

Air Force Institute of Technology

AFIT Scholar

Faculty Publications

Summer 2018

Improvements for Vision-based Navigation of Small, Fixed-wing Unmanned Aerial Vehicles

Robert C. Leishman

Air Force Institute of Technology

Jeremy Gray

John F. Raquet

Air Force Institute of Technology

Adam Rutkowski

Air Force Research Laboratory

Follow this and additional works at: <https://scholar.afit.edu/facpub>



Part of the [Controls and Control Theory Commons](#), and the [Signal Processing Commons](#)

Recommended Citation

Leishman, Robert C.; Gray, Jeremy; Raquet, John F.; and Rutkowski, Adam, "Improvements for Vision-based Navigation of Small, Fixed-wing Unmanned Aerial Vehicles" (2018). *Faculty Publications*. 675. <https://scholar.afit.edu/facpub/675>

This Article is brought to you for free and open access by AFIT Scholar. It has been accepted for inclusion in Faculty Publications by an authorized administrator of AFIT Scholar. For more information, please contact AFIT.ENWL.Repository@us.af.mil.

Improvements for Vision-based Navigation of Small, Fixed-wing Unmanned Aerial Vehicles

Robert C. Leishman¹, Jeremy Gray¹, John Raquet¹, and Adam Rutkowski²

Abstract—Investigating alternative navigation approaches for use when GPS signals are unavailable is an active area of research across the globe. In this paper we focus on the navigation of small, fixed-wing unmanned aerial vehicles (UAVs) that employ vision-based approaches combined with other measurements as a replacement for GPS. We demonstrate with flight test data that vehicle attitude information, derived from cheap, MEMS-based IMUs is sufficient to improve two different types of vision processing algorithms. Secondly, we show analytically and with flight test data that range measurements to one other vehicle with global pose is sufficient to constrain the global drift of a visual inertial odometry-based navigation solution. Further, we demonstrate that such ranging information is not needed at a fast rate; that bounding can occur using data as infrequent as 0.01 Hz.

I. INTRODUCTION

Investigating alternative navigation approaches for use when GPS signals are unavailable is an active area of research across the globe. When multiple sources of precision navigation and timing (PNT) information are available and can be fused together, the resulting navigation solution is more accurate and more robust. There are a variety of areas where alternative navigation solutions are of interest and a variety of phenomenologies upon which alternative navigation estimates can be developed [11], [5], [22]. In this paper we focus on the navigation of small unmanned aerial vehicles (UAVs) that employ vision-based approaches combined with other measurements.

There is a vast array of current and active research on vision navigation approaches for small UAVs, for example [9], [19], [15], [2], [16], [4], [20], [7], [10]. Many of these approaches provide results competitive with or better than estimates computed using GPS signals, over short (hundreds of meters) distances.

The work described in this paper, however, takes a slightly different tack. One current research aim for the ANT Center, of which this paper is a part, is to demonstrate GPS-denied or degraded navigation of small, fixed-wing UAV platforms over longer (tens of kilometers) trajectories. It is clear that a small UAV with only the core onboard sensors: MEMS IMU (accelerometer, gyroscope, and magnetometer) and pressure (barometric altitude and airspeed) cannot maintain

an adequate navigation solution for any sufficient duration; and that something like vision is essential. There are two key questions addressed by this research to further the capabilities of a designed system. The first is, can MEMS-grade IMU data, or rather, attitude estimates derived from that data, be used to benefit vision algorithms? And the second, what additional information is needed to bound the drift in a visual inertial odometry (VIO) solution for a fixed wing UAV? It is known that if there is no loop closure-type data, essentially creating a SLAM algorithm, the drift of a VIO system grows without bounds.

The research discussed in this article demonstrates and details synergies that can be utilized in the development of small UAVs that can navigate with vision sensors without the direct availability of GPS information. It is clear that a small UAV with only the core onboard sensors: MEMS IMU (accelerometer, gyroscope, and magnetometer) and pressure (barometric and airspeed) cannot maintain a good navigation solution for any sufficient duration. As has been discussed, vision navigation methods have demonstrated good utility in providing an alternative to GPS. The contributions made in this article are, first, that MEMS IMU sensors can be utilized to improve the navigation results of at least two different types of vision algorithms; second, it is shown that a range to only one other vehicle with known pose is sufficient to constrain the drift of a VO-based navigation solution.

The paper is organized as follows. First relevant background material is discussed, including the GIVINS project, the two types of vision navigation algorithms, details regarding the autopilots and associated sensors used in the flight tests, a discussion of why attitude information can provide a benefit to vision algorithms and finally an explanation of the ranging concept used in the paper. Section III discusses the filter framework used to combine IMU, barometric, VO and ranging data. Section IV derives the analytic non-linear observability of the filter framework in Section III. A summary of the relevant details regarding the flight test data used in the results is in Section V. Results for attitude aiding of vision algorithms are found in Section VI. Results demonstrating bounding the drift of a VO-based solution using the range to only one other vehicle are in Section VII. Finally, conclusions are offered in Section VIII.

II. BACKGROUND

The material in this section provides some essential background information to support the development and results below. Several topics are covered as part of the background: a major source for the funding of this work, the two types of

*This work was supported by AFRL/RWWI under the GIVINS Program, an international collaboration between AFRL and the Singaporean Ministry of Defense

¹R.C. Leishman, J. Gray, and J. Raquet are with the Autonomy and Navigation Technology (ANT) Center, Air Force Institute of Technology, Wright-Patterson AFB, OH 45433 USA first.last at afit.edu

²A. Rutkowski is with the Air Force Research Laboratory, Munition Directorate, Eglin AFB, FL, 32542 USA adam.rutkowski at us.af.mil

vision navigation methods utilized, commercial small-UAV autopilots and sensors, why attitude aiding for visual navigation algorithms is beneficial, and finally details regarding the ranging scenario and sensors utilized.

A. GIVINS Project

The objective of the GPS, Inertial, Vision Integrated Navigation System (GIVINS) program is to develop an integrated navigation system using a vision sensor, a standard Global Positioning System (GPS) receiver, and a consumer grade Inertial Measurement Unit (IMU). The GIVINS program is a 4 year international project agreement (PA #AF-PA-13-0001) with the Singapore Ministry of Defence (SG MINDEF) that began 30 Oct 2013.

The original plan for the GIVINS program was to implement the system on a ground vehicle and test in an urban environment. Also, the original plan was for the US and SG to develop and use a common set of hardware. However, due to difficulties associated with hardware being commonly available in both countries, it was decided to develop separate hardware solutions. Also, rather than test in an active urban environment, it was decided to shift the focus of the project to aerial applications. SG chose to develop a solution for a quad rotor, while the US chose to work with a fixed wing platform.

The original technical approach was to use zero velocity updates for vision-aided inertial navigation. A zero velocity update, as the name suggests, resets the velocity estimate of the navigation system to zero when it is known that the system is at rest. Zero velocity updates have been effective at reducing the drift of an Inertial Navigation System, particularly for land based systems. Although it is much more difficult to reach zero velocity (i.e. stop) when suspended in the air (and impossible for a fixed wing aircraft that can't hang on the prop), the direction of vehicle motion can be perceived by an onboard visual system. It can then be deduced that there is no component of velocity perpendicular to this direction. In effect, a zero velocity update can be performed in two directions. It was expected that this approach will have similar performance to existing solutions. Furthermore, if GPS is available, then it should be used. If at any time there are not enough GPS satellites visible to compute a full GPS position solution (which commonly happens in urban environments), then it is still possible to use pseudorange measurements to the visible satellites to aid the navigation solution.

The first attempt at using vision-based zero velocity updates was performed by Png (2014). He first developed a simulation environment for GPS pseudoranges, IMU measurements (ΔV and $\Delta\theta$), and visual features. Rather than developing a full scene generation capability, the approach used was to perform feature tracking on images from real datasets (scenarios 2, 6, and 7 of the ASPN dataset). These features were then mapped to 3D feature coordinates to create a 3D map of feature locations. In this way, a more realistic point cloud that was derived from real data was achieved rather than using a pseudo-randomly generated

point cloud of features. This enabled trade studies for varying sensor parameters to simulate different quality sensors and camera look angles. Camera quality was simulated by adding noise (with a standard deviation of a certain number of pixels) to the pixel position of each feature.

For a given normalized translation vector t in East-North-Up coordinates, as determined from visual odometry, the two directions of zero velocity are determined from the full Singular Value Decomposition of t (Png 2014). This produces a matrix U where the first column is t and the last two columns are perpendicular to t . The matrix Σ will be $[1\ 0\ 0]^T$ and V will be 1. Similarly, a full QR factorization could be used to compute the perpendicular directions.

IMU measurements were simulated by inverting the PVA data from the ASPN data sets, then adding FOGM noise according to the desired sensor quality to be simulated.

Rohde (2015) was the first to apply the visual zero velocity update technique to real world data.

Given the success of using this visual and inertial sensor measurements with GPS pseudoranges, the GPS pseudoranges were replaced with actual range measurements to a vehicle in a high flyer / low flyer concept.

B. Absolute Positioning Approach

Two vision navigation approaches are utilized in demonstrating that attitude estimates derived from MEMS IMU can add value to vision processing algorithms. The first is an absolute positioning (AP) algorithm that provides a global position and attitude measurement, illustrated in Figure 1, by comparing the current camera image and a known map of images, e.g. Google Earth. The algorithm used in this research was developed in the ANT Center and is described in detail in [17], [22]. We include here a brief summary of the key points of the algorithm, for reference.

- The algorithm requires a geo-reference database (known map) as an input. This must be prepared beforehand, and is critical to the functionality of the algorithm, see [22] for extensive details.
- The current image and current estimate of global position and orientation (pose) of the vehicle are sent to the algorithm.
- SIFT features are found on the current image.
- A database image tile is retrieved from the map based on the global pose estimate. This tile contains georeferenced SIFT features.
- A brute-force matcher finds the best matching features between the current image and the image tile.
- At this point there are two possible approaches, depending on the selection of an attitude aided solution or a non-attitude aided solution.
 - If the attitude aided solution is selected, the attitude information from the autopilot is used directly to find the translation estimate between the image tile and the current image. This is done using a 3 Degree of Freedom (DoF) perspective n point (PNP) algorithm.

- If a non-attitude aided solution is requested then the full pose information is recovered from the 6 DoF PNP algorithm.

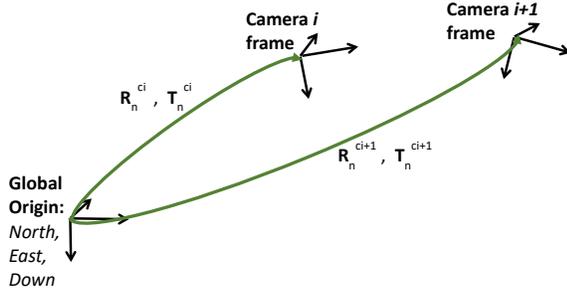


Fig. 1. Global measurement and associated coordinate systems.

C. Visual Odometry Approach

The second vision algorithm used is a visual odometry (VO) algorithm that provides a relative, or incremental, measurement between the previous camera frame and the current camera frame, see Figure 2 below. Consequently, it is often utilized as a velocity update when the position information in a filter state vector is expressed in a global coordinate frame. This is an important distinction between the VO and the Absolute Positioning Approach described above, as the Absolute Positioning approach provides global measurements. This VO algorithm is used in both objectives for this research outlined above.

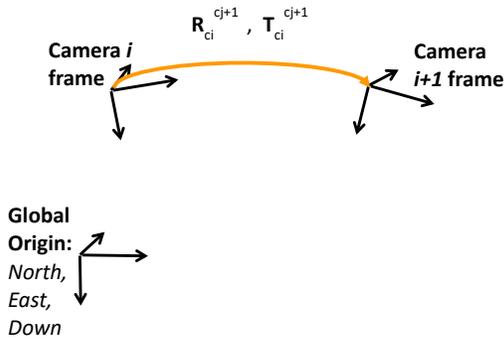


Fig. 2. Relative measurement and associated coordinate systems.

The VO approach used in this work is described in detail in [6]. The algorithm uses detected features, specifically the AKAZE [1] variety, to compute the relative transformation between two images. The following is a quick summary of the algorithm to understand the principle steps.

- The current and just-previous image are sent to the algorithm, along with camera calibration, the body to Navigation rotation matrices for each image and other parameters, like the current barometric height.
- AKAZE features and descriptors are detected on both images. A brute force descriptor matcher, using the Hamming distance, finds the closest matching features between the two images.

- The RANSAC algorithm in the OpenCV "findEssentialMat" function robustly computes an essential matrix and identifies valid matches (inliers) that support the computation.
- At this point, the choice is made on whether to utilize the rotation matrix found as part of the essential matrix (non-attitude aided) or to use the rotation matrix provided by the autopilot (attitude aided). If the non-attitude aided solution is chosen, the algorithm chooses the appropriate rotation matrix extracted from the essential matrix.
- Each of the matched, inlier features are expressed in the navigation frame using the selected rotation matrix. Inherent in this step is the addition of yaw information into the measurement. The source of the yaw information is a magnetometer.
- A depth estimate is computed for each feature, using the height information from the barometer.
- All the inlier measurements, from each image, are averaged together and then the means are subtracted to find the translation vector.
- The computed/selected rotation and computed translation are returned as the six DoF solution by the algorithm.

D. Autopilots

The growth of available low-cost commercial autopilots over the past decade has provided many new avenues for performing previously difficult or even impossible operations. These operations are spread across both commercial and defense domains. Some examples of these operations include infrastructure inspection, reconnaissance, crop inspection, and many others. Currently, the cost of these hobby-grade autopilots range anywhere from \$300 to as inexpensive as \$10 with varying capabilities.

In this research, Pixhawk autopilots were used as the flight controller for each fixed wing as well as a sensor data source. These autopilots contain redundant accelerometers, gyros, barometers, and a single magnetometer, and utilize an external GPS, magnetometer, and airspeed sensor. The sensors contained in the Pixhawk, specifically the accelerometers and gyros, are of a low consumer grade (cost about \$6 for the IMU boards). During the flight tests for constraining the drift on the VIO solution, two Pixhawk autopilots were utilized. The first was connected to GPS and the data from it is used as truth. The second autopilot was purposely not connected to GPS; the data from this autopilot is used for evaluating the filter from a GPS-denied standpoint.

E. Why Attitude Aiding is Helpful

When using vision data for navigation, there is often a very high correlation between attitude and position. Consider, for example, the three pictures shown in Figure 3. The left picture is the original picture, taken approximately 1m from the bookshelf. The center picture was obtained by keeping the camera in the same position, but rotating it approximately 1.7° to the left, and the right picture was obtained by moving

the camera 3cm to the left (but no rotation). It is clear that the center and the right pictures are very similar, which indicates that any vision algorithm which would seek to measure both rotation and translation will have difficulty distinguishing between the two.

In this kind of situation, we are taking advantage of the knowledge of attitude obtained from a Pixhawk autopilot that is not connected to GPS. As a result, the algorithm only needs to solve for position (or delta-position in the case of visual odometry), given a known attitude (or delta-attitude), which should significantly increase the performance of the position/delta-position estimation algorithm. The results in Section VI demonstrate that low-commercial-grade autopilot attitude information can be used to improve vision navigation solutions.

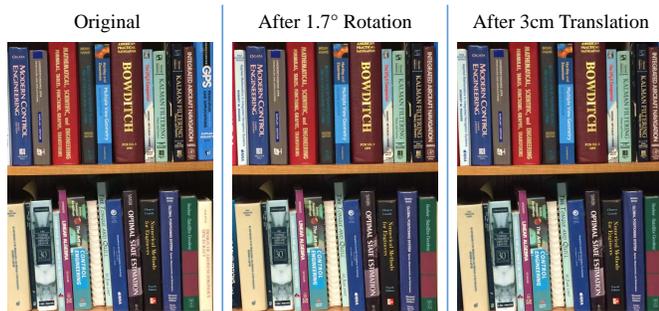


Fig. 3. Demonstration showing correlation between translation and rotation, as demonstrated by similarity between center (rotated) and right (translated) images.

F. Ranging Measurements

The second objective in this study, as outlined above, is to investigate the amount of information that is required to bound the drift of a VIO solution for a UAV. The proposed method for bounding the drift is to utilize range data obtained through a ranging radio to another aircraft/UAV that has global pose information, likely through GPS.

The concept is that one aircraft, the one receiving GPS, is able to fly higher; above the range of the jamming, out of the urban canyon, or out of whatever condition that is limiting the access to GPS signals. The lower vehicle is navigating using IMU and VO information (VIO approach), without GPS. The upper vehicle provides, through the same ranging radio, its position and range to the lower UAV. The concept is illustrated in Figure 4.

Actual flight test data is used to demonstrate the concepts proposed in the paper in addition to an analytic observability analysis in Section IV. Range radios were provided by EN-SCO Corp. through a cooperative research and development agreement.

The research builds on initial work in [8]. All the results in Section VII pertain to estimating the position of the lower UAV, following the filter framework established in Section III immediately below.



Fig. 4. An illustration of the ranging scenario used in the flight test. The upper plane has access to GPS signals, the lower plane does not. It uses autopilot sensors plus visual odometry to navigate, and then it fuses in the range to the upper aircraft. We use a common Navigation frame for both vehicles. We assume that GPS is available at the beginning of the lower planes flight to establish the common frame.

III. VO, IMU, PRESSURE AND RANGE DATA FUSION

We first introduce a couple of the coordinate frames utilized throughout the paper in preparation of the filter derivation, illustrated in Figure 5. The first coordinate frame is the body-fixed frame of the lower UAV, which is fixed to the center of mass of the vehicle with the \hat{i}_b axis out the nose, the \hat{j}_b axis out the right wing, and the \hat{k}_b axis out of the belly. The second is the navigation frame, which is a local frame with the \hat{i}_N axis aligned with magnetic North, the \hat{j}_N axis aligned with East, and the \hat{k}_N axis pointing toward the center of the Earth [21]. The navigation frame sits at a point on the Earth's surface near where the flight occurs and for short, low altitude UAV flights is assumed to be inertial [3]. We utilize the convention described in [3], Chapter 2, including the 3-2-1 Euler angle sequence and the wind triangle.

The extended Kalman filter used to demonstrate bounding of a VIO solution, which fuses IMU, pressure, VO and range measurements, largely follows the development in Chapter 8 of [3]. We have extended the filter by including states for altitude and vertical velocity. The states for the filter are the north position, east position, ground speed, course angle, wind speed in north direction, wind speed in east direction,

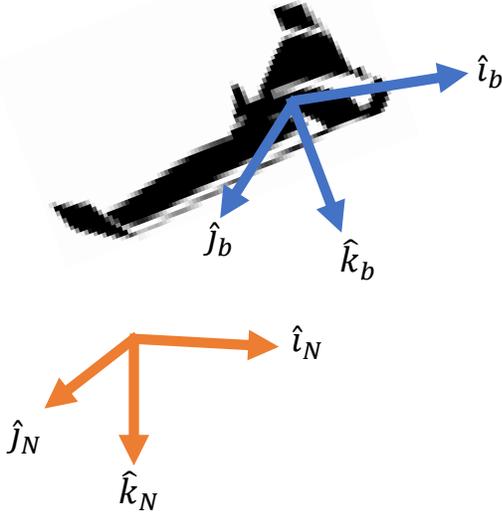


Fig. 5. Coordinate frames used in the filtering derivations. Body-fixed frame is in blue, the inertial frame is in orange.

yaw, altitude, and vertical velocity:

$$\mathbf{x} = [p_n \ p_e \ V_g \ \chi \ w_n \ w_e \ \psi \ p_a \ \dot{p}_a]^T. \quad (1)$$

The north, east and altitude positions are with respect to the navigation frame. Altitude is simply the negative of the position in the down direction. The relationship between ψ and χ is explained using the wind triangle [3]. ψ is the angle between the \hat{i}_N axis and the \hat{i}_b axis in the horizontal plane. χ is the angle between the \hat{i}_b axis and the ground speed vector, in the horizontal plane.

The state \mathbf{x} advances according to the nonlinear model $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u})$, where the input \mathbf{u} is composed of the airspeed, pitch rate, yaw rate, roll angle, pitch angle:

$$\mathbf{u} = [V_a \ q \ r \ \phi \ \theta]^T; \quad (2)$$

$$\mathbf{f}(\mathbf{x}, \mathbf{u}) = \begin{bmatrix} V_g \cos(\chi) \\ V_g \sin(\chi) \\ \frac{(V_a \cos(\psi) + w_n)(-V_a \dot{\psi} \sin(\psi)) + (V_a \sin(\psi) + w_e)(V_a \dot{\psi} \cos(\psi))}{V_g} \\ \frac{g}{V_g} \tan(\phi) \cos(\chi - \psi) \\ 0 \\ 0 \\ q \frac{\sin(\phi)}{\cos(\theta)} + r \frac{\cos(\phi)}{\cos(\theta)} \\ \dot{p}_a \\ 0 \end{bmatrix} \quad (3)$$

The inputs \mathbf{u} for this work were recorded from the Pixhawk autopilot that was purposely **not** connected to a GPS receiver. More details regarding the hardware configuration are provided in Section V below.

A. Measurement Updates

There are several measurement updates used in the developed EKF, they include a wind pseudo-measurement, a velocity update from the visual odometry, an altitude update

from the baro altimeter, and finally the range measurement. The measurement models are described briefly below.

1) *Wind Pseudo-Measurement*: The wind pseudo-measurement update is the same proposed in [3], where the geometry in the wind triangle is used to equate the wind, ground and air speeds using the angles χ and ψ . This measurement is more of a constraint, as the "measurement" should always equal zero.

$$\mathbf{h}_w(\mathbf{x}, \mathbf{u}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} V_a \cos(\psi) + w_n - V_g \cos(\chi) \\ V_a \sin(\psi) + w_e - V_g \sin(\chi) \end{bmatrix} \quad (4)$$

2) *VO Velocity Update*: The VO algorithm [6] utilized for this paper is briefly described above. An important point to recall is the yaw information, from a magnetometer, has been infused into the VO measurements. Consequently, the measurement provides both a ground speed and course angle update for the filter. A speed update is used since the position information derived from the VO algorithm is relative to the image before it and does not provide position information relative to the navigation frame [13]. However, the VO algorithm is provided the best estimate of the rotation between the current body frame and the navigation frame, allowing it to express the velocity in the navigation frame. We can then update both the course angle χ and the ground speed V_g elements of the state with the output of the VO velocity $[V_n, V_e, V_a]$ according to

$$\mathbf{z}_{vo}(\mathbf{x}) = \begin{bmatrix} \tan^{-1}\left(\frac{V_e}{V_n}\right) \\ \sqrt{V_n^2 + V_e^2 + V_a^2} \end{bmatrix}. \quad (5)$$

Consequently, the measurement function is trivially,

$$\mathbf{h}_{vo}(\mathbf{x}) = \begin{bmatrix} \chi \\ V_g \end{bmatrix}. \quad (6)$$

3) *Altitude Update*: The altitude update is quite simple, the information from the barometer in the autopilot directly updates the altitude state

$$h_a(\mathbf{x}) = p_a. \quad (7)$$

4) *Ranging Update*: The measurement of range between the lower vehicle and the upper vehicle includes the actual range in meters between the two vehicles, as well as the GPS coordinates in latitude, longitude, and height of the upper vehicle. We utilize a single GPS measurement of the lower vehicle, taken at the beginning of the flight, to define a common navigation frame of reference for both vehicles, thereafter we only use the lower vehicle GPS measurements for truth data. We do not address the issue of obtaining a common navigation reference frame between two vehicles for the case when GPS is never available for the lower vehicle; in such an instance, an estimate of the upper vehicle's position would likely need to be added to the filter state and a different coordinate system utilized. The measurement function used is

$$h_r(\mathbf{x}, \mathbf{u}_u) = \sqrt{(u_n - p_n)^2 + (u_e - p_e)^2 + (u_d + p_a)^2}, \quad (8)$$

where $\mathbf{u}_u = [u_n, u_e, u_d]^T$ are the coordinates of the upper vehicle in the now-common navigation frame. Recall that u_d and p_a point in opposite directions.

B. Filter Implementation

The above methods were implemented in the Scorpion Estimation Framework, created by the Autonomy and Navigation Technology Center (ANT) at the Air Force Institute of Technology. The implementation used an Extended Kalman Filter and the data was run post-process through the filter. All the results pertain to estimating the states of a lower UAV using the filter described.

IV. OBSERVABILITY ANALYSIS: VO PLUS RANGING

The observability rank condition for nonlinear systems outlined in Chapter 7 of [23], which is also summarized quite well in [12], pertains only to the case where the nonlinear equations can be expressed *linearly in the control*:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \sum_{i=0}^m u_i(t) \mathbf{g}[\mathbf{x}(t)] \quad (9)$$

$$y_j = h_j(\mathbf{x}), \text{ for } j = 1 : l \quad (10)$$

Upon examining the models defined in Section III above, it is apparent that the filter model utilized does not exactly fit this constraint. The process to find the observability rank condition for such a system becomes much more arduous, involving many more uses of the chain rule. We have found that by neglecting second-order and higher derivatives with respect to \mathbf{u} , whose magnitudes are usually small, the same result for the observability index may be achieved, namely

$$\mathcal{O} \approx \begin{bmatrix} \frac{\partial \mathbf{L}_f^0 h_j(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}} \\ \frac{\partial \mathbf{L}_f^1 h_j(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}} \\ \vdots \\ \frac{\partial \mathbf{L}_f^i h_j(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}} \end{bmatrix}, \quad (11)$$

with

$$\begin{aligned} \mathbf{L}_f^0 h_j(\mathbf{x}, \mathbf{u}) &= h_j(\mathbf{x}, \mathbf{u}) \\ \mathbf{L}_f^{i+1} h_j(\mathbf{x}, \mathbf{u}) &= \frac{\partial \mathbf{L}_f^i h_j(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x}, \mathbf{u}). \end{aligned}$$

Where the notation $\mathbf{L}_f h_j(\mathbf{x}, \mathbf{u})$ denotes the Lie derivative of the function $h_j(\mathbf{x}, \mathbf{u})$ with respect to the vector field $\mathbf{f}(\mathbf{x}, \mathbf{u})$, which results in a scalar; and $i \in n - 1$ and $j \in l$. A Lie derivative describes the information within the function $h_j(\mathbf{x}, \mathbf{u})$ that is contained along the vector field $\mathbf{f}(\mathbf{x}, \mathbf{u})$. Then gradients of the Lie derivatives are used to assemble the observability index. The $l = 6$ measurement equations from Eqs.4, 6, 7, 8 form the set of $h_j(\mathbf{x}, \mathbf{u})$ for the system. Note that assuming partial derivatives with respect to \mathbf{u} are zero is actually a more conservative estimate than computing and including those terms.

A. Analytic Observability

The first step in assembling \mathcal{O} in (11) is to compute the gradient of the nonlinear measurement functions, (4), (6), (7), (8), with respect to the state \mathbf{x} . For reference, h_1 and h_2 are from (4), h_3 and h_4 from (6), h_5 from (7) and h_6 from (8) respectively. This is found in (12) below.

(12) provides the first 6 rows to (11); note that the boldfaced $\mathbf{h}(\mathbf{x}, \mathbf{u})$ implies that the partials for each $h_j(\mathbf{x}, \mathbf{u})$ have been stacked to form a matrix. In (12) there is only one state, p_a , that does not receive any information from the measurements directly. However, the rank of (12) is only six, thus we require at least three more linearly independent equations in \mathcal{O} to achieve local observability of the system.

The second row of (11) requests the partial of $\mathbf{L}_f^1 h_j(\mathbf{x}, \mathbf{u})$. Using the chain rule, this becomes

$$\frac{\partial \mathbf{L}_f^1 h_j(\mathbf{x})}{\partial \mathbf{x}} = \left(\frac{\partial^2 h_j(\mathbf{x})}{\partial \mathbf{x}^2} \mathbf{f}(\mathbf{x}) \right)^T + \frac{\partial h_j(\mathbf{x})}{\partial \mathbf{x}} \frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}} \quad (13)$$

where $\frac{\partial^2 h_j(\mathbf{x})}{\partial \mathbf{x}^2}$ is the Hessian of $h_j(\mathbf{x})$. It is useful to look at both of the terms in (13) to see what each part contributes to the observability.

The first half of (13) contains second-order derivatives and is quite sparse. Only measurement functions (4) and (8) contain non-zero terms. The second order gradients of these three measurements, multiplied by $\mathbf{f}(\mathbf{x})^T$, results in three column vectors (14), (15), and (16) respectively.

Equations (14), (15) and (16) are transposed and stacked, along with three zero row-vectors for measurement functions (6) and (7) to form the matrix representation of $\left(\frac{\partial^2 \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}^2} \mathbf{f}(\mathbf{x}) \right)^T$, which is not shown due to size restrictions. Note the first two terms in (16) are non-zero. This information, though second-order, enables the observability of the horizontal position $[p_n, p_e]$. No second-order information contributes to the observability of p_n .

The second half of (13) contains combinations of only first-order derivatives and is expressed in (17).

(17) below clearly adds observability information to p_a , in the final column. However, there is not any additional information added for the horizontal position $[p_n, p_e]$, marked by lack of non-zero terms in the first two columns.

Analytically, by combining the two parts of (13) with (12) we have a 12 by 9 matrix that is of rank 9. Meaning, we do not need any higher-order Lie derivatives to form (11). The reason that the observability matrix is of full rank with range to only one known landmark is that yaw information is provided to the filter through the visual odometry method we have chosen to use. Without the yaw information provided, the observability grammian would be of rank 8 and we would need at least one other known landmark to achieve full observability. Interesting to note, however, is that if we were to follow a more conservative estimate of the observability and ignore any second-order and higher partial derivatives in (11), we would be left with multiples of (17), as shown in [23], with $\frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}} = C$ and $\frac{\partial \mathbf{f}}{\partial \mathbf{x}} = A$. (11) then can be

$$\frac{\partial L_f^0 \mathbf{h}(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}} = \begin{bmatrix} 0 & 0 & -\cos(\chi) & V_g \sin(\chi) & 1 & 0 & -V_a \sin(\psi) & 0 & 0 \\ 0 & 0 & -\sin(\chi) & -V_g \cos(\chi) & 1 & 0 & V_a \cos(\psi) & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \frac{p_n - u_n}{\sqrt{\sigma_1}} & \frac{p_e - u_e}{\sqrt{\sigma_1}} & 0 & 0 & 0 & 0 & 0 & \frac{p_a + u_d}{\sqrt{\sigma_1}} & 0 \end{bmatrix}, \quad (12)$$

where

$$\sigma_1 = (p_e - u_e)^2 + (p_n - u_n)^2 + (p_a + u_d)^2.$$

$$\frac{\partial^2 h_1(\mathbf{x})}{\partial \mathbf{x}^2} \mathbf{f}(\mathbf{x}) = \begin{bmatrix} 0 \\ 0 \\ \frac{g \sin(\chi) \tan(\varphi) \cos(\chi - \psi)}{V_g} \\ g \cos(\chi) \tan(\varphi) \cos(\chi - \psi) + \frac{V_a \sin(\chi) \dot{\psi} (w_e \cos(\psi) - w_n \sin(\psi))}{V_g} \\ 0 \\ 0 \\ -V_a \cos(\psi) \dot{\psi} \\ 0 \\ 0 \end{bmatrix} \quad (14)$$

$$\frac{\partial^2 h_2(\mathbf{x})}{\partial \mathbf{x}^2} \mathbf{f}(\mathbf{x}) = \begin{bmatrix} 0 \\ 0 \\ -\frac{g \cos(\chi) \tan(\varphi) \cos(\chi - \psi)}{V_g} \\ g \sin(\chi) \tan(\varphi) \cos(\chi - \psi) - \frac{V_a \cos(\chi) \sigma_2 (w_e \cos(\psi) - w_n \sin(\psi))}{V_g} \\ 0 \\ 0 \\ -V_a \sin(\psi) \sigma_2 \\ 0 \\ 0 \end{bmatrix} \quad (15)$$

$$\frac{\partial^2 h_6(\mathbf{x})}{\partial \mathbf{x}^2} \mathbf{f}(\mathbf{x}) = \begin{bmatrix} \frac{V_g \cos(\chi)}{2\sigma_1^{3/2}} (\Delta_n + 2\sigma_1) + \frac{V_g \sin(\chi)}{\sigma_1^{3/2}} (\Delta_n + \Delta_e) + \frac{\dot{p}_a}{\sigma_1^{3/2}} (\Delta_n - \Delta_d) \\ \frac{V_g \sin(\chi)}{2\sigma_1^{3/2}} (\Delta_e + 2\sigma_1) + \frac{V_g \cos(\chi)}{\sigma_1^{3/2}} (\Delta_n + \Delta_e) + \frac{\dot{p}_a}{\sigma_1^{3/2}} (\Delta_e - \Delta_d) \\ 0 \\ 0 \\ 0 \\ 0 \\ \frac{\dot{p}_a}{2\sigma_1^{3/2}} (2\sigma_1 - \Delta_d) + \frac{V_g \cos(\chi)}{\sigma_1^{3/2}} (\Delta_n - \Delta_d) + \frac{V_g \sin(\chi)}{\sigma_1^{3/2}} (\Delta_e - \Delta_d) \\ 0 \end{bmatrix} \quad (16)$$

where,

$$\Delta_n = u_n - p_n$$

$$\Delta_e = u_e - p_e.$$

$$\Delta_d = u_d + p_a.$$

expressed just as a linear system observability index

$$\mathcal{O}_l \approx \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}. \quad (18)$$

Following this assumption, (11) would never be full rank and the system would not be observable, even though it should be. The only information available to the filter for horizontal position would be a nonlinear combination of p_n and p_e and

$$\frac{\partial h_j(\mathbf{x})}{\partial \mathbf{x}} \frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}} = \quad (17)$$

$$\begin{bmatrix} 0 & 0 & \frac{V_a \cos(\chi) \sigma_1 \sigma_2}{V_g^2} - \frac{g \sin(\chi) \tan(\varphi) \cos(\chi - \psi)}{V_g} & -\sigma_4 & \frac{V_a \cos(\chi) \sin(\psi) \dot{\psi}}{V_g} - \frac{V_a \cos(\chi) \cos(\psi) \dot{\psi}}{V_g} & \sigma_4 + \frac{V_a \cos(\chi) \dot{\psi} \sigma_3}{V_g} & 0 & 0 \\ 0 & 0 & \frac{g \cos(\chi) \tan(\varphi) \cos(\chi - \psi)}{V_g} + \frac{V_a \sin(\chi) \dot{\psi} \sigma_2}{V_g^2} & \sigma_5 & \frac{V_a \sin(\chi) \sin(\psi) \dot{\psi}}{V_g} - \frac{V_a \cos(\psi) \sin(\chi) \dot{\psi}}{V_g} & \frac{V_a \sin(\chi) \dot{\psi} \sigma_3}{V_g} - \sigma_5 & 0 & 0 \\ 0 & 0 & -\frac{g \tan(\varphi) \cos(\chi - \psi)}{V_g^2} & -\sigma_6 & 0 & \sigma_6 & 0 & 0 \\ 0 & 0 & -\frac{V_a \dot{\psi} \sigma_2}{V_g^2} & 0 & -\frac{V_a \sin(\psi) \dot{\psi}}{V_g} & \frac{V_a \cos(\psi) \dot{\psi}}{V_g} & -\frac{V_a \dot{\psi} \sigma_3}{V_g} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{-\Delta_n \cos(\chi)}{\sqrt{\sigma_1}} - \frac{\Delta_e \sin(\chi)}{\sqrt{\sigma_1}} & \frac{-V_g \Delta_e \cos(\chi)}{\sqrt{\sigma_1}} + \frac{V_g \Delta_n \sin(\chi)}{\sqrt{\sigma_1}} & 0 & 0 & 0 & 0 & \frac{2p_a + 2u_d}{\sqrt{\sigma_1}} \end{bmatrix},$$

where

$$\begin{aligned} \sigma_2 &= w_e \cos(\psi) - w_n \sin(\psi) \\ \sigma_3 &= w_n \cos(\psi) + w_e \sin(\psi) \\ \sigma_4 &= g \sin(\chi - \psi) \sin(\chi) \tan(\varphi) \\ \sigma_5 &= g \sin(\chi - \psi) \cos(\chi) \tan(\varphi) \\ \sigma_6 &= \frac{g \sin(\chi - \psi) \tan(\varphi)}{V_g} \end{aligned}$$

it would not be separable. In Section VII we analyze the rank and condition of the observability index (11) with the flight data for some additional insights in the observability question for this problem.

B. Average Normalized Estimation Error Squared

Another metric that proves useful in analyzing estimators is the Average Normalized Estimation Error Squared (ANEES) [14]. The ANEES metric shows insight into whether an estimator is credible by demonstrating and quantifying how much the results are optimistic or pessimistic. ANEES is defined as

$$\bar{\epsilon} = \frac{1}{M} \sum_{i=1}^M (\mathbf{x}_i - \hat{\mathbf{x}}_i)' \mathbf{P}_i^{-1} (\mathbf{x}_i - \hat{\mathbf{x}}_i), \quad (19)$$

where \mathbf{x} is the true state, $\hat{\mathbf{x}}$ is the estimated state, and M is the number of timesteps.

If the ANEES of the filter is significantly greater than n (the length of \mathbf{x}), the filter is optimistic; meaning that the covariance represented by the filter is too small. When the ANEES is significantly under n the filter is pessimistic and the covariance is too large. When an estimator is optimistic, it is overconfident in the estimate of the covariance bounds, which can lead to filter divergence.

V. FLIGHT TEST

The two data sets analyzed in this research were collected from flight tests performed by AFIT in 2015 and 2017. For each test, a fixed-wing small UAS was flown in either an autonomous or semi-autonomous mode while simultaneously collecting sensor measurements and navigation solutions from an autopilot. Both flights were accomplished at Himsel Field at Camp Atterbury, IN, which consists of a concrete parking apron and runway, grass, trees, and various buildings. Though each flight was accomplished during different

times of the year, the visual seasonal differences were not significant, nor did the structures or other permanent visual cues change in the vicinity of the flight profile.

A. 2015 Flight Test: AP

The flight test conducted in October of 2015 [18] utilized a 12 foot Telemaster model aircraft. The vehicle contained a single Pixhawk running Ardupilot firmware, located at the center of gravity (CG). A payload bay approximately one foot aft of the CG contained a Prosilica camera and an Intel Next Unit of Computing (NUC). The NUC collected sensor data and filter solutions from the Pixhawk along with camera images. An autonomous flight profile was flown at approximately 250 meters above ground level (AGL) and resembled a large loop overflying all the available visual features in the area.

This flight data was used to compute the results of the AP vision algorithm, found in Section VI.

B. 2017 Flight Test: VO

The flight test conducted in July 2017 utilized two Sig Rascal 110 model airplanes with 2.8 meter wingspans, driven by electric motors. Data was collected onboard each plane and then brought together through post processing. Details regarding each plane are below.

Data from this flight test was used to compute the results for the VO algorithms, which are used in the results discussed in Sections VI and VII.

1) *Upper Plane*: The upper plane utilized a Pixhawk 2 as the flight controller and was flown in an autonomous loiter with a radius of 100 meters, at 18 meters per second, at an altitude of about 250 meters AGL. The payload consisted of a small UDOO Quad computer and an ENSCO ranging radio. The computer logged data from the autopilot at about 4 Hz and ranging radio at about 20 Hz. The GPS coordinates were

up-converted to 20 Hz using a cubic-spline interpolation for processing in the filter at maximum ranging rates.

2) *Lower Plane*: The lower plane was flown in a manual mode due to complications with the communications equipment right before takeoff. It was flown over an area directly below the upper plane but within a span of about 500 by 500 meters for about twelve minutes, at a speed of about 22 meters per second, within an altitude band of 75 to 150 meters AGL.

The payload consisted of two Pixhawk autopilots (one disconnected from the GPS receiver, the Payload Pixhawk), an ENSCO ranging radio, an Allied Vision Prosilica GT1290 camera, a five-point ethernet switch, and an Intel NUC computer. Data was logged onboard the NUC computer at variable rates, which proved a little challenging during post-processing: 9Hz monochrome camera images, 20Hz ranging data, 140Hz GPS-denied IMU and altimeter, 5Hz GPS-denied airspeed, 4Hz GPS data.

VI. RESULTS: ATTITUDE-AIDING VO

The following results are partitioned by the vision algorithms. The AP algorithm is evaluated using the flight test data from 2015 on both the non-attitude aided and attitude aided solutions. The VO algorithm is evaluated using flight test data from July 2017 for both non-attitude aided and attitude aided solutions. Position data provided by the autopilots with GPS data are used as truth for evaluating the results.

A. AP Results

Figures 6 and 7 demonstrate the error in the north position estimates for the AP approach. Figure 6 demonstrates the error for the non-attitude aided solution. Figure 7 provides the results for the attitude-aided solution. Note the differences in scale of the vertical axis between the two figures. Comparing the two figures provides insight to the error results when an attitude solution is provided to the AP algorithm from the autopilot. The attitude aided results are significantly improved. Not shown are the errors in the East and Down directions, however they follow a similar pattern. A comparison of the overall 3D RMS error values is in the table below, where it is clearly seen that the RMS is reduced by a factor of four by including the attitude information from the autopilot.

TABLE I

COMPARISON OF THE 3D RMS ERRORS OVER THE WHOLE SET OF DATA (TWO FIGHTS) FOR THE ABSOLUTE POSITIONING VISUAL NAVIGATION APPROACH.

Absolute Positioning	RMS (m)
No Attitude Aiding 3D RMS	69.4
Attitude Aiding 3D RMS	16.2

B. VO Results

After the performance gain was discovered for the AP vision processing algorithm by simply including attitude information from the autopilot, we were eager to try the technique on other vision navigation approaches. The VO algorithm developed in-house was a natural fit. We investigated the limits of the performance improvement a bit more with the VO algorithm than the AP algorithm.

Several pertinent facts were learned or confirmed in initial testing that deserve mentioning, which are not discussed in the results below. We investigated using different numbers of features, different rates of attitude information, different RANSAC tuning parameters and different rates of camera imagery (over the same distance, so wider/narrower gaps). We found that the results stayed fairly constant over the number of features detected, up to a point; the performance obviously dropped off rapidly when little-to-no features were detected, as expected. We confirmed in earlier testing that attitude information must either be time-synchronized with the imagery or taken at a much higher rate (five-to-ten times), otherwise the attitude aiding solution is considerably worse than a vision-only solution. In changing the RANSAC parameters, we noticed a considerable performance change, so we found parameters that maximized the performance without making the solutions brittle and maintained the value constant for all our results. The largest performance effect was found by varying the rate of the imagery, results of which are shown below.

1) *8 Hz VO*: Figure 8 shows results in a different format than Figures 6 and 7 above. Here we show the change in position measurements between successive images from the truth in blue "x"s; the attitude aided VO outputs, in red "+"s; and the vision-only VO solution in black "o"s. A perfect result would be for the black and red marks to overlay the blue "x"s. Note that these change-in-position results will not show drift, as the GPS truth data was converted to relative position transformations between the image timesteps.

As is shown in Figure 8, the attitude aiding solution does well at estimating the correct change in position measurement. Some of the non-attitude aided solutions are also correct, but many of the non-attitude aided solutions report a near-zero result. The reason for the binary behavior, where the solution is either close to correct or near zero, is that the vision-only solution seems to have ascribed many of these small position changes to attitude changes, see Section II-E. This is illustrated by the distinct bias in the change in pitch angle found by the non-attitude aided solution, compared to the change in attitude information between images provided by both the Pixhawk Autopilots, in Figure 9.

Table II reports the RMS errors for the North, East, and Down directions between the attitude and non-attitude aided solutions. The attitude aided solution reduces the RMS in the North and East directions by more than 15 times. The VO solutions in general, whether attitude or non-attitude aided did not estimate changes in Down (or height) very well.

2) *2 Hz VO*: Figures 10 and 11 are the result of running the VO at 1/4th the rate. Increasing the change in time

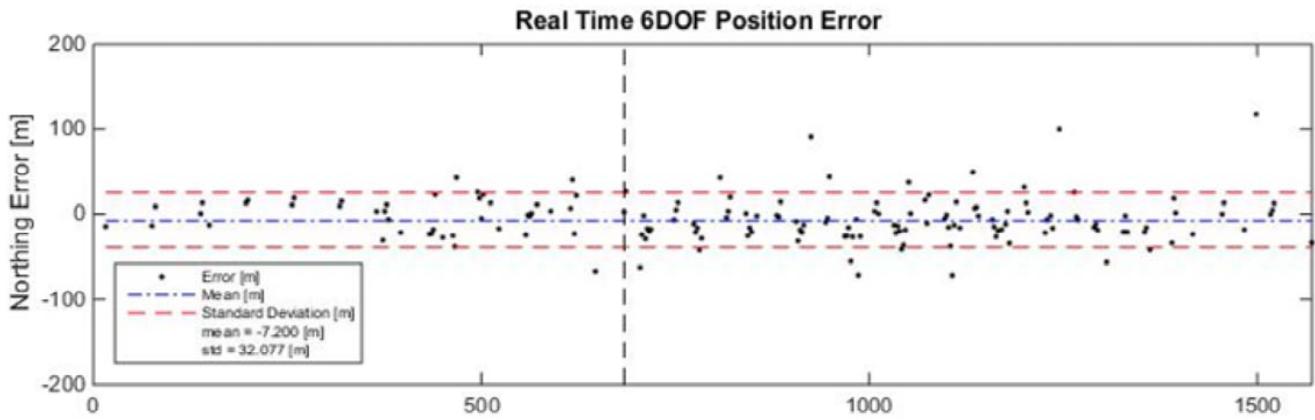


Fig. 6. Error in north position estimates using the Absolute Positioning approach and **no** attitude aiding, with time. East and Down error plots are similar. Error points are in black, one standard deviation bounds marked by the red/brown dashed lines, mean at the blue dot-dashed line. The vertical dashed line marks flight one versus flight two (data was concatenated).

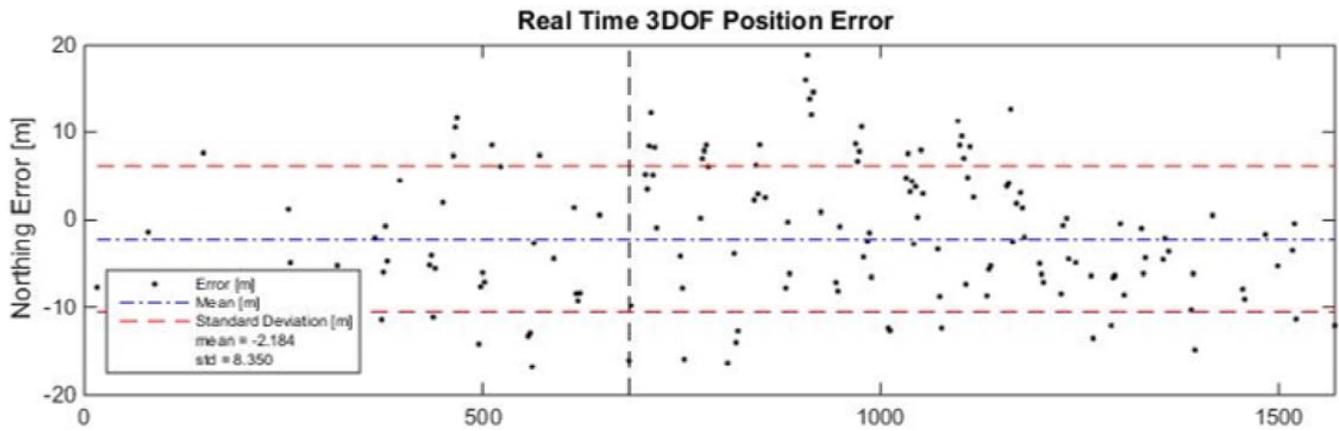


Fig. 7. Error in north position estimates using the Absolute Positioning approach and **attitude aiding**. East and Down error plots are similar. Note the difference in scale of the vertical axis as compared to Figure 6. Error points are in black, one standard deviation bounds marked by the red/brown dashed lines, mean at the blue dot-dashed line. The vertical dashed line marks flight one versus flight two (data was concatenated).

TABLE II

RMS OF THE CHANGES IN POSITION FOR ALL THREE DIMENSIONS, COMPARING THE NON-ATTITUDE AIDED RESULTS TO THE ATTITUDE AIDED RESULTS. THE RMS CHANGES BY MORE THAN A FACTOR OF 15.

Visual Odometry - 8Hz	RMS (m)
No Attitude Aiding North RMS	136.6
No Attitude Aiding East RMS	77.8
No Attitude Aiding Down RMS	0.25
Attitude Aiding North RMS	5.5
Attitude Aiding East RMS	5.3
Attitude Aiding Down RMS	0.25

between images increases the signal to noise of the position measurement, allowing the filter to more easily differentiate between translation and rotation. This is shown in Figure 10 with the Non-Attitude aided solutions holding tighter to the attitude-aided and truth curves. This is also shown by the

80% decrease in the North and East non-attitude aiding RMS. Also notice the solutions in both images at about 135 seconds until the end, where the attitude bias is gone, and all the solutions are close.

TABLE III

RMS OF THE CHANGES IN POSITION FOR ALL THREE DIMENSIONS, COMPARING THE NON-ATTITUDE AIDED RESULTS TO THE ATTITUDE AIDED RESULTS WITH 2HZ CAMERA UPDATES.

Visual Odometry - 2Hz	RMS (m)
No Attitude Aiding North RMS	17.2
No Attitude Aiding East RMS	15.7
No Attitude Aiding Down RMS	1.2
Attitude Aiding North RMS	25.3
Attitude Aiding East RMS	24.1
Attitude Aiding Down RMS	1.2

It is observed that the non-attitude aided solution improves

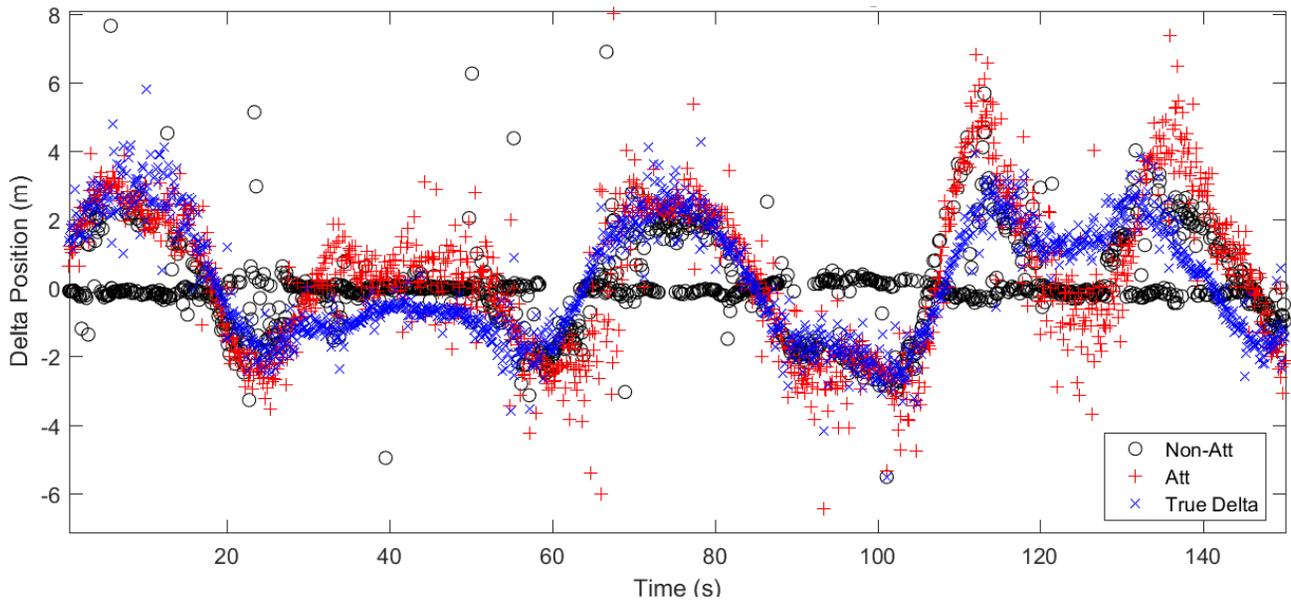


Fig. 8. The change in position in the north direction between images for the truth (GPS) solution, in blue "x"'s; the attitude aided VO solution, in red "+"s; and the VO vision-only solution in black "o"'s. A perfect result would be for the black and red marks to overlay the blue "x"'s. The East plot is similar, changes in down direction were not estimated well by either the attitude aided or non-attitude aided VO algorithms.

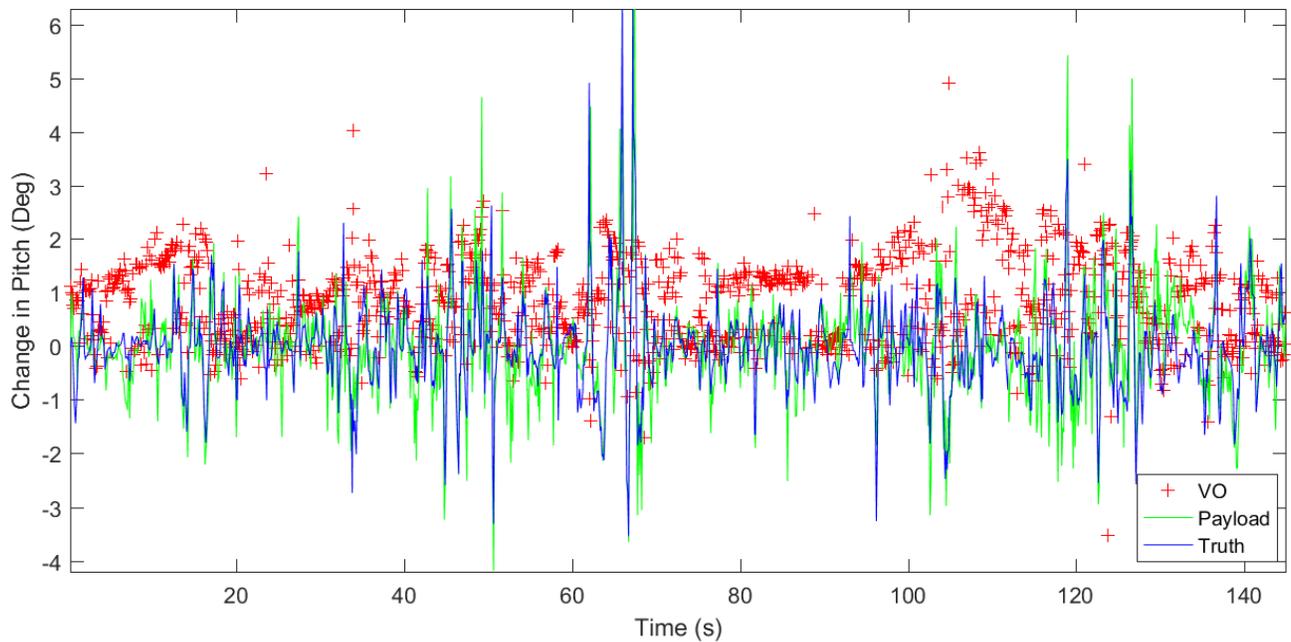


Fig. 9. The **change** in attitude from image to image for the pitch angle from the truth (GPS) Pixhawk, in blue; the Non-GPS Pixhawk, in green, and the non-attitude aided VO solution, in red "+". Note the strong bias of one to two degrees in the red line over time. This bias means that the VO ascribed the difference between the two images to be a change in attitude, not position.

when the image acquisition rate decreases; found by comparing the RMS results in Table II with those in Table III. Also note that the attitude aided solution decreases in performance. The non-attitude aided solution improves due to the overall signal-to-noise increase with a wider baseline. The noise on the images stays relatively constant regardless of the image acquisition rate. However, the signal grows with a longer baseline. This improved signal-to-noise ratio generally

provides the "visibility" to the process to allow the algorithm to extract the differences between translational and rotational motion, giving the approach improved performance. This assumption remains true as long as there is still sufficient overlap to find corresponding features in both images.

The reason for the decrease in performance of the attitude aided solution is due to the quality of the attitude estimates and the fact that a slower image acquisition rate means that

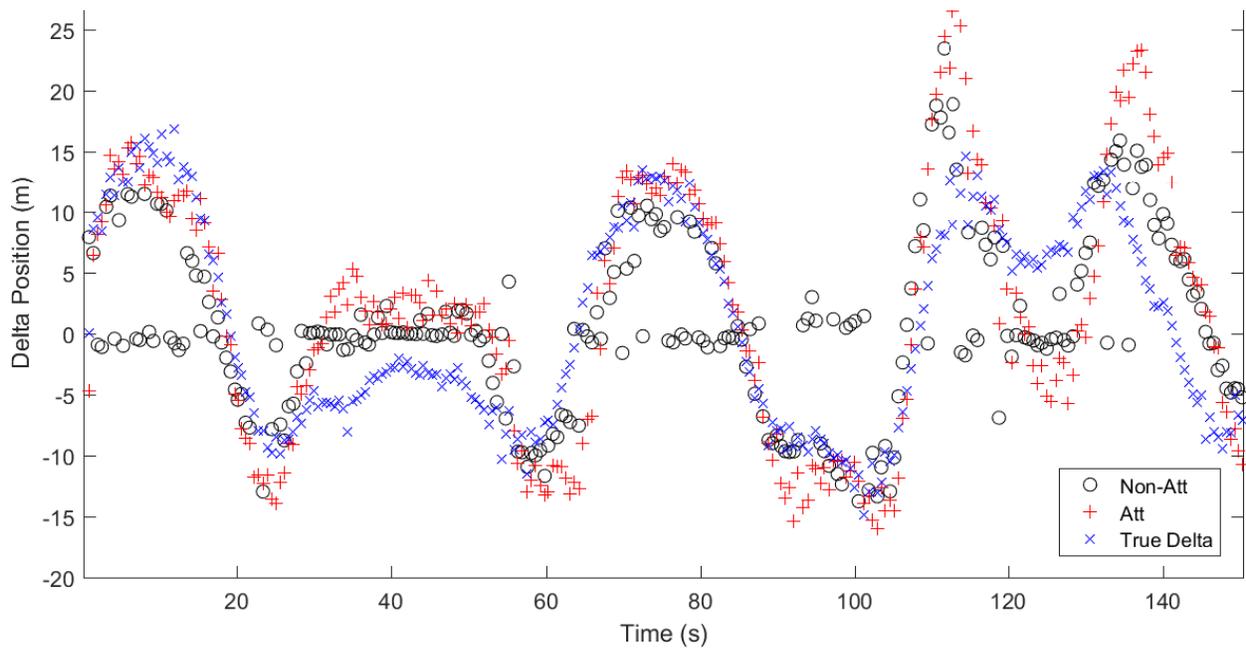


Fig. 10. The change in position in the north direction with images at 2Hz for truth (GPS), in blue "x"'s; the attitude aided outputs, in red "+"s; and the vision-only solution in black "o"'s. The East plot is similar, changes in down direction were not estimated well by either the attitude aided or non-attitude aided algorithms.

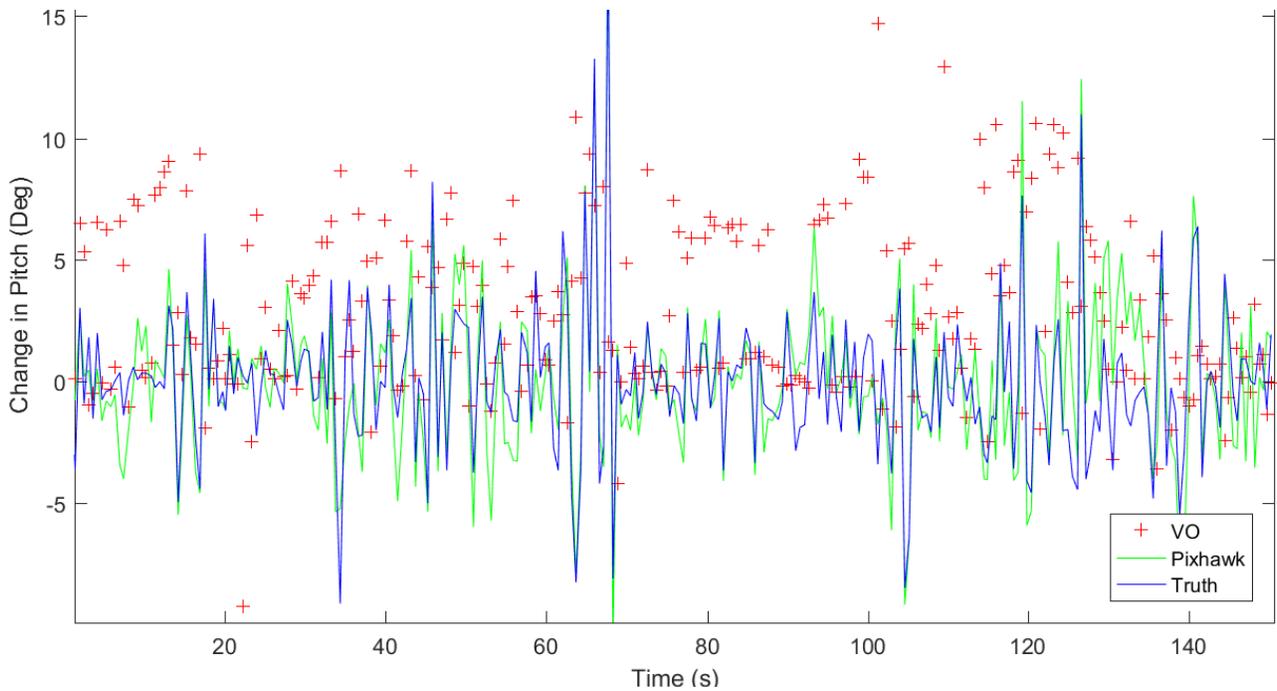


Fig. 11. The **change** in attitude estimates from image to image for the pitch angle with images at 2Hz from the truth (GPS) Pixhawk, in blue; the Non-GPS Pixhawk, in green, and the non-attitude aided VO solution, in red "+". Note the occasional bias in the red "+"s over time. This bias means that the VO ascribed the difference between the two images to be a change in attitude, not position.

the attitudes must be accurate over a longer duration. This can be seen by comparing Figures 9 and 11. In Figure 9 above, the change in attitude for the 8Hz images is estimated similarly between the "truth" Pixhawk and the "payload" Pixhawk autopilots. The comparison between those two autopilots is worse over the longer duration, in Figure 11.

Also, observe the overall attitude estimation performance in Figures 12 and 13. There are dramatic differences over the length of the flight. The poor performance of the "payload" Pixhawk estimates is due to the inability to observe and estimate the accelerometer biases, as no GPS information is available. During short interval comparisons, the attitude

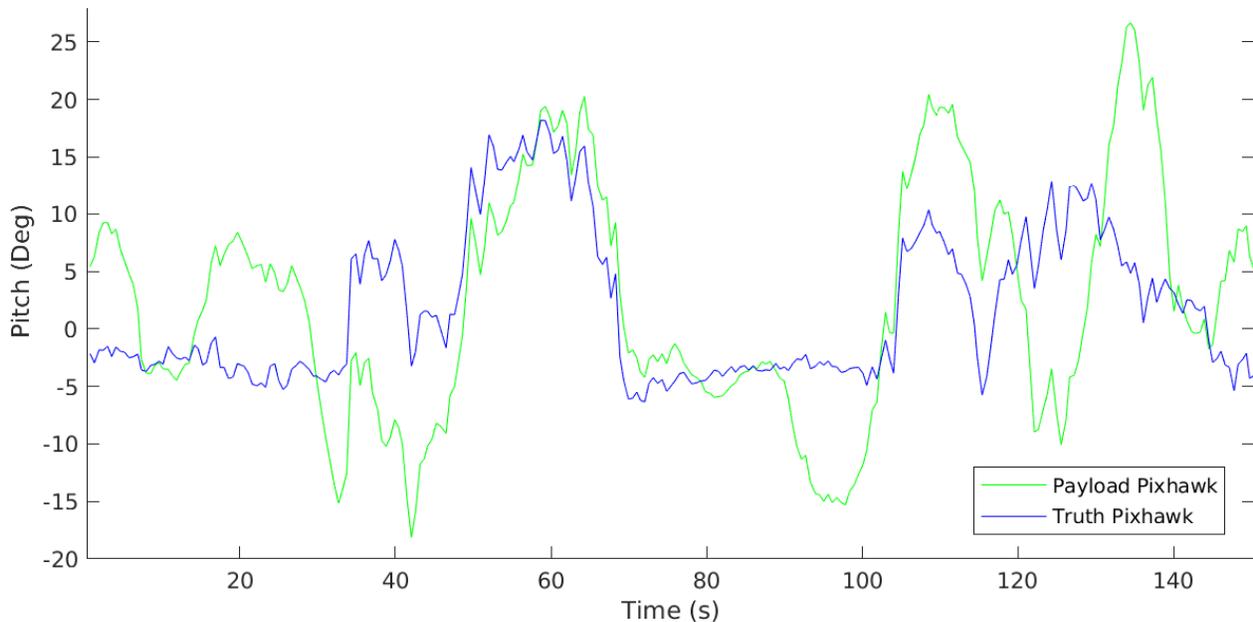


Fig. 12. A comparison of the pitch angle estimated from the two autopilots over the flight. Notice that the Payload Pixhawk did not have access to GPS information, consequently it does not have observability into the accelerometer biases, causing the attitude estimates to drift. In a comparison between this and Figure 9, the small changes in attitude at 8Hz compare well. However, when the imagery is extracted at 2Hz, the larger change in attitude does not align as well, noted by the performance decrease in RMS, Table III.

estimates line up, like those in Figure 9. But as the interval lengthens, the errors in the attitude estimates degrade the solution of the VO algorithm.

C. Altitude Aiding Summary

We have demonstrated that attitude aiding of two different types of vision algorithms can be highly beneficial from a fixed-wing small UAS platform, under 300 meters altitude. This is possible even when the attitude estimates are derived from consumer-grade MEMs sensors and are only accurate to within a few degrees. We found that this relationship is fairly constant over the number of features found in the images (for greater than about 40 to 50 features in each image). The slower the rate that the imagery is acquired, the tighter the requirements on the accuracy of the attitude estimates, particularly for lightweight, highly dynamic, small UAVs. Also, it is critical to either have attitude information acquired much faster than the imagery, by about an order of magnitude, or have the attitude and the imagery synchronized. If either of these conditions is not met, an attitude aided solution will give much worse results than a non-aided solution.

VII. RESULTS: VO COMBINED WITH RANGING

This section transitions to demonstrate the ability to bound the covariance of VIO estimates using range measurements to one other vehicle with known global pose. The same VO algorithm, with attitude aiding, is utilized throughout the research in this section. We begin by reviewing high-level results and then we will descend down into greater detail. First we will review the filter performance using the overall root-mean squared (RMS) error in North and East directions and examine what happens as we change the

update rate for both VO data and Ranging data. Next we examine some approximated ANEES results (19) for some of the filter states, which provides some insight into how consistent the filter is performing. We do not have truth data for all the filter states and some of this data had to be approximated, e.g. velocity information obtained through a dirty-derivative of the position data; nevertheless looking at ANEES provides insight into the performance of the filter. Some of the state estimates from the many generated are next examined in detail to get a better perspective on what the provided overall RMS and ANEES numbers really mean. Finally, we provide numerical analysis of the Observability Index (11) throughout the flight.

A. RMS Results

The filter was run across a wide range of update rates for the VO and Ranging measurements and the overall RMS values were extracted for each run. The VO update rate was varied from 0 to 9 Hz, in integer steps. The Ranging was varied from 0 to 20 Hz in integer steps and then additionally at lower rates between 0 and 1Hz: 0.01Hz to 0.09Hz by 0.01Hz steps and 0.1Hz to 0.9Hz by 0.1Hz steps. Running the filter across this wide swath of measurement update rates allows visibility into trends in the data. Insight into how the filter performance varies with the VO and Ranging update rates is apparent.

We note that the filter was not tuned independently for each of these iterations and so is not optimal at all points shown. Tuning was completed using the first 300 seconds of data from the flight on a few selections of VO and Ranging update rates, predominately at 3 Hz VO and 4 Hz Ranging.

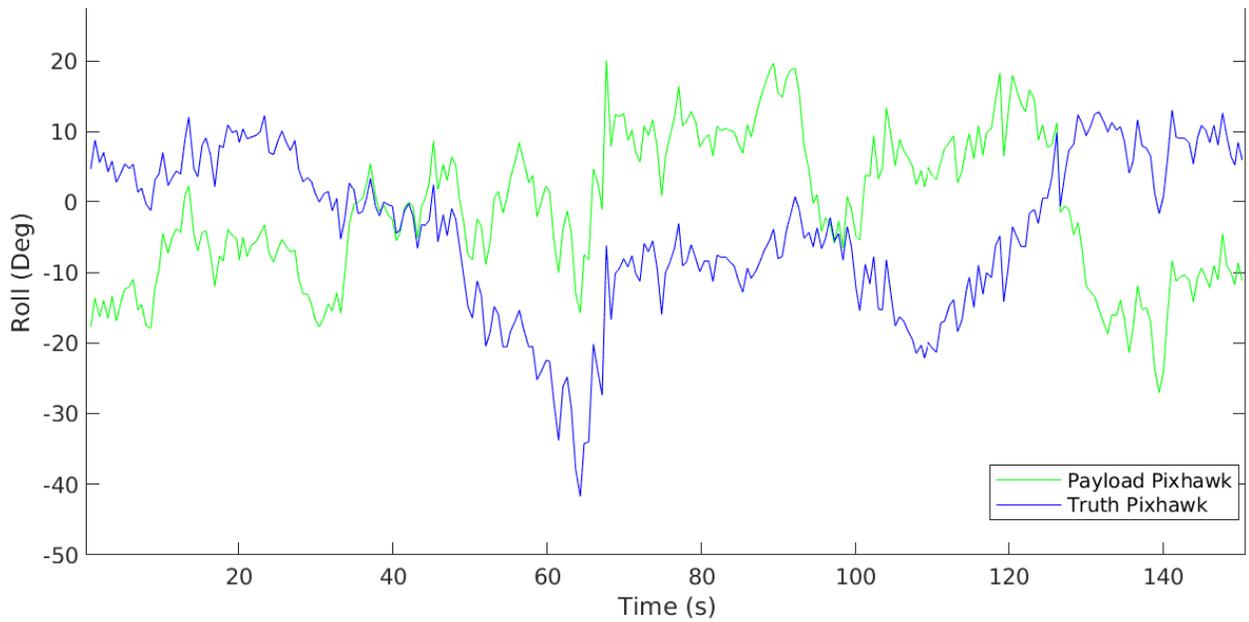


Fig. 13. A comparison of the roll angle estimated from the two autopilots over the flight. Notice that the Payload Pixhawk did not have access to GPS information, consequently it does not have observability into the accelerometer biases, causing the attitude estimates to drift.

Then, the entire batch at all ranging and VO update rates was computed using the same tuning parameters. Again, this approach is sub-optimal but makes the analysis feasible. The VO algorithm was not able to compute solutions over the entire flight for the 1 Hz and 2 Hz settings. The 1 Hz results are excluded, as the duration was only about 130 seconds. The 2 Hz so these results are still shown, but the data reach only 552 seconds into the flight, instead of the full 720. The lines for 0Hz in the figures below signify that VO information was not used, only internal autopilot sensors and range data. The datapoints at 0Hz Ranging correspond to cases where no range data was utilized, only internal sensors and VO measurements.

Figure 14 provides a glimpse of the overall RMS error in the north position of the lower plane. The lines in the figure correspond to different rates of VO measurements, noted by the legend. The x-axis refers to the Ranging measurement update rates. Note how quickly the RMS error drops at the beginning of the figure; it does not take a high rate of Ranging data to begin to bound the drift in the filter. This is shown more clearly in a zoomed-in version of Figure 14 in Figure 15.

By about 3Hz Ranging, there is minimal improvement by including Ranging data any faster. A similar story is found in examining the East RMS error in Figure 16 below.

Figures 14 and 16 also illustrate that there is a significant difference in the performance of the filter when VO data is removed, shown by the line labeled “0” and the performance with some VO data, all the remaining plot lines. This difference could be made more drastic by a state of the art VO algorithm; the algorithm we utilize is a basic comparison of two images and does not do any advanced processing, such as tracking features over multiple frames or bundle

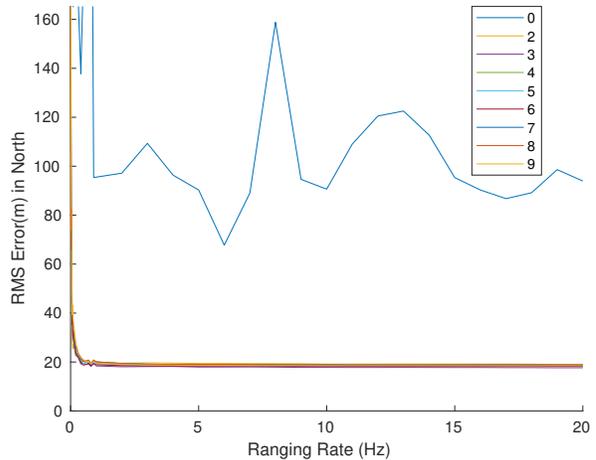


Fig. 14. The overall RMS error is shown in the North direction for position, at different VO update rates (different lines), and for different Ranging update rates (along X axis). Notice how the RMS error drops dramatically with the introduction of range data. By about 3Hz or so, there is minimal improvement for faster range measurement update rates. Note the differences between not having VO data (0 line) and having range data (the other lines). The filter was largely tuned on 3Hz VO data.

adjustment. There is also some difference in performance among the different VO rates, but not by a large margin; 3Hz VO data is a better performer over many range rates in the two figures. Recall that tuning of the filter was completed using the first 300 seconds of data from the flight on a few selections of VO and Ranging update rates, predominately 3 Hz VO and 4 Hz Ranging. Then, the entire batch at all ranging and VO update rates was computed using the same tuning parameters. Consequently, the 3Hz line represents a bounding of the performance of this VO system and any of

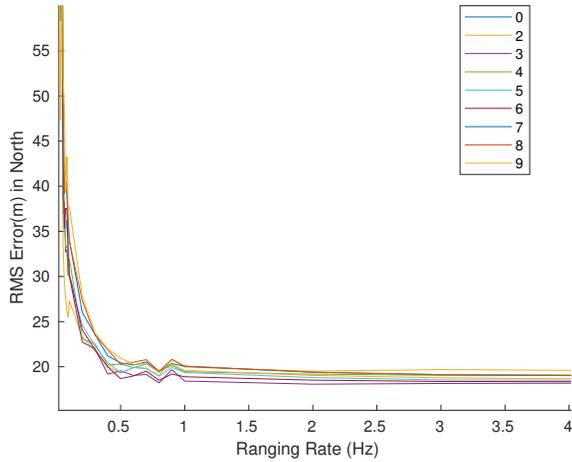


Fig. 15. Zoomed version of Figure 14, showing the steep descent of RMS error with range rate. It was found that even Ranging data at 0.01 Hz (once every 100 seconds) was sufficient to bound the covariance of the filter. Also, for clarity, the line for 0 Hz VO was removed.

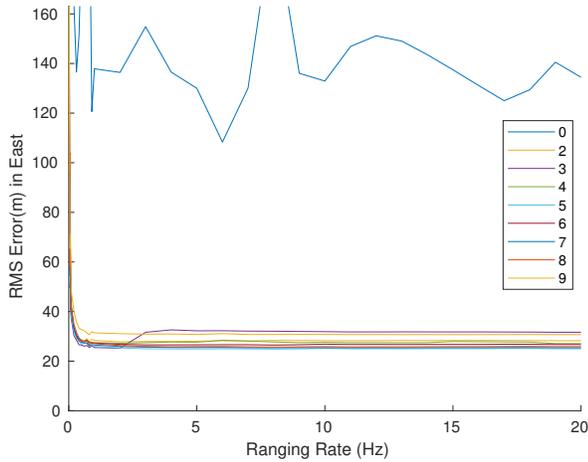


Fig. 16. The overall RMS error in the East direction is shown, at different VO update rates (lines), and for different Ranging update rates (along x axis). Notice how the RMS error drops dramatically with the introduction of range data. By about 3Hz or so, there is minimal improvement for faster range measurement update rates. Note the differences between not having VO data (0 line) and having range data. The filter was largely tuned on 3Hz VO data.

the other VO update rates could be tuned down to close to this performance.

B. ANEES

Figure 17 displays the ANEES over the whole batch of VO and Ranging update rates. This computation of ANEES only considers the 3 position states in the filter p_n , p_e and p_a and the associated reduced covariance. The ANEES for all these lines should therefore be at 3 for all permutations of VO and Ranging updates. On the whole, the filter is inconsistent and the estimated covariance is smaller than the true covariance actually is, with the only exception being the 2 Hz VO results, which are slightly under-confident. However, this inconsistency is not egregious. Detailed state estimates will be shown in the next section, and as will be seen, there are only a few periods of the 12 minute flight that

are responsible for the ANEES to be higher than it ought to be.

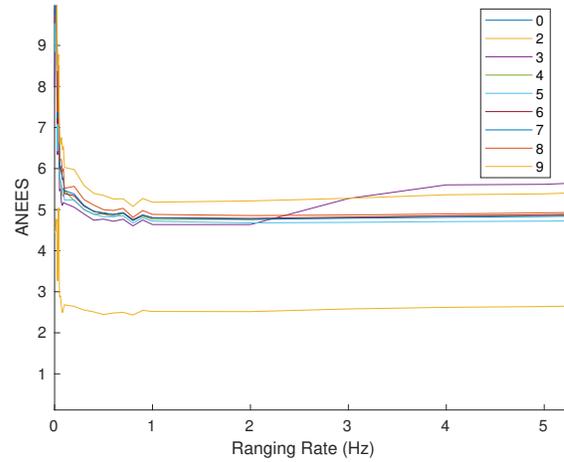


Fig. 17. This figure shows the ANEES for only the 3D position data $[p_n, p_e, p_a]$, at different VO update rates (lines), and for different Ranging update rates (X axis). The line for 0Hz VO is not visible in this figure, as it was zoomed to show detail on the other lines. The ANEES for a consistent filter in this scenario would lie at 3, since there are 3 states considered in this calculation. Only the 2 Hz VO line is below 3, meaning that it is slightly under-confident (covariance is slightly bigger than it actually is).

C. Detailed State Estimates

Detailed results of the filter performance for the state estimates are discussed in this section. First we examine state error estimates for the case of 5 Hz VO and 3 Hz Ranging, one of the best cases shown in Figures 14 and 16. Next the filter results at 8 Hz VO and 0.2 Hz Ranging are shown to demonstrate the results with a slow Ranging update and more middle of the road performance. Next a quick review of the position results at 4 Hz VO and 0.01 Hz Ranging are shown, demonstrating that bounding of the filter covariance occurs even at that low Ranging rate. The results for 0 Hz Ranging and then 0 Hz VO are then shown to demonstrate that the filter performs as expected at the extremes. Overall, the results will demonstrate a converging but slightly overconfident filter; and most importantly, a filter covariance that is bounded by the inclusion of Ranging information to one other vehicle with known pose.

1) *5 Hz VO and 3 Hz Ranging*: Figure 18 shows the errors in the state estimates of the filter over the whole flight, with VO updates at 5 Hz and Ranging updates at 3 Hz. Also shown are 3σ bounds using the estimated covariance. Note that the filter covariances remain bounded over the trajectory even though Ranging data is available to only one landmark. As discussed in Section IV, the complete observability of the state is due to the combination of the yaw information provided by the VO algorithm and the range to the one known vehicle.

2) *8 Hz VO and 0.2 Hz Ranging*: The errors in the states for a different scenario are presented in Figure 19. Here, VO data is received at 8 Hz and Ranging at 0.2 Hz. Similarities to the the case in Figure 18 can be seen, including similar locations in time where the filter estimates are not consistent.

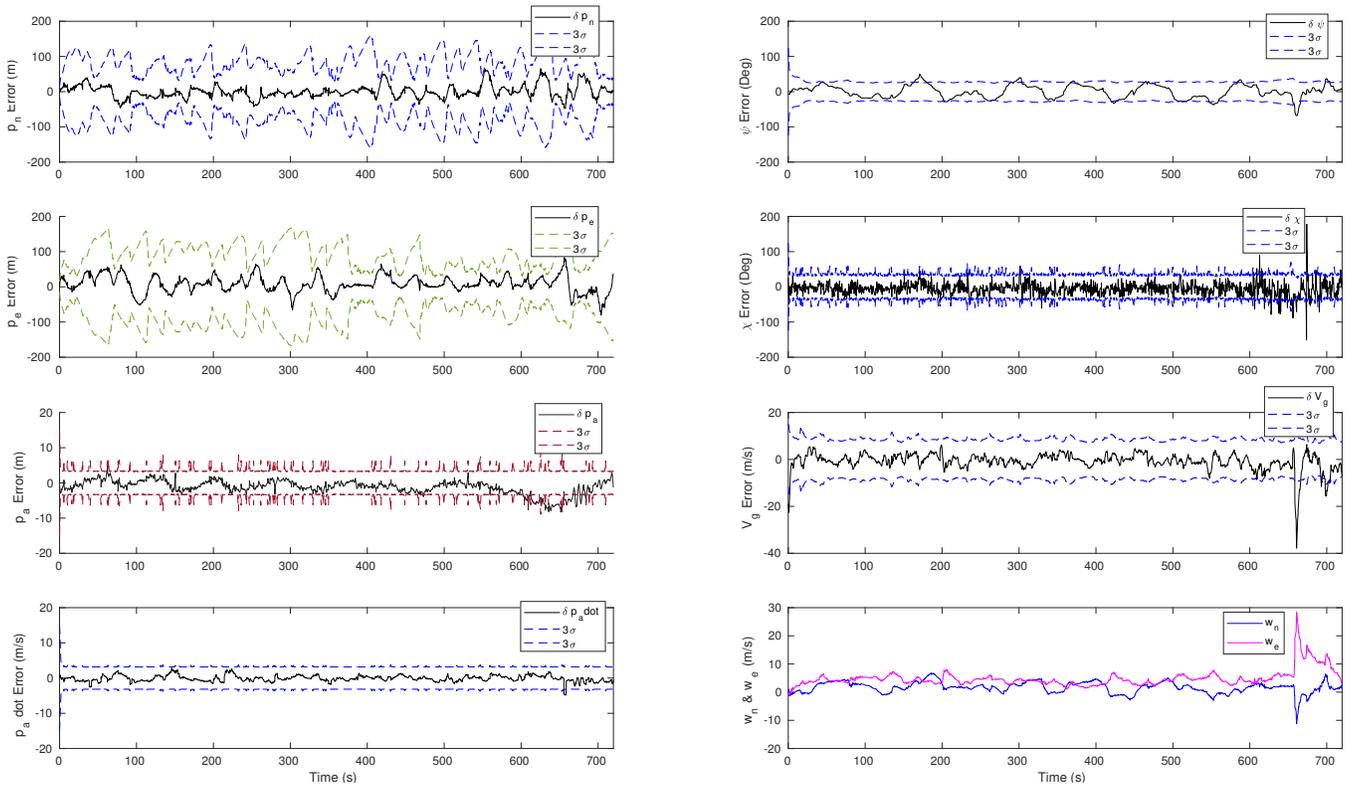


Fig. 18. This figure displays the error in the states (except wind states - no truth information was available) along with 3σ bounds of the covariances for the scenario of 5 Hz VO and 3 Hz Ranging. For the most part, the states are well-behaved, meaning they remain bounded by the 3σ bounds, even though the accuracy is less than desirable. However, there are portions of the trajectory where errors in each state do exceed the bounds; ψ is particularly prone to overconfidence.

Note the larger scale of the y-axes in Figure 19. The 3σ bounds are larger and the performance in p_e is worse than before. The key point here is that the horizontal position is still bounded, even at the low rate of 0.2 Hz Ranging.

3) *0.01 Hz Ranging*: Figure 20 demonstrates that the horizontal position is still bounded at the even lower rate of 0.01 Hz Ranging for 4 Hz VO; although the performance is quite worse than those with higher Ranging rates, as expected from Figure 14. This fact demonstrates that infrequent ranging data from a single known location is sufficient to bound the drift of a VIO estimation approach for fixed-wing UAVs that has access to yaw and pressure information.

4) *Range-only Measurements*: Figure 21 shows the error in the states when VO data is not available, but Ranging data is received at 10 Hz. Figure 21 demonstrates that the divergent-convergent cycle behavior exist even when VO data is not available. Note the drop in yaw (ψ) performance, which demonstrates that the χ update from the VO is providing the yaw information.

5) *VO-only Measurements*: Figure 22 plots the error in the state for 5 Hz VO and 0 Hz Ranging. As expected, the p_n and p_e estimates grow without bound since those positions are unobservable without the range. If the VO were not also providing yaw information through the χ measurement in (6), ψ would also be unobservable.

D. Numeric Observability

An evaluation of the numerical characteristics of the observability index can be completed by an inspection of the rank of the matrix (11) at all timesteps within the flight. If (11) is not full rank, the system is not observable.

Figure 23 shows the rank of (11) evaluated using the filter states and inputs at about 4 Hz throughout the entire flight for the case of 5 Hz VO and 2 Hz Ranging. Note that it is full rank throughout the flight, as was expected from the analytic expression. Range to only one known landmark and the yaw information from the VO are sufficient to fully constrain the covariance of a VIO plus pressure system onboard a fixed-wing UAV. We note that the rank of \mathcal{O} was also full rank for every other combination of VO and Ranging update rates, with the exception of 0 Hz Ranging.

1) *Condition Number*: The problem with only analyzing the rank of (11) is that there is no additional information provided, no ability to see how strong the local observability condition is. The conjecture we pose, without proof as yet, is that if the observability index is poorly conditioned then the flow of information from the measurements to the states will be sensitive to errors and more fragile.

The condition number of a matrix gives an indication of the sensitivity of that matrix to changes in input for a computation like $\mathbf{Ax} = \mathbf{b}$. If the matrix \mathbf{A} is poorly conditioned then small changes (or errors) in the input vector

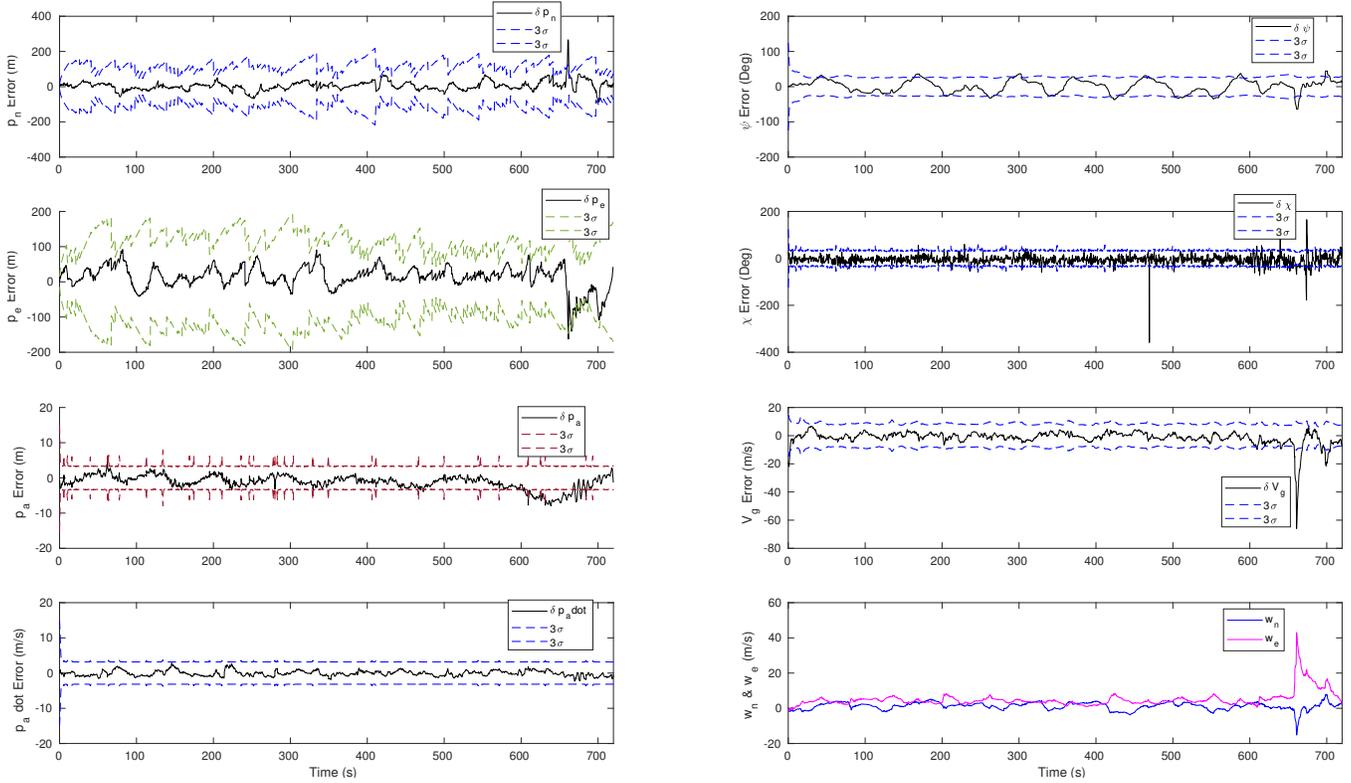


Fig. 19. The error in the states are shown with the exception of wind states, where truth data was not available, along with 3σ bounds of the covariance for the scenario of 8 Hz VO and 0.2 Hz Ranging. The results are fairly similar to those in Figure 18 above. Note differences in the scales on the y-axes. The largest difference from the results in Figure 18 is the performance in p_e .

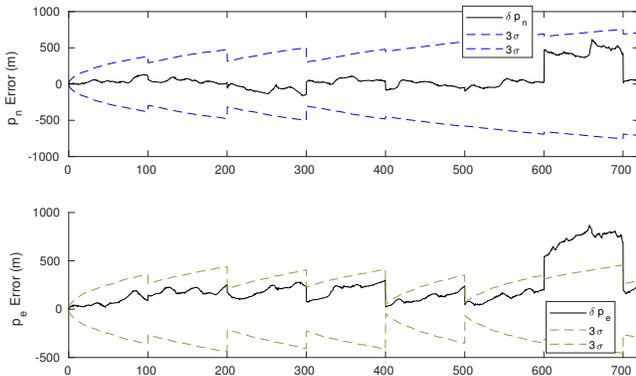


Fig. 20. The error in the horizontal position states are shown along with 3σ bounds of the covariance for the scenario of 4 Hz VO and 0.01 Hz Ranging. The position covariance is bounded, even though ranging data is only received every 100 seconds.

x can result in large changes (or errors) in the output vector b . The condition number is defined as a ratio of the largest to the smallest singular values of a matrix. Ideally, it should be equal, or close to, one for a well-conditioned matrix.

Recall that the observability index \mathcal{O} represents how information from the measurements and the prediction model flows to the states. If the rank of (11) is deficient it means that there is insufficient information from the measurements and model to distinguish the state at one time with the state

at another time.

Analytically, (11) was shown to be full rank in Section IV. However, there are one or two segments of the flight where state estimates begin to diverge and then the filter quickly re-converges. This is clearly seen in the detailed state plots above as well as the attached video of the 3D plots.

Figure 24 plots the condition number of (11) over the whole flight, again for 3 Hz VO update rate and 1 Hz Ranging. The condition is quite poor for some periods of the flight.

Although a little difficult to see from Figure 24, the minimum condition number for this scenario is 303, the mean is $3.1e3$ and the mode is 386. There are several times in the flight when the condition number reaches very high ($\approx 6e4$) and once were it ascends to over $2.2e5$, clearly denoting that the system is approaching being unobservable from a numerical stability standpoint.

The condition number of (11) throughout the flights of all the different range and update rates are similar in nature, denoting that different rates of measurement are not affecting the quality of the observability of the system. To improve the condition number, and thus the observability of the system, more measurements are needed. This result is not surprising, due to the expectations going into the program. Range to another landmark would provide this extra information, changing the condition of \mathcal{O} and improving the performance. However, it is important to realize that range to a single

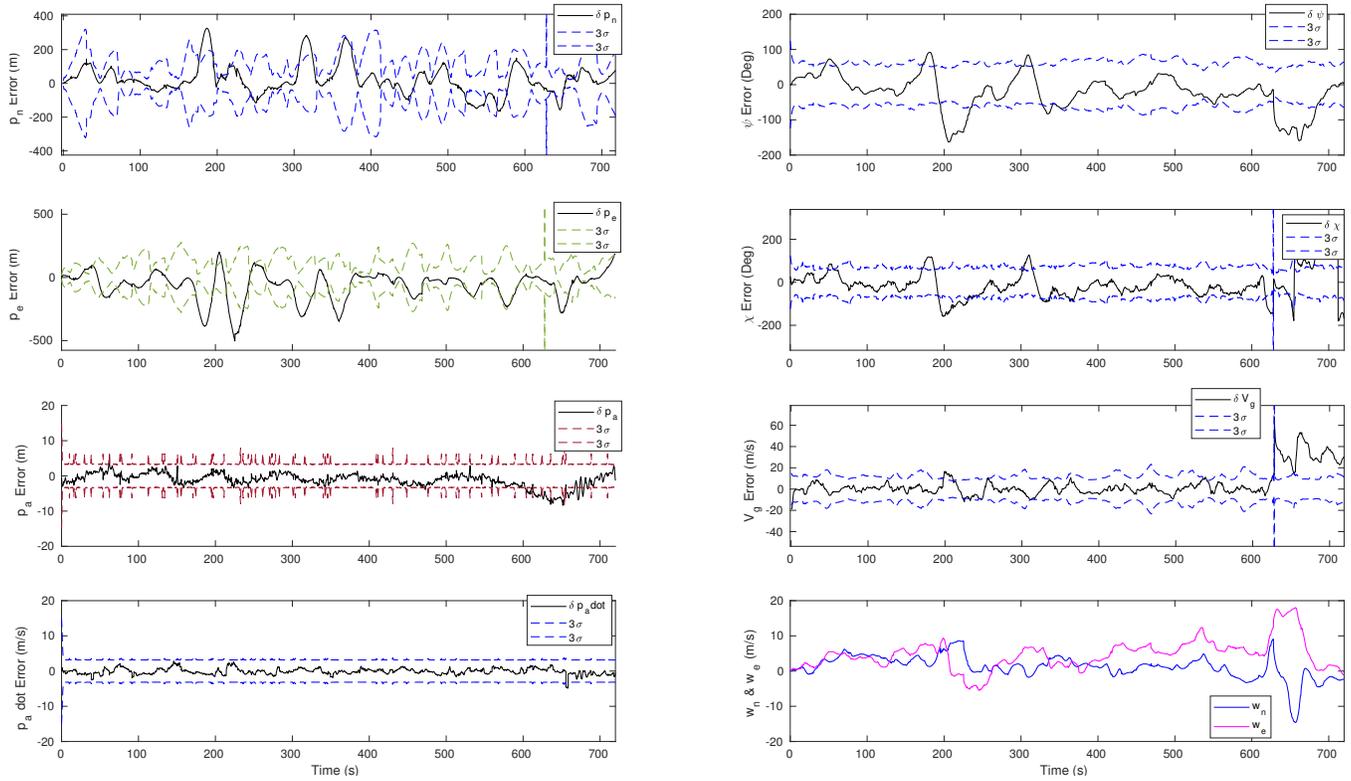


Fig. 21. This figure displays the error in the states, again with exception of the wind states, along with 3σ bounds of the covariance for the scenario of 0 Hz VO and 10 Hz Ranging, in other words, VO data was not used for these results. Note differences in the scales on the y-axes.

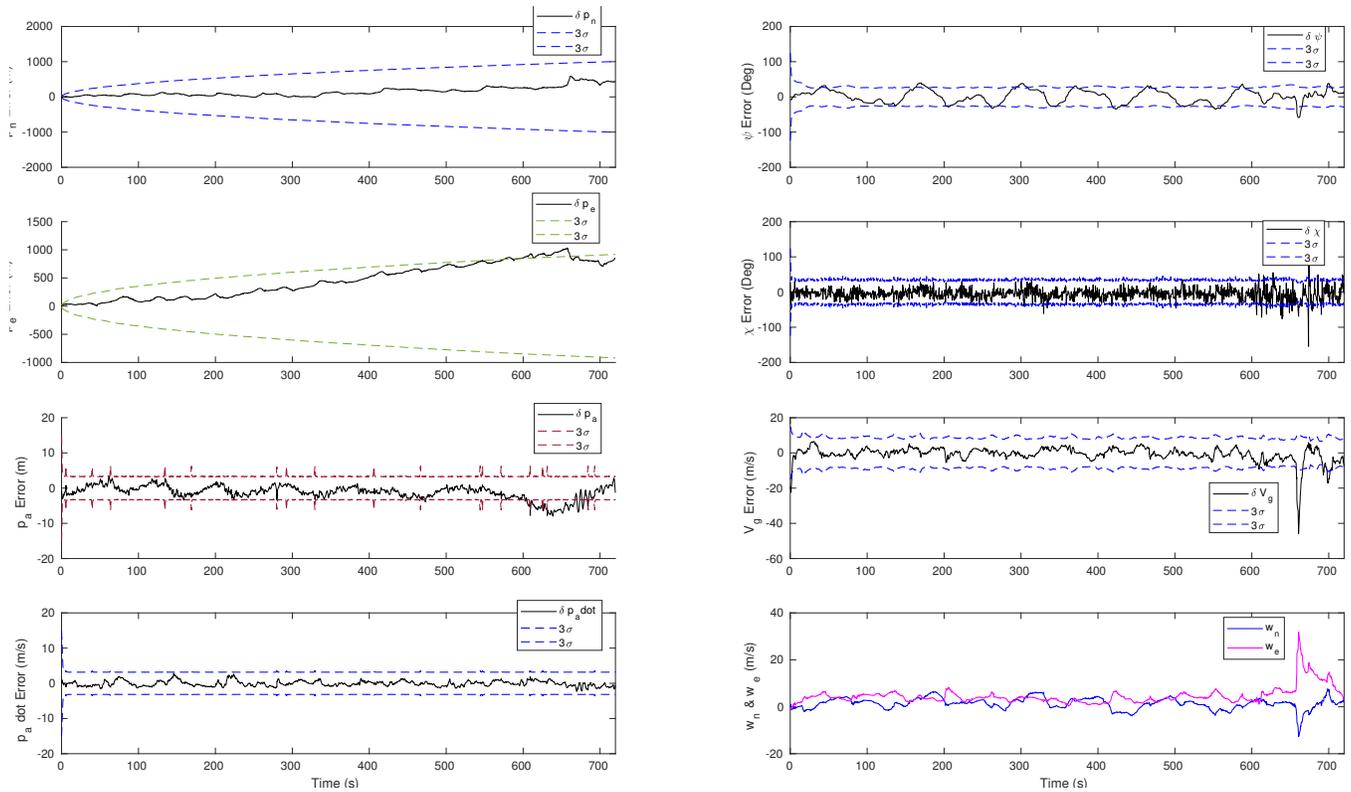


Fig. 22. This figure displays the error in the states along with 3σ bounds of the covariance for the scenario of 5 Hz VO and 0 Hz Ranging, in other words, Ranging data was not used for these results. The horizontal position will grow without bound in this case, as those positions are unobservable without Ranging information.

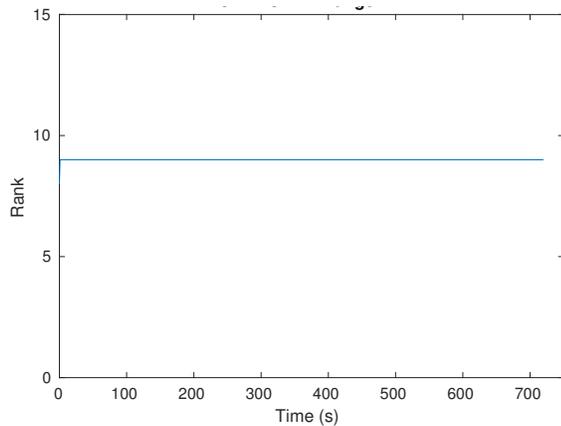


Fig. 23. The rank of (11) for the case of 5 Hz VO and 2 Hz Ranging updates. Note that the matrix is full rank throughout the flight, $\text{length}(\mathbf{x}) = 9$. The \mathcal{O} matrix is also full rank for every other combination of VO and Ranging updates, with the exception of 0 Hz, or no Ranging information.

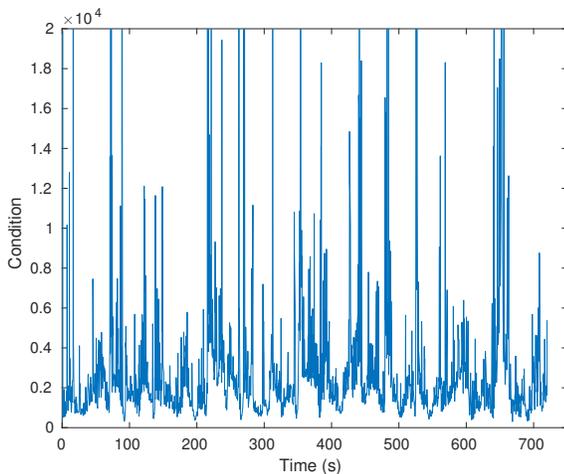


Fig. 24. The condition number of (11) for the case of 5 Hz VO and 2 Hz Ranging updates, note the exponent applied to the y-axis. Even though the matrix is full rank throughout the flight, the matrix is poorly conditioned in several periods of the flight. It is not surprising, based on this result shown here, that there are portions of the trajectory where the filter begins to diverge and then re-converges.

landmark can drastically improve system performance of a GPS-denied, visually guided UAV by bounding the drift in the planar position states.

VIII. CONCLUSIONS

Several key conclusions can be drawn from the results demonstrated above. The first is that attitude data derived from cheap, consumer-grade MEMS IMU sensors is sufficiently useful to improve the performance of vision navigation algorithms. The improvement is useful for both relative algorithms, such as VO, and global algorithms, such as PNP, for comparing imagery to known maps. We found this relationship to be fairly constant over the number of features found. It was also found that the slower the rate of imagery, the tighter the requirements on the accuracy of the attitude estimates are. Also, it is critical to either have attitude information acquired much faster than the imagery,

by about an order of magnitude, or have the attitude and the imagery synchronized. If either of these conditions is not met, an attitude aided solution will provide much worse results than a non-aided solution.

The second key conclusion is that range to a single vehicle with known global position provides a tremendous benefit to a vehicle using a relative vision approach, like VO. The range information allows the vehicle to constrain the global drift and make the vision data begin to match up with the true trajectory, when pressure and magnetometer data are available. A key take-away is that even though the global position is observable in this scenario, the solution is poorly conditioned and can exhibit regions of divergence before the solution re-converges. It is possible that the solution at those points becomes multi-modal, with several positions being near-equally likely. Perhaps the use of a multi-modal filter, like a particle filter, would alleviate some these issues. In the future we plan to investigate the use of a particle filter on this problem, through the Scorpion framework.

The third key conclusion is a follow-on to the second. Ranging data is not needed at high rates to be effective at constraining global drift. It was observed that rates of 0.01 Hz, or once every 100 seconds, was sufficient to bound the filter covariance over the 12 minute flight available for this data analysis. Granted, faster rates than 0.01 Hz do help with the accuracy of the solution, but the improvement is minimal above an update rate of about 3 Hz, for the scenario investigated.

ACKNOWLEDGMENT

We would like to thank Rick Patton and Dave Thacker for their help and expertise with the UAVs. We also acknowledge and are grateful for the previous work completed by Tim Machin and Ben Fain.

DISCLAIMER

The views expressed in this paper are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or U.S. Government.

REFERENCES

- [1] P. F. Alcantarilla, A. Bartoli, and A. J. Davison, "KAZE features," in *European Conference on Computer Vision*, Berlin, Germany, 2012, pp. 214–227.
- [2] H. Bai and C. N. Taylor, "Control-enabled Observability in Visual-Inertial Odometry," in *International Conference on Unmanned Aircraft Systems (ICUAS)*, Miami, FL, 2017.
- [3] R. W. Beard and T. W. McLain, *Small Unmanned Aircraft*. Princeton University Press, 2012.
- [4] M. Bloesch, S. Omari, M. Hutter, and R. Siegwart, "Robust visual inertial odometry using a direct EKF-based approach," *IEEE International Conference on Intelligent Robots and Systems*, vol. 2015-Decem, pp. 298–304, 2015.
- [5] A. J. Canciani and A. Force, "Absolute Positioning Using the Earth's Magnetic Anomaly Field," *Navigation: Journal of The ION*, vol. 63, no. 2, pp. 111–126, 2016.
- [6] D. J. Carson, J. F. Raquet, and K. J. Kauffman, "Aerial Visual-Inertial Odometry Performance Evaluation," in *ION Pacific PNT*. Honolulu: Institute of Navigation, 2016.

- [7] J. Engel, J. Sturm, and D. Cremers, "Scale-aware navigation of a low-cost quadcopter with a monocular camera," *Robotics and Autonomous Systems*, vol. 62, no. 11, pp. 1646–1656, 2014. [Online]. Available: <http://dx.doi.org/10.1016/j.robot.2014.03.012>
- [8] B. M. Fain, "Small Fixed-wing Aerial Positioning using Inter-vehicle Ranging Combined with Visual Odometry," Ph.D. dissertation, Air Force Institute of Technology, 2017.
- [9] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, "On-Manifold Preintegration for Real-Time Visual Inertial Odometry," *IEEE Transactions on Robotics*, vol. 33, no. 1, pp. 1–21, 2017.
- [10] C. Forster, M. Pizzoli, and D. Scaramuzza, "Appearance-based Active, Monocular, Dense Reconstruction for Micro Aerial Vehicles," *Robotics: Science and Systems X*, no. 200021, 2014.
- [11] D. A. Grejner-Brzezinska, C. K. Toth, T. Moore, J. F. Raquet, M. M. Miller, and A. Kealy, "Multisensor Navigation Systems: A Remedy for GNSS Vulnerabilities?" *Proceedings of the IEEE*, vol. 104, no. 6, pp. 1339–1353, 2016.
- [12] Y. Kawano and T. Ohtsuka, *Observability analysis of nonlinear systems using pseudo-linear transformation*. IFAC, 2013, vol. 9, no. PART 1. [Online]. Available: <http://dx.doi.org/10.3182/20130904-3-FR-2041.00100>
- [13] R. C. Leshman, J. Gray, and J. Raquet, "Utilization of UAV Autopilots in Vision-based Alternative Navigation," in *GNSS+*. Portland, OR: Institute of Navigation, 2017.
- [14] X. R. Li, Z. Zhao, and V. P. Jilkov, "Practical measures and test for credibility of an estimator," *Proc. Workshop on Estimation, Tracking, and ...*, pp. 481–495, 2001.
- [15] Y. Lin, F. Gao, T. Qin, W. Gao, T. Liu, W. Wu, Z. Yang, and S. Shen, "Autonomous aerial navigation using monocular visual-inertial fusion," *Journal of Field Robotics*, no. May, pp. n/a–n/a, 2017. [Online]. Available: <http://dx.doi.org/10.1002/rob.21732>
- [16] Y. Ling, T. Liu, and S. Shen, "Aggressive quadrotor flight using dense visual-inertial fusion," in *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 2016-June, 2016.
- [17] T. Machin, J. F. Raquet, D. Jacques, and D. Venable, "Real-time implementation of vision-aided navigation for small fixed-wing unmanned aerial systems," *29th International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS 2016*, vol. 2, pp. 1305–1311, 2016. [Online]. Available: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85017309692partnerID=40md5=c099ffec3f855b77aac407a33384a995>
- [18] T. Machin, "REAL-TIME IMPLEMENTATION OF VISION-AIDED MONOCULAR NAVIGATION FOR SMALL FIXED-WING UNMANNED AERIAL SYSTEMS," Ph.D. dissertation, Air Force Institute of Technology, 2016.
- [19] I. Sa, M. Kamel, M. Burri, M. Bloesch, R. Khanna, M. Popovic, J. Nieto, and R. Siegwart, "Build your own visual-inertial odometry aided cost-effective and open-source autonomous drone," *submitted to IEEE RAM*, pp. 1–21, 2017. [Online]. Available: <http://arxiv.org/abs/1708.06652>
- [20] V. Sazdovskii, A. Kitanov, and I. Petrovic, "Implicit observation model for vision aided inertial navigation of aerial vehicles using single camera vector observations," *Aerospace Science and Technology*, vol. 40, pp. 33–46, 2015. [Online]. Available: <http://dx.doi.org/10.1016/j.ast.2014.09.019>
- [21] D. H. Titterton and J. L. Weston, *Strapdown Inertial Navigation Technology*, 2nd ed. The Institution of Engineering and Technology, 2004.
- [22] D. Venable and J. Raquet, "Large scale image aided navigation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 6, pp. 2849–2860, 2016.
- [23] M. Vidyasagar, *Nonlinear Systems Analysis*, 2nd ed. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2002.