q = P_i^{'q} L_k^{"} - e^{'q} P_i^{"\sim k}$, which can be rewritten in

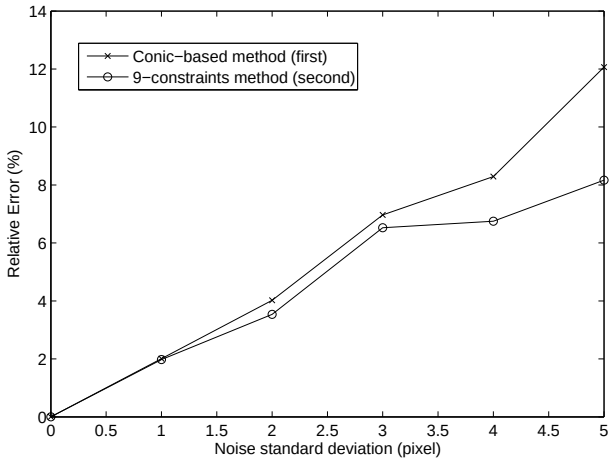

Figure 5.18: Error Performance of the estimated distorted-epipole with different algorithms

vector form as:

$$vec(T) = g(e^{'}, L^{"})vec(P^{'}_{1\sim3}, P^{"}_{1\sim3}) \tag{5.28}$$

where $g(e^{'}, L^{"})$ is a 18×15 matrix composed of the estimated epipole and the distorted-epipole, and $vec(P^{'}_{1\sim3}, P^{"}_{1\sim3})$ is a 15×1 matrix composed of the first three columns of both $P^{'}$ and $P^{"}$.

By replacing the $vec(T)$ in equation 5.23, the original unconstrained minimization problem turns into the linearly constrained problem below. The solution is discussed in appendix A.

$$\min_{vec(T)} \left\| f(x, x^{'}, x^{"})vec(T) \right\| \tag{5.29}$$
$$\text{subject to } \|vec(T)\| = 1 \quad \text{and } vec(T) = g(e^{'}, L^{"})vec(P^{'}, P^{"})$$

Instead of using the epipole and estimated-epipole estimated from the initial RTT, we can also iteratively update them to create a new $g(e^{'}, L^{"})$, which results

in a smaller algebraic error. However, the minimizer from this iterative algorithm is not guaranteed to have a smaller geometric error compared to the non-iterative one in our simulations.

5.4.1.2 Geometrical Minimization Scheme (GMS)

According to section 4.5.3, the extrinsic matrices of each frame can be determined from a RTT as

$$\begin{aligned} P &= \begin{pmatrix} \text{diag}(1,1,1) & 0_{3\times 1} \end{pmatrix} \\ P' &= \begin{pmatrix} T_{ik}^{q} L^{"k} & e' \end{pmatrix} \\ P" &= \begin{pmatrix} P_1^{"} & P_2^{"} & P_3^{"} & e^{"} \end{pmatrix} \end{aligned} \tag{5.30}$$

where $P_i^{"} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} (L^{"} L^{"T} - \text{diag}(1,1)) T_i e'$

Knowing the projective matrices for all three components, the three-dimensional points (X) can then be determined by solving the triangulation of the rays and plane from three components by

$$\min_{X_w} \left\| \begin{pmatrix} (x)_{\times} P \\ (x')_{\times} P' \\ (x^{"})_{\times} P^{"} \end{pmatrix} X_w \right\| \quad \text{subject to } \|X_w\| = 1 \tag{5.31}$$

By re-projecting the estimated X_w back onto each frame, a set of computed points $(\hat{x}, \hat{x}', \hat{x}^{"})$ can be retrieved. By minimizing the geometric error function between the computed points and the measured points using an iterative method, such as the Levenberg-Marquardt algorithm, the best estimator of T_{ik}^{q} can be obtained. Generally, the best geometric error function is the Mahalanobis distance. However,

since the covariance matrix is difficult to be estimated in the real scenario, we instead select the cost function as the Euclidean distance in the proposed algorithm instead, which is defined as

$$d_{res} = \sum_{i}^{n} \|\hat{x}_i - x_i\|^2 + \left\|\hat{x}_i^{'} - x_i^{'}\right\|^2 + \frac{(\hat{x}_i^{"1} x_i^{"2} - \hat{x}_i^{"2} \hat{x}_i^{"1})^2}{\left\|\hat{x}_i^{"}\right\|^2} \tag{5.32}$$

Note that since the image point in the wide-angle camera represents a radial line instead of a point, instead of defining the Euclidean distance as the distance between two points, we define it as the distance from the measured point to the radial line vector formed by the estimation.

Algorithm 5.8 Radial Trifocal Tensor Estimation Algorithm

Require: Find the principal point from equation 5.17 or use the center of the image as the principal point.

Require: Rotate the two-axis gimbal to at least 9 positions and record the transceiver vectors with equation 4.21 and the corresponding image position in the perspective and wide-angle cameras.

Ensure: Notice that the 3D bright points illuminated by the transceiver must not be coplanar.

AMS Algorithm

1: Normalize the measured positions from the perspective camera to 0 mean and an average distance $\sqrt{2}$.
2: Use the normalized image point and other measurements to compute the minimizer of equation 5.23.
3: Find the epipole $(e^{'})$ from equation 5.24 and the distorted-epipole $(e^{"})$ from either equation 5.26 or 5.27.
4: Construct $g(e^{'}, L^{"})$ and apply algorithm A.2 to solve for the minimizer of the equation 5.29, where $L^{"} = \begin{pmatrix} e^{"2} & -e^{"1} \end{pmatrix}$
5: Iteratively update the estimation of $(e^{'}, L^{"})$ in order to minimize the norm of equation 5.29.

GMS Algorithm

6: Estimate the projective matrices from both cameras and the rotary gimbal from equation 5.30.
7: Solve equation 5.31 to estimate the 3D coordinates.
8: Re-project the 3D points back onto each frame.
9: Minimize the Euclidean distance with an iterative numerical algorithm, such as the Levenberg-Marquardt algorithm.
10: Denormalize the minimizer to retrieve the best radial trifocal tensor.

5.4.1.3 Performance Evaluation

The estimation scheme we proposed contains two main sections: 1) algebraic and 2) geometric minimization schemes. The AMS optimizes the calibration by fitting the parameters into the radial trifocal tensor constraints, which serve as initial conditions for the GMS.

This simulation was to verify the robustness of the proposed algorithm with respect to measurement noise. The geometric setup is identical to the simulation 5.5 and plotted in figure 5.8. 24 corresponding triplets were used to estimate the radial trifocal tensor and each image point on both a regular and omnidirectional camera was perturbed by Gaussian noise with a 0 mean and standard deviation (σ) varying from 1 pixels to 3 pixels with a 0.5 pixel spacing. The measurement errors provided by the gimbals are small enough to be neglected. Each different value of the noise was run 50 times and its average residual error and estimation error are plotted in figure 5.19 (a) and (b), respectively:

5.4.2 Pointing Stage

Once we calibrate the RTT and a pair of image points from the perspective and wide-angle camera, the transformation to find the transceiver vector, derived from equation 5.22 is:

$$x^{'q} = T^{q}_{ik} x^{i} x^{"k} \tag{5.33}$$

This gives us a simple procedure to determine the rotation angles of the gimbal. The only exception is when the target appears on the axial plane, where the RTT

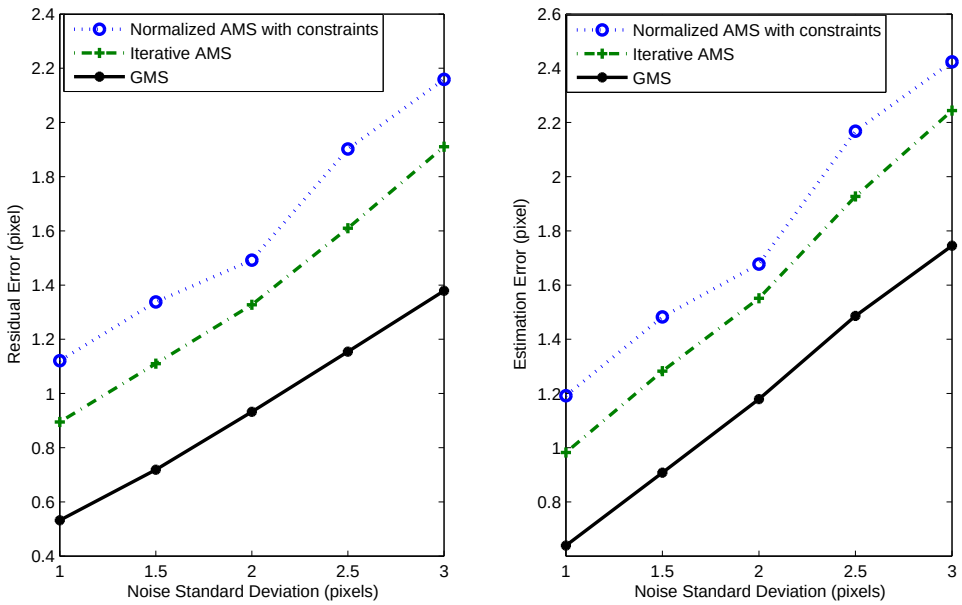

Figure 5.19: (a) Residual error (b)Estimation error with respect to different amount of measurement noise. The reported residual and estimation errors are an average of 50 independent calibrations.

is incapable of estimating the transceiver vector. This is the major price to pay for using an wide-angle camera.

5.4.2.1 Performance evaluation

The second simulation was to demonstrate the pointing results. The radial trifocal tensor was estimated from 24 corresponding triplets with a Gaussian noise of 0 mean and variance 1 pixel. The 24 triplets were selected to lie 5.7 degrees away from the distorted-epipole. After estimating the RTT, we produced another 200 points from both regular and omnidirectional cameras and compared the estimated transceiver vector with respect to the true transceiver vector. These 200 points were also perturbed by the same Gaussian noise used to produce the trifocal tensors.

The average error is 0.7366° for ϕ and 0.2381° for θ. The large difference in the azimuthal direction results from some of the 3D points being too close to the line vector.

For those degenerate points, its azimuth direction can hardly be determined correctly, which results in large errors. If the image points that are too close to the distorted-epipole ($\leq 2.866^\circ$) are removed, the average error becomes 0.2091° for ϕ and 0.1785° for θ. The histogram of the pointing error in both azimuth and elevation direction is plotted in figure 5.20.

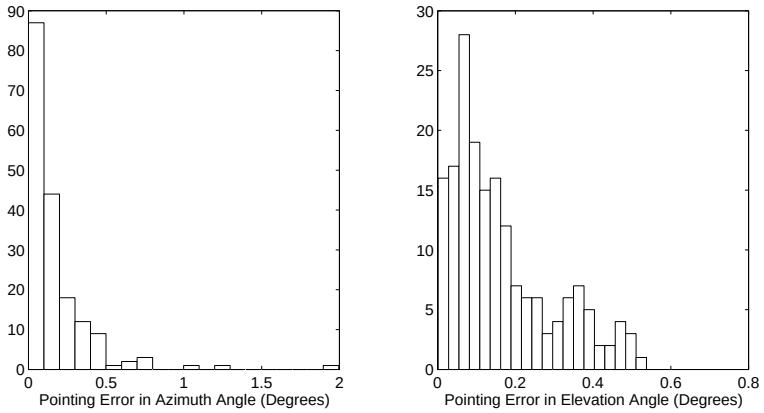

Figure 5.20: Pointing angular error in azimuth and elevation under $N(0, 1)$ noise

5.5 Summary

In this chapter, four CPAT system designs were introduced to fit different specifications. We discovered that the two-axis rotary gimbal and the perspective camera can both be modelled by the perspective projection, which further helps us utilize the projective geometry theorems to model the mapping between different frames. The resulting systems include the single-camera system encoded by a linear homography and the dual-camera system encoded by a trifocal tensor. Both the homography and the trifocal tensor can be calibrated in advance and robust estimation procedures are proposed in algorithms 5.1 and 5.6, respectively. Two optical-system-assisted calibration schemes are also listed to boost the precision of the calibration with respect to long link applications. The selecting criterion (between a single or dual camera system) is listed in equation 5.1, which depends on the link distance, the required angular resolution, and the distance between the transceiver and the camera. Both the simulations and experiments indicated satisfactory

performance of both systems.

One of the advantages of a camera-based system is that the mapping between the camera and the two-axis gimbal is invariant as long as their relative position remains the same. In other words, after the cameras and the gimbal are installed and the whole setup is mounted on a mobile platform (ex: cars, boats), the estimated homography or trifocal tensor only has to be calibrated once and will remain fixed despite the motion of the platform.

Our next focus is to increase the field-of-view of the camera in order to acquire more potential link targets. This idea is based on an observation that most of the cameras provide higher resolution than the required angular resolution of the CPAT system, about $1^\circ \sim 2^\circ$. Also, recent improvements in semiconductor fabrication have greatly increased CCD pixel density (e.g. 12 Mega pixels), which further enhances the angular resolution of the camera. The extra resolution can be traded for a large field-of-view by introducing a wide-angle lens, including dioptric and catadioptric imaging systems. Although the images from these cameras suffer from strong aberrations, since the distortion is static and only affects the radius of the image, it can be corrected by exploiting the geometric constraints embedded in an L-shape calibration object (algorithm 5.7) or by the assistance of a perspective camera (algorithm 5.8).

We can conclude that in a degenerate scenario, such as a planar motion or target is at a far distance, two-dimensional information must be provided in order to uniquely determine the mapping between the transceiver and its own frame, and in a general scenario, three dimensional information is required. This provides us

a relation to determine the minimum number of cameras in a camera-based CPAT system:

$$w + 2p \leq 2 \quad \text{degenerate applications} \tag{5.34}$$

$$w + 2p \leq 3 \quad \text{general applications} \tag{5.35}$$

where w represents number of the wide-angle cameras and p represents number of the perspective cameras. Notice that a corrected wide-angle camera is equivalent to a perspective camera.

Therefore, different camera combinations can be selected, such as the combination between a perspective camera and a wide-angle camera as proposed in section 5.4.

Chapter 6

Conclusions and Future Work

6.1 Summary of Main Contributions

Incorporating geometric invariances into a PAT system design provides three benefits: (1) more flexibility: no deliberate alignment is required, (2) faster acquisition time: the entire pixel array of the acquisition sensor is applicable to target pointing instead of one, and (3) reliable target identification: the target identification scheme can be simpler and more accurate because of the acquisition sensor being static.

If the targets of interest satisfy the distance inequality shown in equations 1.1 and 5.1 for FPAT and CPAT systems, respectively, the combination between one optical sensor and one actuator is sufficient to align the transceivers. Otherwise, two sensors must be applied to provide stereo vision for the pointing.

Through the geometric invariance exploration, three discoveries were made to assist in the PAT design:

1. The axis of propagation of the transceiver must pass through the pivoting point of the actuator, which leads to a gimbal-type design for CPAT systems, shown in figure 5.1 and FPM-based design for FPAT systems, shown in figure 3.3.

2. The aberrations of the optical systems must be minimized to provide a linear projective relationship between the incident angles and the focused positions, which leads to the selections of achromatic lenses from Optosigma and high-resolution imaging lenses from Tameron.

3. Additional optical systems can be designed to produce the desired invariance, even though it is not the one embedded in the current configuration.

The theoretical analyses of three different geometric invariances were presented, namely, linear homography, trifocal tensors, and radial trifocal tensors. The PAT system based on the linear homography can be applied to align the FSO transceivers with a long link distance, which includes most FSO applications; while the PAT system based on the trifocal tensor and radial trifocal tensor is capable of aligning the FSO transceivers without any constraint. Two-step calibration algorithms, algebraic and geometric minimization schemes, were proposed for estimating the three invariances. Simulations confirmed that the residual and estimation errors both approached the theoretical minimum, which demonstrated their robustness. A homography-based CPAT system, consisting of a camera and a two-axis rotary gimbal, was implemented and the pointing accuracy was around 700 μrad, which is on the order of the required precision for an FPAT system.

A homography-based FPAT system was also implemented inside our laboratory, which consisted of an achromatic lens, a lateral effect diode, and a two-axis moving platform with a fiber attached. Compared to the traditional FSM-based FPAT system, the system measures the incoming angles with a higher speed, which

helps in reducing the response time of the system. The average pointing accuracy shown in the experiment was around 170 μrad, while a real ray tracing simulation showed an accuracy around 45 μrad. The inferior performance may result from the self-coupling from the transmitter.

In addition, we demonstrated a distortion-free wide-angle camera with a field-of-view 140°, based on exploiting the geometric invariance (the planar angle is invariant between the camera and image frame). The average residual error after the correction is around 0.3° on a 640×480 pixel camera, which is close to the resolution limit of the camera.

6.2 Future Work

This dissertation is pioneering in its efforts to implement a PAT system capable of realizing a mobile, point-to-point FSO network. This work is just a start and much more effort is still required. Some suggestions are:

- Wavefront distortion evaluation: The current PAT system does not take turbulence effects into consideration. Since turbulence changes the refractive index of the optical paths, the position of the focused spot may change because of additional aberrations induced by the distorted wavefront. A field experiment, provided by Vilcheck [70], indicated the average angular fluctuation ranging from 13 to 25 μrad in the weak turbulence regime. More measurements must be taken to evaluate the pointing errors induced from additional aberrations.

- Integration between FPAT and CPAT systems: When the target is out of the

field-of-view of the FPAT system, the alignment must be handed over to the CPAT system. In the current system, the handover happens with a price that the link will be broken because the FPAT system must be reset. It can be proven that a geometric invariance also exists between the CPAT and FPAT system. If it is pre-calibrated, the FPAT system can drive both the CPAT and FPAT actuators to maintain the link when the handover occurs.

- Data fusion among different FSO transceivers: If many FSO transceivers, equipped with individual FPAT and CPAT systems, are deployed in an area, the acquisition sensors at each transceiver can be related by other invariances. By fusing the image positions from different PAT systems through their invariances, the pointing accuracy can be largely enhanced. Besides, since the fused data contains the coordinates of the target, it can be further integrated with other navigation services. This topic is ongoing research known as camera network calibration in the computer vision community.

Appendix A
Least Squares Optimization

Least squares problems are those in which the cost function may be expressed as the sum of squares. Such problems are very common in Euclidean geometry fitting and can be divided into the two categories as (1) Least Squares Regression and (2) Total Least Squares Regression.

A.1 Least Squares Regression

Least Squares Regression (LS) is the basic form of the least squares problem. If a set of measurements, y_n are gathered for different parameter values, x_n, the LS regression problem can be formulated as

$$\min_p \|y - Xp\| \tag{A.1}$$

where X is a matrix whose column vectors contain the basis vectors.

In general, the number of the equations must be more than the number of parameters, which implies that if X is $n \times m$, then $n > m$. X forms a set of basis to estimate y_n and if the optimizer is achieved, the resulting error is orthogonal to the basis, which is known as the orthogonality principle [54]. The solution is therefore equal to

$$\begin{aligned} p &= X^{\dagger} y \\ &= (X^T X)^{-1} X y \end{aligned} \tag{A.2}$$

where $(.)^{\dagger}$ represents the pseudo-inverse operator.

$(X^T X)^{-1}$ exists if (and only if) X is full rank, which implies the basis vectors are linearly independent. Therefore, the LS regression is appropriate.

A.2 Total Least Squares Regression

Total least squares (TLS) regression fits N-dimensional data with a subspace of dimensionality N-1. The result is equivalent to minimizing the square perpendicular distance from the data to the fitted function. The total least squares problem has the form

$$\begin{aligned} & min_p \|Mp\| \\ \text{subject to } & \|p\| = 1 \end{aligned} \tag{A.3}$$

If additional linear constraints $Cp = 0$ are applied to the TLS problem, it can be viewed as the TLS p is orthogonal to the kernel of the constraint. Therefore, its minimizer is the span of the complementary space of the kernel.

Let the SVD of C be USV^T and has rank r. The solution can be found in two steps:

Algorithm A.1 Unconstrained Optimization

1: Take the SVD of M equal to $U_M S_M V_M^T$.
2: The TLS minimizer p is the last column of V_M corresponding to the smallest singular value.

- Solve the minimizer of the following equation:

$$\min_{p'} \left\| M U_r^{\perp} p' \right\| \quad \text{subject to } p' = 1 \tag{A.4}$$

  where $U_r^{\perp}$ is equal to U without the first r columns.

- Retrieve p by $p = U_r^{\perp} p'$

If the linear constraint is $p = G\hat{p}$ where G has rank r, it represents the minimizer p lying in the span space of G. Let the SVD of the G be USV^T. Following a similar concept as the previous example, the minimizer can be formulated also in two steps as

Algorithm A.2 Linear Constrained Optimization

1: Solve the minimizer of the following equation:

$$\min_{p'} \left\| M U_r p' \right\| \quad \text{subject to } p' = 1 \tag{A.5}$$

   where U_r is equal to the first r columns of U.
2: Retrieve p by $p = U_r p'$

Bibliography

[1] J. Kovalik S. Piazzolla A. Biswas, M. W. Wright. Uplink beacon laser for mars laser communication demonstration (mlcd). *Proc. SPIE*, 5712, 2005.

[2] Anthony Acampora. Last mile by laser. *Scientific America*, Jul. 2002.

[3] P. Amirshai and M. Kaverhrad. Broadband access over medium and low voltage power-lines and use of white light emitting diodes for indoor communications. In *IEEE Consumer Communications and Networking Conference*, 2006.

[4] N. Andreff, R. Horaud, and B. Espiau. On-line hand-eye calibration. In *Second International Conference on 3-D Digital Imaging and Modeling*, 1999.

[5] N. Perez Arancibia, N. Chen, S. Gibson, and T. Tsao. Adaptive control of a mems steering mirror for free-space laser communications. *Proc. of SPIE*, 5892.

[6] J.A. Arnaud and H. Kogelnik. Gaussian light beams with general astigmatism. *Applied Optics*, 8(8), 1969.

[7] S. Baker and S. K. Nayer. A theory of catadioptric imaging formation. In *Proc. of 6th international conference on computer vision*, 1998.

[8] Hynek Bakstein and Tomas Pajdla. Panoramic mosaicing with a 180° field of view lens. In *Proc. OMNIVIS*, 2002.

[9] P.A. Belanger. Beam quality factor of the lp_{01} mode of the stepindex fiber. *Optical Engineering*, 32(9), 1993.

[10] Max Born and Emil Wolf. *Principles of Optics*. Cambridge University Press, 7th edition, 1999.

[11] J.B. Carruther and J.M. Kahn. Angle diversity for nondirected wireless infrared communication. *IEEE Transactions on Communications*, 48(6), 2000.

[12] Robert K. Crane. *Electromagnetic Wave Propagation Through Rain*. John Wiley & Sons, Inc, first edition, 1996.

[13] Christopher C. Davis. *Lasers and Eletro-Optics: Fundamentals And Engineering*. Cambridge University Press, first edition, 1996.

[14] R.C. Dorf and R.H. Bishop. *Modern Control Systems*. Prentice Hall, tenth edition, 2004.

[15] C. Eklund, Roger. B. Marks, Kenneth L. Stanwood, and Stanley Wang. Ieee standard 802.16: A technical overview of the wirelessman air interface for broadband wireless access. *IEEE Communication Magazines*, June, 2002.

[16] B. Epple. Using a gps-aided inertial system for coarse-pointing of free-space optical communication terminals. *Proc. SPIE*, 6304, 2006.

[17] A. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. *CVPR*, 1, 2001.

[18] C. Geyer and K. Daniilidis. Catadioptric projective geometry. *Int'l J. Computer Vision*, 45(3), 2001.

[19] A Emami-Naeini G.F. Franklin, J.D. Powell. *Feedback Control of Dynamic Systems*. Prentice Hall, fifth edition, 2006.

[20] M. L. Workman G.F. Franklin, J.D. Powell. *Digital Control of Dynamic Systems*. Addison Wesley Longman, third edition, 1998.

[21] A. Gibson. Laser pointing technology. *Proceedings. of SPIE*, 4034, 2000.

[22] R. I Hartley. In defense of 8-point algorithm. In *Proc. of the Fifth International Conference on Computer Vision*.

[23] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.

[24] T. Ho, C. C. Davis, and S. Milner. Pointing, acquisition and tracking system with omnivision. *Proc. of SPIE*, 5892:420–431, 2005.

[25] T Ho, C. C. Davis, and S. Milner. Using geometric constraints for fisheye camera calibration. In *Proc. of IEEE OMNIVIS Workshop*, 2005.

[26] T. Ho, C. C. Davis, and S. Milner. Fully optical real-time pointing, acquisition, and tracking system for free space optical link. *Proc. of SPIE*, 5712, Apr 2005.

[27] T Ho and C.C. Davis. Three-dimensional optical pointing system encoded by radial trifocal tensor. In *Proc. of SPIE*, volume 6304, 2006.

[28] http://en.wikipedia.org/wiki/Castigliano's_method.

[29] http://en.wikipedia.org/wiki/Hysteresis.

[30] http://en.wikipedia.org/wiki/Perspective.

[31] http://lasers.jpl.nasa.gov/PAGES/ground.html.

[32] http://www.vision.caltech.edu/bouguetj/calib_doc/index.html.

[33] C.H. Houpis J.J. D'Azzo. *Feedback Control System Analysis and Synthesis*. McGraw-Hill, second edition, 1966.

[34] K. Kagawa, T. Ikeuchi, J. Phta, and M. Nunoshita. An image sensor with a photoreceiver function for indoot optical wireless lans fabricated in 0.8um bicmos technogloy. *Proc. IEEE Sensors*, 2004.

[35] J. Kannala and S. Brandt. Estimation of omnidirectional camera model from epipolar geometry. In *ICPR*, volume 1, 2004.

[36] G. Keiser. *Optical Fiber Communications*. McGraw-Hill, third edition, 2000.

[37] A. Kilpela. Pulsed time-of-flight laser range finder techniques for fast, high precision measurement applications. *Master Thesis*, 2004.

[38] S. Kingsley and S. Quegan. *Understanding Radar Systems*. McGraw-Hill, London UK, 1992.

[39] A. S. Koujelev and A. E. Dudelzak. Three-stage tracking approach in development of a free-space laser communicator. *Proceedings. of SPIE*, 6304, 2006.

[40] K. Kuhnen. Modelling, identification and compensation of complex hysteretic nonlinearities a modified prandtl-ishlinskii approach. *European Journal of Control*, 9, 2003.

[41] Stephen G. Lambert and Williams L. Casey. *Laser Communication in Space*. Artech House, first edition, 1995.

[42] T.S. Low and W. Guo. Modeling of a three-layer piezoelectric bimorph beam withhysteresis. *Journal of Microelectromechanical Systems*, 4, 1995.

[43] Y. Ma, S. Soatto, J. Kosecka, and S.S. Sastry. *An Invitation To 3-D Vision*. Springer, 1st edition, 2004.

[44] Dimitris G. Manolakis, Vinay K. Ingle, and Stephen M. Kogon. *Statistical and Adaptive Signal Processing*. McGraw-Hill, first edition, 2000.

[45] F.M. Margaret. Perspective projection: the wrong imaging model. *Technical Report TR95-01*, 1995.

[46] Issak D. Mayergoyz. *Mathematical Models of Hysteresis*. Springer-Verlag, first edition, 1991.

[47] P.F. McManamon, J. Shi, and P.J. Bos. Broadband optical phase-array beam steering. *Optical Engineering*, 44(12), 2005.

[48] B. Micusik and T. Pajala. Estimation of omnidirectional camera model from epipolar geometry. In *CVPR*, 2003.

[49] S.G. Nash and A. Sofer. *Linear and Nonlinear Programming*. Mcgraw-Hill, 1st edition, 1996.

[50] V. Nikulin, J. Sofka, and R. Khandekar. Performance of free-space laser communication systems as a function of the sampling rate in the tracking loop. *Proc. of SPIE*, 5892.

[51] D. O'Brien and M. Katz. Optical wireless communications within fourth-generation wireless systems. *Journal of Optical Networking*, 4(6), 2005.

[52] Applied Research Defense Electronic Products Radio Corporation of America. *Laser Communication Transmitter Final Report.* NASA http://www.geocities.com/bobandrepont/geminipdf.htm, 1965.

[53] A. Polishuk, U. Peled, M. Nissim, M. Golub, and S. Arnon. Wide-range high-resolution optical scanner for optical wireless communication and laser radar: Laboratory experiment. *Proc. of SPIE*, 5892.

[54] H.V. Poor. *An Introduction to Signal Detection and Estimation.* Springer-Verlag, second edition, 1994.

[55] M. Ribo, A. Pinz, and A. L. Fuhrmann. A new optical tracking system for virtual and augmented reality applications, 2001.

[56] M. Sanchez and D. Gutow. Control laws for a three-element risley prism optical beam pointer. *Proc. of SPIE*, 6304, Sep. 2006.

[57] W. Saw, H. Rafai, and J. Sluss Jr. Free-space optical alignment system: Using gps. *Proc. SPIE*, 5712, 2005.

[58] V.A. Skormin, M.A. Tascillo, and T.E. Busch. Demonstration of jitter rejection technique for free space laser communication. *IEEE Transactions on Aerospace and Electronics Systems*, 33, Apr. 1997.

[59] Warren J. Smith. *Modern Optical Engineering: The Design of Optical Systems.* McGraw-Hill, 2nd edition, 1990.

[60] R.M. Sova, J.E. Sluz, D.W.Young, J.C. Juarez, A. Dwivedi, N.M. Demidovich III, J.E. Graves, M. Northcott, J. Douglass, J. Phillips, D. Driver, A. McClarin, and D. Abelson.

[61] Ferroelectrics Standards Committee of the IEEE Ultrasonics and Frequency Control Society. *An American National Standard: IEEE Standard on Piezoelectricity.* The Institute of Electrical and Electronics Engineers, ANSI/IEEE Std 176-1987, 1987.

[62] Tzung-Hsien Ho Jaime Llorca Sugianto Trisno Stuart D. Milner, Aniket Desai and Christopher C. Davis. Self-organizing broadband hybrid wireless networks. *Journal of Optical Networking*, 4(7), 2005.

[63] SriRam Thirthala and Marc Pollefeys. The radial trifocal tensor: A tool for calibrating radial disortion of wide-angle cameras. In *ICPR*, 2005.

[64] D.H. Titterton and J.L. Weston. *Strapdown Inertial Navigation Technology.* Lavenham Press Ltd, first edition, 1997.

[65] W. Triggs. Auto-calibration and the absolute quadric. In *ICPR*, 1997.

[66] S. Trisno. *Design And Analysis Of Advanced Free Space Optical Communication Systems*. Ph.D Dissertation, 2006.

[67] S. Trisno, T.H. Ho, S.D. Milner, and C.C. Davis. Theoretical and experimental characterization of omni-directional optical links for free space optical communication. In *Proc. IEEE MILCOM*, 2004.

[68] E. Trucco and A. Verri. *Introductory Techniques for 3-D Computer Vision*. Prentice Hall, first edition, 1998.

[69] R.Y. Tsai. A versatile camera calibration technique for high accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses, journal=.

[70] M.J. Vilcheck, A.E. Reed, H.R. Burris, W.J. Scharpf, C. Moore, and M.R. Suite. Multiple methods for measuring atmospheric tuebulence. *Proc. of SPIE*, 4821, 2002.

[71] T. Caudell W. Barfield. *Fundamentals of Wearable Computers and Augumented Reality*. Lawrence Erlbaum Associates, Inc, first edition, Dec. 2000.

[72] W. Wang and H.T. Tsui. A svd decomposition of essential matrix with eight solutions for the relative positions of two perspective cameras. In *ICPR*, 2000.

[73] L.M. Wascizko, I.I. Smolyaninov, and C.C. Davis. Analysis of compound parabolic concentrators and aperture averaging to mitigate fading on free space optics links. *Proc. of SPIE*, 5160, 2003.

[74] W.T. Welford. *Aberrations Of The Symmetrical Optical System*. ACADEMIC PRESS, 1st edition, 1974.

[75] W.T. Welford and R. Winston. *High Collection Nonimaging Optics*. Academic Press, first edition, 1989.

[76] Y.K. Wen. Method for random vibration of hysteretic systems. *Journal of the Engineering Mechanics Division*, 102, 1976.

[77] R. Wilson and S. Shafer. What is the center of the image. *Journal of the Optical Society of America*, 11(11):2946–2955, 1994.

[78] H.J. Woltring. Single- and dual-axis lateral photodetectors of rectangular shape. *IEEE Transactions on Electron Devices*, 22, 1975.

[79] Z. Hu X. Ying. Can we consider central catadioptric cameras and fisheye cameras within a unified imaging model. In *ECCV*.

[80] H. Yuksel, C.C. Davis, and S.D. Milner. Aperture averaging for optimizing receiver design and system performance on free-space optical communication links. *Journal of Optical Networks*, 2005.

[81] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.