Proceedings of Australasian Conference on Robotics and Automation, 2-4 Dec 2014, The University of Melbourne, Melbourne, Australia

Looming Aircraft Threats: Shape-based Passive Ranging of Aircraft from Monocular Vision

Timothy L. Molloy, Jason J. Ford, and Luis Mejias
School of Electrical Engineering and Computer Science
Queensland University of Technology
Brisbane, QLD 4000, Australia
{t.molloy, j2.ford,luis.mejias}@qut.edu.au

Abstract

This paper proposes new techniques for aircraft shape estimation, passive ranging, and shape-adaptive hidden Markov model filtering which are suitable for a monocular vision-based non-cooperative collision avoidance system. Vision-based passive ranging is an important missing technology that could play a significant role in resolving the sense-and-avoid problem in unmanned aerial vehicles (UAVs); a barrier hindering the wider adoption of UAVs for civilian applications. The feasibility of the proposed shape estimation, passive ranging and shape-adaptive filtering techniques is evaluated on flight test data.

1 Introduction

Vision-based collision avoidance appears to be one of the critical enabling technologies for the integration of autonomous unmanned aerial vehicles (UAVs) in civilian airspace [Lai *et al.*, 2011; 2013b; Dey *et al.*, 2011]. The market growth and community benefits derived from the adoption of UAV rely heavily on UAV being certified for operation in non-segregated airspace, which will most likely require a type of non-cooperative collision avoidance (i.e. solution to the sense-and-avoid problem). For many UAV systems, use of conventional active radar based detection and ranging systems for sense-and-avoid currently seems infeasible due to cost, power and weight considerations [Lai *et al.*, 2013b; Dey *et al.*, 2011]. Both passive acoustic based approaches [Finn and Franklin, 2001] and vision-based approaches [Lai *et al.*, 2011; 2013b; Dey *et al.*, 2011] have been proposed as alternative sense-and-avoid technologies for UAV systems, but vision-based approaches seemed to have gained greater attention.

One of the key challenges faced by vision (or other passive sensor) based sense-and-avoid is that detection must occur whilst the potential collision aircraft is still a considerable distance away (due to the high approach speeds of aircraft). This translates into a requirement for detection of small point-like objects in an image, which presents many challenges for monocular camera based detection algorithms. Early contributions on the vision detection problem are presented in [Arnold *et al.*, 1993; Gandhi *et al.*, 2003; Barniv, 1985], whilst more recent flight tests have demonstrated low-false alarm vision-based aircraft detection at the sort of detection distances required for operational use, see [Lai *et al.*, 2011; 2013b]. Other recent work on the vision-based sense-and-avoid problem include [Cho *et al.*, 2013; Dey *et al.*, 2011], with a survey of published vision-based aircraft detection experiments provided in [Lai *et al.*, 2012].

A critical challenge for vision-based collision avoidance that is yet to be progressed is the problem of range (or pseudo-range) estimation from monocular images. The literature on passive ranging is not extensive. The most notable work on passive range estimation for collision avoidance is the manoeuvre based passive ranging study presented in [Shakernia *et al.*, 2005]. Despite its tremendous potential this technique is still unsuitable due to the time it takes the filter to converge and the manoeuvres required to increase range observability. Alternative methods such as stereo based passive ranging have been discussed, however they currently seem unfeasible due to the detection distances involved and the relatively short stereo baseline available on an aircraft. As a further alternative, this paper proposes new shape based Hidden Markov model (HMM) detection filters that extend recently demonstrated vision-based aircraft detection approaches [Lai *et al.*, 2011; 2013b].

HMMs have seen numerous applications in control, image processing, digital communications and bioinformatics [Tonissen and Evans, 1996; Carlemalm *et al.*, 2000; Chung *et al.*, 1991]. In previous works, in the context of vision-based non-cooperative collision avoidance, we have investigated the use of morphological based HMM filters (combining spatial and temporal filter-

ing) for tracking [Lai *et al.*, 2008; Lai and Ford, 2010; Lai *et al.*, 2011; 2013b], for bearing estimation [Lai *et al.*, 2013a], and for aircraft manuever detection [Molloy and Ford, 2012; 2013]. This previous work established the high-reliability low-false alarm detection performance of morphological-HMM filters for detecting aircraft from other real-life aircraft platforms in the presence of real measurement noise artefacts.

The key contribution of this work is to extend these previous morphological-HMM based techniques to allow estimation of the range to an aircraft that has been detected. In particular, we explore the spatial (and temporal) information produced by these detection filters so that the shape of a tracked aircraft can be estimated; shape information which can be used for the purpose of shape-based passive ranging. The three main contributions of this paper are the proposal of: aircraft shape estimation approaches, passive range estimation approaches, and shape-adaptive HMM filters.

This paper is structured as follows: In Section 2 we introduce our HMM representation, and related HMM filters. In Section 3, we propose a shape estimation algorithm, a range estimation approach and our shape-adaptive HMM filter. In Section 4 some implementation details are presented. Results from flight data are presented in Section 5. We provide conclusions in Section 6.

2 Dynamics and Problem Statement

In this section we introduce a representation of aircraft dynamics using Markov chains, and propose a new multi-pixel HMM filter.

2.1 Aircraft Dynamics as a Markov Chain

At each time instant or frame $k \geq 0$, we are interested in an aircraft's location within a $N_h \times N_v$ pixel grey-scale image G_k from a (visual spectrum) camera. As argued in [Lai *et al.*, 2008; 2011; 2013b; 2013a], the aircraft's motion in this image can be modelled as a Markov chain with $N = N_h N_v$ states. Let $X_k \in S_X$ denote this Markov chain, where S_X is the set of indicator vectors e_i for $1 \leq i \leq N$, and e_i is the $\mathbb{R}^N$ is the vector of all zeros expect for 1 in the ith location. Specifically, at time k, we say the aircraft is located at the ith pixel when the ith Markov state is active, or we can write $X_k = e_i$. A natural consequence of this representation of an aircraft location as a Markov chain is that an aircraft's motion can be modelled by the chain's transition probability matrix which describes the probability of transitioning between Markov states (i.e. between pixel locations) [Lai *et al.*, 2013a; Lai and Ford, 2010]. In the following, we let $A \in \mathbb{R}^{N \times N}$ denote the transition probability matrix, where for $1 \leq i, j \leq N$, the ijth element A^{ij} is defined as the probability that the aircraft may transition from pixel j at time k to pixel i at time $k+1$, see [Lai *et al.*, 2013b] and references within for more details.

2.2 Aircraft Observations in an Image

Noise corrupted aircraft measurements are assumed to be provided for $k \geq 0$ by the morphologically processed grey-scale images

$$y_k \triangleq \mathrm{CMO}\left(G_k, S\right),$$

where $\mathrm{CMO}\left(\cdot, S\right) : \mathbb{R}^N \mapsto \mathbb{R}^N$ is the morphological close-minus-open transform of the grey-scale camera image G_k by the structuring element S (here, a 5×5 square of ones, see [Lai *et al.*, 2013b] and references therein for more details). The measurement conditional densities $b^i\left(y_k\right) \triangleq p(y_k \left| X_k = e_i \right.)$ for $1 \leq i \leq N$ describe the probability of observing the (morphological) image y_k given that the aircraft is located at pixel e_i at time k.

We will denote the measurement of the ith pixel at time k as y_k^i. A pixel measurement may either consist of noise, or noise plus an aircraft. Let $p\big(y_k^i\big)$ be the probability density describing the intensities of pixels occupied by an aircraft. Let $q\big(y_k^i\big)$ be the probability density describing the intensities of pixels unoccupied by an aircraft. Following [Lai *et al.*, 2008], we assume that the pixel values of the morphologically processed images are statistically independent in the sense that $p(y_k^m, y_k^n) = p(y_k^m)p(y_k^n)$, and $q(y_k^m, y_k^n) = q(y_k^m)p(y_k^n)$ for all $1 \leq m, n \leq N$, $m \neq n$.

Single Pixel Aircraft

First let us consider a single pixel aircraft. For a single pixel aircraft it follows that

$$y_k^i \sim \begin{cases} p\big(y_k^i\big) & \text{for } X_k = e_i \\ q\big(y_k^i\big) & \text{for } X_k \neq e_i, \end{cases}$$

for all $1 \leq i \leq N$. Hence, we have that

$$\begin{aligned} b_S^i(y_k) &\triangleq p(y_k | X_k = e_i) \\ &= p\big(y_k^i\big) \prod_{\substack{j=1 \\ j \neq i}}^{N} q\Big(y_k^j\Big) \\ &= \frac{p\big(y_k^i\big)}{q\big(y_k^i\big)} \prod_{j=1}^{N} q\Big(y_k^j\Big). \end{aligned}$$

In principle, the values of $b_S^i\left(y_k\right)$ for $1 \leq i \leq N$ can be empirically estimated on the basis of morphological image sequences (see [Lai *et al.*, 2011]). However, in this paper we will use the approximation $b_S^i\left(y_k\right) = y_k^i + 1$ for all $1 \leq i \leq N$ [Lai *et al.*, 2013b]. We let $B_S(y_k)$ be the diagonal matrix with elements $b_S^i(y_k)$.

Multi-Pixel Aircraft

Let us now consider an aircraft that occupies more than one pixel in an image, and allow the aircraft's apparent size to grow or shrink with time. For each $k \geq 0$, we will denote the set of pixels in the image occupied by the aircraft (or a mask) when it is located at e_i as $M_k(i)$ ($M_k(i)$ can be considered to be the set of pixels formed from a shape template shifted to pixel i). When the aircraft is located at e_i, we have that

$$y_k^j \sim \begin{cases} p\left(y_k^j\right) & \text{for } j \in M_k(i) \\ q\left(y_k^j\right) & \text{for } j \notin M_k(i) \end{cases}$$

for all $1 \leq j \leq N$. Under our pixel statistical independence assumption we have that

$$\begin{aligned} b_M^i(y_k, M_k) &\triangleq p(y_k | X_k = e_i, M_k(i)) \\ &= \prod_{\ell \in M_k(i)} p\left(y_k^\ell\right) \prod_{\substack{j=1 \\ j \notin M_k(i)}}^{N} q\left(y_k^j\right) \\ &= \prod_{\ell \in M_k(i)} \frac{p\left(y_k^\ell\right)}{q\left(y_k^\ell\right)} \prod_{j=1}^{N} q\left(y_k^j\right) \\ &= \alpha_k \prod_{\ell \in M_k(i)} b_S^\ell(y_k) \end{aligned}$$

where α_k is defined as

$$\alpha_k \triangleq \left(\prod_{j=1}^{N} q\left(y_k^j\right) \right)^{1-|M_k|}$$

and $|M_k|$ is the number of pixels that the aircraft occupies at time k. For completeness, where the aircraft is a single pixel at pixel i, we can write $M_k(i) = \{i\}$ and $|M_k| = 1$, and note that $b_M^i(y_k, M_k) = b_S^i(y_k)$.

Using properties of the logarithm, we also note that

$$b_M^i(y_k, M_k) = \alpha_k \exp\left(\sum_{\ell \in M_k(i)} \log b_S^\ell(y_k) \right).$$

Importantly, this relationship implies that the multi-pixel likelihoods $b_M^i(y_k, M_k)$ can be calculated for $1 \leq i \leq N$ by scaling the exponential version of a two-dimensional convolution between an image of single pixel log-likelihoods $\log b_S^i(y_k)$ and the aircraft mask M_k. We let $B_M(y_k, M_k)$ be the diagonal matrix with (diagonal) elements $b_M^i(y_k, M_k)$.

2.3 Hidden Markov Model Filter

The HMM filter corresponding to the above models produces estimates $\hat{X}_k \in \mathbb{R}^N$ of the (centre) location of the aircraft in the image given a sequence of images y_k. The conditional probability mass function $\hat{X}_k$ is given by the recursion

$$\hat{X}_k = N_k B_M\left(y_k, M_k\right) A \hat{X}_{k-1} \tag{1}$$

for $k > 0$ where $\hat{X}_0$ is initial or *a prior* estimate, and N_k are scalar normalisation factors so that $\sum_{i=1}^{N} \hat{X}_k^i = 1$ (recalling that $B_M\left(y_k, M_k\right) = B_S(y_k)$ for a single pixel aircraft).

In the problem of vision-based aircraft detection, the normalisation factors provide a useful detection metric (see [Lai *et al.*, 2011; 2013b]),

$$\eta_k = \frac{L-1}{L} \eta_{k-1} - \frac{1}{L} \log N_k$$

where $\eta_0 = 0$ and $L = 10$ has been found suitable. In previous work, an aircraft has been declared detected when η_k exceeds some given threshold h.

The estimate $\hat{X}_k$ can naturally be considered as describing the (likely) aircraft location in the image. In previous work, we have observed that $\hat{X}_k$ will have (small) non-zero probabilities spread over the pixels occupied by multi-pixel aircraft when the approximation $B_S(y_k)$ is used instead of $B_M\left(y_k, M_k\right)$ (for $|M_k| > 1$). Figure 1 illustrates how multi-pixel aircraft can appear as "shaped clusters" of single pixel aircraft in $\hat{X}_k$ (the vector $\hat{X}_k$ has been unwrapped into an image for display purposes). In this paper, we develop novel methods for estimating aircraft shape and range that exploit these mismatch properties of the single pixel HMM filter estimates. Our novel techniques also provide us with insight into the design of new HMM filters for detecting and tracking multi-pixel aircraft.

3 Aircraft Shape and Range Estimation

In this section, we propose a method for estimating the shape and range of a multi-pixel aircraft using the outputs of a single pixel HMM filter. We will then propose a shape-adaptive HMM filtering that uses the estimated shape information.

3.1 Shape Estimation

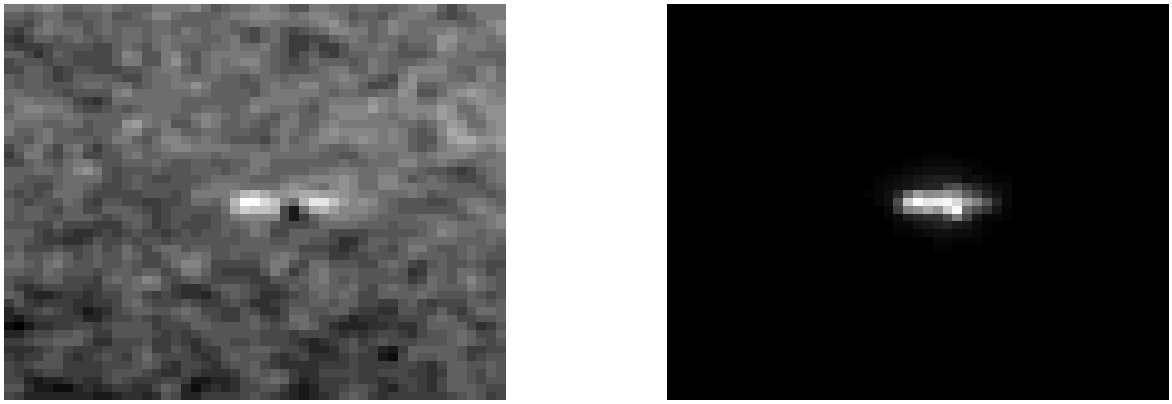
We will now describe the three key stages (illustrated in Figure 2) of our proposed shape estimation approach.

Power Law Image

We first map the elements in the single pixel filter estimate $\hat{X}_k$ to pixel intensities in an $N_h \times N_v$ image I_k using the power law transformation

$$I_k^{i,j} = I_{\max} \left(\hat{X}_k^m \right)^\gamma$$

for all $1 \leq i \leq N_v$ and $1 \leq j \leq N_h$ where $m = i + (j-1)N_v$. Here, $I_{\max}$ is the maximum intensity value (e.g.



(a) (b)

Figure 1: Example of shape information in output of single pixel HMM filters: (a) (cropped) camera image of aircraft, and (b) corresponding (cropped) estimate $\hat{X}_k$ from single pixel HMM filter. All images are cropped to 50×50 pixels.

$I_{\max} = 255$) and $\gamma > 0$ is a constant that controls the spread of the intensities below the maximum intensity value. We have found small γ values (e.g. $\gamma = 0.05$) useful for improving the visibility of multi-pixel aircraft features present in $\hat{X}_k$ (since small γ values assign high intensities to low probabilities).

Our power law images I_k typically contain a dynamic diffuse intensity region surrounding an aircraft. This dynamic diffuse region is a result of the transition dynamics in the single pixel HMM filter, and its mismatched observation model. To reduce the dynamic variations in this diffuse region of interest and to improve the contrast of aircraft, we calculate a cropped exponential moving average intensity image $\bar{I}_k$ with a centre that is aligned with the maximum intensity of I_k. That is, we calculate an average power law image $\bar{I}_k$ with intensities,

$$\bar{I}_k^{i,j} = (1-\lambda) I_k^{i+\delta_i, j+\delta_j} + \lambda \bar{I}_{k-1}^{i,j}$$

for all $1 \leq i, j \leq 40$ where $0 < \lambda < 1$ is a forgetting constant that controls the dynamics in $\bar{I}_k$, and δ_i, δ_j are the vertical and horizontal offsets of the maximum intensity value of I_k from the centre of I_k. We have found $\lambda = 0.95$ as suitable for constructing a relatively slowly time-varying intensity region surrounding aircraft.

Normalised Cross-Correlation Coefficient Filtering

In order to extract weakly contrasting aircraft shapes from diffuse intensity regions, we use a form of normalised two-dimensional filtering. Let T denote a 3×3 square matrix of zeros with 1 in its centre element (i.e., the single pixel shape). We then compute a filtered grey-scale image F_k with elements given by the normalised cross-correlation coefficient in the sense that

$$F_k^{u,v} = \frac{\sum_{i,j} \left(\bar{I}_k^{i,j} - \bar{\mathcal{I}}_k(u,v) \right) \left(T^{i-u,j-v} - \bar{T} \right)}{\sqrt{\sum_{i,j} (\bar{I}_k^{i,j} - \bar{\mathcal{I}}_k(u,v))^2 \sum_{i,j} (T^{i-u,j-v} - \bar{T})^2}}$$

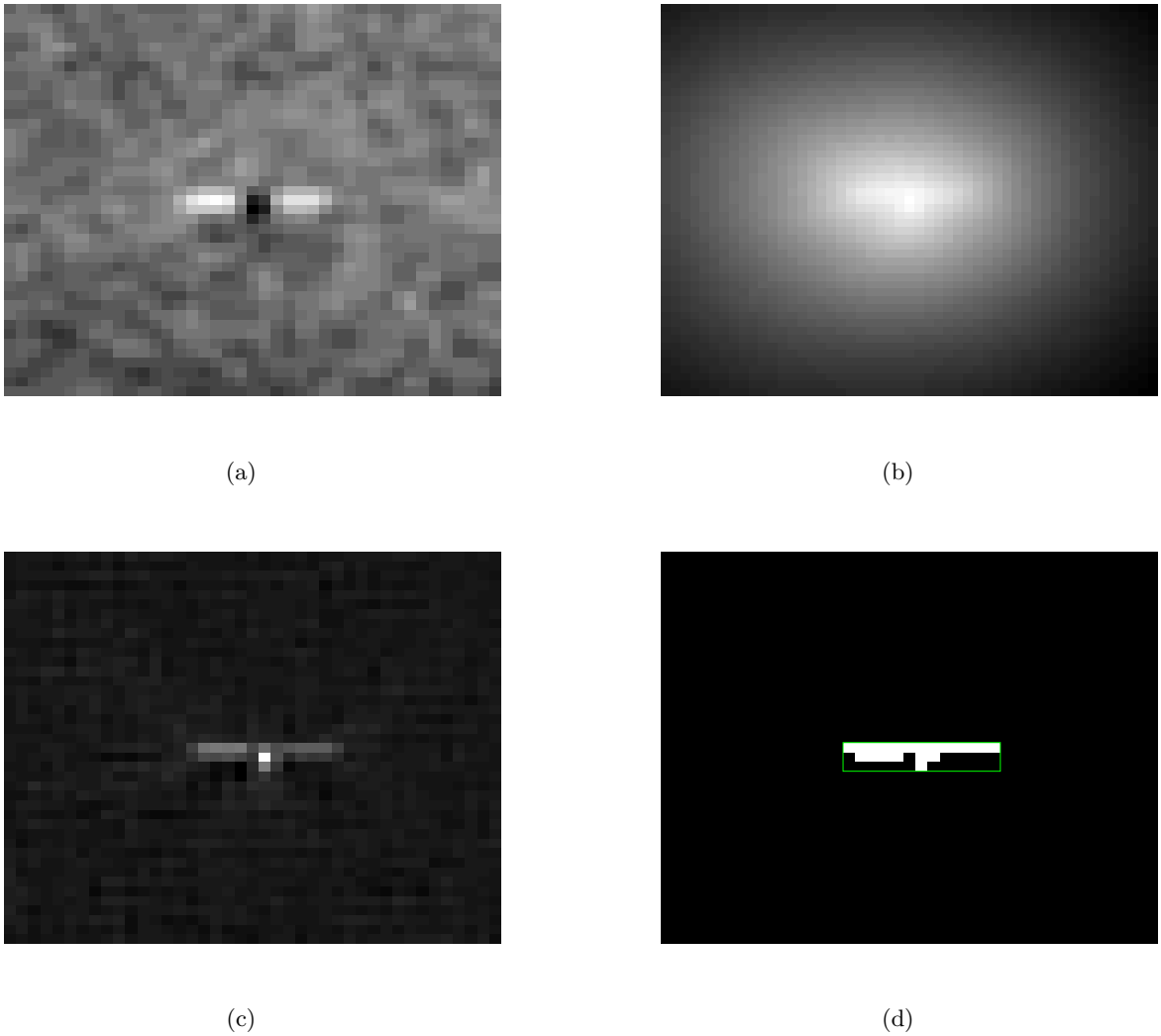
for all $1 \leq u, v \leq 40$ where $\bar{T}$ is the mean of T (here, $1/9$) and $\bar{\mathcal{I}}_k(u,v)$ is the local mean of the pixels of $\left\{ \bar{I}_k^{i,j} : u-1 \leq i \leq u+1, v-1 \leq j \leq v+1 \right\}$.

Shape Analysis

Prior to performing shape analysis, we convert the grey-scale image F_k into a binary version F_k^b using Otsu's method [Otsu, 1979]. Using the contours of the binary image F_k^b we then identify candidate aircraft shapes. The shape created by joining several candidate regions forms an aircraft shape estimate $\hat{M}_k$ (in the form of a binary mask). We perform final checks to reject shape estimates that are larger than aircraft of interest (e.g. 30 pixels in either width or height since aircraft this large normally move too fast for a HMM filter to track), and to reject shape estimates that have large bounding boxes, but sparse binary contents (since these type of "sparse" shape estimates usually correspond to false alarms).

Shape-Adaptive HMM filter

We can then propose a shape-adaptive filter using (1) where $\hat{M}_k$ is the shape estimated from the previous shape analysis.



(a) (b)

(c) (d)

Figure 2: Stages of Proposed Shape Estimator: (a) raw camera image before single pixel HMM filtering, (b) power law image I_k, (c) normalised cross-correlation coefficient filtered image F_k, and (d) estimated aircraft mask $\hat{M}_k$ with bounding box. All images are cropped to 40×40 pixels.

3.2 Aircraft Range Estimation

Recalling the perspective camera model, we estimate the aircraft's range $\hat{r}_k$ by evaluating

$$\hat{r}_k = f \frac{d_k}{p\hat{D}_k} \tag{2}$$

where f is the focal length of the camera, d_k is the physical dimension of the aircraft observed (e.g. wingspan for head-on and tail chase aircraft), $\hat{D}_k$ is the maximum dimension of the aircraft in the image (in pixels determined from $\hat{M}_k$), and p is the physical size of a pixel on the image sensor.

4 Implementation: Shape Estimation, Adaption and Range Declaration

In this section, we describe the implementation of our proposed aircraft shape and range estimators, and our shape-adaptive multi-pixel HMM filter, as described in Figure 3. This implementation is suitable for aircraft approaching head-on (but extension to other geometries is feasible). Our implementation was performed in C++ on a graphics processing unit (GPU) using OpenCV and Nvidia CUDA.

Step 1: Single Pixel HMM Filter Bank: As a method to handle uncertainty in the (image-plane) head-

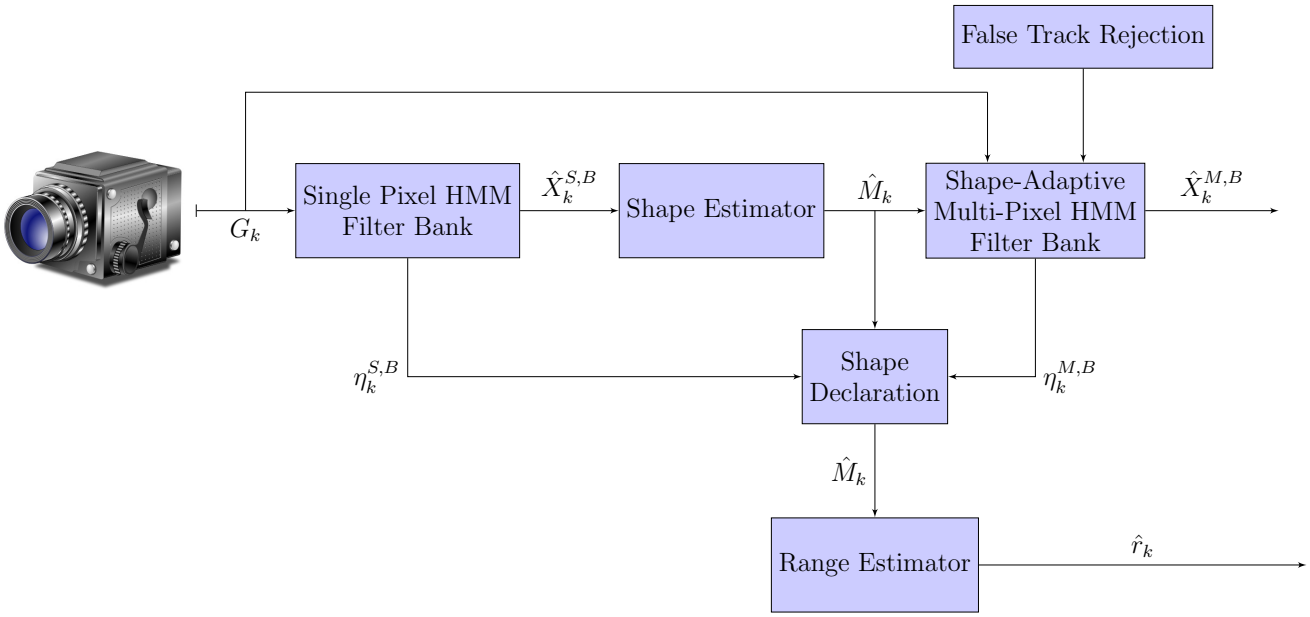


Figure 3: Block Diagram of Implemented Shape and Range Estimation Process with Mulit-Pixel Filtering

ing of aircraft [Lai *et al.*, 2013b], we implemented a bank of 4 single pixel HMM filters, where each HMM filter had a different transition probability matrix (we used the filter bank described in [Lai *et al.*, 2013b]). This bank of single pixel filters has previously been shown to be sufficiently representative of the range of potential aircraft dynamics. For each frame, we calculate $B_S(y_k)$ and estimate by running the bank of single pixel filters (i.e. 4 versions of (1) with $B(y_k) = B_S(y_k)$ indexed by $i = 1, \ldots, 4$) and then select the estimate from the filter with the highest detection statistic $\eta_{k,i}$. We denote this single pixel filter bank estimate as $\hat{X}_k^{S,B}$ and our single pixel filter bank detection statistic as $\eta_k^{S,B} = \max_i \eta_{k,i}$.

Step 2: Shape Estimator: Our proposed shape estimator described in Section 3 was implemented on $\hat{X}_k^{S,B}$ using OpenCV functions to produce an estimated aircraft shape $\hat{M}_k$. If our shape estimator is unable to provide a useful shape estimate (i.e. one of the final checks described in Section 3 failed), then we set $\hat{M}_k = T$ as the aircraft shape (that is, the single pixel shape as given in Section 3.1).

Step 3: Shape-Adaptive Multi-Pixel HMM Filter Bank: We then calculate a refined estimate of aircraft location from a bank of 4 adaptive-shape multi-pixel HMM filters based on the estimated aircraft shape $\hat{M}_k$, where each branch of the adaptive-shape multi-pixel HMM filter bank shares its transition scheme with the corresponding branch of the single pixel HMM filter bank. Here we use $\hat{M}_k$ to calculate the required $B_M(y_k)$ from $B_S(y_k)$ using two-dimensional convolution as described in Section 2.

Again we select the filter estimate with the highest detection statistic. We denote this shape-adaptive multi-pixel filter estimate as $\hat{X}_k^{M,B}$ and our shape-adaptive detection statistics as $\eta_k^{M,B} = \max_i \eta_{k,i}$.

Step 4: Shape Declaration: Since $\eta_k^{M,B}$ will be much larger than $\eta_k^{S,B}$ when the estimated aircraft mask $\hat{M}_k$ is well matched to the aircraft shape, we declare a range estimate only when $\eta_k^{M,B} - \eta_k^{S,B} > \Delta\eta$ (we have experimentally found $\Delta\eta = 40$ to give reasonable results). We also only declare a range estimate if the mask is rectangular (since in this case the wings of an aircraft are likely to be visible).

False Track Rejection Check: To avoid the shape-adaptive multi-pixel HMM filters prolonging a track of false aircraft target (e.g. cloud edges), we reinitialise the shape-adaptive multi-pixel HMM filter bank with the estimate from the single pixel filter bank when the pixel location of the maximum intensity in the power law image I_k changes by more than 30 pixels (this criteria was experimentally determined).

5 Data Collection and Results

In this section, we demonstrate the performance of our proposed range estimator and shape-adaptive filters on image sequences of an approaching Cessna aircraft (note that these image sequences were previously analysed for

a different purpose in the field-of-view study provided in [Lai *et al.*, 2012]). As described in [Lai *et al.*, 2012], we captured five sequences of a Cessna 172 flying towards a ground-based camera (to simulate a head-on collision scenario). We have previously conducted both airborne flight test and ground-based camera tests of morphological-HMM based detection approaches, and established that the aircraft detection performance in ground-based testing provides good indication of airborne behaviour (the primary difference relates to the requirement for compensation of the sensor or platform ego-motion that occurs during flight, which tends to mildly reduce detection range), see [Lai *et al.*, 2011; 2013b] for details of flight tests. Two of our sequences (labeled G3 and G4) contained strong cloud features, and the other three had weak cloud features (G1, G2 and G5). For the purpose of range truth data, we also captured synchronised navigation solutions (GPS/INS position) of the aircraft and camera.

All of our image sequences were captured in 8-bit greyscale at a resolution of 1024×768 pixels with a Basler Vision Technology Scout series (scA1300-32f) camera. The camera's physical pixels are square with side length $p = 3.75\mu$m. Sequences G1, G2 and G3 were captured with a $f = 16$mm focal length lens, G4 was captured with a $f = 5$mm focal length lens, and G5 was captured with a $f = 50$mm focal length lens. For the purpose of range estimation, we assumed prior knowledge that the aircraft is approaching head-on so that the physical dimension used to calculate our range is the Cessna's wingspan $d_k = 11$m (admittedly, this is a strong assumption).

5.1 Range Estimation and Shape-Adaptive Multi-Pixel HMM Filter Illustrative Example

To illustrate our range estimator and shape-adaptive filtering, we processed 2058 frames of the G3 sequence with our filter bank implementation. Figure 4 shows a plot of the aircraft range estimate we inferred from our aircraft shape estimator. Figures 5, 6 and 7 illustrate the shape-adaptive estimate, detection statistic and sample standard deviation in position estimate respectively.

From Figure 4, we see that the range estimates are often within 1-2 pixel error of the GPS/INS truth. Initially (prior to $k = 3042$), the aircraft range estimate is not declared whilst the single pixel filter bank is tracking a cloud and the shape estimate exceeds the maximum dimensions of an admissible aircraft. At other time instants, a range estimate is not declared because the detection statistics of the shape-adaptive filter bank are not sufficiently different from the detection statistics of the single pixel filter bank (suggesting that the shape estimate is poor). A range estimate is unavailable at the end of the sequence due to failure of the shape checks.

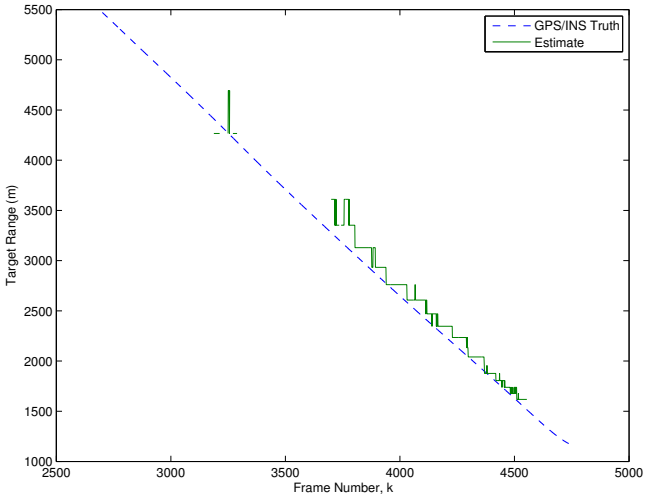


Figure 4: Illustrative Example G3: Range estimate from ground-based camera with GPS/INS Truth. The apparent quantisation in range estimates corresponds to aircraft size being quantised to pixels.

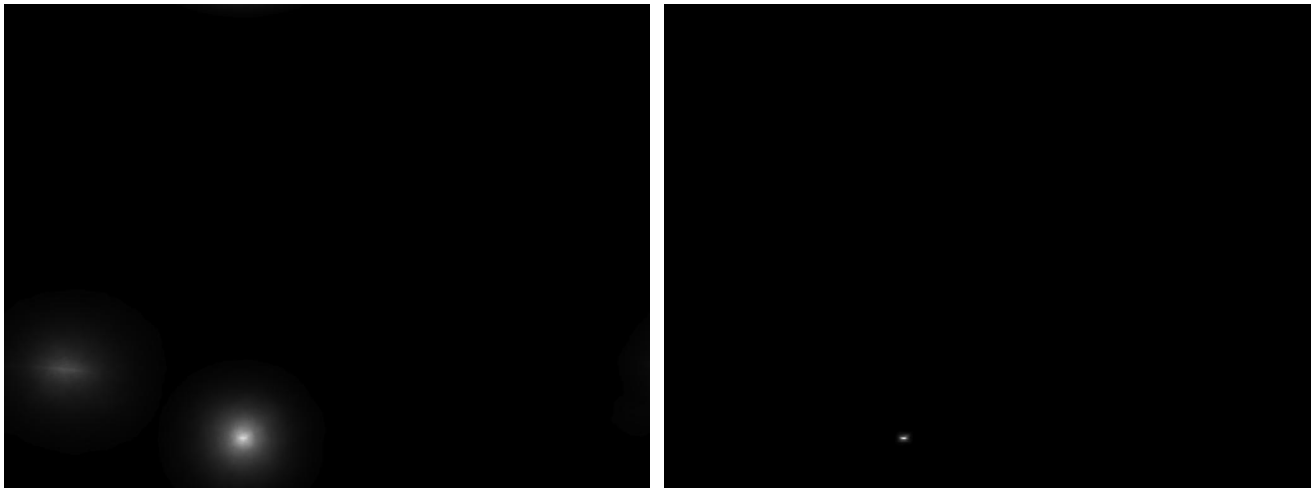
Figure 5 highlights that other aircraft-like image features (e.g. cloud), can be ignored by the shape-adaptive HMM filters proposed here. We can quantify the reduction in uncertainty about aircraft location in a given frame by calculating the sample standard deviation of the aircraft location conditional mean estimate. We see from Figure 7 that the standard deviation is reduced by several pixels with the use of a shape-adaptive filter. Importantly, we claim that concentration of filter probabilities enables the shape-adaptive filters to avoid aircraft tracks being lost when aircraft pass close to other (false) aircraft-like features. The detection statistics in Figure 6 also highlight that the shape-adaptive filters often have significantly higher detection statistics compared to single pixel filters (which highlights that the shape-adaptive filters are more confident of their aircraft tracks).

5.2 Range Estimation and Shape-Adaptive Multi-Pixel HMM Filter Performance Study

We now further investigate the performance of our proposed methods using the four other image sequences: G1, G2, G4 and G5.

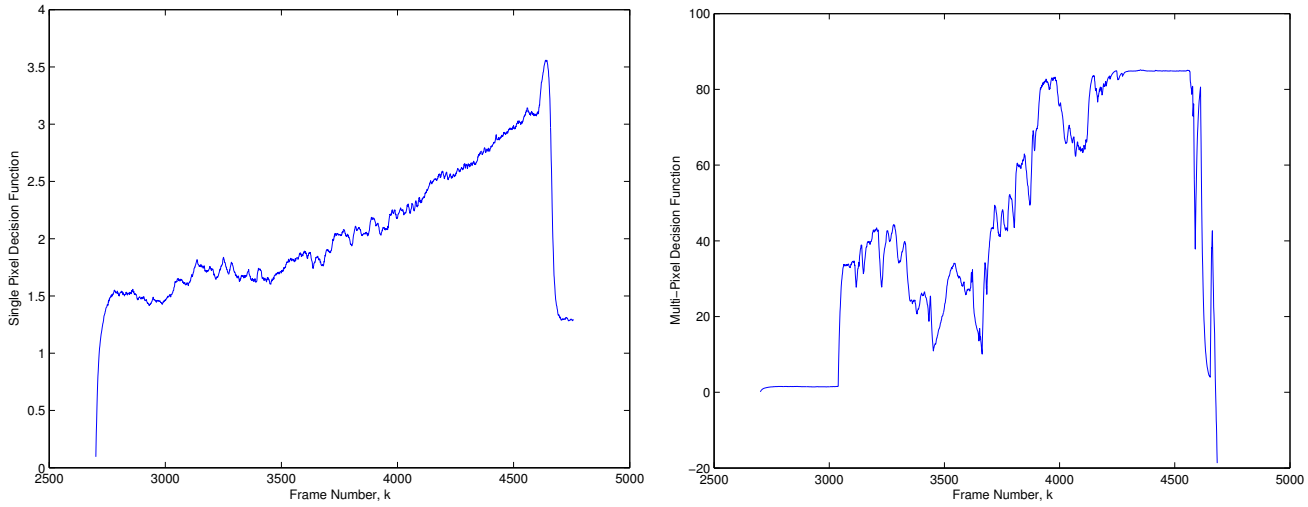
Range Estimation

From the set of experiments we present two of the most representative range estimates in Figures 8 and 9, for sequences G1 and G5 respectively. In sequence G1 the range estimator performs well, with range errors (when the shape estimates are suitable) being within a 1-2 pixel tolerance of the GPS/INS truth. The range estimate in sequence G5 seems also to perform within 1-4 pixels of



(a) (b)

Figure 5: Illustrative Example G3: Filter Estimate $\hat{X}_{3130}$ from: (a) Single Pixel Filter Bank and (b) Proposed Shape-Adaptive Filter Bank. The aircraft is the bright feature in both images. The bright feature on the left side of (a) is a cloud. As expected the aircraft estimate is more distinct in the shape-adaptive filter.



(a) (b)

Figure 6: Illustrative Example G3: Detection Statistic η_k from: (a) Single Pixel Filter Bank and (b) Proposed Shape-Adaptive Filter Bank. As expected the aircraft detection metric is more distinct in the shape-adaptive filter.

the GPS/INS truth. We note that the large focal length used to capture G5 enables us to estimate the range of the more distant aircraft.

Although not plotted here, we can report that the performance of our range estimator was poor in the two other sequences (G2 and G4). In the G2 sequence, only part of the aircraft's wingspan was visible in the raw camera images due to lighting conditions (and so the estimated shape was only the visible part of the wingspan). Since our conversion from shape to range (2) is based on the assumption that d_k is the full wingspan, the range estimates in G2 were much larger than the GPS/INS truth (but the poorly performing shape estimate was still able to suggest that the aircraft was approaching). In the G4 sequence, the single pixel filter bank tracked cloud features instead of the aircraft. Our shape estimator was able to recognise that these cloud features were too large to be an aircraft of interest, and range estimates were only produced when the single pixel filter bank tracked the (much smaller) aircraft. The few range estimates

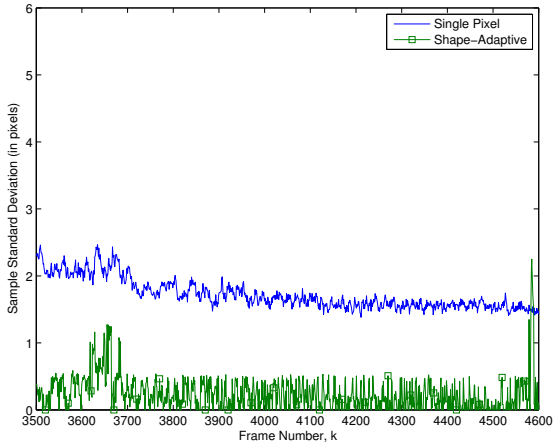


Figure 7: Illustrative Example G3: Standard deviation (in pixels) of the aircraft's x position component of the conditional mean estimate in each frame (higher standard deviation means that the filter was more uncertain about the pixel location of the aircraft). The (not shown) y component standard deviation performance was similar.

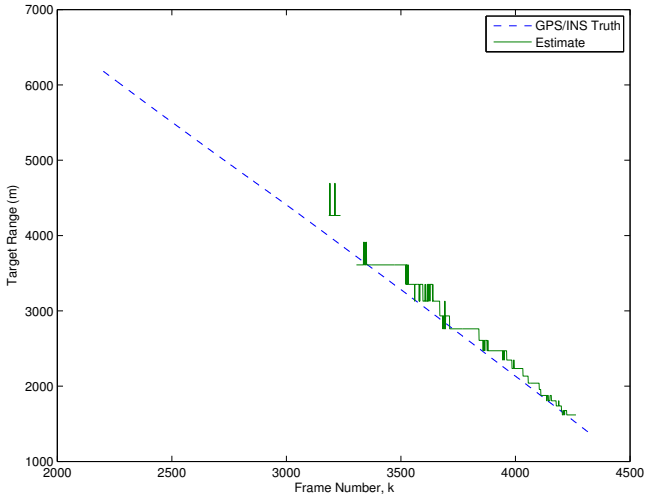


Figure 8: Image Sequence G1: Range estimate from ground-based camera with GPS/INS Truth. The apparent quantisation in range estimates corresponds to aircraft size being quantised to pixels.

(less than 50) produced for G4 appeared reasonable.

Shape-Adaptive Multi-Pixel HMM Filter Performance

To characterise the difference in filter performance we calculated a mean "distinctness" signal-to-noise ratio (DSNR) value for each sequence, see [Lai *et al.*, 2011]. The DSNR value provides a description of the filter confidence in the location of an aircraft and is defined as:

$$DSNR \triangleq 20 \log \frac{P^T}{P^N}$$

where P^T and P^N are the peak aircraft and non-aircraft probabilities respectively (at the output of the filter banks). The DSNR values are summarised in Table 1. These high DSNR results suggest shape-adaptive filtering results in very distinct or clear aircraft estimates. We note that in all of the sequences, the DSNR of the shape-adaptive filter was beyond numerical precision for some frames, and so the mean DSNR quoted is a minimum value.

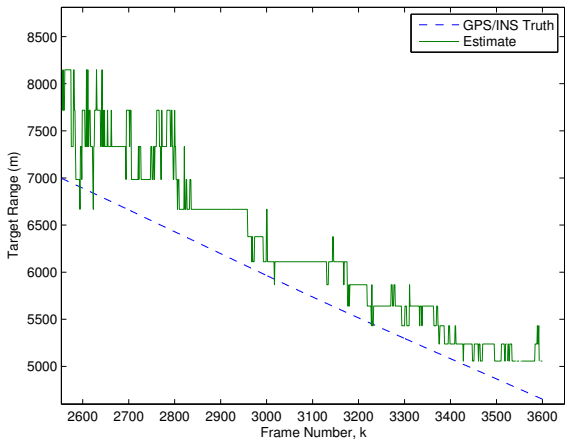


Figure 9: Image Sequence G5: Range estimate from ground-based camera with GPS/INS Truth. The apparent quantisation in range estimates corresponds to aircraft size being quantised to pixels.

Table 1: Average DSNR (in dB) of Single Pixel and Shape-Adaptive HMM Filter Banks

Filter	**G1**	**G2**	**G3**	**G4**	**G5**
Pixel	87.9	116	111	96.8	60.5
Shape	>240	>389	>438	>188	>307

6 Conclusions and Future Work

In this paper, we proposed and investigated new techniques for aircraft shape and range estimation from image sequences. We also proposed a shape-adaptive HMM filtering approach. Our preliminary investigation has demonstrated the initial feasibility of extracting both shape and range information from the output of HMM aircraft filters, but also illustrates some of the challenges in this problem.

Acknowledgements

This research was supported under the Australian Research Council's Linkage Projects Scheme (project number LP100100302).

References

[Arnold *et al.*, 1993] J. Arnold, S.W. Shaw, and H. Pasternack. Efficient target tracking using dynamic programming. *Aerospace and Electronic Systems, IEEE Transactions on*, 29(1):44 –56, 1993.

[Barniv, 1985] Y. Barniv. Dynamic programming solution for detecting dim moving targets. *Aerospace and Electronic Systems, IEEE Transactions on*, AES-21(1):144 –156, January 1985.

[Carlemalm *et al.*, 2000] C. Carlemalm, H. V. Poor, and A. Logothetis. Suppression of multiple narrowband interferers in spread-spectrum communication system. *IEEE J. Sel. areas Commun.*, 18(1365-1374), August 2000.

[Cho *et al.*, 2013] S. Cho, S. Huh, D. Hyunchul Shim, and H. Sik Choi. Vision-based detection and tracking of airborne obstacles in a cluttered environment. *Journal of Intelligent & Robotic Systems*, 69(1-4):475–488, January 2013.

[Chung *et al.*, 1991] S. H. Chung, V. Krishnamurthy, and J. B. Moore. Adaptive precessing techniques based on hidden markov models characterizing very small channel currents buried in noise and deterministic interferences. *Philos. Trans. R. Soc. Lond. B.*, 334:357–384, December 1991.

[Dey *et al.*, 2011] D. Dey, S Geyer, S. Singh, and M. Digionia. A cascaded method to detect aircraft in video imagery. *The International Journal of Robotics Research*, 30(12):1527–1540, October 2011.

[Finn and Franklin, 2001] Anthony Finn and Stephen Franklin. Acoustic sense & avoid for uav's. In *Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2011 Seventh International Conference on*, pages 586–589, 2001.

[Gandhi *et al.*, 2003] T. Gandhi, Mau-Tsuen Yang, R. Kasturi, O. Camps, L. Coraor, and J. McCandless. Detection of obstacles in the flight path of an aircraft. *Aerospace and Electronic Systems, IEEE Transactions on*, 39(1):176 – 191, January 2003.

[Lai and Ford, 2010] J. Lai and J.J. Ford. Relative Entropy Rate Based Multiple Hidden Markov Model Approximation. *Signal Processing, IEEE Transactions on*, 58(1):165 –174, January 2010.

[Lai *et al.*, 2008] J. Lai, J.J. Ford, P. O'Shea, and R. Walker. Hidden Markov Model Filter Banks for Dim Target Detection from Image Sequences. In *Computing: Techniques and Applications, 2008. DICTA '08.Digital Image*, pages 312 –319, dec 2008.

[Lai *et al.*, 2011] John Lai, Luis Mejias, and Jason J. Ford. Airborne vision-based collision-detection system. *Journal of Field Robotics*, 28(2):137–157, 2011.

[Lai *et al.*, 2012] J. Lai, J. J. Ford, Luis Mejias, Alexander Lloyd Wainwright, Peter J. O'Shea, and R. Walker. Field-of-view, detection range, and false alarm trade-offs in vision-based aircraft detection. In *Proc. International Congress of the Aeronautical Sciences,*, September 2012.

[Lai *et al.*, 2013a] J. Lai, J. Ford, P. O'Shea, and L. Mejias. Vision Based Pseudo-Bearing Rate for Airborne Collision Avoidance. *IEEE Transactions on Aerospace and Electronic Systems.*, 49(4):2129–2145, 2013.

[Lai *et al.*, 2013b] John Lai, Jason J. Ford, Luis Mejias, and Peter O'Shea. Characterization of Sky-region Morphological-temporal Airborne Collision Detection. *Journal of Field Robotics*, 30(2):171–193, 2013.

[Molloy and Ford, 2012] Timothy Molloy and Jason J. Ford. HMM triangle relative entropy concepts in sequential change detection applied to vision-based dim target manoeuvre detection. In *Proc. 15th International Conference on Information Fusion*, Singapore, July 2012.

[Molloy and Ford, 2013] Timothy L. Molloy and J.J. Ford. HMM relative entropy rate concepts for vision-based aircraft manoeuvre detection. In *Australian Control Conference*, November 2013.

[Otsu, 1979] N Otsu. A threshold selection method from gray-level histograms. *Systems, Man and Cybernetics, IEEE Transactions on*, 9(1):62–66, Jan 1979.

[Shakernia *et al.*, 2005] O. Shakernia, W. Z. Chen, and V. M. Raska. Passive ranging for UAV sense and avoid applications. In *Proc. AIAA Infotech@Aerospace Conf.*, pages 1–10, Arlington, Virginia, 2005.

[Tonissen and Evans, 1996] S.M. Tonissen and R.J. Evans. Peformance of dynamic programming techniques for track-before-detect. *Aerospace and Electronic Systems, IEEE Transactions on*, 32(4):1440 – 1451, October 1996.