AIAA SciTech Forum
7-11 January 2019, San Diego, California
AIAA Scitech 2019 Forum

10.2514/6.2019-1055

Optical Geolocation for small Unmanned Aerial Systems

Chester V. Dolph,[1] Robert G. McSwain[2], William M. Humphreys, Jr[3], David P. Lockard[4], Mehdi R. Khorrami[5]
NASA Langley Research Center, Hampton, VA, 23681

This paper presents an airborne optical geolocation system using four optical targets to provide position and attitude estimation for a sUAS supporting the NASA Acoustic Research Measurements (ARM), where the goal is to reduce nuisance airframe noise during approach and landing. A large precision positioned microphone array captures the airframe noise for multiple passes of a Gulfstream III aircraft. For health monitoring of the microphone array, the Acoustic Calibration Vehicle (ACV) sUAS completes daily flights with an onboard speaker emitting tones at frequencies optimized for determining microphone functionality. An accurate position estimate of the ACV relative to the array is needed for microphone health monitoring. To this end, an optical geolocation system using a downward facing camera mounted to the ACV was developed. The 3D positioning of the ACV is computed using the pinhole camera model. A novel optical geolocation algorithm first detects the targets, then a recursive algorithm tightens the localization of the targets. Finally, the position of the sUAS is computed using the image coordinates of the targets, the 3D world coordinates of the targets, and the camera matrix. A Real-Time Kinematic GPS system is used to compare the optical geolocation system.

I. Nomenclature

2D	=	two Dimensional
3D	=	three Fimensional
ACTE	=	Adaptive Compliant Trailing Edge
ACV	=	Acoustic Calibration Vehicle
AFB	=	Air Force Base
ARM	=	Acoustic Research Measurements
CMOS	=	Complementary metal–oxide–semiconductor
CMT	=	Civil Mean Twilight
GPS	=	Global Positioning System
f	=	ratio of focal length to diameter of aperture, also called f-stop or f-number
MEMS	=	Micro-Electro-Mechanical Systems
microSD	=	Micro-Secure Digital
NASA	=	National Aeronautics and Space Administration
NeAR	=	Networked Array Recorder
RTK	=	Real-Time Kinematic
SALT	=	Structural Acoustic Loads and Transmission
sUAS	=	small Unmanned Aerial System

II. Introduction

The aim of this work is to provide an alternate positioning system for the Acoustic Calibration Vehicle (ACV), which is a small Unmanned Aerial System (sUAS) supporting the NASA Acoustic Research Measurements (ARM) project. The overall objective of the ARM project is to validate NASA design tools which reduce aircraft noise.

[1] Aerospace Engineer, Aeronautics Systems Engineering Branch, AIAA Member
[2] Aerospace Engineer, Aeronautics Systems Engineering Branch, AIAA Member
[3] Research Engineer, Adv. Measurements and Data Systems Branch, Associate Fellow AIAA
[4] Aerospace Technologist, Computational AeroSciences Branch, Senior Member, AIAA
[5] Aerospace Engineer, Computational AeroSciences Branch, Associate Fellow AIAA

Several aircraft modifications were designed and fabricated to reduce airframe noise at low altitudes for take-off and landing. Aircraft modifications include landing gear fairings, landing gear cavity treatments, and ACTE flaps. Flight testing to validate predicted acoustic effects was conducted by two NASA research aircraft (NASA 804, NASA 808) flying over a 185 microphone array [1] [2] with varying configurations (e.g. gear-up/down, flaps- 0/30). Acoustic propagation is sensitive to changes in environmental conditions, therefore a method of characterizing the environmental effects is needed. At the beginning and end of each flight test, a speaker mounted to the ACV emits three specific tones for three different altitudes to provide insight into the current environmental conditionals for sound propagation and to verify the array is functioning properly. This requires a highly accurate position and attitude for the sound source mounted on the ACV, which currently utilizes an RTK GPS position and MEMS sensor attitude.

This work tests a proposed optical geolocation system design which includes optical targets and a camera. The optical system may serve as a complimentary 3D position source to GPS for the flight controller. The optical targets are designed as unique navigational markers. These visual markers may be installed on rooftops or parking lots and serve as additional navigational aides for autonomous sUAS flights. Specifically this can complement navigation with GPS or GPS-denied environments, such as urban canyons. A sUAS with an onboard calibrated camera may use the visual targets for the 3D positioning during take-off, landing, or navigation. An image differencing approach that subtracts the image frame from the known visual appearance of the optical target may potentially establish whether a foreign object (e.g. person) is on the optical target. In this way, the visual targets may be used an autonomous landing zone for applications such as package delivery.

This conference paper presents the ARM configuration of the optical geolocation system, the algorithm used for computing 3D coordinates of the ACV, and a comparison of optical geolocation versus RTK GPS for the ACV.

III. ARM III Phased Array System

Description of Array: The microphone phased array deployed for the ARM III flight test consisted of 185 hardened microphones (Figure 1) placed on the overrun area of inactive runway 24 at Edwards AFB. The microphone pattern chosen for the array (Figure 2) arranged the sensors into 12 spiral arms with 15 or 16 microphones populating each arm. The innermost 49 microphones of the array were mounted on a 1.82 meter (6 ft) diameter aluminum central mounting plate with the remainder of the microphones mounted on individual 0.31 meter (1 ft) Plexiglas ground plates. The microphones were connected to a central data acquisition system via individual coaxial cables providing power to the sensors using a 4-mA current loop connection. Details of the array construction and performance can be found in [1]. Twelve additional prototype Networked Array Recorder (NeAR) microphones (red squares in Figure 2) were included in the array pattern for ARM III. These are described in [3].

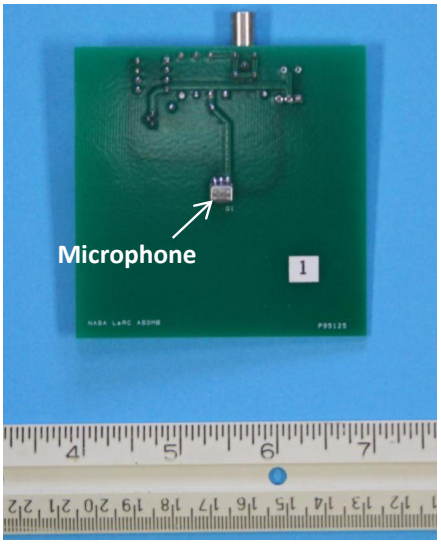


Figure 1: ARM III array microphone.

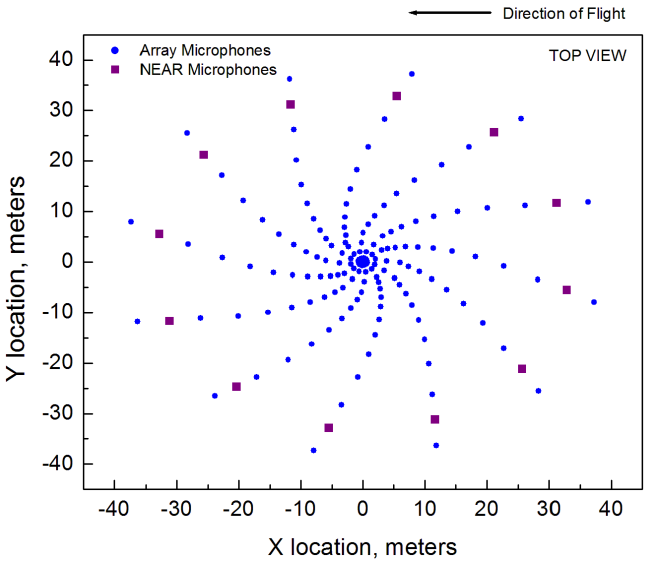


Figure 2: ARM III array pattern.

Array Health Monitoring Concept: A critical aspect of the array operation for the ARM III deployment was the inclusion of array health monitoring methods whereby daily in-situ array sensor response (i.e., sensitivity) tracking and channel performance data could be obtained for the duration of the deployment. This is a non-trivial task, given the constraints imposed by the ambient environment (i.e., weather conditions, background noise) as well as the large size of the array aperture. In general, it is not practical to perform a daily calibration of each individual microphone in the array. Therefore, global methods for monitoring the response of all of the array elements must be employed. Two methods were utilized during ARM III for health monitoring:

1. Multiple ground-based sound sources were embedded in the array at known locations that could be operated either independently or in unison to excite the array elements on demand. Comparisons between baseline and repeat measurements allowed the performance and health of individual sensors to be tracked on a daily to weekly basis. [2] contains a detailed description of this health monitoring method.

2. A unique in-situ health monitoring system was developed using an aerial sound source attached to a hovering vehicle that could be launched over the array. This health monitoring technique is described in detail in this section.

Aerial Sound Source: In principle, use of an aerial sound source should allow comparison of recorded microphone output levels during flights to be conducted on a daily or weekly basis, as long as the effects of speaker directivity, sound propagation (including meteorological conditions) and vehicle station keeping are taken into account. A customized sUAS vehicle was chosen as the platform for the aerial source, denoted here as the Acoustic Calibration Vehicles (ACV). The salient features of the vehicle utilized during ARM III are summarized in Table I in section IV.

The vehicle payload was a suspended, vehicle-powered Anchor® MiniVox Lite public address speaker and MP3 player, shown attached to the undercarriage of the vehicle in Figure 3. The speaker provided both tonal and white noise excitation of the array at a number of vehicle altitudes spanning 45 to 106 meters (150 to 350 feet). The speaker tones were validated in the Langley SALT anechoic chamber using the speaker and a B&K microphone. Note that the frequencies were validated as well as the speaker directivity; however, the amplitude levels were dependent on the MP3 player volume setting as well as the speaker volume setting during individual flights.

Aerial Source Health Monitoring Procedure: The following procedure was followed to perform health monitoring using the ACV:

1. The sound source on the ACV was programmed to produce a 4-kHz tone and the vehicle was launched, centered over the array, and brought to three designated altitudes 106 , 76 , and 45 meters (150, 250, and 350 feet).
2. Time history data was captured for all of the array and NeAR microphones over a nominal 10-second recording cycle with the ACV at the three altitudes.
3. For the initial run, the pressure squared values for each microphone were computed at a 4-kHz narrowband frequency and the equivalent sound pressure level (SPL) was calculated. Full corrections for speaker directivity and atmospheric attenuation as described in [2] were applied to the SPL values. This defined the baseline SPL for each microphone.
4. Absolute output changes across the array due to speaker drift as a function of time must be taken into account. Thus, for subsequent aerial source runs, the current microphone outputs had a reference level, dB_{ref}, subtracted from them. dB_{ref} was obtained via one of two methods depending on the location of a microphone in the array: (1) for those sensors located within 7.32 meter (24 feet) of the center of the array, a mean dB was computed for the inner 49 microphones in the array with dB_{ref} defined as the difference between the means for the current and baseline acquisitions, or (2) for those sensors located outside the 7.32 meter radius from the center, the mean was computed from the nearest 8 neighboring microphones, again with dB_{ref} defined as the difference between the means for the current and baseline acquisitions.
5. Each subsequent calibration produced a ΔdB_{mic} by subtracting the baseline level from the current calibration via

$$\Delta \boldsymbol{dB_{mic}} = (\boldsymbol{dB_{mic}} - \boldsymbol{dB_{ref}})_{\boldsymbol{calibration}} - (\boldsymbol{dB_{mic}} - \boldsymbol{dB_{ref}})_{\boldsymbol{baseline}} \quad \textbf{(1)}$$

A running history of the ΔdB levels was maintained to observe trends and track the microphone responses.

Health Monitoring Example: An example of array health monitoring using the aerial source can be seen in Figure 4. In the figure a total of 12 individual flights of the ACV covering a total time span of 48 days are represented and used for the health monitoring (a few outliers have been identified and removed from the figure). The particular ACV flights shown were chosen based on low ambient wind conditions and indications of good atmospheric sound propagation during a flight. Fig. 4 shows response tracking data for four microphones located on the array central plate and four microphones located on individual ground plates near the central plate, with the ΔdB levels computed according to Eqn. (1). A remarkable consistency of response is observed for the majority of these microphones over the 48-day time span, demonstrating that array health monitoring is viable using a hovering aerial sound source. Note that microphone failures will manifest as sharp or sustained deviations in observed dB levels relative to the baseline and can thus guide the array operators in determining when sensors need to be replaced during the deployment.

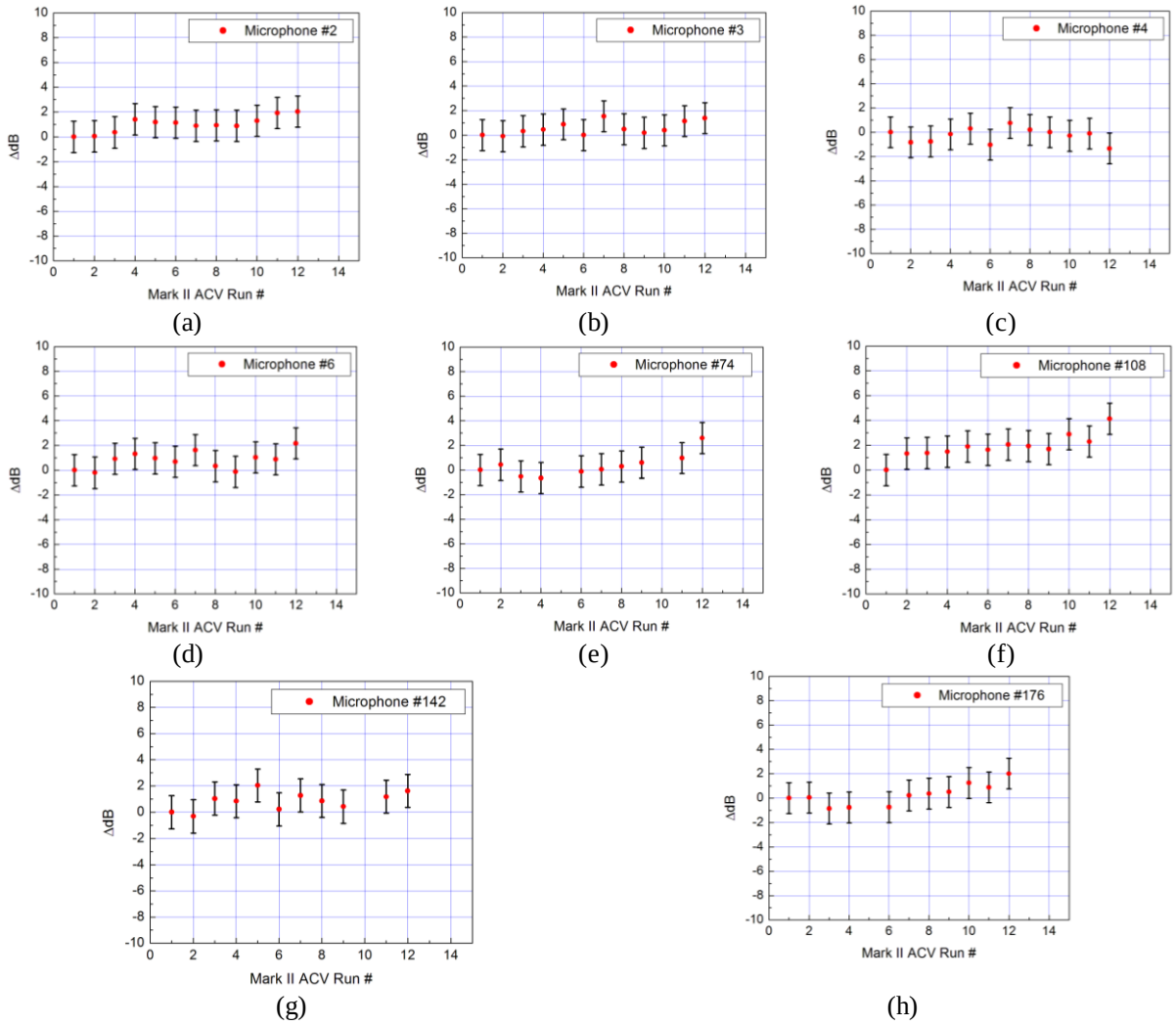


Figure 4. Tracking of measured microphone levels for 12 separate ACV runs conducted at Edwards AFB. From Reference 3.

(a) – (d) are microphones on central plate, (e) – (h) are individual inner microphones

IV. Acoustic Calibration Vehicle

The microphone health monitoring necessitated that the ACV hover for three minutes at three altitudes at 106 m, 76 m, and 45 m while emitting acoustic tones. To this end, the ACV needed to have a strong position hold capability, a high accuracy position estimate, and an acoustic system that delivered the tones. The ACV platform evolved over four deployments and replaced a previous method using a scissor lift. The ACV current configuration is shown Figure 3 and Figure 4. The technical specifications are detailed in Table 1 and the research payload is listed in Table 2. The research payload provides two unique capabilities for the phased microphone array used for acoustic airframe testing. In collaboration with the Armstrong Flight Research Center Model Shop, 44 ACV flights were completed in support of the ARM III project between March 6th, 2018 and May 3rd, 2018.

Figure 3: ACV during Functional Check Flights at NASA Langley Research Center

Figure 4: Calibrated sound source located beneath the ACV

Table 1: ACV Specifications.

sUAS Type	Multi-Rotor, 8 Brushless Motors
Diagonal Length	41.1 inches
Take-Off Weight	22 lbs
Speed	0 – 42.5 knots
Battery	6S 22,000 mAhr LiPo
Endurance	19 minutes w/20% Capacity Reserve
Command and Control	2.4 GHz Remote Control Transmitter (3 mile range)
Command and Control Telemetry	900 MHz Radio
Research Payload Telemetry	900 MHz Radio

Table 2: Research Payload Components

Commercial Component	Function
Teledyne Nano Camera	Video Data Input
Tamron 50mm Lens	Video Data Input
NVIDIA Jetson Tx2	Video Data Logging
Antcom L1/L2 GPS Antenna	Primary Position Data (RTK GPS)
GPS/Compass Module	Secondary Position Data (GPS)
Freewave MM2T 900MHz Radio	Primary Position Data (RTK GPS)
NovAtel OEM 615 Receiver	Primary Position Data (RTK GPS)
Anchor Audio AN-MINI Speaker	Calibrated Sound Source
Eclipse MP3 Player	Audio File Player
Dimension Engineering Battleswitch	Speaker Power Relay

The first capability focused on providing a calibrated sound source at a designated targeted fly-over altitude. This was important because the aircraft to be tested would need to fly over the center of the microphone array at a specific altitude. The ACV allowed the calibrated sound source in Figure 4 to be positioned over the center of the array at 350ft. AGL with an accuracy of 3m laterally, and 10m vertically. This position accuracy was based on a standard GPS position solution provided to the auto-pilot from the baseline aircraft flight system.

The second capability focused on providing an accurate position of the sound source. This was accomplished by provided two independent position solutions based on RTK GPS and machine vision. These position systems are now presented.

A. RTK Geolocation System

RTK GPS data provides high accuracy real time position data. This is accomplished by having two independent GPS receivers share satellite range data to determine their relative position to one another. This eliminates some of the sources of error since it is a relative position and both receivers will have the same errors associated with location on the Earth and atmospheric effects. When this is added to a known surveyed location of one of those GPS receivers, the other GPS receiver position now can be as accurate as the relative position as seen in Figure 5.

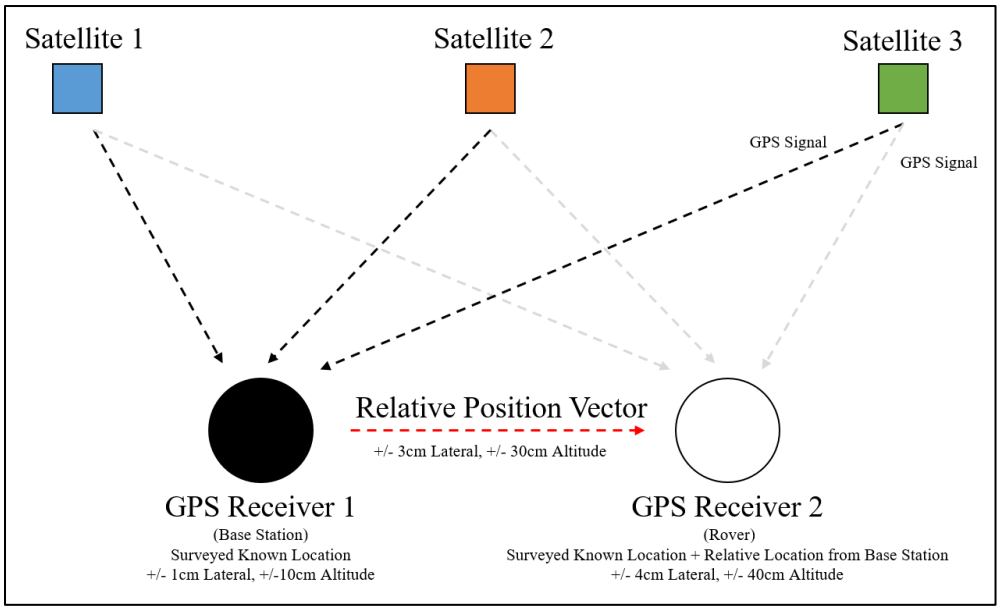


Figure 5: RTK GPS diagram

The RTK GPS approach consisted of a NovAtel OEM 615 receiver integrated into a 3D printed nylon enclosure labeled ACV Data System Enclosure as seen in Figure 6 and the base station shown in Figure 7. This enclosure housed the amplifier for the speaker, the RTK GPS receiver and a 900MHz telemetry radio.

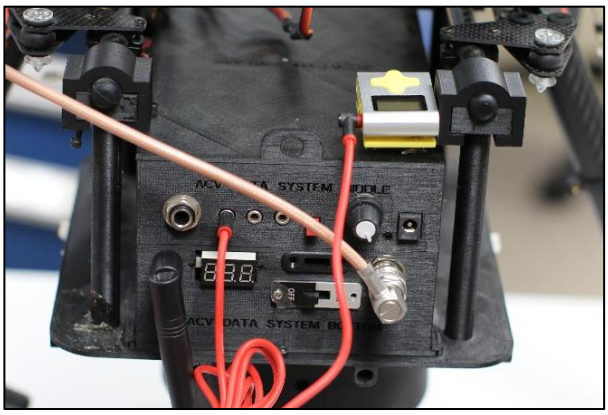


Figure 6: ACV Rover and Data System.

Figure 7: The RTK base station is mounted on a tripod and positioned over a surveyed location using a plum-bob.

B. Optical Geolocation Hardware

The camera uses a global shutter and has a 4112 by 3008 pixel resolution with a 3.45 micron pixel size CMOS sensor. The lens selected is a 50 mm prime lens supporting 12 MP with an *f*/1.8 aperture. The camera is hard-mounted to the bottom of the sUAS so that it points in the same direction as the speaker. The image frames are transmitted via Ethernet to an onboard NVIDIA Jetson TX2 that records the images onto a microSD using lossless compression.

Four visual targets are surveyed into position with mm accuracy and positioned in the microphone array shown in Figure 8. The targets are rugged polypropylene black and white interlocking tiles. Each tile has a dimension of 0.3048 by 0.3048 m (1 by 1 ft.) and each optical target has dimensions of 2.7432 by 2.7432 m (9 by 9 ft.). Geolocation accuracy was improved by surrounding the optical targets with a black border for an overall dimension of 3.3528 by 3.3528 m (11 by 11 ft.). The black border reduced the number of false positives.

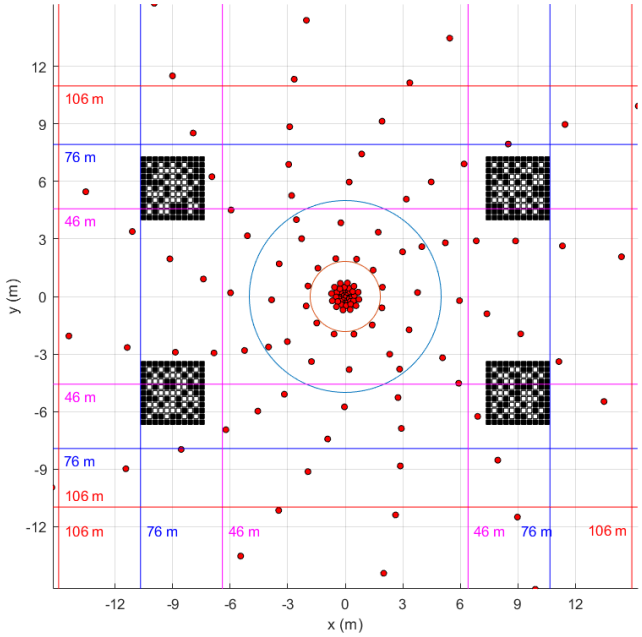


Figure 8: Aerial view of the interior of the microphone array showing the location of the optical targets. The vertical lines show the theoretical field of view for the camera at nadir orientation while centered over the array at altitudes 106 meter, 76 meters, and 46 meters.

V. Optical Geolocation Theory

A. Pinhole camera and weak perspective models

The pinhole camera model, also known as the projective camera model, is defined by an image plane and a 3D focus point, *c*, as shown in Figure 9. The pinhole camera assumes that the aperture is a point and thus ignores distortions caused by lens shape.

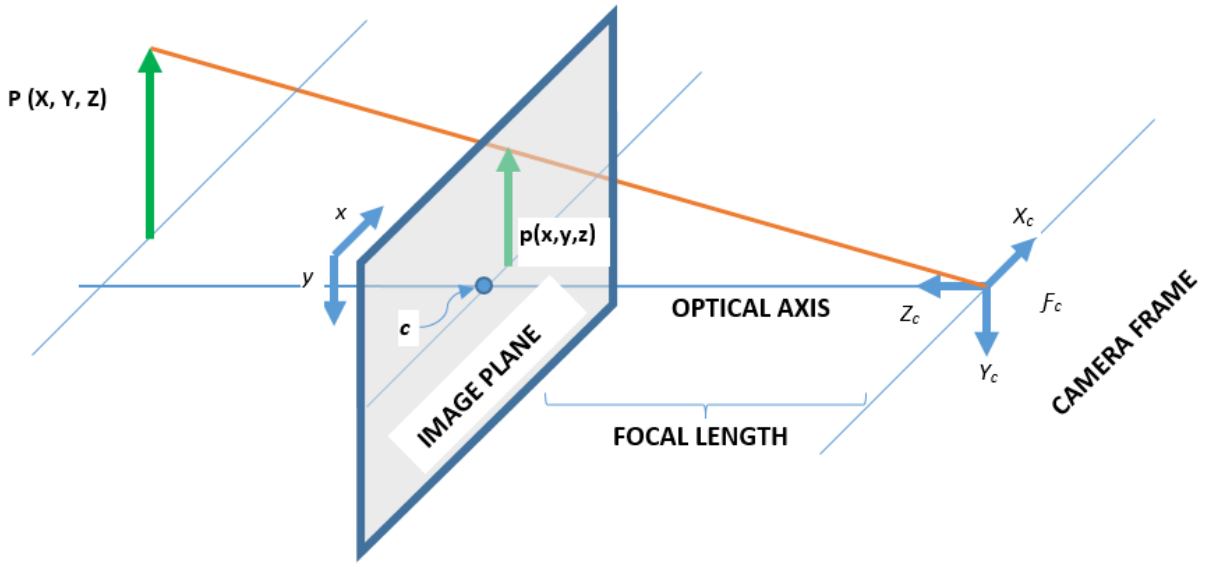


Figure 9: Pinhole Camera model.

The 3D coordinates are defined here with respect to the camera reference frame F_c in the pinhole model. The optical axis is perpendicular to the camera reference frame. The distance between the camera reference and the image planes is the focal length, *f*. Thus the image of the point $\mathbf{P} = [\, X, Y, Z \,]^T$ is given by $\mathbf{p} = [\, x, y, z \,]^T$ in the camera frame. The fundamental equations of the perspective model are written as:

$$x = f\,\frac{X}{Z} \tag{2}$$

$$y = f\,\frac{Y}{Z} \tag{3}$$

This may be simplified using the weak-perspective model by assuming the difference between any two scene points in the *Z* dimension is much smaller the average distance $\bar{Z}$.

$$x = f\,\frac{X}{Z} \approx \frac{f}{\bar{Z}}X \tag{4}$$

$$y = f\,\frac{Y}{Z} \approx \frac{f}{\bar{Z}}Y \tag{5}$$

B. Spatial resolution from weak perspective model

The spatial resolution for varying altitudes is computed by determining the pixel equivalency using the weak perspective model. Re-arranging equation 4 to solve for X to determine ground spatial resolution:

$$X \approx x\frac{\bar{Z}}{f} \tag{6}$$

X spatial resolution may be computed by using the sensor parameters:

$$X = x\frac{\bar{Z}}{f} = \frac{w_p}{N_p} * \frac{\bar{Z}}{f} = \frac{14.1864\ mm}{4112\ pixels} * \frac{\bar{Z}}{50\ mm} \tag{7}$$

Repeating for Y spatial resolution:

$$Y = y\frac{\bar{Z}}{f} = \frac{h_p}{N_p} * \frac{\bar{Z}}{f} = \frac{10.3914\ mm}{3012\ pixels} * \frac{Z}{50\ mm} \tag{8}$$

Due to the nature of spatial sampling, the spatial resolution needs to be doubled and is plotted in Figure 10

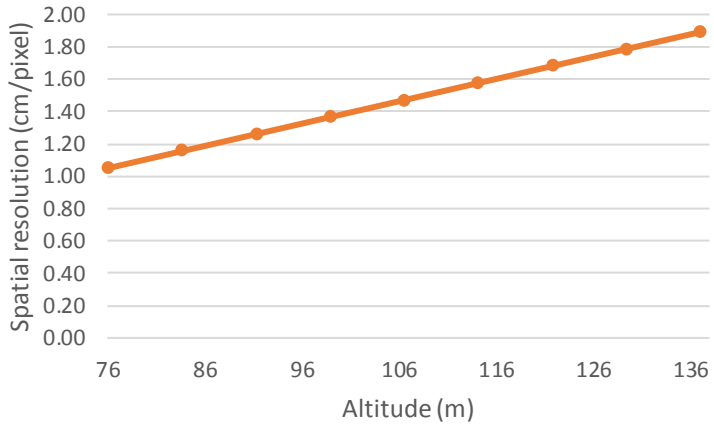


Figure 10: Optical Spatial resolution.

The spatial resolution is ~1.5 cm/pixel at the typical max altitude for the ACV is 106 meters.

C. Camera Model

The weak perspective model assumes that the camera reference frame is known, however, since the objective is to estimate the UAV position using a ground reference it is necessary to use the world reference frame. The pinhole camera model is extended for optical geolocation using intrinsic and extrinsic parameters. The intrinsic parameters are computed during camera calibration and include radial distortion, center pixel coordinates, focal length, and scaling. The extrinsic parameters include translation and rotation relative to the target. The pinhole model establishes the relationship between the 3D world-coordinates and 2D image coordinates as described by equations 9 and 10:

$$s\,m' = A[R|t]M' \tag{9}$$

where s is a scaling factor, m' is the 2D image coordinates, A is the camera matrix, R is 3D camera rotation, t is camera translation, and M' is the 3D world coordinates of m'. The pinhole camera model may also be written as:

$$s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (10)$$

where s is scaling factor, u and v are the 2D image coordinates, f_x and f_y are the intrinsic focal lengths, c_x and c_y are the optical center, r_{ii} is the 3D rotations matrix, t_i is the translation of the camera, and X,Y,Z are the 3D world coordinates of the 2D image point. Camera calibration was completed using the OpenCV library and 552 images of varying poses. The average reprojection error was 3.91e-01.

VI. Optical Geolocation Algorithm

An overview of the optical geolocation algorithm is shown in Fig. 11. The geolocation algorithm was implemented using functions from the OpenCV library [4]. The first step is to correct for lens aberrations determined from the camera calibration process. Next an image descriptor is generated using the SURF image feature detector [5], which is an extension of SIFT detector [6]. The rows in the image descriptor are the keypoints in the image (e.g. corners of optical key) while the columns contain scale and orientation information of the keypoints. Now that the keypoints and their corresponding points of interest exist, the 2D image location of the current image frame's keypoints are matched with the reference image using the FLANN [7]. Only matches that meet a thresholding criteria are kept as successful matches. This matching threshold was empirically determined as 3 times the Euclidean distance of the best keypoint match. The thresholding technique strived to eliminate bad matches generated from the microphone cabling and the paint on the runway. Initially during development of this geolocation algorithm, the 3D coordinates are immediately computed by solving equation 10 after this first round of thresholding at step 3 in Fig. 11. The rectangular coordinates used in the image frame for equation 10 are shown in the sample result shown in Figure 12, where there is an apparent error in the location of the bounding box on the bottom left. The white circles are the matched keypoints used in generating the bounding box. Many keypoints exist on the runway, indicating wrong matches exist with the reference image and lost keypoints for generating an accurate bounding box. The bounding box for the optical keys is not aligning with the bottom left corner. Also note that using an individual key results in weak bounding for the upper left two optical keys.

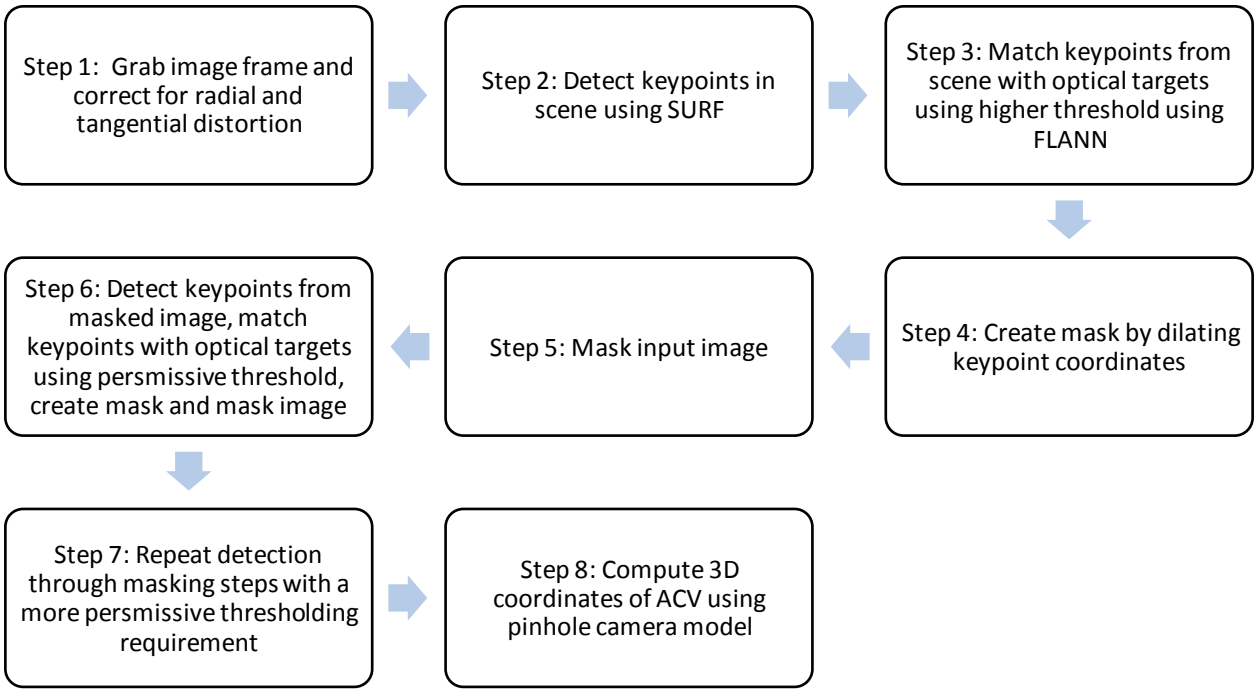


Fig. 11: Optical Geolocation Algorithm

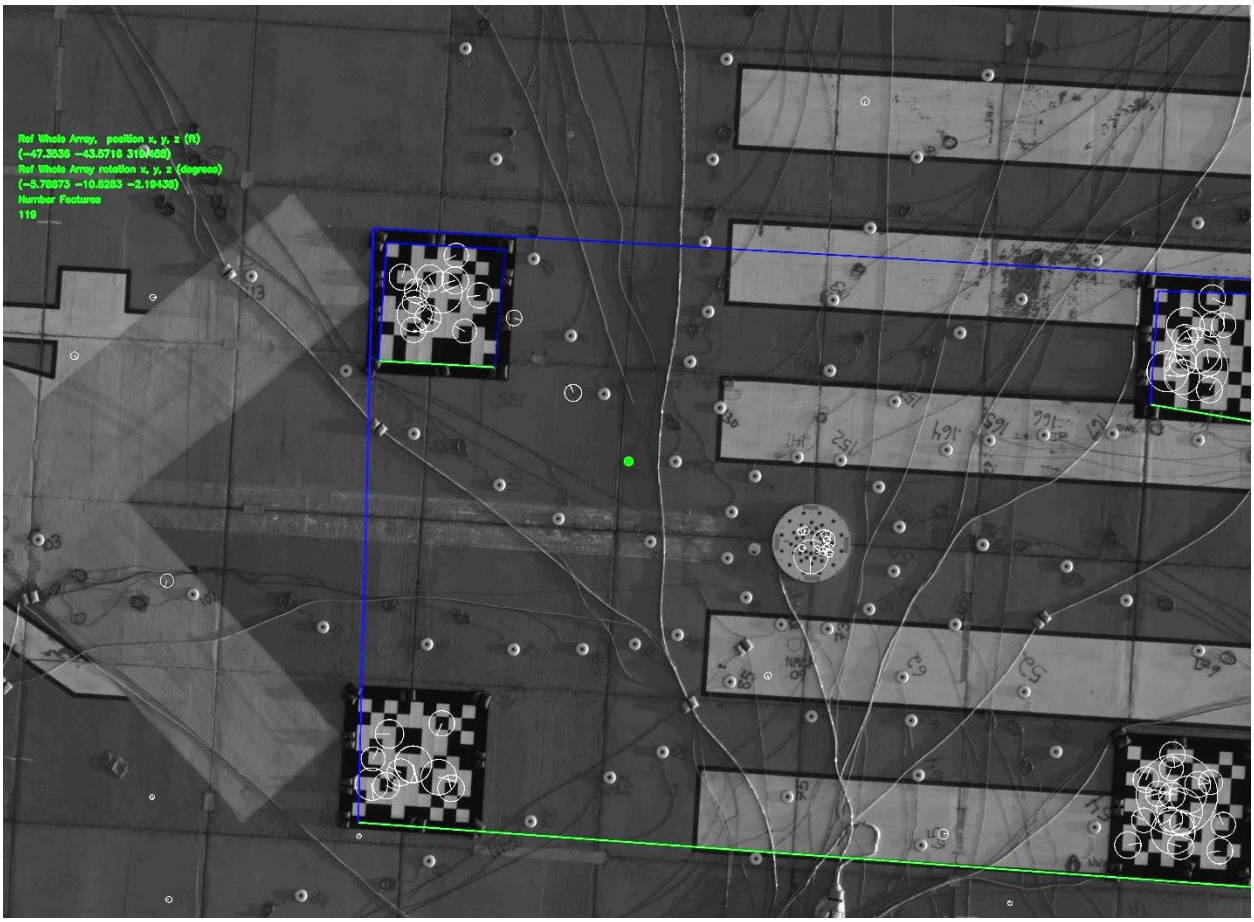


Figure 12: Optical Geolocation Result without recursive filtering.

The rejected matches result in a loss of information for the geolocation computation because position from the reference image is not being used. Equation 10 is sensitive to small errors in image space and therefore, a technique for increasing the number of keypoints for matching was developed. Instead of computing 3D position, a mask is created by dilating the matches from the current image using a circle with a radius of 300 pixels. The threshold here is set to 5 times the Euclidean distance of the best keypoint match, allowing more keypoints to be used in the geolocation calculation while also allowing for a more constrained mask of 200 pixels in the next iteration. The result for one masking iteration is shown in Figure 13.

After the final masking, an empirical threshold value of 0.34 is used for matching. This threshold is generally more permissive than previous threshold values and passes more matches for the generation of the rectangular bounding box. 244 matched points are utilized in creating the bounding box for the final geolocation calculation as shown in Figure 14.

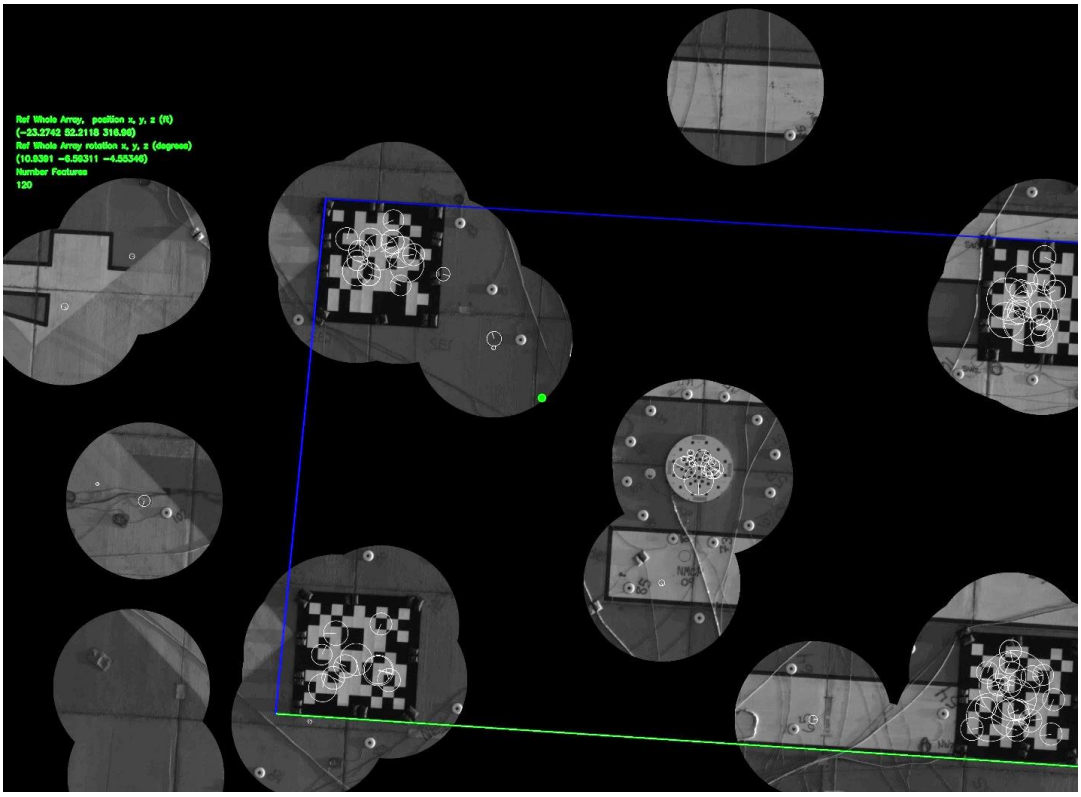


Figure 13: Geolocation result after masking the input image and generating new keypoints.

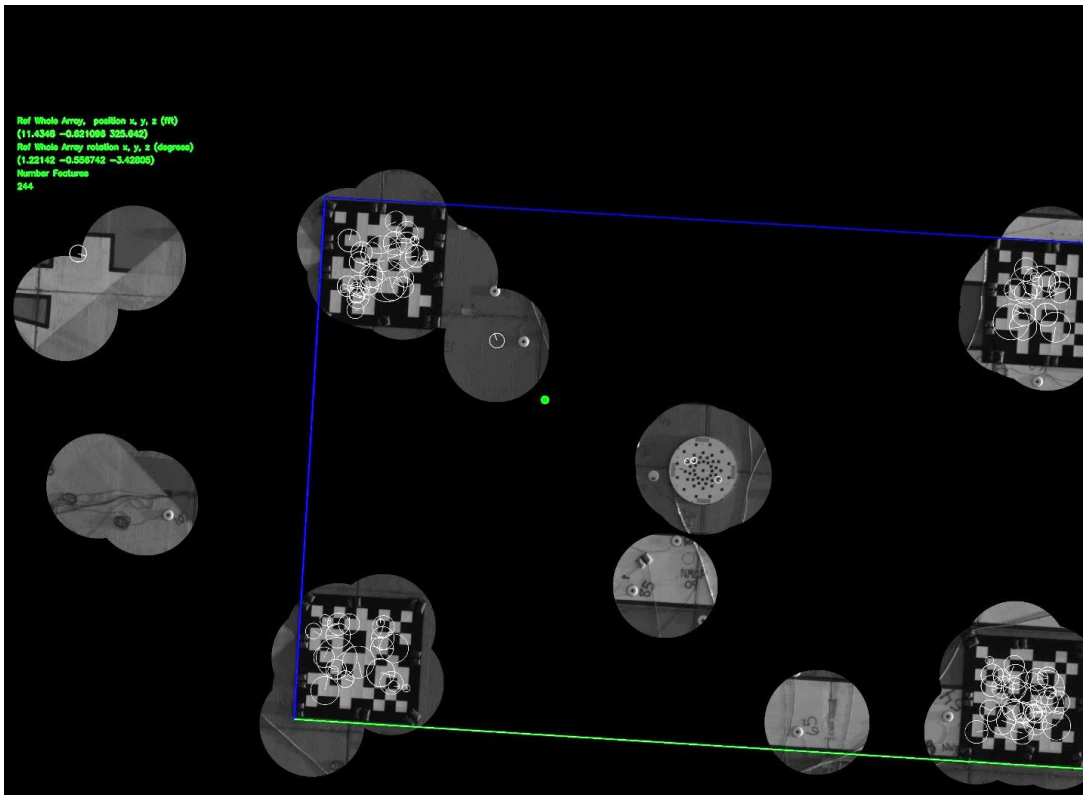


Figure 14: Geolocation results after final masking and with permissive keypoint threshold.

VII. Optical Geolocation Results

Two sample flights are presented below in detailed analysis from civil mean twilight (CMT) and midmorning of the same day. The midmorning flight was windier and gustier as shown by the wind plots in Figure 15 and Figure 16: Wind conditions for Midmorning Flight. Generally windier and gustier at the midmorning flight then the CMT flight. Figure 17 and Figure 18 show the raw optical position estimate, median filtered optical result, and the RTK result. The differences between the median filter optical geolocation and RTK computations are shown in Figure 19 and Figure 20.

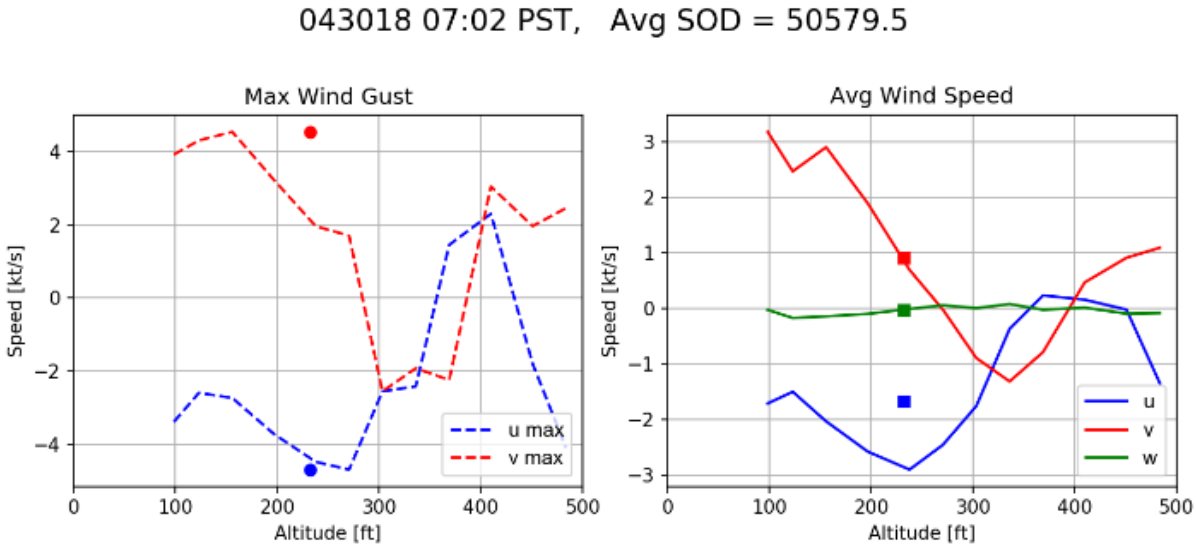


Figure 15: Wind Conditions for CMT Flight.

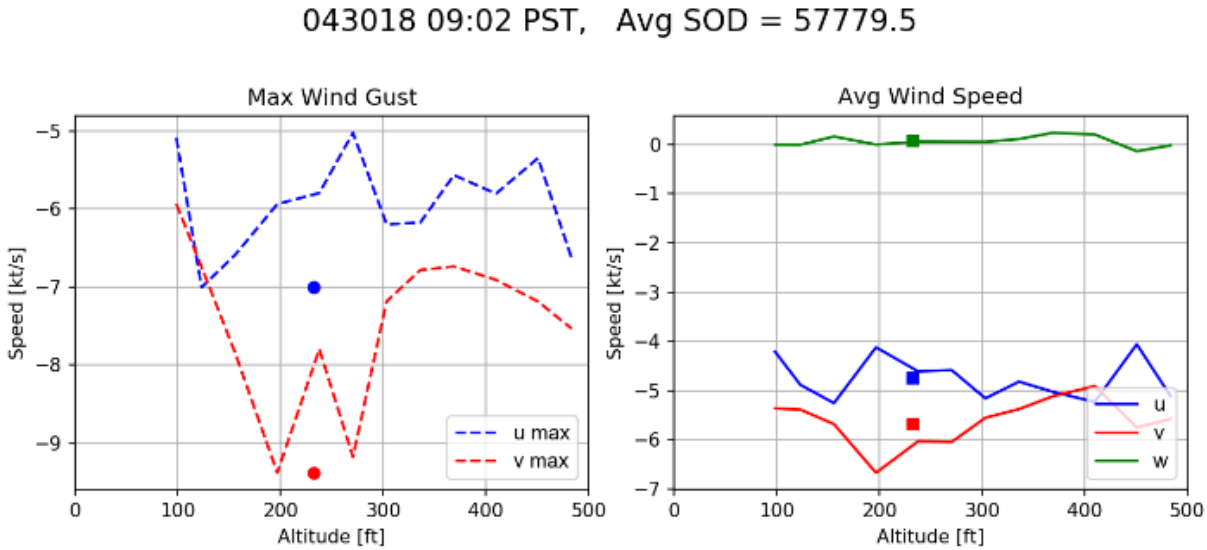


Figure 16: Wind conditions for Midmorning Flight.

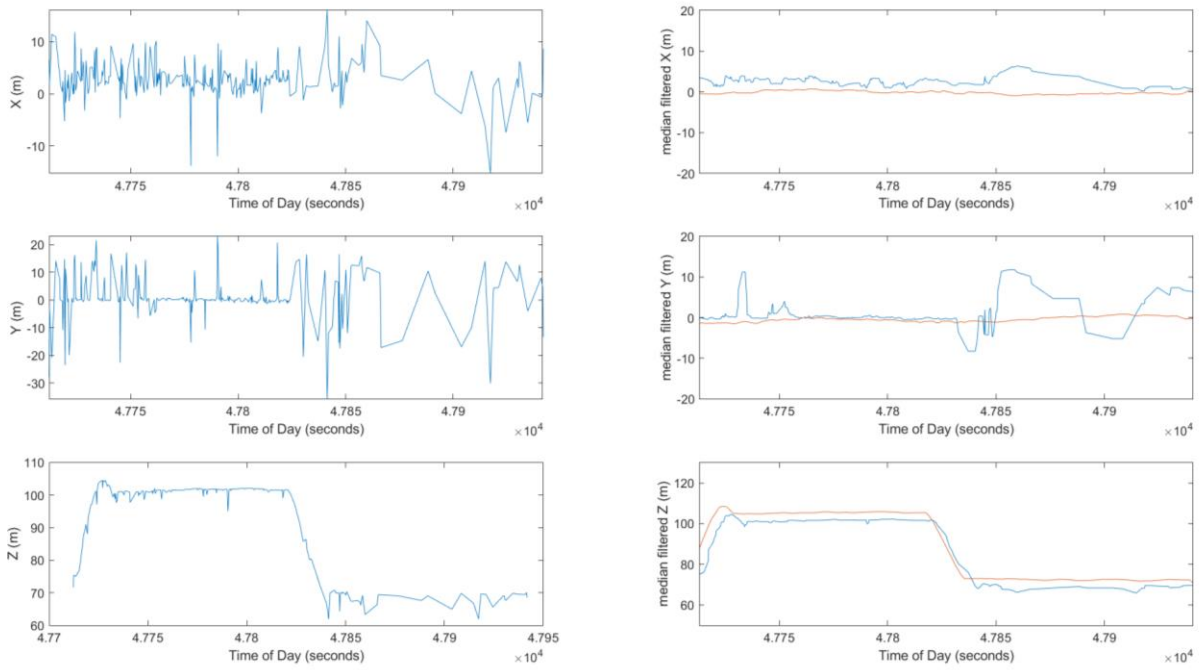


Figure 17: CMT Available. Left shows the raw unfiltered in optical computation for position and right shows the optical result filtered (blue) using median filtering with RTK result (orange).

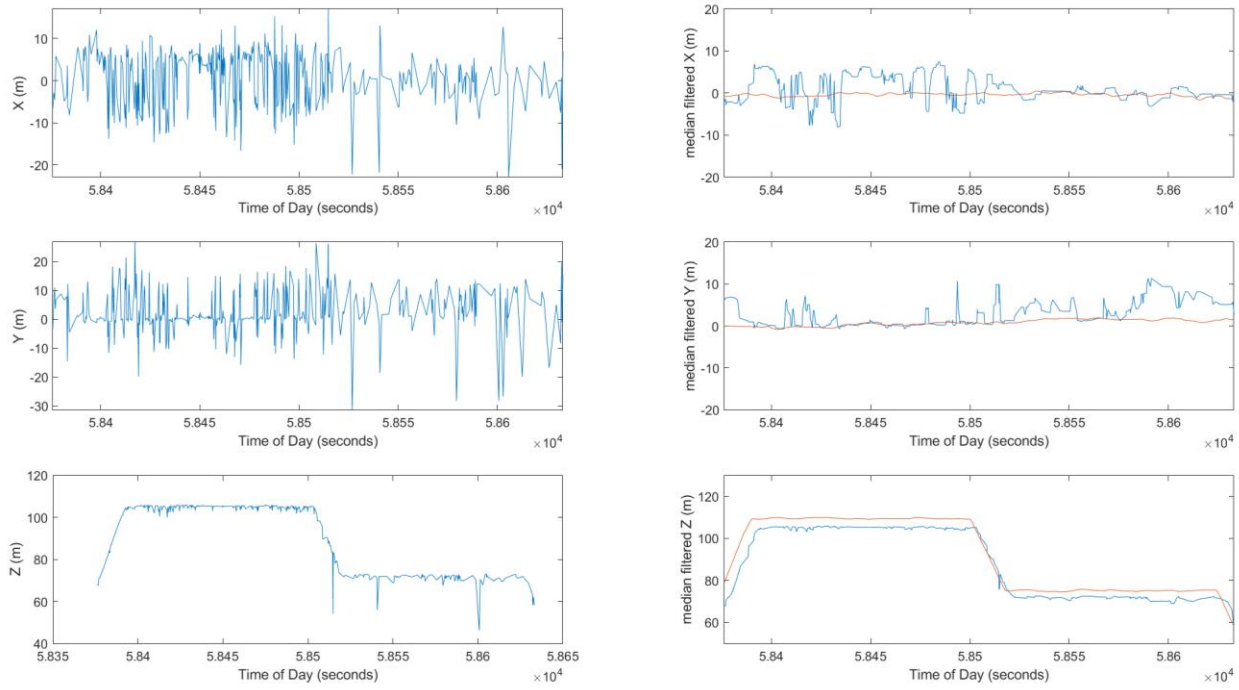


Figure 18: Midmorning Available. Left shows the raw unfiltered in optical computation for position and right shows the optical result filtered (blue) using median filtering with RTK result (orange).

The raw geolocation result shows substantial noise in the plots on the left in Figure 17 and Figure 18. The geolocation calculations are sensitive to small errors in detection of the keys. To improve geolocation calculations, a

median filter is used to suppress high frequency changes in the 3D positon estimates as shown in the plots on the right in Figure 17 and Figure 18. The difference between RTK and optical geolocation 3D coordinates is shown in Figure 19 and Figure 20 with the RTK standard deviation and attitude data from the flight controller. The RTK standard deviation is below 0.25 meters for these two flights.

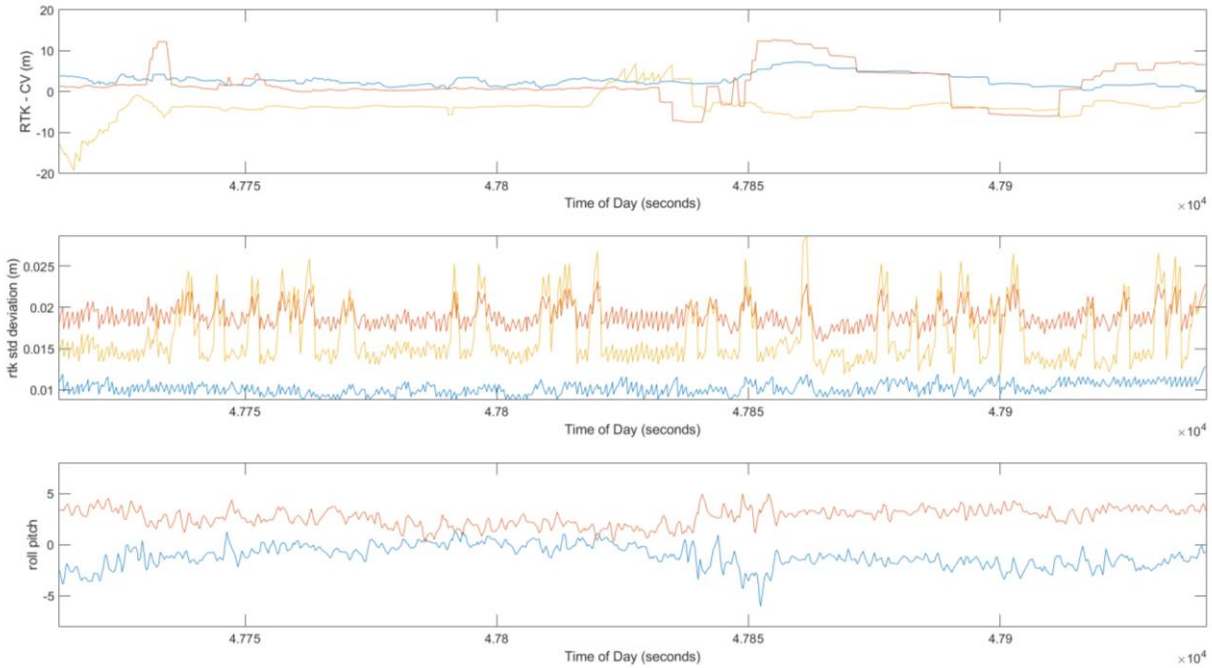


Figure 19: CMT all available keys 3D Position Plot. X direction is blue, Y is red, and Z yellow, roll is blue, and pitch is red.

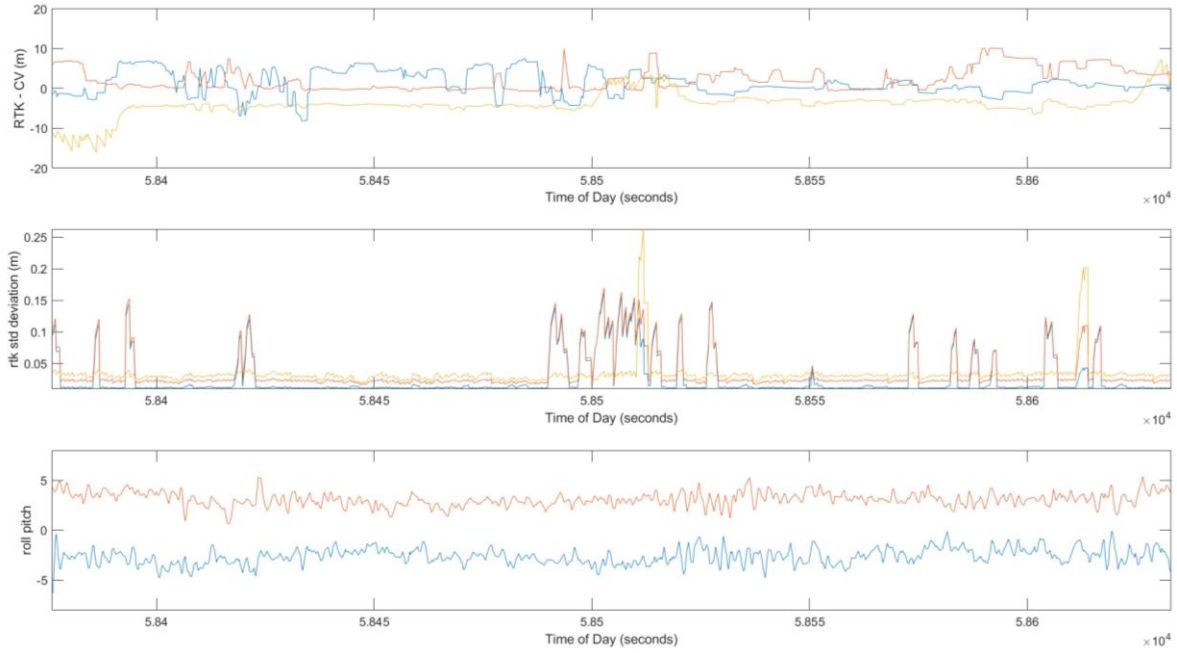


Figure 20: Midmorning all available keys 3D Position Plot. X direction is blue, Y is red, and Z yellow, roll is blue, and pitch is red.

The noise of the 3D position signal in Figure 19 and Figure 20 is largely generated when one or more optical keys exit the FOV. Figure 21 and Figure 22 are generated by only computed 3D position using frames that have all four optical keys completely in the FOV. The four key 3D position plot is substantially more stable.

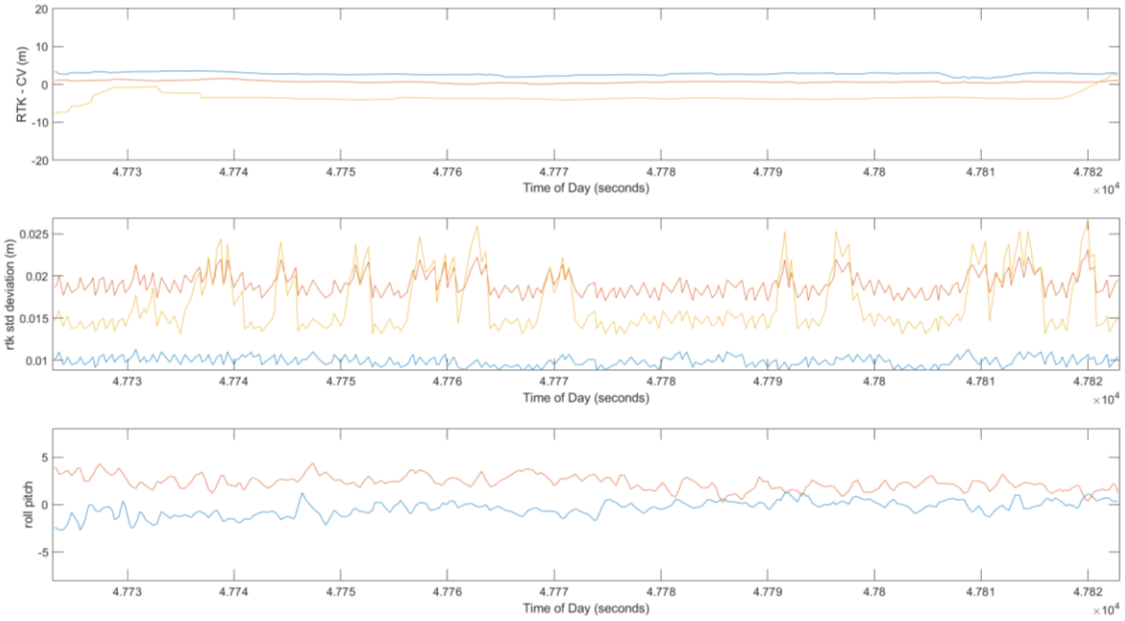


Figure 21: CMT four key 3D position plot. X direction is blue, Y is red, and Z yellow, roll is blue, and pitch is red.

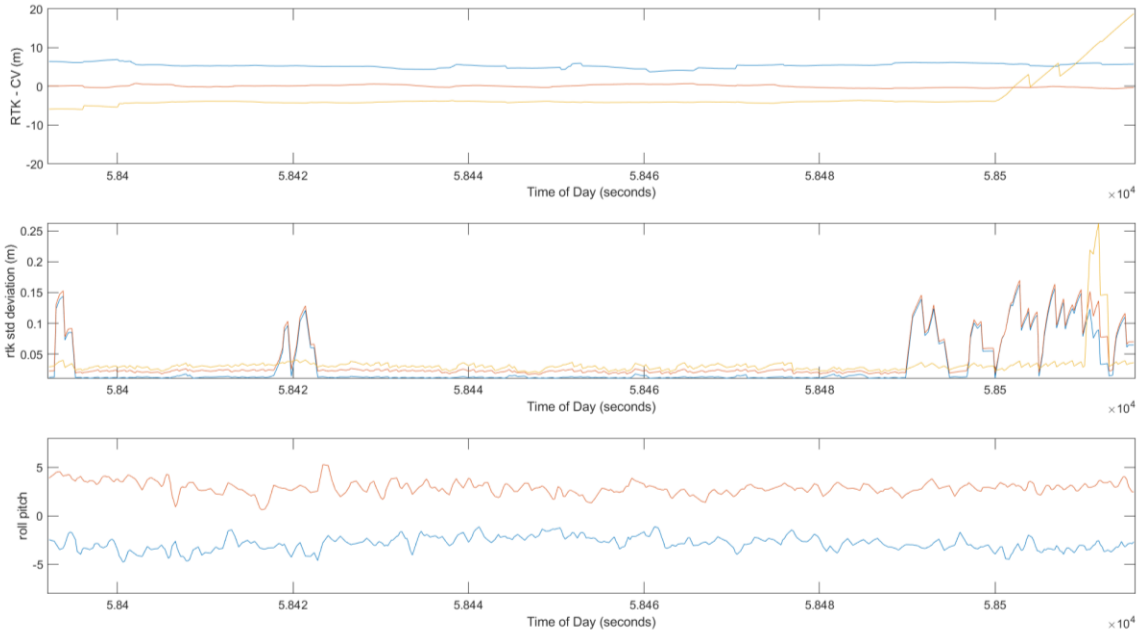


Figure 22: Midmorning four key 3D position plot. X direction is blue, Y is red, and Z yellow, roll is blue, and pitch is red.

Table 3 and Table 4 below show statistical data for the CMT and Midmorning flights. The input frames are subcategorized by the number of complete targets in the image to assess the contribution of the optical keys. The available category uses all targets present in the imagery for the geolocation calculation. Statistical results indicate noisier position estimates when using fewer optical keys as given by the standard deviation of the difference between the optical and RTK position estimates for both flights. The standard deviation of the difference between optical and rtk increases in the midmorning results relative to CMT, possibly due to increased wind during the second flight as presented in Figure 15. Also note that higher wind conditions existed at the altitude of 250 meters and that three or more optical keys were less often in the FOV for the 250 meter altitude calculations.

Table 3: Statistical Analysis of CMT Flight.

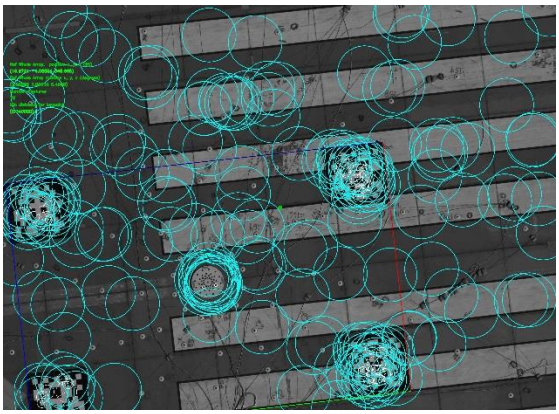
		Average Difference between Optical and RTK Position			Standard Deviation between Optical and RTK Position			Number of Image Frames
		X (m)	Y (m)	Z (m)	X (m)	Y (m)	Z (m)	
Raw	Four Keys	2.64	0.78	-3.27	1.60	0.97	1.94	62
	Three Keys	1.91	0.58	-3.86	2.61	2.26	1.29	64
	Two Keys	3.05	3.52	-4.67	3.65	8.50	3.44	108
	One Key	3.74	3.67	-2.49	2.94	5.63	2.57	7
	Available	2.82	1.61	-4.11	3.56	7.76	2.80	291
Median Filtered	Four Keys	2.66	0.64	-3.25	0.48	0.28	2.02	62
	Three Keys	2.09	0.62	-3.82	0.59	0.30	1.58	64
	Two Keys	2.64	2.65	-4.43	1.01	3.18	3.54	108
	One Key	2.81	1.12	-0.56	0.21	0.41	5.82	7
	Available	2.53	1.61	-3.91	1.17	3.03	2.84	291

Table 4: Statistical Analysis of Midmorning Flight.

		Average Difference between Optical and RTK Position			Standard Deviation between Optical and RTK Position			Number of Image Frames
		X (m)	Y (m)	Z (m)	X (m)	Y (m)	Z (m)	
Raw	Four Keys	5.10	0.27	-3.99	2.31	0.93	1.41	43
	Three Keys	2.71	0.96	-4.34	6.48	6.32	1.86	197
	Two Keys	-2.31	3.96	-4.52	6.52	8.22	3.56	108
	One Key	-0.60	2.31	-4.05	5.52	9.58	4.12	234
	Available	0.99	1.95	-4.24	6.47	7.51	2.99	455
Median Filtered	Four Keys	5.52	0.13	-3.92	0.79	0.31	1.86	43
	Three Keys	4.49	0.27	-4.24	2.06	1.11	2.28	197
	Two Keys	-3.22	4.11	-4.10	3.06	3.44	3.18	108
	One Key	-0.16	3.76	-3.49	1.66	2.70	3.56	106
	Available	1.74	1.55	-3.97	3.46	2.40	2.49	455

A sample four image sequence is shown in Figure 23 to help explain the noise levels in the raw 3D position plots and corresponding raw standard deviation results in Table 3 and Table 4. Initially, three complete optical keys are in the FOV with portions of the bottom left optical key visible (a). In (b), 0.604 seconds later there are only two complete optical keys available with portions of the other keys present. The x accuracy improves, the y accuracy decreases,

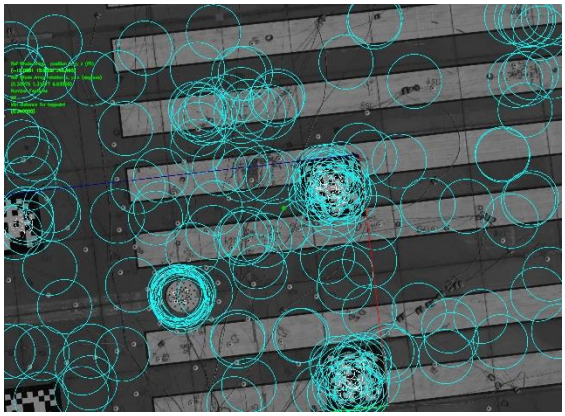
and the z accuracy remains roughly the same at 0.604 seconds. Combining the magnitude of the X and Y accuracy indicates that the X-Z accuracy remained approximately the same at ~7.9 m in (a) and (b). At 1.8 seconds from the image in (a), there are only two keys present in the image as shown in (c) and all three position accuracies decrease sharply using only two optical keys. Accuracies at 2.105 seconds are similar to 1.8 seconds and are shown in (d). During this time sequence, the ACV is under the influence of wind and thus flight controller corrects for the change in position via change of motor rpm, resulting in a change of attitude. The change of attitude impacts the optical approach substantially at ~15 meters change in x and y estimates over the image sequence while the change in RTK position estimate is ~10 cm.



Time= 0.000 seconds

	X (m)	Y (m)	Z (m)
RTK	-0.89	-0.36	109.55
unfiltered optical	6.09	-1.25	105.66
Difference	-6.98	0.88	3.89

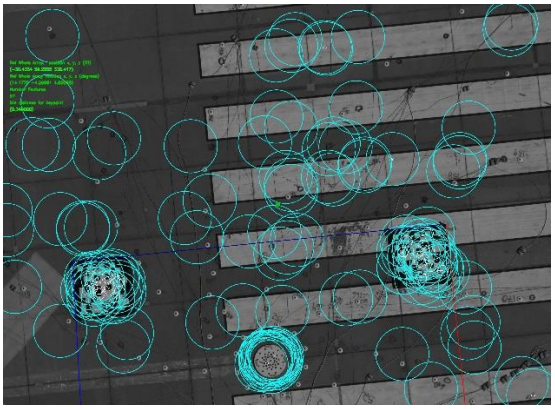
(a)



Time = 0.604 seconds

	X (m)	Y (m)	Z (m)
RTK	-0.90	-0.28	109.50
unfiltered optical	-3.68	4.85	105.60
difference	2.78	-5.13	3.90

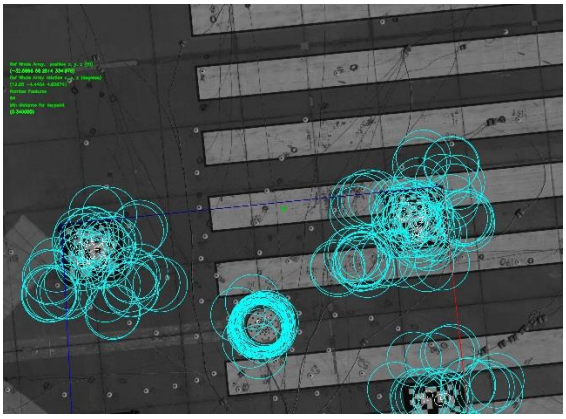
(b)



Time = 1.800 seconds

	X (m)	Y (m)	Z (m)
RTK	-0.93	-0.28	109.43
unfiltered optical	-9.28	17.76	102.54
difference	8.35	-18.03	6.89

(c)



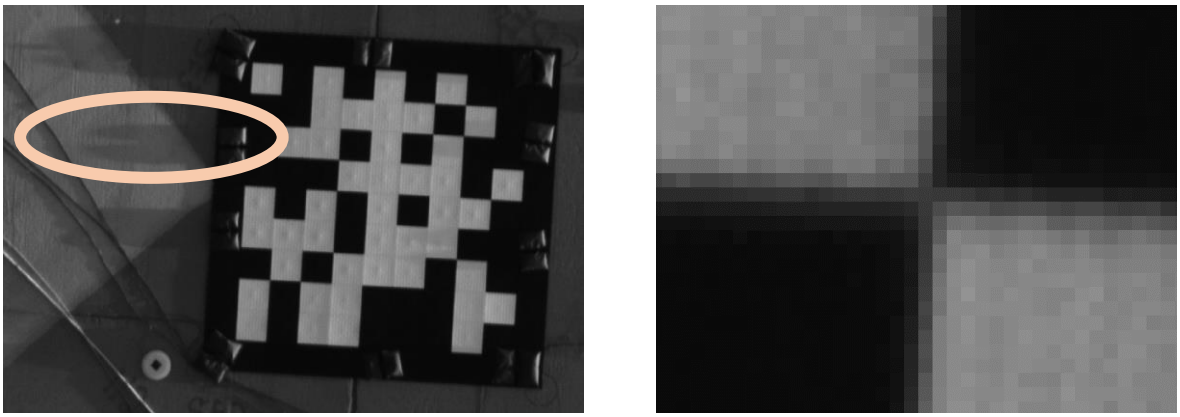
Time = 2.105 seconds

	X (m)	Y (m)	Z (m)
RTK	-0.93	-0.27	109.42
unfiltered optical	9.93	17.15	102.10
difference	-10.85	-17.42	7.31

(d)

Figure 23: Sample four image sequence over 2 seconds and corresponding optical computations. (a) and (b) have portions of the four optical keys in the images. (c) and (d) only have two complete optical keys in the images. The blue circles show the mask from the previous layer.

When examining the case of four keys in the FOV for the CMT and midmorning flights, the median filtered X standard deviation between Optical and RTK Position are 0.48 m and 0.79 m while the Y median filtered standard deviation between Optical and RTK Position are 0.28m and 0.31 m. The theoretical spatial resolution is 0.015 m, thus the results obtained are between one and two orders of magnitude higher than the theoretical accuracies. Despite using a high exposure time, there is a shadowing in the CMT flight and motion blur within the image frame as shown in Figure 24.



(a) (b)

Figure 24: Sample zoom in with key from an image taken at an altitude of 106m. (a) shows the extent of the shadow from a sandbag on the optical key in the oval and (b) shows that the lines on the blocktiles are not crisp. Interpolation helped with placing the keypoint in the correct location but may have contributed to decrease geolocation accuracy.

VIII. Conclusion

A novel optical geolocation algorithm was presented and evaluated for sUAS altitudes of 106 and 76 meters. A single optical key was substantially less stable for geolocation calculations using the pipeline discussed. The most stable 3D position results were obtained when four optical keys were in the FOV. Low lighting conditions and wind conditions may have resulted in motion blur in the imagery, resulting in decreased quality of the optical geolocation computation. A narrower FOV prime lens may have increased accuracy provided the optical targets fit within the phased array without interfering with microphone functionality. Framerate for this work was bounded by the write-speed of the microSD. Geolocation accuracy would improve with a higher frame rate due to increased filtering performance.

IX. Future Work

The optical geolocation work here suffered when the optical targets were not within the FOV of the camera. Instead of four 9 by 9 optical keys used in this work, a greater number of smaller optical targets such as the 4 by 4 keys in Figure 25 would provide an improved geolocation computation because the keypoints are more distributed in the FOV of the camera. The method in [8] maximizes the hamming distance of the binary bits used in the targets for improved and faster recognition of the targets.

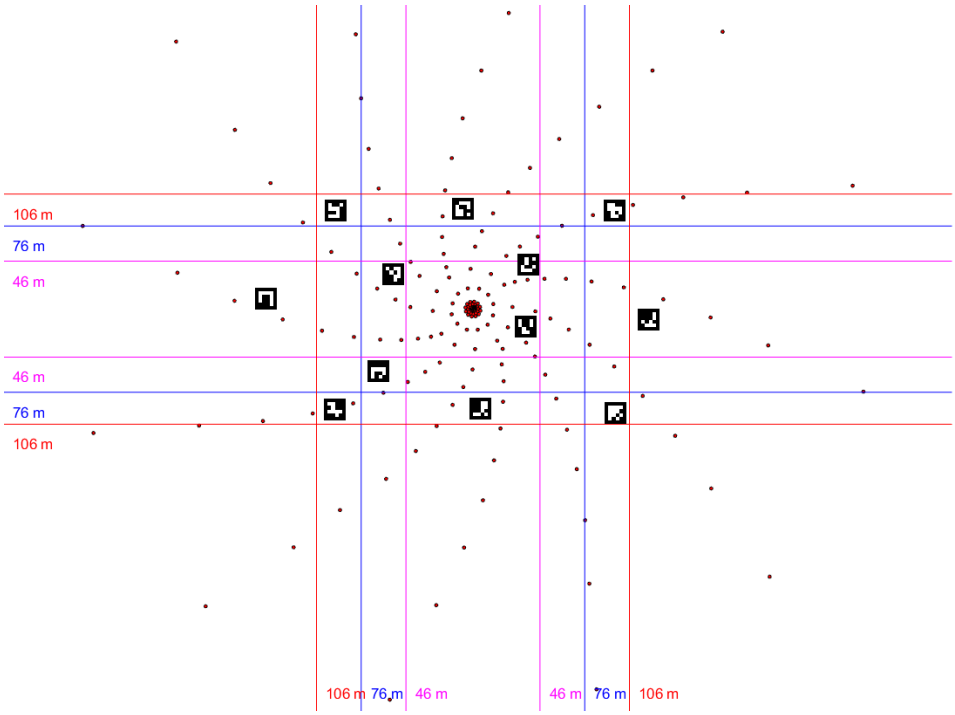


Figure 25: The microphone array with twelve 4 by 4 Aruco tags.

Acknowledgements

We would like to thank NASA Armstrong project support of Erin Waggoner, Claudia Herrera, and Kevin Weinert, Jacobs project support of Angel Guilloty, sUAS pilots A.J. Jaffe and Alex Flock, and model shop members Derek Abramson and Red Jensen.

X. References

[1] W. M. Humphreys, Jr., D. P. Lockard, M. R. Khorrami, W. G. Culliton, R. G. McSwain, P. A. Ravetta and Z. R. Johns, "Development and Calibration of a Field-Deployable Microphone Phased Array for Propulsion and Airframe Noise Flyover Measurements," in *AIAA/CEAS Aeroacoustics Conference*, Lyon, France, 2016.

[2] W. M. Humphreys, Jr., D. P. Lockard, M. R. Khorrami, W. G. Culliton and R. G. McSwain, "Evaluation of Methods for In-Situ Calibration of Field-Deployable Microphone Phased Arrays," in *AIAA/CEAS Aeroacoustics Conference*, Denver, 2017.

[3] W. G. Culliton and W. M. Humphreys, Jr., "Networked Array Recorder (NeAR) Microphones for Field-Deployed Phased Arrays," in *AIAA Aviation Conference*, Dallas, TX, 2019.

[4] OpenCV, Open Source Computer Vision, 13 February 2018. [Online]. Available: https://docs.opencv.org/3.4.1/index.html.

[5] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.

[6] E. Rosten and T. Drummond, "Machine learning for highspeed corner detection," in *European Conference on Computer Vision*, Graz, 2006.

[7] M. Muja and D. G. Lowe, "Fast approximate nearest neighbors with automatic algorithm configuration (2009)," in *VISAPP International Conference on Computer Vision Theory and Applications*, Lisboa, 2009.

[8] F. Romero-Ramirez, R. Munoz-Salinas and R. Medina-Carnicer, "Speeded Up Detection of Squared Fiducial Markers," *Image and Vision Computing*, vol. 76, pp. 38-47, 2018.