

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

Debadeepta Dey, Christopher Geyer, Sanjiv Singh, Matt Digioia

## 1 Introduction

Unmanned Aerial Vehicles (UAVs) typically fly blind with operators in distant locations. Most UAVs are too small to carry a traffic collision avoidance system (TCAS) payload or transponder. Collision avoidance is currently done by flight planning, use of ground or air based human observers and segregated air spaces. US lawmakers propose commercial unmanned aerial systems access to national airspace (NAS) by 30th September 2013. 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 (1) proposed requirements which include a field of regard of  $220^\circ$  (horizontal)  $\times 30^\circ$  (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

---

Debadeepta Dey  
Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, USA e-mail: debadeep@cs.cmu.edu

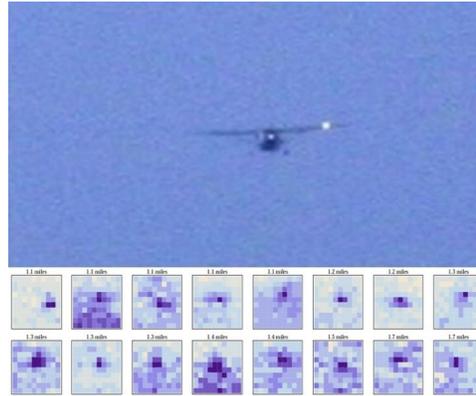
Christopher Geyer  
Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, USA e-mail: cgeyer@cs.cmu.edu

Sanjiv Singh  
Carnegie Mellon University, 5000 Forbes Avenue e-mail: ssingh@ri.cmu.edu

Matthew Digioia  
The Penn State Electro-Optics Center, 222 Northpointe Blvd Freeport, Freeport, USA e-mail: mdi-  
gioia@eoc.psu.edu

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.

**Fig. 1** Selection of  $11 \times 11$  subwindows showing the image of the Piper Archer II which was used as the intruder 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  $30^\circ(H) \times 21^\circ(V)$ .



cameras can be used to cover the wide field-of-regard. A typical image of an aircraft at a range of 3 miles is a few pixels in diameter. Fig.1 shows a  $11 \times 11$  window around the image of the intruder 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 (2) for UAVs. Passive vision provides a low cost, low power solution albeit at the cost of a relatively high false positive rate. Using an approach based on morphological filters augmented with a trained classifier we have been able to obtain 98% detection rate out to 5 statute miles and a false positive rate of 1 every 50 frames.

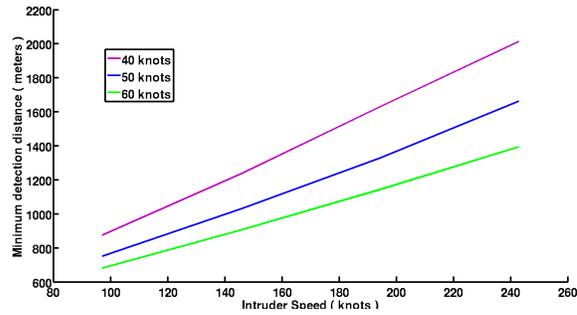
In section 2 the related work in sense and avoid systems is outlined. In section 3 we discuss the details of the vision based aircraft detection algorithm. In section 4 we outline our efforts to collect imagery of flying aircraft, and details about the result of our algorithm on the corpus of real ground truthed imagery of aircraft. Finally in section 5 we discuss the path forward towards a field deployable sense and avoid system.

## 2 Related Work

Utt et al. (20) describe a fielded vision-based sensory and perception system. McCandless (14) proposes an optical flow method for detecting moving aircraft. The use of morphological filtering is popular in computer vision based sense and avoid systems (11; 4). This approach generates a significant number of false positives. Petridis et al. use AdaBoost (17) to detect aircraft in low resolution imagery. Track-Before-Detect (TBD) (9) is an approach used mainly on infrared imagery. Defence Research Associates have implemented a vision based sense and avoid system on a Predator UAV system (13).

Efforts to directly model the range of atmospheric conditions under VFR remain untouched. A field deployable sense and avoid system must be able to operate in a variety of atmospheric conditions including fog, haze and directly against the glare of the sun. The operation of the system must not degrade beyond an acceptable level under all these conditions. We have developed an image formation model which ac-

**Fig. 2** The minimum detection distance required to guarantee collision avoidance for varying intruder speeds and ownship speeds of 40, 50 and 60 knots. For the best case scenario the minimum detection distance is 700 meters and for the worst case scenario the minimum detection distance is 2100 m



counts for the various atmospheric conditions. We used this model to predict the signal to background ratio of the image of the aircraft. The image formation model is described in detail in (12). The model also allows the determination of 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.

## 2.1 Requirements

The range requirements of an aircraft detection system are influenced by two main factors: regulations developed by the FAA and the maneuvering capabilities of the UAV.

Duke et al.(6) and Schaeffer et al.(18) outline core competencies needed by a human equivalent system. The human equivalence mandated by OSD (15) and ACC (7) require vehicles to avoid non-cooperative vehicles without such systems. Shakernia et al. (19) leverage work of Utt et al. (20) on using maneuvers to reduce the intrinsic uncertainty about range when using an image based detector.

In order to decide the range requirements of the system we opted for a collision avoidance system by Frazzoli et al. (10) that enables an UAV to aggressively maneuver without breaching its envelope. Further details are in (12).

## 3 Approach

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 and low computational complexity.

We have developed a multi-stage method that starts with a large number of candidates and winnows these down. We start with a morphological filter 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. Below we discuss each "stage" of detection in detail.

### 3.1 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:

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

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,  $\mathcal{M}^-$ , swaps min's for max's, and negates the expression. From  $\mathcal{M}^+$  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  $\mathcal{M}^-$ . Fig.3 shows an example aircraft image and the result of the morphological filtering on the example image.

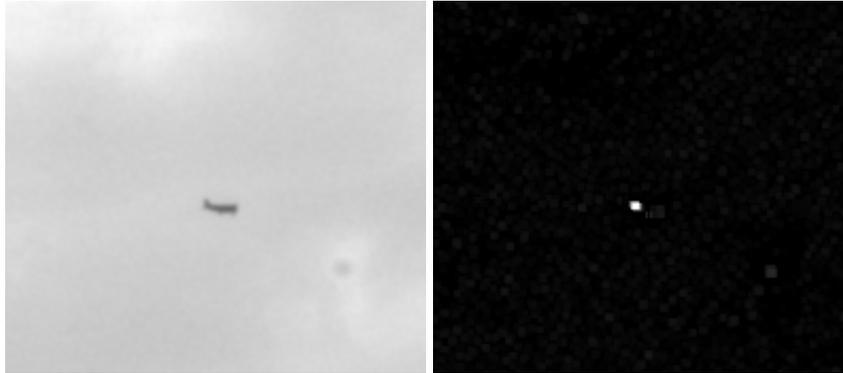


Fig. 3: The image on the left shows part of the image of the Piper Archer II at a range of 2.87 miles. The image on the right shows the result of the morphological operation of the left image in Stage 1 of the processing pipeline. The dark aircraft image shows up as bright white spot.

### 3.2 Stage 2: Construction of a Shape Descriptor

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

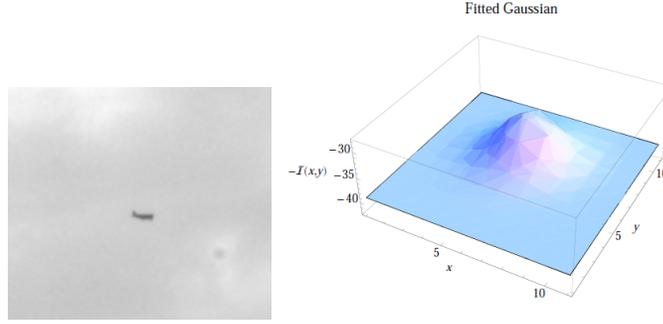


Fig. 4: The image on the left shows an image of the aircraft at 2.87 miles. The figure on the right shows the two dimensional Gaussian function fitted on the  $15 \times 15$  subwindow centered on the image of the aircraft in the right image in Stage 2 of the processing pipeline.

computed from the residual image. We use an axis-aligned Gaussian, parameterized as follows:

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

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  $\mathcal{G}(\cdot; \xi)$ , 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  $\mathcal{G}$ 's over a range of  $\sigma_x$  and  $\sigma_y$  pairs, with  $(b, s) = (0, 1)$ , and a set of finite difference approximations to the partial derivatives of  $\mathcal{G}$  with respect to  $\sigma_x$  and  $\sigma_y$ . Fig.4 shows an example aircraft image and the fitted two dimensional Gaussian window centered on the image of the aircraft.

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

$$S_{\text{UL}}^+ = \sum_{\substack{1 \leq x \leq w \\ 1 \leq y \leq w}} \max[0, \mathcal{G}^*(x, y) - \mathcal{I}(x, y)]$$

...

$$S_{\text{C}}^- = -\sum_{\substack{w/2 < x < 3w/2 \\ w/2 < y < 3w/2}} \min[0, \mathcal{G}^*(x, y) - \mathcal{I}(x, y)]$$

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

$$\text{SBR} = \frac{|b| + |s|/2\pi\sigma_x\sigma_y}{|b|} \quad (3)$$

Finally, the shape descriptor we use is:

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

We associate this 19-dimensional vector with each detection.

### 3.3 Stage 3: SVM-based Classification of Potential Targets

The next stage of the algorithm is to pass the computed descriptor  $\mathbf{f}$  for each detection through a *support vector machine* (SVM) (5). We trained the SVM using shape descriptors computed for positive and negative examples taken from a sequence of hand-labeled images. For negative examples we used the false negatives produced by the morphological filter. We used OpenCV's (3) implementation of an SVM (3), and chose to use radial basis functions for the classifier.

OpenCV's SVM implementation returns a hard classification: positive or negative depending on the sign of a summation, e.g.  $y = \text{sign } x$ , where  $x = \sum_i f_i(\mathbf{d})$ . We want to associate a probability with each detection, so as to make them comparable, and so we modified the implementation to use the value of  $x$  to estimate the probability that the detection is from a true target. 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  $p_{\min}$ .

### 3.4 Stage 4: Tracking

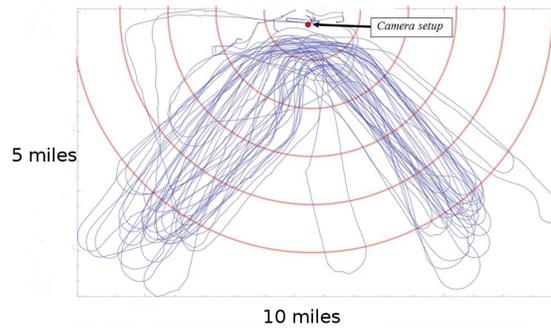
The purpose of this stage is to track detections over time, associating detections to a list of tracked targets. Since many of the false positives are intermittent, we also use tracking to reduce the false positive rate. We arrived at a simple procedure for target tracking that provides a full screen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



**Fig. 5** 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. The circles show the distance to the cameras in miles.

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 (16) to find this matching. For the number of targets we typically have, usually less than 200, this computation can be computed in less than 8 milliseconds.

#### 4 Data Collection and Results

Collaborating with the Penn State Electro-Optics Center (EOC), we collected ground to air imagery of aircraft with ten different camera/lens combinations. Four different infra-red cameras were also used to acquire imagery. The EOCs Payload Development Center, a comprehensive aerial systems integration lab (located on the field at the Jimmy Stewart Airport, Indiana, PA), was utilized to provide hardware, testing and technical support, and flight operations.

We tracked a Piper Archer II flying in the air from the ground with cameras mounted on pan tilt unit and synchronised with the geolocation of the aircraft so that it always remained in field of view of the cameras.

In Fig.5 we show the pattern the intruder aircraft flew as we gathered imagery. Till date we have collected 2.5 terabytes of imagery of which in 2 terabytes the position of the aircraft has been picked out manually for ground truth purposes. This corpus of real imagery has been used to analyze the performance of our algorithm.

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.

Stage 3 improves the false positive rate by a factor between 6 and 17 depending on the detection rate over Stage 1. Refer Table.1 . We get a vast improvement with tracking in Stage 4. In the case of both Stage 1 and Stage 3 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 3, the threshold is the probability according to the SVM classifier, that the detection is a target.

Fig.6 on the left shows the ROC curve for Stage 1, Stage 3 and Stage 4 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

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

TP %	Stage 1 FP/frame	Stage 3 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×

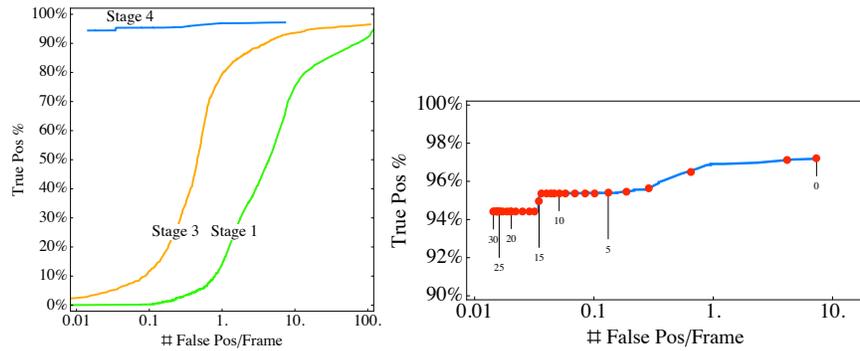


Fig. 6: Figure on the left shows the ROC curve for true positive and false positive for the 3 main stages of the algorithm on 4 mp imagery. The curve for Stage 4 shows almost perfect detection rate with a false positive rate of as low as 0.02 per frame. Figure on the right shows the effect of varying the minimum number of frames that a potential target has to be tracked for before being declared as a target in Stage 4. This curve is very flat as most true positives have long tracks and false positives have short tracks.

of Stage 4 is higher than the best overall detection rate of Stage 3, even though it is based on the output of Stage 3. We believe that this 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.6 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 92%. 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.1 the results of Stage 4 and present them in Table.2. The entries for detection 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 4 improves the false positive rate by a factor of over 500 times over Stage 3.

Table 2: Performance of various stages of the algorithm. Stage 4 achieves a false positive reduction rate of 571 times over Stage 1

TP %	Stage 1 FP/frame	Stage 3 FP/frame	Stage 4 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	-	-

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 detection rate was 95%, the false positive rate was 0.05 false positives per frame. The detection rate is nearly 100% between 2.5 and 3.75 miles.

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.

We have developed and demonstrated a vision based algorithm that achieves a reasonable false positive rate of approximately 98% out to 5 statute miles and a false positive rate of 1 in every 50 frames which exceeds the FAA (8) regulatory requirement of reliable detection out to 3 statute miles.

## 5 Future Work

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 and collecting below horizon imagery are areas we need to address.

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.

## References

- [1] Standard specification for design and performance of an airborne sense-and-avoid system, f2411-04. ASTM International, West Conshohocken, PA (2004)
- [2] Bernier, R., Bissonnette, M., Poitevin, P.: Dsa radar - development report. In: UAVSI, 2005 (2005)
- [3] Bradski, G.: The OpenCV Library. *Doctor Dobbs Journal* **25**(11), 120–126 (2000)
- [4] Carnie, R., Walker, R., Corke, P.: Image processing algorithms for UAV "sense and avoid". In: *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*, pp. 2848–2853 (2006)
- [5] Cristianini, N., Shawe-Taylor, J.: *An Introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge University Press (2000)
- [6] Duke, E., Vanderpool, C., Duke, W.: Turning pinocchio into a real boy: A turing test for UAV operations. In: *AIAA Infotech at Aerospace 2007 Conference and Exhibit* (2007)
- [7] Ebdon, M.D., Regan, M.J.: Sense-and-avoid requirement for remotely operated aircraft (roa) (2004)
- [8] Federal Aviation Administration: Order 7610.4 (2006)
- [9] Fernandez, M., Aridgides, T., Bray, D.: Detecting and tracking low-observable targets using IR. *Proceedings of SPIE* **1305**, 193 (1990)
- [10] Frazzoli, E., Dahleh, M., Feron, E.: Real-time motion planning for agile autonomous vehicles. *Journal of Guidance Control and Dynamics* **25**(1), 116–129 (2002)
- [11] Gandhi, T., Yang, M.T., Kasturi, R., Camps, O.I., Coraor, L.D., McCandless, J.: Detection of obstacles in the flight path of an aircraft. In: *CVPR*, pp. 2304–2311. IEEE Computer Society (2000)
- [12] Geyer, C., Dey, D., Singh, S.: Prototype Sense-and-Avoid System for UAVs. Technical Report CMU-RI-TR-09-09, Robotics Institute, Carnegie Mellon University (2009)
- [13] McCalmont, J., Utt, J., Deschenes, M.: Detect and Avoid Technology Demonstration. In: *Proceedings of the American Institute of Aeronautics and Astronautics Infotech* (2002)
- [14] McCandless, J.: Detection of aircraft in video sequences using a predictive optical flow algorithm. *Optical Engineering* **38**, 523 (1999)
- [15] Office of the Secretary of Defense: Airspace integration plan for unmanned aviation (2004)
- [16] Papadimitriou, C., Steiglitz, K.: *Combinatorial Optimization: Algorithms and Complexity*. Courier Dover Publications (1998)
- [17] Petridis, S., Geyer, C., Singh, S.: Learning to detect aircraft at low resolutions. In: A. Gasteratos, M. Vincze, J.K. Tsotsos (eds.) *ICVS, Lecture Notes in Computer Science*, vol. 5008, pp. 474–483. Springer (2008)
- [18] Schaeffer, R.: A standards-based approach to sense-and-avoid technology. In: *AIAA 3rd Unmanned Unlimited Technical Conference, Workshop and Exhibit* (2004)
- [19] Shakernia, O., Chen, W., Raska, V.: Passive ranging for UAV sense and avoid applications. In: *AIAA's Infotech at Aerospace* (2005)
- [20] Utt, J., McCalmont, J., Deschenes, M.: Development of a sense and avoid system. In: *AIAA Infotech at Aerospace* (2005)