Detection and Tracking of Aircraft in the Far-Field from Small Unmanned Aerial Systems

Chester V. Dolph,[1]
NASA Langley Research Center, Hampton, VA, 23681, USA

Cyrus Minwalla,[2]
Scoped Systems, Ottawa, ON, K2G 7A5, Canada

Louis J. Glaab[3], Michael J. Logan[4], B. Danette Allen,[5]
NASA Langley Research Center, Hampton, VA, 23681, USA

and
Khan M. Iftekharuddin[6]
Old Dominion University, Norfolk, VA, 23529, USA

Onboard far-field aircraft detection is needed for safe non-cooperative traffic mitigation in autonomous small Unmanned Aerial System (sUAS) operations. Machine vision systems, based on standard optics and visible light detectors, possess the ideal size, weight, and power (SWaP) requirements for sUAS. This work presents the design and analysis of a novel aircraft detection and tracking pipeline based on optical sensing alone. Key contributions of the work include a refined range inequality model based on sensing and detection with FAA well-clear separation assurance distances between aircraft in mind, a detector fusion method to maximize the benefit of two image detectors, and a comparative analysis of Linear Kalman-filtering and Extended Kalman-filtering to seek optimal tracking performance. The pipeline is evaluated offline against multiple intruder platforms, using two types of flight encounters: multirotor sUAS vs. fixed-wing sUAS and multirotor sUAS vs. general aviation (GA) plane. Analysis is restricted to the rate-limiting head-on and <mark>departing collision volume</mark> cases vertically separated for safety. Results indicate that it is feasible to use the proposed optical spatial-temporal tracking algorithm to provide adequate alerting time to prevent penetration of well-clear separation volumes for both sUAS and GA aircraft.

---

[1] Aerospace Engineer, Aeronautics Systems Engineering Branch, AIAA Member.
[2] Principal, Scoped Systems.
[3] Assistant Branch Head, Aeronautics Systems Engineering Branch.
[4] Aerospace Engineer, Aeronautics Systems Engineering Branch, AIAA Associate Fellow.
[5] Senior Technologist for Intelligent Flight Systems, AIAA Associate Fellow.
[6] Associate Dean for Graduate Research and Graduate Programs.

Nomenclature

d_{pix} = pixel pitch

D_{GTS} = Number of Ground Truth Detected in Sortie

D_S = Total Number of Detections in Sortie

D_{UGTS} = Number of Unique Ground Truth Detected in Sortie

f = focal length

GT_S = Total Number of Ground Truth Labels in Sortie

R_{DAA} = detect and avoid range

R_{detect} = detecting range

R_{evade} = evading maneuver range

R_0 = initial detection range

R_{warn} = warning range

t_{avoid} = time to compute the avoidance maneuver

t_{DAA} = time to complete detect and avoid maneuver

t_{detect} = time of detection

t_{evade} = time to complete evading maneuver

t_0 = initial detection time

T_{UGTS} = number of Unique Ground Truth Tracked in Sortie

x,y = pixel coordinates

X,Y,Z = global coordinates

v_{close} = aircraft closing velocity

I. Introduction

THE rapid expansion of small Unmanned Aerial System (sUAS) commercial applications such as package delivery, aerial inspection, and emergency response has created a need for a safe and on-board traffic avoidance systems [1]. The number of private (or "model") sUAS and commercial sUAS in the United States are expected to reach 1.39 to 1.59 million and 0.60 to 1.20 million by 2024, respectively [2]. sUAS traffic management is a challenging multi-faceted problem because of low altitudes and the density of shared airspace for general aviation, sUAS, and

recreational vehicles operated by the general public. The problem is further exacerbated by commercial interests such as package delivery, commercial videography, and performance art, all leading to a high density of sUAS operating in a low altitude urban environment. To this end, the Federal Aviation Administration (FAA) and National Aeronautics and Space Administration (NASA) are working together, in partnership with private industry, to develop a Concept of Operations for UAS Traffic Management (UTM) that focuses on traffic management operating below 400 ft.

Multiple approaches of Detect and Avoid (DAA) methods are being investigated for sUAS traffic management, such as, wireless communication links [3], radar [4], infrared sensors [5], and vision systems [6]. Each technique has strengths and weaknesses. On-board wireless communication will suffer from signal dropout in dense urban environments where line-of-sight may be disrupted by skyscrapers along with GPS signal. Ground-based wireless communication systems (e.g., cellular networks) may have too much latency for safety-critical applications of sUAS missions. Radar sensors tend to be heavy relative to sUAS payloads. Infrared and optical sensors require line-of-sight and produce degraded information in weather conditions such as fog or cloud cover.

This work proposes a novel method of detection and tracking for low-altitude far-field sUAS and General Aviation (GA) aircraft using image-differencing and morphological detectors, which are evaluated against a well-clear definition. The detections are used as input to a Kalman filter for tracking of intruder aircraft. To the best of our knowledge, there is no work in the literature that has integrated the difference detector and horizon detector for an application as in this paper. The pipeline is evaluated for two cases: 1) fixed-wing sUAS vs. multirotor camera carrier, and 2) GA SR22 vs. multirotor camera carrier. This paper is organized as follows: Section II present background summary of alternate sUAS detection techniques and prior work completed at NASA, Section III provides experimental design describing aircraft, hardware, flight paths, and the detection and tracking pipeline, Section IV discusses results, Section V provides conclusions, and Section VI describes future work.

II.   Background

Despite two decades of research, DAA for sUAS remains an open area of research. Optics-based DAA systems using cameras and computer vision techniques receive continued attention in the literature. The authors in [7] compared optical based tracking systems to pulsed radar in manned aircraft DAA using 16 head-on encounter scenarios with varying horizontal separation distances, vertical separation distances, and weather conditions. Extended Kalman Filter (EKF) tracking was employed throughout. Results indicated that 1) radar had decreased sensitivity to changes in weather conditions, 2) Electro-Optical (EO) systems had superior range and sensitivity in favorable

conditions, 3) first detection distances for both radar and EO system varied substantially across the encounter scenarios, which they attributed to the variations in flight encounter approaches and weather conditions, and 4) EO systems detected the intruder before the radar system in 6 of the 16 encounter scenarios. While this work demonstrates the feasibility of using onboard sensing for DAA or GA intruder and GA ownship encounters, it does not investigate DAA problem for sUAS multirotor vs. sUAS fixed-wing or sUAS vs. manned aircraft encounters, which is needed to ensure safe autonomous operations.

The authors in [8] propose that electro-optical sensors are well-suited for DAA due to their SWaP properties. Their DAA system was mounted in a ScanEagle (military grade sUAS) and the intruder aircraft was a Cessna 172R. The DAA system operated in real-time, where the detection algorithm used GPS/INS based image stabilization to mitigate for aircraft vibration, wind gusts, and turbulence. Point features were extracted from the stabilized imagery using bottom-hat morphological filter, and then Hidden Markov Models (HMM) were applied for temporal filtering. A total of 22 head-on collision-course encounters were conducted, with average detection ranges of 1.48 km for a 5mm lens and 2.14 km for an 8 mm lens, which corresponded to 19.2 and 27.7 seconds until collision respectively for the 150 knots closing speed. These detection results were expected given the angular resolution of the system. Greater ranges would be possible with longer focal length lenses at the expense of field of view. An altitude separation of 500 ft was used in the experiments and the intruder aircraft was always flown above the horizon. The large altitude separation used in this work is inadequate for evaluating sUAS vs. other aircraft encounters because it does not closely model a collision geometry.

The authors in [9] collected video data of manned aircraft for three clear and four cloudy days for head-on and tail-chase scenarios. Their method achieved no false detections in 6.6 hours out of 14.1 hours of flight data. Two distinct approaches to temporal processing were employed: Viterbi-based filtering and HMM for object detection. However, the technique uses GA vs. GA encounters with much larger vertical separation of 152 meters to satisfy safety need for GA vs. GA flight-ops.

The authors in [10] proposed separate detection strategies for above- and below- horizon targets. For above the horizon cases, morphological close-minus-open filters were used, whereas differencing was used below the horizon, in both cases to extract keypoints followed by an EKF for inter-frame target tracking. The pipeline was tested on manned aircraft in 4 head-on and 6 crossing scenarios. The first detection ranged from 1350 to 2960 meters and their algorithm achieved tracking at distances ranging from 980 to 2588 meters. The authors found that the morphological feature detector generated many false positives when used below the horizon and the difference pipeline achieved

poor range performance. Data collected onboard sUAS is needed to develop onboard sUAS collision avoidance systems because they have different maneuvering capabilities, vibrational environment impacting sensors, more dynamic pitching, and tend to operate at lower altitudes with closer proximity to clutter.

The authors in [11] applied a deep learning technique on data set acquired from a delta wing sUAS imaging up to 8 sUAS simultaneously, with 1920×1080 or 1280×960 resolution cameras without describing the programmable focal length setting for the fisheye lenses. Detection consisted of image differencing followed by a Shi-Tomasi corner extractor on images stabilized by an inter-frame homography calculation. Classification was performed via a convolutional neural network on image patches extracted around the target. This deep-learning approach was not developed with consideration to an avoidance model, which is needed to meet an aviation authority's definition of well-clear to keep aircraft safely separated.

The work in [12] surveyed 8 different feature detectors applied to imagery of a Harvard Mark IV trainer from a Bell 206 helicopter ownship equipped with three 5 MP cameras and analyzed with respect to true positive rate, false positive rate, earliest detection range, and average run-time per frame based on server with 256 GB DRAM, SSD, and two Intel Xeon 16 core 2.2 GHz CPUs. For head-on, the SIFT feature detector performed with the highest true positive rate, Shi-Tomasi with the lowest false positive rate, and the FAST feature detector with maximum detection range of 2.0 km. For a 10-degree offset collision geometry, the FAST feature detector performed with the highest true positive rate and maximum detection range, and the Harris corner detector performed with the lowest false detection rate. The FAST detector had the fastest average run-time per frame by two orders of magnitude at 0.002 seconds. This work does not analyze the low-altitude environment and does not include sUAS intruder aircraft.

The NASA prior work in [6] utilized fisheye lens with 4k action cameras 170º FOV from aerial sUAS vs. sUAS detection and that dataset is available for download here [13]. Camera carrier sUASs included multirotor and fixed-wing sUAS. The intruder sUAS included multirotor sUAS and two types of fixed-wing sUAS. The intruder sUAS was not visible for a large portion of this dataset due to wide-angle lens on the action camera and the testing environment aimed at detecting of the intruder at distances over 1000 feet. An image differencing approach utilizing generated limited detection results. Most detections occurred when the intruder sUAS was performing a high-profile maneuver above the horizon and near the minimum separation distance. The work in [14] reports a high sUAS detection rate with a very high false positive rate for one of the videos using the NASA dataset. Similar flight tests were repeated using 41-degree FOV lenses and results demonstrated substantial improvement with detection in [15].

III. Contributions

This work presents an improved NASA ==machine vision== detection and tracking pipeline developed in [9] [11] and [15] with the introduction of fusing the morphological and difference detections, and the implementation of an Extended Kalman Filter (EKF). In addition, robust analysis is conducted to test the efficacy of the pipeline based on a range inequality model. Furthermore, sensitivity analysis on the detection and tracking parameters is conducted. Unlike much of the prior vision-based detection and tracking work in sUAS DAA, this work tests both head-on collision geometry for GA and sUAS-fixed wing from ownship sUAS multirotor with onboard sensors, evaluates the detection and tracking system using well-clear separation distances, and systems level performance metrics described in the experimental methodology section. This work tests for head-on collision geometries at low-altitudes and long ranges (3.5 to 1 km), which is a typical sUAS operating environment. ==The head-on collision geometries are the worst-case scenario from a collision sensing perspective and also where image-based trackers are at their worst because changes in image space are very small.==

IV. Techniques

A. Range Inequality Models

A set of minimum operational performance standards (MOPS) issued by a civil aviation authority is necessary to establish a minimum performance requirement for detection of aircraft in a DAA system. At present, no civil aviation regulatory agency has adopted or published a performance requirement for electro-optical detection systems. This work proposes a quantitative approach towards establishing requirements and bounds on the detection subsystem of DAA. The concept, originally defined in [16], is extended here in multiple ways.
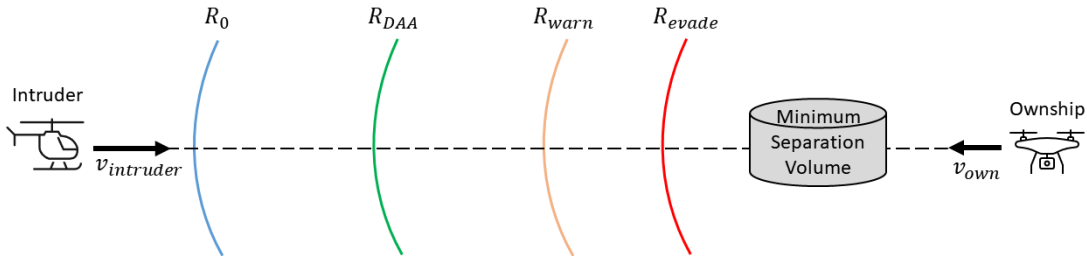

Fig. 1 Range Inequality Model.

Fig. 1 defines four ranges as critical distance metrics, namely R_0, R_{DAA}, R_{warn} and R_{evade}. These ranges (or distances) are linked by an inequality relationship (Eq. 1), which quantitatively ties the design characteristic of the sensing system (R_0) to the performance of the detection algorithm (R_{DAA}) in relation to the performance of the avoidance algorithm (R_{warn} and R_{evade}). The lower bound to the DAA algorithm is defined by the time at which a warning must be issued to the pilot in command or autopilot system (R_{warn}), which must happen before a collision maneuver can be safely initiated (R_{evade}) given as,

$$R_0 \geq R_{DAA} \geq R_{warn} \geq R_{evade}. \tag{1}$$

In Eq. 1, R_0 is the principal design parameter for the sensing system driven by angular and range resolution parameters of a given sensor. It is typically limited by system transfer functions including sensor point-spread effects, platform motion, and atmospheric effects such as attenuation and scattering. For an electro-optical system, R_0 is a function of the angular resolution and target contrast. The former is derived from the lens focal length and pixel pitch of the detector array, while the latter is limited by optical and environmental transfer functions, including atmospheric effects such as photon extinction and scattering at the range limit. R_{DAA} is the performance characteristic of the DAA system. Specifically, R_{DAA} is the range at which a target is reported as a collision-course intruder by the processing algorithm and a suitable maneuver is calculated by the avoidance system. Specifically, R_{DAA} is defined as follows (Eq. 2),

$$R_{DAA} = R_0 - v_{close} \cdot (t_{detect} + t_{avoid}). \tag{2}$$

Here, v_{close} is the closing rate between ownship and intruder, t_{detect} is the time required by the detect and track algorithm to produce a collision course intersection with sufficient confidence, and t_{avoid} is the time required by the avoidance algorithm to select an optimal maneuver to successfully evade a collision. v_{close} is a function of the collision geometry, determined by ownship and intruder headings and airspeeds. There exists an inherent tradeoff between t_{detect} and t_{avoid} as the more time dedicated to the detection algorithm means less time to compute a suitable avoidance solution. A fast algorithm may reduce t_{detect} at the expense of reduced track precision, track accuracy, and a higher number of false tracks, which may prove detrimental to the efficacy of the avoid solution. Therefore, careful tradeoffs must be made in a given DAA pipeline to balance accuracy with speed.

R_{evade} is the range at which a maneuver must commence if the ownship is to successfully avoid the collision, and the corollary t_{evade} is the time interval between the start of the maneuver and the closest point of approach. R_{evade} and t_{evade} are dependent on the ownship's aerodynamic performance envelope and the minimum separation distance requirement. R_{warn} introduces a buffer around R_{evade} to account for communication delays to the ground station and pilot in command. In 2015, the Transport Canada Working Group recommended an $t_{warn} = 2 \cdot t_{evade} + 15$ for typical UAS operations in Canadian airspace. However, an optimal maneuver may be automatically be undertaken at any point between R_{warn} and R_{evade} without violating the inequality, therefore the value for R_{warn} is driven entirely by operational considerations. By factoring in v_{close} for a particular collision geometry, one can define ranges that establish lower bounds on the DAA system (Eq. 3),

$$R_{evade} = v_{close} \cdot t_{evade}, \qquad R_{warn} = v_{close} \cdot (2 \cdot t_{evade} + 15). \qquad (3)$$

Given the definitions of R_{warn} and R_{evade}, the spatial and temporal inequalities are denoted as follows (Eq. 4),

$$R_{DAA} \geq R_{warn} \geq R_{evade}$$

$$= R_0 - v_{close} \cdot t_{DAA} \geq v_{close} \cdot (2 \cdot t_{evade} + 15 \text{ sec}) \geq v_{close} \cdot t_{evade};$$

$$\therefore t_0 - t_{DAA} \geq 2 \cdot t_{evade} + 15 \text{ sec} \geq t_{evade}, \qquad \text{where } t_0 = \frac{R_0}{v_{close}}. \qquad (4)$$

In summary, the model presented herein quantitatively links the collision geometry, the DAA system's and the ownship's aerodynamic performance envelope. For the DAA system in particular, the upper bound is established by R_0 and the lower bound by R_{warn}, captured temporally in Eq. 4. Any proposed DAA system may be benchmarked by computing t_{detect} and t_{avoid}, and ensuring that the inequalities (Eq. 4) are satisfied.

B.   Computer Vision Detection Pipeline

The detection and tracking pipeline, shown in Fig. 2, enables detection of different objects above the horizon. This pipeline extracts objects using image contrast and shape, creates tracks and assigns detections to existing tracks, and uses thresholds to extract good tracks. The main steps are to extract the location of the horizon, extract objects using

morphological or difference detector, cluster detections, track using Kalman filtering, and use thresholds to establish good tracks. Each of these components are described in detail the subsections below. <mark>The detection and tracking stages update at the camera frame rate of 30Hz.</mark>
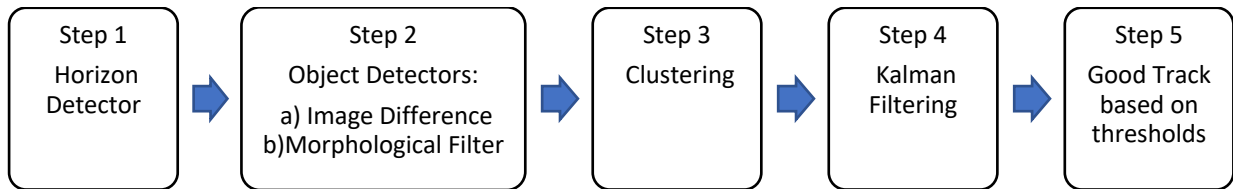
Step 1 Horizon Detector	→	Step 2 Object Detectors: a) Image Difference b)Morphological Filter	→	Step 3 Clustering	→	Step 4 Kalman Filtering	→	Step 5 Good Track based on thresholds

Fig. 2 Detection and Track Pipeline.

1. Horizon Detector

A gradient-based approach was used to extract the horizon contour from collected imagery. Given that the camera was mounted to the ownship gimbal mount, and moved independently from the aircraft frame, attitude information from the flight controller could not be used to stabilize imagery. The technique operates as follows: First, the input image is converted from color to grayscale, Otsu's binary thresholding method and Canny Edge detector are utilized to extract the horizon contour [17] [18] by creating a below-horizon mask. The mask is then dilated with a gaussian blur 5x5 kernel and then eroded with a 100 x100 kernel to reduce jagged edges in the mask and reduce lines parallel to the horizon from terrain and clouds. Finally, the horizon is reduced by changing the edge of the mask from a jagged contour to a 2D line using the average vertical positions of the left and right sides of the bisected mask: Point 1 (left image horizontal midpoint, left image average vertical position) and Point 2 (right image horizontal midpoint, right image average vertical position). This ensemble of horizon reduction techniques overcame the challenges from low-lying clouds, haze, multiple tree-lines, fields, and roads that were parallel to horizon.

The proposed technique was adequate for horizon segmentation in representative datasets consisting of relatively flat terrain but would not be considered suitable for flight operations over terrain containing roads, rivers, mountain ranges, and other features that may generate long contours unrelated to the horizon location. Future flight test plans include the integration of an Inertial Navigation System (INS) mounted rigidly to a calibrated camera system, allowing the horizon to be directly computed from the attitude information and projected into the captured imagery.

2. Object Detectors

a) Image Differencing Detector

The image differencing method subtracts the previous frame from the current frame as shown in Fig. 3. First the contrast in the image is enhanced using Contrast Limited Adaptive Histogram Equalization (CLAHE) [19] and then decreased using a gaussian blur to suppress intensity variations unrelated to objects. Next, image keypoints are detected using the FAST feature detector for the previous frame [20]. Next, the keypoints detected in the previous frame are localized in the current frame via a pyramidal implementation of Lucas Kanade Optical Flow tracking model [21]. The image is stabilized by computing the homography between the previous frame and the current frame as a perspective transformation, and then applied to all pixels in the previous frame. Once the image is stabilized, the absolute difference is computed between the current frame and the transformed previous frame.
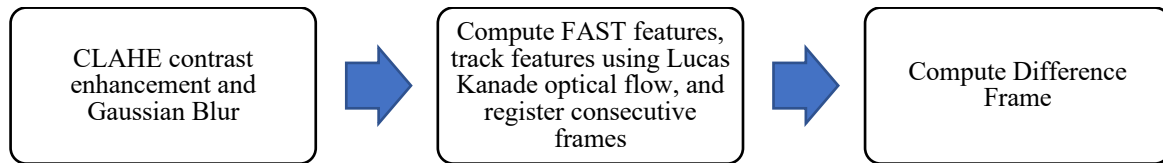
| CLAHE contrast enhancement and Gaussian Blur | → | Compute FAST features, track features using Lucas Kanade optical flow, and register consecutive frames | → | Compute Difference Frame |

Fig. 3 Image Differencing Based Detection.

The difference detections are extracted using the FAST Feature detector. Differenced images are cluttered with minor changes in the scene and thus a technique for removal of stationary objects is required. Moving objects have higher contrast than non-moving objects in the difference image, thus the FAST feature detector is able to extract them from the background. The FAST threshold value is initialized to 10 and incremented until the number of keypoints is less than the max difference detection threshold per that frame

b) Morphological sUAS Detector

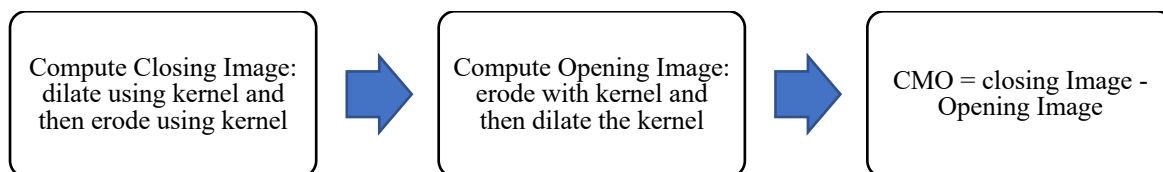| Compute Closing Image: dilate using kernel and then erode using kernel | → | Compute Opening Image: erode with kernel and then dilate the kernel | → | CMO = closing Image - Opening Image |

Fig. 4 Morphological Based Detection.

The morphological sUAS detector, depicted in Fig. 4, is a close-minus-open transform. Note that filter processing is restricted to above-horizon region of interest in this work. The filter is computed as follows: First, the above the horizon region is converted to greyscale, then a closing image is computed, where closing consists of a dilation step followed by an erosion step with a known filter kernel. In this work, the cross shaped kernel developed in [10] is used and shown here:

$$K = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \tag{5}$$

Next, the opening image is computed in an erosion step followed by a dilation step. The opening filter reduces image noise. Finally, the difference image is computed by subtracting the closing image from the opening image. Detections are denoted as high-contrast features and extracted by the FAST feature detector employing the dynamic thresholding technique described previously for the differencing detector.

This morphological filtering approach works well for objects on low textured backgrounds, e.g. clear sky, as the close-minus-open process generates high contrast for the aircraft against the background.

3. Clustering

Detections from individual morphological and difference detectors were combined with the fused morphological/difference detector using the following clustering technique. Detection centroids are clustered as one detection if they are within a radius of 10 pixels of each of other, which is equivalent to clustering detections within 0.107° azimuth and 0.148° elevation with this camera configuration

4. Tracking

For a given frame, each detection is either assigned to a new track or an existing track by generating a cost matrix and then determining whether to assign the current detection to an existing track or create a new track. The cost matrix is defined as the Euclidean distance between each detection centroid and each existing track prediction; thus, the cost matrix dimensions are the number of existing tracks by the number of current frame detections. The Munkres [22] implementation of the Hungarian assignment algorithm [23] assigns detections to a track with a cost of non-assignment of 10. The Kalman filter updates the predicted location for each existing track. The filter assumes a constant velocity with empirically determined parameters: an initial estimate uncertainty variance of 10 x 10 pixels, motion noise of 20 x 20 pixels, and measurement noise of 20 pixels.

Each track is described by the current location in pixel coordinates, Kalman Filter (KF) parameters, age (number of frames since initial detection), total number of frames where the object was detected, and total number of consecutive invisible counts. An Extended Kalman Filter (EKF) assuming constant velocity is also tested and a comparative analysis between KF and EKF is shown in the Results section.

5. Good Track

A high-confidence track is defined as having at least 30 detections with a maximum of 5 consecutive dropouts (1/6th of a second) and with a visibility greater than 0.6, where visibility is total visible count divided by the track age count [15]. Once a track meeting the threshold requirements is identified, it would be reported to a collision avoidance system. This work provides detailed analysis on these high-confidence tracks in the results section.

6. Sample Results

Sample pipeline results are shown in Fig. 5. The morphological detector generated the results shown in the left column, the difference detector in the center column, and the centroids from the detectors are shown in the right column where yellow is the morphological detector result, red is the difference detector result, and green indicates an overlap of the difference and morphological detector. The Close-minus-open process enlarges the intruder aircraft in image space beyond the gaussian blurring described the pre-processing stage for the morphological detector. The difference detector generates silhouetting as the homography matrix for the interframe change is not perfect.

—	Morphological Detector	Difference Detector	Raw Image
GA Sortie 4			
GA Sortie 10			
Tempest Sortie 3			
Tempest Sortie 6			

Fig. 5: Sample Detection Segmentation Results: Morphological Detector, Difference Detector, and the respective Detector centroids on the raw image where yellow is the morphological detection result, red is the difference detector result, and green indicates that the morphological and difference detector provide the same centroid coordinates within precision of a pixel.

C. Benchmarking

The pipeline's processing rate must correlate to the DAA system performance requirement as established by t_{detect} in t_{DAA}. Note that t_{detect} is bounded by the amount of time required to develop a solid track. Since the current pipeline is designed for offline processing, it operates at a rate of 1 frame per second. Furthermore, the tracking stage in the present pipeline requires a minimum of 30 hits to establish a track with sufficient confidence. The real-time version is expected to be 5 to 10x faster, pending implementation and benchmarking. For convenience, a conservative estimate of a real-time variant operating at 5 frames/sec is reasonable, leading to a $t_{detect} = 6\,\textbf{sec}$. Assuming that the avoidance algorithm takes an equivalent amount of time, i.e. $t_{avoid} = 6\,\textbf{sec}$, the total DAA processing time is established at $t_{DAA} = 12\,\textbf{sec}$.

The benchmarking implementation is as follows. Detection was implemented in C++ using OpenCV on a desktop with dual Intel 5220 CPU x 18 cores with 256 GB DDR4 memory and two NVIDIA 2080 Ti GPU. The tracking pipeline was completed on a laptop in MATLAB with a quadcore i7-6829HQ CPU with 32 GB memory. The path to onboard DAA starts with design using Minimum Operational Performance Standards (MOPS), next selecting a camera, developing an algorithm-based data collected in flight, finalizing the algorithm after revisions to camera and flight-testing design, and finally developing dedicated flight hardware. The dedicated flight hardware will likely use Application Specific Integrated (ASIC) because it maximizes computational performance while minimizing Size Weight and Power (SWAP). A conservative expectation for algorithm processing time improvement is 10x to 100x when switching to a dedicated ASIC.

V. Experimental Methodology

A. Description of Experiments

The flight tests were completed in support of the NASA UTM project where different sensors were evaluated for onboard collision avoidance using Detect and Avoid Alerting Logic for Unmanned Systems (DAIDALUS) [24]. These sensors include Automatic Dependent Surveillance–Broadcast (ADS-B), RADAR, and Dedicated Short-Range Communications (DSRC) [3]. A multirotor is used as the ownship for UTM flight testing with fixed wing sUAS and GA aircraft as the intruder aircraft at Beaverdam Airpark in Elberton, Virginia, USA. Flights were recorded over two days in autumn of 2018. This airfield is surrounded by a tree farm with neighboring agricultural fields on a coastal plain. DAIDALUS did not execute an avoidance maneuver for the data used in this work and the range inequality

model proposed in IV.A is used for offline analysis of the vision system. The following subsections describe the aircraft, camera, collision avoidance objectives, flight test methodologies, dataset, and the evaluation methods.

B. Aircraft

The aircraft used in this work are shown in Table 1. The fixed-wing intruder sUAS is a tempest. Aircraft like the tempest sUAS may be used for LIDAR mapping after natural disasters or for small package delivery such as medical supplies [25] as they have flight time of approximately an hour at 20 m/s and can reach altitudes over 1000 feet. The tempest was equipped with a Pixhawk for these flight operations. The GA plane selected in this work is the NASA SR22 with tail number NASA 501. The SR22 is a good example of intruder aircraft that an sUAS may encounter because it is representative of size and shape of commonly flown GA planes. The SR22 is equipped with a Status 2S, which logged the GPS coordinates used in this work. The ownship multirotor is a DJI S1000 and is named: ISAAC [26]: ICAROUS Sense and Avoid Characterization, where ICAROUS is Independent Configurable Architecture for Reliable Operations of Unmanned Systems [27]. ISAAC was equipped with a Pixhawk flight controller.

Table 1: Aircraft.

Aircraft	Type	Image	Dimensions w x l x h (m)	Manufacture	Ground Speed (m/s)	Max Takeoff Weight (kg)
Tempest (intruder)	fixed wing sUAS		3.2 x 1 x 0.3	UASUSA	18	20
SR22 (intruder)	GA		11.68 x 7.92 x 2.72	Cirrus	53	1,633
ISAAC (ownship)	sUAS multirotor		1 x 1 x 0.3	DJI	10	11

C. Camera

A Sony™ FDR x1000V action camcorder equipped with an 8.2 MP image sensor and retrofitted with an 8.25 mm 41° Horizontal Field of View (HFOV) by 32 Vertical Field Of View (VFOV) lens was mounted to the ownship for this work. The camera was equipped with a rolling shutter sensor and is shown in **Fig. 6** and specifications in Table 2. Given that the ownship was a flying multirotor and subject to vibrations, rolling-shutter artifacts were captured in the imagery, degrading image quality.

Table 2: Camera Specifications.

Specification	Measurement
Resolution	3840 x 2160
Horizontal Field of View	41°
Vertical Field of View	32 °
Pixels Per Degree	93.7
Approximate Mass	408g

Fig. 6 Sony Action Camera with 8.25 mm lens. The hot glue on the enclosure secured the focus.

D. Range Inequality Sensor Performance

An estimate of R_0 is required to estimate the upper bound of the collision avoidance model. The model may be constrained within the pinhole camera approximation and to Gaussian optics. 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 7. The pinhole camera assumes that the aperture is a point and thus ignores distortions caused by lens shape.
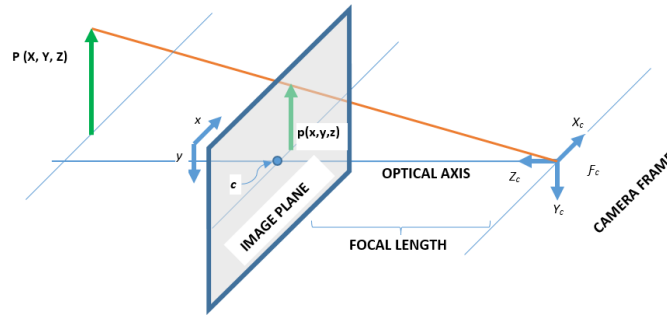

Figure 7: Pinhole Camera model.

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 $\mathbf{P} = [\, X, Y, Z\,]^{\mathrm{T}}$ is given by $\mathbf{p} = [\, x, y, z\,]^{\mathrm{T}}$ in the camera frame.

It can be further assumed for simplicity that all transfer functions related to optics, motion and, in particular, atmospheric attenuation and scattering, are absent. Note that relaxing these assumptions will impact the target's contrast, further reducing the R_0 quantity. As such, the value computed here represents the upper bound on R_0 in the absence of transfer function effects. Given the dimensions of the image sensor, 6.17x4.55 mm, the pixel pitch can be computed as follows,

$$d_{pix} = \frac{6.17 \text{ mm}}{3840 \text{ pixels}} = \frac{4.55 \text{ mm}}{2160 \text{ pixels}} = 1.6 \text{ μm}. \tag{6}$$

For convenience, it is assumed that the target subtends one pixel as the minimum requirement for detection. Note that existing transfer functions in a real system will result in an equivalent point spread function larger than a single pixel, therefore this is the limiting case. Under these assumptions, the relationship tying a target of known dimensions $X, Y, Z,$ the focal length, f, and the imaged dimension x, can be stated as follows,

$$\frac{x}{f} = \frac{X}{Z} \quad , \quad \frac{y}{f} = \frac{Y}{Z} \tag{7}$$

Imposing the constraint that the target subtends exactly one pixel allows x and y to be substituted by the pixel pitch. Isolating for Z, the following relationship is ascertained,

$$Z_x = \frac{X.f}{d_{pix}} \quad , \quad Z_y = \frac{Y.f}{d_{pix}} \tag{8}$$

For the remainder of this analysis, the tempest serves as the intruder aircraft. The maximum closing rate between the ownship and the intruder may be computed assuming a head-on collision scenario as the limiting case, with velocities obtained from Table 1, as follows,

$$v_{close} = v_{TEMPEST} + v_{ISAAC} = 28 \ m/s \tag{9}$$

To compute R_0, the wingspan and the fuselage cross-section are considered as single dimensions in X and Y axes respectively. The minimum of the two is taken as a conservative estimate of R_0, which establishes the upper bound on R_{DAA},

$$Z_X = \frac{X.f}{d_{pix}} = \frac{3.2 \times 8.25E - 3}{1.6 \ E - 6} = 16500 \text{ m}, \quad Z_Y = \frac{Y.f}{d_{pix}} = \frac{1.0 \times 8.25E - 3}{1.6 \ E - 6} = 5156 \text{ m}, \tag{10}$$

$$\therefore R_{0,GEOMETRIC} = \min(Z_X, Z_Y) = 5156 \text{ m}, \quad t_{0,GEOMETRIC} = \frac{R_0}{v_{close}} = 184 \text{ sec}. \tag{11}$$

Note that the geometric R_0 is an estimate where the target subtends exactly one pixel in the absence of intermediate mediums such as atmospheric attenuation, ownship motion and vibration, optical and imager transfer effects (including point spread function, aberrations and pixel quantum efficiency) that ultimately limit the contrast of real targets. Of the preceding transfer functions, atmospheric attenuation tends to dominate for targets much beyond 1 km,

especially under clear-sky conditions where the target is darker than the background, where irradiance from the target is lower than the irradiance due to atmospheric scattering effects [28]. In the absence of explicit knowledge of the transfer functions, the empirical R_0 may be attained from captured imagery by visual inspection. Sample empirical $R_{0,TEMPEST}$ and $R_{0,SR22}$ are shown in Fig. 8 were extracted by visual inspection using the following technique: the target was identified at close range, then the sequence of images run backwards until the target could not be distinguished from the background. It can be observed that the separation distance between the ownship and the Tempest is 1136m and the Tempest at this range approximately is approximately 4 by 3 pixels. For the SR22, the separation distance is 3406m and the aircraft is 3 by 1 pixels.

(a)   (b)

Fig. 8 Empirical R_0 with aircraft indicated by red circle: (a) Sortie 1 Tempest (b) Sortie 4 SR22.

For both types of intruder, the empirical R_0 is substantially lower than the geometric R_0, which is to be expected due to the aforementioned contrast limitations and serves as a much more meaningful metric for the present dataset.

$$\therefore R_{0,TEMPEST} = 1136\ m\ , t_{0,TEMPEST} = \frac{R_0}{v_{close}} = 42\ \text{sec}$$

And,

$$\therefore R_{0,SR22} = 3406\ m\ , t_{0,SR22} = 42\ \text{sec.}$$

The lower bounds on R_{DAA}, namely R_{warn} and R_{evade}, are derived from the avoidance sub-system. To remain as generic on the avoid component as possible, the simplest avoidance system is considered for the ownship used in these experiments. The intruder is assumed to be operating in level flight on a non-maneuvering flight path, with the ownship and intruder on a co-altitude, head-on collision course. The avoid maneuver is a 90-degree right-hand turn away from traffic. Given that the ownship is a multi-rotor, turn mechanics are simplified and estimated to take 3 seconds to change heading and resume at 10 m/sec. Note that a multirotor may also vertically descend to evade a collision, although this is often sub-optimal given that descent rates are limited to avoid vortex ring disturbances.

An avoidance maneuver requires a well-clear definition to ascertain the minimum distance requirements. In this work, the well-clear criterion is derived from minimum separation requirements of 1000 ft (300 m) lateral and 500 ft (150 m) vertical, as per FAA, Transport Canada and EASA advisories on UAS operating in GA space (10000 ft and below). For a right-hand turn at an airspeed of 10 m/s, the ISAAC requires $t_{evade} = \mathbf{30\ sec}$ to maintain well-clear separation from the intruder under the assumptions outlined. Furthermore, the ownship executes an avoidance maneuver autonomously as necessary, and informs the pilot in command at the start of the maneuver, equating R_{warn} to R_{evade}, thus,

$$t_{warn} = t_{evade} = \mathbf{30\ sec}, \qquad R_{warn} = R_{evade} = v_{close} \cdot t_{evade} = \mathbf{840\ m} \tag{12}$$

$$\mathbf{42\ sec} \leq t_{DAA} \leq \mathbf{30\ sec}, \qquad \mathbf{1168\ m} \leq R_{DAA} \leq \mathbf{840\ m} \tag{13}$$

Therefore, given upper and lower bounds, the time interval (Δt_{DAA}) available to detect, track, and compute an avoid solution may be computed as per Eqs. 2 and 4,

$$\therefore \Delta \mathbf{t_{DAA}} = \Delta(t_{detect} + t_{avoid}) \simeq \mathbf{12\ sec} \tag{14}$$

Table 2 illustrates the results with the computation repeated under identical assumptions for the GA aircraft. Note that the empirical R_0 is extracted from all sorties and presented as a range of values. This range takes into considerations differences in the environmental conditions including visibility and sun angle that affect the perceived range.

Note that $R_{warn} = R_{evade}$ and $t_{warn} = t_{evade}$ are assumed for convenience, and would in reality be distinct quantities based on the avoid solution employed.

Intruder	v_{close} [m/s]	Geometric R_0 [m]	Empirical R_0 [m]	t_0 [sec]	R_{evade}^{*} [m]	t_{evade}^{*} [sec]	Δt_{DAA} [sec]
SUAS – Tempest	28	5156	1168-1006	$42 - 36$	840	30	$12 - 6$
GA – SR22	63	14025	3406-2401	$54 - 38$	1890	30	$24 - 8$

As shown in Table 2, the GA aircraft is the limiting case, and depending on the design of the optical sub-component, can present as few as 15 seconds for $R_0 = 4.7$ km to the processing algorithm to detect, track and report a collision course intruder to the avoidance subsystem.

E. Flight Test Methodologies

For each day, fixed-wing sUAS vs. multirotor sorties were the first ones collected. The fixed-wing sUAS was launched and established as a waypoint pattern prior to the multirotor takeoff. The fixed-wing sUAS had about an

hour of flight time while the multirotor had 15 minutes. The fixed-wing sUAS was launched via rail in manual mode and then switched to auto mode by sUAS pilot command. Each sUAS pilot had a supporting ground control station monitoring onboard sensor data. The SR22 departed from Langley Air Force Base in Hampton, Va and flew to Beaverdam Airpark. The SR22 pilots communicated when it was ready for the multirotor to launch and to begin taking collision encounter data. A representative subset of 10 sorties across two days and two intruder types was selected for analysis in this work.

F. Dataset

Table 3 shows the dataset selected for use in this analysis. The testing environment was altitude separated for safety using 30 meters for sUAS encounters and 170 to 180 meters for GA encounters. Every other frame was labeled for Ground Truth (GT). On 10/25 flight conditions were partly cloudy, however, Visual Flight Rules (VFR) applied with the aircraft observed against a cloud background. Flight trajectories are superimposed on Google Maps as shown in Fig. 9. The first three sorties were head-on between a fixed-wing sUAS vs. a multirotor sUAS. The second set of two sorties were between GA head-on and multirotor with a slight lateral offset. On 10/30 flight conditions were clear with very few clouds in the sky with the aircraft observed against a sky background. The first two sorties featured the fixed-wing sUAS departing the collision volume of the multirotor. The next three sorties were head-on between GA and multirotor.

Table 3 Dataset

| Date and time Takeoff (EST)	Sortie	Ownship	Ownship Altitude (m)	Intruder	Intruder Altitude (m)	Number GT Labels
2018/10/25 09:40 AM	1	S1000	200	Fixed-wing sUAS head-on	230	393
	2	S1000	200	Fixed-wing sUAS head-on	230	298
	3	S1000	200	Fixed-wing sUAS head-on	230	367
2018/10/25 11:46	4	S1000	200	GA head on	380	639
	5	S1000	200	GA head on, camera offset	380	186
2018/10/30 9:29	6	S1000	200	Tempest Departing Collision Volume	230	664
	7	S1000	200	Tempest Departing Collision Volume	230	269
2018/10/30 12:41	8	S1000	200	GA head on	370	599
	9	S1000	200	GA head on	370	611
	10	S1000	200	GA head on	370	717

(a)

(b)

(c)

(d)
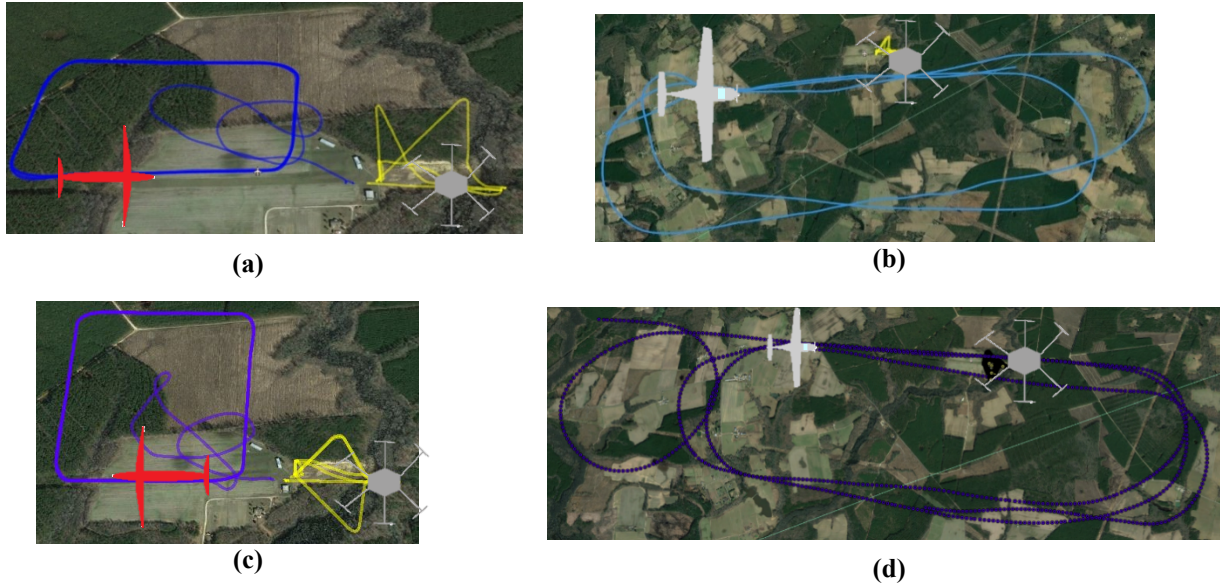
Fig. 9 Example Flight Trajectories: a) Sorties 1-3. Fixed wing sUAS flight path in blue with ownship multirotor in yellow, b) Sorties 4-5. Fixed wing GA plane flight path in blue with ownship multirotor in yellow, c) Sorties 6-7. Fixed wing GA plane flight path in blue with ownship multirotor in yellow, d) Sorties 8-10. Fixed wing GA plane flight path in blue with ownship multirotor in yellow.

G. Evaluation Methodology

The computer vision detection system is evaluated at the detector and tracker levels using the performance measures and metrics in Table 4 with definitions as described. The *Number of Ground Truth Detected in Sortie*, D_{GTS}, is a measure that includes all detections of ground truth by the detector, where the ground truth was manually extracted by visual inspection. The localization algorithm clusters detections within a 100-pixel radius, therefore detections are always unique. Difference and morphological detections are also fused together using the 100-pixel radius. The *Unique Ground Truth Detected in Sortie*, D_{UGTS}, is a measure includes only one detection per ground truth label. The difference detector and morphological detector may detect the same target twice, thus the definition of Unique is required. In any given frame, there is only 1 ground truth detection. Similarly, the *Unique Ground Truth Tracked in Sortie*, T_{UGTS}, includes only one tracker prediction per ground truth label.

For any of the aforementioned detection or prediction metrics, the Euclidean distance of the centroid must be within 100 pixels of the ground truth centroid for the intruder to be considered present for D_{GTS} and T_{UGTS} calculations. This equates to less than an azimuth angular threshold of 1.07 degrees and elevation of 1.48 degrees to be considered a hit. A 100-pixel radius corresponds to 1 degree in azimuth and 1.5 degree in elevation for the chosen camera configuration and is considered to be reasonable for the geometries under consideration in this detect and avoid scenario. Under the blue-sky theory assumption, it is unlikely that two aircraft are present in such a narrow voxel of space.

Table 4: Performance Measures and Metrics.

Performance Measure or Metric	Definition
True Positive	A detection with a centroid within 100 pixels of the Ground Truth Centroid
False Positive or False Alarm	A detection with a centroid that is 100 pixels or more from the Ground Truth Centroid
GT_S	Total Number of GT Labels in Sortie
D_{GTS}	Number of True Positives Detected in Sortie
D_S	Total Number of Detections in Sortie
D_{UGTS}	Number of Unique True Positives Detected in Sortie
T_{UGTS}	Number of Unique Ground Truth Tracked in Sortie
Detection Precision	$\dfrac{D_{GTS}}{D_S}$
Detection Recall	$\dfrac{D_{UGTS}}{GT_S}$
Tracker Ground Truth Accuracy	$\dfrac{T_{UGTS}}{GT_S}$
Average Tracker Ground Truth Accuracy	$\dfrac{\sum_{i=1}^{Number\ of\ Sorties} Tracker\ Accuracy_i}{Number\ of\ Sorties}$
Valid Tracks	Exist at least for 30 frames and 80% or more of tracker predictions are within 100-pixel Euclidean distance ground truth centroid
False Tracks	Exist at least for 30 frames and less than 80% are within a 100-pixel Euclidean distance ground truth centroid
Tracker Precision	$\dfrac{Valid\ Tracks}{Valid\ Tracks + False\ Tracks}$
Average Tracker Precision	$\dfrac{\sum_{i=1}^{Number\ of\ Sorties} Tracker\ Precision_i}{Number\ of\ Sorties}$
Tracker Robustness	$\dfrac{Length\ of\ Longest\ Track}{GT_S}$

Detector precision and recall are useful metrics for confirming that detectors are extracting the aircraft from the video and determining sources of false positives. *Tracker Accuracy* is a metric of how often the intruder aircraft is correctly tracked. *Tracker Precision* measures the valid track rate. *Tracker Robustness* is an indicator of how often the intruder aircraft is tracked continuously with just one track. Tracker accuracy is reported for each sortie to determine the influence of different detection thresholds and to compare KF and EKF tracker performance.

VI.   Results

The metrics defined in section V.G were utilized to analyze the performance of the detection and tracking stages of the pipeline, including sensitivity to key parameters and range analysis, and a comparison between KF and EKF trackers. The overall vision system is evaluated against the collection avoidance model described in the methods section above.

Note that the detection and tracking results are not scale invariant because: the aircraft have different aspect ratios, morphologies, and visual-cross sections. Image-based feature extractors have been shown to perform differently depending on the object, thus not all are the same from the perspective of detectability. Further, the aircraft operate at different speeds relative to the sampling rate of the sensor and the tracking algorithm will respond differently and have different Kalman filter parameters.

A.   Detection Threshold Sensitivity Analysis

The detection threshold is a parameter that influences tracker performance by including and excluding potential hits from the tracking stage. Careful selection of the threshold can provide a good balance between track robustness and false detections. Fig. 10 shows tracker performance in terms of Average Tracker Precision across all sorties and Average Tracker Ground Truth Accuracy across all sorties to determine the optimum detection threshold for KF and EKF trackers. A detection threshold of 5 would provide up to 5 detections for a given frame, which typically is 4 false detections and 1 ground truth detection for a given frame. Ideal performance would be 1.0 for Average Tracker Precision across all sorties and 100% for Average Tracker Ground Truth Accuracy across all sorties because all GT labels would be captured with the tracker without false track generation. The elbow of precision-accuracy is selected because it represents a good trade-off between precision and accuracy. Thresholds values ranging from 1 through 4 provided higher tracker precision than threshold 5, and similar Average Tracker Ground Truth Accuracy, however, a vision-based collision avoidance system needs to be able to track multiple aircraft. Therefore, the detection threshold 5 was selected for detailed analysis as it provides strong ground truth accuracy across all sorties while achieving high Average Tracker Ground Truth Accuracy. Fig. 11a shows individual detector performance as a function of sortie and detection recall for a detection threshold of 5. Here, the morphological detector consistently outperformed the difference detector across all sorties. The combined detector yielded the highest detection recall across all sorties.
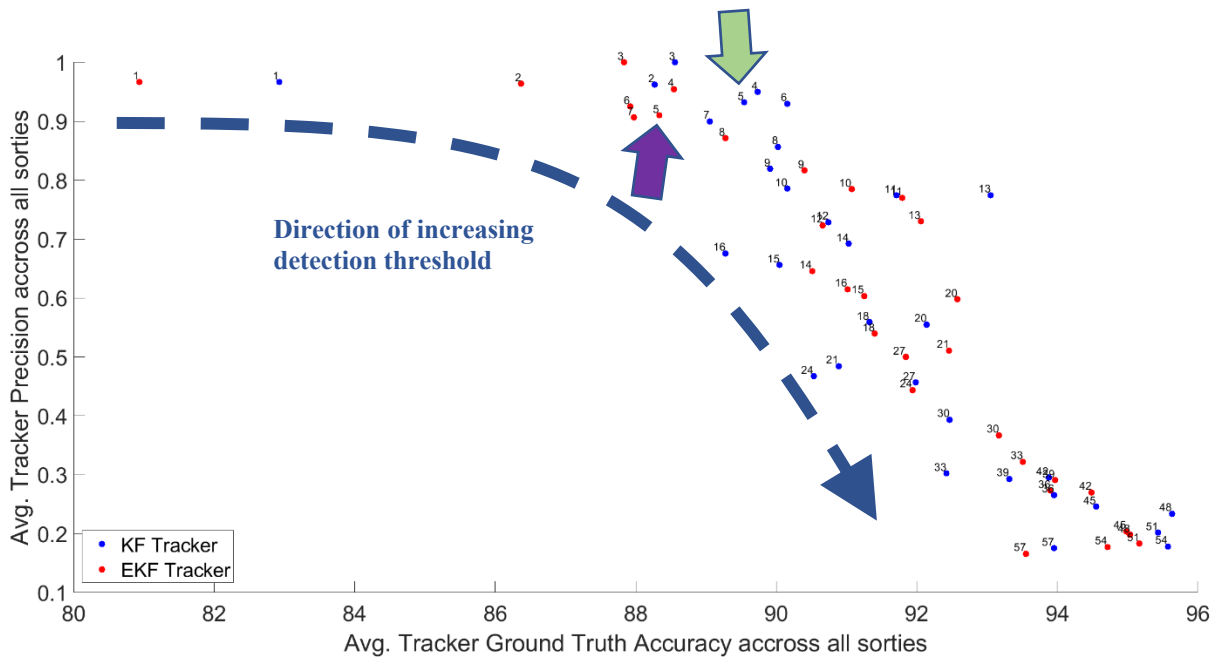
Fig. 10 Detection Threshold for Tracking using the Linear Kalman Filter (blue) and the Extended Kalman Filter (red). Selected detection thresholds of 5 for EKF and KF shown by green and purple arrows respectively. The dashed line with arrow shows the direction of increasing detection threshold, where precision decreases and ground truth accuracy increases.

Fig. 11b shows the impact of different detection thresholds to tracker accuracy without respect to tracker precision.

Tracker performance is analyzed by plotting accuracy for KF and EKF trackers across sorties to determine then

influence of detect thresholds. Table 5 shows the tracking accuracy in relation to precision, sortie length, number of
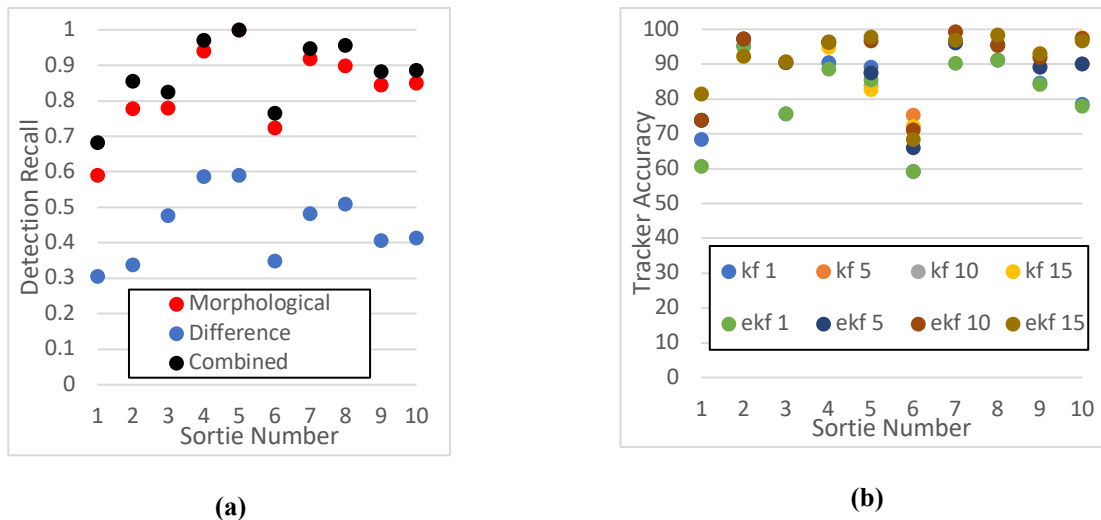
false tracks, and total tracks.

(a)

(b)

Fig. 11 (a) Detector Performance threshold 5 and (b) Tracker Performance for Different Detection Thresholds, where KF is the Kalman Filter tracker results and EKF is the Extended Kalman Filter tracker for the combined detections.

Tracking was the weakest in the first part of Tempest Sortie 1, as the intruder aircraft emerged from a cloudy background with varying low contrast and thus generated unstable detection at maximum ranges. For Tempest sorties 2 and 3, the background contrast had less variation at maximum range and tracking was stable across all frames with tracker accuracy exceeding 90% in nearly all cases. In GA Sortie 4, the ownship turned left and right three times, changing the location of the GA plane in image space, which generated short periods of instability in tracking as the ownship movement offsets were not integrated into the computer vision algorithm. GA sortie 5 had an offset camera relative to the ownship heading and thus the encounter was short in terms of time and frames as well as maximum available detection range, as the intruder entered and exited from the side of the frame, instead of the appearing from the background as in all other sorties. The trackers performed weakly in tempest sortie 6 when the intruder aircraft moved into a low-contrast backgrounds closer to the horizon. Tracking was stronger when the ownship overtook tempest sortie 7. GA sorties 8 through 10 consistently had tracking accuracy over 90% for all trackers with detection thresholds higher than 1.

B. Precision and Recall

Computer vision and machine learning benchmarking challenges utilize different evaluation metrics for object detection [28]. One frequently used metric is the Received Operator Characteristic (ROC). ROC is frequently used in RADAR detection research, where true negative is typically RADAR signal returns from non-aircraft sources that are correctly identified as noise. ROC was considered in this work, however, a definition for true negatives is not practical because all pixels in an image without an object would need a label as true negatives. For example, consider a detection threshold 5 where there are 4 false positives and 1 true positive, representing five 10 x 10 pixel regions. In this case, all pixels that are not associated with a detection would be considered a true negative. For the ~8.3 MP imagery in this work, this would mean that ~7.8 MP would be labeled as true negatives, 0.4 MP would be false positives, and 0.01 MP would be true positive. However, ~7.8 consistently drives a sensitivity calculation to always be 1.0 because ROC False Positive Rate = True Negatives / (True Negatives + False Positives). Instead of ROC analysis, precision-recall in useful because precision is a metric of the detector's ability to extract the objects of interest and recall is a metric of detector's ability to detect all the labeled ground truths. For this aviation application, precision and recall is used to evaluate detectors because it captures each detector ability to find the intruder aircraft with respect false positive generation.

A sensitivity analysis was conducted on detection threshold for each sortie using precision and recall. Fig. 12 illustrates a sensitivity analysis for detectors by sortie for detection threshold and annotated by aircraft, trajectory, and
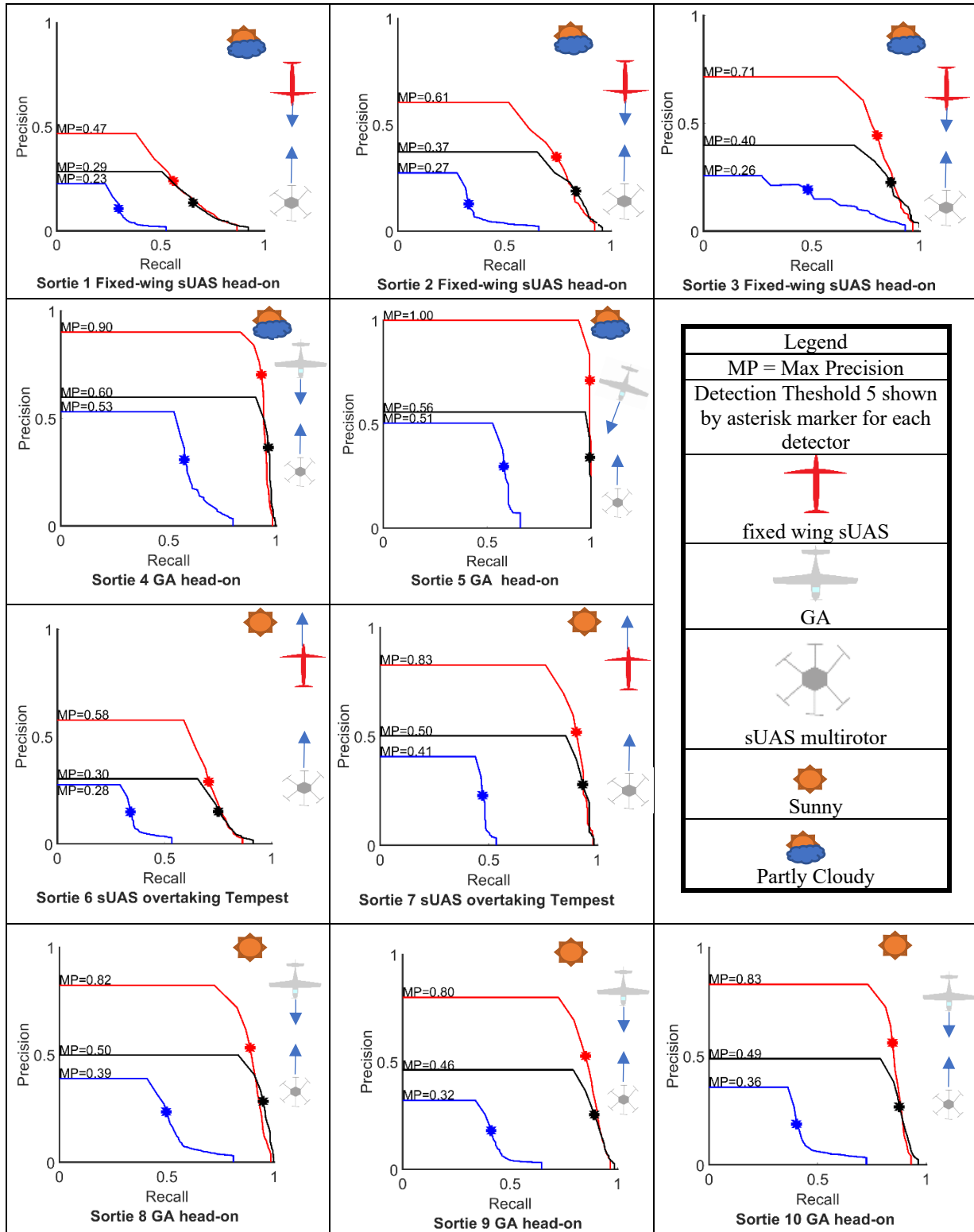
Fig. 12 Precision-Recall for Detectors by Sortie where blue is the difference detector, red is the morphological detector, and black is the combined morphological and difference detectors. MP denotes the maximum precision for the detectors and the asterisk show precision-recall values for the detection threshold 5 for each detector.

weather condition. For all sorites, the blue line is the difference detector, the red line is the morphological detector, and the black line is the combined detectors. The x-axis is recall and the y-axis is precision. Each point on the graph represents a different detection threshold between 1 and 57. The detection threshold drives the precision-recall calculations as increasing the detection threshold tends to add false alarms at a faster rate than the additional ground truth detections, which tends to decrease precision and increase recall. The asterisk shows the precision-recall value for detection threshold 5, which was selected for detailed analysis for tracking. The maximum precision value is shown by extending the precision-recall curve to the x-axis.

Generally, a strong precision-recall curve would start with high max-precision nearly equaling 1 for recall values between 0 and 0.9. Precision then would steeply decrease from 1 to nearly 0 for recall values between 0.9 and 1.0. In this way, the model quickly captures most of the ground truth detections and only starts producing false alarms at a high rate after the detection threshold is increased substantially, thus driving the precision-recall curves towards a precision value of 0 and a recall value of 1. Precision-recall curves do not perform ideally in this work because of the challenge of detecting very small, low-contrast objects with differing lighting conditions.

For Sortie 1, the morphological curve has the highest maximum precision value of 0.47, which corresponds to detection threshold of 1 and indicates that about half of the detections were ground truth. The difference detector has a maximum precision value 0.23 for the detection threshold of 1, indicating that about one fifth of the detections in the sortie captured the ground truth. The combined detector has a maximum precision of 0.29 for threshold 1, which corresponds to1 detection from morphological and 1 detection from difference detector. For this sortie, all detector curves decrease in a convex manner indicating a delayed gain in recall relative to precision loss. One possible explanation for the convex curve behavior is that small intruder aircraft are difficult to detect because of their small size and low contrast, thus increasing detection threshold tends to add false detections from other regions of the image at a higher rate relative to the rate of additional new ground truth detections on the intruder aircraft. The convex behavior extends to all difference detector results across all sorties. Small, low-contrast aircraft generate weak signatures in the difference frame. Typically, the detection thresholds between 30 and 60 show the most improvement in recall for the difference detector. The morphological detector curve shows strong improvement for recall as precision decreases. The morphological detector can extract nearly all intruder aircraft with an increasing the number of false alarms. The higher precision-recall performance for the morphological detector suggests that the cross-hair convolution filter performs well at extracting small aircraft targets. This may be related to specific morphological shape of fixed-wing aircraft with the wing, tail, and fuselage providing good activation for cross-hair convolution

while suppressing false positives from clouds and MJPEG encoding artifacts. Saturating the detection threshold generates a higher recall, however, the simultaneous inundation of false alarms decreases tracker robustness. Combining the detectors decreased precision while increasing recall. Precision decreases when combining detectors because the morphological detector generates both false alarms and ground truth while the difference detector generates fewer ground truth detections and more false alarms compared to the morphological detector. This imbalance drives the precision calculation lower than the morphological detector's precision, however, the additional unique ground truths captured by the difference detector increases recall by nearly 10% for detection threshold of 5 in sortie 1.

The precision-recall results from this work show a higher precision for detecting GA aircraft to the fixed-wing sUAS and an improvement in precision-recall later in the day for sUAS sorties. The combined precision-recall curve often exceeds the morphological detector precision-recall curve, showing that the difference detector is extracting detections that the morphological detector missed. Tempest precision-recall curves were stronger on day 2, however, it is not clear whether this is due to differing weather conditions or over-taking vs. head-on collision geometries. Precision and recall increased during each day for Tempest flights for all detectors. This may be related to higher sun elevation angles providing more light on the tempest, thus improving contrast between the tempest and its background.

C.   Tracker Comparison

EKF trackers typically outperform KF trackers for non-linear data, however, the tracker parameters such as dropout, track age, and invisibility may explain similar performance for both trackers in the dynamic sUAS flight environment. The comparison between KF and EKF is shown in Table 5 examines the Detection Threshold 5 in detail at the tracking-level. The EKF and KF trackers generated similar performances in terms of precision and tracker accuracy averages across all sorties. GA aircraft were tracked with greater accuracy over sUAS, which may be related to their larger physical dimensions and higher altitudes relative to atmospheric attenuation and scattering effects. The false tracks for both trackers in sortie 6 occur simultaneously in the same area of the frame, slightly above the horizon, and the false track in sortie 7 appears at the horizon.

Tracking robustness was similar across the EKF and KF trackers with minor variations for sorties 1, 6, 7, and 9. Sorties where the Robustness is similar to Tracker Accuracy indicates that the tracker was able to maintain one track for the majority of the frames in the sortie, which also correlates to a lower number of tracks for that sortie. However, discontinuities in tracking occurred in sortie 1 as the Tempest moves across an area of light atmospheric attenuation

and scattering effects, sortie 6 and 7 when the Tempest turned at the closest corner, and sortie 9 when the SR22 entered

an area of low contrast.

Table 5: Tracking Performance for Detection Threshold 5.

Sortie Number and Description	Num. of Frames	KF Track. Precision	Num. KF False Tracks	Total Num. KF Tracks	KF Track. Acc.	KF Track. Robust.	EKF Track. Precision	Num. EKF False Tracks	Total Num. EKF Tracks	EKF Track. Acc.	EKF Track. Robust.
1 Tempest Head-on	393	1	0	2	74.0	38.2	1	0	2	74.0	38.2
2 Tempest Head-on	298	1	0	1	97.3	97.0	1	0	1	97.3	97.0
3 Tempest Head-on	367	1	0	3	90.5	80.7	1	0	4	90.5	80.7
4 GA Head-on	639	1	0	1	96.4	96.2	1	0	1	96.4	96.2
5 GA Head-on, near miss	186	1	0	2	87.6	77.4	1	0	3	87.6	87.1
6 ==Tempest Departing Collision Volume==	664	0.57	3	7	75.5	55.1	0.5	3	6	66.1	53.6
7 ==Tempest Departing Collision Volume==	269	0.75	1	4	99.3	52.4	0.8	1	5	96.3	52.4
8 GA Head-on	599	1	0	3	95.4	79.0	0.8	1	5	95.7	79.0
9 GA Head-on	611	1	0	4	89.2	39.8	1	0	4	89.2	39.8
10 GA Head-on	717	1	0	4	90.2	82.2	1	0	4	90.2	82.2
Avg GA	550.4	1	0	2.8	91.8	74.9	0.96	0.2	3.4	91.8	76.9
Avg. sUAS	385.2	0.86	0.8	3.4	87.3	64.7	0.86	0.8	3.6	84.8	64.4
Avg. across all sorties	474.3	0.93	0.4	3.1	89.5	69.8	0.91	0.5	3.5	88.3	70.6

Table 6: Localization Error Detection Threshold 5.

Sortie Number and Description	AVG Difference Localization Error	AVG Morph Localization Error	AVG KF Tracker Localization Error	AVG EKF Tracker Localization Error
1 Tempest Head-on	10.2	11.1	10.4	10.5
2 Tempest Head-on	9.5	6.2	8.4	8.4
3 Tempest Head-on	8.0	9.2	9.8	9.1
4 GA Head-on	19.7	23.6	23.4	23.4
5 GA Head-on, near miss	32.5	32.3	31.1	30.4
6 ==Tempest Departing Collision Volume==	12.5	14.5	15.1	11.6
7 ==Tempest Departing Collision Volume==	15.3	16.7	17.1	18.0
8 GA Head-on	16.6	18.7	16.7	17.0
9 GA Head-on	19.7	20.5	19.3	19.3
10 GA Head-on	15.1	16.7	16.1	16.1

Detectors and trackers performed with localization errors ranging from 8.0 to 32.5 pixels as shown in Table 6. The

KF and EKF trackers performed similarly with respect to localization error. Typically, the average tracker localization

errors were between the morphological and difference localization errors. However, the tracker makes predictions

based off the Kalman Filter, and improves localization error for sortie 5, which happens to be a head-on collision that begins at closer range. Sorties 1, 2, 4, 9, and 10 generated nearly identical average KF and EKF tracker localization errors. Sorties 3, 7, and 8 differed in average localization error by less than 1 pixel. The average EKF localization error for Sortie 6 is lower than average KF localization error, however, the EKF tracker accuracy is 66.1 compared to 75.5 for KF from Table 5, thus supplying less intruder tracking coverage. Given that average tracker localization errors are within 1 pixel for 9 of 10 sorties, the decision on which tracker should be based on other metric such as tracker accuracy and tracker robustness. The KF tracker achieves 1.2 % higher tracker accuracy and is 0.8% lower in tracker robustness, thus making it slightly preferred based on tracker accuracy, robustness, and localization error.

D. Range Estimation

Table 7 shows the vision system results with respect to the DAA model described on page 15. Recall that R_{det} is the initial tracking range and that R_{evade} is the minimum distance to prevent a well-clear violation by executing an avoidance maneuver. Detectors performed similarly for the first detection range with the morphological detector achieving higher R_{det} in a few instances. EKF and KF trackers achieved the same R_{det}. R_{det} surpassed R_{evade} in 9 of 10 sorties, with the exception of sortie 5, where the camera was offset relative to the ownship and GA plane encounter. Sortie 5 results demonstrate that this system can track a GA plane on a tangential but not intersecting course, which is useful for identifying aircraft that may cross the FOV without being on a direct collision course.

Table 7: DAA analysis Detection Threshold 5.

Sortie Number and Description	GT [m]	POT for GT w x h	Diff first detection [m]	Morph first detection [m]	KF R_{det} [m]	EKF R_{det} [m]	R_{evade} [m]	EKF and KF t_{detect} [sec]
1 Tempest Head-on	1136	4x3	1136	1136	1022	1022	840	36.5
2 Tempest Head-on	1107	3x3	1093	1107	1078	1078	840	38.5
3 Tempest Head-on	1006	3x4	999	999	980	980	840	35.0
4 GA Head-on	3406	3x1	3406	3406	3247	3247	1890	51.5
5 GA Head-on, near miss	858	608x161	812	858	721	721	1890	11.4
6 ==Tempest Departing Collision Volume==	1067	8x10	1067	1067	987	987	840	35.3
7 ==Tempest Departing Collision Volume==	1168	4x10	1166	1166	1150	1150	840	41.1
8 GA Head-on	2401	4x3	2401	2401	2297	2297	1890	36.5
9 GA Head-on	2514	3x3	2514	2514	2404	2404	1890	38.2
10 GA Head-on	3053	3x2	2994	2994	2936	2936	1890	46.6
Avg GA	2446	NA	2425	2435	2321	2321	1890	37.3
Avg. sUAS	1097	NA	1092	1095	1043	1043	840	36.8

The initial detection results are on par with the Aerial Object Tracking pipeline in [29], where established tracks ranged between 2830 to 980 meters for GA vs. GA detection and track. The average R_{det} for GA for this work is 2435m compared to 1822m in [29], however, the closure speed for GA vs. GA is much higher than the SR22 vs. ISAAC.

VII. Conclusion

Reliable onboard collision avoidance systems for sUAS are needed for safe autonomous sUAS operations. This work demonstrates that fixed-wing sUAS and GA planes above the horizon can be detected and tracked in the far-field from a multirotor in a dynamic sUAS environment using a machine vision system. The computer vision approach was developed thoroughly by testing different aircraft detectors over different flight conditions and with two types of intruder aircraft. Combining the morphological and difference detectors provided the best tracking performance. The detection thresholds were thoroughly explored by heuristically determining values for this sUAS tracking. The EKF and KF trackers performed equally to the nearest meter for the initial detection ranges. Average KF tracker accuracy across all sorties was slightly higher than average EKF tracker accuracy, however, there was a slight decrease in tracking robustness. Success criteria was defined using standard well-clear definitions and a refined range inequality model that incorporates sensor characteristics. The offline vision system exceeds the well-clear definitions commonly used in sUAS and GA for the head-on and overtaking sUAS and GA aircraft sorties. The vision system provides enough time for the flight controller to perform an avoidance maneuver while enabling communication to cooperative aircraft in the surrounding area.

Future work includes porting the current system to an embedded board for a real-time demonstration of the computer-vision sensing pipeline. This would include the integration of attitude information from the inertial navigation system and use of global shutter cameras to improve vision system performance. Track stability and robustness in the presence of ownship and intruder motion will also be analyzed. Another goal is the development of a perception system to classify different kind of intruders such as birds, GA plane, multirotor sUAS, and fixed wing sUAS. Classifying the intruder will give insights into its performance envelope, i.e. it will help constrain the choices of avoidance maneuvers. But it is unlikely to impact the performance of detection and tracking of systems. Any DAA system highly dependent on classification will be quite fragile in the field because of the constant upgrades as new platforms enter the marketspace. Future avenues explore ranging and also combining with other sensors and exploring how the vision sensor can perform within a dynamic aircraft environment, such as keeping track of existing intruders during an evasive maneuver.

References

[1] Federal Aviation Administration, "Unmanned Aircraft Systems (UAS) Traffic Management (UTM) Concept," 2018.

[2] Federal Aviation Administration, "FAA Aerospace Forecast: Fiscal Years 2020-2040," 2020.

[3] G. Glaab, C. Dolph, S. Young, N. Coffey, R. McSwain, M. Logan and D. Harper, "Small Unmanned Aerial System (UAS) Flight Testing of Enabling Vehicle Technologies for the UAS Traffic Management Project," NASA, Hampton, 2018.

[4] S. Szatkowski, A. Kriz, L. Ticatch, R. C. J. Briggs and C. Morris, "Airborne Radar for sUAS Sense and Avoid," in *Digital Avionic Systems Conference*, San Diego, 2019.

[5] P. Andrasi, T. Radisic, M. Mustra and J. Ivosevic, "Night-time Detection of UAVs using Thermal Infrared Camera," *Transportation Research Procedia,* vol. 28, pp. 183-190, 2017.

[6] C. Dolph, M. Logan, L. Glaab, T. Vranas, R. McSwain, Z. Johns and K. Severance, "Sense and Avoid for Small Unmanned Aircraft Systems," in *AIAA Information Systems-AIAA Infotech@ Aerospace*, Grapevine, 2017.

[7] G. Fasano, D. Accardo, A. Tirri and A. Moccia, "Experimental Analysis of Onboard Non-Cooperative Sense and Avoid Solutions Based on Radar , Optical Sensors, and Data Fusion," *IEEE A&E SYSTEMS MAGAZINE,* vol. 31, no. 7, pp. 6-14, 2016.

[8] M. Wilson, D. Ryan, D. Bratanov, A. Wainwright, J. Ford, L. Cork and M. Brouckaert, "Flight Test and Evaluation of a Prototype Sense and Avoid System Onboard a ScanEagle Unmanned Aircraft," *IEEE A&E Systems Magazine ,* vol. 10, no. 1109, pp. 6-15, 2016.

[9] J. Lai, J. Ford, L. Mejias and P. O'Shea, "Characterization of Sky-region Morphological-temporal Airborne Collision Detection," *Journal of Field Robotics,* vol. 30, no. 2, pp. 171-193, 2013.

[10] A. Nussberger, Aerial Object Tracking from an Airborne Platform, Zurich: ETH Zurich Research Collection, 2015.

[11] D. Ye, J. Li, Q. Chen, J. Wachs and C. Bouman, "Deep Learning for Moving Object Detection and Tracking from a Single Camera in Unmanned Aerial Vehicles (UAVs)," *Society for Imaging Science and Technology,* vol. 466, pp. 1-6, 2018.

[12] C. Minwalla, D. Tuplan, N. Belacel, F. Famili and K. Ellis, "Detection of Airborne Collision-Course Targets for Sense and Avoid on Unmanned Aircraft Systems Using Machine Vision Techniques," *Unmanned Systems,* vol. 4, no. 4, pp. 1-18, 2016.

[13] "Small Aircraft Flight Encounters (SAFE) Dataset Repository," 2019. [Online]. Available: https://sites.google.com/vt.edu/safe-repository.

[14] H. Seviil, A. Dogan, K. Subbarao and B. Huff, "Evaluation of Extant Computer Vision Techniques for Detecting," in *International Conference on Unmanned Aircraft Systems (ICUAS)*, Miami, 2017.

[15] C. Dolph, L. Glaab, B. Allen, M. Consiglio and K. Iftekharuddin, "An Improved Far-Field Small Unmanned Aerial System Optical Detection Algorithm," in *Digital Avionic Systems Conference*, San Diego, 2019.

[16] C. Minwalla and K. Ellis, "Experimental Evaluation of PICAS: An Electro-Optical Array for Non-Cooperative Collision Sensing on Unmanned Aircraft Systems," in *AIAA SciTech*, Grapevine,Tx, 2017.

[17] N. Otsu, "A Treshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics,* vol. 9, no. 1, pp. 62-66, 1979.

[18] J. Canny, "A Computational Approach to Edge Detection," *IEEE Trans. Pattern Anal.. Machine Intell.,* vol. 8, pp. 679-698, 1986.

[19] S. M. Pizer, E. Amburn, J. Austin, R. Cromartie, A. Geselowitz, T. Gree, B. Romeny, J. Zimmerin and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Computer Vision, Graphics, and Image Processing,* vol. 39, no. 3, pp. 355-368, 1987.

[20] 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.

[21] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *International Joint Conference on Artificial Intelligence*, British Columbia, 1981.

[22] J. Munkres, "Algorithms for Assignment and Transportation Problems," *Journal of the Society for Industrial and Applied Mathematics,* vol. 5, no. 1, 1957.

[23] H. Kyhn and B. Yaw, "The hungarian method for the assignment problem.," *Naval. Res. Logist. Quart,* pp. 83-97, 1955.

[24] C. Munoz, A. Narkawicz, G. Hagen, J. Upchurch, A. Dutle and M. Consiglio, " DAIDALUS: Detect and Avoid Alerting Logic for Unmanned Systems," in *Proceedings of the 34th Digital Avionics Systems Conference* , Prague, 2015.

[25] E. Landhuis, "Medical Cargo Could Be The Gateway For Routine Drone Deliveries," National Public Radio, 10 March 2018. [Online]. Available: https://www.npr.org/sections/health-shots/2018/03/10/592059175/medical-cargo-could-be-the-gateway-for-routine-drone-deliveries. [Accessed 31 Jul 2020].

[26] M. Consiglio, L. Glaab and C. Munoz, "Sense and Avoid Characterization of the Independent Configurable Architecture for Reliable Operations of Unmanned Systems," in *Air Traffic Management Research* , Vienna, Austria, 2019.

[27] S. Balachandran, C. Munoz, M. Consiglio, M. Feliu and A. Patel, "Independent Configurable Architecture for Reliable Operation of Unmanned Systems with Distributed Onboard Services," in *38 IEEE/AIAA 37th Digital Avionics Systems Conference* , London, UK, 2018.

[28] R. Padilla, S. Netto and E. . da Silva, "Survey on Performance Metrics for Object-Detection Algorithms," in *International Conference on Systems, Signals and Image Processing (IWSSIP)*, Brazil, 2020.

[29] A. Nussberger, H. Grabner and L. Van Gool, "Robust Aerial Object Tracking in Images with Lens Flare," in *IEEE International Conference on Robotics and Automation*, Seattle, 2015.

[30] C. Minwalla, P. Thomas, K. Ellis, R. Hornsey and S. Jennings, "Range performance evaluation from the flight tests of a passive electro-optical aircraft detection sensor for unmanned aircraft systems," *Journal of Unmanned Vehicle Systems,* vol. 4, no. 2, pp. 96-114, 2016.

[31] Radio Technical Commission for Aeronautics, *SC-228, Minimum Performance Standards for Unmanned Aircraft Systems,* Washington DC: Radio Technical Commission for Aeronautics, 2020.