Passive, long-range detection of Aircraft: Towards a field deployable Sense and Avoid System

Debadeepta Dey Christopher Geyer Sanjiv Singh Matthew Digioia

May 1, 2011

Abstract

Unmanned Aerial Vehicles (UAVs) have played vital roles recently in both military and non-military applications. One of the reasons UAVs today are unable to routinely fly in U.S. National Air Space (NAS) is because they lack the sense and ability to avoid other aircraft. Although certificates of authorization can be obtained for short term use, it entails significant delays and bureaucratic hurdles. Therefore there is a great need to develop a sensing system that is equivalent to or has greater performance than a human pilot operating under Visual Flight Rules (VFR). This is challenging because of the need to detect aircraft out to at least 3 statute miles, while doing so on field-of-regard as large as 30 degrees by 220 degrees and within the payload constraints of a medium-sized UAV. In this paper we report on recent progress towards the development of a field deployable sense and avoid system and concentrate on the detection and tracking aspect of the system. We tested a number of approaches and chose one that resulted in 100% detection rate (over about 40 approaches) and 98% tracking rate out to 5 statute miles and a false positive rate of 1 every 50 frames. Within a range of 3.75 miles we can achieve nearly 100% tracking rate.

1 Introduction

Medium and small Unmanned Aerial Vehicles (UAVs) [2, 3, 4, 1] are typically commanded via waypoints with the operators at remote locations. Such UAVs generally do not have the payload to carry radar systems or a traffic collision avoidance system (TCAS) or transponders. Collision avoidance is currently done by flight planning, use of ground or air based human observers and segregated air spaces. Lack of an autonomous sense-and-avoid system along with absence of regulations is preventing commercial UAVs from flying in the national airspace (NAS). The Radio Technical Commission for Aeronautics, Special Technical Committee 203 has been given the task to prepare Sense and Avoid Minimum Performance Standards for Unmanned Aircraft Systems [6] by December, 2013. These proposals will be used by FAA

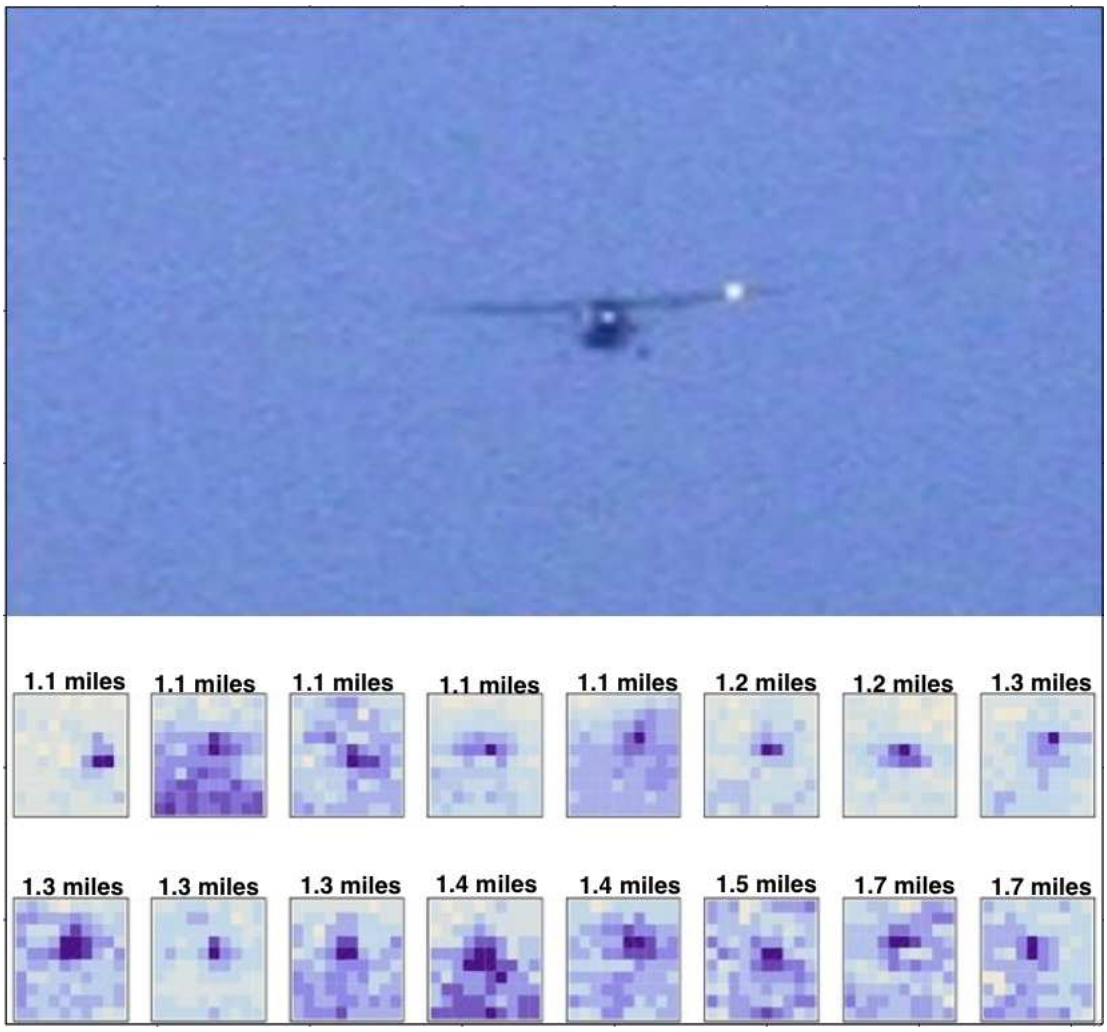

Figure 1: Selection of 11×11 subwindows showing the image of the Piper Archer II which was used as the approaching aircraft for collecting imagery, at a range of 1.5 miles. The camera and lens used had 0.41 milliradian/pixel resolution and a field of view of 30°(H) $\times$ 21°(V).

to determine final regulations. UAVs must not degrade the existing safety of the NAS, but the metrics that determine this have to be fully determined yet. It is still possible to state functional requirements and determine some performance minimums. For both manned and unmanned aircraft to fly safely in the same airspace UAVs will need to detect other aircraft and follow the same rules as human pilots.

Key specifications of the international committee F38 on UAS systems standard F2411-04 [5] proposed requirements which include a field of regard of 220°(horizontal) $\times$ 30°(vertical), minimum detection range of 3 statute miles under visual flight rules and a required miss distance of 500 ft. Without this capability, widespread utilization of UAVs will not be possible.

This paper focuses on the sensing of aircraft with passive vision. Small size, low weight

and power requirement make cameras attractive for this application. Multiple cameras can be used to cover the wide field-of-regard. A typical image of an aircraft at ranges on the order of a few miles is a few pixels in diameter. Fig.1 shows a 11×11 window around the image of the approaching aircraft at various ranges. Part of the challenge in detecting aircraft in such a wide field of regard reliably is the low signal to background ratio. Active sensors like radar are not feasible because of their prohibitive power and size requirements [8] for UAVs. Passive vision provides a low cost, low power solution albeit at the cost of a relatively high false positive rate.

The contributions of our work include the formulation of a novel method for long-range detection of aircraft for UAV sense and avoid purposes, the evaluation of the approach on a large corpus of real imagery of flying aircraft vs range. Our approach uses a novel descriptor for low signal to noise ratio targets. This descriptor is then used to train a supervised classifier using a closed-loop training method to distinguish between true and false detections. The closed-loop training method uses the false positives from the commonly used morphological filtering stage as the negative examples. This suppresses a large percentage of the false positives even before the tracking stage of our approach. We achieve 100% detection (over about 40 approaches) and 98% tracking rate out to 5 statute miles and a false positive rate of 1 every 50 frames. This is beyond the 3 statute miles limit imposed by proposed regulations.

In section 2 the related work in vision based sense and avoid systems for UAVs is discussed. In section 3 we detail the requirements imposed on a complete sense and avoid system by regulations. In section 4 we discuss the details of the vision based aircraft detection algorithm. In section 5 we outline our efforts to collect imagery of flying aircraft. In section 6 we present details about the result of our algorithm on the corpus of real ground truthed imagery of aircraft. Finally in section 7 we discuss the path forward towards a field deployable sense and avoid system.

2 Related Work

Utt et el. describe a fielded vision-based sensory and perception system with potential for implementation on small UAVs [23]. McCandless proposes an optical flow method for detecting aircraft [18]. This is suitable only for moving objects and therefore is not useful for a target on a collision course, which will not appear to be moving. The use of morphological filtering is popular in computer vision based sense and avoid systems [14, 10]. However this approach generates a significant number of false positives and requires tracking of the features over a large number of frames. Petridis et al. use AdaBoost to detect aircraft in low resolution imagery [21]. Track-Before-Detect (TBD) is an approach used especially on infrared imagery [13, 7]. Defence Research Associates have implemented a vision based sense and avoid system on a Predator UAV system, which can track targets using three cameras and custom hardware [17].

A field deployable Sense and Avoid system must be able to operate in a variety of atmospheric conditions including fog, haze and even directly against the glare of the sun. The operation of the system must not degrade beyond an acceptable level under all these conditions. We developed an image formation model which accounts for the various atmospheric conditions. We used this model to predict the signal to background ratio of the image of the aircraft. The signal to background ratio is a measure of the detectability of the aircraft in the image. The image formation model is described in detail in [15]. The image formation model

also allows us to determine the suitability of any sensor combination before using the sensor and also to determine the minimum derived resolution for achieving a specified performance. The performance of the image formation model has been validated by the vast corpus of real imagery of flying aircraft that we collected during the course of this project.

3 Summary of regulations

Both civil and defense authorities have been involved in drafting regulations for UAV integration into NAS. This is in light of the fact that a large number of UAV operations in NAS will be for security and defense purposes. Hence a summary of the proposed regulations from both civil and defense organizations is presented in this section.

In July 2003, manufacturers, members of the Association for Unmanned Aerial Vehicle Systems International (AUVSI) and other interested parties voted to create through American Society for Testing and Materials (ASTM International) the International Committee F38 on Unmanned Aerial Systems (UAS). In November, 2004 the Office of the Secretary of Defense (OSD) [19] released a plan for airspace integration for unmanned aviation. This document established top-level timelines and program milestones in relation to the Federal Aviation Authority's (FAA) strategic plans. This is in accordance with the mandate by Congress for the defense and civil authorities to work out the timelines and standards for full integration of UAVs into manned airspace. In November, 2006 the FAA [12] updated rules to establish protocols and guidelines for obtaining Certificates of Authorization (CoA) for temporary UAV operation in manned airspace. The definition of Remotely Operated Aircraft (ROA) was also incorporated to account for the fact that most UAVs are remotely operated.

In 2004 the F38 committee released standard F2411-04 which proposed requirements for sense-and-avoid systems [5]. These proposed standards serve as a baseline for current and future design of sense-and-avoid systems and the most relevant points are presented here. F2411 defined classes of sense and avoid systems, as well as functional and non-functional requirements for collision detection. The F2411 standard also defines three classes of sense and avoid systems based on their sense and avoid capabilities and where and how they apply them:

Class 1 (Pilot-in-the-loop): A system that warns a remote operator of a potential collision with another vehicle. The remote operator is responsible for evasive maneuvers.

Class 2 (Automated Air): In addition to detecting threats, a Class 2 system initiates a maneuver to avoid a potential mid-air collision or near mid-air collision autonomously.

Class 3 (Automated Air and Surface): Class 3 systems have the additional capability of detecting and avoiding collisions with vehicles while taxiing on the runway.

Table 1 lists requirements as specified in the ASTM F2411 standard, as well as the Highly Capable UAVs (HCU) Payloads planning document. Where they differ, we note the two different requirments. These are only a subset of the requirements, that are most likely to affect sensing requirements.

4 Aircraft Detection and Tracking

We experimented with a number of different approaches to detecting small targets with low signal to background ratios with an emphasis on methods that have both high detection rates

Table 1: Basic requirements from the ASTM F2411 standard definition and Highly Capable UAV Payloads Planning Document (HCU)

F2411 Class	Class 3: Autonomous air and ground sense and avoidance (HCU 7.1.4.2.4, HCU 7.3.5)
Required Miss Distance	500 ft (F2411 4.2.1)
Field of Regard	270°(H) × 40°(V) (HCU 7.3.1.1.1) 220°(H) × 30°(V) (F2411 4.2.2)
Minimum detection range	3 statute miles (HCU 7.3.1.1.2); "at a range to allow a resolution maneuver that results in a required miss distance of 500 ft or greater." (F2411 4.2.1)
Angular Resolution	0.2 mrad or 0.011° (HCU 7.3.1.1.2.1)
Environmental	Day, night, bright light, and any weather as long as there is 3 statute miles visiblity (HCU 7.3.1.1.2.1, HCU 7.3.1.2.4)
Accuracy and reliability	False alarm rates, false positive rates to be decided (HCU 7.3.1.2.3)

and low computational complexity.

We have developed a multi-stage method that starts with a large number of candidates and winnows these down. The approach is a combination of existing approaches demonstrated in the literature. We start with a morphological filter [14] that looks for high contrast regions in the image that are most likely to be aircraft. Next we use a classifier that has been trained on positive and negative examples and finally we track the candidates over time to remove false positives. We chose the tracking algorithm based on Shafique and Shah [22] at the recommendation of Yilmaz et al.'s object tracking survey [24]. Neither (extended) Kalman filters nor particle filters were appropriate based on the number of false positives present. Other candidates were Multiple Hypothesis Tracking (MHT) and Joint Probability Data Association (JPDA). It was the simplest method to implement and best adapted to situations of the kind we ran into: large number of distracting false positives; non-gaussian. Below we discuss each "stage" of detection in detail.

4.1 Stage 1: Morphological Filtering

In the first stage, we apply a morphological filter that detects deviations from the background intensity. We use two types, one favors dark targets against lighter backgrounds (positive), and the other favors light targets against darker backgrounds (negative). The positive morphological filter takes the form:

$$\mathscr{M}^{+}(x,y) = \mathscr{I}(x,y) - \max\{\max_{|i|\leq w}\min_{|j|\leq w}\mathscr{I}(x+i+j,y), \max_{|i|\leq w}\min_{|j|\leq w}\mathscr{I}(x,y+i+j)\} \quad (1)$$

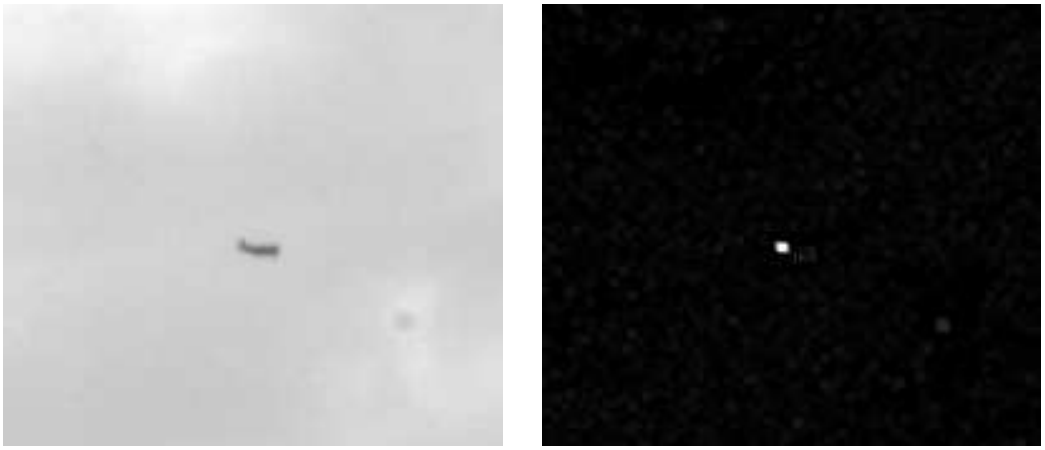

Figure 2: The image on the left shows part of the image of the Piper Archer II at a range of 2.87 miles. The image on the right shows the result of the morphological operation of the left image in Stage 1 of the processing pipeline. The dark aircraft image shows up as bright white spot.

Here $\mathscr{M}^+(x,y)$ refers to the output of the morphological operator at pixel location x,y in the image $\mathscr{I}$ and w is half the length of the window side being considered around the pixel location x,y. As long as no $2w+1$ sub-window (we used $w=2$) contains all target pixels (higher intensity) and no background pixels (lower intensity), then all sub-windows will contain at least one (darker) background pixel. Since the background could be noisy, the max's have the effect of finding a conservative greatest lower-bound for the background intensity. The difference, then, yields an estimate of the difference between signal and background for the pixel. The negative morphological filter, $\mathscr{M}^-$, swaps min's for max's, and negates the expression. From $\mathscr{M}^+$ we choose the top n_+ pixels above a threshold T_+, while suppressing local non-maxima, and construct a list of detections. We do the same for $\mathscr{M}^-$. Fig.2 shows an example aircraft image and the result of the morphological filtering on the example image. Fig.3 shows all the points of interest (indicated by black boxes) detected by the morphological operator in an image. Although the true location (indicated by the bold black box) of the aircraft is detected by the morphological operator, a large number of false positives can be seen in the image. The subsequent classification and tracking stages reduce the false positives significantly.

4.2 Stage 2: Construction of a Shape Descriptor and SVM-based Classification of Potential Targets

4.2.1 Stage 2 a: Construction of a Shape Descriptor

In order to discriminate between the actual image of the aircraft and the large number of false positives that are generated by the morphological filter in Stage 1 it is necessary to find a representation of the target that distinguishes it from false positives. In this section we describe the 19-attribute shape descriptor that we calculate for each subwindow in an image frame

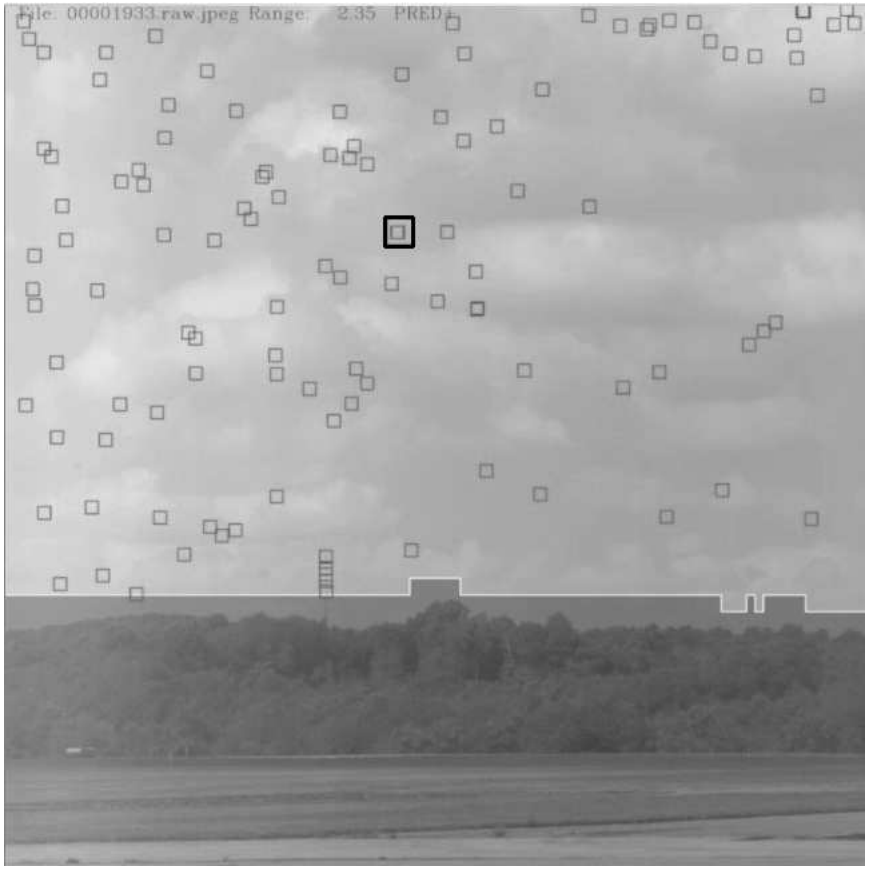

Figure 3: Stage 1 output on example image. Example output of the morphological filter on an image. The black boxes are the points of interest detected by the detector. The bold black box represents the manually located aircraft. The below the horizon part of the image is automatically detected and not processed further. Although the position in the image of the approaching aircraft has been correctly detected by the morphological filter a large number of false positives are present in the frame

which contains the aircraft image. Using such a representation allows us to train a classifier (See Section 4.2.2) to detect false positives and remove them from further consideration.

For each detection we fit a Gaussian function to its $(2r+1)\times(2r+1)$ sub-window (we settled on $r=7$) and construct a shape descriptor for the detection. Through trial and error we found a descriptor that was a good discriminator in test sequences. The descriptor encodes the parameters of the fitted Gaussian, as well as statistics computed from the residual image. We use an axis-aligned Gaussian, parameterized as follows:

$$\mathscr{G}(x,y;\sigma_x,\sigma_y,b,s) = b + \frac{s}{2\pi\sigma_x\sigma_y} e^{-\frac{x^2}{2\sigma_x^2}-\frac{y^2}{2\sigma_y^2}} \tag{2}$$

Here σ_x,σ_y are the standard deviations of the Gaussian along the x,y directions in the image, b is the background intensity and s is the intensity of the image of the aircraft or target.

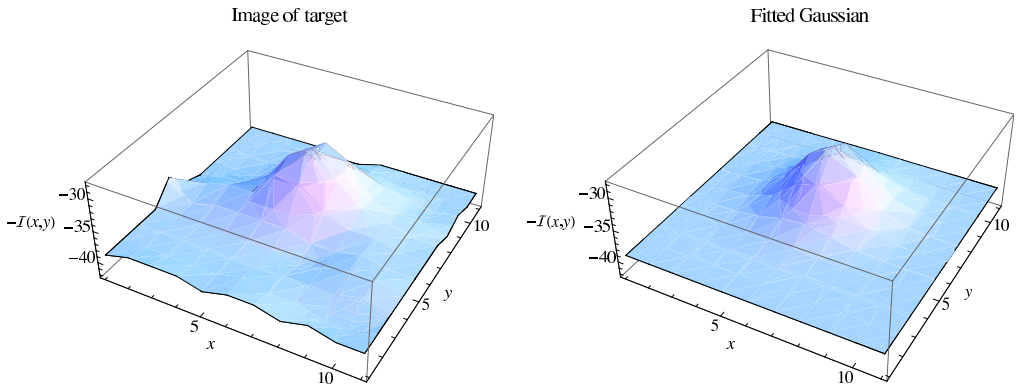

Figure 4: (Left) An example of an image of a target shown with inverted intensity.The aircraft image is the lump in the middle of the height field. This is an 11×11 subwindow centered around the location of the image of the aircraft. (Right) Shows the fitted Gaussian shape to the intensity profile of the aircraft image in the left figure.

We center the Gaussian at the the pixel with the largest absolute deviation from the window's mean intensity. We use gradient descent to minimize the sum of square errors between the input sub-window and $\mathscr{G}(\cdot;\xi)$, minimizing over $\xi = (\sigma_x, \sigma_y, b, s)$. To do this efficiently, we avoid repeated calls to the exponential function by pre-computing both a set of template $\mathscr{G}$'s over a range of σ_x and σ_y pairs, with $(b,s) = (0,1)$, and a set of finite difference approximations to the partial derivatives of $\mathscr{G}$ with respect to σ_x and σ_y. Fig.4 shows an example aircraft image's intensity profile and the fitted two dimensional Gaussian window centered on the image of the aircraft.

Using the best fitting Gaussian $\mathscr{G}^*$, we compute a shape descriptor from the residual difference between the input image and $\mathscr{G}^*$ in upper-left (UL), upper-right (UR), lower-left (LL), lower-right (LL), and center (C) regions. We construct both positive and negative half-sign sums. For example:

$$S^+_{\mathrm{UL}} = \sum_{\substack{1\le x\le w \\ 1\le y\le w}} \max\left[0, \mathscr{G}^*(x,y) - \mathscr{I}(x,y)\right]$$

$$\cdots$$

$$S^-_{\mathrm{C}} = -\sum_{\substack{w/2<x<3w/2 \\ w/2<x<3w/2}} \min\left[0, \mathscr{G}^*(x,y) - \mathscr{I}(x,y)\right]$$

Then, we construct min's and max's of positive and negative half-sign sums, e.g., $S^+_{\max} = \max\left(S^+_{\mathrm{UL}}, \ldots, S^+_{\mathrm{C}}\right)$, and for each statistic we take its log normalized by the background intensity b, e.g., $s^+_{\max} = \log\left(S^+_{\max}/b\right)$. We also compute the estimated signal to background ratio:

$$\mathrm{SBR} = \frac{|b| + |s|/2\pi\sigma_x\sigma_y}{|b|} \tag{3}$$

Finally, the shape descriptor we use is:

$$\mathbf{d} = \big(b, s, \sigma_x, \sigma_y, \mathrm{SBR}, s^+_{\min}, s^+_{\max}, s^-_{\min}, s^-_{\max}, \\ s^+_{\mathrm{UL}}, s^+_{\mathrm{UR}}, s^+_{\mathrm{LL}}, s^+_{\mathrm{LR}}, s^+_{\mathrm{C}}, s^-_{\mathrm{UL}}, s^-_{\mathrm{UR}}, s^-_{\mathrm{LL}}, s^-_{\mathrm{LR}}, s^-_{\mathrm{C}}\big)$$

We associate this 19-dimensional vector with each detection.

4.2.2 Stage 2 b: SVM-based Classification of Potential Targets

The computed shape descriptors, **f** are then used to train a supervised classifier. We chose a *support vector machine* (SVM) [11] as the supervised classifier. The SVM was trained using descriptors from positive and negative examples taken from a sequence of hand-labeled image. For negative examples we used the false negatives produced by the morphological filter. We used radial basis functions as the kernel and obtained a probability of detection from a true target using the distance of an example from the classification boundary. During training we construct empirical densities of x for positive (p^+_x) and negative (p^-_x) classes using a mixture of Gaussians, and store a log-likelihood ratio function $\ell(x) = \log p^+_x / p^-_x$ in a look-up table keyed on x. We choose the kernel bandwidth just large enough to make the odds monotonic in x.

We keep only those detections whose odds exceed a minimum value of $p_{\min}$. Fig.5 shows an example image where as compared against Fig.3 a large percentage of the false positives have been eliminated. The few remaining false positives are eliminated in the tracking stage.

4.3 Stage 3: Tracking

The purpose of this stage is to track detections over time, associating detections to a list of tracked targets. Since many of the false positives are intermittent, we also use tracking to reduce the false positive rate. We arrived at a simple procedure for target tracking that provides a full screen tracking system for high definition imagery.

First, we always maintain a list of targets, and in steady state, it is the job of the tracker to associate to every existing target a detection. With any remaining detections, it also decides whether to create new targets.

For each existing target we consider a set of candidate detections, which are chosen from a wide search area around the predicted position of the target. Detections outside this area are not considered for pairing with that particular target. This gating [16] technique prevents targets from being associated with unlikely candidate detections. For each potential matching detection, we evaluate the likelihood that the target and detection are associated given their respective descriptors.

Then, given a list of the likelihoods for the possible pairings we construct a graph with a node for each target and each detection, and edges between possible pairings, where the weights are the log likelihoods of pairings.

We construct a cost matrix, whose rows correspond to targets, columns to detections, and entries are the log likelihoods of the potential pairing, with $-\infty$ given to non-candidate pairings. The goal is to choose entries from the matrix, no more than one from every row and no more than one from every column such that the sum of probabilities is a maximum. We use the Hungarian algorithm to find this matching [20]. For the number of targets we typically have, usually less than 200, this computation can be computed in less than 8 milliseconds.

Figure 5: Stage 2b output on example image. The points of interest detected by the morphological filter in Stage 1 are assigned probabilities by a trained Support Vector Machine (SVM) of being true positives. Only those detections which exceed a minimum value are kept. This step eliminates a large percentage of the false positives. As compared to Fig.3 the above image contains only a few false positives. The true location of the aircraft in the image is represented by the bold black box and the remaining points of interest after SVM based false positive elimination are shown by black boxes. In spite of the elimination of a large percentage of the false positives, the true positive (inside the bold black box) is correctly preserved.

Fig.6 shows the result of tracking consecutive frames on an example image. The track history is overlaid as a black trail on the image. The bold black box represents the actual location of the aircraft and the black box inside the bigger box the correctly tracked location. The other black boxes represent detected dead pixels and dust particles on the camera lens. Dead pixels and dust particles are very persistent and we detect such objects and remove them from further consideration. A list of known locations of dead pixels in an imager can also be supplied.

5 Data Collection

We collected ground to air imagery of aircraft with ten different camera/lens combinations at Penn State Electro-Optics facilities at Jimmy Stewart Airport, Indiana, Pennsylvania, USA. The data was collected on four separate days of July $2^{nd}, 16^{th}, 22^{nd}$ and September 5^{th}, 2008. Four different infra-red cameras were also used to acquire imagery. The infra-red imagery has

Figure 6: Stage 3 output on example image. The detected track history of the actual aircraft in the image has been overlaid as a black trail on the image. The actual aircraft position is represented by the bold black box and the detected and tracked position by the smaller black box which is shown positioned on the groundtruth location. The other black boxes represent detected dead pixels and detected dust particles on the camera lens. There are no false positives remaining in Stage 3 in contrast to Stages 1 and 2. See Fig.3 and Fig.5. The inset image on the bottom left of the figure shows the aircraft zoomed in on the example image as detected by the algorithm.

not been fused with the visual spectrum imagery yet. The total list of combinations of visual spectrum cameras we tested is listed below in 2.

For the ground-to-air data collection we used a Directed Perception D-100 pan tilt unit (PTU) with a geo-pointing module (GPM). We mounted the cameras on the PTU and used a Piccolo autopilot to send the GPS stream from the approaching aircraft to the GPM. The GPM automatically computes the angles the PTU needs to be pointing at in order to always have the aircraft in the field of view of the cameras. A Piper Archer II was used as the approaching aircraft during the data collection. Fig.7 shows the pan tilt unit and cameras setup used at the airport.

In Fig.8 we show the pattern the approaching aircraft flew as we gathered imagery on 22^{nd} July, 2008. The approaching aircraft flew out to 5-6 miles each pass. Till date we have collected 2.5 terabytes of imagery of which in 2 terabytes the position of the aircraft has been picked out by hand for ground truth purposes. This corpus of real imagery has been used to analyze the performance of our algorithm.

Camera	Lens	Res.	mpix	FOV
IPX- 4M15	Nk105mm	2048×2048	4.2	11°×11°
IPX- 16M3	Nk105mm	4872×3248	15.8	25°×17°
IPX- 4M15	Zs85mm	2048×2048	4.2	13°×13°
IPX- 16M3	Zs85mm	4872×3248	15.8	31°×21°
IPX- 4M15	Zs50mm	2048×2048	4.2	22°×22°
IPX- 16M3	Zs50mm	4872×3248	15.8	50°×34°
Lu125	CmpVF@36mm	1280×1024	1.3	18°×14°
IPX- 2M30	CmpVF@36mm	1600×1200	1.9	24°×18°
Lu125	Cmp25mm	1280×1024	1.3	25°×20°
IPX- 2M30	Fujinon 25mm	1600×1200	1.9	33°×25°

Table 2: This table lists the camera/lens pairs that we tried during testing. In order, this table lists in each column: (i) the camera; (ii) the lens; (iii) the resolution of the camera; (iv) the number of megapixels (megapixel); (v) the field-of-view of the lens.

Figure 7: Directed Perception D-100 pan tilt unit was used along with a geo-pointing module to automatically keep the approaching aircraft in the field of view of the mounted cameras. The pan tilt unit tracked the location of the approaching aircraft using the GPS stream over a radio modem from the aircraft.

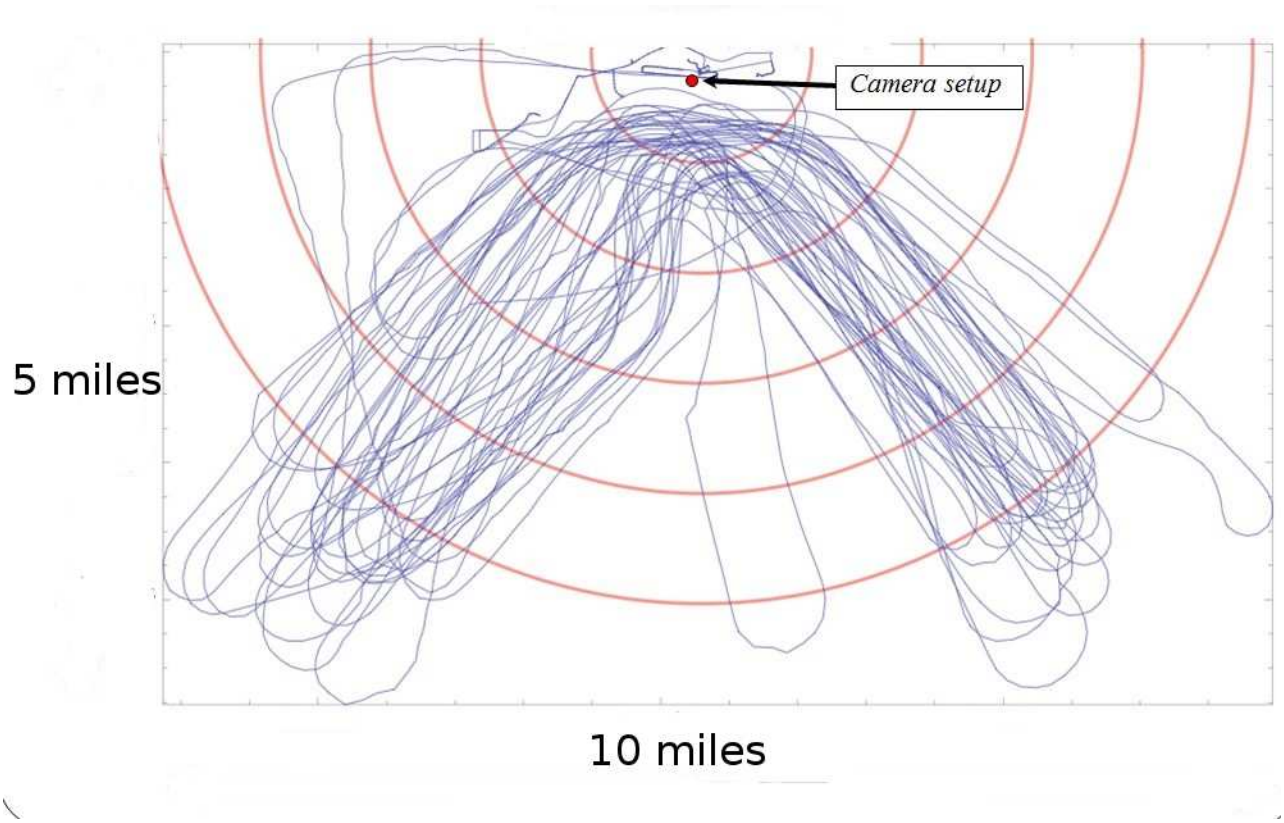

Figure 8: A manned aircraft equipped with a GPS was flown in a series of flights such that it was in the field of view of the ground based cameras on 22^{nd} July, 2008.
The circles show the distance to the cameras in miles.

6 Detection and Tracking Results

We evaluated the performance of each stage of the algorithm using receiver operator characteristic curves (ROC) curves, which measure specificity (ability to reject outliers) and sensitivity (ability to detect true target) of a detector on about 2 terabytes of imagery of above the horizon flying aircraft. Here we present the results using the imagery from the Imperx 4 megapixel camera with Zeiss 85 mm lens which includes data from four days of data collection.

Stage 2 improves the false positive rate by a factor between 6 and 17 depending on the detection rate over Stage 1. Refer Table.3. We get a vast improvement with tracking in Stage 3. In the case of both Stage 1 and Stage 2 the variable affecting rates is a threshold. For Stage 1, the threshold is the value returned by the morphological filter at the detection. For Stage 2, the threshold is the probability according to the SVM classifier, that the detection is a target.

Fig.9 on the left shows the ROC curve for Stage 1, Stage 2 and Stage 3 of the algorithm. Whereas before the value affecting performance was a threshold on the output of a filter or classifier, in this case the threshold is the number of frames for which a target has been tracked. It is to be noted that the best overall detection rate of Stage 3 is higher than the best overall detection rate of Stage 2, even though it is based on the output of Stage 2. We believe that this

Table 3: Shows the number of false positives per frame for Stage 1 and Stage 2 as a function of the true positive percentage. Stage 2 reduces the false positive rate by a factor between 6 and 17 times as Stage 1.

TP %	***Stage 1*** **FP/frame**	***Stage 2*** **FP/frame**	**FP Reduction Factor**
95%	120	20	5.9×
90%	66	3.9	17×
80%	14	1.0	14×
70%	8	0.66	12×
60%	6.2	0.56	11×

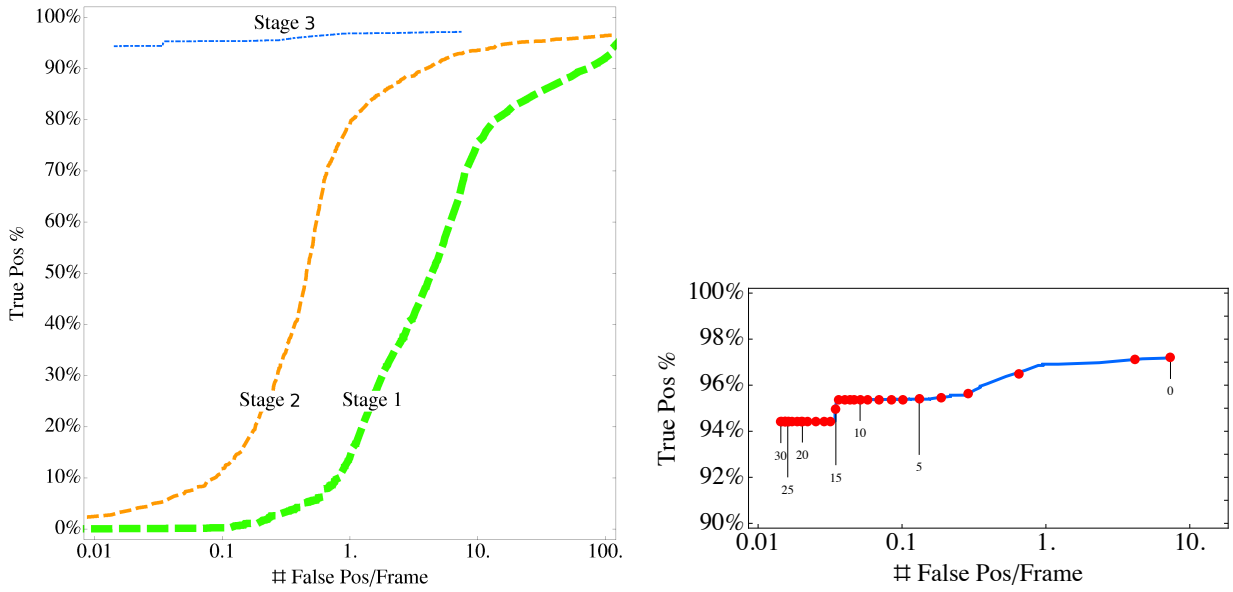

Figure 9: Figure on the left shows the ROC curve for true positive and false positive for the 3 main stages of the algorithm on 4 mp imagery. The curve for Stage 3 shows almost perfect tracking rate with a false positive rate of as low as 0.02 per frame. Figure on the right shows the effect of varying the minimum number of frames that a potential target has to be tracked for before being declared as a target in Stage 3. This curve is very flat as most true positives have long tracks and false positives have short tracks.

is a temporal effect, in that detections that are intermittently below threshold, are picked up by the tracker. The detection rate decreases slightly at closer ranges. This is due to the fact that the algorithms were not optimized for close ranges.

Fig.9 on the right shows the effects of the variance of the minimum number of frames that

a potential target has to be tracked for before being declared as a target. The points on the curve are the number of frames that a target has to have been tracked for, for it to be declared a possible target. In our experiments we let this threshold go up to 30 frames, at which point the false positive rate was 0.014 FP/frame and detection rate was 95%. It is to be noted that this curve is very flat. Most of the true positives have long tracks and almost all outliers have short tracks.

Overall there is a significant decrease in the number of false positives per frame. We add to Table.3 the results of Stage 3 and present them in Table.4. The entries for tracking rates below 95% are not filled in because we chose not to evaluate the threshold frames beyond 30. If we had, the detection rate would have eventually fallen. We find that tracking in Stage 3 improves the false positive rate by a factor of over 500 times over Stage 2.

Table 4: Performance of various stages of the algorithm. Stage 3 achieves a false positive reduction rate of 571 times over Stage 1

TP %	***Stage 1*** **FP/frame**	***Stage 2*** **FP/frame**	***Stage 3*** **FP/frame**	**FP Reduction Factor**
97%	-	-	7.3	-
95%	120	20	0.035	571×
90%	66	3.9	-	-
80%	14	1.0	-	-
70%	8	0.66	-	-
60%	6.2	0.56	-	-

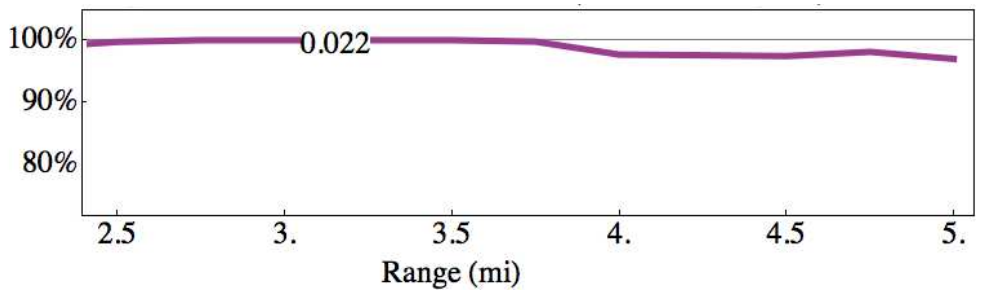

Figure 10: Tracking rate vs range (miles) for a false positive rate per frame of 0.022. Between the range of 2.5 and 3.75 miles the tracking rate is nearly 100%.

We found a reasonable compromise in false positive and true positive rate when we insisted that targets be tracked for at least 10 frames. Then the overall tracking rate was 95%, the false positive rate was 0.05 false positives per frame. From Fig.10 the tracking rate is nearly 100% between 2.5 and 3.75 miles. Since the signal-to-background ratio increases as the range to

the intruder aircraft decreases it can be assumed that our detection and tracking algorithm will continue to have nearly 100% tracking rate at ranges lower than 2.5 miles. If we let the targets be tracked for 25 frames then the false positive rate is reduced to about 0.02 false positives per frame (1 false positive in every 50 frames).

About 80% of the false positives that made it through the tracking of at least 10 frames were items that are of interest to collision avoidance. Most of the false positives were birds or landmarks on the ground that were not segmented out by the horizon detector (e.g. an antenna in the distance). These targets are of interest and could be considered useful. If we account for the useful targets detected then the false positive rate at the 10 frames tracking threshold drops from 0.05 to 0.01 false positives per frame (1 false positive in every 100 frames). At the 25 frames tracking threshold the false positive rate drops to 0.004 false positives per frame (1 false positive every 250 frames).

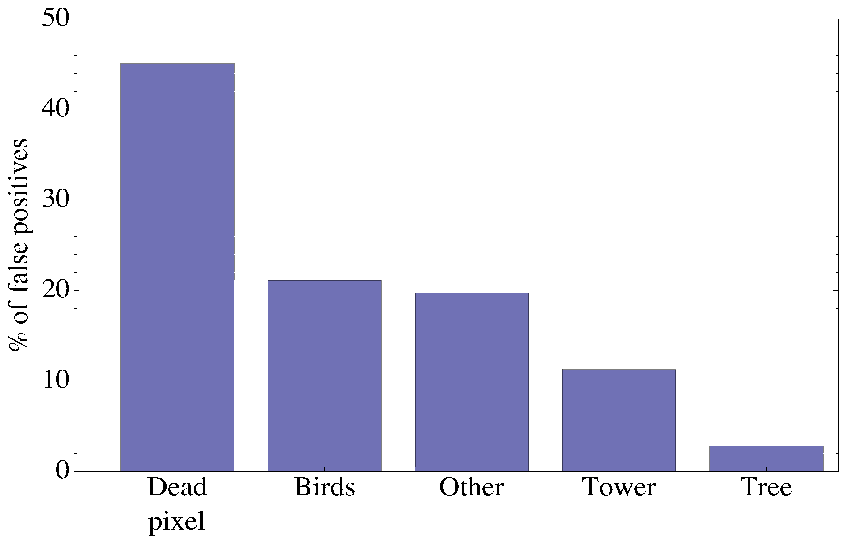

Figure 11: About 80% of detected false positives which are not aircraft are not really false positives and can be of advantage to the overall system. The false positives under the 'Other' category in the figure were due to image regions that looked like aircraft at a distance and were not discriminative enough to be rejected by the classifier. This group can be considered as real false positives.

It is to be emphasized that the tracking rate is the accuracy of detecting the aircraft in each of the image frames containing the aircraft for a given approach of the aircraft averaged over all the approaches of the intruder aircraft towards the ground station (See Section 5). The data we have used contains about 40 such approaches. Our detection algorithms were able to successfully detect an incoming aircraft on every approach starting at 5 miles. Before this system can be fielded extensive testing on several thousand hours of data taken under different atmospheric conditions and geographic locations will have to be carried out.

We have developed and demonstrated a vision based algorithm that achieves a reasonable true positive rate of approximately 98% out to 5 statute miles and a false positive rate of 1 in every 50 frames which exceeds the minimum range requirement of 3 statute miles imposed by regulatory requirement.

7 Future Work

Detecting aircraft below the horizon presents different challenges. It is not possible to eliminate below the horizon part of the scene and search for aircraft in the sky in scenes where the aircraft is against the ground and hence can have a much more confusing background. Collecting such below-the-horizon imagery of aircraft and evaluating detection performance is essential for a complete system.

Currently, our system detects bearing to targets that must be avoided. An important extension will be to estimate range to the target so that precise maneuvers can be planned. We are currently investigating active ranging systems that can be pointed at potential targets, to estimate range and further reduce false positives.

Fusing infrared imagery with visible spectrum imagery is another area that we need to address that can help further reduce false positives.

Although the current algorithm takes about 0.8 seconds per 4 mp frame on an AMD Athlon X2 3800+ processor, most of the computation is image processing and hence amenable to parallelization. Specialized hardware like Digital Signal Processors are promising.

All of the above issues affect how a collision detection and warning system should be designed so as to cover the desired field-of-regard.

8 Funding

This material is based upon work supported by the Unique Missions Division, Department of the Army, United States of America under Contract No. W91CRB-04-C-0046. Any opinions, findings and conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Department of Army or Penn State EOC.

A Appendix: Index to Multimedia Extensions

The multimedia extensions to this article are at `http://www.ijrr.org`

Extension	Type	Description
1	Video	Illustration of multi-stage aircraft detection algorithm (multimedia_extension.mov)

References

[1] Aai corporation. http://www.aaicorp.com/products/uas/uas_main.html.

[2] Aerovironment, inc. http://www.avinc.com/uas/small_uas/.

[3] Aerovironment, inc. http://www.avinc.com/uas/small_uas/.

[4] Micro uav.com. http://www.microuav.com/tigershark.html.

[5] Standard specification for design and performance of an airborne sense-and-avoid system, f2411-04. ASTM International, West Conshohocken, PA, July 2004.

[6] Radio technical commission for aeronautics, inc. http://www.rtca.org/CMS.../SC-203-TOR-PMC20\%Approved-04-26-2010.pdf, 2010.

[7] J.F. Arnold and H. Pasternack. Detection and tracking of low-observable targets through dynamic programming. *Proceedings of SPIE*, 1305:207, 1990.

[8] R. Bernier, M. Bissonnette, and P. Poitevin. Dsa radar - development report. In *UAVSI, 2005*, 2005.

[9] G. Bradski. The OpenCV Library. *Doctor Dobbs Journal*, 25(11):120–126, 2000.

[10] R. Carnie, R. Walker, and P. Corke. Image processing algorithms for UAV "sense and avoid". In *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*, pages 2848–2853, 2006.

[11] Nello Cristianini and John Shawe-Taylor. *An Introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge University Press, 2000.

[12] Federal Aviation Administration. Order 7610.4, November 2006.

[13] M.F. Fernandez, T. Aridgides, and D. Bray. Detecting and tracking low-observable targets using IR. *Proceedings of SPIE*, 1305:193, 1990.

[14] Tarak Gandhi, Mau-Tsuen Yang, Rangachar Kasturi, Octavia I. Camps, Lee D. Coraor, and Jeffrey McCandless. Detection of obstacles in the flight path of an aircraft. In *CVPR*, pages 2304–2311. IEEE Computer Society, 2000.

[15] Christopher Geyer, Debadeepta Dey, and Sanjiv. Singh. Prototype Sense-and-Avoid System for UAVs. *Technical Report CMU-RI-TR-09-09, Robotics Institute, Carnegie Mellon University*, 2009.

[16] M.O. Kolawole. *Radar systems, peak detection and tracking*. Newnes, 2002.

[17] J. McCalmont, J. Utt, and M. Deschenes. Detect and Avoid Technology Demonstration. In *Proceedings of the American Institute of Aeronautics and Astronautics Infotech*, 2002.

[18] J.W. McCandless. Detection of aircraft in video sequences using a predictive optical flow algorithm. *Optical Engineering*, 38:523, 1999.

[19] Office of the Secretary of Defense. Airspace integration plan for unmanned aviation, Nov 2004.

[20] Christos Papadimitriou and Kenneth Steiglitz. *Combinatorial Optimization: Algorithms and Complexity*. Courier Dover Publications, 1998.

[21] Stavros Petridis, Christopher Geyer, and Sanjiv Singh. Learning to detect aircraft at low resolutions. In Antonios Gasteratos, Markus Vincze, and John K. Tsotsos, editors, *ICVS*, volume 5008 of *Lecture Notes in Computer Science*, pages 474–483. Springer, 2008.

[22] K. Shafique and M. Shah. A noniterative greedy algorithm for multiframe point correspondence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(1):51–65, 2005.

[23] J. Utt, J. McCalmont, and M. Deschenes. Development of a sense and avoid system. In *AIAA Infotech at Aerospace*, Sep 2005.

[24] A. Yilmaz, O. Javed, and M. Shah. Object tracking: A survey. *ACM Computing Surveys (CSUR)*, 38(4):13, 2006.