Purdue University
Purdue e-Pubs

Open Access Theses | Theses and Dissertations

2013

Autonomous Orientation and Geolocation via Celestial Objects

Cheng Liu
Purdue University, sohpl2001@gmail.com

Follow this and additional works at: https://docs.lib.purdue.edu/open_access_theses

Part of the Electrical and Electronics Commons, and the Mechanical Engineering Commons

Recommended Citation
Liu, Cheng, "Autonomous Orientation and Geolocation via Celestial Objects" (2013). *Open Access Theses*. 33.
https://docs.lib.purdue.edu/open_access_theses/33

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

Graduate School ETD Form 9
(Revised 12/07)

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By CHENG LIU

Entitled

Autonomous Orientation and Geolocation via Celestial Objects

For the degree of Master of Science in Mechanical Engineering

Is approved by the final examining committee:

KARTIK B. ARIYUR
Chair

GALEN B. KING

XINYAN DENG

To the best of my knowledge and as understood by the student in the *Research Integrity and Copyright Disclaimer (Graduate School Form 20)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Approved by Major Professor(s): KARTIK B. ARIYUR

Approved by: David C. Anderson 09/17/2013
Head of the Graduate Program Date

AUTONOMOUS ORIENTATION AND GEOLOCATION VIA CELESTIAL OBJECTS

A Thesis

Submitted to the Faculty

of

Purdue University

by

Cheng Liu

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science in Mechanical Engineering

December 2013

Purdue University

West Lafayette, Indiana

To my parents,

for their selfless love and care.

To the stars shining in the sky,

for igniting my curiosity.

ACKNOWLEDGMENTS

I would like to sincerely thank my advisor, Dr. Kartik Ariyur, for helping with the signal processing algorithms and other theoretical issues; John Barnes, for doing a helpful pioneer work; Feichen Shen, for pointing out the extraordinary usage of camera blooming; Adam Willats, for helping me with the programming and image processing; Sunghun Jung, for helping me with the hardware installation and experiments of the solar compass.

TABLE OF CONTENTS

Page

LIST OF TABLES vi

LIST OF FIGURES vii

SYMBOLS ix

ABBREVIATIONS xi

ABSTRACT xii

1. INTRODUCTION 1

2. PRINCIPLES OF CELESTIAL NAVIGATION 8
   - 2.1 Geographic Coordinate System 8
   - 2.2 Equatorial Coordinate System 9
   - 2.3 Horizontal Coordinate System 10
   - 2.4 Navigation via Celestial Objects 11
     - 2.4.1 Celestial Orientation 11
     - 2.4.2 Celestial Geolocation 13
   - 2.5 Remarks on Celestial Navigation 16

3. HEMISPHERICAL SENSOR GEOMETRY AND SUN VECTOR ESTIMATION VIA LIGHT INTENSITY SENSORS 19
   - 3.1 Hemispherical Sensor Geometry 19
   - 3.2 Estimation of the Sun Vector via Light Intensity Sensors 22
   - 3.3 Rotation Matrices and Attitude Compensation 28
   - 3.4 Sensor Operation and Field Experiments 31
     - 3.4.1 Experiment for the Sun Vector Measurement 33
     - 3.4.2 Absolute Orientation via a Solar Compass 34

4. ESTIMATION OF CELESTIAL VECTORS VIA CAMERAS 39
   - 4.1 Camera Modeling and Calibrations 42
     - 4.1.1 Camera Modeling 42
     - 4.1.2 Camera Calibration and Undistortion 46
   - 4.2 Image Processing for Celestial Object Detection 48
     - 4.2.1 Sun Detection 48
     - 4.2.2 Moon Detection 57
   - 4.3 Experimental Results and Analysis 62
     - 4.3.1 Estimation of the Sun Vector 62
     - 4.3.2 Estimation of the Moon Vector 66

Page

4.3.3 Suggestions for Improvements 67

5. GENERALIZED INTERCEPT METHOD FOR GEOLOCATION 69
5.1 Generalized Intercept Method 70
5.2 Iterative Position Matching Algorithm 78
5.3 Applications and Tests 85
5.4 Uncertainty Analysis and Improvements 88

6. CONCLUSIONS AND FUTURE WORK 92
6.1 Conclusions 92
6.2 Future Work 93

LIST OF REFERENCES 94

LIST OF TABLES

Table		Page
3.1	α and ζ values for all sensing pixels.	21
4.1	Average RGB color values of all pixels in two areas.	59
4.2	Detection results and the corresponding sun vectors.	63
4.3	Intersection points of blooming lines with azimuth and zenith values.	64
4.4	Final result of the sun vector after weighting.	65
4.5	Detection results and the corresponding moon vectors.	66
4.6	Final result of the moon vector after weighting.	67
5.1	Geolocation inputs from Table 4.4.	85
5.2	Iterative steps for the geolocation of Table 5.1.	86
5.3	Possible source of uncertainties for celestial geolocation.	88
5.4	Second group of inputs for solar geolocation.	90

LIST OF FIGURES

Figure		Page
1.1	Pattern formed by three stars.	6
2.1	Geographic coordinate system.	8
2.2	Equatorial coordinate system.	9
2.3	Celestial reference vector of a celestial object in HCS.	10
2.4	Illustration of the celestial compass.	12
2.5	Illustration of the geolocation approach via coordinate transformations.	14
2.6	The intercept method and LOPs with respect to two different celestial objects.	16
3.1	Geometry of the hemispherical sensor.	20
3.2	Theoretical frequency distribution across sensor surface when $\alpha_s = 20°$, $\zeta_s = 75°$.	24
3.3	Three Euler angles of yaw (ψ), pitch (θ) and roll (φ).	29
3.4	Experiment result of the sun vector estimation.	34
3.5	Integrated UAV system with the solar compass.	35
3.6	Experiment result of three compass systems.	37
4.1	Imaging system of a camera.	39
4.2	Cameras on the hemispherical sensor, with one on the top and eight are evenly distributed at 45° zenith.	41
4.3	Illustration of all coordinate systems.	43
4.4	Images of the checkerboard used for calibration.	46
4.5	Corner extraction of the checkboard.	47
4.6	A sample image of the sun taken by camera #2.	49
4.7	Procedure for the sun and blooming detections.	50
4.8	Preprocessing step 1: RGB to binary image.	51

Figure Page

4.9 Preprocessing step 2: Eliminate small noises that are fewer than 500 pixels. 51

4.10 Preprocessing step 3: Eliminate undesired features and find the edges. . 52

4.11 Example Hough line and Hough matrix for the blooming feature. . . . 53

4.12 Detection result of Figure 4.6 via HT. 54

4.13 Sun and blooming detection results with disturbances. 56

4.14 Prepocessing of moon's images to eliminate disturbances. 58

4.15 Lunar phases in one cycle of lunar month. 60

4.16 Two possible results of the moon detection via HT. 61

4.17 Hemispherical sensor and other devices during the experiments. 62

4.18 Original images and the results of sun detection. 63

4.19 Inverse projection of blooming lines with azimuth and zenith angles. . . 64

4.20 Original images and the results of moon detection. 66

5.1 Simulated SOPs of the sun with $120^\circ \pm 1^\circ$ azimuth and $45^\circ \pm 1^\circ$ zenith at UTC time noon, March 20th, 2012. 72

5.2 SOPs when two overlapped areas exist. 72

5.3 Illustration of the sSOP. 73

5.4 Global sun vector distribution at UTC time noon, March 20th, 2012. . 79

5.5 Iterative point matching by narrowing the GoL. 81

5.6 Diagram of the iterative matching algorithm. 82

5.7 Example azimuth and zenith SOPs with a large overlapped area. 83

5.8 Global distribution of sun azimuth after compensating the magnetic declination. (UTC time at noon on March 20th, 2012) 84

5.9 Geolocation result of Table 5.1. 87

5.10 Geolocation results with different measurement errors. 87

5.11 Geolocation result of Table 5.4. 91

5.12 Intersected geolocation result of Table 5.1 (red) and Table 5.4 (blue). . 91

SYMBOLS

Ref	celestial reference
A	azimuth of a celestial reference
Z	zenith of a celestial reference
Dec_M	magnetic declination
$\mathbf{e_s}$	sun vector
$\mathbf{e_i}$	unit position vector of the ith pixel
α_i	azimuth angle of the ith pixel in sensor's spherical coordinate
ζ_i	zenith angle of the ith pixel in sensor's spherical coordinate
Φ	unit vector matrix of sensor pixels
I_s	solar irradiance
I_i	irradiation experienced by the ith pixel
μ	sensitivity gain of the light intensity sensor
f_i	output frequency of the ith pixel
F	frequency matrix of sensor pixels
(XYZ)	stationary world coordinates
(x_s, y_s, z_s)	sensor coordinates
(x_c, y_c, z_c)	camera coordinates
(O', u, v)	image coordinates
G	location of Greenwich, UK on Earth's surface
O	center of the Earth
P	arbitrary point on the Earth's surface
λ	geocentric longitude
ϕ	geocentric latitude
δ	geographic angular distance
h	geographic altitude from the sea level

t	time information
φ	Euler roll angle
θ	Euler pitch angle
ψ	Euler yaw angle
$\mathcal{R}_i(\theta)$	rotation matrix for rotating an angle of θ about the ith axis
$\mathbb{E}$	expectation operator
Var	variance operator
ε	error
(f, u_0, v_0, S_u, S_v)	camera parameters
$(k_{1,2\ldots}, p_{1,2})$	lens distortion parameters

ABBREVIATIONS

AHRS	Attitude and Heading Reference System
CCD	Charge-Coupled Device
CMOS	Complementary Metal Oxide Semiconductor
CRV	Celestial Reference Vector
Dec	Declination
DOF	Degree of Freedom
ECS	Equatorial Coordinate System
FOV	Field of View
GCS	Geographic Coordinate System
GMST	Greenwich Mean Sidereal Time
GoL	Grid of Locations
GNSS	Global Navigation Satellite Systems
GPS	Global Positioning System
GS	Geographic Size
HCS	Horizontal Coordinate System
HT	Hough Transform
LOP	Line of Position
PN	Position Nodes
RA	Right Ascension
SHA	Sidereal Hour Angles
SOP	Set of Position
sSOP	successive Set of Position
UAV	Unmanned Aerial Vehicle
UTC	Coordinated Universal Time
WGS84	World Geodetic System (1984 model)

ABSTRACT

Liu, Cheng M.S.M.E., Purdue University, December 2013. Autonomous Orientation and Geolocation via Celestial Objects. Major Professor: Kartik B. Ariyur, School of Mechanical Engineering.

Based on a hemispherical sensor geometry, a novel celestial navigation system is developed to use celestial objects to determine the absolute location and orientation information without the aid of satellites via two different approaches.

The first approach employs a hemispherical arrangement of light intensity sensors to determine the vector to the dominant light source. We present the sensing system to measure the sun vector via least squares method and achieve the application of a low-cost, small-sized solar compass. The system is shown to work well under ideal conditions but is susceptible to noise and uncertainties in some situations.

The second approach uses camera instead of light sensor, enabling the detection of celestial objects in a much more accurate and flexible fashion. An elaborate camera calibration was conducted to mitigate lens distortion and explore the transformation from image pixel coordinates to stationary world coordinates. With suitable image processing strategies, the system is able to use images of the sun and moon for the purpose of obtaining azimuth and zenith angles in spite of various disturbances.

Given the results measured with our sensing systems, a generalized geolocation method is presented to estimate the absolute location on the earth. The approach, inspired by the traditional manual intercept method, automates all of its steps in an iterative fashion. It derives both the geolocation estimates and the error intervals based on measurement noise levels. This method is superior to most traditional approaches in that it derives the estimates even with lower quality sensors.

1. INTRODUCTION

Navigation, in common usage, is to accurately determine:

- Location, such as longitude, latitude and altitude information.
- Orientation, either on local North-East coordinates or from an arbitrary datum.

Since ancient times, breakthroughs in navigation have corresponded to great expansion of commerce and exchange. Using natural aids of the earth's magnetic field, the sun or other bright celestial bodies, and local landmarks or plants, a navigator can apply navigation principles to arrive at his destination [1]. However, these methods via the natural aids suffered from various limitations: prerequisite of an experienced operator, limited data acquisition, low accuracy due to human operation, and influences of environmental uncertainties.

The development of Global Navigation Satellite Systems (GNSS), especially the Global Positioning System (GPS) has dramatically improved navigation methods over the past four decades [2]. The completion of the satellite constellation and reduction in the prices of receivers enabled diverse applications. Although the accuracy is influenced by the number of available satellites, the type and quality of the receiver, and the local environment, an uncertainty of ten meters in location measurements can often be guaranteed for ordinary usage [3]. Users with limited navigation skills can use GPS devices to easily determine location and orientation.

The great success of GPS has almost obscured all other navigation aids in civilian use. In colloquial terms, it is not much of an exaggeration to say that navigation is virtually synonymous with GPS [4]. However, the data acquired from GNSS depends upon the maintenance of a large number of operational satellites. Their relatively weak signals on the ground make them vulnerable to local electromagnetic disturbances [5] and may be easily jammed or spoofed by hostile forces [6]. The terrain

and physical nature of the local environment also affect the reception of a GPS receiver, and solar flares are known to strongly disrupt GPS signals [7]. In addition, because different GNSS systems such as GPS, Galileo and GLONASS are operated by different governments or organizations, interoperability and complementarity are difficult to achieve [8]. Nowdays, due to the ubiquity of GNSS in navigation markets, individuals without training are likely to get lost when GNSS navigation is no longer available. Hence, there is a need for an independent, world-wide and self-contained navigation system as a back up to GNSS [9]. Celestial navigation is perhaps the best candidate for this purpose.

Celestial Navigation is the art and science of navigating by celestial objects such as the sun, moon, stars and planets[1], and in one form or another, is one of the oldest operational skills in human history. Unlike man-made satellites, most astronomical objects are reliable, universal and persistent navigation aids when the local weather permits. They move in ways which can be calculated and predicted. Using accurate measurements of celestial objects, it is possible to determine geolocation and orientation. Repeated geolocation via celestial references allows the calculation of dynamic velocity, acceleration, etc. Celestial navigation skills have been taught since the Academy was established. Greek astronomers constructed geographic maps and estimated the circumference of the earth using celestial references [10]. "The Almagest", a comprehensive ancient treatise on astronomy, was compiled by Claudius Ptolemy with his astronomical models in convenient tables to compute the future or the past position of the planets [11]. However, in ancient times, celestial navigation techniques were only comprehended by a small group of experts and the accuracy was not great because of their incorrect geocentric model of the universe.

The Age of Discovery comes with the florescent of colonialism from the early 15th century to the 17th century. During that time, a navigator with proficient celestial navigation skills was treated as one of the most important prerequisites on a ship. Navigational instruments such as the compass, the astrolabe, the backstaff (Davis

[1]There are many other objects that can be observed, such as comets in space. But they cannot be used for navigation because they are either unpredictable or not persistent.

Quadrant) and the sextant aided the discovery and exploration [12]. The mariner's astrolabe and backstaff were in a form of inclinometer used to determine the latitude of a ship at sea by measuring the altitude of a celestial body, in particular, the sun or the moon. The sextant provided precision in measuring a celestial object's position relative to the horizon. Unlike the earlier sea astrolabe or backstaff, the sextant can be directly used to observe the stars due to its doubly reflecting optical designs, making it available at night for multiple targets [1]. Another remarkable accomplishment was the first completion of a global star chart. Johannes Hevelius proposed his Firmamentum Sobiescianum star atlas in 1690 with improved accuracy for the southern stars [13].

During the Renaissance, the geocentric system of Ptolemy was replaced by the heliocentric system of Copernicus, Galileo and Kepler [10]. Later, the law of universal gravitation, developed by Isaac Newton, directly linked physics and astronomy with neat formulas and acted as the cornerstone for astronomical computations. From the mathematical side, Carl Friedrich Gauss successfully predicted the future location of the dwarf planet[2] Ceres with an accuracy less than half a degree when it was rediscovered from the glare of the sun. His method was further developed as a theory for the motion of planetoids [14]. Systematic astronomical computation was then gradually established in a much more efficient and accurate framework. The development of mathematics and the formulation of classical physics enabled the tabulation of the ephemerides of a large number of astronomical objects [15, 16].

This model of our universe became mature in the 19th century and further improved in the early 20th century with the development of modern physics. Some special cases, e.g. anomalous precession of the perihelion of Mercury, were finally explained via the General Relativity. Given the time, a celestial object's position in the sky can be mathematically calculated at the observer's location, and vice versa, with a rather satisfactory accuracy at that time. However, the operation of traditional celestial measurement devices and the computation of astronomy algorithms still needed

[2] A celestial object with enough mass and orbiting around the sun, but not as huge as normal planets.

experienced specialists and did not provide measurements with the continuity and speed required for modern navigation systems.

From the mid 20th to the 21th century, autonomous celestial navigation appeared along with the explosive growth of modern sensor and information technologies. The accuracy and availability of measuring celestial objects have been significantly improved due to the technological developments of sensor hardware and computational devices. Compared to traditional navigation instruments, automated systems with electronic devices provide better accuracy, faster data acquisition and flexibility of integrating with other devices [4]. With more accurate sensing devices to measure celestial objects and faster computers to avoid the tedious astronomical computations, celestial navigation is theoretically accessible to ordinary users even with limited astronomical knowledge.

Autonomous celestial navigation is especially widespread in outer space, where no direct human operation is available. It has been used on satellites and spacecraft since the early age of space exploration. For example, many satellites use star sensors to determine the attitude [17]. Deep space explorers or land rovers may use star or sun sensors for location determination [18–20]. Some missiles are also using star positioning to tune the accuracy of the inertial guidance system after launch.

Digital imaging sensors such as the charge-coupled device (CCD) or complementary metal-oxide semiconductor (CMOS) are widely used for detecting celestial objects. The CCD sensor uses a biased PN junction to create a potential well that accumulates a charge proportional to the incident light and then drains into a register at the end of the integration period. The CMOS sensor, on the other hand, integrates a photodiode or photogate, a capacitor, and up to three transistors on a circuit, either accumulating or draining the charge on capacitors during the integration period [21, 22]. A number of studies have been performed to use CCD or CMOS to enable construction of advanced sun sensors or star sensors in miniature sizes [17, 23, 24].

The most common approach to a sun sensor is to use a mask with a number of small pinholes in front of a flat 2D array of CCD or CMOS pixels. Sunlight is allowed to pass through the pinholes and create a pattern of light and dark spots on the sensor. Since the focal distance and pattern of pinholes in the mask is known, the distribution of light and dark spots can be analysed to determine the sun's position.

Unlike the sun sensor, star trackers were developed for multi-object detection capability. Because the number of references that can be used for star navigation is usually more than one, it is necessary to have a process to autonomously identify different stars based on the brightness information and the pattern of separations. An algorithm provided by Liebe uses a selected star and two closest stars to form a triangular star pattern featured by the their angular distances and the intersection angle as shown in Figure 1.1 [25]. By checking this pattern within an existing star catalogue and database, all three stars can be ideally identified and used for navigation purposes. Other image processing algorithms like "grid algorithm" and "neural networks" were also applied in the area of automatic star pattern recognition. The usage of "binary search trees" and "Search Less Algorithm" also largely accelerated the speed of star pattern matching. Potential accuracy that can be achieved by a star tracker was analysed to be within 1 arcsecond (0.00028°) when a high resolution camera was used and statistical techniques were applied [17]. This is equivalent to about 30 meters on the surface of the earth, which is comparable to the standard positioning accuracy of GPS.

As earlier mentioned, celestial navigation failed to continue its prosperity in civilian market since the appearance of GNSS. Now, countless GPS embedded devices can be found on the market at reasonable prices while celestial navigation devices are either stored and exhibited, or limited in space exploration and military purposes. Little work has been done for its usage on ordinary mobile platforms such as aircraft, vehicles or personal devices.

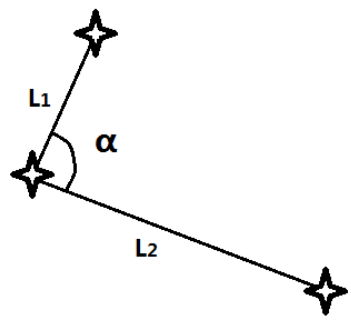


Figure 1.1. Pattern formed by three stars.

The disappearance of celestial navigation as a continuation of this trend is impossible from my point of view. This is because the following properties make it an irreplaceable navigation aid:

- Mature theoretical and practical basis are available to everyone.
- Persistent and global coverage when the weather permits, and applicability in outer space.
- Low cost and no need for maintenance.
- Not dependent upon any human-controlled factors, cannot be jammed and does not give off any signals.

It can be anticipated that in the future, celestial navigation will maintain its existence. The implementation of small-sized, low cost and fully autonomous celestial navigation devices will perhaps stimulate its future prospects.

Aiming to replicate the classical celestial navigation algorithms with modern technologies and achieve a reliable and accurate navigation system on various mobile platforms, I constructed a sensor platform that exploits the same principles for those classical navigational aids, and can potentially augment GNSS without the need for skilled operators.

The organization of this thesis follows a natural structure based on my research process. Chapter 2 introduces the basic astronomical knowledge needed for celestial navigation and some remarks on the applications of autonomous celestial navigation. Chapter 3 describes the design and construction of a hemispherical sensor structure for supporting digital sensors and cameras to measure celestial objects. A revised algorithm based on a previous work [26] is used to measure the sun vector under clear skies and tested via light intensity sensors. Chapter 4 replaces the light intensity sensor with cameras, offering a higher accuracy and more flexibility through image processing techniques. The camera calibration model and the object detection methods are also presented and tested experimentally for both sun and moon detections. Chapter 5 presents an approach of using the position of celestial objects to determine the location information via iterative comparisons and matches, together with the experimental results and uncertainty analysis. Conclusions are then drawn in Chapter 6 with some thoughts on the overall performance of the proposed hardware and navigation algorithms, as well as a potential roadmap for improvements of both.

2. PRINCIPLES OF CELESTIAL NAVIGATION

2.1 Geographic Coordinate System

The Geographic Coordinate System (GCS) is the most wildly used and well known coordinate system to indicate a location on the earth. It is commonly acknowledged as the only criterion on geographic maps and navigation applications. As shown in Figure 2.1, the GCS consists of three independent indicators: latitude, longitude and altitude. The equator is at 0 degree of latitude. All other latitude lines are parallel to the equator circle with 0° to 90° for the north latitudes and 0° to −90° for the south. Longitude lines are the vertical half circles connecting the north pole and the south pole. 0° locates at the Greenwich, England, increasing until 180° longitude to the east and −180° longitude to the west. The altitude is measured from the sea level with a positive value indicates a location higher than the sea. In GCS, any location on the earth is exclusively pinpointed by three fixed values.

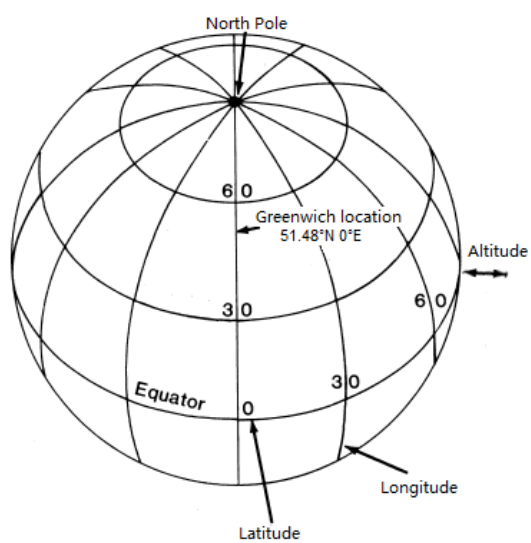


Figure 2.1. Geographic coordinate system.

2.2 Equatorial Coordinate System

The Equatorial Coordinate System (ECS) is an analogous system to GCS, specialized for specifying the positions of celestial objects [27]. An celestial object's position in ECS are often expressed in a pair of angles, Declination (Dec) and Right Ascension (RA) as shown in Figure 2.2. Through projecting the latitude and longitude lines of GCS directly onto the celestial sphere, latitudes correspond lines of declination and longitudes correspond lines of right ascension, or the "hour circles" in general expressions.

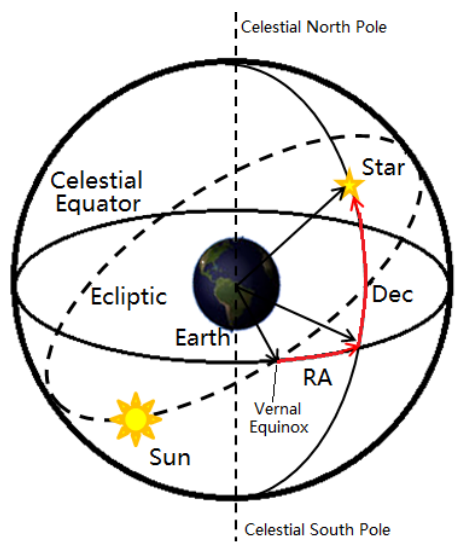


Figure 2.2. Equatorial coordinate system.

Declination indicates the angular distance measured perpendicularly northward(+) or southward(−) from the celestial equator, i.e. the direct projection of the earth's equator onto the celestial sphere. RA is the angle measured eastward from the vernal equinox[1] to the hour circle which passes the measured celestial body. Unlike declination which ranges from $-90°$ of the celestial south pole to $90°$ of the celestial north pole, RA is usually expressed in sidereal hours, minutes and seconds instead of degrees. This is to simplify the calculation of time when celestial objects cross the

[1]One of the two points where the celestial equator intersects the ecliptic.

meridian. An RA of 24h represents 360° eastward around the entire celestial equator thus 15° equals to one hour of right ascension.

There are more ways to express ECS besides the notation of Dec/RA. Instead of RA, Nautical Almanacs[2] use Sidereal Hour Angles (SHA), which is the explementary angle of RA (RA+SHA=360°) to indicate hour circles. Since the celestial sphere is considered to be infinite in radius, coordinates of ECS are assumed to be the same for all observers regardless the time and locations. For most stars, Dec/RA or Dec/SHA are nearly constant values[3] to locate a celestial body on the celestial sphere. For the sun, moon and planets, the values change along with their relative positions from the earth, but can always be mathematically determined.

2.3 Horizontal Coordinate System

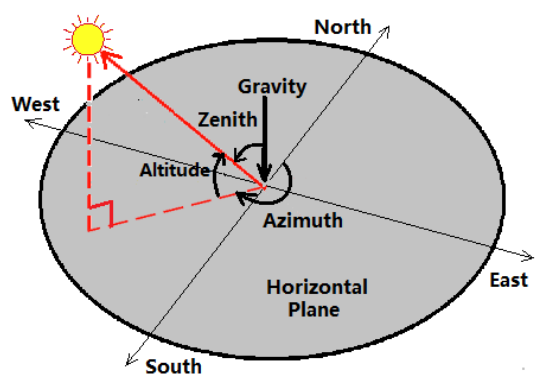


Figure 2.3. Celestial reference vector of a celestial object in HCS.

Since most celestial objects are far away from us, the light rays from them can be approximated as parallel to a high degree of accuracy. At a specific location, the Celestial Reference Vector (CRV) is defined as the normalized orientation vector which points to the celestial object and aligns with its light rays. In the Horizontal Coordinate System (HCS), the vector can be defined by two angles of azimuth and

[2]A book of astronomical tables published annually to specify the positions of the sun, moon, planets and selected stars on the celestial sphere during the whole year.
[3]Due to precession of the earth's axis, the Dec/RA of a star may change slowly

altitude, both in degrees [28]. As shown in Figure 2.3, the azimuth angle is measured around the horizon plane, usually measured from the geographic true north, increasing toward the east from 0° to 360°. The altitude ranges from −90° to 90° and indicates the vertical angle of objects from horizon. Zenith angle, the complementary angle of altitude, can also be used to represent higher altitudes with smaller values from 0° to 180°. To avoid the confusion of celestial altitude and geographic altitude, the rest of this thesis will use Azimuth/Zenith notation.

Unlike the previous two coordinate systems, HCS is fully constructed from the observer's perspective and varies from time to time, place to place. HCS is especially useful for indicating the motion of a celestial object in the sky. The change of zenith directly shows whether it's rising or setting. Since all celestial objects are subject to east-west diurnal motions on the celestial sphere, HCS also correlates their positions with orientation indicators.

2.4 Navigation via Celestial Objects

2.4.1 Celestial Orientation

Most of the time, the earth's magnetic field can provide a natural compass to determine orientations. But magnetic sensors used near the ground are susceptible to interferences from iron in buildings, iron bearing minerals and electrical appliances. In addition, there is a difference referred as the Magnetic Declination (Dec_M) between the geographic true north and the magnetic north because the true north pole (90°N) and the magnetic north pole (85.9°N 147.0°W in 2012) are at different locations. The value of Dec_M varies from place to place and can be more than ten degrees at some locations. The distribution of magnetic declination also keeps changing slowly [29]. Considering those facts, it is very difficult to determine the absolute orientation using merely a magnetic sensor.

Orientation via celestial objects can perform as an alternative for modern orientation systems. The principle of celestial orientation is rather straightforward and has

been used by human beings for thousands of years. It is a common sense that almost all celestial objects rises from the east and sets to the west due to rotation of the earth, following a predictable schedule and path. It has been researched that some animals are also able to use celestial objects to determine their migration directions [30].

The easiest way to do celestial orientation is to find the north celestial pole or the south celestial pole in ECS since these two fixed points are direct projections of the geographic north and south poles of GCS. In the northern hemisphere, Polaris, commonly the North Star, can be used to find the approximate north celestial pole[4]. Since the southern sky lacks an easily targeted polar star, the Crux constellation can be used by tracing a line from Gacrux to Acrux to a point close to the southern celestial pole. However, aids from stars are not applicable under bright skies during the daytime.

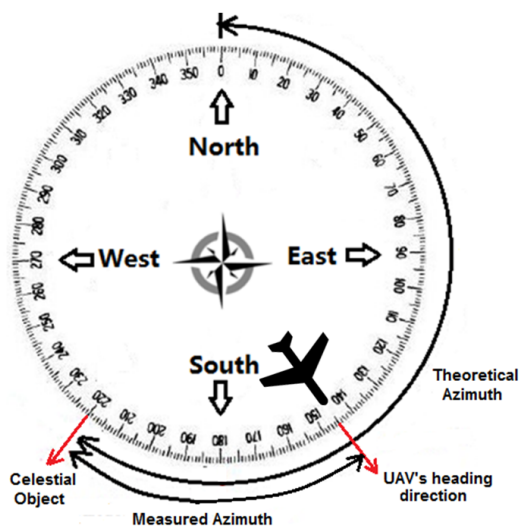


Figure 2.4. Illustration of the celestial compass.

The issue can be solved by a celestial compass since HCS indicates the position of a celestial object with respect to local orientations. If an observer on the earth knows his accurate location information (latitude, longitude and altitude) and exact time, the theoretical azimuth value of an observable celestial object in HCS coordi-

[4]Polaris is about 0.7° away from the ture north pole.

nates can be accurately calculated. On the other hand, a relative azimuth angle can also be estimated between the celestial object and the observer's heading direction through measurements. A comparison between those two azimuth angles yields the orientation. For example, as shown in Figure 2.4, if the theoretical azimuth angle of a celestial object is 220° clockwise to the true north and a UAV measures the relative azimuth angle as 75° clockwise to its heading direction. The observer is thus facing 220°- 75°=145° clockwise to the north.

In Chapter 3, we will apply the principle and present a low-cost, small-sized solar compass that can potentially augment the heading estimation of a UAV when it flies around the areas with large electromagnetic disturbances.

2.4.2 Celestial Geolocation

Celestial geolocation is the method of using celestial objects to determine the geographic location of a point on the earth, usually expressed in GCS notations of latitude, longitude and altitude. In other words, it is the method of using positions of celestial objects in ECS and HCS to locate a point in GCS. Since altitude can be easily measured with an altimeter, this thesis will focus on determining latitude and longitude values.

The relationship between the three coordinate systems is shown in Figure 2.5 [22]. The center of the earth is denoted as $\mathbf{O}$. $\mathbf{G}$ is the position of Greenwich, England at 0 hour Universal Time and the origin of longitude. $\mathbf{P}$ is an arbitrary point on the earth's surface and located by that point's latitude ϕ and longitude λ in GCS coordinates.

Assuming a celestial reference $\mathbf{Ref}$, two CRVs $\mathbf{Ref}^H$ and $\mathbf{Ref}^E$ can be used to indicate the position of $\mathbf{Ref}$ in HCS and ECS respectively. $\mathbf{Ref}^H$, the vector measured

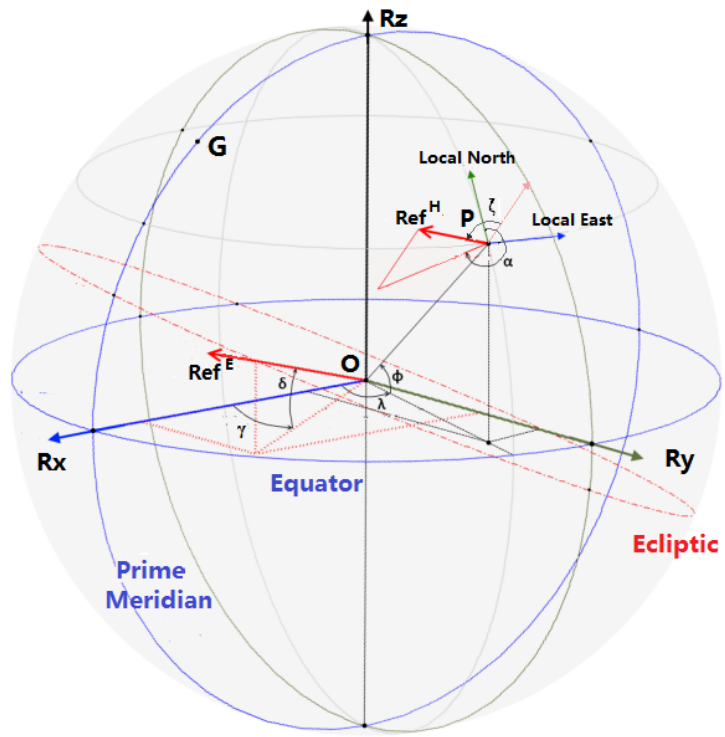


Figure 2.5. Illustration of the geolocation approach via coordinate transformations.

in local HCS, can be expressed in a Cartesian form by using the angles of azimuth $\alpha_{\mathbf{Ref}}$ and zenith $\zeta_{\mathbf{Ref}}$ as,

$$\mathbf{Ref}^H = \begin{pmatrix} \cos(\alpha_{\mathbf{Ref}})\sin(\zeta_{\mathbf{Ref}}) \\ \sin(\alpha_{\mathbf{Ref}})\sin(\zeta_{\mathbf{Ref}}) \\ \cos(\zeta_{\mathbf{Ref}}) \end{pmatrix}. \tag{2.1}$$

In ECS, the same celestial reference can also be specified by the right ascension RA and declination δ. Defining a relative angle γ between the hour circle of **Ref** and the projection of prime meridian[5] on the celestial sphere such that $\gamma = RA - GMST$, where $GMST$ is the Greenwich Mean Sidereal Time that can be found precisely in

[5]Prime meridian is the line of locations where longitude is 0°.

almanacs. The vector from the center of the earth $\mathbf{O}$ to the celestial reference $\mathbf{Ref}^E$ thus can be expressed in the Cartesian form as,

$$\mathbf{Ref}^E = \begin{pmatrix} \cos(\gamma)\cos(\delta) \\ \sin(\gamma)\cos(\delta) \\ \sin(\delta) \end{pmatrix}. \tag{2.2}$$

From $\mathbf{Ref}^E$ to $\mathbf{Ref}^H$, the transformation involves at least two rotations, one rotates by latitude ϕ and the other one rotates by longitude λ as,

$$\mathbf{Ref}^H = \mathcal{R}_{(\phi,\lambda)}\mathbf{Ref}^E. \tag{2.3}$$

Thus it is possible to use the local measurement of $\mathbf{Ref}^H$ and the theoretical value of $\mathbf{Ref}^E$ to mathematically calculate the latitude and longitude values.

The "Intercept Method" by finding the "Line of Position (LOP)" was first discovered in 1837 by Thomas Hubbard Sumner and later developed by Marcq St Hilaire. It was considered as a much more straightforward geolocation approach and employed by sailors for a long period of time [31]. The basic idea of the method is illustrated in Figure 2.6. Given the estimated zenith angle of a celestial object at an unknown location, it is most likely to find other places where the same angle can also be estimated since the earth has a spherical shape. With aid of a Nautical Almanac, a LOP can be drawn on the map to indicate all these selected locations [32]. With more LOPs from other celestial objects, the desired location can be directly yielded as the intersection of two or more such LOPs. Usually, two circular LOPs may intersect at two separated locations that are far away from each other. The correct solution, with both latitude and longitude values, can be determined with other apparent criteria, i.e. on the sea or land, northern or southern hemisphere. With more LOPs, the method will ideally yield a unique solution and the accuracy can be further improved via statistical analyses.

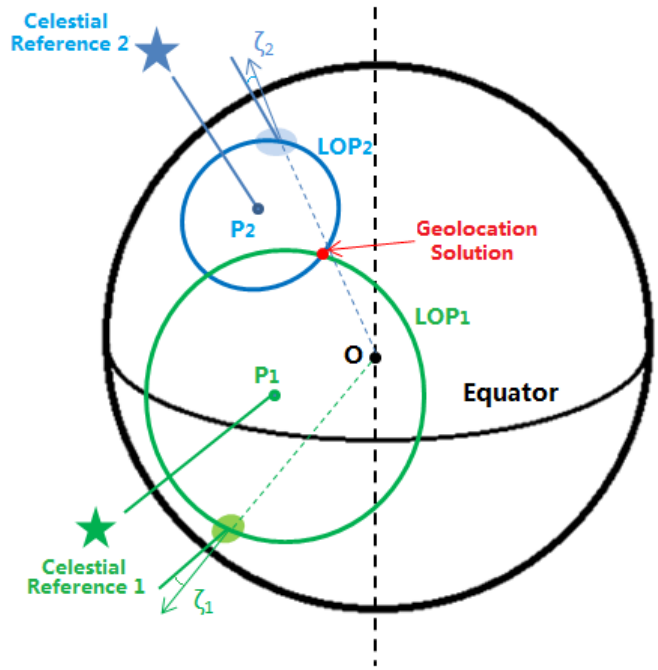


Figure 2.6. The intercept method and LOPs with respect to two different celestial objects.

In practice, the intercept method uses a celestial object's zenith rather than azimuth angle to draw the LOP. This is because most navigational instruments, e.g. sextant, perform best for measuring altitude or zenith angles. The measurement of azimuth, however, needs the exact knowledge of north direction which cannot be always guaranteed. Thus an apparent disadvantage of the method is that it can only be applied when there are two or more observable celestial objects.

In Chapter 5, a generalized intercept method will be presented to implement the autonomous celestial geolocation under more circumstances and to get better accuracies.

2.5 Remarks on Celestial Navigation

Celestial navigation is subject to kinds of uncertainties. It is necessary to point out some remarks on celestial navigation that should be noticed.

- Good weather condition is apparently needed to achieve a better celestial navigation. A cloud cover or bad weather may prohibit celestial observations by a day or more. But since most commercial jets cruise ranging from 30,000 to 37,000 feet[6], the influence of clouds and extreme weather happens at low chance to enable stable celestial navigations upon this height.

- Due to the tremendous size of the earth (a mean radius of 6,371 km), 1°'s error in the measurement of latitude or longitude implies about 111 km of distance error on the ground. Thus celestial navigation, especially geolocation, has a high demand on the measurement accuracy.

- Celestial navigation requires exact time information with down to the seconds. The accuracy of time directly influence the result of celestial navigation. Time measured from a marine chronometer or a quartz clock is subject to the accumulated error along with the usage (about 5 to 20 seconds per year) and should be synchronized timely with an atomic clock. To be precisely defined, the time frame should be in the form of Coordinated Universal Time (UTC). UTC is the primary time standard by which the world regulates clocks and time [33].

- The atmospheric effects, especially the atmospheric refraction [34], influence the observations of celestial objects. The worst case happens when the sun has a small altitude during the dawn and sunset. Celestial objects that are very close to the horizon may appear about half a degree higher than they would if there were no atmosphere. The simulated model for compensating their influences is available for most common weather conditions [22].

- The stars and constellations are usually assumed to be static references without any change in ECS. However, because the earth's orientation axis is not quite fixed, the precession causes a slow but continuous turning of the celestial poles and changes the positions of stars in a period of about 26,000 years or 1° every

[6]Military aircraft are able to fly considerably higher to an altitude of 50000-80000 feet.

72 years [35]. Additionally, the stars themselves keep moving far away or closer from each other. Their projections on the celestial sphere, i.e. the constellations, may look in very different shapes after a long time scale [36]. Thus applications of celestial navigation need timely corrected almanacs and database to maintain the accuracy of celestial references.

- The shape of earth cannot be easily assumed as a perfect sphere with equal radius everywhere. It actually approximates an oblate spheroid flattened along the axis and bulged around the equator. The erroneous usage of a simplified model will introduce noticeable errors in celestial navigation. The earth is more accurately represented in the World Geodetic System (WGS84) [37], the same model used by GPS. It uses an ellipsoid shape of the earth with a major radius of 6,378,137 m at the equator and an oblateness of $\frac{1}{298.257223563}$. The error is believed to be within half a meter.

- The usage of HCS needs a well defined horizontal plane, either geometric or gravitational, to estimate the CRV of difference celestial references. The horizontal plane sometimes is simply assumed to be perpendicular to local gravity vector. However, in reality, due to the spheroid shape of the earth, uneven distribution of underground materials, and gravitational effects from the moon, the sun or other planets, horizontal plane measured from a plumb or accelerometer devices may not be parallel to the actual horizontal plane at that location. To solve this, the earth's gravitational model and its effects are included in the newest version of WGS84 for more accurate compensations.

3. HEMISPHERICAL SENSOR GEOMETRY AND SUN VECTOR ESTIMATION VIA LIGHT INTENSITY SENSORS

Most traditional sun sensors and star trackers use a planar surface containing photosensitive elements to sense light from celestial objects [23, 24]. The sensing ability of these sensors is restricted by the limited field of view (FOV) and the optical aberrations due to simple, few-component imaging optics. A wider lens, though produces a larger FOV for processing, also compromises the image quality due to worse distortion and a curved focus plane [38].

Different from the traditional sense of a flat sensor geometry, a hemispherical structure with a number of distributed sensors yields a wider FOV of up to 360° sensing ability. Each sensor on the hemispherical surface can be treated as a "pixel" to receive signals independently. Current sensor manufacturing technologies have enabled such applications for the electronic eye cameras [39]. A previous work from John Barnes proposed a novel hemispherical sensor geometry to determine the sun's position upon certain idealized assumptions [26]. In this chapter, based on his work, the real sensor structure is designed and constructed for practical usage [40]. The least squares method is applied to use light intensity sensors to determine the CRV of a dominant light source, e.g. the sun. An iterative sensor operational procedure is also proposed to improve the sensor's performance. Using the sun vector, a low-cost, small-sized solar compass is designed and compared with GPS compass and magnetic compass to improve the orientation performance of a UAV [41].

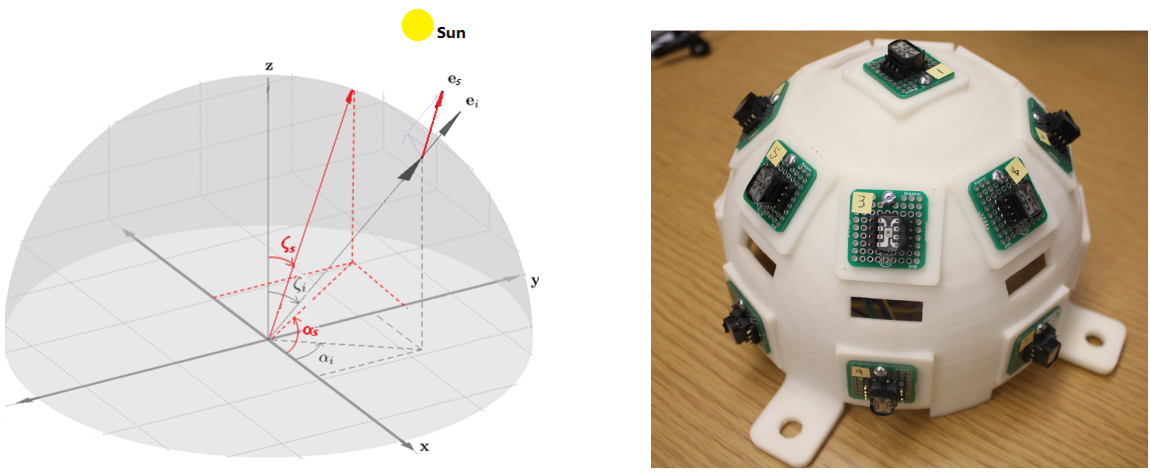
3.1 Hemispherical Sensor Geometry

The origin of a hemispherical sensor is defined as the center of the circle which constitutes the hemisphere's base. The radius of the sensor is denoted by R and there

are total N sensing pixels mounted on the hemispherical sensor's surface. Each pixel can be modeled as a very small flat surface with a unit vector $\mathbf{e}_i$ perpendicular to its surface. Since the radius of the hemispherical sensor is fixed, the location of the ith pixel in the sensor's spherical coordinate system can be specified by two angles α_i and ζ_i (Figure 3.1(a)).

Definition 1 *The pixel unit vector* $\mathbf{e}_i$ *is the normalized position vector pointing from the origin of the hemisphere to the* i^{th} *pixel directly, specified by the angles* α_i *and* ζ_i *in a Cartesian form of,*

$$\mathbf{e}_i(\alpha_i, \zeta_i) = \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \begin{pmatrix} \cos(\alpha_i)\sin(\zeta_i) \\ \sin(\alpha_i)\sin(\zeta_i) \\ \cos(\zeta_i) \end{pmatrix}. \tag{3.1}$$



(a) Sensor's spherical coordinate system.

(b) Hemispherical sun sensor with 17 sensing pixels.

Figure 3.1. Geometry of the hemispherical sensor.

To realize the above sensor geometry, a hemispherical sensor structure with a radius of 80 millimeters was designed and built by a 3D printer as shown in Figure 3.1(b). The main frame has 17 mounting positions to install light intensity sensors or video cameras. There are also four supporting stands for connecting the sensor

with mobile platforms. The inside hollow can be used for arranging wires, microcontroller and other sensors. The ideal unit vectors of all pixel positions in the spherical coordinate system are listed in Table 3.1.

Table 3.1. α and ζ values for all sensing pixels.

Sensor No.	α	ζ	Sensor No.	α	ζ
1	Undefined	0°	10	0°	90°
2	0°	45°	11	-45°	90°
3	-45°	45°	12	-90°	90°
4	-90°	45°	13	-135°	90°
5	-135°	45°	14	-180° or 180°	90°
6	-180° or 180°	45°	15	135°	90°
7	135°	45°	16	90°	90°
8	90°	45°	17	45°	90°
9	45°	45°			

The light sensor we employed is the TSL230R high-resolution programmable light-to-frequency converter with sensitivity and output scale adjustments[1]. Ideally, it outputs a square wave (50% duty cycle) with its frequency directly related to the light intensity.

It needs to be clarified that the sensor structure doesn't necessarily need to be an exact hemisphere. As long as the unit vector $\mathbf{e}_i$ of a sensing pixel is known for sure, there is always an equivalent point can be found in the above sensor's spherical coordinate system.

[1] `http://www.ams.com/eng/Products/Light-Sensors`

3.2 Estimation of the Sun Vector via Light Intensity Sensors

As illustrated in Section 2.3, a celestial object's position in HCS can be expressed by two angles of azimuth and zenith. A time and location-varying CRV which is defined as the sun vector is exploited to indicate the position of the sun in HCS and correlate the sun's position with orientations.

Definition 2 *The sun vector* $\mathbf{e}_s(A, Z)$ *is the instant angular measurement of the sun's center in the local Horizontal Coordinate System, and is defined by two angles,*

- *Sun azimuth A, the angle between the projection of the sun's rays on the horizon plane and the* ***True North****. It varies over the range of [0°, 360°],*
- *Sun zenith Z, the angle between the sun's rays and the local vertical (defined as the negative direction of the acceleration due to gravity at that point). It varies over the range of [0°, 180°], with* $Z > 90°$ *below the horizon.*

At any time and location, all direct solar rays are antiparallel to the instantaneous sun vector.

Assuming that the sensor structure is subjected to direct solar irradiation, the sun's position in the sensor's spherical coordinate system can also be expressed as angles of α_s and ζ_s without loss of generality (Figure 3.1(a)),

$$\mathbf{e_s}(\alpha_s, \zeta_s) = \begin{pmatrix} \cos(\alpha_s)\sin(\zeta_s) \\ \sin(\alpha_s)\sin(\zeta_s) \\ \cos(\zeta_s) \end{pmatrix}. \tag{3.2}$$

On the hemispherical structure, a pixel located at (α_s, ζ_s) will experience greater irradiation than any other pixels because its unit vector is parallel to the sun vector (i.e. it faces the sun directly). On the other hand, a second pixel offsets from the first pixel by 90° in any direction will experience zero solar irradiation because its unit vector is perpendicular to the sun vector. Any pixel between these two pixels

will experience an intermediate level of irradiation. We can obtain the angles of (α_s, ζ_s) by exploiting this principle in one or more ways using a finite number of sensors.

Definition 3 *The incident irradiance I_i is the scalar value of direct irradiation experienced by i^{th} pixel as,*

$$I_i = I_s \max\big(\mathbf{e}_i \cdot \mathbf{e}_s, 0\big), \tag{3.3}$$

where I_s is the instant local solar irradiance when directly measured toward the sun (i.e. the maximum irradiance at (α_s, ζ_s)), scaled in W/m^2 to show the power of electromagnetic radiation per unit area. The max *operator is used to ensure that the unilluminated pixels are not experiencing negative irradiance, an impossible result.*

Assuming a constant sensor sensitivity of μ (Hz per W/m^2) for all light intensity sensors, the theoretical frequency output of ith pixel is given by,

$$f_i = \mu I_s \max\big(\mathbf{e}_i \cdot \mathbf{e}_s, 0\big). \tag{3.4}$$

When the number of pixels $N \to \infty$, the entire surface of hemispherical sensor becomes saturated with pixels and Equation (3.4) may be plotted as a continuous surface with all azimuth and zenith values. An example distribution map when $\alpha_s = 20°$ and $\zeta_s = 75°$ is shown in Figure 3.2. Note that the maximum frequency (assuming no pixel saturation) occurs at (α_s, ζ_s).

When N ($N > 3$) pixels with known unit vectors are illuminated by the direct solar radiation, the number of equations needed to calculate α_s and ζ_s is over determined. Thus the least squares method is employed to minimize the influence of possible disturbances.

In order to account for the influence of noise and disturbances, Equation (3.4) must be augmented by adding a measurement error ε_i as,

$$f_i = \mu I_s \max\big(\mathbf{e}_i \cdot \mathbf{e}_s, 0\big) + \varepsilon_i, \tag{3.5}$$

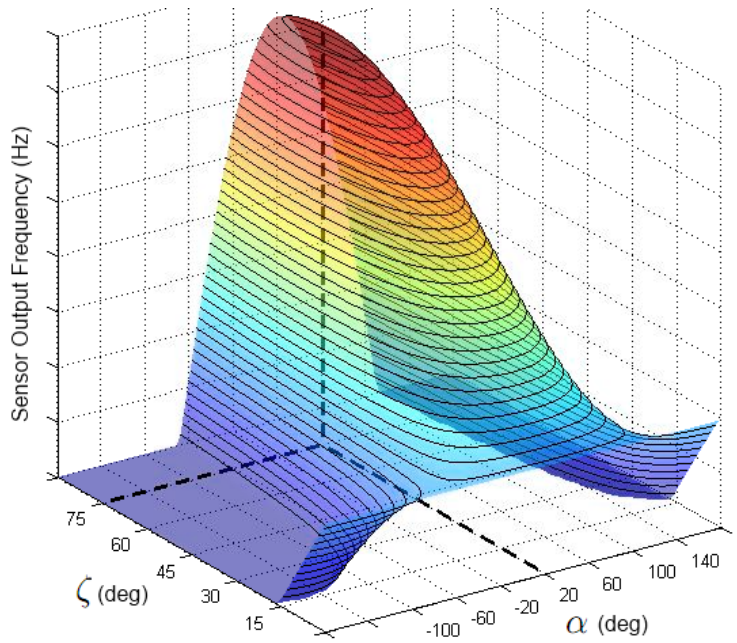


Figure 3.2. Theoretical frequency distribution across sensor surface when $\alpha_s = 20°$, $\zeta_s = 75°$.

where $i = 1, 2, ..., N$. For sensors with positive outputs, the dot product of two vectors in Equation (3.5) can be rewritten into a Cartesian form as,

$$\begin{aligned} f_i(\alpha_s, \zeta_s) &= \mathbf{e}_i^T \cdot \mathbf{s} + \varepsilon_i \\ &= x_i x_s + y_i y_s + z_i z_s + \varepsilon_i, \end{aligned} \tag{3.6}$$

where

$$\mathbf{s} = \begin{pmatrix} x_s \\ y_s \\ z_s \end{pmatrix} = \mu I_s \mathbf{e_s} = \mu I_s \begin{pmatrix} \cos(\alpha_s)\sin(\zeta_s) \\ \sin(\alpha_s)\sin(\zeta_s) \\ \cos(\zeta_s) \end{pmatrix},$$

which is the resized sun vector in sensor's coordinate system by a magnitude of μI_s.

If the sampling of all sensing pixels can be finished almost simultaneously, the resized sun vector (x_s, y_s, z_s) can be approximated as a constant for all sensing pixels during one measurement of $f_1, f_2, ..., f_N$. Equation (3.6) can be stacked into a matrix form to represent the output frequencies for all pixels as,

$$\mathbf{F} = \mathbf{\Phi s} + \varepsilon, \tag{3.7}$$

where

$$\mathbf{F} = \begin{pmatrix} f_1 \\ \vdots \\ f_N \end{pmatrix}, \mathbf{\Phi} = \begin{pmatrix} \mathbf{e}_1^T \\ \vdots \\ \mathbf{e}_N^T \end{pmatrix} = \begin{pmatrix} x_1 & y_1 & z_1 \\ \vdots & \vdots & \vdots \\ x_N & y_N & z_N \end{pmatrix}, \quad \text{and} \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{pmatrix}.$$

Proposition 3.2.1 *Sun Vector Estimation via Least Squares.*
With adequate measurement of frequency matrix $\mathbf{F}$ *and unit vector matrix* $\mathbf{\Phi}$*, the best linear estimation of the resized sun vector* $\overline{\mathbf{s}}(\overline{x}_s, \overline{y}_s, \overline{z}_s)$ *is given by,*

$$\overline{\mathbf{s}} = (\mathbf{\Phi}^T\mathbf{\Phi})^{-1}\mathbf{\Phi}^T\mathbf{F}, \tag{3.8}$$

with

$$\mathbb{E}[\overline{\mathbf{s}}] = \mathbf{s} \tag{3.9}$$

and

$$Var[\overline{\mathbf{s}}] = (\Phi^T\Phi)^{-1}\sigma^2 \tag{3.10}$$

and

$$\sqrt{N}(\overline{\mathbf{s}} - \mathbf{s}) \to \mathcal{N}(0, (\Phi^T\Phi)^{-1}\sigma^2), \tag{3.11}$$

provided the following conditions are satisfied:

- *The measurement error has zero mean,* $\mathbb{E}[\varepsilon] = 0$.
- *The measurement error has finite variance,* $Var[\varepsilon] = \sigma^2 < \infty$.

Proof The result follows directly from the Gauss-Markov Theorem [42] which states that an ordinary least squares estimator is the best linear unbiased estimator (BLUE).

■

Remark 3.2.1 *From Equation (3.11), it is clear that when there are more sensor pixels on the hemispherical structure, i.e. when N is larger, the deviation between the estimator* $\bar{\mathbf{s}}$ *and the true value of* $\mathbf{s}$ *has a smaller variance. So in practical use, a sufficiently large number of sensor pixels is preferred.*

Corollary 3.2.2 *Given* $\bar{\mathbf{s}}$*, the best estimations of* $\overline{\alpha}_s$ *and* $\overline{\zeta}_s$ *in sensor's coordinate system are given by,*

$$\begin{aligned} \overline{\alpha}_s &= \operatorname{atan2}(\overline{y}_s, \overline{x}_s), \\ \overline{\zeta}_s &= \operatorname{atan2}\left(\sqrt{\overline{x}_s{}^2 + \overline{y}_s{}^2}\ ,\ |\overline{z}_s|\right)\ . \end{aligned} \tag{3.12}$$

where the special arctangent function $\operatorname{atan2}(y, x)$ *is defined as,*

$$\operatorname{atan2}(a, b) = \begin{cases} \arctan(\frac{a}{b}) & (b > 0) \\ \arctan(\frac{a}{b}) + 180^\circ & (a \geq 0, b < 0) \\ \arctan(\frac{a}{b}) - 180^\circ & (a < 0, b < 0) \\ 90^\circ & (a > 0, b = 0) \\ -90^\circ & (a < 0, b = 0) \\ \text{undefined} & (a = 0, b = 0). \end{cases} \tag{3.13}$$

The only two unknown coefficients needed for the estimation are the magnitude of solar irradiance I_s and the sensor sensitivity coefficient μ. As shown in Equation (3.13), both I_s and μ only influence the magnitude of resized sun vector and do not influence the angles of $\overline{\alpha}_s$ and $\overline{\zeta}_s$. So our approach does not require the knowledge or calibration of I_s or μ.

For a hemispherical sun sensor with many pixels, an important consideration is the minimum number of pixels necessary to operate a sun sensor with a desired degree of accuracy. The thought leads to the following corollary.

Corollary 3.2.3 *At any given time, the signals from any three illuminated pixels are adequate to estimate the sun vector, provided the following conditions are satisfied:*

1. *The three pixels are illuminated only by the sun's direct radiation.*

2. *Few noise or interference sources are present.*

3. *The plane containing the three pixels does not intersect the origin of the sensor.*

Proof Conditions one and two require an environment in which the sensor outputs are uncontaminated to guarantee an accurate result as,

$$\mathbb{E}[\bar{\mathbf{s}}] \approx \mathbb{E}[\mathbf{s}] \qquad \textit{and} \qquad Var[\bar{\mathbf{s}}] \approx 0. \tag{3.14}$$

The proof of condition three relies upon a further exploration of Equation (3.8). It is clear that the solution given by Equation (3.8) only exists when the matrix $\mathbf{\Phi}^T\mathbf{\Phi}$ is invertible, i.e. when it has non-zero determinant, or equivalently a rank of three. For a sensor composed by N pixels,

$$\text{Rank}(\mathbf{\Phi}^T\mathbf{\Phi}) \equiv \text{Rank}(\mathbf{\Phi}) = \text{Rank}\begin{pmatrix} x_1 & y_1 & z_1 \\ \vdots & \vdots & \vdots \\ x_N & y_N & z_N \end{pmatrix} \leq 3. \tag{3.15}$$

When $N = 1$ and $N = 2$, the rank of $\mathbf{\Phi}^T\mathbf{\Phi}$ is 1 and at most 2 respectively. Thus $\mathbf{\Phi}^T\mathbf{\Phi}$ has no inverse and Equation (3.8) has no solution. When $N = 3$ and the plane containing the unit vectors of three pixels intersects the origin (i.e. condition three does not hold), there must be a set of (a, b, c) that can be found as,

$$\begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{3.16}$$

For non-trivial cases, the constants a, b, and c cannot be identically zero, meaning the determinant of the matrix must satisfy,

$$|\mathbf{\Phi}| = 0. \tag{3.17}$$

Thus, when the plane containing the three pixels passes through the origin, the matrix $\mathbf{\Phi}^T\mathbf{\Phi}$ has no inverse and there is no solution for Equation (3.8). ■

The above corollary puts a lower bound on the number of pixels needed to estimate the sun vector. Specifically, at least three pixels are required to make the least squares calculation possible. For a sensor with N $\geq$ 3, the rank of $\mathbf{\Phi}^T\mathbf{\Phi}$ must be three in order to make it invertible. This merely requires that all the sensors in $\mathbf{\Phi}$ do not lie on the same plane with the origin and can be easily avoided with an intentional pixel arrangement.

3.3 Rotation Matrices and Attitude Compensation

Since the sensor is mostly used on mobile platforms, the sun vector $\mathbf{e}_s(A, Z)$ in HCS is usually not equal to $\mathbf{e}_s(\alpha_s, \zeta_s)$ in sensor's body fixed coordinate system. To implement celestial navigation, sensor's attitude needs to be measured and compensated for the transformations between these two coordinates.

In a typical reference frame, the motion of a rigid body in three-dimensional space consists of six degrees of freedom: forward/backward, up/down, left/right and rotations about three perpendicular axes. Since most celestial objects are far away from us, only rotation affects the measurement of CRVs. Considering two coordinate systems: one is the stationary coordinate system (X,Y,Z) fixed in the space and the other one is the body-fixed coordinate system (x,y,z) along with the sensor's attitude. As shown in Figure 3.3, the transformation between two coordinates can be approached by a sequence of three specific body-fixed rotations by the Euler angles [43] as,

- A yaw rotation of ψ about the Z axis;
- A pitch rotation of θ about the y' axis;
- A roll rotation of φ about the x'' axis;

Angles ψ, θ and φ are measured positively in the counter-clockwise direction.

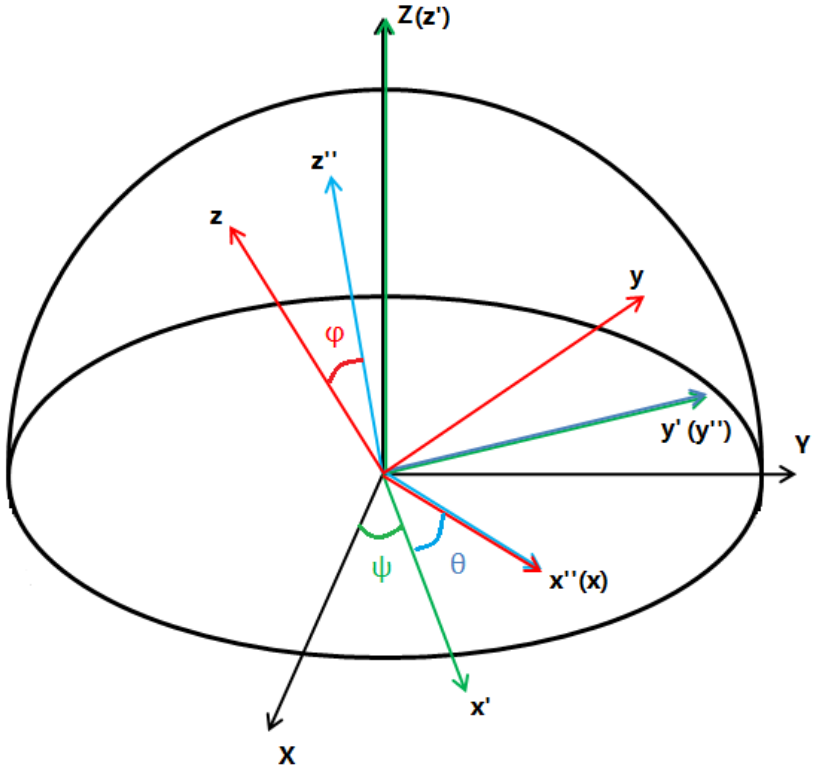


Figure 3.3. Three Euler angles of yaw (ψ), pitch (θ) and roll (φ).

If a celestial reference **Ref** measured in the sensor's body-fixed coordinate system is expressed as $\mathbf{Ref}^{(x,y,z)}$, it's corresponding position in the stationary coordinate, normally in HCS, can be denoted as $\mathbf{Ref}^{(X,Y,Z)}$. The coordinate transformation between $\mathbf{Ref}^{(x,y,z)}$ and $\mathbf{Ref}^{(X,Y,Z)}$ can be processed as,

$$\begin{aligned}\mathbf{Ref}^{(x,y,z)} &= \mathcal{R}_{(X,Y,Z)\to(x,y,z)}(\psi,\theta,\varphi)\,\mathbf{Ref}^{(X,Y,Z)} \\ &= \mathcal{R}_3(\varphi)\,\mathcal{R}_2(\theta)\,\mathcal{R}_1(\psi)\,\mathbf{Ref}^{(X,Y,Z)},\end{aligned} \tag{3.18}$$

where $\mathcal{R}_1(\psi)$, $\mathcal{R}_2(\theta)$ and $\mathcal{R}_3(\varphi)$ are individual rotation matrices as:

$$\mathcal{R}_1(\psi) = \begin{bmatrix} \cos(\psi) & \sin(\psi) & 0 \\ -\sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{3.19}$$

$$\mathcal{R}_2(\theta) = \begin{bmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & \cos(\theta) \end{bmatrix}, \tag{3.20}$$

$$\mathcal{R}_3(\varphi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\varphi) & \sin(\varphi) \\ 0 & -\sin(\varphi) & \cos(\varphi) \end{bmatrix}. \tag{3.21}$$

Inversely,

$$\begin{aligned} \mathbf{Ref}^{(X,Y,Z)} &= \mathcal{R}_{(x,y,z)\to(X,Y,Z)}(\psi,\theta,\varphi)\,\mathbf{Ref}^{(x,y,z)} \\ &= [\mathcal{R}_{(X,Y,Z)\to(x,y,z)}(\psi,\theta,\varphi)]^{-1}\ \mathbf{Ref}^{(x,y,z)} \\ &= [\mathcal{R}_1(\psi)]^{-1}[\mathcal{R}_2(\theta)]^{-1}[\mathcal{R}_3(\varphi)]^{-1}\ \mathbf{Ref}^{(x,y,z)} \\ &= [\mathcal{R}_1(\psi)]^{T}[\mathcal{R}_2(\theta)]^{T}[\mathcal{R}_3(\varphi)]^{T}\ \mathbf{Ref}^{(x,y,z)} \\ &= [\mathcal{R}_{(X,Y,Z)\to(x,y,z)}(\psi,\theta,\varphi)]^{T}\ \mathbf{Ref}^{(x,y,z)} \end{aligned} \tag{3.22}$$

since a rotation matrix commutes with its transpose as a normal matrix.

In our sensing systems, the 3DM-GX3-35® high performance, miniature Attitude Heading Reference System (AHRS) is used to measure the Euler angles[2]. It incorporates a triaxial accelerometer, triaxial gyro, triaxial magnetometer and an on-board processor running a fusion algorithm to provide orientation and inertial measurements with a static accuracy of $\pm 0.5°$.

Since there's no permanent indicator of the true north, most AHRS use the magnetic north and the earth's gravity to determine Euler angles. However, the desired sun vector with respect to the true north is undetermined when the local magnetic declination cannot be provided. This may cause an unexpected bias in celestial geolocation when using the magnetic azimuth angle instead of the true azimuth. The issue will be further discussed and solved by our algorithm in Chapter 5.

[2] `http://www.microstrain.com/inertial/3dm-gx3-35`

3.4 Sensor Operation and Field Experiments

The light intensity sensors are subject to several sources of noise and interferences, many of which can seriously compromise the sensor accuracy if not properly addressed.

In practice, since the sky scatters sunlight to some extent, the sun cannot be simply assumed as a point dominant light source. A background irradiance is then assumed as the approximation of scattered sunlight. All sensor data should subtract this background irradiance before the calculation of Equation (3.8).

Reflections and shadowing affect all optical sensors. When the sensor is shadowed or influenced by the reflection, it will receive erroneous amount of sunlight. Both reflections and shadowing can be easily identified by checking the continuity of sensor outputs since both the sensor platform and the sun move in a continuous fashion. A sudden change of sensor outputs caused by reflections or shadowing can be monitored by microcontrollers and eliminated as incorrect results. Most erroneous sensor data may also be detected and compensated via a comparison between the real sensor outputs and the theoretical frequency (sensor output) distribution as shown in Figure 3.2.

We supply our sensor's operational procedure below. This is based on our experience and is aimed at maximal mitigation of uncertainties.

1. Place the sensor device and get the sensor initial attitude data (Euler angles) to find the rotation matrices in Equation (3.18) between sensor's coordinate system and HCS.

2. Measure the outputs of all sensors, $\{f_i(t)\}$, and eliminate inconsistent and discontinuous values that may caused by shadowing or reflection.

3. Employ the least squares method to obtain the initial estimation of the sun vector $\overline{\mathbf{e}}_s^0(t)$ by using at least three sensors with large outputs.

4. Use $\overline{\mathbf{e}}_s^0(t)$ to find the illuminated area (where the sensor outputs should be positive) on the hemispherical sensor structure. For those sensor pixels out of the area, calculate the average output value $\overline{f}(t)$ as the background irradiance.

5. Subtract $\overline{f}(t)$ from $\{f_i(t)\}$. Use the remaining positive outputs $\{f_i^+(t)\}$ and the corresponding pixel unit vectors $\{\mathbf{e}_i^+(t)\}$ to redo the least squares calculation to estimate the sun vector $\overline{\mathbf{e}}_s(t)$.

6. Repeat Steps 3 through Step 5 for different initial estimations of $\overline{\mathbf{e}}_s^0(t)$ until find the best $\overline{\mathbf{e}}_s(t)$ that lies within a suitable tolerance by defining a cost function as,

$$J = \left(\sum_{i=1}^{m} \sum_{j=1}^{m} \left(\frac{(\overline{\mathbf{e}}_s \cdot e_i^+) f_j^+}{(\overline{\mathbf{e}}_s \cdot e_j^+) f_i^+} - 1 \right)^2 \right)^{\frac{1}{2}}, \tag{3.23}$$

   where m is the total number of sensors in the illuminated area.

7. Repeat Steps 1 through 6 for the next measurement of $\{f_i(t + \delta t)\}$ and check the continuity of sun vector. If the change is outside the expected bound (depending on the time, the speed of the vehicle and sampling interval), ignoring the measurement and continuing to the next. Otherwise compensate the sensor attitude with rotation matrices and calculate the desired sun vector in HCS.

8. Repeat Steps 1 through 7 for all sample instants.

Remark 3.4.1 *The duration of one sample instant should be fast enough to make sure the motion of the sensor or the sun won't influence the estimations.*

Remark 3.4.2 *In Step 3, the reason for choosing sensors with large values is because these sensors are most likely to face the sun in relative small angles and minimally influenced by disturbances. With more sensing pixels on the structure, we will have a larger number of accurate sensor data to do the least squares calculations.*

Remark 3.4.3 *The cost function J in Equation (3.23) is used to evaluate the correctness of estimations by comparing the theoretical distribution of sensor outputs as*

in Figure 3.2 with the real sensor data. J is zero when $\overline{\mathbf{e}}_s = \mathbf{e}_s$ since from Equation (3.4),

$$\frac{\mathbf{e}_s \cdot e_i^+}{\mathbf{e}_s \cdot e_j^+} = \frac{f_i^+}{f_j^+}, \tag{3.24}$$

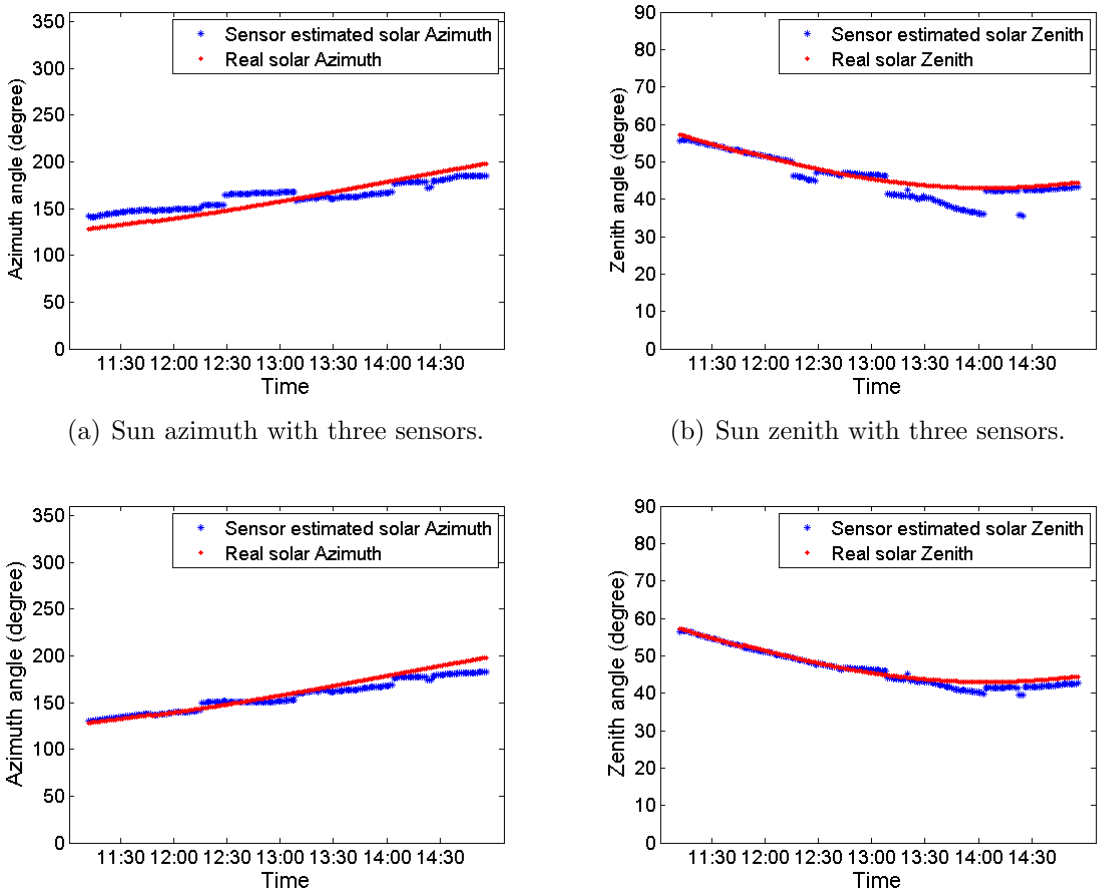
if the background irradiance and disturbances are negligible. A smaller J indicates a better estimation.

3.4.1 Experiment for the Sun Vector Measurement

The field experiment was operated on the top of the Purdue Northwestern Avenue Parking Garage. The approximate location data (longitude, latitude, altitude) was at (-86.91°, 40.43°, 209 m). Experiment time was continuous from local time 11 : 11 to 14 : 47 on March 13th, 2012 under clear skies. A microcontroller received the light intensity signals from all sensors and stored them in a computer for processing. All data was collected and processed through MATLAB®. Given the GPS and time information, the algorithm in [44] was employed to find out the theoretical values of sun vector for comparisons.

To simplify the experiment, the sensor was positioned horizontally and kept stationary for all the time. The sensor's coordinate system was also aligned with the natural directions. Due to the limited number of sensing pixels (only 17), at most five sensing pixels with reliable outputs were available for the least squares calculations. The estimated sun vector and the theoretical sun vector are compared for the azimuth and zenith respectively in Figure 3.4. We also explore the influence of pixel amount by comparing the results of using only three "brightest" pixels (i.e. pixels with largest values) and the average results of different combinations with three pixels out of five in Steps 3 through 6. Comparisons of azimuth and zenith angles are shown in Figure 3.4(a) and Figure 3.4(b) for the case of using three pixels while Figure 3.4(c) and Figure 3.4(d) show the result with five pixels.

Overall, the graphs show that our sensor performs the capability of measuring the sun vector, especially for the zenith angle. The average error is 5.05° for the



(a) Sun azimuth with three sensors.

(b) Sun zenith with three sensors.

(c) Sun azimuth with five sensors.

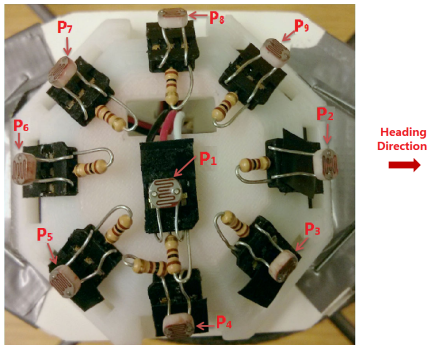
(d) Sun zenith with five sensors.

Figure 3.4. Experiment result of the sun vector estimation.

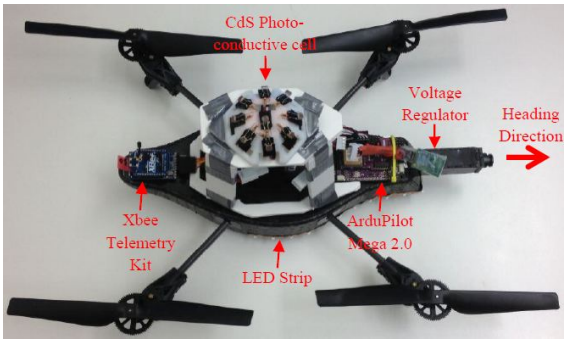
azimuth and 1.28° for the zenith with five sensing pixels. This is not adequate for accurate geolocation but the sensor is adequate for applications such as orientation or sun tracking. It is also notable that the result of using five sensing pixels is better than using three pixels. So increasing the number of sensing pixels is preferred to get more accurate results as we proved in Remark 3.2.1.

3.4.2 Absolute Orientation via a Solar Compass

As illustrated in Section 2.4.1, given the measured sun vector $\mathbf{e}_s(\alpha_s, \zeta_s)$ in sensor's coordinate and the theoretical sun vector $\mathbf{e}_s(A, Z)$ in HCS, it is possible to determine



(a) The hemispherical sensor structure with 9 pixels. (One on the top and eight are evenly distributed at 45° zenith.)



(b) UAV with the sun sensor mounted.

Figure 3.5. Integrated UAV system with the solar compass.

the absolute orientation via a celestial compass. Using the same principle, the absolute orientation of a Unmanned Aerial Vehicle (UAV) is achieved by a solar compass system designed upon our sensor geometry and algorithm [41].

Our solar compass sensor is shown in Figure 3.5(a) with nine CdS photoconductive cells distributed similarly to the previous structure. Since the resistance of a CdS changes along with the light illuminance, it is possible to estimate the illuminance on the ith pixel and then calculate the sun vector via the same least squares algorithm. A small quadcopter UAV is used to support an ArduPilot Mega microcontroller board and the solar compass sensor (Figure 3.5(b)). The overall system retains a small size, light weight and low cost but is strong enough for ordinary missions.

The outdoor experiment was performed at Intramural Black Playing Fields, located in the southwestern area of Purdue University. The approximate location data (longitude, latitude, altitude) is at (-86.9319°, 40.4225°, 192.6 m). The time was from 17:45 to 17:47 on July 12, 2013 and the weather was rather good to eliminate the influence of clouds. Using the same astronomical algorithm, the theoretical sun azimuth was calculated in values from 267.38° to 267.77° during the experiment. The magnetic declination is about -4.0321° at the location based on the current model of global magnetic field.

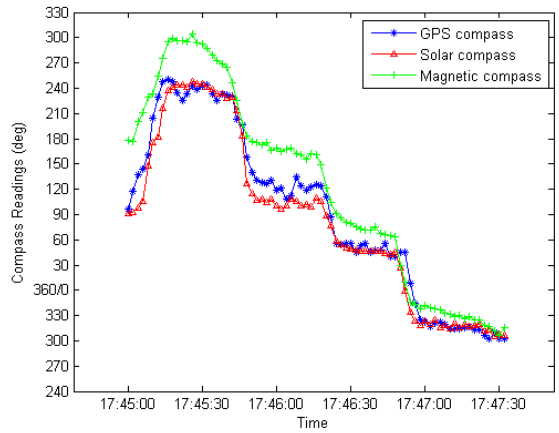
Since there is no absolute orientation aid during the daytime, we use GPS data to determine the orientation by assuming the UAV keeps its moving direction during two successive measurements and use the vector from the previous GPS point to the latter GPS point to approximate orientations. If the GPS sensor has a sufficiently high accuracy, the result of GPS compass should be very close to the true value and qualified as the criterion for evaluating the performance of other compass systems.

For the data analysis, the least squares calculation of the sun vector was performed by six pixels at the same time, i.e. $N = 6$ in Φ. Among them, five adjacent sensors with $\zeta = 45°$, e.g. (3,4,5,6,7), (7,8,9,2,3), were chosen if the five sensors received the maximum summation of illuminance when compared to other combinations. The last sensor is always the sensor 1 on the top such that these six sensors are facing the sun in relative smaller angles and should get positive values in Equation (3.3).

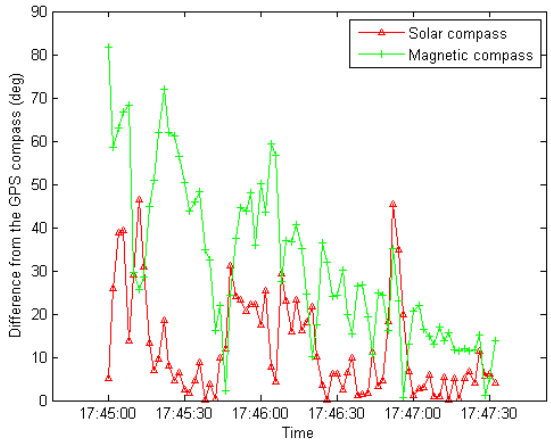
During the experiment, the UAV followed a path as shown in Figure 3.6(a). The data from three orientation systems, the GPS compass, the solar compass, and the magnetic compass (after compensating the magnetic declination) were collected together for comparisons. As shown in Figure 3.6(b), there are clearly some fluctuates of GPS compass due to GPS sensor's error. But when referring to the path in the map, GPS compass basically indicates a correct range of orientations. The solar compass follows GPS compass closely and retains a smooth shape with few fluctuates. On the other hand, the result of magnetic compass varies a lot from other two compasses, especially at the beginning location where a huge stadium lamp stands nearby. The error maybe due to the surrounding electromagnetic noises and the low quality magnetic sensor on our UAV.

Using the GPS compass as the criterion, Figure 3.6(c) quantitatively indicates the absolute differences between GPS compass and other two compasses for all measurements. It is clear that the solar compass performs much closer to the GPS compass than the magnetic compass during the experiment, especially when the UAV moved in straight directions. Some large differences between the solar compass and GPS compass happen at the turnings. This is caused by the change of sensors in Φ for the

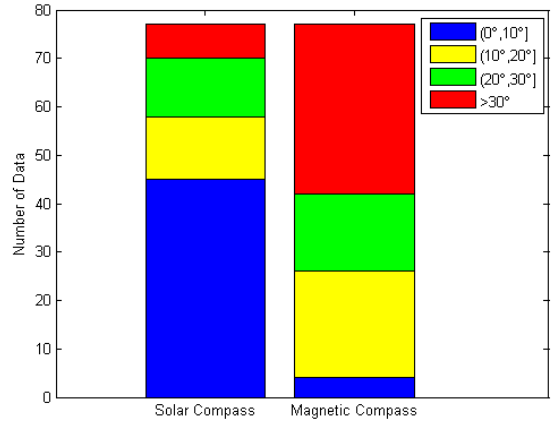
(a) Experiment location and UAV's path.



(b) Compass readings of GPS compass, solar compass and magnetic compass.



(c) Absolute differences between GPS compass and other two compasses.



(d) Error distributions of solar compass and magnetic compass.

Figure 3.6. Experiment result of three compass systems.

least squares calculation when the UAV changes its attitude. The statistical distributions of the values in Figure 3.6(c) are plotted in Figure 3.6(d) for the solar compass and magnetic compass respectively. The solar compass has most uncertainties less than 10° and retains a smaller amount of large error while the magnetic compass has more undesired large errors.

Depending on the graphs, all three compasses can estimate the orientations. The GPS compass we used, though reliable, may has unexpected fluctuates when a large

GPS error happens. It also requires a sufficient motion of the platform to compare two successive GPS data. The magnetic compass, though simple and straightforward, can be easily disturbed by noises. It is also unable to determine the true north without knowing the magnetic declination. Our solar compass works very well under good whether conditions. It may perform even better than the magnetic compass in an environment filled with electromagnetic disturbances. As proved mathematically, the accuracy of our solar compass can be improved by increasing the number of sensing pixels for better least squares calculations. Therefore, the absolute orientation via the sun is applicable and qualified as a good augment for modern orientation systems.

However, it is undeniable that light intensity sensors are vulnerable to disturbances such as clouds or other light sources. It can only be used to determine a dominant light source. Current uncertainties in azimuth and zenith measurements are still not adequate for celestial geolocation. To deal with these issues, we developed a hemispherical camera sensing system in the next chapter for better estimations.

4. ESTIMATION OF CELESTIAL VECTORS VIA CAMERAS

Camera, in a common sense, is a device to record images directly from objects. A typical structure of camera is presented in Figure 4.1. There is a lens, or a set of lenses, in the front of the camera to converge light beams from an object to the image plane perpendicular to the lens axis. The distances from the object to the lens and from the lens to the image sensor, L_1 and L_2 respectively, satisfy the thin lens formula as,

$$\frac{1}{L_1} + \frac{1}{L_2} = \frac{1}{f}, \tag{4.1}$$

if the lens has a negligible thickness and with a focal length of f. The image sensor at L_2 will produce a sharp image of the object at L_1 as long as $L_1 > f$ [38].

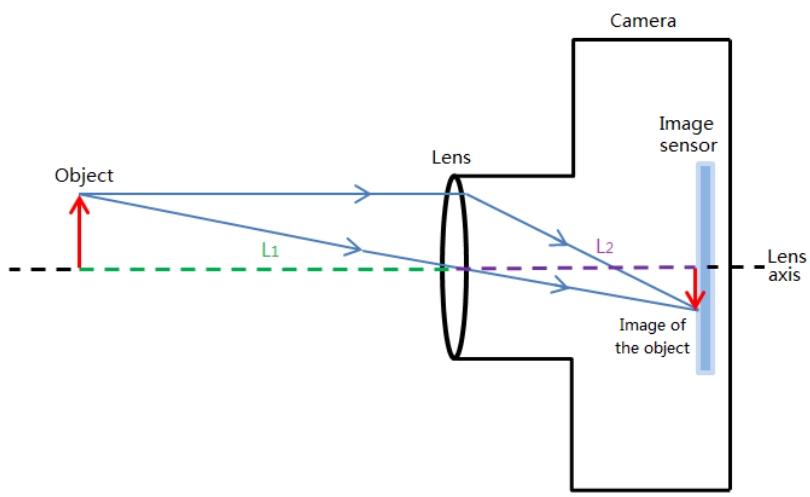


Figure 4.1. Imaging system of a camera.

Compared to the light intensity sensors used in Chapter 3, a camera sensing system has the following advantages:

- Reliable and abundant information in images.

- Robust to different disturbances.
- Various image processing techniques are available.
- Applicable to more celestial objects such as the moon and stars.
- Extensible to other applications like object detection and astrophotography.

The quality of optical lenses and image sensor largely influences the performance of a camera. A lens that produces few optical aberrations and a sensitive, high resolution image sensor with few noise are highly preferred. To capture celestial objects, the field of view (FOV) is also an important consideration since it describes the angular range that can be imaged by the camera. With a wider FOV, more celestial objects are likely to be captured in the image for a better processing.

The FOV of a rectilinear image sensor can be calculated based on the dimensional size of the sensor (d) and effective focal length (f) as,

$$\alpha = 2 \arctan \frac{d}{2f}. \tag{4.2}$$

FOV thus can be increased by using a larger image sensor or a wide-angles lens with a shorter focus length [45]. However, a larger image sensor will increase the camera weight and cost. A wide-angles lens also aggravate aberrations issues.

Alternatively, a wider FOV can be achieved by a multi-vision camera system. Some insects, e.g. fly, have very efficient visual systems composed of several thousands of distributed visual neurons referred as the ommatidium. The structure is able to provide a omnidirectional vision [46]. Scientists in EPFL have developed a camera system that can take panoramic pictures or films and reconstruct the pictures in 3D space [47].

In this chapter, nine cameras are mounted on the hemispherical structure we built in Chapter 3 to construct a panoramic camera system for detecting celestial objects (Figure 4.2). The camera we used is a programmable CCD camera with adjustable

features for different light conditions[1]. The system is connected with a computer to trigger shooting and store images. Ideally, all nine cameras face outward the hemispherical structure with their unit vectors aligned with the lens optic axis.

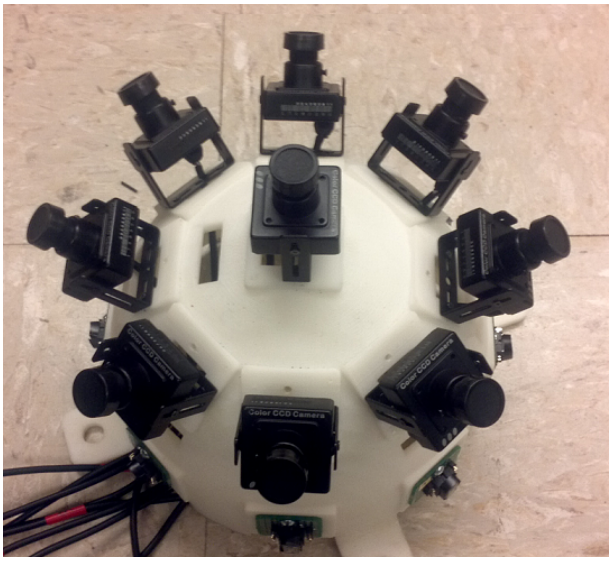


Figure 4.2. Cameras on the hemispherical sensor, with one on the top and eight are evenly distributed at 45° zenith.

Since the cameras are programmable, the settings can be configured manually for different celestial objects. For bright targets such as the sun, it is very important to use a shorter shutter and lower sensitivity to avoid saturation and protect image sensor. On the contrary, a high sensitivity and longer shutter time are employed for taking pictures of dim objects like the planets, stars or the moon at small phases.

In this chapter, a well established camera model and the detailed camera calibration procedures are presented to accurately determine the transformation between the image coordinate system and HCS. The image processing algorithms for detecting the sun and moon are introduced and shown with examples. The field experiments together with the analysis are also presented.

[1] `http://store.3drobotics.com/products/programmable-ccd-camera-480tv`.

4.1 Camera Modeling and Calibrations

To apply celestial navigation for an observable celestial object, it is necessary to build the full transformation from its projection in the image to its corresponding position in HCS. Thus, the model of all coordinate systems and the corresponding transformation coefficients need to be established and calibrated.

4.1.1 Camera Modeling

In our camera sensing system, five coordinate systems:

- Stationary coordinate system (X, Y, Z), fixed in the space and represents a picturable point's position in HCS[2],
- Sensor coordinate system (x_s, y_s, z_s), aligned with the hemispherical sensor to denote the position in sensor's spherical coordinate system[3],
- Camera coordinate system (x_c, y_c, z_c), the camera centered coordinate which represents the object's position from the camera's perspective,
- Image plane coordinate system (O', u, v), to indicate the object's image projection on the image plane,
- Image coordinate system (u_p, v_p), to define the object's position in the digital images after transformed to pixel values,

and, two sets of parameters:

- Extrinsic Parameters, to indicate the transformation between the object's position in stationary coordinate system (X, Y, Z) and camera coordinate system (x_c, y_c, z_c),
- Intrinsic Parameters, to represent a camera's imaging property from camera coordinate system (x_c, y_c, z_c) to the final image coordinate system (u_p, v_p),

[2]This corresponds to the stationary coordinate system in Section 3.3.
[3]This corresponds to the body-fixed coordinate in Section 3.3.

describe a well established camera model [48, 49].

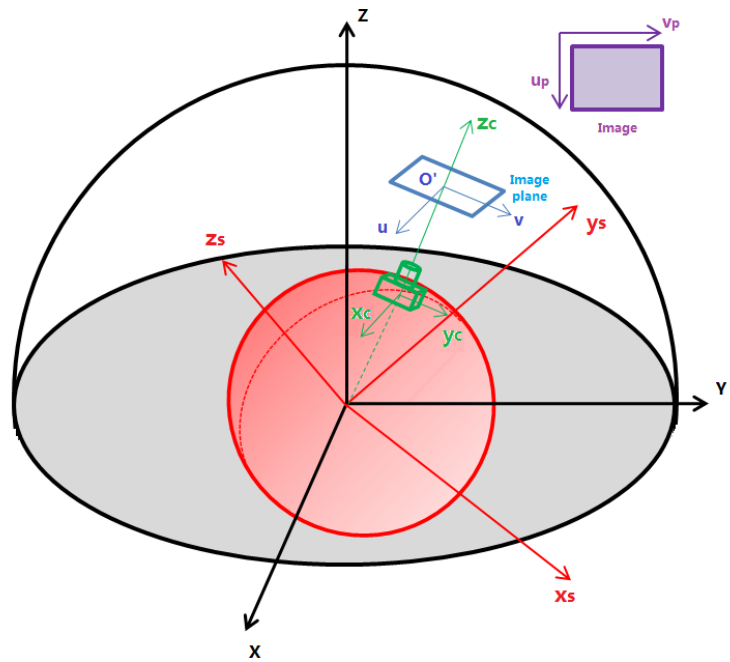


Figure 4.3. Illustration of all coordinate systems.

The illustration of these coordinate systems is shown in Figure 4.3. Following Equation (3.18), given a celestial reference **Ref**, the transformation between the sensor coordinates (x_s, y_s, z_s) and the stationary coordinates (X, Y, Z) includes three sequential rotation matrices by Euler angles as,

$$\begin{aligned} \mathbf{Ref}^{(x_s,y_s,z_s)} &= \mathcal{R}_{(X,Y,Z)\to(x_s,y_s,z_s)}(\psi,\theta,\varphi)\,\mathbf{Ref}^{(X,Y,Z)} \\ &= \mathcal{R}_3(\varphi)\,\mathcal{R}_2(\theta)\,\mathcal{R}_1(\psi)\,\mathbf{Ref}^{(X,Y,Z)} \end{aligned} \tag{4.3}$$

where ψ, θ and φ are the yaw, pitch, and roll angles respectively.

The origin of camera coordinate system (x_c, y_c, z_c) is usually assumed to locate at the optical center of the camera and the z_c axis coincides with the optical axis, pointing outward from the camera. The transformation between the sensor coordinates (x_s, y_s, z_s) and the camera coordinates (x_c, y_c, z_c) is based on the camera's unit

vector $\mathbf{e}_i(\alpha_i, \zeta_i)$. At most two similar rotation matrices from Equations (3.19)-(3.21) are enough to describe the transformation as,

$$\begin{aligned}
\mathbf{Ref}^{(x_c,y_c,z_c)} &= \mathcal{R}_{(x_s,y_s,z_s)\to(x_c,y_c,z_c)}(\alpha,\zeta)\,\mathbf{Ref}^{(x_s,y_s,z_s)} \\
&= \mathcal{R}_{(x_s,y_s,z_s)\to(x_c,y_c,z_c)}(\alpha,\zeta)\,\mathcal{R}_{(X,Y,Z)\to(x_s,y_s,z_s)}(\psi,\theta,\varphi)\,\mathbf{Ref}^{(X,Y,Z)} \\
&= \mathcal{R}_{(X,Y,Z)\to(x_c,y_c,z_c)}(\alpha,\zeta,\psi,\theta,\varphi)\,\mathbf{Ref}^{(X,Y,Z)} \\
&= \mathcal{R}_5(\zeta)\,\mathcal{R}_4(\alpha)\,\mathcal{R}_3(\varphi)\,\mathcal{R}_2(\theta)\,\mathcal{R}_1(\psi)\,\mathbf{Ref}^{(X,Y,Z)},
\end{aligned} \tag{4.4}$$

where $\mathcal{R}_5(\zeta)$ and $\mathcal{R}_4(\alpha)$ should be determined by the relative position between (x_c, y_c, z_c) and (x_s, y_s, z_s) and may have different forms.

Inversely,

$$\begin{aligned}
\mathbf{Ref}^{(X,Y,Z)} &= \mathcal{R}_{(x_c,y_c,z_c)\to(X,Y,Z)}(\alpha,\zeta,\psi,\theta,\varphi)\,\mathbf{Ref}^{(x_c,y_c,z_c)} \\
&= (\mathcal{R}_{(X,Y,Z)\to(x_c,y_c,z_c)}(\alpha,\zeta,\psi,\theta,\varphi))^T\mathbf{Ref}^{(x_c,y_c,z_c)}.
\end{aligned} \tag{4.5}$$

Values of $\alpha, \zeta, \psi, \theta, \varphi$ are actually the extrinsic parameters needed for the modeling. α, ζ are fixed for all cameras while the Euler angles of ψ, θ, φ need to be measured by a AHRS.

The second transformation is between the camera coordinate system (x_c, y_c, z_c) and the image coordinate system (u_p, v_p). For an ideal pinhole camera model without any optical aberration, the transformation can be expressed as two successive steps.

Firstly, the projection of the point (x_c, y_c, z_c) on the image plane is expressed as,

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{f}{z_c}\begin{pmatrix} x_c \\ y_c \end{pmatrix}, \tag{4.6}$$

regarding to the thin lens formula in Equation (4.1).

Next, the image sensor senses the object's projection on the image plane and digitize it into a picture scaled by pixels. Since the principal point of the image plane

O' may not coincide with the center of image sensor, the corresponding pixel position in image coordinates can be considered as,

$$\begin{pmatrix} u_p \\ v_p \end{pmatrix} = \begin{pmatrix} S_u u \\ S_v v \end{pmatrix} + \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}, \tag{4.7}$$

where (u_0, v_0) denotes the position of camera's principal point O' in the image and S_u, S_v are two scaling factors for changing units from length to pixels on two dimensions respectively. The origin of the image pixel coordinate is usually at the upper left corner of the image pixel array.

However, the pinhole model is not adequate in realistic since it is only an approximation of the real image projections. As the lens is an optical device with some degree of aberrations, camera does not produce perfect images. There are basically two types of aberrations: chromatic and monochromatic aberrations. The former performs as a function of light frequency, i.e. the color, while the latter occurs for any kinds of light regardless the color of light. There are aberrations such as the spherical aberration, coma and astigmatism that deteriorate the image by making it unclear, and field curvature or distortions that deform the image [38]. Among these aberrations, distortion is the one that must be dealt with for object detections since it deforms images and affects coordinate transformations. There are usually two types of distortion: the radial distortion that displaces radially on the image plane with an expression as,

$$\begin{pmatrix} \delta u^r \\ \delta v^r \end{pmatrix} = \begin{pmatrix} u(k_1 r^2 + k_2 r^4 + k_3 r^6 \ldots) \\ v(k_1 r^2 + k_2 r^4 + k_3 r^6 \ldots) \end{pmatrix}, \tag{4.8}$$

and the decentering distortion which includes both radial and tangential components as,

$$\begin{pmatrix} \delta u^t \\ \delta v^t \end{pmatrix} = \begin{pmatrix} 2p_1 uv + p_2(v^2 + 3u^2) \\ p_1(u^2 + 3v^2) + 2p_2 uv \end{pmatrix}, \tag{4.9}$$

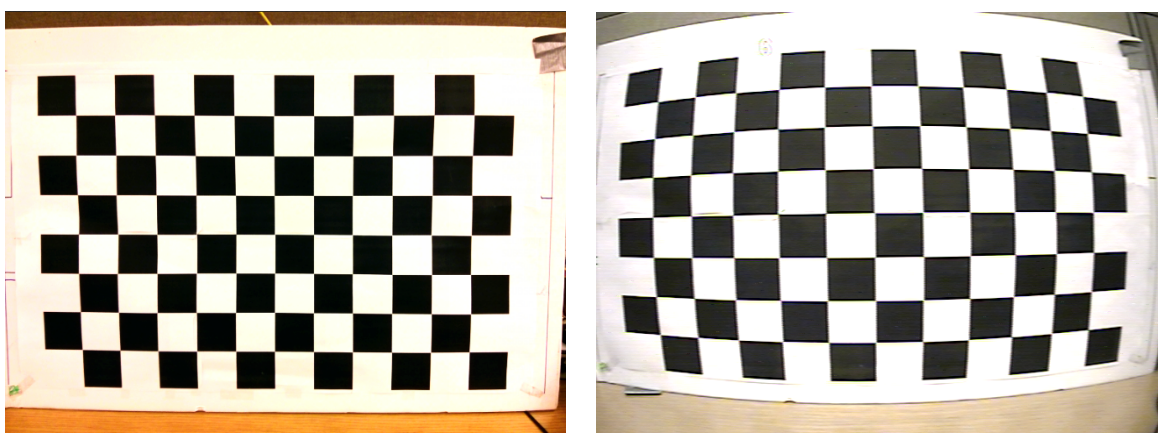
where $r = \sqrt{u^2 + v^2}$ and $k_{1,2\ldots}$, $p_{1,2}$ are the distortion parameters [49–51].

Combining Equations (4.7)-(4.9), the transformation with the consideration of distortions can be summarized as,

$$\begin{pmatrix} u_p^d \\ v_p^d \end{pmatrix} = \begin{pmatrix} S_u(u + \delta u^r + \delta u^t) \\ S_v(v + \delta v^r + \delta v^t) \end{pmatrix} + \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}. \tag{4.10}$$

The values of f, u_0, v_0, S_u, S_v, together with parameters $k_{1,2...}$, $p_{1,2}$, are the intrinsic parameters which define the camera's imaging properties. Unlike the extrinsic parameters, the intrinsic parameters need to be provided by the camera's manufacturer or just like what has been done in this thesis, a camera calibration is needed to figure out these values.

4.1.2 Camera Calibration and Undistortion



(a) Normal checkerboard image.

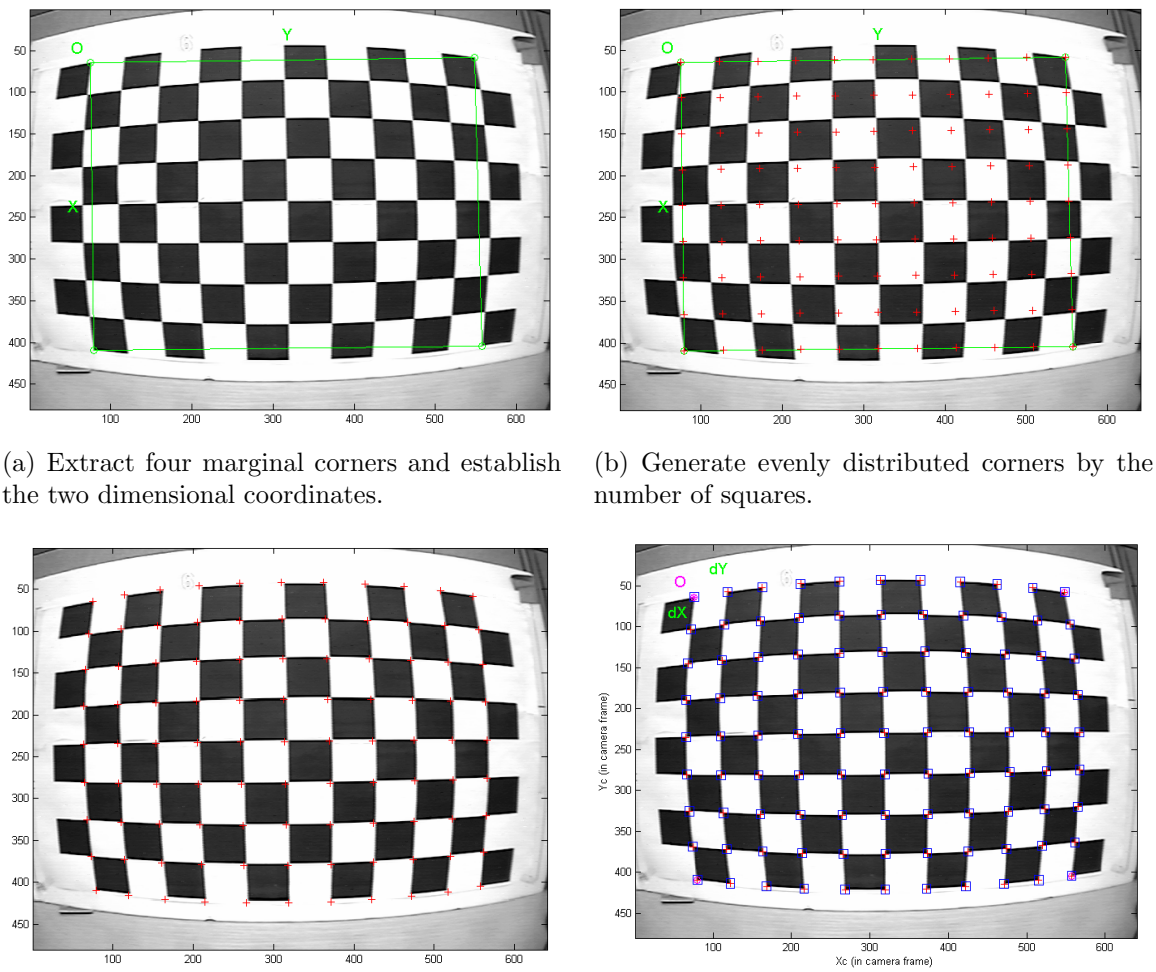
(b) Distorted checkerboard image taken by the calibrated camera.

Figure 4.4. Images of the checkerboard used for calibration.

The basic procedure and algorithm of our camera calibration follows the Camera Calibration Toolbox for Matlab® [52]. The calibration needs a black and white planar checkerboard as shown in Figure 4.4(a), which is taken by a less distorted cell phone camera. Figure 4.4(b) shows the image of the same checkerboard taken by the camera

we used for the project. It is clear that the camera has serious distortion issues that need to be addressed.

The first step of the camera calibration is the corner extraction. Figure 4.4 shows the basic procedures. For every corner extracted between the black and white squares on the checkerboard, the program saves its pixel position in the image and run the algorithms to calculate the best set of parameters that minimizes the two dimensional pixel errors.



(a) Extract four marginal corners and establish the two dimensional coordinates.

(b) Generate evenly distributed corners by the number of squares.

(c) Redistribute the corners by an initial guess of distortion coefficients.

(d) Autonomous detection and extraction of corners.

Figure 4.5. Corner extraction of the checkboard.

The function $normalize(pixelposition, fc, cc, kc, alpha_c)$ in the toolbox can be employed to find out the projection (x, y) on the image plane as,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{x_c}{z_c} \\ \frac{y_c}{z_c} \end{pmatrix}, \tag{4.11}$$

in the pinhole camera model.

Since celestial objects can be assumed to locate infinitely far away from camera sensor, the value of z_c actually is meaningless. It is possible to choose an arbitrary value, i.e. $z_c = 100$ in this case, and use,

$$\begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} = \begin{pmatrix} x \times z_c \\ y \times z_c \\ z_c \end{pmatrix}, \tag{4.12}$$

as the equivalent Cartesian position of a celestial object in the camera coordinate system. Finally, Equation (4.5) can be employed to find the corresponding CRV in HCS.

4.2 Image Processing for Celestial Object Detection

In this section, MATLAB® and some image processing methods are used to detect the sun and moon in images, and with special considerations for some possible disturbances.

4.2.1 Sun Detection

A sample image of the sun taken by camera #2 is shown in Figure 4.6. It can be noticed clearly that there is a vertical bright line crossing the whole image and dividing the sun into two parts. This line should attribute to a phenomenon called blooming that occurs on almost all CCD image sensors under the condition when the maximum charge that the CCD can collect and transfer is reached and the overflow

of excess charges from a pixel wells into adjacent ones. Without special designs to drain those excess charges, all the adjacent pixels will be saturated and produce a blooming line on all over-charged pixels [53].

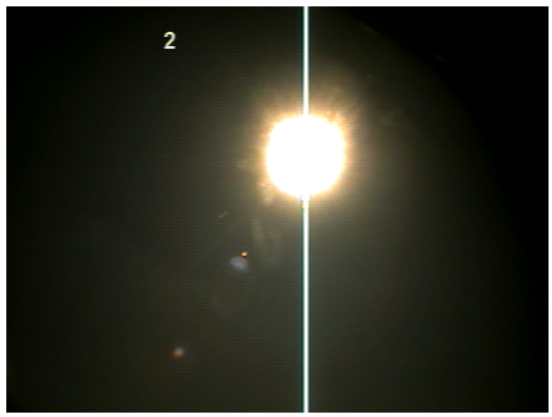


Figure 4.6. A sample image of the sun taken by camera #2.

The blooming effect usually deteriorates the CCD photography. However, it actually contributes a very important approach which can be used to improve the sun detection results. Since the source of excess charges is the sun in images, the blooming line usually happens across the center of the sun, i.e. where the maximum brightness achieves. This property limits the horizontal position of sun center for detections. With more blooming lines, the intersection point of these blooming lines in HCS should be the sun's center exactly. Combined with the traditional image processing method, two different approaches thus can be integrated to improve the accuracy of detections.

Figure 4.7 provides the basic procedure of doing sun and blooming detections. A 2D array *label* is created to store all information. The row index of *label* indicates the camera number from which the image is taken so that the corresponding camera parameters can be used. The first column shows the existence of the sun and blooming. The second and third columns indicate the detected pixel locations (u_p^d, v_p^d) of the sun center. The rest columns can be used for other information, if needed.

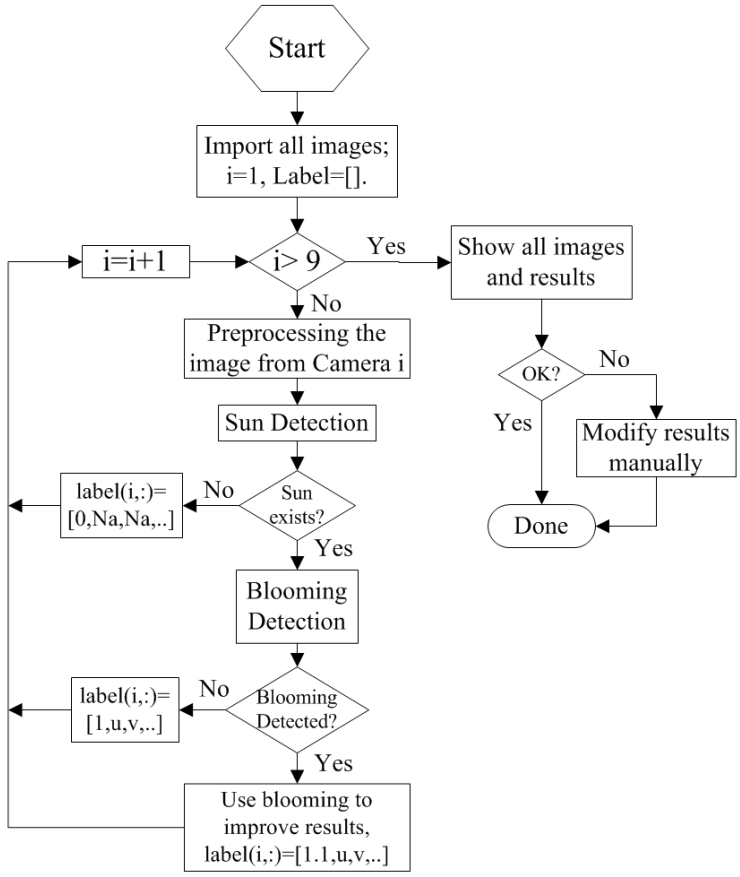


Figure 4.7. Procedure for the sun and blooming detections.

Preprocessing of Sun's Images

Using Figure 4.6 as an example, the first step of preprocessing needs to change the original image from RGB colorspace to a binary black/white image as shown in Figure 4.8. Figure 4.8(a) with a large threshold is used for the sun detection while Figure 4.8(b) with a smaller threshold will be used for the blooming detection.

The results are shown in Figure 4.8 with some noises due to the flare, image number or other disturbances. It is possible to eliminate all the isolated and small noises by a threshold of size. The result is shown in Figure 4.9.

Next, the linear features, sags and crests in Figure 4.9(a) are undesired for sun detection since they influence the roundness of the sun in images. These features can

(a) Binary image with a threshold of 0.9.

(b) Binary image with a threshold of 0.3.

Figure 4.8. Preprocessing step 1: RGB to binary image.

(a) B/W image for sun detection after eliminating noises.

(b) B/W image for blooming detection after eliminating noises.

Figure 4.9. Preprocessing step 2: Eliminate small noises that are fewer than 500 pixels.

be deleted via morphological and erosion operations. For the detection of blooming lines, however, it is necessary to keep all linear features. The final edges of the sun and blooming lines after the preprocessing are shown in Figure 4.10(a) and Figure 4.10(b) separately.

(a) Sun's edge shown in original image.

(b) Edges for blooming detection.

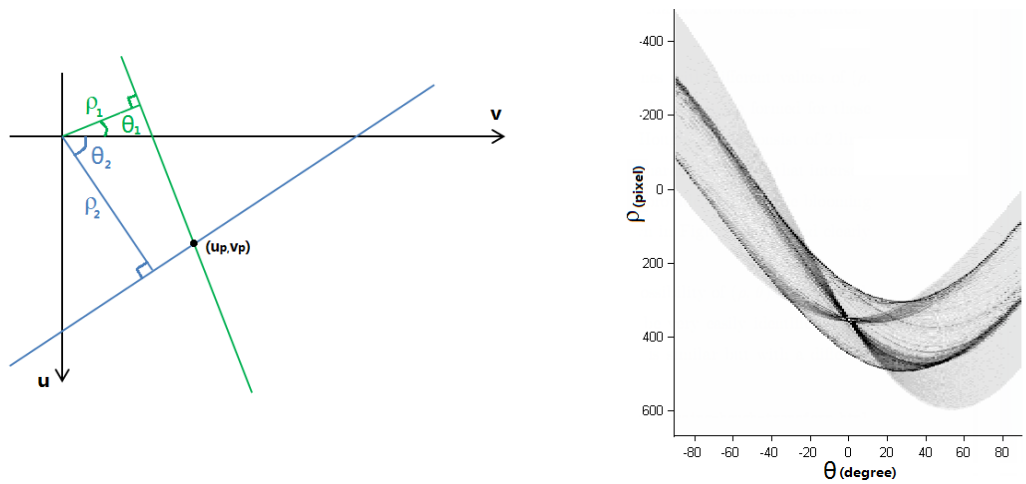
Figure 4.10. Preprocessing step 3: Eliminate undesired features and find the edges.

Feature Extraction via Hough Transform

The edges found in Figure 4.10 include countless pixels needed for the object detection. The circular shape of sun's edge should be detected to get the position of sun center. For the blooming lines, we need to extract the linear features that crosses the whole image vertically. To achieve these feature extractions, we employed the Hough Transform (HT).

The Hough transform was first introduced by Paul Hough in a patent filed in 1962 [54] and latter developed as a method of detecting complex patterns of points in binary images. The Hough transform has long been recognized as a universal technique for shape analysis in images and usually has good performance even with noisy, missing, and extraneous data [55]. The basic idea of Hough transform is to characterize different features with specific values of parameters. Duda and Hart [56] suggested a very useful notation by parameterizing a point in the image (u, v) into

the distance, ρ, and the orientation, θ, of the normal vector to the line from the image origin as shown in Figure 4.11(a).



(a) Parameterizing lines cross a specific point.

(b) Hough matrix of the blooming feature in Figure 4.10(b).

Figure 4.11. Example Hough line and Hough matrix for the blooming feature.

For each pixel in the image, there are infinite lines with different values of (ρ, θ) crossing the point. A corresponding curve called the "Hough Matrix" can be drawn in (ρ, θ) plane for those possible "Hough lines". If there are two points, the Hough Matrix will have two separate lines intersect at a point (ρ_l, θ_l) which characterize the desired line that crosses both of two points in the (u, v) plane [4]. If more pixel points are provided such as the blooming feature in Figure 4.10(b), the Hough Matrix is shown in Figure 4.11(b) and clearly with some peaks of intersections. Applying a voting rule to find the maximum possibility of (ρ, θ) in the parameter space, the corresponding feature in the image can be easily identified. The detection of a circular object is very similar but with a different notation of parameter space [57].

The HT algorithm works very well for both sun and blooming detections in our case as shown in Figure 4.12. The blue circle represents the best fit of the sun and

[4] `http://basic-eng.blogspot.com/2005/12/object-detection-using-hough-transform.html`

the blue star indicates the estimated sun center. The green lines with red arrows are the estimated edges of the blooming lines.

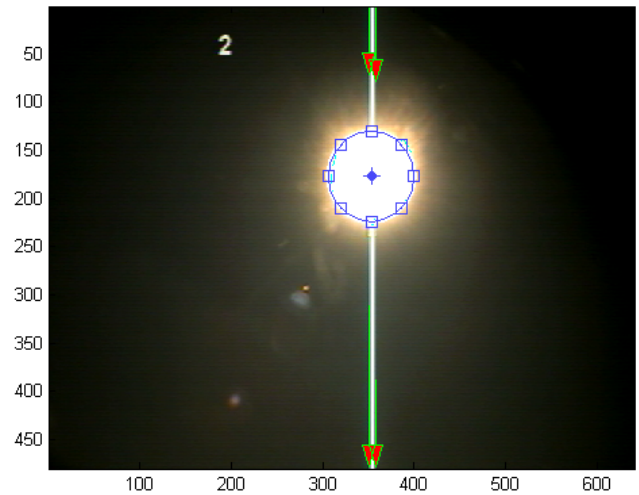


Figure 4.12. Detection result of Figure 4.6 via HT.

Disturbances and Their Mitigations

Both camera hardware and the Hough Transform provide a much better robustness to the disturbances than normal light sensors. However, due to disturbances such as clouds, light flares, obstacles in front of the camera, or simply the sun only shows partially in the image, the accuracy may become worse if these noises are not properly addressed. To improve sensor performance, we use the following criteria to determine desired features:

- Blooming lines:
  - Blooming lines are almost vertical lines.
  - At least one left and one right lines are needed to identify the blooming edges.
  - The horizontal width of blooming lines is in the range of 10 pixels.

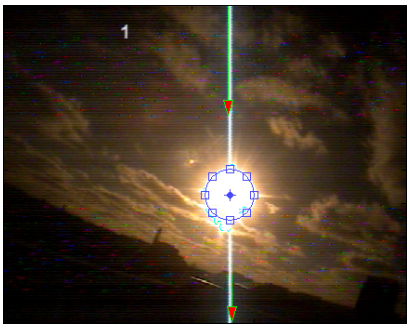
  - Blooming lines that have similar horizontal positions should be considered as the same line.

- The sun:
  - The longest and shortest diameter of the sun should be about the same.
  - The area of the sun should near the result from formula πR^2, where R is the estimated radius from the Hough Transform.
  - The perimeter of the edge should approximately follow the result of $2\pi R$.
  - If the blooming line exists, the sun center should locate in the thin rectangular area restricted by the blooming edges.

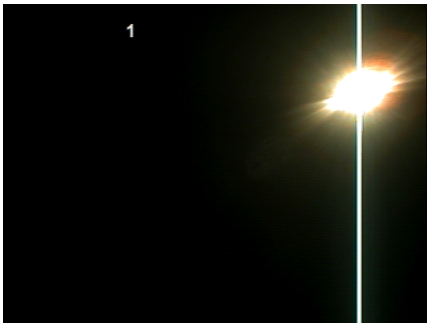
Additionally, different thresholds can also be adjusted for converting binary image during the preprocessing until the sun and blooming features are accurately detected. Figure 4.13 are some examples to show the detection results with different disturbances. It can be seen that the Hough Transform also performs very well in these situations.

The program may fail to detect the desired features if the image is seriously disturbed. For the cases when automatic detection fails, the blooming lines and circular sun can also be chosen by hands. After the processing of all images, the positions of sun center and blooming lines in HCS can be yielded through an undistortion process with intrinsic parameters and a coordinate transformation with extrinsic parameters by Equation (4.5).

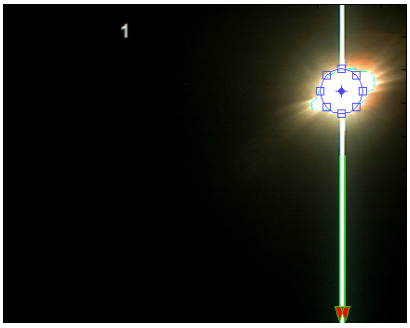
(a) Original image with clouds.



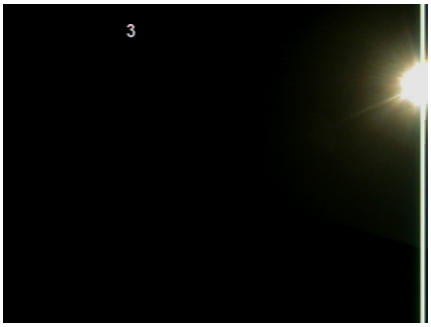
(b) Detection result with clouds.



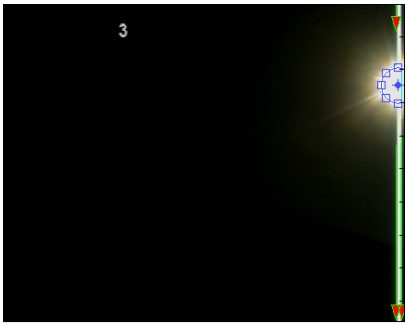
(c) Original image with flares.



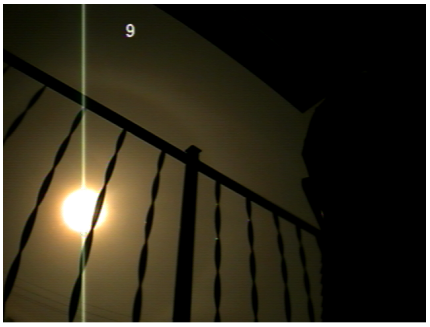
(d) Detection result with flares.



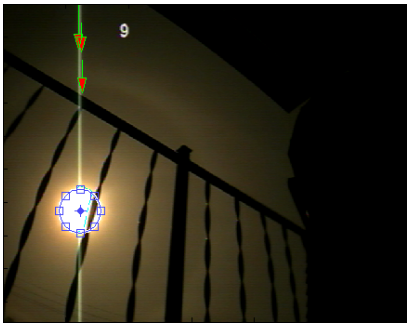
(e) Original image of a partial sun.



(f) Detection result of a partial sun.



(g) Original image with obstacles.



(h) Detection result with obstacles.

Figure 4.13. Sun and blooming detection results with disturbances.

4.2.2 Moon Detection

Except the sun, traditional celestial navigation uses the moon as a common aid since it's the second largest celestial object in the sky and can be easily identified. At clear night, when the sun is not available, sometimes the moon by itself is also qualified as an apparent celestial object for celestial navigation.

Similarly to the sun, the moon is a circular object and has its own moon vector $\mathbf{e}_m(A_m, Z_m)$ to indicate its position in HCS. Thus the processing of moon's images and the detection of moon vector basically follow the same procedure. However, the apparent differences between the daytime and night environments and the moon's periodic change need some special considerations.

Preprocessing of Moon's Images

Unlike the sun, which is the dominant light source during the daytime, the moon light is usually not sufficient to cause CCD saturation (blooming) and may be seriously influenced by other ordinary light sources such as the street lamps, flashes from aircraft and reflections etc. Some features, e.g. the camera number shown at the upper-left part of the image, may also compromise the detection result.

There are several possible approaches to eliminate these disturbances. For the camera number, it can be adjusted in camera settings thus the position remains fixed in all images. It is possible to crop the smallest area that encloses the number and fill the area with black color instead as shown from Figure 4.14(b) to Figure 4.14(c). Next, a further filtration can be implemented to eliminate other disturbances based on the following three criteria:

- Size. Given the time and camera property, the moon's size is relative predictable in images. A possible candidate of the moon can be selected if its area lies in an acceptable range.

- Shape. If an areas has very apparent linear features and unusual shapes, it is likely to be a noise and should be eliminated.
- Color. Unlike artificial lights with various colors, the moon in the image at clear night is usually in white or yellow color[5].

(a) Original image with a street lamp.

(b) Binary image after a binary threshold of 0.3.

(c) Binary image after cropping the number area.

(d) Binary image after filtering out the noises.

Figure 4.14. Prepocessing of moon's images to eliminate disturbances.

As shown in Figure 4.14(d), after eliminating the small and liner features, there are two areas left. Compared with Figure 4.14(a), area 2 is the moon and area 1 is a street lamp with an apparent red color. Table 4.1 lists the average RGB colorspace values of all pixels in two areas. The value is from 0 to 255 for red, green and blue

[5]This may be influenced by the setting of camera white balance or the weather condition.

separately to construct a composite color with different combinations [6]. It is clear that area 1 has more weight of red color than both green and blue thus it must show redder in the image. The moon of area 2, however, has more balanced values. Thus the area 1 can be determined as an artificial light and deleted as an incorrect result.

Table 4.1. Average RGB color values of all pixels in two areas.

Area	Red	Green	Blue
1	207.2318	147.0795	88.3775
2	192.5443	182.4051	167.3924

If the images of the moon are taken continuously on a mobile platform, the artificial light sources can also be identified by comparing successive images. Since the moon is far away from the sensor, a short linear movement shouldn't change the moon vector thus the moon is more likely to remain at the same position in images. However, for nearby light sources, a linear motion of the mobile platform may significantly change their projections in images.

Detection of the Moon with Phases

After preprocessing, next step is to determine the moon vector by detecting the moon's center in images. Since the moon is also a circular object in the sky, the Hough Transform is also employed for feature extraction to achieve the best performance.

Different from the sun, the moon is not a self-luminous object and it is always half illuminated by direct sunlight. During a cycle of one lunar month (approximately 29 days), the relative orbital positions of the sun, moon and earth change everyday thus only a certain degree of the illuminated area can be seen from the earth. Beginning at the first day of a lunar month, the moon's appearance changes from invisible through

[6] `http://www.rapidtables.com/web/color/RGB_Color.htm`

partially illuminated to a fully illuminated circular disk, then back through partially illuminated to invisible again as shown in Figure 4.15. The degree to which the moon is illuminated and the geometric appearance of the illuminated part can be designated by different lunar phases in an order of New moon(a), Waxing crescent moon(b), First quarter moon(c), Waxing gibbous moon(d), Full moon(e), Waning gibbous moon(f), Third (last) quarter moon(g), Waning crescent moon(h) and Dark moon(i) during one cycle [7].

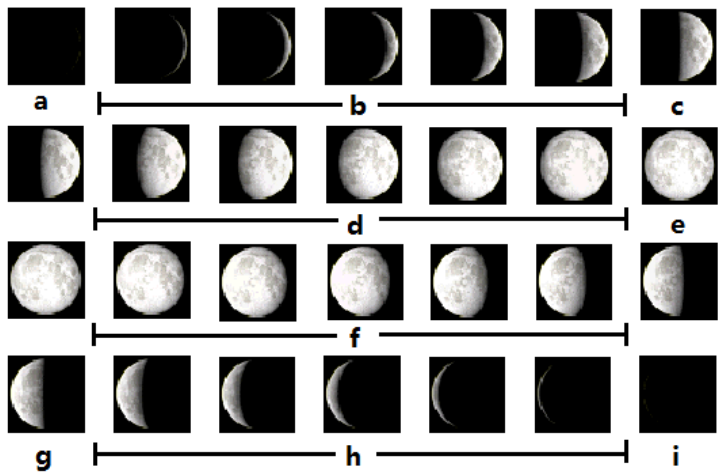


Figure 4.15. Lunar phases in one cycle of lunar month.

The lunar phases, except the full moon, make the moon not a perfect circular shape in images. A partially illuminated moon has an irregular edge that consists of two arcs with different radiuses as shown in Figure 4.16. If the phase cannot be properly identified, the Hough Transform may return an erroneous result.

To quantitatively describe the lunar phases, a ratio k_m of the illuminated area to the total area of lunar disk can be established. k_m is 0% for the new moon, 50% for the first and last quarters and 100% for the full moon. The value of k_m actually can also be accurately predicted and tabulated for any date in the future [8].

[7] http://en.wikipedia.org/wiki/Lunar_phase
[8] http://aa.usno.navy.mil/data/docs/MoonFraction.php

(a) Original image of a waxing gibbous moon. (b) Two possible solutions via HT.

Figure 4.16. Two possible results of the moon detection via HT.

After the preprocessing, the area of the moon $Area_m$ in the image can be measured directly. If the corresponding k_m is known for sure, the approximate radius R_m of the moon in the image can be calculated as,

$$R_m = \sqrt{\frac{Area_m}{\pi k_m}}. \tag{4.13}$$

As shown in Figure 4.16(b), the incorrect arcs has a larger radius than the desired lunar disk. Using the approximate radius R_m as a condition for the Hough Transform, the detection result can be significantly improved.

Limitations of Lunar Navigation

The usage of the moon as a celestial navigation aid has some limitations that should be noticed. Firstly, the moon is not a permanent celestial aid for all nights. When near the new moon or dark moon, the moon is so close to the sun that it is saturated by sun light and could hardly be observed. After the beginning of a lunar cycle, the moon also rises later and later. A moon at the first quarter may be available for observations during the first half of night but it disappears below the horizon and becomes unobservable after midnight. Secondly, when the lunar phase is close to the

new moon or dark moon, the moon is too small to be accurately identified. After applying the Hough Transform, the two possible results of the lunar disk may also have similar radiuses, making it hard to be distinguished. Thus lunar navigation is more reliable during the period from the waxing crescent moon to the waning crescent moon.

4.3 Experimental Results and Analysis

The field experiment was performed separately for the sun and the moon on the top of Purdue Northwestern Avenue Parking Garage with location parameters of latitude: 40.42935°, longitude: −86.91061° and altitude: 197m on April 22nd, 2013. The weather condition was very good with few clouds. Figure 4.17 shows the overall sensor system we used during the experiment. All images, sensor's attitude data (ψ, θ, φ), and UTC time were collected and processed by MATLAB®.

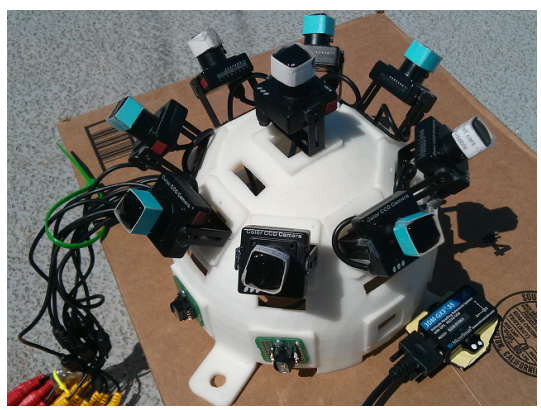


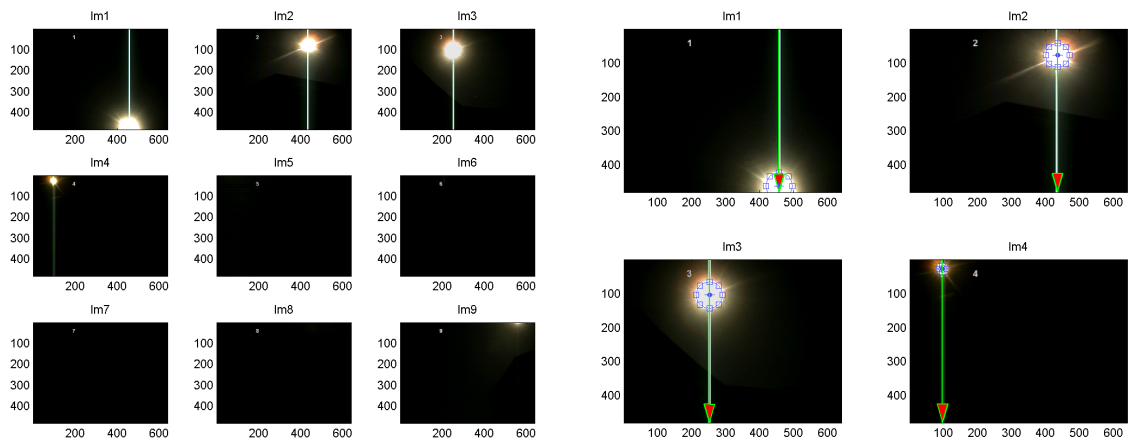
Figure 4.17. Hemispherical sensor and other devices during the experiments.

4.3.1 Estimation of the Sun Vector

As an example, Figure 4.18(a) shows the images acquired at local time 01:49 pm (UTC 05:49 pm). Following the procedure and image processing strategies presented

in Section 4.2.1, the results are shown in Figure 4.18(b) for four images from camera #1, 2, 3 and 4.

To accurately transform the sun center from the pixel position $[u_p^d, v_p^d]$ to the sun vector $\mathbf{e}_s(A, Z)$ in HCS, the distortion model in Section 4.1.2 and the coordinate transformation developed in Section 4.1.1 are employed together with the AHRS data (Yaw: $2.6098rad$, Roll: $0.0193rad$ and Pitch: $0.0122rad$). The results are listed in Table 4.2 for the pixel position of sun center, the corresponding sun vector in HCS and the pixel range of blooming line respectively.



(a) All nine images of the sun before processing. (b) Detection results.

Figure 4.18. Original images and the results of sun detection.

Table 4.2. Detection results and the corresponding sun vectors.

Image No.	Sun center (u_p^d, v_p^d)	Sun vector (A,Z)	Blooming $[v_{pl}^d, v_{pr}^d]$
1	(461,457.25)	(187.1045°,26.6401°)	[455.0,460.5]
2	(77,435.25)	(184.9939°,28.3754°)	[430.3,437.0]
3	(104,252.25)	(185.8093°,28.8089°)	[249.0,256.3]
4	(27,97.1875)	(185.9179°,28.4688°)	[95.7,100.8]

Additionally, using the blooming features detected in Table 4.2, the projections of these blooming lines in HCS are plotted in Figure 4.19. Ideally, all the blooming lines should intersect at the same point where is the sun vector exactly. Figure 4.19 shows a very good result except the intersection of line 1 and 2 since they are closely tangent. Table 4.3 lists the intersections for all pairs of blooming lines.

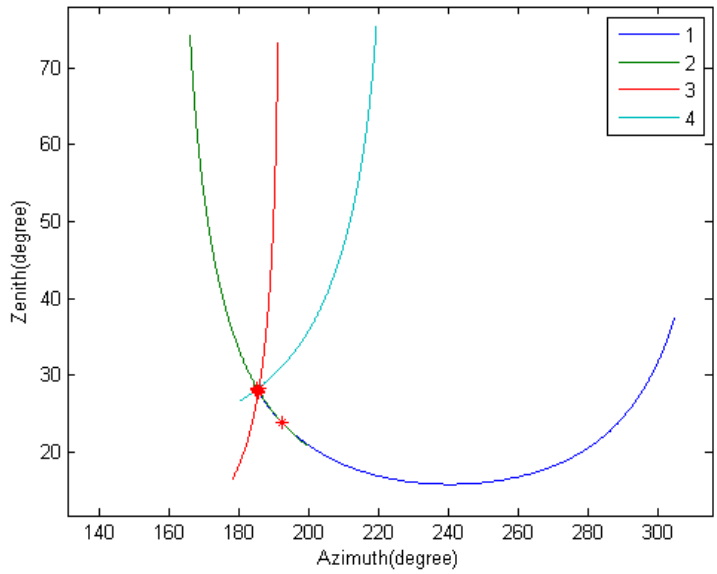


Figure 4.19. Inverse projection of blooming lines with azimuth and zenith angles.

Table 4.3. Intersection points of blooming lines with azimuth and zenith values.

Image No.	Sun azimuth	Sun zenith
1 and 2	192.0749°	23.8715°
1 and 3	185.4382°	27.7731°
1 and 4	184.9516°	28.1275°
2 and 3	185.5117°	27.9717°
2 and 4	185.1994°	28.2136°
3 and 4	185.6576°	28.3754°

Since there are usually more than one result via different approaches and from different images, a proper weighting strategy is needed to evaluate their credibilities:

- Since the distortion issue becomes more serious when closer to the edge of image, the result should be more accurate when the sun is near the image center.
- If any of the disturbances introduced in Section 4.2.1 happens (e.g. Image 1 in Figure 4.18 shows only a part of the sun.), the detected sun center should contribute less to the final result.
- The result via the intersecting of blooming lines should be more accurate than the result via Hough Transform.
- Blooming lines that are close to each other (e.g. line 1 and 2 in Figure 4.19) may introduce an incorrect result and should be excluded.
- If any of the features was chosen manually, the result should be more credible.

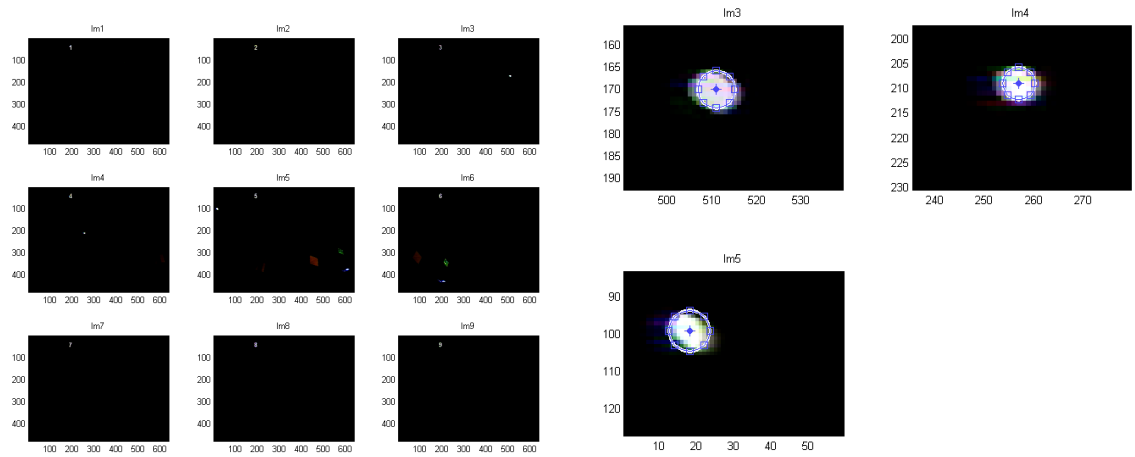
Finally, the sun vector after weighting is presented and compared with the theoretical result in Table 4.4. As illustrated in Section 2.4.1, a magnetic declination of $-4.0313°$ should be compensated.

Table 4.4. Final result of the sun vector after weighting.

Sun vector	Weighted result	After compensation	Theoretical result	Error
Azimuth	185.5873°	181.5560°	181.5758°	0.0198°
Zenith	28.1016°	28.1016°	28.0034°	0.0982°

4.3.2 Estimation of the Moon Vector

The estimation of moon vector basically follows the same procedure as the sun. Figure 4.18(a) shows the original images taken at 09:58 pm, April 22nd (UTC 01:58 am, April 23rd). Cameras #3, 4 and 5 have the moon captured in their images.



(a) All nine images of the moon before processing.

(b) Detection result.

Figure 4.20. Original images and the results of moon detection.

Table 4.5. Detection results and the corresponding moon vectors.

Image No.	Moon center (u_p^d, v_p^d)	Moon Vector (A_m, Z_m)
3	(170,511)	(156.0291°,48.7316°)
4	(209,257)	(157.1419°,47.6790°)
5	(99,18)	(153.5324°,46.8578°)

Considering the lunar phase, the illuminated fraction k_m of the moon was about 84%-87% [9]. The size, i.e. the total number of the pixels in the area, of the moon lies in a range of $[30, 100]$ in images. The results after the preprocessing and the

[9] `http://aa.usno.navy.mil/data/docs/MoonFraction.php`

Hough Transform are shown in Figure 4.20(b). After applying the undistortion and coordinate transformation with the AHRS data (Yaw: $1.2743rad$, Roll: $0.1474rad$ and Pitch: $0.0069rad$), the moon vector results for all three images are listed in Table 4.5. The final result after weighting and the comparison with the theoretical moon vector are listed in Table 4.6.

Table 4.6. Final result of the moon vector after weighting.

Moon vector	Weighted result	After compensation	Theoretical result	Error
Azimuth	155.8423°	151.8107°	150.0238°	1.7869°
Zenith	47.8386°	47.8386°	47.1946°	0.644°

4.3.3 Suggestions for Improvements

The experimental results of zenith angles usually have a high accuracy within 1° while the azimuth results sometimes have a larger error due to the incorrect data of yaw angle. Since most AHRS use local magnetic field for measuring yaw angles, the electromagnetic noises in the environment, e.g. buildings, vehicles or underground minors, may introduce a large bias on azimuth measurements.

The estimation of the moon vector appears to be less accurate than the sun vector because the image sensor usually has more noise in a dark environment. Without a good camera, the moon is too small and blurry to be accurately detected. The improvement via blooming is also not applicable to the moon since the moon is not bright enough to cause CCD saturation.

We present the following suggestions to mitigate possible effects and to improve results:

- Use zenith angles as the priority. If the magnetic azimuth is used, try to avoid the place filled with electromagnetic noises.

- To statistically reduce the error of yaw angle, the AHRS sensor can be employed several times with different sensor attitudes to get an averaged result.
- The camera used for celestial navigation needs to have a higher pixel resolution and less distortion to improve accuracy.
- When taking images of the sun, camera settings should be adjusted adequately to cause proper blooming lines.
- When taking images of the moon, image sensor should have less noise and retain the natural color of moon light.

5. GENERALIZED INTERCEPT METHOD FOR GEOLOCATION

In most common situations, geolocation can be easily achieved by a GNSS receiver with a satisfactory accuracy. But in the cases when GNSS is no longer reliable, the position of the sun or the moon can also accomplish the same task or help to augment the measurements.

As shown in Section 2.4.2, it is possible to use the measurements of local sun or moon vector to mathematically determine the values of latitude and longitude. However, in practice, estimation of the azimuth angle needs the accurate knowledge of geographic true north. This is usually unattainable since there is no permanent natural indicator for the true north[1]. The approximate orientation to the north is usually measured via a magnetic compass which may introduce degrees of error due to the magnetic declination. Without compensating the magnetic declination by the World Magnetic Model, the error on geolocation result will be unacceptable. However, the magnetic model needs accurate location information to provide the theoretical value of local magnetic declination [58]. This is actually impossible since location information is the unknown result we want to estimate. Thus the geolocation approach via Equation (2.3) is impractical if the azimuth angle can not be measured directly from the true north.

Alternatively, the traditional "Intercept Method" by finding the "Line of Position (LOP)" uses the altitudes (or zenith) of two or more celestial objects thus avoids the issue of azimuth measurement. However, it cannot be used for the sun during the daytime since the saturation of sunlight prohibits the measurements of other celestial objects. Thus the application of Intercept Method is limited and cannot be successfully implemented in all situations.

[1]The north star is only available at clear starry night.

This chapter introduces a generalized approach of celestial geolocation developed and improved from the traditional Intercept Method. The method achieves the all-time available geolocation by introducing more available "Set of Position (SOP)" instead of LOP. A fully autonomous algorithm based on an iterative position matching is presented to apply the method directly with computer programs. The method also maintains its robustness to the magnetic declination when using azimuth angles. The experiment results and uncertainty analysis are also presented.

5.1 Generalized Intercept Method

Ideally, an LOP should be a 3D curve on the earth's surface to cross all locations where share the same value of a celestial reference. For instance, a zenith LOP may be approximately drawn as a circle on the map. However, if considering the influences of local terrain and the earth's ellipsoid shape, accurate drawing of LOP will be extremely difficult. Additionally, an LOP only represents a two dimensional feature without considerations for the error and uncertainty that may exist in measurements. To deal with these issues and apply the method for more general cases, define a "Set of Position (SOP)" to replace LOP as,

Definition 4 *The "Set of Position (SOP)" of a measurable celestial reference "***Ref***" at time "t" is an uncountable location set that comprises all qualified locations as,*

$$\mathbf{SOP}(\mathbf{Ref}, t) = \{P(\lambda, \phi, h) \in \mathbb{L} : |f(P, t) - \mathbf{Ref}| \leq \varepsilon\}, \tag{5.1}$$

where $\mathbb{L}$ is a pre-defined set of locations, $f(P, t)$ is an applicable function to use the location and time to determine the theoretical value of **Ref***, and ε denotes the tolerance of measurement error.*

Remark 5.1.1 *SOP is a subset of $\mathbb{L}$ and described by the longitude(λ), latitude(ϕ), and altitude (h) in GCS. In practical usage, $\mathbb{L}$ is preferable to be a set of locations at the same altitude.*

Remark 5.1.2 *A full set of SOP is obtainable if and only if $f(P,t)$ exists and is applicable for all locations in $\mathbb{L}$. Most $f(P,t)$ are directly given by existing astronomical algorithms with a high accuracy. It does not necessarily need to be invertible or unique.*

Remark 5.1.3 *SOP is an empty set when* **Ref** *is impossible to be measured at any location in $\mathbb{L}$.* **Ref** *cannot be unrealistic values such as an azimuth of 420° or a zenith below the horizon.*

Remark 5.1.4 *ε defines the estimated measurement error of* **Ref**. *A smaller value of ε reduces the area of SOP until SOP is equivalent to LOP when $\varepsilon = 0$.*

Instead of using merely zenith SOP, our generalized intercept method employs two more kinds of available SOPs. Considering the azimuth angle of a celestial object, one more SOP can be drawn to indicate all the locations where share the same azimuth. Figure 5.1 is an example graph to show the SOPs of a specific sun vector on the date of March 20th, 2012 (Vernal Equinox), UTC time 12 : 00 pm. The red and blue points indicate the SOPs at sea level ($h = 0$) where the azimuth angle lies in $120° \pm 1°$ and the zenith angle lies in $45° \pm 1°$ respectively. The overlapped area must be the desired location corresponding to the estimated sun vector. Similarly to the traditional intercept method, there may be more than one solutions for a certain time and sun vector as shown in Figure 5.2.

Additionally, if the mobile platform performs a measurable movement between two measurements, the SOP measured formerly can be assumed as moving together with the platform to provide an updated SOP for the next measurement.

Definition 5 *The "successive Set of Position (sSOP)" of a previous SOP after a movement of $\vec{l}$ can be defined as,*

$$\mathbf{sSOP}(\mathbf{Ref}, t, \vec{l}) = \{P'(\lambda, \phi, h) : \ \forall P \in \mathbf{SOP}(\mathbf{Ref}, t), \ P' = P + \vec{l}\}. \tag{5.2}$$

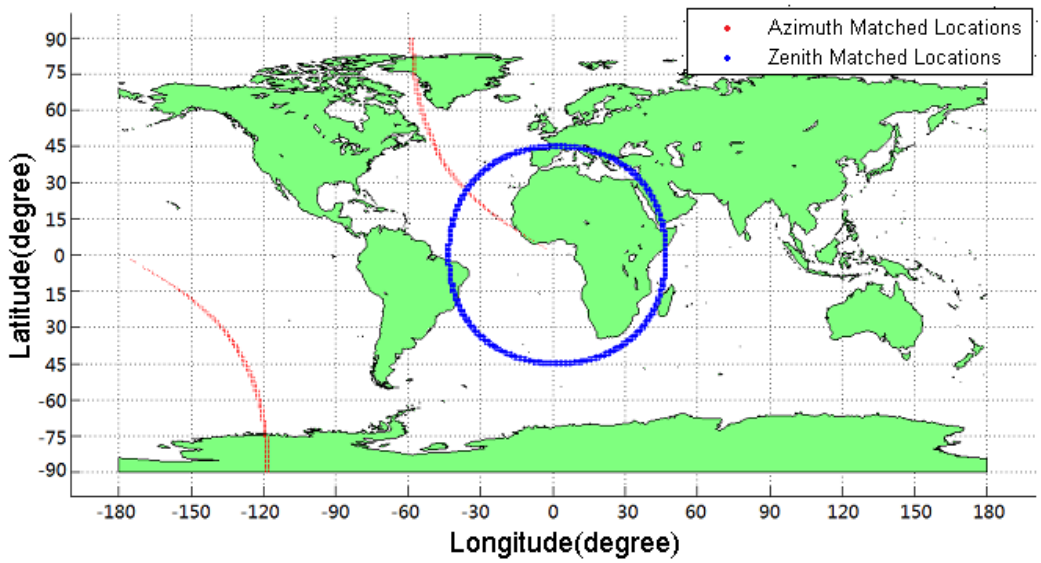


Figure 5.1. Simulated SOPs of the sun with $120° \pm 1°$ azimuth and $45° \pm 1°$ zenith at UTC time noon, March 20th, 2012.

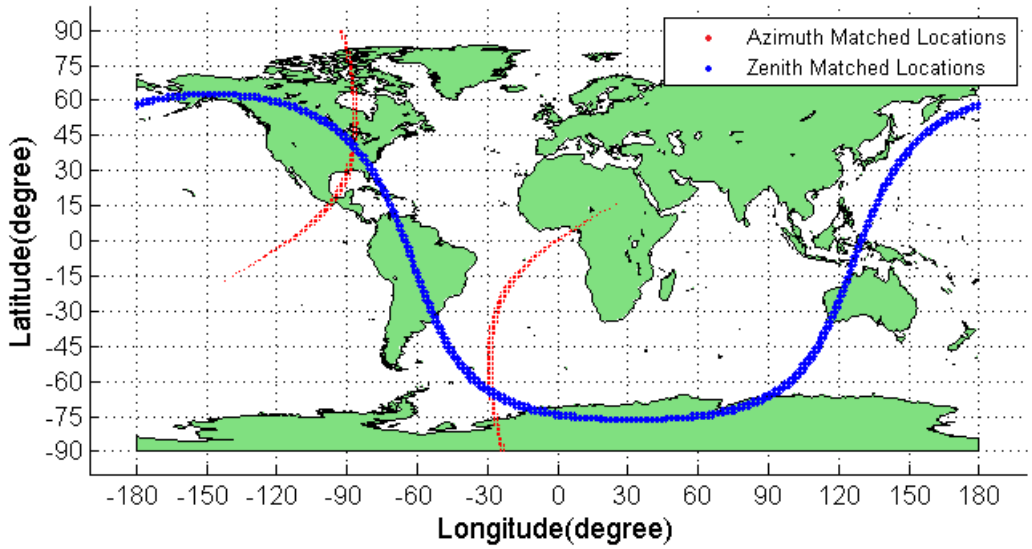


Figure 5.2. SOPs when two overlapped areas exist.

Remark 5.1.5 *The SOP measured at a different location cannot be used directly unless its sSOP at the current location is obtainable. The current location, where* $\vec{l} = \mathbf{0}$*, should be assumed as the origin to determine all other sSOPs.*

Remark 5.1.6 $\vec{l}$ *defines the motion vector of sSOP from the previous SOP in 3D space, with both distance and direction. If only the distance* r *is known for the movement, sSOP can be modified as,*

$$\mathbf{sSOP}(\mathbf{Ref}, t, r) = \{P'(\lambda, \phi, h) : \ \forall P \in \mathbf{SOP}(\mathbf{Ref}, t) \text{ and } |\vec{l}| = r, \ P' = P + \vec{l}\}. \quad (5.3)$$

Thus sSOP is expanded from SOP as shown in Figure 5.3.

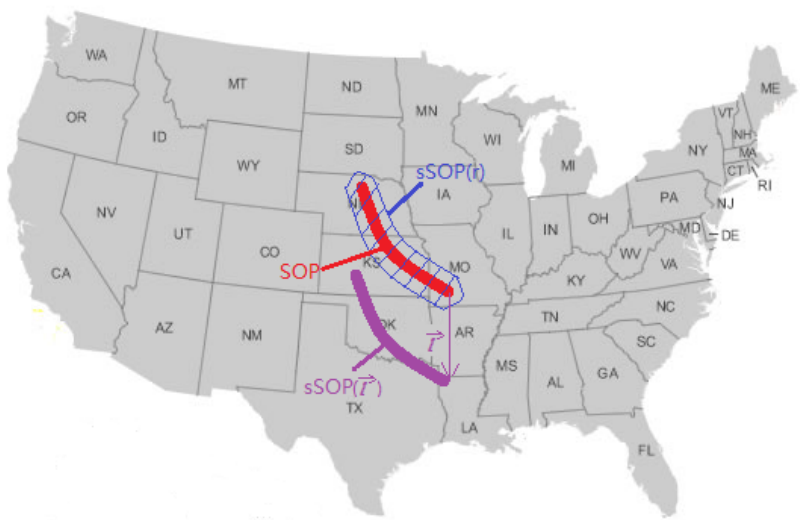


Figure 5.3. Illustration of the sSOP.

Given the above SOP and sSOP, an all time available "Generalized Intercept Method" thus can be presented to apply celestial geolocation in more general cases.

Theorem 5.1.7 *Generalized Intercept Method*
If n *independent SOP and* m *sSOP can be determined at an unknown location* $P(\lambda, \phi, h)$ *and satisfy,*

$$\mathbb{P} = \mathbf{SOP}_1 \cap \mathbf{SOP}_2 \cap \ldots \cap \mathbf{SOP}_n \cap \mathbf{sSOP}_1 \cap \mathbf{sSOP}_2 \cap \ldots \cap \mathbf{sSOP}_m \neq \emptyset, \quad (5.4)$$

then $P(\lambda, \phi, h)$ can be located inside a certain area as,

$$P(\lambda, \phi, h) \in \mathbb{P}. \tag{5.5}$$

Proof Depending on the definition of SOP, if $\mathbf{Ref}_i$ is measured at location $P(\lambda, \phi, h)$ as,

$$|f_i(P, t_i) - \mathbf{Ref}_i| \leq \varepsilon_i, \tag{5.6}$$

then $P(\lambda, \phi, h)$ must locate in the SOP of $\mathbf{Ref}_i$ as,

$$P(\lambda, \phi, h) \in \mathbf{SOP}_i(\mathbf{Ref}_i, t_i),\ i = 1, 2, \ldots, n. \tag{5.7}$$

Similarly, if a previous $\mathbf{SOP}_j(\mathbf{Ref}_j, t_j)$ is measured at the location P_j and the displacement between P_j and the current location P is known as,

$$P = P_j + \vec{l_j}, \tag{5.8}$$

the definition of sSOP gives,

$$P(\lambda, \phi, h) \in \mathbf{sSOP}_j(\mathbf{Ref}_j, t_j, \vec{l_j}),\ j = 1, 2, \ldots, m. \tag{5.9}$$

Applying Equation (5.7), Equation (5.9) and the property of set intersection,

$$P(\lambda, \phi, h) \in \mathbb{P}, \tag{5.10}$$

given $\mathbb{P}$ is the intersection of all SOP and sSOP as,

$$\mathbb{P} = \mathbf{SOP}_1 \cap \mathbf{SOP}_2 \cap \ldots \cap \mathbf{SOP}_n \cap \mathbf{sSOP}_1 \cap \mathbf{sSOP}_2 \cap \ldots \cap \mathbf{sSOP}_m. \tag{5.11}$$

■

Definition 6 *The "Geographic Size (GS)" of intersection* $\mathbb{P}$ *is defined as the smallest volume in the 3D space that comprises all locations in* $\mathbb{P}$ *as,*

$$GS(\mathbb{P}) = \iiint\limits_{\mathbb{P}} g(\lambda, \phi, h) d\lambda \, d\phi \, dh, \tag{5.12}$$

where $g(\lambda, \phi, h)$ *is the function to use geographic coordinates to calculate the volume in 3D space and,*

$$GS(\mathbb{P}) \geq 0 \ with \ GS(\mathbb{P}) = 0 \ if \ and \ only \ if \ \mathbb{P} = \emptyset. \tag{5.13}$$

Remark 5.1.8 *When one or more values of* λ*,* ϕ *or* h *is known for sure,* $GS(\mathbb{P})$ *can also represent the area on two dimensional space or the length of a geographic line on the earth.*

Corollary 5.1.9 *Corollary A of the Generalized Intercept Method*
At location $P(\lambda, \phi, h)$*, if adding a new SOP or sSOP of* $\mathbf{S}$ *and* $\mathbf{S}$ *is not a superset of* $\mathbb{P}$ *as,*

$$\mathbb{P}' = \mathbb{P} \cap \mathbf{S} \neq \mathbb{P}, \tag{5.14}$$

the geographic size of the intersected area will become smaller,

$$GS(\mathbb{P}') = GS(\mathbb{P} \cap \mathbf{S}) < GS(\mathbb{P}), \tag{5.15}$$

but never reach zero as,

$$GS(\mathbb{P}') > 0. \tag{5.16}$$

Proof Since all SOP and sSOP include the location $P(\lambda, \phi, h)$,

$$P \in \ all \ \mathbb{P}' \Rightarrow GS(\mathbb{P}') > 0. \tag{5.17}$$

Since $\mathbf{S}$ is not a a superset of $\mathbb{P}$, define a set $\mathbf{B}$ to contain all the locations that exist in $\mathbb{P}$ but not in $\mathbf{S}$ as,

$$\mathbf{B} = \{P(\lambda, \phi, h) : P \in \mathbb{P} \; and \; P \notin \mathbf{S}\} = \mathbb{P} \cap \mathbf{S}^c. \tag{5.18}$$

Applying the Distributive Laws of set,

$$\mathbf{B} \cap \mathbb{P}' = (\mathbb{P} \cap \mathbf{S}^c) \cap (\mathbb{P} \cap \mathbf{S}) = \mathbb{P} \cap (\mathbf{S} \cap \mathbf{S}^c) = \mathbb{P} \cap \emptyset = \emptyset, \tag{5.19}$$

$$\mathbf{B} \cup \mathbb{P}' = (\mathbb{P} \cap \mathbf{S}^c) \cup (\mathbb{P} \cap \mathbf{S}) = \mathbb{P} \cap (\mathbf{S} \cup \mathbf{S}^c) = \mathbb{P} \cap \mathbf{U} = \mathbb{P}. \tag{5.20}$$

Thus,

$$\begin{aligned} GS(\mathbb{P}) &= \iiint\limits_{\mathbf{B} \cup \mathbb{P}'} g(\lambda, \phi, h) d\lambda \, d\phi \, dh \\ &= \iiint\limits_{\mathbf{B}} g(\lambda, \phi, h) d\lambda \, d\phi \, dh + \iiint\limits_{\mathbb{P}'} g(\lambda, \phi, h) d\lambda \, d\phi \, dh - \iiint\limits_{\mathbf{B} \cap \mathbb{P}'} g(\lambda, \phi, h) d\lambda \, d\phi \, dh \\ &= GS(\mathbf{B}) + GS(\mathbb{P}') - GS(\emptyset) \\ &= GS(\mathbf{B}) + GS(\mathbb{P}'). \end{aligned} \tag{5.21}$$

Since,

$$GS(\mathbf{B}) > 0, \tag{5.22}$$

it is obvious that,

$$GS(\mathbb{P}') < GS(\mathbb{P}). \tag{5.23}$$

■

From Equation (5.21), it is clear that a larger $GS(\mathbf{B})$ yields a preferred smaller $\mathbb{P}'$. After adding more SOP or sSOP to intersect with $\mathbb{P}$, it is possible to find a sufficient small area for the need of geolocation purpose.

Definition 7 *If there is a point $P_o(\lambda_o, \phi_o, h_o)$ in $\mathbb{P}$ can be found as,*

$$\forall P(\lambda, \phi, h) \in \mathbb{P} : |P_o - P| \leq \delta, \tag{5.24}$$

with a sufficient small δ, $P_o(\lambda_o, \phi_o, h_o)$ is a "Representative Location" of $\mathbb{P}$ and qualified as the final geolocation result.

Remark 5.1.10 *There may be countless P_o that satisfy Equation (5.24) or none P_o exists at all. If $GS(\mathbb{P})$ is small enough, any P_o that is close to other locations in $\mathbb{P}$ within an acceptable small δ is qualified as a Representative Location. However, when $\mathbb{P}$ is a widespread area on the earth, P_o cannot be determined within a desired accuracy. In practice, if a small $\mathbb{P}$ is achieved, we can usually use the geographic center of $\mathbb{P}$ as P_o.*

Narrowing a smaller $\mathbb{P}$ is the ultimate goal of our method. Thus **Ref** with distinct SOP and sSOP should be chosen and intersected together to make the overlapped area converge quickly. In fact, celestial objects can provide diverse choices to make this possible.

Corollary 5.1.11 *Corollary B of the Generalized Intercept Method*
Ideally it is possible to use, but not limited to, the following choices of at least two SOPs to apply the Generalized Intercept Method.

- *Zenith (altitude) angles of two independent celestial objects at the same location:* $\mathbf{SOP}(Z_1, t_1)$ *and* $\mathbf{SOP}(Z_2, t_2)$.
- *Both azimuth and zenith angles of a single celestial object at the same location:* $\mathbf{SOP}(A_1, t_1)$ *and* $\mathbf{SOP}(Z_1, t_2)$.
- *Two successive measurements at different locations after a measurable movement:* $\mathbf{sSOP}(\mathbf{Ref}, t_1, \vec{l})$ *and* $\mathbf{SOP}(\mathbf{Ref}, t_2)$.

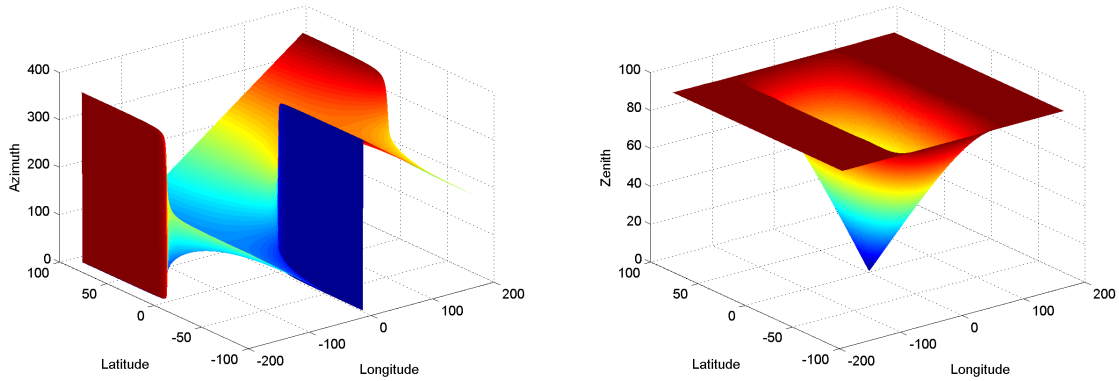
Remark 5.1.12 *When more than one celestial objects are available, choice 1 is preferred since zenith angle is more likely to be measured accurately than azimuth. However, choice 2 can be applied for situations even with only one celestial object (e.g. the sun during the daytime and the moon at cloudy night etc.). Choice 3 can be used to augment the results of other choices when* $\vec{l}$ *is obtainable. It is optional but highly recommended to use more combinations to improve the result.*

5.2 Iterative Position Matching Algorithm

Like the traditional Intercept Method, manually conducting the Generalized Intercept Method needs experienced operators. The usage of SOP instead of LOP significantly increases the computational burden, making it tedious and almost impossible as a manual work.

The approach, however, can be easily implemented by efficient computers. Standard astronomical algorithms for calculating the local sun vector and moon vector are already available in [16, 44]. The corresponding computer program and database are also provided in [59,60] with uncertainties of $\pm 0.0003°$. Similarly, calculations of CRV for most bright stars and planets can also be easily programmed. These algorithms and computer programs enable fast and reliable calculations of $f(P, t)$ in Equation (5.1).

SOP is usually an uncountable set contains infinite location elements. Given a celestial reference value **Ref** and its uncertainty range ε, we need to find all locations that satisfy Equation (5.6) to form its SOP. It is easy to use $f(P, t)$ to figure out the theoretical **Ref** at a given location while it is impossible to apply $f(P, t)$ for all locations on the earth. If dividing the earth by a $1°$ interval on both longitude and latitude, there will be total $360 \times 180 = 64800$ executions of Equation (5.6) to determine whether the locations are in the SOP or not. Considering the fact that there are several SOPs and an accuracy within $0.01°$ is desired for geolocation, finding SOP via enumeration is inefficient and should be avoided.



(a) Distribution of the sun azimuth.

(b) Distribution of the sun zenith.(All zenith angles that are larger than 90° are revalued to 90° since the sun is under the horizon.)

Figure 5.4. Global sun vector distribution at UTC time noon, March 20th, 2012.

Figure 5.4(a) and Figure 5.4(b) are two graphs to show the global sun vector distributions of both the azimuth and zenith angles at zero altitude on March 20th, 2012 (Vernal Equinox), UTC time 12 : 00 pm. The distributions on other dates are also similar. It is clear that the distribution of zenith and most part of azimuth are smooth surfaces and follow an increasing or decreasing tendency in certain ranges. This stimulates the following idea of an iterative matching algorithm on map grids.

Definition 8 *Given $\mathbb{L}$ is a pre-defined area of locations on the same altitude h with longitude ranges in $[\lambda_{min}, \lambda_{max}]$ and latitude ranges in $[\phi_{min}, \phi_{max}]$,*

- *The "Grid of Locations (GoL)" of $\mathbb{L}$ by $\delta_{GoL}(\delta_\lambda, \delta_\phi)$ is the map grid that separates $\mathbb{L}$ into small areas by an interval of δ_λ on $[\lambda_{min}, \lambda_{max}]$ and by δ_ϕ on $[\phi_{min}, \phi_{max}]$.*

- *The "Position Nodes (PN)" is the set of intersected locations in the GoL as,*

$$PN(i,j) = (\lambda(i), \phi(j), h) = (\lambda_{min} + (i-1)\delta_\lambda,\ \phi_{min} + (j-1)\delta_\phi,\ h), \quad (5.25)$$

where $i = 1, 2, \ldots, (\frac{\lambda_{max}-\lambda_{min}}{\delta_\lambda} + 1)$ and $j = 1, 2, \ldots, (\frac{\phi_{max}-\phi_{min}}{\delta_\phi} + 1)$.

With proper values of δ_λ and δ_ϕ, a GoL thus can be built to divide the geographic map into small areas (Figure 5.5). Since (λ, ϕ, h) of all PN are known with certain values in the GoL, it is possible to use $f(P, t)$ to determine the theoretical value of **Ref** at every $PN(i, j)$. Given the zenith and azimuth distribution graphs in Figure 5.4, the theoretical values of **Ref** should be close for neighboring locations. If $\mathbf{Ref}_k$ is used to draw the SOP, a threshold ε'_k can be used to select the qualified location nodes in PN to form a set of PN_k as,

$$PN_k = \{PN(i, j) : |f_k(PN(i, j), t) - \mathbf{Ref_k}| \leq \varepsilon'_k\}, \tag{5.26}$$

for all (i, j).

If total K $(K \geq 2)$ **Ref** are employed, a smaller set of PN' can be created by taking the intersection of all PN_k to comprise the location nodes that are qualified for all **Ref**,

$$PN' = PN_1 \cap PN_2 \cap \ldots \cap PN_K. \tag{5.27}$$

Since the theoretical values of $\mathbf{Ref}_{1,2,\ldots,K}$ at the location nodes in PN' are closer to the measured values than all other location nodes in PN, the estimated location $P_o(\lambda_o, \phi_o, h)$ is most likely to be enclosed in the area $[\lambda'_{min}, \lambda'_{max}, \phi'_{min}, \phi'_{max}]$ where,

$$\lambda'_{min} = min(\lambda_{PN'}), \tag{5.28}$$

$$\lambda'_{max} = max(\lambda_{PN'}), \tag{5.29}$$

$$\phi'_{min} = min(\phi_{PN'}), \tag{5.30}$$

$$\phi'_{max} = max(\phi_{PN'}). \tag{5.31}$$

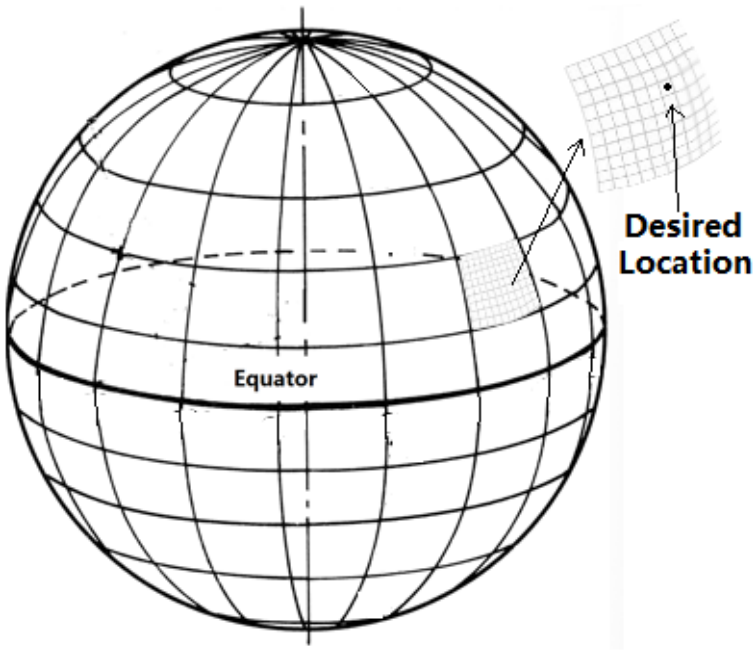


Figure 5.5. Iterative point matching by narrowing the GoL.

Among the area of $[\lambda'_{min}, \lambda'_{max}, \phi'_{min}, \phi'_{max}]$, a new GoL and a new set of PN thus can be yielded by a smaller δ_{GoL} and $\varepsilon'_{1,2\ldots,K}$. Repeating the previous steps iteratively until,

- $[\lambda'_{min}, \lambda'_{max}, \phi'_{min}, \phi'_{max}]$ becomes small enough to find a $P_o(\lambda_o, \phi_o, h)$ in Equation (5.24) or,
- All $\varepsilon'_{1,2,\ldots,K}$ achieves the desired accuracy for $\mathbf{Ref}_{1,2,\ldots,K}$.

The final geolocation result thus can be described either as a location of $P_o(\lambda_o, \phi_o, h)$ or the area of PN as $[\lambda'_{min}, \lambda'_{max}, \phi'_{min}, \phi'_{max}]$. Figure 5.6 shows the general procedure of how to implement our iterative matching algorithm.

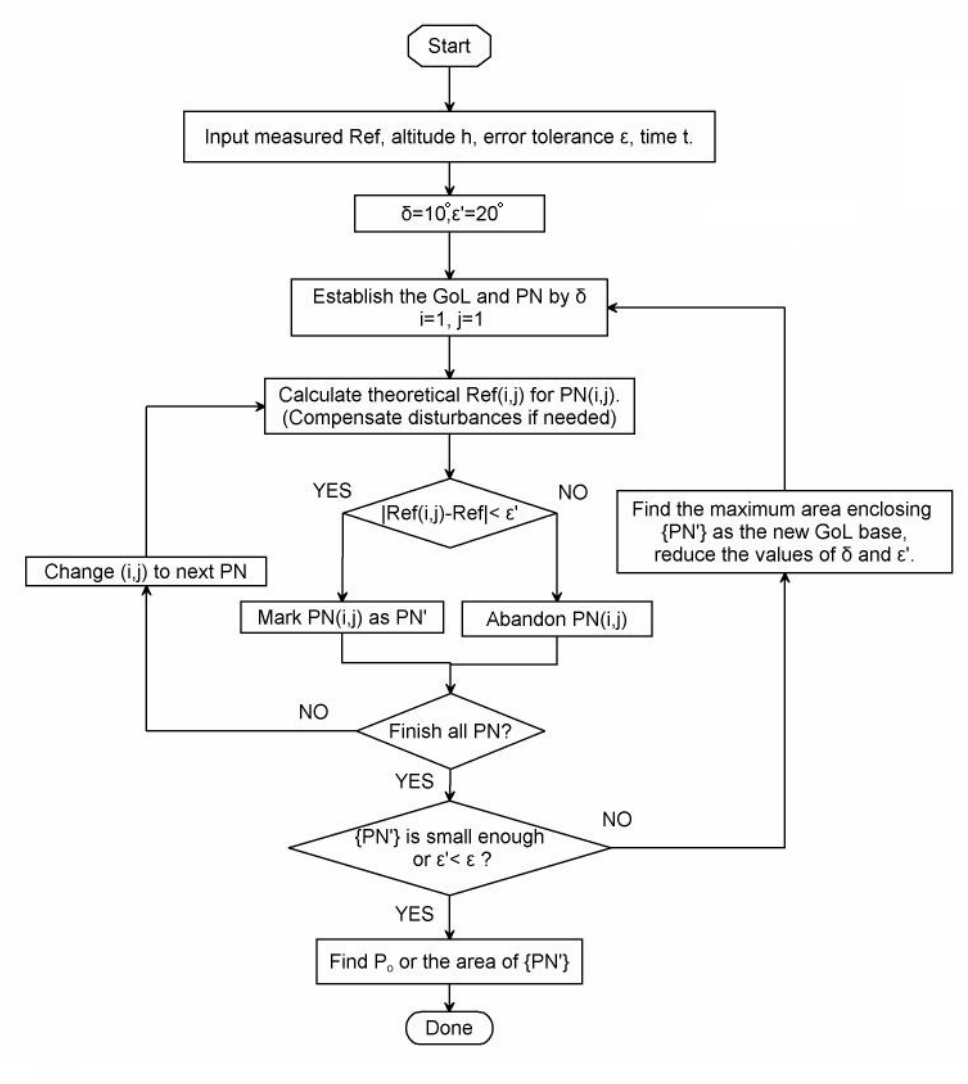


Figure 5.6. Diagram of the iterative matching algorithm.

Remark 5.2.1 *If there are more than one solutions as shown in Figure 5.2, it can be easily found by checking the continuity of* $\lambda_{PN'}$ *or* $\phi_{PN'}$*. From* λ'_{min} *to* λ'_{max} *or from* ϕ'_{min} *to* ϕ'_{max}*, if the interval between two successive location nodes is not equal to* δ_{GoL}*, the original GoL can be divided into two sub-GoLs upon this gap. Two solutions thus can be determined separately in two sub-GoLs through the same iterative matching algorithm.*

Remark 5.2.2 *The final result, if described as* $[\lambda_{min}, \lambda_{max}, \phi_{min}, \phi_{max}]$*, is a rectangular area and may not accurately represent the overlapped area of all SOPs. In very limited cases, the estimated location can be outside the rectangular area even all the nodes in* PN *are qualified locations.*

Remark 5.2.3 *Two SOPs may have a large overlapped area as shown in Figure 5.7. In this case, the iterative matching algorithm may give an undesired large rectangular area instead of a small area that could be represented by* $P_o(\lambda_o, \phi_o, h)$*. Thus more SOP or sSOP is needed to improve the result.*

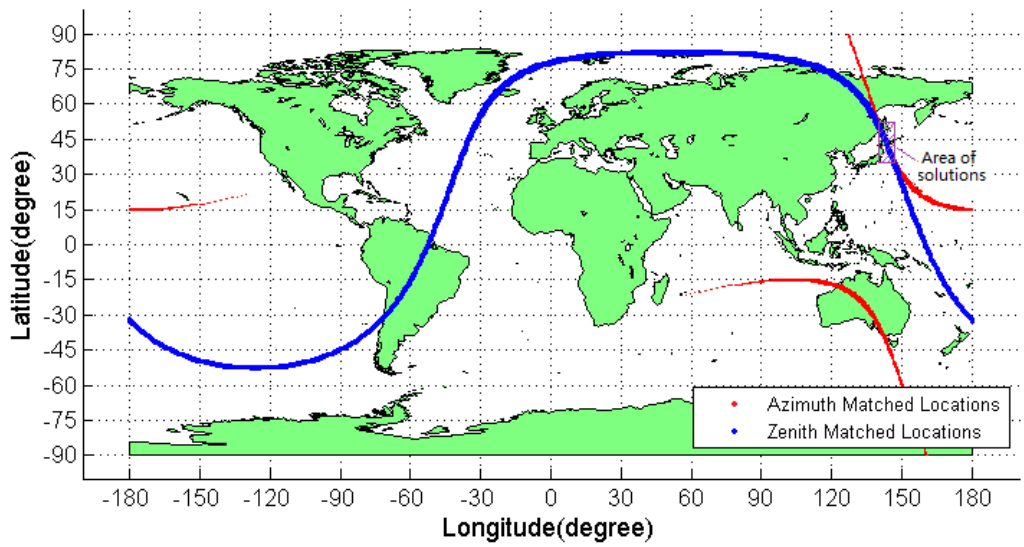


Figure 5.7. Example azimuth and zenith SOPs with a large overlapped area.

Remark 5.2.4 *If azimuth is used as the* **Ref***, it is necessary to consider the magnetic declination when the north is measured by magnetic sensors. The iterative matching algorithm can solve the issue easily since the theoretical magnetic declination* Dec_M *for all location nodes can be compensated directly given the time and* $P(\lambda, \phi, h)$[2]*. The criterion in Equation (5.26) thus becomes,*

$$|f_A(PN(i,j),t) - Dec_M(PN(i,j),t) - A| \leq \varepsilon_A, \tag{5.32}$$

[2] `http://www.mathworks.com/help/aerotbx/ug/wrldmagm.html`

for all steps to maintain the correctness of result.

However, after considering the magnetic declination, the global azimuth distribution of Figure 5.4(a) changes to Figure 5.8. This may introduce some locations where the continuity of azimuth angle no longer exists.

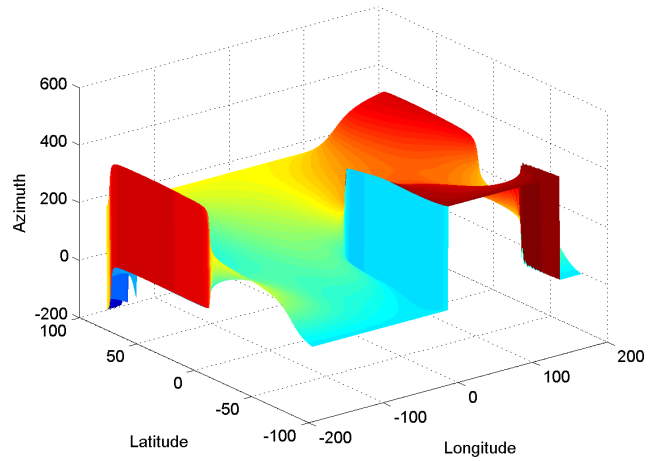


Figure 5.8. Global distribution of sun azimuth after compensating the magnetic declination. (UTC time at noon on March 20th, 2012)

Compared to other geolocation approaches, the generalized intercept method through iterative matching algorithm has following advantages:

- Fast. For each iteration, only a limited number of location nodes in PN needs to be calculated, which largely reduces the computational burden.
- Accurate. With consideration of measurement error ε for each **Ref**, the approach uses an area instead of a single point to represent the solution, which is very useful for the uncertainty analysis.
- Robustness to disturbances. Some disturbances, e.g. the atmospheric refraction and magnetic declination, can be quantitatively compensated in Equation (5.26). Other undetermined errors can be included in ε thus the final geolocation area always gives an estimation of the desired location.

- Flexible usage with different choices of celestial references as long as $f(P,t)$ in Equation (5.1) is well defined.
- Fully autonomous and programmable. All iterations, calculations of **Ref**, and compensations for disturbances can be easily programmed into computer functions for direct use.

5.3 Applications and Tests

In this section, the results from Section 4.3 will be used to illustrate our method and to show the geolocation results.

Directly applying the results from Table 4.4, Table 5.1 lists the geolocation inputs by the aid of the sun. The sun vector $\mathbf{e}_S(A_{Mag}, Z, t)$ with azimuth angle measured from the local magnetic north is used as **Ref**. The measurement errors $(\varepsilon_A, \varepsilon_Z)$ indicate the possible biases between the estimated sun vector and the true sun vector values due to disturbances. Altitude and the UTC time are also needed and should be estimated via other approaches.

Table 5.1. Geolocation inputs from Table 4.4.

Inputs	A_{Mag}	Z	$(\varepsilon_A, \varepsilon_Z)$	Altitude	UTC Time
Values	185.5873°	28.1016°	(0.2°,0.1°)	197 m	17:49 04/22/2013

The program begins with a global grid and continuously follows Equation (5.32) and Equation (5.28) to narrow the size of grid until the final error thresholds $(\varepsilon'_A, \varepsilon'_Z)$ become the same, or even smaller, as the measurement error of $(\varepsilon_A, \varepsilon_Z)$. The results for all iterations are given in Table 5.2. To avoid the discontinuity at longitude -180° and 180°, the method uses $[180°, 360°]$ to represent the longitudes in $[-180°, 0°]$. The intervals $\delta_{GoL}(\delta_\lambda, \delta_\phi)$ for dividing the grid should always be small enough to divide the area into at least 5×5 sub areas.

Table 5.2. Iterative steps for the geolocation of Table 5.1.

Step	$[\lambda'_{min}, \lambda'_{max}]$	$[\phi'_{min}, \phi'_{max}]$	$(\delta_\lambda, \delta_\phi)$	$(\varepsilon'_A, \varepsilon'_Z)$
1	[0°,360°]	[-90°,90°]	(10°,10°)	(20°,20°)
2	[270°,280°]	[30°,60°]	(2°,5°)	(10°,10°)
3	[270°,276°]	[35°,50°]	(1°,2.5°)	(5°,5°)
4	[271°,274°]	[37.5°,45°]	(0.5°,1.25°)	(2.5°,2.5°)
5	[272.5°,273.5°]	[38.75°,42.5°]	(0.2°,0.625°)	(1.25°,1.25°)
⋮	⋮	⋮	⋮	⋮
8	[273°,273.15°]	[40.3125°,40.78°]	(0.025°,0.078°)	(0.2°,0.15625°)
9	[273.025°, 273.15°]	[40.390°,40.625°]	(0.0125°,0.039°)	(0.2°,0.1°)
Result	[273.025°, 273.15°]	[40.4297°,40.625°]		

After normalization of longitudes, the geolocation result of an area described by $[-86.975°, -86.85°, 40.4297°, 40.625°]$ is achieved with nine iterations. The area and all the nodes in final PN are shown in Figure 5.9 and the arrow indicates the experiment location.

An accurate geolocation needs better estimations of celestial objects, i.e. small ε, to reduce uncertainties. Ideally, if the estimated sun vector equals the true value[3] and the errors are as small as $(0.01°, 0.01°)$, the geolocation result is shown in Figure 5.10(a) with an area of $[-86.9141°, -86.9078°, 40.41992°, 40.43457°]$. Since all the locations in the area are within one mile's distance, it is possible to find out a "Representative Location" in Equation (5.24) to represent the final result. For simplicity, we can use the geographic center of the area $(40.4272°, -86.9109°)$ as P_o. The result is only about 0.15 mile away from the real experiment location.

[3]The theoretical sun vector is $(A, Z) = (181.5758°, 28.0034°)$. Considering the magnetic declination of $-4.0313°$, $(A_{Mag}, Z) = (185.6071°, 28.0034°)$.

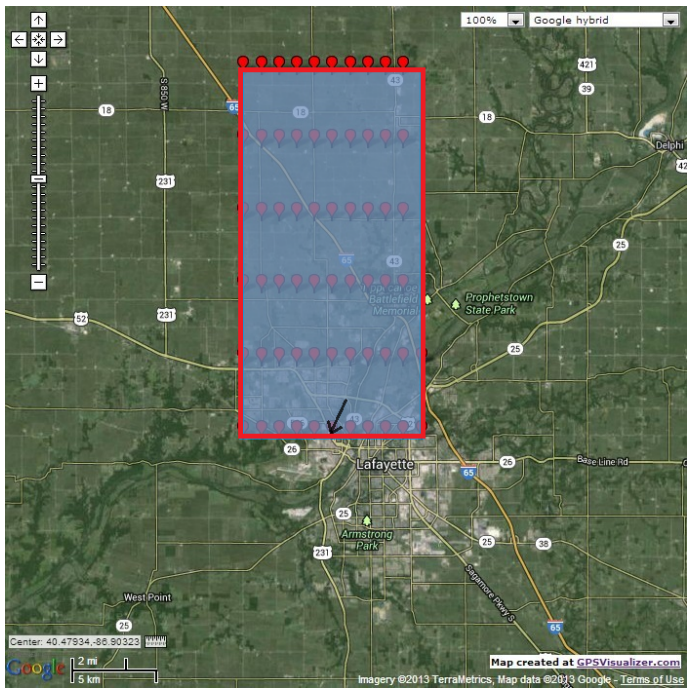
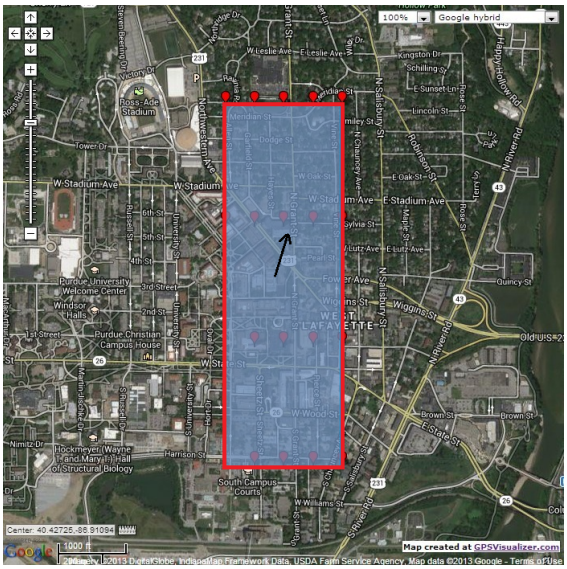
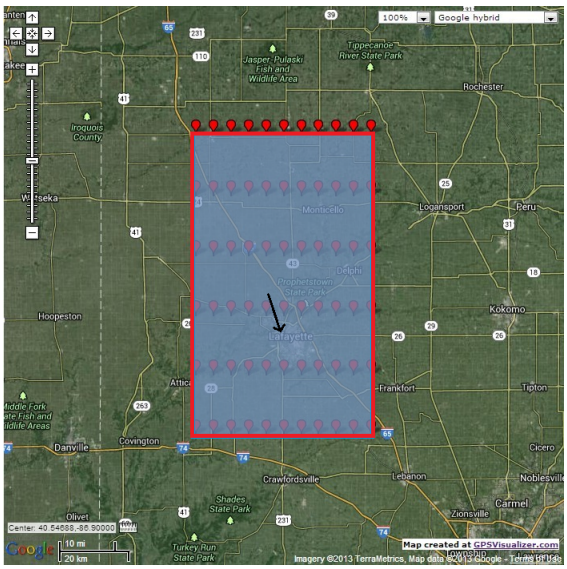


Figure 5.9. Geolocation result of Table 5.1.



(a) Geolocation result of the true sun vector with $\varepsilon_A = 0.01°$ and $\varepsilon_Z = 0.01°$.



(b) Geolocation result of Table 5.1 with $\varepsilon_A = 1°$ and $\varepsilon_Z = 0.5°$.

Figure 5.10. Geolocation results with different measurement errors.

However, in real situations, the system must contain some extent of errors due to various disturbances. Figure 5.10(b) provides the geolocation result of the same inputs as in Table 5.1 but with larger measurement errors of $(\varepsilon_A, \varepsilon_Z) = (1°, 0.5°)$. The result is an area of $[-87.2°, -86.6°, 40.156°, 40.937°]$ with a larger size. The area is too widespread to find the P_o but still accurately indicates an approximate estimation that includes the desired location.

5.4 Uncertainty Analysis and Improvements

A good navigation approach requires an exhaustive consideration of various uncertainties. Table 5.3 lists the possible source of uncertainties in our system.

Table 5.3. Possible source of uncertainties for celestial geolocation.

Source of uncertainties	Notation	Causes
Camera sensor	ε_{cam}	Poor quality, inaccurate calibration, or misalignment of cameras.
AHRS	ε_{AHRS}	System error from the accelerometer and magnetic sensor, noises in the environment.
Clock	ε_T	UTC time is incorrect due to an inaccurate clock.
Altitude	ε_{al}	Incorrect measurement of altitude from the sea level.
Image processing	ε_{im}	Anything that deteriorates image quality.
Atmospheric effect	ε_{at}	Atmospheric refraction, dispersion, turbulences etc.
Geometry of the earth	ε_{geo}	Earth's spheroid shape, local terrains.
Computational error	ε_{cal}	Inaccurate computation of **Ref** via algorithms.
All others	ε_{ot}	NA.

In the table, ε_{cam}, ε_{AHRS}, ε_T and ε_{al} mainly exist for the hardware devices. The most straightforward way to reduce those errors is to use more precise devices. Because most **Ref** changes slowly, most modern clocks with an error less than 10 seconds are adequate to provide an almost negligible ε_T. Considering the tremendous distance

to celestial objects, the ε_{al} influenced by the altitude measurement is also minimized when the error is less than 100 meters. To mitigate ε_{cam} and ε_{AHRS}, both good sensors and a clean environment with minimum amount of noise are highly desired.

To secure a small ε_{im} in the image processing, a good camera sensor is the prerequisite. Image with a large resolution and minimal noise is extreme useful. If a celestial object is captured in more images by increasing the number of camera or through more measurements, the averaged result of these individual images should become more credible. Section 4.2 has introduced the basic procedure and methods to achieve accurate detections.

Atmospheric effects cause the light from celestial objects to be absorbed, distorted or scattered by the particles in atmosphere, adding the difficulty of accurate celestial measurements. Celestial navigation on the earth prefers a clear sky to avoid the influence of atmospheric effects. When using the sun or the moon as the reference, light dispersion and turbulences contribute minimally to ε_{at} as the sun and the moon are much brighter than the dispersed light. However, the turbulence may blur the stars and make them twinkle or even cause small random motions in images. Among all atmospheric effects, the refraction contributes ε_{at} mostly, especially when the celestial object is close to the horizon. Using the model developed in [22], the effect can be compensated theoretically. In fact, the algorithms used in Section 5.2 for calculating the sun and moon vector have convenient built-in compensations for general atmospheric refractions. The mitigation for ε_{geo} is similar since both algorithms use an accurate model of the earth to eliminate geographic uncertainties.

The computational error ε_{cal} is an unavoidable error during computations of **Ref**. For instance, the calculation of magnetic declination relies on the Global Magnetic Model, which is continuously influenced by the unpredictable geomagnetic activities and needs to be revalued for every five years. In addition, the round-off error caused by reducing digits may happen during the computer processing. The error may accumulate, sometimes even override the results, if a sequence of computational error happens during the whole process. But since more mature algorithms and powerful

computers have been employed for scientific usages, ε_{cal} usually can be limited to a negligible range.

In summary, for uncertainties caused by the misalignment of cameras, lens distortion, clock or altitude measurements, most of their influences can be eliminated after an exhaustive calibration and adjustment of hardware devices. Effects of the atmospheric refraction, the earth's geometry and the uncertainty in algorithms may be quantitatively determined and compensated during the calculation of **Ref**. Other uncertainties which cannot be determined directly, e.g. ε_{im} and ε_{AHRS}, should be included in ε as measurement errors. Regardless the values of these errors, our algorithm always gives a solution area has the desired location inside. The area is preferred to be small but will become larger when ε increases.

Sometimes the value of ε is too large to yield a usable geolocation area. Under this scenario, if more than one measurements can be conducted independently, the uncertainty of geolocation result may be significantly reduced by intersecting the solution areas.

For instance, Table 5.4 is another group of data measured at the same location. Through the same approach, a result of $[-86.9531°, -86.8125°, 40.2734°, 40.4688°]$ is achieved as shown in Figure 5.11. Since both two areas in Figure 5.9 and Figure 5.11 are valid results, the desired location must lie in the intersected area as proved in Equation (5.4). The result after intersection is shown with black color in Figure 5.12 as a much smaller geolocation result of $[-86.9531°, -86.85°, 40.4297°, 40.4688°]$.

Table 5.4. Second group of inputs for solar geolocation.

Inputs	A_{Mag}	Z	$(\varepsilon_A, \varepsilon_Z)$	Altitude	UTC Time
Values	200.09°	28.755°	(0.2°,0.1°)	197 m	18:17 04/22/2013

Figure 5.11. Geolocation result of Table 5.4.

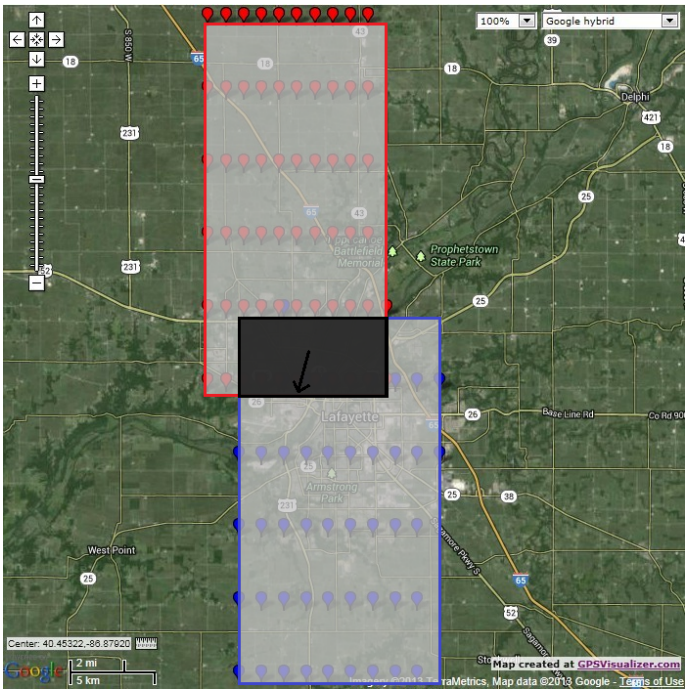


Figure 5.12. Intersected geolocation result of Table 5.1 (red) and Table 5.4 (blue).

6. CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

We have shown that, upon a hemispherical sensor structure, an arbitrary array of light intensity sensors or cameras can be used to detect celestial objects and estimate their exact positions in the sky. The array with light intensity sensors, though susceptible to noise and disturbance, achieves an accuracy of about 5° for the sun azimuth and 1° for the sun zenith during experiments and was successfully applied as a solar compass. Due to its simple structure, low cost and weight, the system can be easily mounted on a small-sized mobile platform or device. The camera based sensing system, however, uses image processing to accurately detect celestial objects in images even with influences of noise and disturbance. Using an elaborately calibrated camera model, the target can be estimated to a high accuracy, especially for the zenith angle.

Also presented is the generalized intercept method of using celestial objects to determine exact locations on the earth even without the need for skilled operators. Based on the experimental results, it accurately narrows the desired location to inside an area through a limited number of iterations. An impressive accuracy of 0.15 mile can be achieved if the sun azimuth and zenith could be measured to within an error of 0.01°. Geolocation to within a small area may also be achieved by intersecting the results from multiple individual measurements over time.

Applications of the system span a diverse array of areas. It can be used in the field of solar electricity generation to aid solar panels facing toward the sun to receive maximum solar energy. Since the system is fully independent of satellites, it can be widely used as a supplement of modern GNSS systems on marine vessels, aircraft, ground vehicles, and other mobile platforms.

6.2 Future Work

- Undeniably, the weather conditions limit the usage of celestial navigation. Our future work needs to explore the possibility of observing celestial objects even through the clouds.

- Since the AHRS sensor, especially the magnetic sensor for measuring the yaw angle, is the main source of uncertainties, it is necessary to explore more aids to accurately determine the true/magnetic norths and the gravity vector.

- The array includes a total of nine individual cameras for a 360° vision which significantly increases the cost and system complexity. A fisheye lens and a single imaging sensor may be sufficient to accomplish the same task.

- Although the algorithms for calculating the sun vector and moon vector are provided and available as programs, similar functions for bright planets and stars still need to be formulated if they are to be employed for celestial navigation.

LIST OF REFERENCES

LIST OF REFERENCES

[1] Benjamin Dutton. *Dutton's Nautical Navigation*. Naval Institute Press, 2004.

[2] Christopher J Hegarty and Eric Chatre. Evolution of the global navigation satellite system (gnss). *Proceedings of the IEEE*, 96(12), 2008.

[3] Michael G Wing, Aaron Eklund, and Loren D Kellogg. Consumer-grade global positioning system (gps) accuracy and reliability. *Journal of Forestry*, 103(4):169–173, 2005.

[4] George H Kaplan. New Technology for Celestial Navigation. *Proceedings: Nautical Almanac Office*, 1999.

[5] Alan Grant, Paul Williams, Nick Ward, and Sally Basker. GPS Jamming and the Impact on Maritime Navigation. *Journal of Navigation*, 62(02):173, March 2009.

[6] Jon S Warner and Roger G Johnston. GPS spoofing countermeasures. *Homeland Security Journal*, 2003.

[7] EL Afraimovich, VV Demyanov, NS Gavrilyuk, AB Ishin, and G Ya Smolkov. Malfunction of satellite navigation systems GPS and GLONASS caused by powerful radio emission of the Sun during solar flares on December 6 and 13, 2006, and October 28, 2003. *Cosmic Research*, 47(2):126–137, April 2009.

[8] Unidroit Secretariat. An instrument on third party liability for Global Navigation Satellite System (GNSS) services. Technical Report January, INTERNATIONAL INSTITUTE FOR THE UNIFICATION OF PRIVATE LAW, 2010.

[9] George H Kaplan. Determining the position and motion of a vessel from celestial observations. *Navigation*, pages 1–15, 1995.

[10] James Evans. *The history and practice of ancient astronomy*. Oxford University Press, 1998.

[11] Gerald J Toomer et al. Ptolemy's almagest. *New York: Springer-Verlag, 1984.*, 1, 1984.

[12] WFJ Mörzer Bruyns and Richard Dunn. *Sextants at Greenwich: A Catalogue of the Mariner's Quadrants, Mariner's Astrolabes, Cross-staffs, Backstaffs, Octants, Sextants, Quintants, Reflecting Circles and Artificial Horizons in the National Maritime Museum, Greenwich.* Oxford University Press, 2009.

[13] Nick Kanas. *Star maps: history, artistry, and cartography*. Springer, 2007.

[14] Carl Friedrich Gauss. *Theory of the motion of the heavenly bodies moving about the sun in conic sections: a translation of Carl Frdr. Gauss" Theoria motus": With an appendix. By Ch. H. Davis*. Little, Brown and Comp., 1857.

[15] Eva Germaine Rimington Taylor, Commodore K St B Collins, and Joseph Needham. *The haven-finding art: a history of navigation from Odysseus to Captain Cook*. Hollis and Carter for the Institute of Navigation, 1971.

[16] Jean H Meeus. *Astronomical Algorithms*. Willmann-Bell, Incorporated, 1991.

[17] Carl Ch Liebe. Star trackers for attitude determination. *Aerospace and Electronic Systems Magazine, IEEE*, (June), 1995.

[18] Fabio Cozman and Eric Krotkov. Robot localization using a computer vision sextant. In *Proceedings of 1995 IEEE International Conference on Robotics and Automation*, volume 1, pages 106–111. Ieee, 1995.

[19] Fabio Cozman, Eric Krotkov, and Carlos Guestrin. Outdoor visual position estimation for planetary rovers. *Autonomous Robots*, 27, 2000.

[20] Xiaolin Ning and Jiancheng Fang. A new autonomous celestial navigation method for the lunar rover. *Robotics and Autonomous Systems*, 57(1):48–54, January 2009.

[21] Courtney Peterson. How it works: The charged-coupled device or CCD. *Journal of young investigators.*, 2010.

[22] John Barnes. Autonomous navigation via measurement of natural optical signals. MS Thesis, 2011.

[23] Sohrab Mobasser and Carl Christian Liebe. Micro sun sensor for spacecraft attitude control. 2004. Jet Propulsion Laboratory, Califomia Institute of Technology.

[24] Marcello Buonocore, Michele Grassi, and Giancarlo Rufino. APS-based miniature sun sensor for earth observation nanosatellites. *Acta Astronautica*, pages 3–6, 2005.

[25] Carl Ch Liebe. Pattern recognition of star constellations for spacecraft applications. *Aerospace and Electronic Systems Magazine, IEEE*, (January), 1992.

[26] John Barnes and Kartik B. Ariyur. Miniaturizing the spherical sundial: A hemispherical sensor for orientation and positioning with respect to point sources of light. *2011 IEEE International Conference on Control Applications (CCA)*, pages 662–667, September 2011.

[27] P Kenneth Seidelmann. *Explanatory Supplement to the Astronomical Almanac: A Revision to the Explanatory Supplement to the Astronomical Ephemeris and the American Ephemeris and Nautical Almanac*. University Science Books, 2005.

[28] James Schombert. Earth Coordinate System. University of Oregon Department of Physics. Retrieved 19 March 2011.

[29] Stefan Maus, Susan Macmillan, Susan McLean, B Hamilton, A Thomson, M Nair, and C Rollins. The us/uk world magnetic model for 2010-2015. 2010.

[30] Frank C Bellrose. Celestial orientation by wild mallards. *Bird-banding*, 29(2):75–90, 1958.

[31] Michael Vanvaerenbergh and Peter Ifland. *Line of Position Navigation: Sumner and Saint-Hilaire, the Two Pillars of Modern Celestial Navigation*. Unlimited Publishing LLC, 2003.

[32] US Naval Observatory. History of The Nautical Almanac, 2008. Available at `http://aa.usno.navy.mil/publications/docs/na_history.php`.

[33] P Kenneth Seidelmann and John H Seago. Time scales, their users, and leap seconds. *Metrologia*, 48(4):S186, 2011.

[34] Lawrence H Auer and E Myles Standish. Astronomical refraction: computational method for all zenith angles. *The Astronomical Journal*, 2000.

[35] James G Williams. Contributions to the earth's obliquity rate, precession, and nutation. *The Astronomical Journal*, 108:711–724, 1994.

[36] Robert Hurt. What Will the Constellations Look Like in 50,000 Years?, 2012. Available at `http://news.discovery.com/space/history-of-space/changing-constellations-50000-years.htm`.

[37] IEN Eurocontrol. WGS 84 Implementation Manual. Technical report, 1998.

[38] Eugene Hecht. *Optics (4th Edition)*. Addison-Wesley, 2001.

[39] Heung Cho Ko, Mark P Stoykovich, Jizhou Song, Viktor Malyarchuk, Won Mook Choi, Chang-Jae Yu, Joseph B Geddes, Jianliang Xiao, Shuodao Wang, Yonggang Huang, and John a Rogers. A hemispherical electronic eye camera based on compressible silicon optoelectronics. *Nature*, 454(7205):748–53, August 2008.

[40] Cheng Liu and Kartik B. Ariyur. Construction of a Sun Sensor for Geolocation and Orientation. In *Proceedings of the 25th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2012)*, pages 3627–3635, Nashville, TN.

[41] Cheng Liu, Sunghun Jung, and Kartik B Ariyur. Absolute Orientation for a UAV using Celestial Objects. In *AIAA Infotech@Aerospace 2013 Conference*, Boston, 2013.

[42] C. Radhakrishna Rao, Helge Toutenburg, Shalabh, and Christian Heumann. *Linear Models and Generalizations: Least Squares and Alternatives*. Springer, 2007.

[43] James Diebel. Representing attitude: Euler angles, unit quaternions, and rotation vectors. *Matrix*, 2006.

[44] Ibrahim Reda and Afshin Andreas. Solar position algorithm for solar radiation applications. *Solar energy*, (January), 2004.

[45] Ernest McCollough. Photographic topography. *Industry: A Monthly Magazine Devoted to Science, Engineering and Mechanic Arts*, (54):65, 1893.

[46] Michael F Land. Visual acuity in insects. *Annual review of entomology*, 42(1):147–177, 1997.

[47] Hossein Afshari, Laurent Jacques, Luigi Bagnato, Alexandre Schmid, Pierre Vandergheynst, and Yusuf Leblebici. Hardware implementation of an omnidirectional camerawith real-time 3d imaging capability. In *3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 2011*, pages 1–4. IEEE, 2011.

[48] Juyang Weng, Paul Cohen, and Marc Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(10):965–980, 1992.

[49] Janne Heikkila and Olli Silven. A four-step camera calibration procedure with implicit image correction. In *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, pages 1106–1112. IEEE, 1997.

[50] Duane C Brown. Decentering Distortion of Lenses. *Photogrammetric Engineering*, 32(3):444–462, 1966.

[51] Jianhua Wang, Fanhuai Shi, Jing Zhang, and Yuncai Liu. A new calibration model of camera lens distortion. *Pattern Recognition*, 41(2):607–615, February 2008.

[52] Jean-Yves Bouguet. Camera calibration toolbox for matlab. 2004.

[53] Thomas J. Fellers and Michael W. Davidson. CCD Saturation and Blooming. Available at `http://learn.hamamatsu.com/articles/ccdsatandblooming.html`.

[54] Paul VC Hough. Method and means for recognizing complex patterns, December 18 1962. US Patent 3,069,654.

[55] John Illingworth and Josef Kittler. A survey of the Hough transform. *Computer vision, graphics, and image processing*, 44(1):87–116, 1988.

[56] Richard O. Duda and Peter E. Hart. Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15, January 1972.

[57] Raymond K.K. Yip, Peter K.S. Tam, and Dennis N.K. Leung. Modification of hough transform for circles and ellipses detection using a 2-dimensional array. *Pattern Recognition*, 25(9):1007–1022, September 1992.

[58] Apply World Magnetic Model in Matlab. Available at `http://www.mathworks.com/help/aerotbx/ug/wrldmagm.html`.

[59] Vincent Roy. sun_position.m. URL: `http://www.mathworks.com/matlabcentral/fileexchange/4605-sunposition-m` [cited 13 July 2013].

[60] Darin Koblick. Lunar Azimuth and Altitude Estimation Algorithm. URL: `http://www.mathworks.com/matlabcentral/fileexchange/22992-lunar-azimuth-and-altitude-estimation-algorithm` [cited 13 July 2013].