A Method for the Visual Estimation and Control of 3-DOF Attitude for UAVs

Richard J. D. Moore, Saul Thurrowgood, Dean Soccol, Daniel Bland & Mandyam V. Srinivasan
The University of Queensland, Australia

Abstract

This study describes a novel, vision-based method for automatically obtaining the 3-DOF attitude of an aircraft. A very wide-angle view of the environment is captured and the input images are classified into fuzzy sky and ground regions using the spectral and intensity properties of the pixels. A novel approach to obtaining the 2-DOF attitude of the aircraft from the classified images is described, which is used to generate a stabilised panoramic image of the horizon and sky. This 'visual compass' is then used to determine the heading direction of the aircraft. We present results from recent flight testing that demonstrates the abilty of this vision-based approach to outperform an inexpensive inertial unit and provide real-time guidance for a UAV.

1 Introduction

For an unmanned aerial vehicle (UAV), the ability to estimate 3-DOF attitude (roll and pitch angles, and heading direction) accurately is crucial for applications such as mapping, landmark localisation, or augmented reality, where even small attitude errors can lead to failures in feature matching or misalignments during reprojection. Traditionally, UAV attitude is sensed by integrating the outputs from rate gyroscopes, but this method is susceptible to noise-induced drift [Rohac, 2005]. This shortcoming can be mediated by augmenting the system with a direct measure of absolute orientation. For example, 3-axis accelerometers or magnetometers may be used to sense the direction of gravity or the local magnetic field respectively. However, accelerometers also measure the linear and centripetal accelerations of the aircraft, which confound the process of estimating the gravity vector when the aircraft is manoeuvring [Rohac, 2005], and magnetometers will not sense rotations about an axis that is parallel to the direction of the local magnetic field [Merhav, 1996]. Furthermore, the local magnetic field is subject to disturbances due to environmental factors. The limitations of these approaches mean that they are not ideal for applications where the attitude of the aircraft must be known precisely.

In [Moore *et al.*, 2011], we describe a vision-based method for sensing the 2-DOF attitude of an aircraft. The approach involves capturing a very wide-angle view of the environment, including the horizon. An adaptive classifier is used to segment the scene into sky and ground regions, using the spectral and intensity properties of the pixels. The position and orientation of the horizon are then obtained and used to infer the roll and pitch angle of the aircraft. Here, we present a method that builds upon this approach to estimate the 3-DOF attitude of an aircraft. Using the estimated 2-DOF attitude, a stabilised panoramic image of the horizon profile and sky is generated. The stabilised horizon image is independent of the roll and pitch of the aircraft, but the image undergoes a horizontal shift as the heading direction of the aircraft changes (see Figure 1). This so called 'visual compass' can therefore be used to determine the heading direction of the aircraft, relative to any prior point in time.

Methods for obtaining the 3-DOF attitude of an aircraft from a visual input have been described before [Mondragón *et al.*, 2010]. However, the novelty of the approach described here is that it is able to adapt continuously to changing environmental conditions and therefore does not require training off-line prior to use. Additionally, we show that this method is both robust and

This work was supported partly by US Army Research Office MURI ARMY-W911NF041076, Technical Monitor Dr Tom Doligalksi, US ONR Award N00014-04-1-0334, ARC Centre of Excellence Grant CE0561903, and a Queensland Smart State Premier's Fellowship.

Authors are associated with the Queensland Brain Institute & School of Information Technology and Electrical Engineering, University of Queensland, St Lucia, QLD 4072, Australia and ARC Centre of Excellence in Vision Science, Australia. Correspondence via r.moore1@uq.edu.au

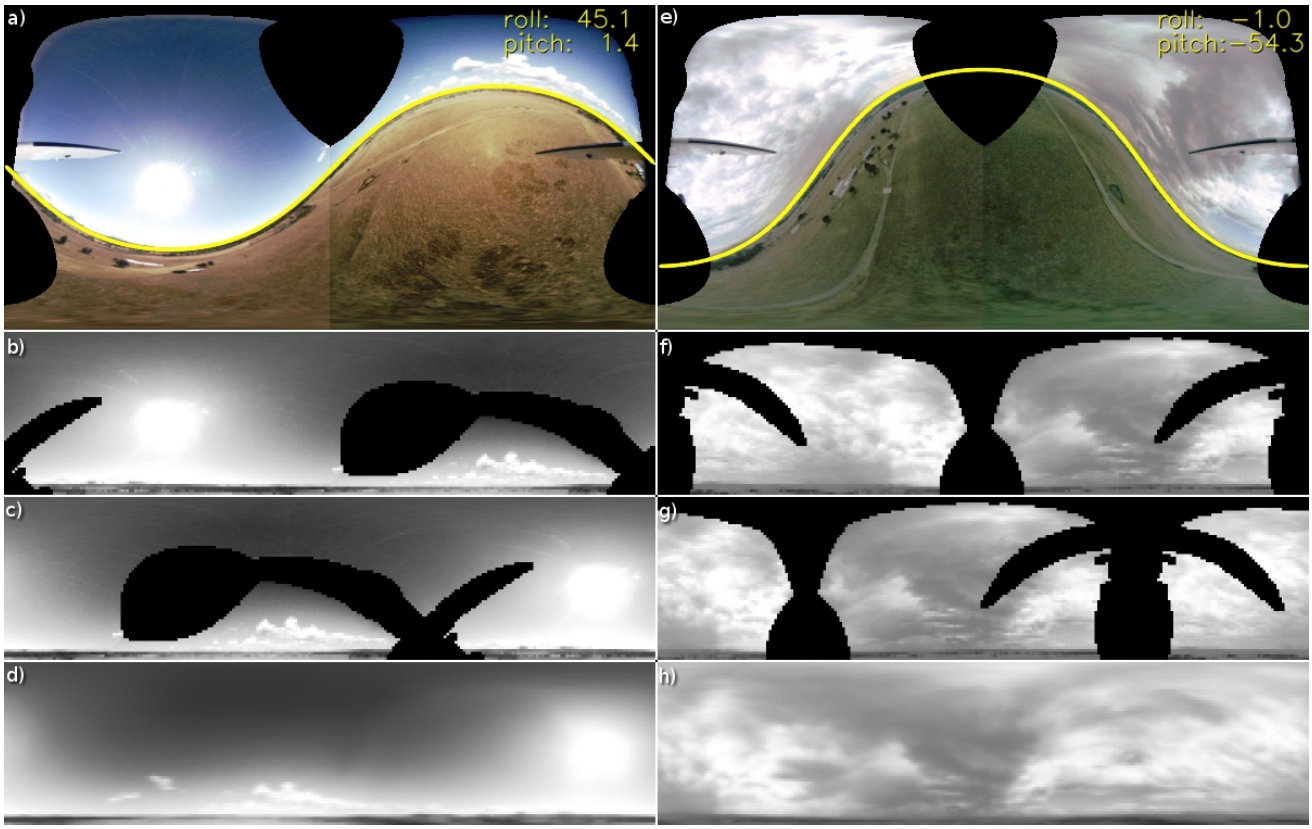

Figure 1: An illustration of how the visual compass is generated and used to estimate heading direction. Panels (a) and (e) show raw stitched images captured by the vision system during two different flight tests, overlayed with the estimated position and orientation of the horizon (yellow). Panels (b) and (f) show the stabilised horizon images, extracted from (a) and (e) respectively. The accumulated reference horizon images are shown in panels (d) and (h). At each frame the stabilised horizon image is shifted and matched against the reference image to obtain the corresponding difference in heading direction between the two. The best matching shifted horizon images for these two frames are shown in panels (c) and (g). The FOV of the raw image is $360° \times 180°$, and the FOV of the remapped horizon images is $360° \times 90°$. The black regions in panels (a) and (e) represent regions of the view sphere not imaged by the vision system, and additionally regions of the view sphere that correspond to the aircraft in the other panels.

relatively computationally efficient to implement, when compared with other state-of-the-art approaches. We also describe and implement a simple closed-loop control system that allows this method to provide guidance to a small-scale UAV performing a real-world task. Finally, we present results from a recent closed-loop flight test that demonstrate the ability of this vision-based approach to outperform an inexpensive inertial system.

2 3-DOF Visual Attitude

In the proposed approach, the full 3-DOF attitude of the aircraft is estimated in two steps. Firstly, the distribution of sky and ground regions in the input image is used to compute the position and orientation of the horizon, or identically, the roll and pitch angle of the aircraft. Using this information, we then produce a stabilised 'visual compass', from which the heading direction of the aircraft is obtained. We introduce and describe our procedure for estimating the 2-DOF attitude of an aircraft as well as analysing the performance of that system in [Moore *et al.*, 2011]. We do not repeat the results of that analysis here, but we summarise the operation of that system in section 2.1 for completeness. Then, in section 2.2, we describe our method for estimating the heading direction of the aircraft, which forms the main contribution of this article.

2.1 Visually Estimating 2-DOF Attitude

Automatic horizon detection schemes typically attempt to find an n-dimensional contour in colour, intensity, and/or texture space that separates the input scene into two distinct regions – one corresponding to the ground, and the other to the sky. Therefore, the separation of the input pixels into two classes is arguably the most cru-

cial step in automatic visual horizon estimation and this problem has been tackled previously in various ways.

[Cornall and Egan, 2004; Cornall *et al.*, 2006] use a static transformation of the RGB colour space to enhance the contrast between the sky and ground regions before applying a threshold to distinguish the two. [Mondragón *et al.*, 2010; Thurrowgood *et al.*, 2009; 2010] also apply a static transform to the RGB or YUV colour space but additionally define adaptive thresholding techniques that allow some limited variability between the appearance of the sky and ground and the static representations. These approaches can be computationally efficient. However, because the representations of sky and ground are largely predetermined and fixed, these approaches can fail when the appearance of the sky or ground in the input image does not closely match the global average representation of sky or ground for all scenes. [Dusha *et al.*, 2007] do not explicitly label sky and ground regions but instead directly search for the horizon contour in RGB colour space by employing the Hough transform on an edge image, while [Demonceaux *et al.*, 2006; Ettinger *et al.*, 2002a; 2002b; Todorovic and Nechyba, 2004; Todorovic *et al.*, 2003; McGee *et al.*, 2005] all describe approaches that actively search for the parameterised horizon contour that best separates the input image into two classes. While these approaches are robust to dramatic variations in the appearance of the sky and/or ground, they can be computationally intensive to execute and they must also maintain a representation of the sky and ground classes in order to distinguish the two.

In [Moore *et al.*, 2011] we propose an *adaptive* approach to the classification of sky and ground regions in the input image. In this approach, the classifier maintains a record of the spectral and intensity properties of the sky and ground regions, which is continuously updated online. This does not require a fixed, predetermined representation of the sky and ground classes and hence does not require the classifier to be trained offline prior to use. Additionally, the implementation is computationally efficient and hence fast to execute.

The proposed horizon detection scheme involves four stages (Figure 2):

1. *Classification* – Spectral and intensity properties (i.e. the YUV values) are used by the classifier to assign a weight to each pixel in the colour input image. The weights represent the likelihood that the corresponding input pixel belongs to either the *sky* or *ground* class, or a combination of the two. The 2D array of pixel weights forms the classified image.
2. *Matching* – The classified image is reduced to two 1D arrays (we refer to these arrays as a kernel), which hold the row and column averages of the classified image. The classified image kernel is then matched against a database of reference kernels, which represent every possible combination of roll and pitch, to find the current 2-DOF attitude.
3. *Masking* – A sky/ground mask is generated using the current estimate of roll and pitch. The mask is used to compute the desired spectral and intensity weights within the classifier.
4. *Training* – The weights in the classifier are updated using the sky/ground mask and an online reinforcement strategy.

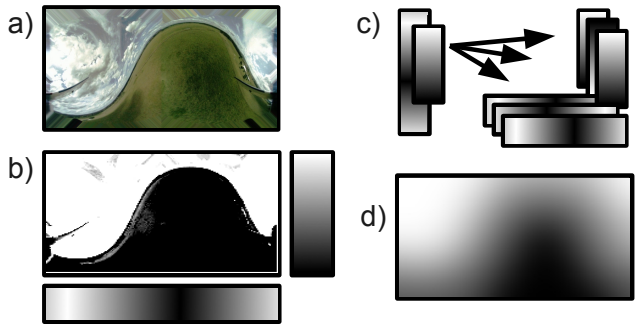

Figure 2: An illustration of the 2-DOF attitude estimation process. First, the raw input image (a) is processed to give the classified image (b). Then, the classified image is reduced to two 1D arrays (a kernel) which contain the row and column averages. The input arrays are matched to a pre-computed database of attitude kernels (c) to retrieve the current roll and pitch angles. Finally, a sky/ground mask (d) is generated, using the estimated 2-DOF attitude, and used to update the weights in the classifier.

Each of these steps is performed in turn on each input image, so that over time the classifier builds up a representation of the ground and sky classes. Because the classifier is trained continuously, it is able to adapt to the changing appearance of the ground and sky. Each stage is described in greater detail in [Moore *et al.*, 2011], and a summary is given in the following sub-sections.

Classifying sky and ground

The role of the classification stage is to label the sky and ground regions in the input image. Our sky and ground classes are fuzzy subsets (first described by [Zadeh, 1965]) of the YUV colour space. Therefore, each pixel in the input image may be labeled as definitely sky or ground, or some uncertain combination of the two. To represent the range of possible classifications, each input pixel, i, is assigned a weight, $w_i(y, u, v)$, which is continuous and defined over the range $[-1, +1]$. The weight is a function of Y, U, and V, and is a measure of the

likelihood that a particular combination of spectral and intensity properties represents a sky region ($w = +1$), or a ground region ($w = -1$).

In practice, the classifier is simply a three dimensional structure in YUV space, where each element, $w_{y,u,v}$, within the structure stores a classification weight for a particular YUV combination. To classify an input image at run-time, the Y, U, and V values for each input pixel are used simply to address an element within the classifier and retrieve the appropriate weight.

Matching against stored attitude kernels

To obtain the current 2-DOF attitude of the aircraft, the classified input image is exhaustively matched against a database of ideal image classifications, which represent all possible combinations of roll and pitch. The database is pre-computed offline by generating the ideal classification images for an observer above an infinite ground plane, over the range of all possible attitudes.

To reduce the complexity of the matching process, and without loss of generality, each classified image is reduced to two 1D arrays (a kernel), which hold the row and column averages of the classified image. An error score for each candidate 2-DOF attitude is computed from the sum of absolute differences (SAD) between the classified input kernel and each of the stored reference kernels. In practice, the reference kernels are computed at some finite number of attitudes (in this study we use an angular resolution of 5°), and hence the true current attitude is estimated by applying an equiangular fit [Shimizu and Okutomi, 2003] to the minimum and neighbouring SAD scores.

Generating the sky/ground mask

The classifier weights, $w_{y,u,v}$, are updated online using a reinforcement strategy that is similar to the well-known delta rule [Rumelhart, 1985],

$$\Delta w_{y,u,v} = \alpha_w \cdot \delta_{y,u,v}, \tag{1}$$

where, $\Delta w_{y,u,v}$ is the change applied to a weight within the classifier, α_w is the global learning rate (we use $\alpha_w = 0.05$), and $\delta_{y,u,v}$ is the difference between the desired output and the current output for each element. However, here we wish only to maximise the difference between the classifier weights for the sky and ground regions. Hence, we may choose $\delta_{y,u,v}$ arbitrarily, such that it acts as a velocity term that drives the corresponding weight towards the desired class.

To identify the desired class (sky or ground) for each YUV combination, we first identify the desired class for each pixel, i, in the input image:

$$m_i = \hat{\mathbf{v}}_i \cdot \hat{\mathbf{n}}, \tag{2}$$

where m_i and $\hat{\mathbf{v}}_i$ are the mask value and unit view vector (obtained from the camera calibration) respectively for the i^{th} pixel in the input image, and $\hat{\mathbf{n}}$ is the unit vector direction of up[1] given in the camera frame, which is obtained from the current 2-DOF attitude estimate.

Training the classifier

Once we have identified the desired class for each pixel in the input image using (2), we can compute the desired class for each YUV element within the classifier according to

$$\delta_{y,u,v} = \frac{1}{k} \sum_{i \in K} m_i,$$

where K is the subset of pixels in the input image whose Y, U, and V values correspond to the appropriate YUV element within the classifier, and k is the cardinality of K.

Then, we may update the classifier weights,

$$w^{t+1}_{y,u,v} = w^{t}_{y,u,v} + \Delta w_{y,u,v},$$

where $w^{t+1}_{y,u,v}$ is the new classifier weight for the element YUV and $\Delta w_{y,u,v}$ is defined above in (1).

Initially, the weights within the classifier are $w_{y,u,v} = 0$, for all Y, U, and V. However, as training examples are presented to the classifier, it quickly learns to separate the sky and ground classes and ideally $|w_{y,u,v}| = 1$ for all Y, U, and V, as $t \to \infty$.

2.2 Visually Estimating Heading Direction

The 'visual compass' has been described previously for the purpose of extracting relative heading direction from a visual input [Mondragón *et al.*, 2010; Labrosse, 2006; Scaramuzza and Siegwart, 2008]. Typically, this approach operates on the principle that a rotation around the yaw axis of the vehicle corresponds to a simple left or right column shifting of a 360° panoramic image taken around the same axis.

[Scaramuzza and Siegwart, 2008; Labrosse, 2006] describe implementations of the visual compass for ground vehicles, where the yaw axis of the vehicle is in general well aligned with the inertial frame up vector, hence they neglect the rolling and pitching motion of the vehicle. [Mondragón *et al.*, 2010] extend the approach to include aerial vehicles, but also neglect roll and pitch. Their approach is limited, therefore, to estimating the change in visual heading direction between consecutive frames, between which the roll and pitch angles of the vehicle are similar. Absolute heading direction is then obtained through integrating the estimated deltas.

In this work, we describe a method for obtaining the heading direction of a UAV using a visual compass that accounts for the roll and pitch angle of the aircraft. Therefore, this method should be more accurate than

[1] We define up as a vector that is perpendicular to the surface of the Earth.

other previously described approaches, and additionally the visual compass should be valid over much longer periods of time. This means that absolute heading can be computed directly (relative to any prior heading), rather than by integrating the intervening changes in heading direction.

The proposed approach involves two stages. First, the estimated orientation and position of the horizon is used to extract a stabilised horizon image. Then, this 'visual compass' is used to determine the relative heading direction by left or right column shifting the stabilised image (or equivalently rotating the panoramic image around the inertial frame *up* vector) to match a stored reference image. These two steps are illustrated in Figure 1 and described in more detail in the following sub-sections.

Generating the visual compass

The visual compass is a stabilised panoramic image of the horizon and sky. The panorama extends 360° around the inertial frame *up* vector in the horizontal image axis and from 5° below the horizon to 85° above the horizon in the vertical image axis. We ignore image regions that are well below the horizon plane because they are likely to contain objects that are close by, and hence transient in the field of the view as the aircraft translates. Nearby objects that protrude through the horizon plane will affect the accuracy of the visual compass if they subtend a large enough angular size and move significantly through the aircraft's field of view as it translates (i.e. if the aircraft flies closely past a tall city building). However, for our visual compass we use a greyscale image with dimensions 80px × 20px. Therefore, given the small resolution and large field of view of the visual compass, the aircraft must be very close to a very large object before it significantly affects the visual heading estimate. Furthermore, because our reference horizon image is accumulated slowly over time (see *Determining heading direction*), nearby objects will not persist for long in the field of view of the aircraft as it translates, and hence their impact on the visual heading estimate will be mitigated. It is important to note that no significant increase in accuracy is observed when higher resolution horizon images are used, although the execution time of the algorithm is increased significantly.

To produce the visual compass, the estimated orientation and position of the horizon is used to generate a transform that remaps pixels from the raw input image to the stabilised horizon image. The transform is given by

$$\begin{bmatrix} s_x^i \\ s_y^i \\ s_z^i \end{bmatrix} = \begin{bmatrix} c(\phi) & s(\phi)\cdot c(\theta) & s(\phi)\cdot s(\theta) \\ -s(\phi) & c(\phi)\cdot c(\theta) & c(\phi)\cdot s(\theta) \\ 0 & -s(\theta) & c(\theta) \end{bmatrix} \begin{bmatrix} r_x^i \\ r_y^i \\ r_z^i \end{bmatrix},$$

where $s()$ and $c()$ represent the sine and cosine functions respectively, $\mathbf{s^i}$ and $\mathbf{r^i}$ are the view vectors in the source and remapped images respectively that correspond to the i^{th} pixel in the remapped image, and ϕ and θ are the estimated roll and pitch angles of the aircraft respectively. The relationship between the view vector and the pixel coordinates for a particular pixel is determined from the camera calibration parameters.

Determining heading direction

By removing the influence of the aircraft's rolling and pitching motions on the visual compass (i.e. by stabilising the horizon image), the only remaining degree of freedom is a 1D shift along the horizontal image axis that corresponds to the change in heading direction of the aircraft. Therefore, by defining a reference horizon image, the relative heading direction of the aircraft can be determined by shifting the instantaneous horizon image to match the reference image.

An error score for each candidate heading direction is computed from the sum of absolute differences (SAD) between the instantaneous horizon image and the stored reference image. In practice, a discrete number of candidate heading directions are tested (in this study we use an angular resolution of 4°) and the true heading direction is estimated by performing an equiangular fit [Shimizu and Okutomi, 2003] on the minimum and neighbouring SAD scores.

In this study, the reference horizon image is not static but is accumulated from each matched horizon image according to

$$d_i^t = \alpha_d \cdot c_i^t + (1 - \alpha_d) \cdot d_i^{t-1},$$

where d_i and c_i are the i^{th} pixels in the accumulating reference image and the instantaneous matched horizon image respectively, and α_d is the accumulation rate of the reference image (in this study we use $\alpha_d = 0.01$, which corresponds to $\tau_d \sim 4\text{s}\,@\,25\text{Hz}$). Regions in the matched horizon image that correspond to the body of the aircraft or areas not imaged by the vision system are not used to update the reference horizon image.

The estimated relative heading of the aircraft was not observed to drift significantly during any of the flight tests conducted for this study. However, because the reference horizon image is not static, it is possible that noise or biases in the matching process, or the motion of the sun, cloud patterns, or other distinctive features in the visual compass may cause the reference image to drift over long periods of time. Therefore, in order to counter this unwanted drift, and also to give physical meaning to the relative heading estimates, an angular offset between the estimated relative heading direction and magnetic north could be estimated continuously during flight. However, magnetometer measurements are typically much higher latency and lower accuracy than the approach described here. Hence, the visual heading

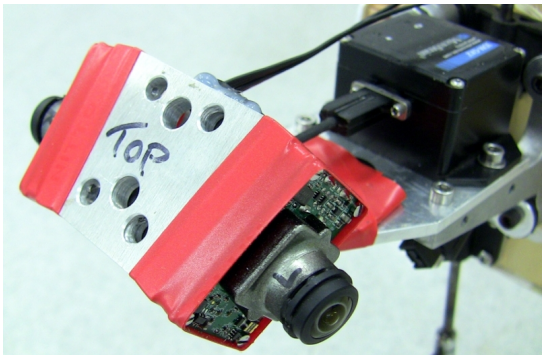

Figure 3: The dual-camera vision system mounted on the nose of the aircraft. The on-board IMU is visible in the background.

estimate should only be very loosely coupled with the magnetic estimate. We do not apply any such correction to the visual heading estimate for the results presented here.

3 Closed-loop Flight Testing

To validate the performance of the proposed approach in the real world, a closed-loop flight test was performed in which a trainer aircraft was commanded to perform a series of 90° turns, whilst maintaining altitude. To facilitate these requirements, a simple closed-loop flight control system was implemented using PID feedback control loops. The PID control loops were cascaded to give control of both altitude and heading direction, as well as roll and pitch angle.

During the flight test, images were captured at 25Hz by a dual-fisheye vision system that was mounted on-board the test aircraft (Figure 3). The cameras were positioned back-to-back so that the vision system had an almost spherical view of the environment (see Figure 1). A MicroStrain inertial measurement unit (IMU) was mounted rigidly to the vision system and its output was logged during the flight test to provide a comparitive attitude reference. The flight platform and vision system are discussed in more detail in [Thurrowgood *et al.*, 2010].

The vision system provided control feedback for the heading direction, and the roll and pitch angle of the aircraft, while the height control feedback was provided by a u-blox GPS unit that was mounted on-board the aircraft.

In order to quantify the accuracy of the proposed approach for automatically computing heading direction, a ground-truth measure of the aircraft's heading direction during the flight test was obtained using a novel procedure that involved manually tracking the location of the aircraft's shadow in the imagery captured from on-board the test aircraft. Using the cameras' calibration parameters, camera-frame view vectors to the apparent positions of the shadow were recorded at regular intervals throughout the flight test. For each frame, the visually estimated roll and pitch angles were used to transpose the view vector corresponding to the shadow into the (rotating) inertial frame (we use the NED coordinate system), and the ground-truth heading estimates were computed according to

$$\Psi^t_{\text{true}} = \arctan(\frac{g^t_y}{g^t_x}),$$

where Ψ^t_{true} is the ground-truth heading direction at time t, and g^t_x and g^t_y are the x and y components of the inertial frame view vector that corresponds to the position of the centre of gravity of the aircraft's shadow at time t.

It was shown in [Moore *et al.*, 2011], using the proposed approach, that the aircraft's 2-DOF attitude is able to be estimated with an average angular error of $\sim 1.49°$, and here we estimate the average error for the *manual* tracking of the aircraft's shadow to be $< 1.5°$. Therefore, we consider that the ground-truth measure described here accurately represents the true heading direction of the aircraft during the flight test. Frames in which the aircraft's shadow is not visible are omitted from analysis.

3.1 Flight Test Results

The flight test used for analysis is shown in the accompanying video[2]. The analysed section of data covers a complete flight, from take-off to touch-down, and lasts approximately 250s. The flight test data is shown in Figure 4. The performance of the visual horizon detection algorithm, used to estimate the 2-DOF attitude of the aircraft, is analysed in [Moore *et al.*, 2011] and we do not repeat those results here.

It can be seen from Figure 5 that the error between the visually estimated heading direction and the ground truth is consistently $< 10°$ for the entire flight. This indicates that there is very little or no drift of the accumulated reference horizon image, even after a flight time of 250s. The average angular error for the visually estimated heading direction was computed as 2.47°, using the ground truth measure described in section 3. For comparison, the average angular error for the IMU heading direction was 13.7° during the same period. Additionally, the maximum error for the IMU heading direction is much worse than that for the visual estimate. These results indicate that the visual estimates of heading direction are much more accurate than those derived from the IMU.

[2]Available at http://youtu.be/zKHh1kndWs4

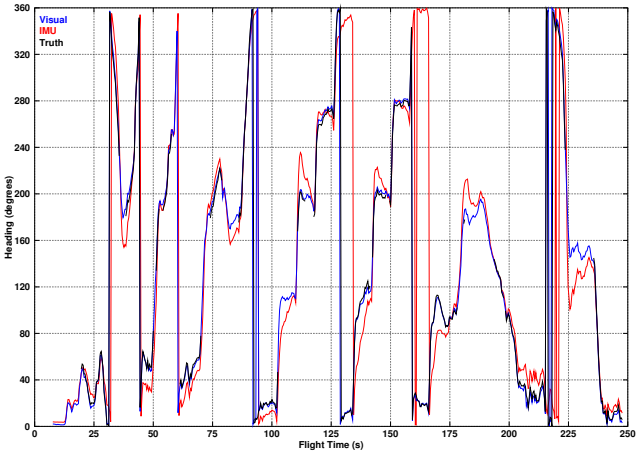

Figure 4: The heading direction during the flight test as measured automatically from the visual compass using the proposed approach (blue), by the IMU (red), and manually for ground truth (black). Gaps in the ground truth data represent frames in which the aircraft's shadow was not visible.

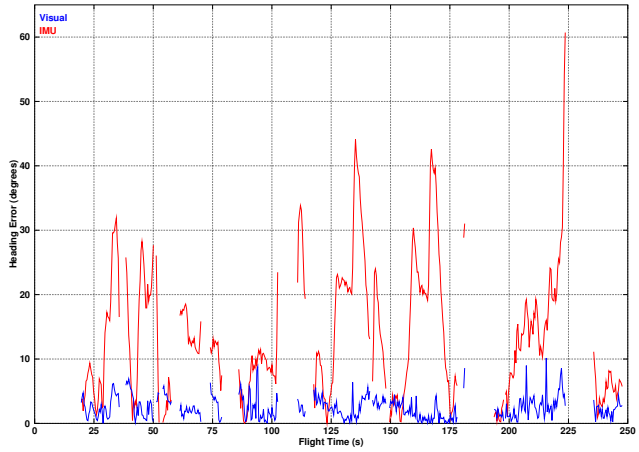

Figure 5: The absolute angular error between the visually estimated heading direction and the ground truth (blue), and the heading direction measured by the IMU and the ground truth (black). Gaps in the data represent frames in which the ground truth was unable to be measured.

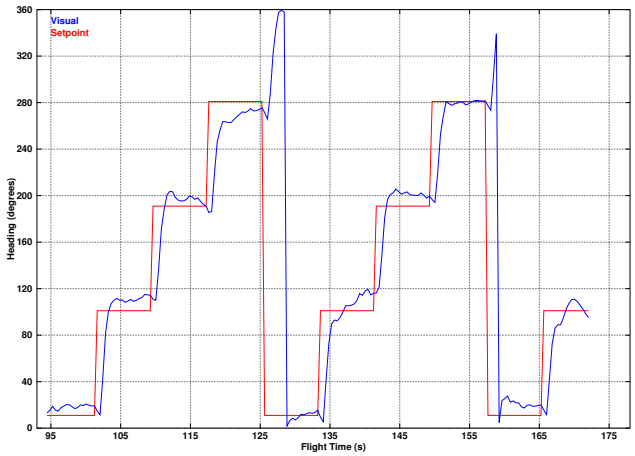

Figure 6: The visually estimated heading direction (blue) and the set-point (red) during the autonomous section of the flight test. The aircraft was under full autonomous control during the period 95s $\sim$ 173s.

This supports the results presented in [Moore *et al.*, 2011], where we found that the horizon-based approach for estimating the 2-DOF attitude of the aircraft – described here in brief – is significantly more accurate than an inexpensive IMU. Taken together with the results presented here, this indicates that the proposed approach is better able to provide accurate estimates of 3-DOF attitude than an inexpensive IMU.

The period of flight during which the aircraft was under autonomous control (95s $\sim$ 173s in Figure 4) is shown in Figure 6. It can be seen that the closed-loop control system was able to control the aircraft towards the desired heading direction effectively, although some improvement in the accuracy of the steady-state response of the aircraft to a heading direction command could be achieved by more finely tuning the heading direction PID.

4 System Performance

Due to the difficulty of obtaining proper ground-truth measurements of aircraft attitude during flight, it is hard to quantitatively compare the performance of the proposed visual 3-DOF attitude estimation scheme with other previously published approaches. However, we have implemented several of the state-of-the-art visual horizon detection schemes, and we show in [Moore *et al.*, 2011] that our approach for automatically computing 2-DOF attitude compares favourably, with an average angular error of 1.49°.

The approach described by [Mondragón *et al.*, 2010] for computing 3-DOF attitude from visual input is similar in purpose to ours. They compare the performance of their visual heading direction estimate to the output

from an IMU and quote an average RMSE of 5.8° from approximately 180s of short flight segments (compared to an average angular error of 2.47° for our visual heading estimate). However, we have shown here that inexpensive IMUs do not necessarily provide an accurate ground truth for heading direction. Furthermore, our visual compass is stabilised with a 2-DOF attitude estimate. Hence, we would expect the proposed system to function more accurately over a wider range of aircraft attitudes than the approach described by [Mondragón *et al.*, 2010]. In fact, our aircraft reached roll and pitch angles of ±80° during the flight tests analysed here, compared with ±15° for [Mondragón *et al.*, 2010].

Finally, in [Moore *et al.*, 2011] we show that our algorithm for computing 2-DOF attitude from the visual horizon is computationally efficient compared with other state-of-the-art approaches, requiring 2.0ms @ 1.0GHz to execute. Here, we compute the time required to visually estimate heading direction to be 2.4ms @ 1.0GHz, which gives a total execution time for visually estimating 3-DOF attitude as 4.4ms @ 1.0GHz. This indicates that the proposed approach is ideally suited to providing fast and accurate estimates of the 3-DOF attitude for an aircraft performing autonomous flight, and the results from our closed-loop flight testing (section 3.1) confirm this.

Limitations

We discuss the limitations of the proposed visual horizon detection algorithm in detail in [Moore *et al.*, 2011]. Briefly, however, vision systems with a very wide-angle field of view are able to image the majority of the horizon, irrespective of the roll and pitch angle of the aircraft, which greatly enhances the robustness of the 2-DOF attitude estimation procedure. Here also, the robustness of the procedure for estimating heading direction is greatly enhanced by utilising an omnidirectional vision system.

The main failure mode for the proposed approach for visually estimating heading direction is when there is insufficient texture or features for the visual compass to uniquely describe the heading direction of the aircraft. This may occur on days which are completely overcast, or on blue-sky days when the only significant feature is the Sun, which can be obscured by the body of the aircraft. However, by not explicitly extracting prominent features to track, and by matching the stabilised horizon images using an appearance-based algorithm instead, the visual compass is able to make use of other heading cues present in the environment – such as intensity gradients in the sky, which are present even when there are no prominent features.

For example, Figure 1e-h shows a dataset collected on a completely overcast day. However, the visual compass is able to match the stabilised horizon image correctly, despite the lack of distinctive features and large areas of the horizon and sky being obscured by the body of the aircraft. Additionally, it can be seen from the flight data in the attached video that the visual compass is able to make use of the intensity gradient over the whole sky to match the stabilised horizon image correctly when the only distinctive feature (the Sun) is obscured by the body of the aircraft.

To achieve stable estimates of heading direction over the course of hours or even days, an angular offset between the visually estimated heading direction and the direction of magnetic north can be estimated as described in section 2.2 and used to correct the visual heading direction estimate.

5 Conclusion

This study has described a novel, vision-based method for automatically obtaining the 3-DOF attitude of an aircraft. This approach involves capturing a very wide-angle view of the environment, including the horizon. A classifier is used to segment the scene into sky and ground regions, using the spectral and intensity properties of the input pixels. The aircraft's 2-DOF attitude is estimated by comparing the classified images with a database of reference image classifications and used to generate a stabilsed panoramic image of the horizon profile and sky.

This so called 'visual compass' is then used to determine the heading direction of the aircraft, relative to a reference horizon image. An appearance-based matching procedure allows the aircraft's heading direction to be estimated visually, even when there are no prominent features in the input image. Furthermore, by coupling the visual heading direction estimation with the visual horizon attitude estimation the full 3-DOF attitude of the aircraft is able to be determined robustly, irrespective of the orientation of the aircraft.

We have presented results previously ([Moore *et al.*, 2010]) that demonstrate the ability of this vision-based approach to outperform the 2-DOF attitude estimates obtained from an inexpensive IMU. Here, we present results from a recent flight test that demonstrate that the proposed approach is also superior for estimating heading direction. Furthermore, we have shown that the accuracy of this approach is better than that of other state-of-the-art, vision-based approaches, while being relatively computationally efficient and fast to execute. This system's robustness to varying environmental conditions and aircraft motions will make it a convenient option for any UAV application where the aircrafts 3-DOF attitude must be known precisely.

Acknowledgements

Sincere thanks to Mr. David Brennan, who owns and maintains the airstrip at which the flight testing was done, and to Mr. Samuel Baker, who performed the flight testing.

References

[Cornall and Egan, 2004] T. D. Cornall and G. K. Egan. Measuring horizon angle from video on a small unmanned air vehicle. In *2nd International Conference on Autonomous Robots and Agents*, pages 339–344. Citeseer, 2004.

[Cornall *et al.*, 2006] T. D. Cornall, G. K. Egan, and A. Price. Aircraft attitude estimation from horizon video. *Electronics Letters*, 42(13):744–745, 2006.

[Demonceaux *et al.*, 2006] C. Demonceaux, P. Vasseur, and C. Pegard. Omnidirectional vision on uav for attitude computation. In *Proc. IEEE International Conference on Robotics and Automation*, pages 2842–2847. IEEE, 2006.

[Dusha *et al.*, 2007] D. Dusha, W. W. Boles, and R. Walker. Fixed-wing attitude estimation using computer vision based horizon detection. In *Proc. 12th Australian International Aerospace Congress*, pages 1–19, Melbourne, Australia, 2007.

[Ettinger *et al.*, 2002a] S. M. Ettinger, M. C. Nechyba, P. G. Ifju, and M. Waszak. Towards flight autonomy: Vision-based horizon detection for micro air vehicles. In *Florida Conference on Recent Advances in Robotics*, volume 2002. Citeseer, 2002.

[Ettinger *et al.*, 2002b] S. M. Ettinger, M. C. Nechyba, P. G. Ifju, and M. Waszak. Vision-guided flight stability and control for micro air vehicles. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 3, pages 2134–2140. IEEE, 2002.

[Labrosse, 2006] F. Labrosse. The visual compass: Performance and limitations of an appearance-based method. *Journal of Field Robotics*, 23(10):913–941, 2006.

[McGee *et al.*, 2005] T. G. McGee, R. Sengupta, and K. Hedrick. Obstacle detection for small autonomous aircraft using sky segmentation. In *Proc. IEEE International Conference on Robotics and Automation*, pages 4679–4684. IEEE, 2005.

[Merhav, 1996] S. Merhav. *Aerospace sensor systems and applications*. Springer Verlag, 1996.

[Mondragón *et al.*, 2010] I. F. Mondragón, P. Campoy, C. Martinez, and M. Olivares. Omnidirectional vision applied to Unmanned Aerial Vehicles (UAVs) attitude and heading estimation. *Robotics and Autonomous Systems*, 58(6):809–819, 2010.

[Moore *et al.*, 2010] R. J. D. Moore, S. Thurrowgood, D. Bland, D. Soccol, and M. V. Srinivasan. UAV altitude and attitude stabilisation using a coaxial stereo vision system. In *Proc. IEEE International Conference on Robotics and Automation*, Anchorage, AK, May 2010.

[Moore *et al.*, 2011] R. J. D. Moore, S. Thurrowgood, D. Bland, D. Soccol, and M. V. Srinivasan. A fast and adaptive method for estimating uav attitude from the visual horizon. In *Proc. IEEE International Conference on Intelligent Robots and Systems*, San Francisco, CA, September 2011.

[Rohac, 2005] J. Rohac. Accelerometers and an aircraft attitude evaluation. In *IEEE Conference on Sensors*, page 6. IEEE, 2005.

[Rumelhart, 1985] D. E. Rumelhart. Learning internal representations by error propagation. Technical report, Institute for Cognitive Science, University of California San Diego, 1985.

[Scaramuzza and Siegwart, 2008] D. Scaramuzza and R. Siegwart. Appearance-guided monocular omnidirectional visual odometry for outdoor ground vehicles. *Robotics, IEEE Transactions on*, 24(5):1015–1026, 2008.

[Shimizu and Okutomi, 2003] M. Shimizu and M. Okutomi. Significance and attributes of subpixel estimation on area-based matching. *Systems and Computers in Japan*, 34(12), 2003.

[Thurrowgood *et al.*, 2009] S. Thurrowgood, D. Soccol, R. J. D. Moore, D. Bland, and M. V. Srinivasan. A vision based system for attitude estimation of UAVs. In *Proc. IEEE International Conference on Intelligent Robots and Systems*, St Louis, MO, October 2009.

[Thurrowgood *et al.*, 2010] S. Thurrowgood, R. J. D. Moore, D. Bland, D. Soccol, and M. V. Srinivasan. UAV Attitude Control using the Visual Horizon. In *Proc. Australasian Conference on Robotics and Automation*, Brisbane, Australia, December 2010.

[Todorovic and Nechyba, 2004] S. Todorovic and M. C. Nechyba. A vision system for intelligent mission profiles of micro air vehicles. *Vehicular Technology, IEEE Transactions on*, 53(6):1713–1725, 2004.

[Todorovic *et al.*, 2003] S. Todorovic, M. C. Nechyba, and P. G. Ifju. Sky/ground modeling for autonomous MAV flight. In *Proc. IEEE International Conference on Robotics and Automation*, volume 1, pages 1422–1427. IEEE, 2003.

[Zadeh, 1965] L. A. Zadeh. Fuzzy sets*. *Information and control*, 8(3):338–353, 1965.