3-19-2020

Maximizing Accuracy through Stereo Vision Camera Positioning for Automated Aerial Refueling

Kirill A. Sarantsev

MAXIMIZING ACCURACY THROUGH STEREO VISION CAMERA POSITIONING FOR AUTOMATED AERIAL REFUELING

THESIS

Kirill A Sarantsev, Captain, USAF

AFIT-ENG-MS-20-M-059

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENG-MS-20-M-059

MAXIMIZING ACCURACY THROUGH STEREO VISION CAMERA POSITIONING FOR AUTOMATED AERIAL REFUELING

THESIS

Presented to the Faculty

Department of Electrical and Computer Engineering

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Computer Science

Kirill A Sarantsev, B.S.C.S.

Captain, USAF

March 19, 2020

AFIT-ENG-MS-20-M-059

MAXIMIZING ACCURACY THROUGH STEREO VISION CAMERA POSITIONING FOR AUTOMATED AERIAL REFUELING

THESIS

Kirill A Sarantsev, B.S.C.S.
Captain, USAF

Committee Membership:

Clark N. Taylor, Ph.D
Chair

Douglas D. Hodson, Ph.D
Member

Scott L. Nykl, Ph.D
Member

Abstract

Aerial refueling is a key component of the U.S. Air Force strategic arsenal. When two aircraft interact in an aerial refueling operation, the accuracy of relative navigation estimates are critical for the safety, accuracy and success of the mission. Automated Aerial Refueling (AAR) looks to improve the refueling process by creating a more effective system and allowing for Unmanned Aerial Vehicle(s) (UAV) support. This paper considers a cooperative aerial refueling scenario where stereo cameras are used on the tanker to direct a "boom" (a large, long structure through which the fuel will flow) into a port on the receiver aircraft. The analysis focuses on the effects of camera positioning with the rear-facing stereo vision system. In particular, the research seeks the optimal system design for the camera system to achieve the most accurate navigational estimates. The testing process consists of utilizing a simulation engine and recreating real world flights based on previously collected Global Positioning System (GPS) data. Using the pose estimation results and the ground truth information, the system computes the error between the incoming aircraft's position in the virtual world and its calculated location based on the stereo matching algorithm. The testing process includes both un-obscured scenarios and cases where the boom causes significant occlusions in the camera images. The results define the improvements in position and orientation estimation of camera positioning from the consolidated simulation data. Conclusions drawn from this research will propose and help provide recommendations for future Air Force acquisition and development of aerial refueling systems.

Table of Contents

Page

List of Figures

List of Tables

MAXIMIZING ACCURACY THROUGH STEREO VISION CAMERA POSITIONING FOR AUTOMATED AERIAL REFUELING

I. Introduction

Aerial refueling expands the longevity and reach of aircraft and serves the Air Force goal of air superiority and global reach [2, 3]. The process requires precision and high-frequency updates in order to provide the aircrews with centimeter level accuracy for two objects moving at great velocities. There are two primary methods currently utilized by the United States military: the Air Force's boom method and the Navy's probe-and-drogue method as seen in Figure 1 [4]. With the boom method, the receiver must move to a position close enough to the tanker. Once the receiver is in position, the tanker aircraft has a "boom" that is moved to precisely the correct location to enable the refueling hose to touch the receptacle on the receiver aircraft [2]. Currently, this method involves an Aerial Refueling Officer (ARO), or boom operator, manually controlling the device [5]. Instead of utilizing a boom, the probe-and-drogue method uses a flexible hose with a drogue on the end which acts as a funnel for the receiver to insert its probe into the hose to begin refueling [6]. Probe-and-drogue, enables multiple aircraft the ability to attach simultaneously, however

Figure 1: Refuelling methods: (Left) probe-and-drogue and (Right) boom [1]

substantially more effort is required from the pilot of the receiver, who must precisely control the speed and altitude of the aircraft [1]. Other advantages of the boom over prove-and-drogue include the fuel transfer speeds and adaptability of the hardware, boom tankers can be converted in the field to accommodate probe-quipped aircraft, this cannot be done the other way around [2]. Therefore, this research focuses on the boom refueling method.

1.1 Problem

There are two key problems related to current aerial refueling operations; crew safety and Unmanned Aerial Vehicle(s) (UAV) support. In the past decade, UAV have become critical to U.S. military operations [7], the current aerial refueling process does not work with them due to communication latency between the ground controller operator and the UAV [8, 9, 10]. Therefore, Automated Aerial Refueling (AAR) has been proposed as the way forward in order to overcome the latency issue. The goal of the system is to create an AAR system, capable of instant adjustments during the receiver aircraft approach without requiring input from a human operator. AAR also looks to improve the safety of the process in order to avoid accidents such as KC-10 refueling boom mishap [11] and the KC-130 crash over the pacific [12] further referenced in the chapter II.

Previous work on AAR has used Differential GPS (DGPS) to provide the precise relative positioning required [1, 13]. While DGPS can provide the centimeter level accuracy needed for an AAR approach, DGPS relies on external signals and is vulnerable to environmental conditions, blockage, and jamming [14]. Therefore, the interest is in finding alternatives which do not rely on DGPS capabilities [15]. Prior work has demonstrated that stereo imagery can be used to achieve relative pose estimates in real time [14]. Other work conducted at the Air Force Institute of Technology

(AFIT), has shown that using a stereo camera system on the tanker aircraft allows for high accuracy estimates of the receiver aircraft position [9].

The final system envisioned uses tanker cameras to track the incoming receiver aircraft, determining its current position and orientation and providing real-time updates to the tanker and control commands to the receiver and boom. With this system, the boom provides another layer of challenge due to its occlusion of the receiver aircraft when viewed by the cameras. The problem then consists of relaying the position of the receiver with the best precision possible, even when it is occluded, in a system which can execute the refueling process without additional input from the operator.

1.2 Research Objectives

This research builds on the prior AFIT work described in [16, 14, 17, 9] and aims to find improvements to the stereo vision system in order to achieve the best accuracy possible in both full visibility and occluded scenarios. Specifically, analyzing if increasing the camera separation baseline improves the system's accuracy.

This thesis addresses two issues; first, based on prior work, what limits does the current camera baseline pose on the overall performance of the system and what results could be achieved if the system was not limited to a single type of stereo camera configuration. Without considering the effect that a particular system choice has on the overall accuracy of the system, there cannot be a consensus that the capabilities of the design are fully realized. The second issue that this thesis addresses is analyzing the impact of occlusions between the tanker and receiver aircraft. The objective is to determine how different camera system designs reduce or increase the error introduced by occlusion.

To test the system and analyze how to mitigate these issues, the research will

utilize a prior Iterative Closest Point (ICP) algorithm [16] for determining relative navigation within a simulation engine previously developed for AAR at AFIT [17]. Modifying this environment and adjusting baseline separation in order to gather data through simulations and find the most accurate relative navigation estimates, both with and without a boom present in the imagery. The simulation is based off of realistic flight motions from prior recorded flights. Recreating the realistic flight motions provides an accurate approach and expected aircraft behavior within the testing process and allows for data aggregation across a myriad of different scenarios.

Analyzing the data will help answer three questions:

1. What is the impact of camera separation on the accuracy of the relative position estimation system?

2. How large of an impact does occlusion make in the error of the pose estimation?

3. Is there an optimal camera separation value which overcomes the increased error due to occlusion?

1.3 Contribution

The contribution of this research is to provide data that supports changes to the current system in order to improve its accuracy and therefore is feasibility in becoming the first AAR system utilized by the United Sates Air Force (USAF).

1.4 Document Overview

The remainder of this thesis consists of four chapters: chapter II, chapter III, chapter IV, and chapter V. Chapter II defines the background and key concepts integral to the research. Chapter III describes the details of the simulation environment and the specific scenarios tested to determine the optimal system configurations for

the stereo vision system. Chapter IV presents the data accumulated across a large variety of operating scenarios. Chapter V concludes the paper and provides a way ahead for future work.

II. Background and Literature Review

This chapter outlines the background information necessary in order to understand the thesis research objectives and analysis of camera separation effects on relative navigation performance. Section 2.1 describes aerial refueling history, terms, methods, evolution, current limitations, and Automated Aerial Refueling (AAR) research. Section 2.2 describes the key concepts needed to understand how a stereo vision system functions. Finally, section 2.3 defines the background components related to the simulations engine utilized in this research.

2.1 Aerial Refueling

2.1.1 History

Aerial refueling dates back to 1921, starting with Russian–American Alexander de Seversky and following demonstrations, experiments, improvements through the 1920s, 30s, and 40s [1]. The initial process began with an operator physically climbing from one aircraft to another with a can of fuel and refueling the aircraft [18]. Over time the systems evolved to the more advanced processes seen today. The first routinely used air to air refueling system utilized a looped hose, a modification to the original rubber hose method, pictured in Figure 1, which attached additional connectors and fittings to allow for a much faster hookup process [1]. Table 1, compiled from [1, 18, 19], outlines the history of aerial refueling and the events which proved its viability in humanitarian and warfare operations.

The United States military has adapted aerial refueling in order to expand its global reach and power [19]. Air refueling came about from the need to expand the range and maintain the safety of military aircraft; by refuelling in the air aircraft can takeoff from airfields outside of the conflict areas and strike deep inside enemy

territory [19]. From its inception aerial refueling proved its significance to the U.S. military, playing an especially significant role in modern conflicts such as the Vietnam War and Operation Desert Storm [19]. Today's primary approaches are the boom and probe-and-drogue methods, evolutions of the original idea of the looped hose method of attaching two aircraft [20].

2.1.2 Boom Refueling Method

Developed by Boeing in the 1950s, the flying boom is utilized to refuel all current Air Force fixed-wing aircraft. The boom is a rigid tube operated by the boom operator, also known as the Aerial Refueling Officer (ARO), in order to insert the line into the receptacle of the receiver aircraft. Boom-equipped tankers all have a single boom and can only refuel a single aircraft at a time [2]. Boom equipped tankers were originally created in order to refuel long-range bombers which required large amounts of fuel due to their size and payload's weight. One of the advantages of the boom method is the amount of fuel it is able to transfer, 6,000 lbs per minute, versus the probe-and-drogue's 2,000 lbs per minute.

Because the boom is able to transfer fuel at a much faster rate, it is extremely effective in refueling bombers and large aircraft. However, most modern fighter aircraft cannot match the throughput, being able to handle, at maximum, 3,000 lbs per minute of fuel, therefore some of the efficiency of the boom method is lost [2]. Even with this limitation, the boom is the most versatile system. Boom equipped tankers are able to reduce fuel dispense speeds when refueling fighters, while tankers utilizing the probe-and-drogue method cannot increase fuel speeds to accommodate large aircraft [2].

The boom refueling method has been utilized by the United Sates Air Force (USAF) with the KB-29 Superfortress, KC-97 Stratofreighter, KC-135 Stratotanker,

Table 1: Aerial Refueling History

Year	Event
1917	Alexander de Seversky proposes aerial refueling as a means of increasing combat aircraft range
1921	Alexander de Seversky files a patent for air-air refueling
1921	An airborne Navy lieutenant used a hook to snatch a can of gas from a barge
1921	Weslay May climbed from the wing of a Lincoln Standard to a Curtiss JN-4 with a can of gas
1923	U.S. Army begins experimenting with refueling by lowering a hose from one aircraft to another
1923	Using the hose technique to refuel, a record of a 37 hour endurance flight is set
1929	The Questing Mark flight is conducted with an aircraft staying airborne over Los Angeles for 150 hours and 40 minutes using a total of 5,660 gallons of gasoline
1934	Royal Air Force Lieutenant Richard Atcherly developed and patented a looped-hose aerial refueling system. This system would be later refined by adding a drogue to the hose
1948	As the head of Strategic Air Command (SAC), General Curtis LeMay made aerial refueling a major goal, recognizing its need with new jet-powered bombers
1949	Flight Refueling Limited tested its newly developed probe-and-drogue system with the U.S. Air Force expressing interest in the technology
1950	Boeing developed the flying boom system for the B-29s with the Air Force ordering 100 units and designating the aircraft as the KB-29P
1953	SAC commands nearly 30 aerial refueling squadrons flying the KC-97 with an improved boom system
1957	SAC begins procuring Boeing KC-135 Stratotankers as the main refueling aircraft
1964	KC-135s are utilized in trans-Atlantic fighter deployments maturing the idea of rapid mobility and global reach
1972	Aerial refueling enables the F-105 to strike targets in North Vietnam and return to secure bases away from the conflict areas
1973	Yom Kippur War kicked off Operation Nickel Grass, with C-141s and C-5s flying 22,318 tons of cargo. This conflict demonstrated to the Pentagon the need for aerial refueling for transport aircraft
1974	SAC begins aerial refueling training and qualification for C-5 crews
1990	During Operation Desert Storm tankers refueled an average of 125 airlift missions per day transferring over 725 million pounds of fuel to combat sortie aircraft
1992	Air Mobility Command is established with a focus on airlift and aerial refueling
2001	KC-135 and KC-10 Extenders supported combat air patrols after the September 11 attacks
2011	Boeing KC-46 Pegasus selected as the replacement for the Boeing KC-135 Stratotankers

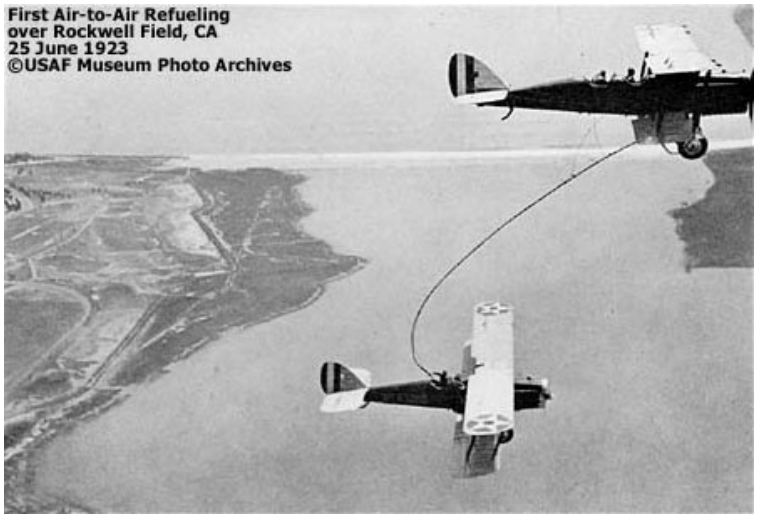


Figure 2: June 25, 1923: The first air-to-air refueling [21]

and the KC-10 Extender tankers [22].

The most recently constructed tanker is the Boeing KC-46 Pegasus, Figure 3, which is the next generation, multi-role, air refueling tanker aircraft for the USAF [24]. The KC-46 utilizes the boom refueling approach and is compatible with the A-10, B-52, C-17, KC-10, KC-135, KC-46, F-15E, F-16 and F/A-18 aircraft.

2.1.3 Modern Limitations and Issues

While there have been many advances and positives from utilizing aerial refueling, there are still many issues with the process and accidents which endanger the crews and aircraft still occur. One such incident occurred on November 1, 2016 in which an aerial refueling boom detached from a KC-10 extender tanker during a training flight [11]. When the boom operator lowered the boom, the boom began to move erratically and the operator lost control of the device with the boom eventually fully detaching from the fuselage. No lives where lost but this incident cost the government $6.52 million [11]. The investigation into the accident later revealed that the primary cause of the accident was structural damage to the boom rotary crank, however had the boom operator turned off the boom flight control switch before the damage became

Figure 3: Aerial refueling connection: KC-46 boom to fighter receptacle [23]

critical, the boom would not have detached from the aircraft. The advantage of AAR in this case would have been to remove the operator error and activate the necessary safety measures as soon as an error is sensed.

Another major refueling accident occurred on December 6, 2018, when a F/A-18 fighter jet collided with the KC-130 refueling aircraft during a refueling training exercise over the ocean near western Japan [12]. This incident lead to loss of life and the investigation found a lack of training and experience as the root cause of the crash, with no detected equipment issues in either aircraft [25]. A more robust system, such as the proposed AAR stereo vision system, can support these missions and help stop unnecessary loss of life.

2.1.4 UAVs and Aerial Refueling

Unmanned Aerial Vehicle(s) (UAV) are aircraft with no physical pilot on board. UAV are either remotely controlled – Remotely Piloted Aircraft (RPA) – by operators at a ground base, often hundreds of miles away, or fly autonomously based on pre-programmed flight plans [26]. UAV are currently used for a number of missions, including reconnaissance and attack roles. UAV use and procurement has been grow-

ing at a fast rate over the years with the FY 2019 Depeartment of Defense (DoD) budget consisting of approximately $9.39 billion in drone-related procurement, research and development, and construction funding, a 26 percent increase over the FY 2018 request of $7.5 billion [27]. The majority of USAF procurement focuses on the MQ-9 Reaper and the RQ-4 Global Hawk [27].

One of the biggest limitations of UAV is the inability to utilize aerial refueling [28, 10]. The issue lies with the communication latency between the ground controller and the UAV, does not allow for the quick response movement necessary for aerial refiling [3]. The goal is then to utilize AAR in order to overcome the latency issue by communicating with the receiver directly (bypassing the remote operator) to allow for boom connection and movement.

2.1.5 Automated Aerial Refueling

AAR consists of the same principles as aerial refueling, however the process attempts to take the next step and automate the execution. The key improvements are the decrease of manpower required to operate the boom, less error due to the judgment of controllers, and an improvement in the boom attachment speed during the docking process.

With the USAF refueling boom system, the vision is to extend this system to refuel UAV. In order to accomplish this task, the system must be able to accurately measure the location and orientation of the incoming UAV in order to track when the UAV is in range and where to move the boom. The KC-46 already has the cameras and equipment necessary to utilize stereo vision as show in Figure 4.

Figure 4: KC46 Camera and Boom Placement [29]

2.1.6 AAR Research

There has been a lot of research into the AAR from academic institutions and governments. Many systems such as [30, 31] focus on the probe-and-drogue system, while the research in this thesis focuses on the boom method due to its advantages and relevance to the Air Force. Research into probe-and-drogue differs because, unlike the boom, no refueling arm needs to be moved to make contact with the receiver, instead most of the process lies with the receiver pilot correctly positioning and moving the receiver to make contact with the hose. For boom refueling, proposed solutions include Global Positioning System (GPS) and machine vision systems [1, 32].

As mentioned earlier, Differential GPS (DGPS) technologies are suitable for this task, however a limitation of this approach is the possible DGPS signal distortion by the tanker air frame [10]. [33, 28] discuss combinations of machine vision systems and GPS. Further the systems discussed and depend on markers on the tanker with cameras attached to the roof of the receiver aircraft. Limitations of this research include the reliance on markers and the need to add them to military aircraft as well

as the positioning of UAV camera systems. [32] introduces a machine vision system which can also function with AAR, however this approach requires modification to the receiver aircraft to have painted target images near the fuel receptacle. The amount of modifications required and approval necessary to apply such markers to all military aircraft diminishes the feasibility of this approach.

A pure vision based approach which does not rely on external systems or markers is then ideal for an Air Force solution to AAR. Air Force Institute of Technology (AFIT) researches have been working on this problem for the past several years. The research forms the building blocks for the testing and extermination conducted in this thesis. Specifically, [17] outlines the virtual world computer vision pipeline process of an aerial refueling scenario and how to recreate a physically realistic aerial refueling approach to utilize in a simulation environment. The paper provides a novel and cost effective way to simulate the system and includes the methodology for the simulation process. The research demonstrates how virtual imagery can be used to generate point clouds for pose estimation in real time, with the results providing positive support for the viability of a stereo vision system for AAR. The outline disparity map generation, point cloud reprojection, and point-to-point Iterative Closest Point (ICP) implementation serve as the basis for the methodology behind the simulations outlined in this thesis. This research works to build upon the system outlined by including additional analysis into the stereo vision system configuration as well as further study into how occlusion impacts the accuracy of the system. By incorporating the new recorded flight data and aggregating the error between the flights, this research is able to expand further on the viability and negative impacts of occlusion the the stereo vision system. They key aspect of filtering out points generated on the boom is also incorporated in the simulations discussed in this research.

2.2 Stereo Vision

In order to understand the approach used in this research, the basic concepts of stereo vision are defined in this section. Stereo Vision consists of utilizing multiple cameras in order to find the depth information lost when converting to a 2D image from the 3D world. Section 2.2.1, section 2.2.2, section 2.2.3, section 2.2.4, and section 2.2.5 outline the concepts vital to the stereo vision process.

2.2.1 Pinhole Camera Model

The pinhole camera model is the basic arrangement for a camera. This is the type of camera the thesis will use throughout its experimentation. F_c corresponds to the pinhole camera's focal center and lies in the center of the camera. The center plane represents the 2D image plane with *(u,v)* corresponding to point P in the 3D world. As seen from the conversion the Z depth information is lost.

2.2.2 Camera Calibration

Camera calibration consists of finding the intrinsic and extrinsic camera parameters in order to correctly align the points between the two camera images. The intrinsic parameters include the focal length, the optical center, also known as the principal point, and the skew coefficient. The camera intrinsic matrix, K, is defined as:

$$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

where f_x and f_y are the camera focal lengths and and c_x and c_y are the optical centers. For distortion, OpenCV takes into account the radial, Figure 6, and tangential,

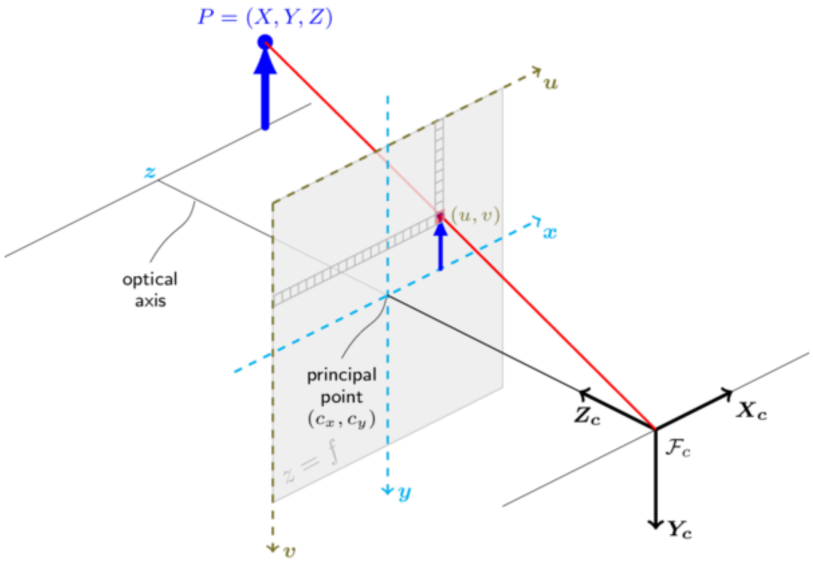


Figure 5: Pinhole Camera Model [34]

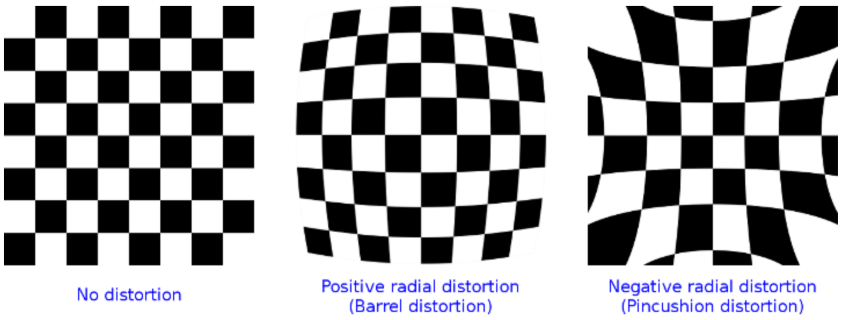


Figure 6: Radial Distortion [34]

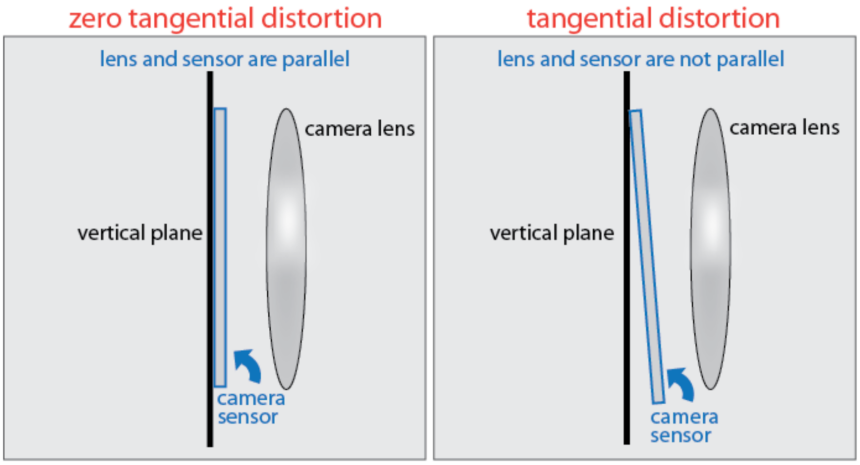


Figure 7: Tangential Distortion [35]

Figure 7, factors [36] using the formulas:

$$x_{corrected} = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)$$

$$y_{corrected} = y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)$$

for radial distortion and

$$x_{corrected} = x + [2p_1 xy + p_2(r^2 + 2x^2)]$$

$$y_{corrected} = y + [p_1(r^2 + 2y^2) + 2p_2 xy]$$

for tangential distortion. OpenCV represents the five distortion parameters as a one row, five column matrix

$$Distortion_{coefficients} = (k_1\ k_2\ p_1\ p_2\ k_3)$$

The rotation and translation matrix R and t are the extrinsic parameters [34]. These matrices translate coordinates of a point, X, Y, Z, to a coordinate system, fixed with respect to the camera, x,y,z [36].

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + t$$

2.2.3 Epipolar Geometry

With the conversion from 3D to 2D when taking an image, the depth information is lost. The question then becomes how can we tell the distance an object was away from the camera in the image. Stereo vision works to solve this issue by utilizing

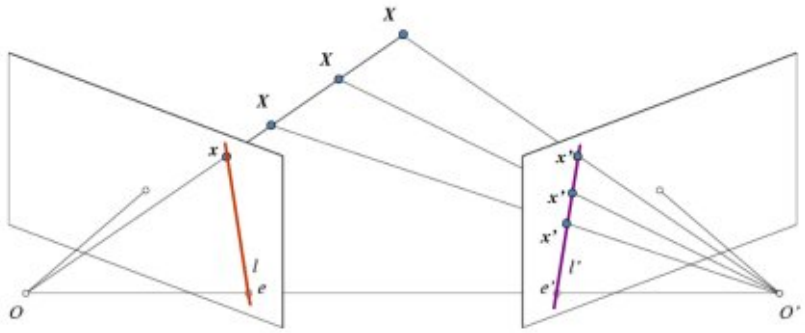


Figure 8: Epipolar Geometry: Left and Right Cameras [37]

two cameras by aligning up the object in both images, the distance can then be determined, this is known as epipolar geometry, Figure 8. *oxo'* is called the epipolar plane. When only using the left camera or image, x can appear anywhere on the line ox [37]. However, when the right camera is included, the ability to triangulate the point is gained. The projection from the right camera onto ox creates the epiline l'. Then x exists somewhere on l', this is called the epipolar constraint. This information is then used to find the position of x with the lost depth information.

2.2.4 Disparity Map

A disparity map or depth map shows the pixel difference between two images. It contains the depth information of every pixel stored in it. The disparity map returns the Z value information lost when converting from 3D to 2D as seen in Figure 5. Pixels with smaller disparities are farther from the camera while pixels with larger disparities are closer.

2.2.5 Occlusion

One issue which can interfere with the stereo vision algorithm is boom occlusion. Occlusion results from a lack of visibility of the object being tracked from the view of the cameras. Specifically in this research scenario, the boom is an operator controlled

Figure 9: Grayscale Left and Right Images

Figure 10: Generated Disparity Map

arm which attaches to the receivers fuel receptacle and the current placement of the cameras, 5.5 meters behind the boom, means that the boom occludes the receiver as it is in motion. The amount of occlusion will depend on the flight path of the receiver, however in the worst cases the boom can occlude a large portion of the receiver for the entire duration of the refueling approach. This occlusion cause issues in the tracking algorithm calculations and ability to match features creating a larger error in the pose estimation [17]. Figure 11 illustrates boom occlusion as seen from each camera.

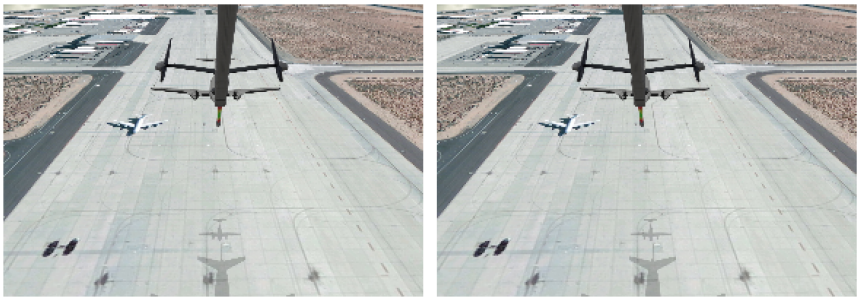


Figure 11: Boom occlusion of receiver aircraft as seen from the left and right cameras

2.3 Virtual Simulation

Virtual simulations are a cost effective and efficient way to generate data while maintaining the same testing conditions for each scenario [1]. The simulation provides a large amount of control over the many variables which can impact system performance reducing the risk of getting false results due to an unknown factor. Also, due to the scheduling time for physical aircraft and the amount of test runs necessary for this research means that virtual simulations are the best course of action. Testing was conducted using the AFTR Burner engine to recreate aerial refueling approaches. The independent variable was a changing camera baseline separation.

2.3.1 AFTR Burner Engine

This thesis utilizes the AFTR Burner engine to create the environment and research simulations for camera separation. The AFTR Burner engine is an OpenGL based renderer created at AFIT Autonomy and Navigation Technology (ANT) center in order to create realistic synthetic imagery representative of the real world [38]. The engine allows for the replay of real flight data in the virtual world and the ability to compare the truth data to the estimated position. The realistic motion and lighting provides a good test bed for new techniques and algorithms.

2.3.2 3D Models

3D models are the objects used in the virtual simulation consisting of a polygon mesh. The mesh defines the shape of the object defines the shape of the object through a collection of vertices, edges and faces. The models are representative of their real world counterparts in relative scale and appearance [9]. The simulation utilizes a KC-46 textured model, Figure 12, for the tanker and a C-12 textured model, Figure 13, for the receiver.

2.3.3 Flight Replay Data

The flight test data files contain the timing, orientation, position, and imagery information gathered at Edwards Air Force Base in 2019. The test flights consisted of two C-12s aircraft, one acting as the tanker and the other as the receiver. The aircraft execute several different approaches representative of typical refueling scenarios. This data can be replayed in the virtual world in use with the 3D models to recreate the flight and motion of the aircraft.

Figure 12: KC-46 Tanker 3D Model

Figure 13: C-12 Receiver 3D Model

2.3.4 OpenCV

Open Source Computer Vision Library (OpenCV) is an open source computer vision software library utilized for its computer vision algorithms within the simulation code. It functions with C++ and the AFTR Burner engine. The library provides resources for camera calibration, and creation of disparity maps from input images as well as reprojection of the point clouds to 3D coordinates [34]. This research specifically utilizes the Camera Calibration and 3D Reconstruction functions in order to take the images from the simulation and generate a point cloud for the estimated pose of the receiver.

2.3.5 Point Cloud

Point clouds are generated from the disparity map information. The pixels are translated into the 3D world as points forming a larger cloud. These points align with the visible features matched in the two images. In the simulation environment the points should cluster around the receiver to create a cloud which can be matched to the reference point cloud of the C-12.

(a) Generated point cloud (b) Red reference point cloud matched

Figure 14: ICP with occlusion

2.3.6 Iterative Closest Point

ICP is the algorithm which has been utilized with the stereo vision system in order to estimate the location of the receiver based on the generated point cloud. ICP functions by minimising distance between two point clouds, also known as point set registration, in order to assign correspondences between two sets of points and to recover the transformation that maps one point set to the other [39]. In the simulation, the ICP algorithm matches the red reference point cloud of the receiver to yellow sensed point cloud created from the disparity map, Figure 14. The system utilizes a shell model for the receiver aircraft model which allows for a reduction in time for the pose estimation phase of the stereo vision pipeline [17].

III. Methodology

This research utilizes a simulation environment to evaluate the possible utility of changing the camera separation baseline between the stereo cameras. This simulation environment is capable of simulating the approach of receiver aircraft to a tanker aircraft, the imagery that stereo cameras view as the aircraft approaches, and the occlusion of the boom, when necessary, in that imagery. The process used to compute pose, introduced in section 2.2 and section 2.3 is run in real-time within this simulation. This chapter describes in more detail the simulation environment and the types of simulations used to test the different baselines between the stereo cameras.

3.1 Components

The simulation components include the tanker, receiver, red reference point cloud, refueling boom, and two rear-facing cameras which function as the stereo vision system shown in Figure 15 and Figure 16. Within the simulation, the location of the different objects are computed as follows:

- The tanker and receiver positions and orientations are taken from flight test Global Positioning System (GPS) data recorded at Edwards Air Force Base. Because the KC-46 replaces the C-12 that was originally used in the real world flights, the tanker GPS location in the local body frame is changed to (-5.0633, 0.0, 1.6751) in order to maintain the same separation between the KC-46 and the C-12 as there was in the original run with the C-12 aircraft.
- The boom and stereo cameras are attached to the tanker and therefore follow the motions of the KC-46 while the simulation is in motion.
- In this research the boom is maintained in a set position angled 30 degrees down in the direction of the receiver as demonstrated in Figure 15.

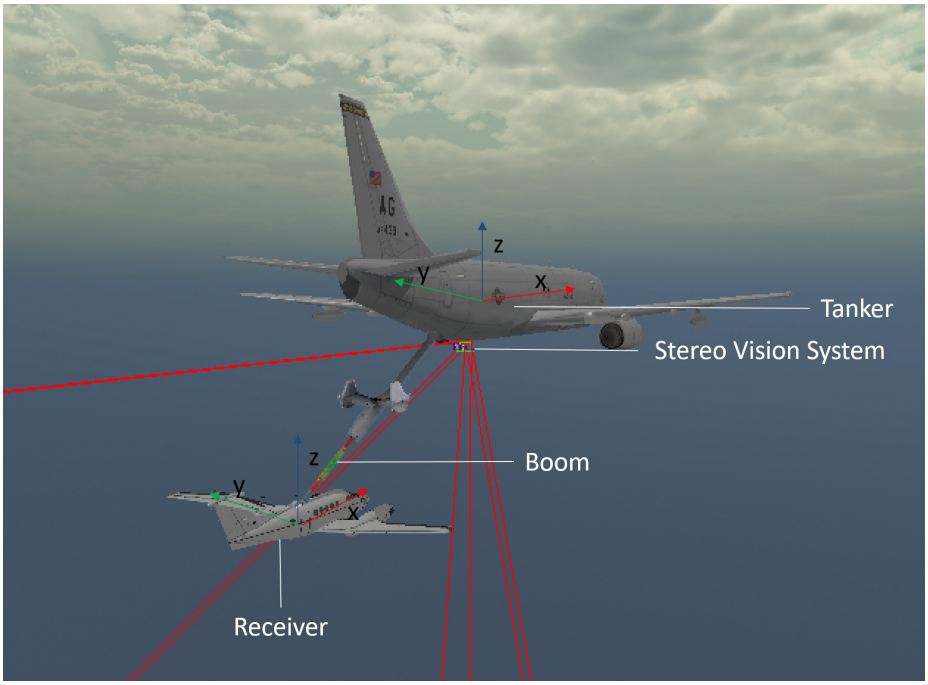


Figure 15: Simulation components

The stereo camera parameters are outlined in Table 2. The roll, pitch and yaw are in the tankers body frame, therefore the cameras face 25 degrees down and are rotated around 180 degrees to face in the opposite direction of the tanker's nose. The stereo vision system is placed at the anchor point, in the tanker space, with each camera being offset by half of the baseline separation in both directions. The boom's anchor point is then set 5.5 meters away from the stereo anchor point. This configuration is the starting position, however as the simulation runs, the baseline separation and camera distance from the boom are modified. The boom's visibility is disabled in simulations which test the systems performance without occlusion.

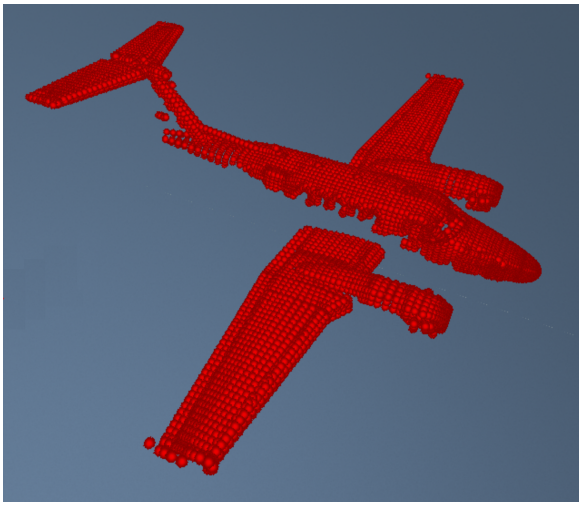


Figure 16: Red Reference Point Cloud

Table 2: Stereo Vision Camera Parameters

Image Width	1280
Image Height	960
Lens Focal Length	16.44 mm
Horizontal FOV	56
Baseline Separation	0.5 m
Roll	0
Pitch	-25
Yaw	180
Anchor Point X	-8.311 m
Anchor Point Y	0.0 m
Anchor Point Z	-7.031 m

3.2 Simulation Process

Each simulation follows the general process outlined in [17] and illustrated in Figure 17. The process is split into two parts; calibration and creating of the reprojection matrix (denoted by green in Figure 17), which is done once per flight replay, and the computer vision pipeline (shown in blue in Figure 17), which is repeated over the course of a flight. Each portion of Figure 17 will be described in more detail below.

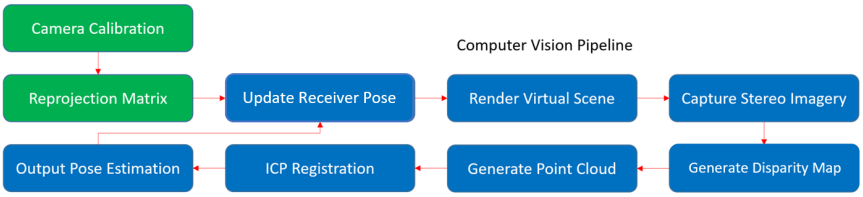


Figure 17: Overview of the flight simulation process

3.2.1 Camera Calibration

The process starts with calibration of the two pinhole cameras. In order to use OpenCV's *stereoRectify()* and get accurate results when running the simulation, the left and right camera intrinsic and extrinsic parameters were collected using the checkerboard method in the virtual world, the example images of the calibration process are shown in Figure 18. Because the same camera models are used for all tests, these parameters do not change and are read from a file. The remaining parts necessary to create the reprojection matrix are the relative rotation and translation going from the left camera frame to the right camera frame. For this simulation, the following parameters were used:

Intrinsic matrix for both cameras:

$$K = \begin{bmatrix} 1.204 \times 10^3 & 0 & 6.395 \times 10^2 \\ 0 & 1.204 \times 10^3 & 4.793 \times 10^2 \\ 0 & 0 & 1 \end{bmatrix}$$

Rotation matrix:

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

0.5 meter baseline separation OpenCV translation vector:

$$t = \begin{bmatrix} -500 \\ 0 \\ 0 \end{bmatrix}$$

3.2.2 Reprojection Matrix

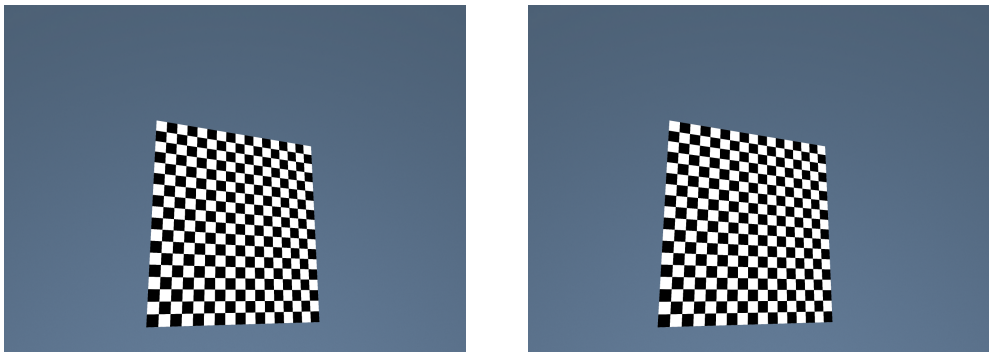
Once all of the calibration parameters are read, OpenCV's *stereoRectify()* is called to get the reprojection matrix, rectification transform (rotation matrix) for the first camera, rectification transform (rotation matrix) for the second camera, and the left and right projection matrices in the new (rectified) coordinate systems [34]. Then the undistortion and rectification transformation map for the left and right cameras is created using OpenCV's *initUndistortRectifyMap()*, the output of which will be used with OpenCV's *remap()* to create the rectified images from the left and right image captures.

3.2.3 Update Receiver Pose

With the calibration complete, the simulation begins replaying the flight data. For each frame, the system reads the orientation and position of the tanker and receiver based on the number of the current frame.

3.2.4 Render Virtual Scene

The scene is rendered with the call to AFTRBurner's *updateWorld()* and the cameras textures are updated with the new positions of the tanker, boom, and receiver. In order to get accurate results and keep the simulation running, the sensed yellow point clouds, red reference point clouds, and camera frustum are made invisible to the



(a) Left Image (b) Right Image

Figure 18: Checkerboard process utilizes different positions for the checkerboards during camera calibration

cameras before the images for the next frame are taken. In this way the performance of the simulation is visible to a viewer without interfering with the next system step in the process.

3.2.5 Capture Stereo Imagery and Generate Disparity Map

In this step, the system captures the left and right camera images and converts them to grayscale. Calling OpenCV's *remap()* with the two images and setting the number of disparities and the SAD window size, for the *StereoBM* class and calling *StereoBM - compute()* to generate the depth information needed to calculate the location of the receiver. Based on previous research [9] the following values are utilized for best results with the system:

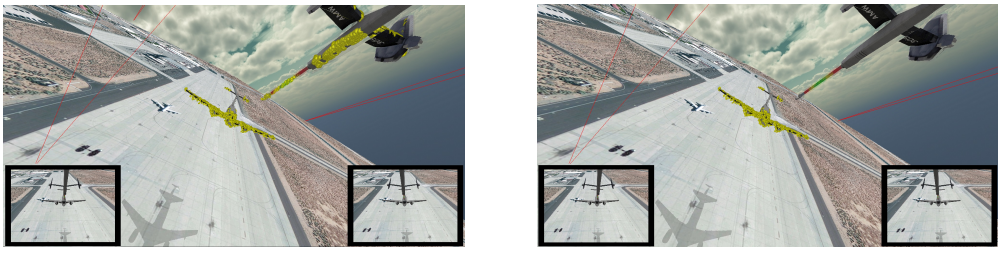
Table 3: Disparity and Filter Speckles Parameters

Number of Disparities	256
SAD Window Size	11
Disparity Value	0
Max Speckle Size	210
Maximum Difference	5

3.2.6 Generate Point Cloud

The points from the depth map are then reprojected using *reprojectImageTo3D()* to create a set of 3D coordinates which are used to generate the yellow sensed point cloud in the virtual world. In the simulation these points are represented by the yellow dots which form the yellow point cloud and represent the estimated location of the receiver.

When simulating flights with boom occlusion, an intermediate step must be added between "generate point cloud" and "ICP registration" to filter out the points created from the disparity map that are part of the boom. Without this step, Iterative Closest Point (ICP) will generate large errors in the estimated pose of the receiver aircraft. Figure 19 illustrates the difference between the simulations when performing filtering, note the absence of yellow dots on the boom after the filtering process. To filter out the boom points, each point is tested by creating a ray between the point and the left camera, if the ray intersects any part of the boom then the point is thrown out. Because this process requires additional computation, it slows down the overall performance of the system. In order to decrease this performance hit, only points which are within the distance from the tip of the boom to the tanker are tested. With this filtering method and ICP, the system is able to match the red reference point cloud to the yellow point cloud generated.



(a) Unfiltered (b) Filtered

Figure 19: Occlusion filtering

3.2.7 ICP Registration

Using ICP, the reference point cloud model of the receiver is matched to the generated point cloud. The algorithm functions by iteratively converging to the minimum mean squared distance between two point clouds [40, 3].

3.2.8 Output Pose Estimation

Once the clouds are matched (registered) together, the final pose is output. The output pose includes x, y, and z axis values and roll, pitch, and yaw angles for the receiver position and orientation.

For each ICP iteration, the system generates the error data using the difference between the computed position and the true receiver model position. The X,Y, and Z error is saved individually with the distance from the camera of the receiver in order to understand which part of the flight this error corresponds to as well as understand the correlation between distance and error. The same process is accomplished for the orientation error with the individual roll, pitch and yaw error saved for future processing. We also record the total number of points generated for the yellow point cloud.

After a full simulation is complete for a camera baseline, the individual raw data is saved and graphs are generated for the run, plotting all position and orientation errors. To evaluate the accuracy of any given simulation setup, the position and orientation errors are averaged using Root Mean Square Error (RMSE)

$$RMSE = \sqrt[2]{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n}}$$

P_i represents the predicted value, the estimated pose of the receiver. O_i is the observed value, the truth data for the receiver pose. The total number of poses for

the flight is represented by n.

This average is computed independently on x, y, z, yaw, pitch, and roll, however the total combined error for position and orientation is generated as well. The standard deviation of the error is also computed for each pose element. All of this provides as much data as possible for future analysis. The pose and number of points in the point cloud are stored for each run of the data. For these tests, there are 9 runs of recorded data representing 9 different approaches of a receiver aircraft to the tanker.

3.3 Simulation Parameters

In order to capture all of the data necessary and answer the research questions for this thesis, several different representative flights were utilized. There were two phases to this process. The first phase consisted of data set selection. This focused on baselining the system and finding starting and stopping points in the flight data files. The flights needed to represent a realistic approach and provide data within the ranges of 30-75 meters. Due to the length of the flights and the amount of simulations needed to fully explore the domain, cutting down the time by removing unnecessary data gathering points is key. Therefore, the data was analyzed and the flight start and stop points chosen to covers a moderate distance and provides a good snapshot of a refueling maneuver. All of the initial experimentation and data gathering focused on narrowing down these points. The final playback pulses are shown in Table 4.

The next phase consisted of using the flight data to run the simulations. The tests consisted of continuous runs through the same flight with modified camera parameters. These tests move the cameras to the necessary locations and set the needed calibration parameters before each execution. The simulations start with the 0.5 meter baseline and increase by .2 meters for each run until reaching the maximum

useful data distance of 6.1 meters.

The final set of simulations focused on several key separations and how the system performs when moving the camera anchor point towards and away from the boom anchor point. This test looks to find if there is any correlation between camera distance from the boom and system accuracy.

Table 4: Simulation Flight Parameters

Flight Number	Data File	Starting Pulse	Ending Pulse
1	2019.Mar.13_18.33.52.4719014.UTC_Flt5_2.log_relnav	1	5001
2	2019.Mar.13_18.33.52.4719014.UTC_Flt5_2.log_relnav	6100	9100
3	2019.Mar.13_18.33.52.4719014.UTC_Flt5_2.log_relnav	14700	17700
4	2019.Mar.14_20.05.26.8065267.UTC.log_relnav	14000	19000
5	2019.Mar.14_20.05.26.8065267.UTC.log_relnav	25000	28000
6	2019.Mar.14_20.05.26.8065267.UTC.log_relnav	30300	32800
7	2019.Mar.14_20.05.26.8065267.UTC.log_relnav	46000	50000
8	2019.Mar.14_20.05.26.8065267.UTC.log_relnav	55000	59000
9	2019.Mar.14_20.05.26.8065267.UTC.log_relnav	60000	66000

IV. Results and Analysis

With the simulation environment described in the previous section, the test flights provide the data necessary to determine the effect different camera baselines have on the accuracy of the end-to-end relative localization system. This section presents the results obtained, focusing on the position and orientation error data and how mean error grows or shrinks across the flights. The data is initially presented in graphs to show the visual change as separation grows, with additional tables providing the specific values of the overall mean and standard deviation for the data. The first three sections, section 4.1, Section 4.2, section 4.3 analyze the data form a specific flight and compare the un-occluded and occluded results. Section 4.4 analyses data across several flights and how distance, separation and error interact with each other. The final sections provide some interesting results found during the testing process; section 4.5 discusses what happens as the stereo vision system is moved towards or away from the boom anchor point, and section 4.6 analyzes how separation impacts point cloud density.

4.1 Position and Orientation Error without Occlusion

Key simulation results for flight 1 are shown in Figure 20, Figure 21, Figure 22 and Figure 23. This data was gathered without boom occlusion interference and illustrates the effect of increasing the baseline between the stereo cameras. Each dot in the graphs represents one relative pose estimated by the Iterative Closest Point (ICP) algorithm, each dot in position error has a counterpart in orientation error.

Figure 20 and Figure 21 shows the error on the x (red), y (green), and z (blue) position estimates for each relative pose output by the algorithm as described in III. In these plots, the x axis is the distance of the receiver away from the tanker cameras,

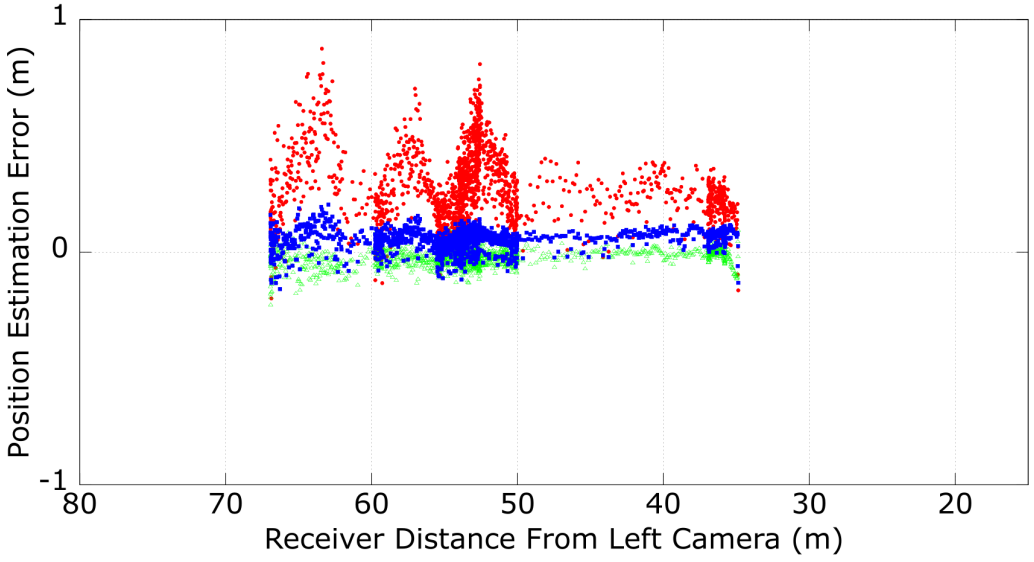
where the distance decreases going from left to right in the graph, and the y-axis is the error ranging from -1 to 1 meters.

Figure 22 and Figure 23 show the error for the roll (red), pitch (green), and yaw (blue) orientation estimates for each relative pose output by the algorithm. For these plots, the x axis is the distance of the receiver away from the tanker cameras, where the distance decreases going from left to right in the graph, and the y-axis is the error ranging from -5 to 5 degrees.

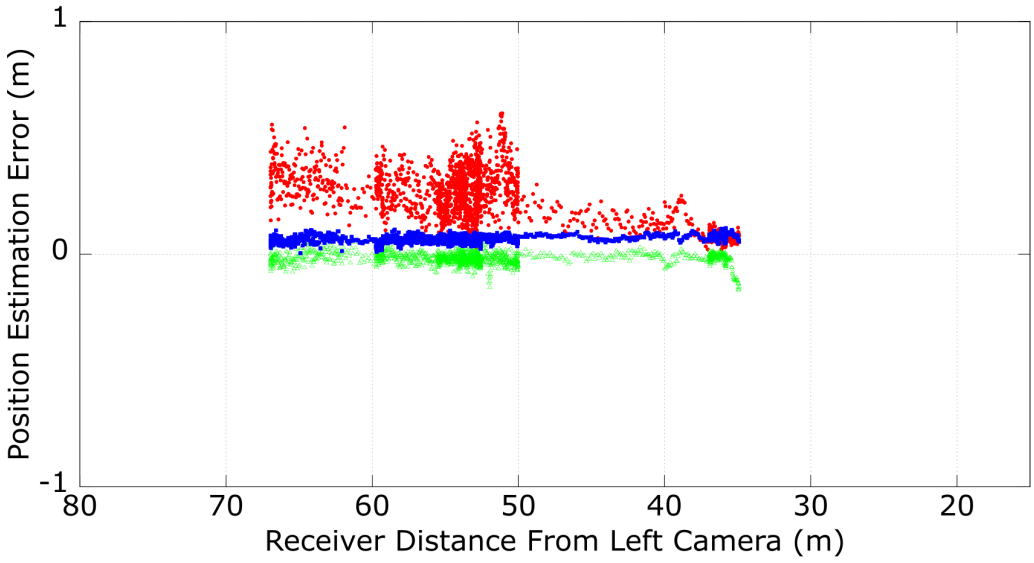
The graph on the left correspond with a 0.5m separation, while the graph on right has the cameras separated by 1.5m, with the two lower graphs representing the increased separation to 2.5m and 3.5m respectively. The resulting errors grow significantly smaller when moving from the leftmost graph down to the rightmost graph, demonstrating the potential advantage of increasing the baseline between cameras. For position results, the largest area of error corresponded to x, with positive x extending in the direction of the aircrafts nose, measuring the forwards and backwards pose estimation results. Comparatively, y, left and right, and z, up and down, saw much smaller errors and decreases in error. Position error is visibly larger at further distances and decreases as the receiver moves closer to the cameras.

The orientation graphs do not have a visible correlation between distance and error. However, as separation increases, the points in orientation graphs grow closer together and form more solid lines. The overall error for each orientation graph shows a visible decreases as separation increases.

To understand the details of camera separation impact, Table 5 presents a summary of the simulation data for multiple camera separation distances. The columns outline the separation in meters and the corresponding position and orientation error, as calculated by Root Mean Square Error (RMSE), for that simulation. The table also provides the standard deviation for each simulation. The “Baseline change” col-

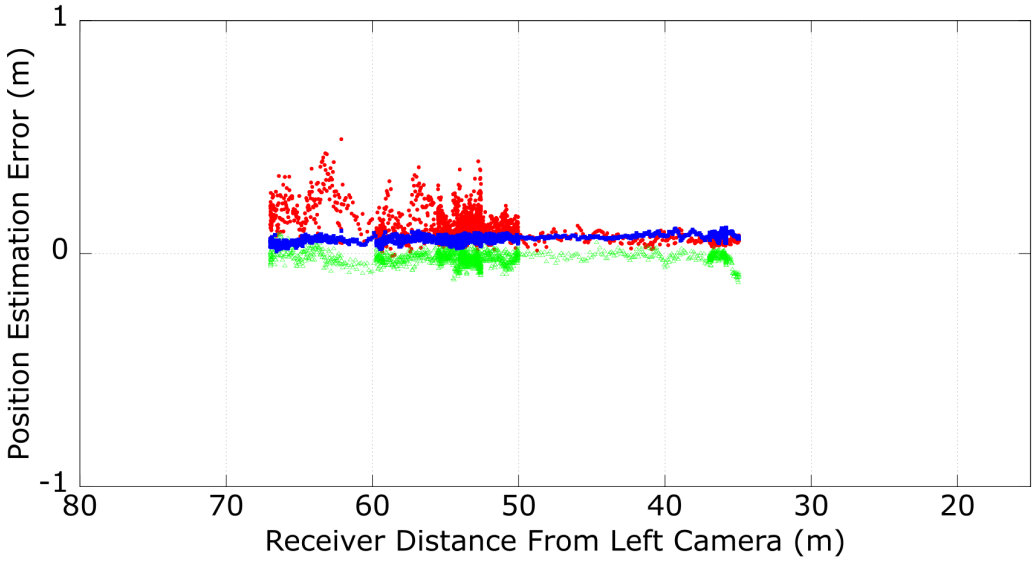


(a) 500mm separation

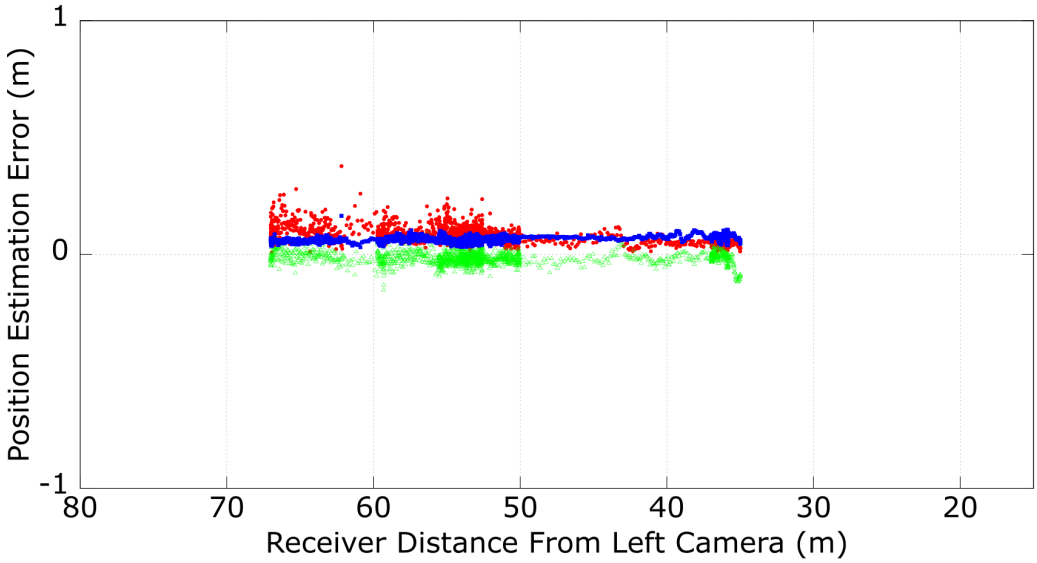


(b) 1500mm separation

Figure 20: Flight 1 (500mm, 1500mm) - position error comparison without occlusion (x-red, y-green, z-blue)

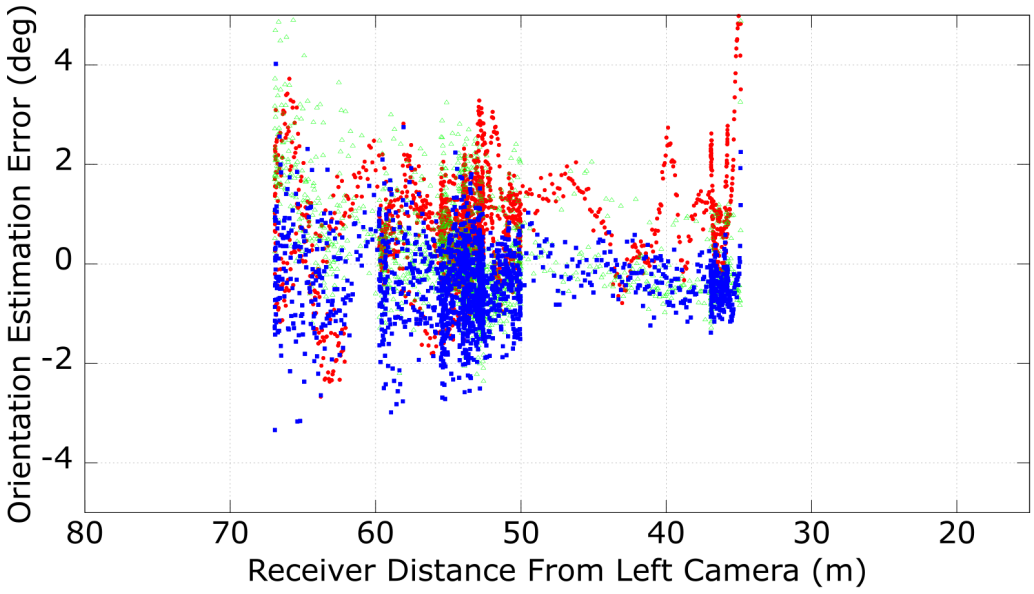


(a) 2500mm separation

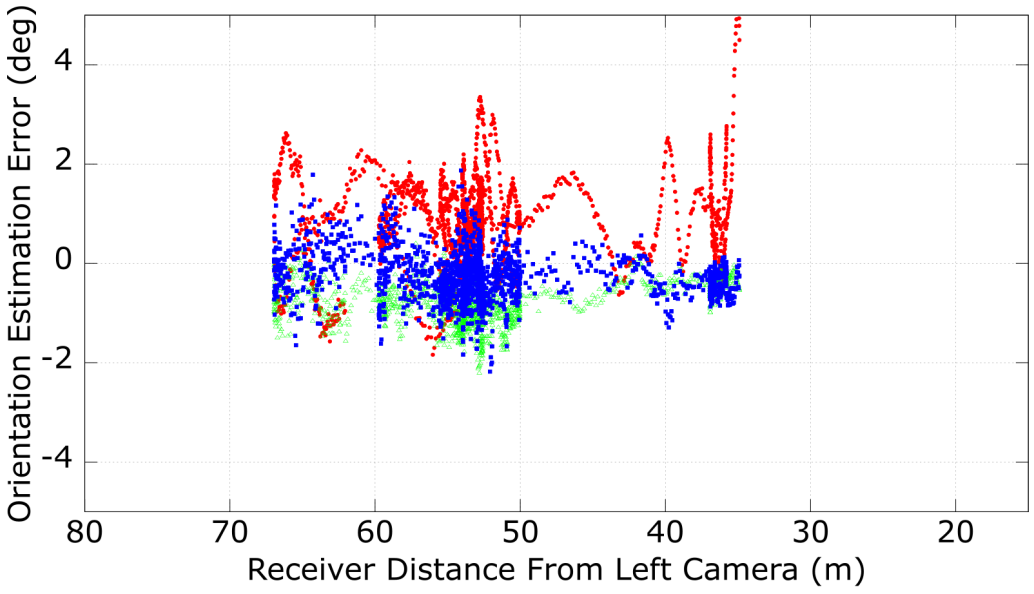


(b) 3500mm separation

Figure 21: Flight 1 (2500mm, 3500mm) - position error comparison without occlusion (x-red, y-green, z-blue)

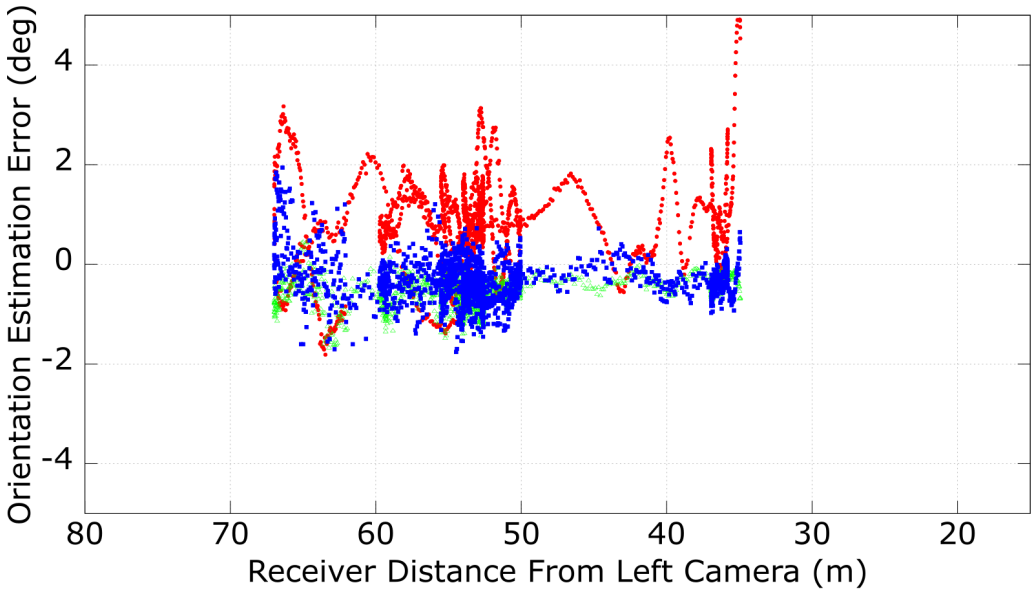


(a) 500mm separation

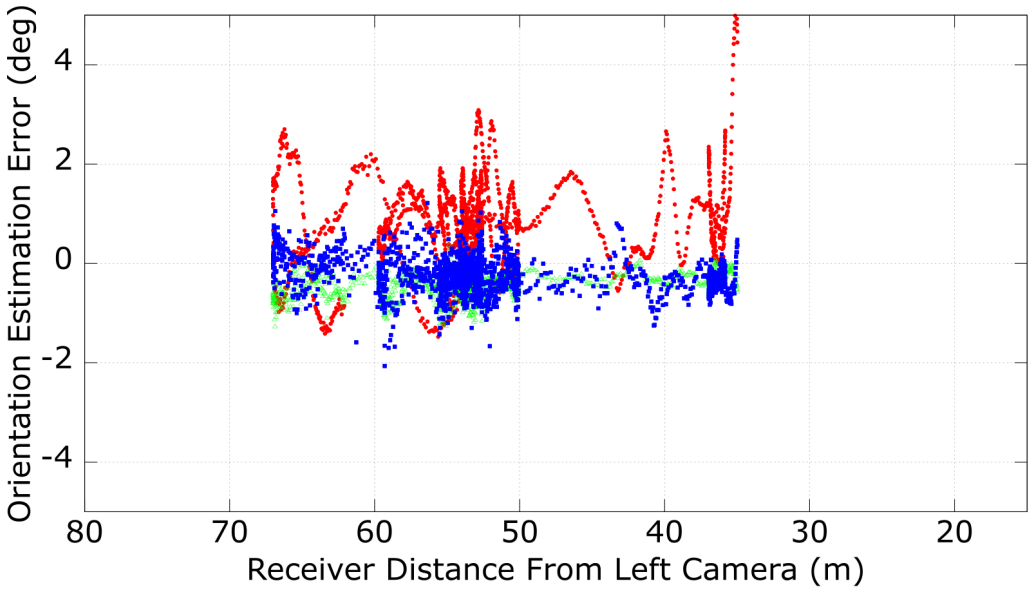


(b) 1500mm separation

Figure 22: Flight 1 (500mm,1500mm) - orientation error without occlusion (roll-red, pitch-green, yaw-blue)



(a) 2500mm separation



(b) 3500mm separation

Figure 23: Flight 1 (2500mm,3500mm) - orientation error without occlusion (roll-red, pitch-green, yaw-blue)

umn shows the level of improvement as the fraction of the current separation error over the 0.5 meter baseline separation. Analyzing the trends in these four categories, provides insightful information into how separation impacts the stereo vision system.

As separation increases, the mean error decreases in both position and orientation pose estimation. Standard deviation follows the same trend. The largest decreases occur up until the 3 meter mark, after which the error decrease slows dramatically. Based on these results, the optimal separation lies close to 3 meters.

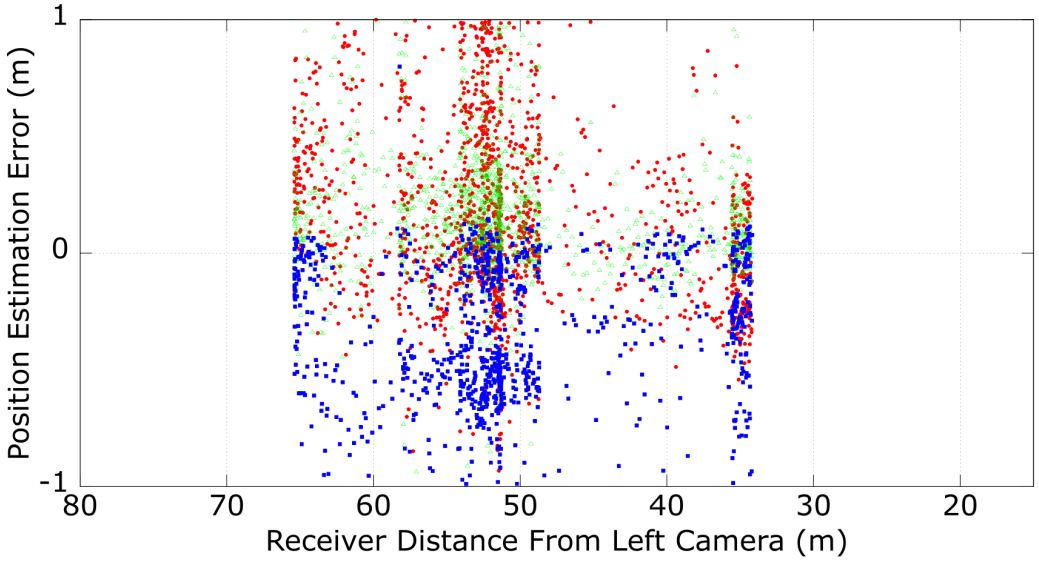
Table 5: Flight 1 - position and orientation error mean and baseline change without boom

Separation (m)	Position Mean Error (m)	Baseline Change	Position Standard Deviation	Orientation Mean Error (degree)	Baseline Change	Orientation Standard Deviation
0.5	0.14272	N/A	0.242963	0.857128	N/A	1.63038
0.7	0.145064	1.02	0.132542	0.740581	0.86	1.1593
0.9	0.126729	0.89	0.0921424	0.762931	0.89	1.21038
1.1	0.123637	0.87	0.0790038	0.775768	0.91	1.17029
1.3	0.109583	0.77	0.0745573	0.753543	0.88	1.131
1.5	0.112706	0.79	0.0779227	0.771298	0.90	1.16408
1.7	0.0971804	0.68	0.0776579	0.722796	0.84	1.09725
1.9	0.0914787	0.64	0.0723979	0.73326	0.86	1.10889
2.1	0.0785932	0.55	0.0677291	0.72267	0.84	1.08021
2.3	0.0744252	0.52	0.064473	0.686836	0.80	1.01176
2.5	0.0709098	0.50	0.0765352	0.693128	0.81	1.0994
2.7	0.0678361	0.48	0.0592771	0.658626	0.77	0.964544
2.9	0.0682052	0.48	0.0610952	0.67439	0.79	1.01075
3.1	0.0604539	0.42	0.0538516	0.651078	0.76	0.942308
3.3	0.0589868	0.41	0.052083	0.608641	0.71	0.865696
3.5	0.057397	0.40	0.0508924	0.602663	0.70	0.864637
3.7	0.057844	0.41	0.0558312	0.612981	0.72	0.880795
3.9	0.0560572	0.39	0.0525336	0.607139	0.71	0.864995
4.1	0.0565778	0.40	0.0534593	0.602176	0.70	0.865908
4.3	0.0533347	0.37	0.0511624	0.587367	0.69	0.830677
4.5	0.0543009	0.38	0.0500564	0.58409	0.68	0.836789
4.7	0.0527193	0.37	0.0511995	0.575761	0.67	0.814586
4.9	0.0537193	0.38	0.0514234	0.578177	0.67	0.823434
5.1	0.0515539	0.36	0.0511899	0.566712	0.66	0.803949
5.3	0.0507589	0.36	0.0492589	0.555105	0.65	0.779517
5.5	0.0508381	0.36	0.0488412	0.543169	0.63	0.765516
5.7	0.0526832	0.37	0.0515654	0.559616	0.65	0.797931
5.9	0.0523546	0.37	0.0598181	0.558619	0.65	0.795961
6.1	0.0534953	0.37	0.0557641	0.557617	0.65	0.802492

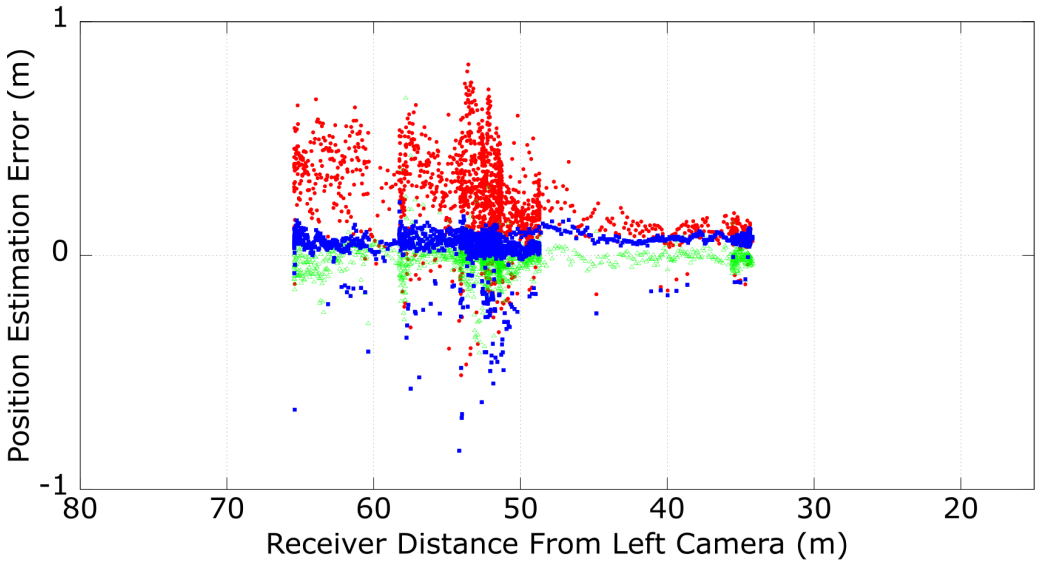
4.2 Position and Orientation Error with Occlusion

This set of data includes boom occlusion into the simulation. The boom model is set to visible and the boom filtering of the generated point cloud is enabled. When the boom is introduced into the simulation, there is a large increase in error. Figure 24, Figure 25, Figure 26, and Figure 27 show results of flight 1 with the boom present. Both the table and graph follow the same principles outlined in section 4.1.

When a boom is present, there also appears to be an "optimal" separation of around 3m, after which there is little to no advantage to increasing baseline displacement. This "optimal" may be due to more of the aircraft being out of the field of view of one of the two (or both) cameras when the baseline is larger. The data shows that not every separation saw a direct improvement over the baseline with some values hovering near the same mean, with diminishing returns in both sets of data as the distance grew between the cameras.

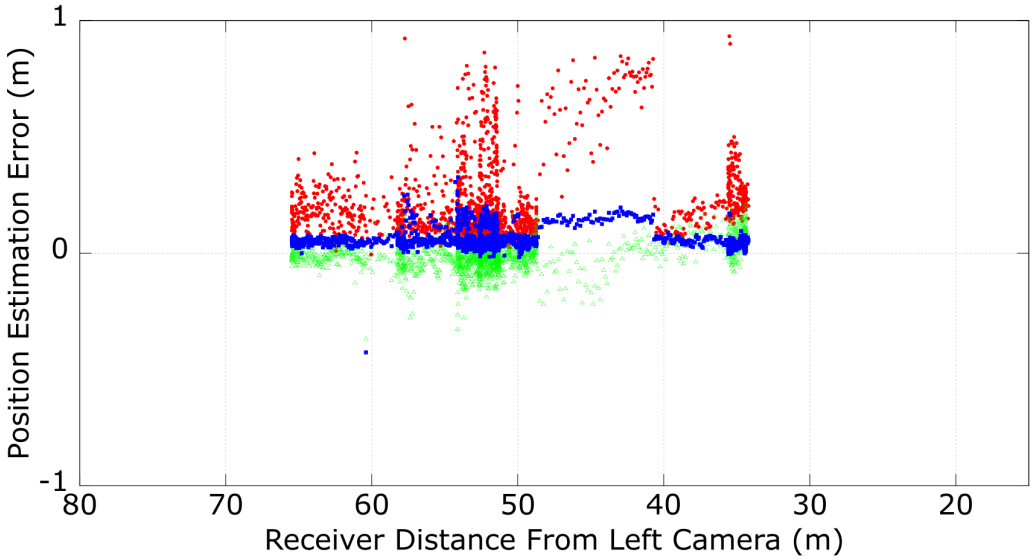


(a) 500mm separation

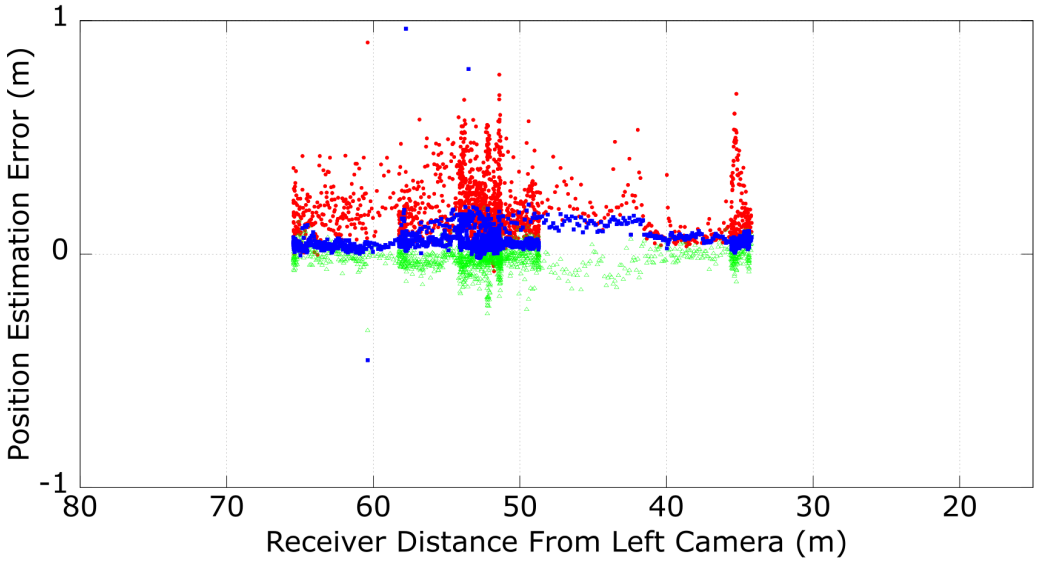


(b) 1500mm separation

Figure 24: Flight 1 (500mm,1500mm) - position error comparison with occlusion (x-red, y-green, z-blue)

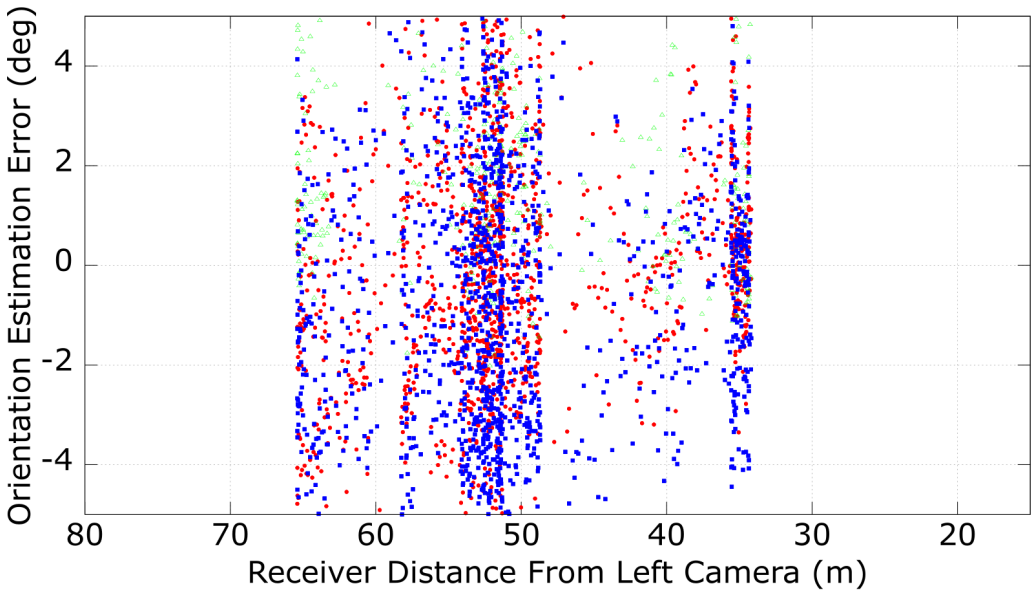


(a) 2500mm separation

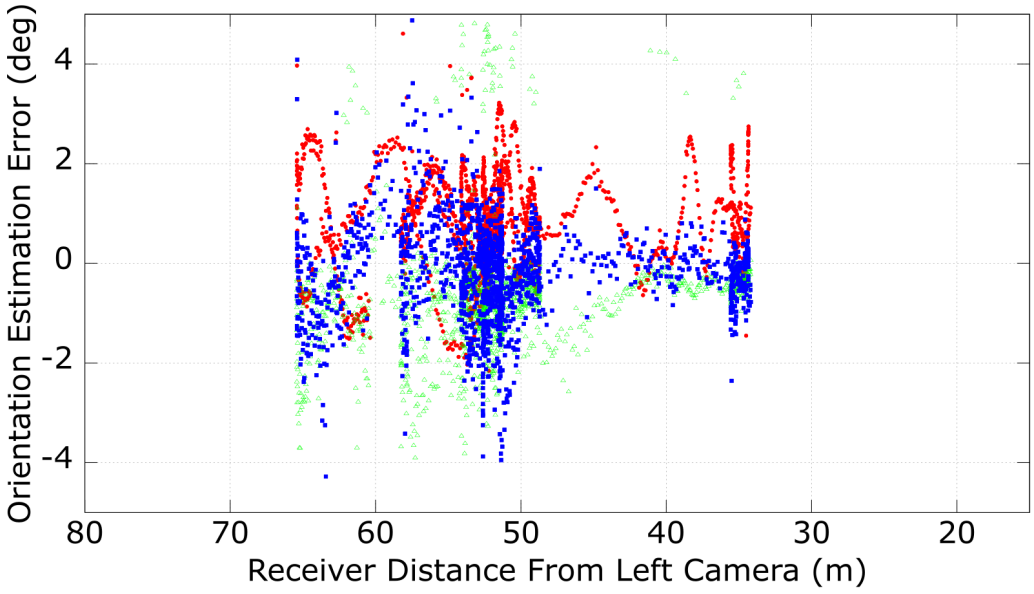


(b) 3500mm separation

Figure 25: Flight 1 (2500mm,3500mm)- position error comparison with occlusion (x-red, y-green, z-blue)

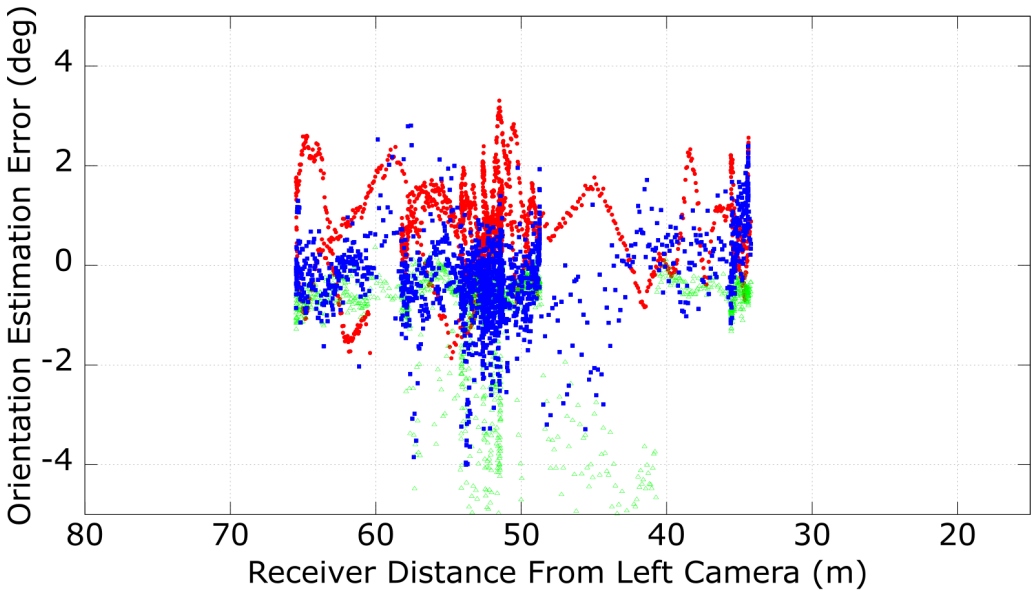


(a) 500mm separation

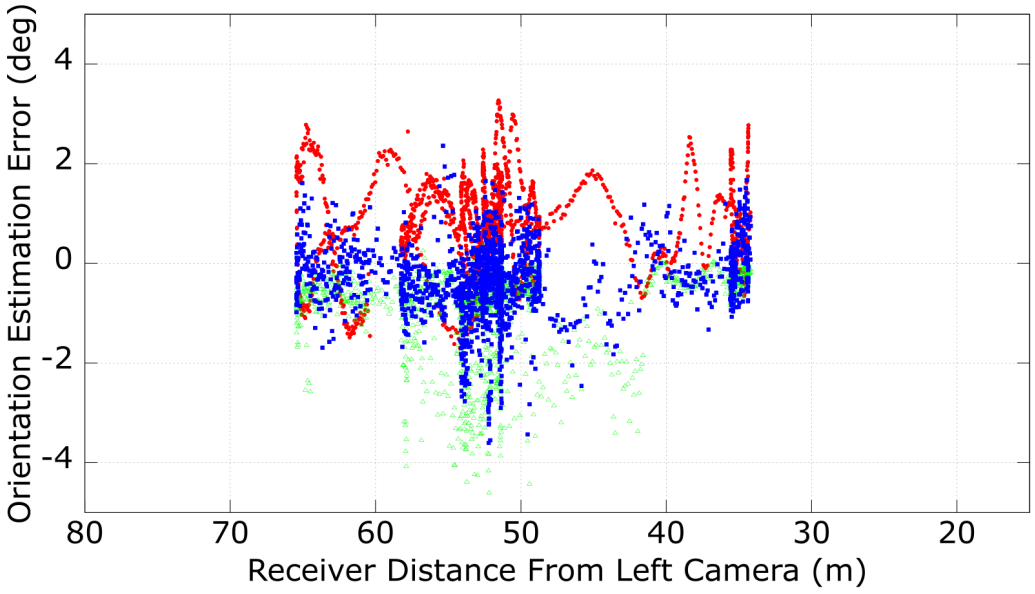


(b) 1500mm separation

Figure 26: Flight 1 (500mm,1500mm) - orientation error with occlusion (roll-red, pitch-green, yaw-blue)



(a) 2500mm separation



(b) 3500mm separation

Figure 27: Flight 1 (2500mm,3500mm) - orientation error with occlusion (roll-red, pitch-green, yaw-blue)

Table 6: Flight 1 - position and orientation error mean and baseline change with boom

Separation (m)	Position Mean Error (m)	Baseline Change	Position Standard Deviation	Orientation Mean Error (degree)	Baseline Change	Orientation Standard Deviation
0.5	0.698941	N/A	1.53455	12.0346	N/A	18.0018
0.7	0.483838	0.69	0.977582	5.88848	0.49	11.6422
0.9	0.263942	0.38	0.493233	3.06562	0.25	6.90704
1.1	0.318252	0.46	0.564652	2.41715	0.20	5.40076
1.3	0.303534	0.43	0.53132	2.3719	0.20	5.32402
1.5	0.319371	0.46	0.553935	1.29628	0.11	2.49102
1.7	0.304388	0.44	0.545747	2.61989	0.22	3.49863
1.9	0.325734	0.47	0.54043	2.63835	0.22	3.54646
2.1	0.286646	0.41	0.46701	2.76529	0.23	3.00157
2.3	0.215401	0.31	0.42791	2.48998	0.21	3.07719
2.5	0.230845	0.33	0.425443	1.98693	0.17	2.38375
2.7	0.103552	0.15	0.132508	0.836648	0.07	1.51365
2.9	0.0937251	0.13	0.111156	0.843231	0.07	1.50108
3.1	0.13128	0.19	0.281811	1.17753	0.10	2.5353
3.3	0.201726	0.29	0.430713	1.72879	0.14	3.82547
3.5	0.118752	0.17	0.251461	0.985216	0.08	2.03996
3.7	0.102698	0.15	0.118319	0.9003	0.07	1.58127
3.9	0.11305	0.16	0.129542	0.968611	0.08	1.76417
4.1	0.127038	0.18	0.167152	1.02454	0.09	1.93086
4.3	0.171775	0.25	0.250412	1.32595	0.11	2.67008
4.5	0.175044	0.25	0.214379	1.35817	0.11	2.72782
4.7	0.168141	0.24	0.217463	1.29985	0.11	2.60309
4.9	0.170111	0.24	0.225639	1.30024	0.11	2.65544
5.1	0.200102	0.29	0.254681	1.44396	0.12	2.93925
5.3	0.253201	0.36	0.280094	1.72721	0.14	3.48987
5.5	0.265333	0.38	0.282712	1.79339	0.15	3.69256
5.7	0.266704	0.38	0.284908	1.80079	0.15	3.8212
5.9	0.224789	0.32	0.260835	1.50969	0.13	3.12675
6.1	0.185316	0.27	0.251238	1.28245	0.11	2.7825

4.3 Effect of Increasing Baseline

Both with and without boom occlusion, the data shows a reduction in error as stereo camera separation is increased from the 0.5 meter baseline. Table 5 and Table 6 demonstrate that the error mean and standard deviation generally improves sub-linearly with separation.

Analyzing the 3500mm separation when a boom is present (Figure 21(b)) and the other results, shows that increasing separation causes a significant decrease in error when a boom is present in the images. The new error, however, is still not smaller than the 3500mm separation without a boom, showing that larger separation does not completely overcome effect of having a boom in the imagery.

4.4 Aggregate Error

Figure 28 is the aggregate error over all of the flights, including the previous data in Table 5 and the additional data in Appendix A and Appendix B. The figure represents the mean error at specific distance for each camera separation and demonstrates the relationship between separation, receiver distance from the camera, and error. The key trends are how error decreases and levels out as separation increases as well as how error grows with receiver distance from the camera. From the gathered data, going beyond a three meter separation does not provide any major benefit.

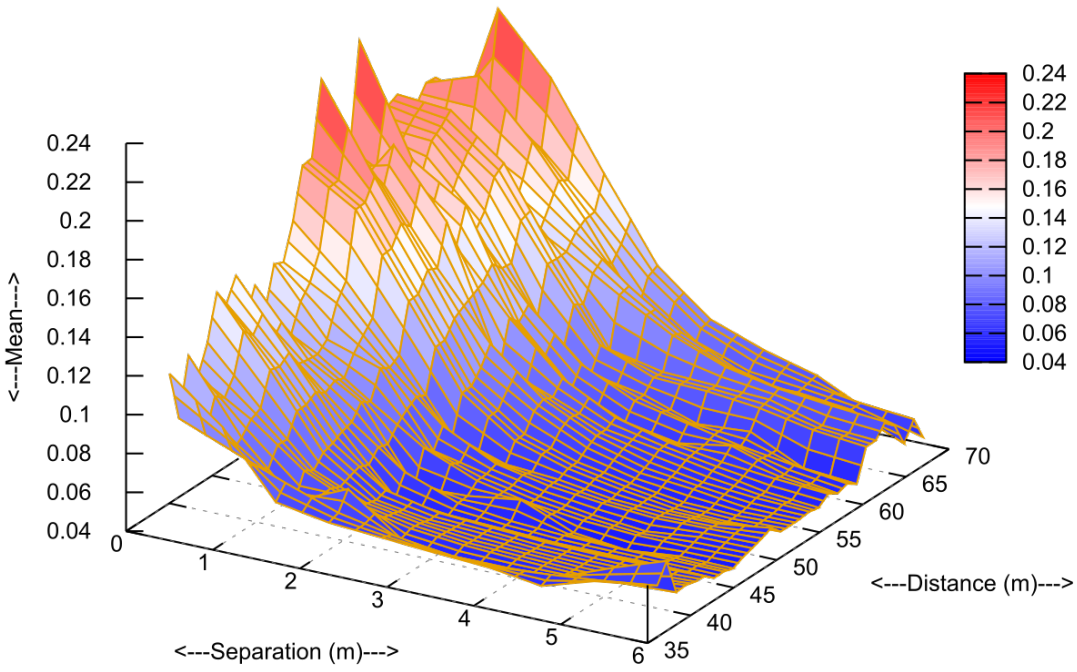


Figure 28: Flight 1-9 aggregate position error (without occlusion)

4.5 Stereo Camera Placement

One component of the experiment involved testing if moving the cameras closer or further from the boom improved the performance of the system. Table 7 and Table 8 outlines the results of the experiment for several key camera separations. The test focused on the 0.5, 1.5, 2.5, and 3.5 meter separations and consisted of moving the stereo vision system anchor point from its original position; 5.5 meters away from the boom anchor point. In the first half of the test, the stereo vision system anchor point was moved 1 meter at a time in the positive X, in the tanker frame, up to 9 meters (14.5 meters away from the boom anchor point). The second half of the test moved the camera system closer to the boom, once again starting at 5.5 meters and ending 3.5 meters past the boom anchor point. The chart distance utilizes the 3D world coordinates centred on the boom anchor point, positive values move the camera anchor point towards the tanker's nose while negative values are closer to the tail of the tanker. Note that any placement of value less than 0 meters is actually "behind" the boom on the aircraft, as opposed to its current location in front of the boom. The flight data remains the same for each simulation and based on the camera position the receiver is viewed as closer or farther away, shifting all of the flight data. The results demonstrate greater position error at farther distances and less error when the cameras are closer to the receiver. Looking at the orientation data and how there is no specific trend, much of the improvement appears to be due to the accuracy improvement of the receiver getting closer, rather than due to the system being less affected by boom occlusion.

Table 7: Flight 5 - camera position movement relative to the boom anchor point error results (0.5 and 1.5 meter separation)

	Separation (m)			
	0.5		1.5	
X Distance From Boom Anchor Point (m)	Position Mean Error (m)	Orientation Mean Error (degree)	Position Mean Error (m)	Orientation Mean Error (degree)
-3.5	0.697337	9.4605	0.0844706	0.746237
-2.5	0.66288	9.52098	0.0873317	0.658521
-1.5	0.6998804	10.50895	0.083894	0.684528
-0.5	0.659331	10.9381	0.10894	0.76418
0.5	0.651451	9.2237	0.110907	0.78443
1.5	0.542367	8.6961	0.123047	0.74921
2.5	0.54538	8.80360	0.118128	0.770488
3.5	0.56596	10.5208	0.119267	1.00491
4.5	0.545698	8.24269	0.146377	1.42339
5.5	0.558933	11.6565	0.137135	1.39769
6.5	0.551606	12.5563	0.110329	0.850919
7.5	0.587937	7.13598	0.112643	0.849904
8.5	0.582971	9.33242	0.114061	0.930782
9.5	0.592089	6.68037	0.122674	1.01899
10.5	0.60265	7.95929	0.13413	1.12487
11.5	0.63067	10.4768	0.183129	1.74243
12.5	0.70745	11.2179	0.241222	2.29963
13.5	0.77450	11.2135	0.418964	1.89604
14.5	0.83026	11.8676	0.505824	2.62058

Table 8: Flight 5 - camera position movement relative to the boom anchor point error results (2.5 and 3.5 meter separation)

	Separation (m)			
	2.5		3.5	
X Distance From Boom Anchor Point (m)	Position Mean Error (m)	Orientation Mean Error (degree)	Position Mean Error (m)	Orientation Mean Error (degree)
-3.5	0.080148	0.571429	0.0789425	0.551725
-2.5	0.091244	0.550063	0.088453	0.636821
-1.5	0.0807767	0.550537	0.0868785	0.57023
-0.5	0.105244	0.745276	0.085152	0.698271
0.5	0.104749	0.81958	0.087737	1.77316
1.5	0.0973745	0.98638	0.0755669	0.674535
2.5	0.0833458	0.993809	0.0753535	0.663379
3.5	0.129752	0.986193	0.0863783	0.743388
4.5	0.19535	0.90528	0.10438	0.966175
5.5	0.117385	0.10523	0.908261	0.89695
6.5	0.10761	0.817268	0.0910895	0.701105
7.5	0.105215	0.791185	0.088024	0.74421
8.5	0.113812	0.791419	0.100593	0.788512
9.5	0.121549	0.81780	0.11159	0.837996
10.5	0.12415	0.857519	0.121645	0.912832
11.5	0.119189	0.823122	0.113753	0.885182
12.5	0.11780	0.837525	0.111248	0.931774
13.5	0.113649	0.963077	0.115194	0.835091
14.5	0.11535	0.974307	0.11717	0.845131

4.6 Sensed Points

One factor that initially seemed to have a significant impact on performance as baseline increased turned out to be inconsequential: the number of points found from stereo triangulation. When running the simulations, a gradual decrease in the amount of sensed points as the camera separation grew was observed. These are the points which form the yellow point cloud generated by from the disparity map and reprojected into the virtual world. Figure 29 illustrates the decrease in relation to separation and distance. Generally more features were matched at closer ranges and with the amount of points falling off as the distance between the tanker and receiver grew.

Table 9 further breaks down the data showing the downward trend for a separation in each distance category. The table consists of the mean across several flights at specific distances. The largest reduction seen being 50% of the points from the 0.5 meter baseline separation point cloud. In the case of this system the decrease in the amount of sensed points did not increase the position or orientation pose estimation error. However, if a future system requires larger point clouds, this is one trade off to consider when looking at camera separation as a means of increasing the accuracy of the system.

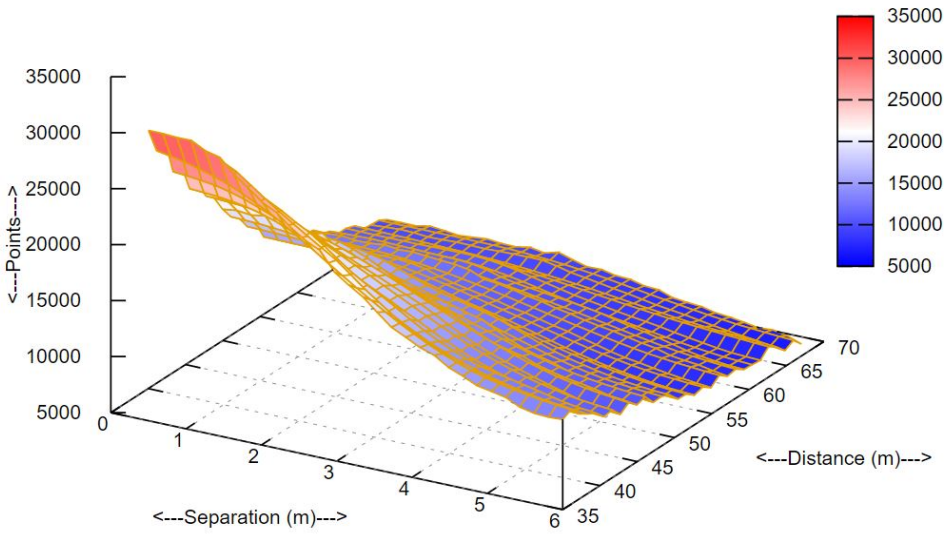


Figure 29: Sensed Points Comparison (without occlusion)

Table 9: Flight 1-9 mean sensed points per distance

	Distance (m)						
Separation(m)	35	40	45	50	55	60	65
0.5	31303	23408	19622	14793	12897	10840	9762
0.7	30923	23398	19843	14992	12851	10803	9752
0.9	30819	23360	19650	14718	12992	11002	9974
1.1	30875	23132	19923	14779	12823	10685	9952
1.3	30219	23128	19406	14710	12966	10731	9647
1.5	28925	22236	19761	14788	12827	10852	9800
1.7	28177	22395	19233	14487	12886	10700	9871
1.9	27291	21440	18907	14851	12798	10850	9747
2.1	26279	20747	18296	13839	12745	11110	9746
2.3	25539	19949	17688	13551	12326	10640	9942
2.5	24553	19247	16979	13221	11869	10469	9730
2.7	23531	18763	16447	12881	11641	10357	9968
2.9	22302	17882	16134	12694	11327	10343	9809
3.1	21243	17442	15732	12313	11148	9740	9255
3.3	20371	16767	15265	11735	10736	9457	9187
3.5	19076	15964	14714	11256	10569	9343	8902
3.7	17993	15103	14111	11119	10203	9041	8719
3.9	17374	14631	13434	10610	9905	8661	8461
4.1	16733	13969	12821	10039	9693	8608	8310

V. Conclusions

There are three primary conclusions drawn from the simulation data:

1. Separating the cameras leads to an increase in accuracy of the relative position estimation system.

2. While there is some decrease in error without the boom, the increase in accuracy when a boom is present is almost required. In the occluded scenario, 0.7 meters of error at 0.5 meter separation is beyond the acceptable thresh hold, but errors approximately the same magnitude as the 0.5 meter separation without a boom can be achieved by increasing the displacement to around 3 meters.

3. The data supports an optimal separation of the cameras at around 3 meters for both the un-occluded and occluded scenarios.

While previous stereo vision systems utilized a fixed camera displacement and some occlusions analysis, the conducted experiments work to understand the impact of widening the stereo baseline, especially in the presence of significant occlusion. By replaying representative flight data and reprojecting 3D point clouds for use with Iterative Closest Point (ICP), the system is able to record the error between the estimated pose and the receiver in the virtual world. Results showed a reduction in error in position and orientation as a result of increasing the baseline camera separation in the stereo vision system. These improvements, were sub-linear with diminishing returns. There appears to be an optimal separation value though that should be evaluated for each scenario. In this case, the optimal separation is around 3 meters when large occlusions due to the boom are present.

The research data and conclusions help guide future Automated Aerial Refueling (AAR) efforts utilizing a stereo vision system. The separation between cameras be-

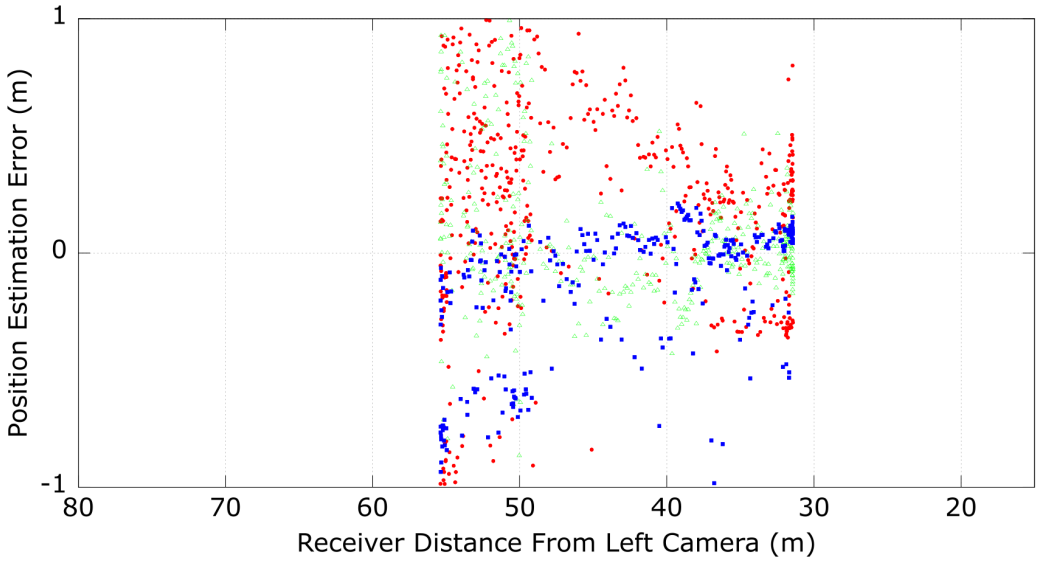
comes critical to achieve the overall accuracy required for relative navigation-based systems, especially in the case when boom occlusion becomes a serious factor.

5.1 Future Work

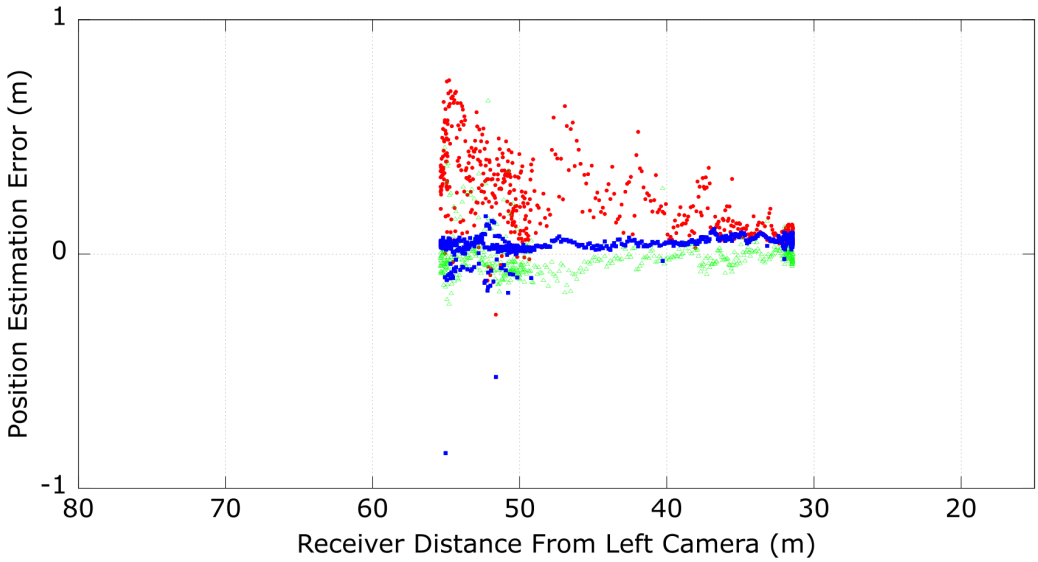
There are several avenues for future work on the AAR stereo vision configuration. One area entails testing the separation impact using real world imagery to further verify the validity of the simulation results. Test flights or ground camera testing with a few key chosen camera separations would help show how the additional factors, in the real world, impact the stereo vision system and the error reduction gained from increased camera separation.

Other testing required before full implementation of AAR includes testing other changes to the stereo system such as the impact of different cameras and different sized aircraft. Wingspan and aircraft size can have an impact on the accuracy of the system. Finally, further testing utilizing a moving boom is necessary. The current tests focused on a static boom, in the final system the boom with move, changing what portions of the aircraft are occluded. Tests with a moving boom will provide a more realistic simulation scenario.

Appendix A. Additional Results - Graphs

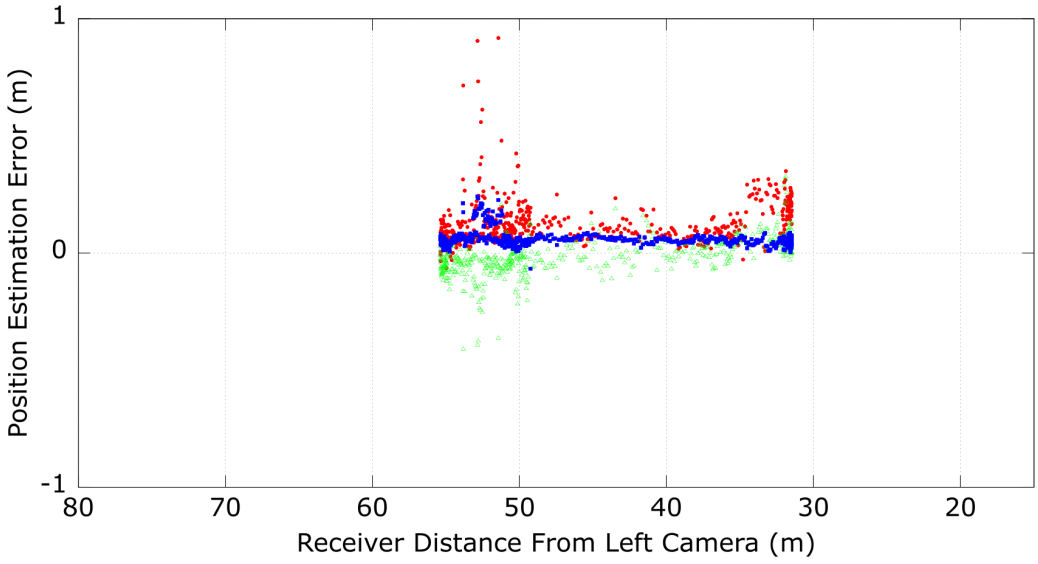


(a) 500mm separation

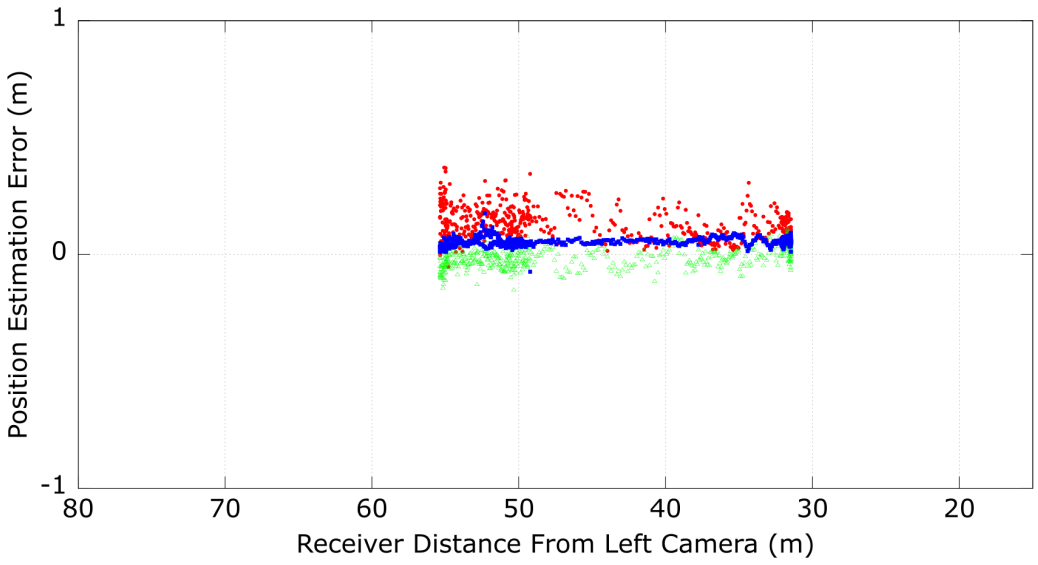


(b) 1500mm separation

Figure 30: Flight 6 (500mm,1500mm) - position error comparison with occlusion (x-red, y-green, z-blue)

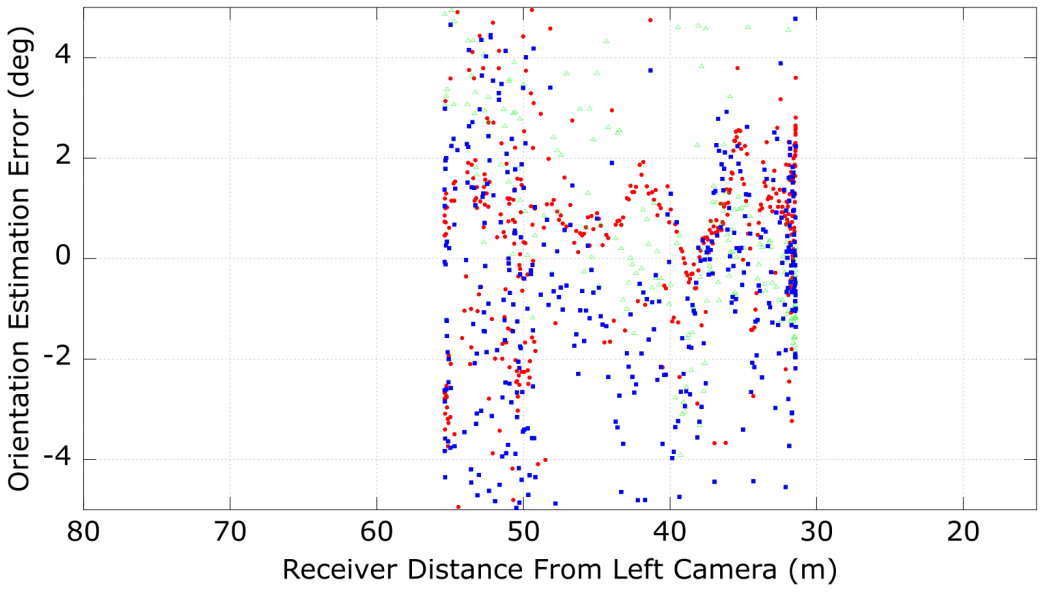


(a) 2500mm separation

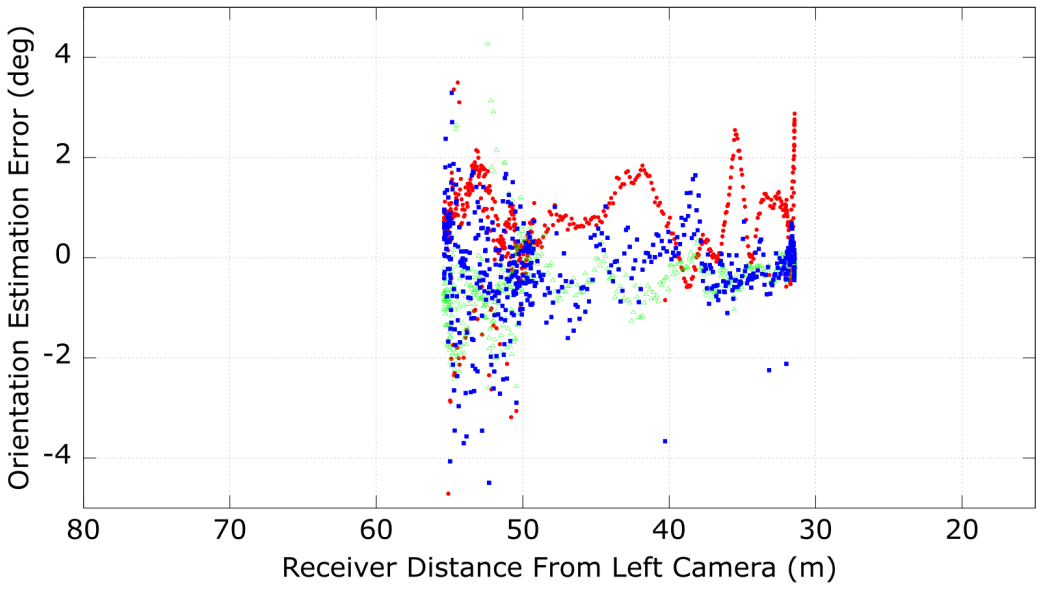


(b) 3500mm separation

Figure 31: Flight 6 (2500mm,3500mm) - position error comparison with occlusion (x-red, y-green, z-blue)

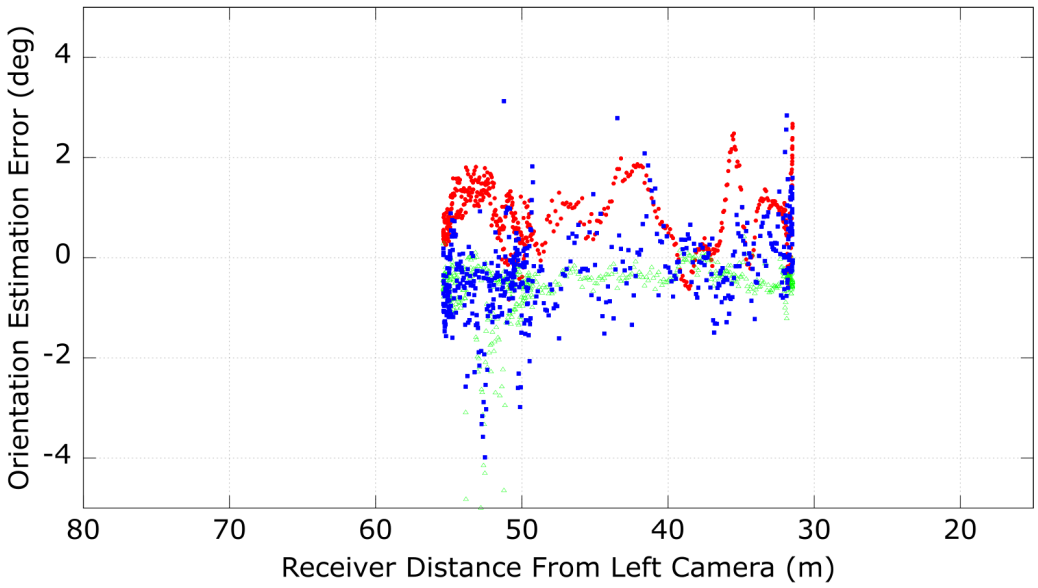


(a) 500mm separation

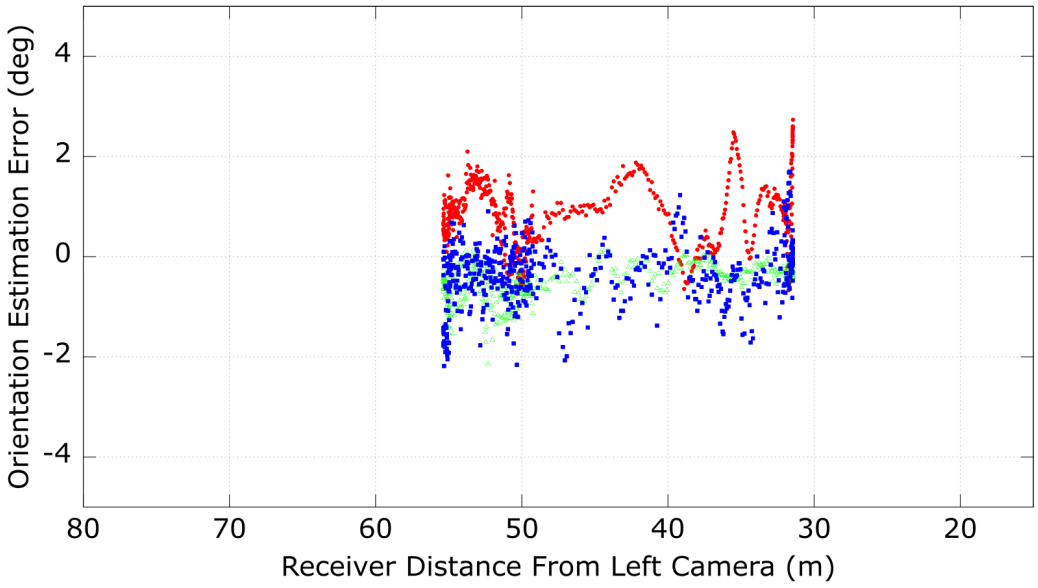


(b) 1500mm separation

Figure 32: Flight 6 (500mm,1500mm) - orientation error with occlusion (roll-red, pitch-green, yaw-blue)



(a) 2500mm separation



(b) 3500mm separation

Figure 33: Flight 6 (2500mm,3500mm) - orientation error with occlusion (roll-red, pitch-green, yaw-blue)

Appendix B. Additional Results - Tables

Table 10: Flight 6 - Position and orientation error mean and baseline change with boom

Separation (m)	Position Mean Error (m)	Baseline Change	Position Standard Deviation	Orientation Mean Error (degree)	Baseline Change	Orientation Standard Deviation
0.5	0.551606	N/A	1.23306	10.0346	N/A	13.7658
0.7	0.44538	0.81	0.985701	5.84093	0.59	10.5804
0.9	0.272641	0.50	0.5859	1.90031	0.19	5.03549
1.1	0.199913	0.36	0.28079	1.04245	0.10	3.98175
1.3	0.141387	0.26	0.18162	0.948224	0.09	2.94083
1.5	0.154179	0.28	0.217048	0.770488	0.08	2.12536
1.7	0.118128	0.21	0.179376	0.812008	0.08	1.87948
1.9	0.125155	0.23	0.145822	0.755505	0.08	1.62521
2.1	0.129989	0.24	0.127795	0.71065	0.07	1.33968
2.3	0.0962992	0.17	0.119918	0.670152	0.07	1.51929
2.5	0.09068	0.16	0.0749114	0.663379	0.07	1.20521
2.7	0.077994	0.14	0.0591981	0.658738	0.07	1.02728
2.9	0.075797	0.14	0.0572016	0.682547	0.07	0.935417
3.1	0.0753535	0.14	0.0514878	0.686425	0.07	0.959137
3.3	0.0731486	0.13	0.0526902	0.712565	0.07	0.944487
3.5	0.0719288	0.13	0.0540936	0.745276	0.07	0.929704
3.7	0.0769078	0.14	0.0596403	0.77062	0.08	1.02771
3.9	0.0806263	0.15	0.0971555	0.80146	0.08	1.36334
4.1	0.0833458	0.15	0.0824319	0.80547	0.08	1.17052

Bibliography

1. Peter R. Thomas, Ujjar Bhandari, Steve Bullock, Thomas S. Richardson, and Jonathan L. Du Bois. Advances in air to air refuelling. *Progress in Aerospace Sciences*, 71:14–35, 2014.

2. Christopher Bolkcom and Jon D Klaus. Air force aerial refueling methods: flying boom versus hose-and-drogue. LIBRARY OF CONGRESS WASHINGTON DC CONGRESSIONAL RESEARCH SERVICE, 2005.

3. Kyle P. Werner. Precision relative positioning for automated aerial refueling from a stereo imaging system. 2015.

4. Mario L Fravolini, Marco Mammarella, Giampiero Campa, Marcello R Napolitano, and Mario Perhinschi. Machine vision algorithms for autonomous aerial refueling for uavs using the usaf refueling boom method. In *Innovations in Defence Support Systems–1*, pages 95–138. Springer, 2010.

5. Charles J Lloyd and Steven G Nigus. Effects of stereopsis, collimation, and head tracking on air refueling boom operator performance. In *Proceedings of the Image Society Annual Conference, Scottsdale, Arizona*, 2012.

6. Ujjar Bhandari, Peter R Thomas, and Thomas S Richardson. Bow wave effect in probe and drogue aerial refuelling. In *AIAA Guidance, Navigation, and Control (GNC) Conference*, page 4695, 2013.

7. M. M. Mostafa, A. M. Moussa, N. El-Sheimy, and Abu B. Sesay. A smart hybrid vision aided inertial navigation system approach for UAVs in a GNSS denied environment. *Navigation, Journal of the Institute of Navigation*, 65(4):533–547, 2018.

8. SC De Vries. Uavs and control delays. Technical report, TNO DEFENCE SECURITY AND SAFETY SOESTERBERG (NETHERLANDS), 2005.

9. William E. Dallmann. Infrared and electro-optical stereo vision for automated aerial refueling. 2019.

10. Marco Mammarella, Giampiero Campa, Marcello R. Napolitano, and Mario L. Fravolini. Comparison of point matching algorithms for the UAV aerial refueling problem. *Machine Vision and Applications*, 21(3):241–251, 2010.

11. Report details cause of Travis KC-10 refueling boom mishap, 2017. Online; accessed 25-November-2019.

12. Mari Yamaguchi. 5 Missing, 2 Found After F/A-18 Fighter, Refueling Tanker Crash Off Japan, 2018. Online; accessed 15-November-2019.

13. Steven M Ross. Formation flight control for aerial refueling. Technical report, AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL OF ENGINEERING AND , 2006.

14. Daniel T Johnson, Scott L Nykl, and John F Raquet. Combining stereo vision and inertial navigation for automated aerial refueling. *Journal of Guidance, Control, and Dynamics*, 40(9):2250–2259, 2017.

15. Graham Warwick. Afrl advances autonomous aerial refueling. *Aviation Week*, 2008.

16. Jace Robinson, Matt Piekenbrock, Lee Burchett, Scott Nykl, Brian Woolley, and Andrew Terzuoli. Parallelized iterative closest point for autonomous aerial refueling. In *International Symposium on Visual Computing*, pages 593–602. Springer, 2016.

17. Christopher Parsons, Zachary Paulson, Scott Nykl, William Dallman, Brian G. Woolley, and John Pecarina. Analysis of Simulated Imagery for Real-Time Vision-Based Automated Aerial Refueling. *Journal of Aerospace Information Systems*, 16(3):77–93, 2019.

18. Erin Lasley. Refueling through the century. Online; accessed 13-December-2019.

19. Stanley J Dougherty. Air refueling: The cornerstone of global reach-global power. Technical report, AIR WAR COLL MAXWELL AFB AL, 1996.

20. Mario L. Fravolini, Marco Mammarella, Giampiero Campa, Marcello R. Napolitano, and Mario Perhinschi. Machine vision algorithms for autonomous aerial refueling for UAVs using the USAF refueling boom method. *Studies in Computational Intelligence*, 304:95–138, 2010.

21. First Air-to-Air Refueling. Online; accessed 18-December-2019.

22. Roshawn Elizabeth Bowers. *Estimation algorithm for autonomous aerial refueling using a vision based relative navigation system.* PhD thesis, Texas A&M University, 2005.

23. Air Force accepts KC-46A, 2019. Online; accessed 2-January-2020.

24. Tracy J. Barnidge and Joseph L. Tchon. 3D display considerations for rugged airborne environments. *Display Technologies and Applications for Defense, Security, and Avionics IX; and Head- and Helmet-Mounted Displays XX*, 9470:947009, 2015.

25. Four Marine officers fired in wake of deadly 2018 collision during in-air refueling, 2018. Online; accessed 25-November-2019.

26. Ghozali S Hadi, Rivaldy Varianto, B Trilaksono, and Agus Budiyono. Autonomous uav system development for payload dropping mission. *The Journal of Instrumentation, Automation and Systems*, 1(2):72–22, 2014.

27. Dan Gettinger. Study: Drones in the FY 2019 Defense Budget, 2018. Online; accessed 10-November-2019.

28. Mario Luca Fravolini, Giampiero Campa, Marcello Napolitano, Antonio Ficola, Information Engineering, and Giampiero Campa. Evaluation of Machine Vision Algorithms for Autonomous Aerial Refueling for Unmanned Aerial Vehicles Submitted to : AIAA Journal of Aerospace Computing , Information and Communication Evaluation of Machine Vision Algorithms for Autonomous Aerial Refueling. (April 2005):1–23.

29. Tyler Rogoway and Joseph Trevithick. This Is What The Boom Operator's Station On The New KC-46 Tanker Actually Looks Like, 2019. Online; accessed 5-January-2020.

30. Bin Huang, Yong-rong Sun, Ling Wu, Jian-ye Liu, and Qing-hua Zeng. Monocular vision navigation sensor for autonomous aerial refueling. 2017.

31. John Valasek, Douglas Famularo, and Monika Marwaha. Fault-tolerant adaptive model inversion control for vision-based autonomous air refueling. *Journal of Guidance, Control, and Dynamics*, 40(6):1336–1347, 2017.

32. James Doebbler, Theresa Spaeth, John Valasek, Mark J Monda, and Hanspeter Schaub. Boom and receptacle autonomous air refueling using visual snake optical sensor. *Journal of Guidance, Control, and Dynamics*, 30(6):1753–1769, 2007.

33. Giampiero Campa, Mario Luca Fravolini, Antonio Ficola, Marcello Napolitano, Brad Seanor, and Mario Perhinschi. Autonomous Aerial Refueling for UAVs Using a Combined GPS-Machine Vision Guidance. (November 2015), 2012.

34. Camera Calibration and 3D Reconstruction. Online; accessed 1-October-2019.

35. Camera Intrinsics. Online; accessed 15-September-2019.

36. Camera calibration With OpenCV OpenCV 2.4.13.7 documentation. Online; accessed 13-December-2019.

37. Epipolar Geometry. Online; accessed 1-December-2019.

38. Zachary Paulson, Scott Nykl, John Pecarina, and Brian Woolley. Mitigating the effects of boom occlusion on automated aerial refueling through shadow volumes. *The Journal of Defense Modeling and Simulation*, 16(2):175–189, 2019.

39. Andriy Myronenko and Xubo Song. Point set registration: Coherent point drift. *IEEE transactions on pattern analysis and machine intelligence*, 32(12):2262–2275, 2010.

40. Paul J Besl and Neil D McKay. Method for registration of 3-d shapes. In *Sensor fusion IV: control paradigms and data structures*, volume 1611, pages 586–606. International Society for Optics and Photonics, 1992.

Acronyms

AAR Automated Aerial Refueling. iv, 2, 3, 4, 6, 9, 10, 11, 12, 13, 54, 55

AFIT Air Force Institute of Technology. 2, 3, 13, 19

ANT Autonomy and Navigation Technology. 19

ARO Aerial Refueling Officer. 1, 7

DGPS Differential GPS. 2, 12

DoD Depeartment of Defense. 10

GPS Global Positioning System. iv, 11, 12, 24

ICP Iterative Closest Point. 3, 13, 22, 30, 31, 34, 54

OpenCV Open Source Computer Vision Library. 20

RMSE Root Mean Square Error. 31, 35

RPA Remotely Piloted Aircraft. 10

UAV Unmanned Aerial Vehicle(s). iv, 2, 10, 11, 12

USAF United Sates Air Force. 4, 7, 9, 10, 11

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

1. REPORT DATE *(DD–MM–YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From — To)*
19–03–2020	Master's Thesis	Sept 2018 — Mar 2020

4. TITLE AND SUBTITLE

Maximizing Accuracy Through Stereo Vision Camera Positioning for Automated Aerial Refueling

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Kirill, Sarantsev A, Capt

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT-ENG-MS-20-M-059

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Ba T Nguyen
Aerospace Systems Directorate, Air Force Research Laboratory
2210 8TH ST
WPAFB OH 45433-7765
COMM (937) 938-7765
Email: ba.nguyen@us.af.mil

10. SPONSOR/MONITOR'S ACRONYM(S)

AFRL/RQ

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A:
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

This work is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

14. ABSTRACT

Aerial refueling is a key component of the U.S. Air Force strategic arsenal. AAR looks to improve the refueling process by creating a more effective system. This thesis focuses on an AAR scenario where stereo cameras are used on the tanker to direct a boom into a port on the receiver aircraft. The research seeks the optimal design for the camera system to achieve the most accurate navigational estimates. The testing process consist of simulations with both un-obscured scenarios and cases where the boom causes significant occlusion. The results define the improvements in position and orientation estimation of camera positioning from the consolidated simulation data. Conclusions drawn from this research will propose and help provide recommendations for future Air Force acquisition and development of aerial refueling systems.

15. SUBJECT TERMS

Aerial Refueling, Automated Aerial Refueling, Stereo Vision, Computer Vision

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Captain Kirill A. Sarantsev, AFIT/ENG
U	U	U	UU	77	19b. TELEPHONE NUMBER *(include area code)* (937) 255-3636 x4395; kirill.sarantsev@afit.edu