A 3D sky compass to achieve robust estimation of UAV attitude

Tasarinan Jouir, Reuben Strydom and Mandyam V. Srinivasan
Queensland Brain Institute at The University of Queensland, Brisbane, Australia
tas.jouir@uq.net.au

Abstract
This study describes a novel, pure vision-based, method of obtaining the 3DOF attitude of an aircraft through the use of panoramic images of the sky. An iterative optimisation algorithm is applied to estimate the rotation of the aircraft in 3DOF by minimising the sum of absolute differences between two video frames of the sky. This method is unaffected by cumulative integration error, as all estimations are computed with respect to an initial reference image. Furthermore, the adverse effects resulting from movement of objects in the sky or changes in light intensity are ameliorated through the use of an adaptive reference image. Overall, the results illustrate the potential of this system to produce attitude estimates with an accuracy rivalling and in some cases surpassing those of two alternative vision-based methods, and an industry standard IMU.

1 Introduction
Efficient and accurate attitude estimation is a requirement for the autonomous navigation of unmanned aerial vehicles (UAVs). This is due to the challenging scenarios that UAVs face while performing complex navigation tasks such as search and rescue, or the exploration of remote terrain. The vehicle must be able to manoeuvre reliably under difficult conditions, with minimal or no direct control from a human operator.

Traditionally, attitude estimation is achieved using an inertial measurement unit (IMU), which utilizes a combination of accelerometers, gyroscopes and magnetometers. Unfortunately, both accelerometer and gyroscope based attitude estimation rely on integration over time which leads to noise-induced drift [Rohac, 2005]. Furthermore, magnetometers are incapable of functioning accurately in areas of magnetic irregularity, or when the rotational axis of the aircraft is parallel to the direction of the local magnetic field [Merhav, 2012]. The consequence of IMU drift is an ever increasing discrepancy between the self-determined orientation of the vehicle, and its true bearing. Ultimately this means that as the duration of the journey increases, the probability of the UAV reaching the intended destination is substantially reduced, potentially jeopardizing the mission.

A solution to this problem is to utilize the readily available information provided by the visual surroundings of the UAV, to complement and validate the IMU data. The use of visual information to determine attitude during flight is illustrated elegantly in nature. For example, flying insects stabilize their attitude by sensing the position of the horizon through their specialized light-sensitive organs, termed Ocelli [Srinivasan, 2011]. Some insects, such as honeybees, dung beetles and desert ants, also derive compass information from the sun and the polarized light pattern of the sky [Srinivasan, 2011; Dacke et al., 2003]. These elegant techniques in nature can be utilized to produce purely vision-based algorithms for attitude estimation in robots.

Attitude estimation using computer vision has been explored extensively in the field of robotics [Shabayek et al., 2012]. A number of techniques have been developed for UAVs that utilize the visual horizon to determine roll and pitch angles in a similar fashion to human pilots. Horizon detection was achieved by Dusha et al. [2007] through the use of a tri-color edge detection procedure followed by an application of the Hough transform. Candidate horizon profiles were tracked over time and statistical methods were implemented to weight the potential choices to determine the most likely profile. A similar technique involving line detection was used by Dumble and Gibbens [2012], however they extended the scope of the horizon line to take a more general shape rather than a straight line, to account for distant features on the horizon line such as hills, buildings and trees. The choice between candidate horizons was based on a comparison with a prediction utilizing the current roll and pitch angles. Moore et al. [2011a] describe a horizon de-
tion method based on an adaptive segmentation technique. This designated each pixel within a video frame to either the sky or ground region, with the horizon line forming the boundary between the two areas. These horizon-based methods, however, require a view of the horizon, the presence of a flat horizon devoid of hills, trees or man-made structures, and a robust way to identify the horizon line. Additionally, the horizon line does not provide any indication of a heading direction. To overcome these limitations, an alternative, vision-based approach is to use an image of the sky to estimate the attitude of an aircraft.

A study by Zeil et al. [2003] used a robotic gantry to illustrate the potential of image differencing to determine location in outdoor scenes due to the distinct pixel intensity minimum when comparing a current frame to a reference image. They also describe the negligible position change of distant objects during pure translation, and the change that occurs irrespective of distance during rotation. An image differencing algorithm implemented by Labrosse [2006] used sequential color images from a ground based robot using panoramic catadioptric video footage to determine the heading of the vehicle. The algorithm used the Euclidean distances between two consecutive frames in combination with a localized gradient descent optimisation. The captured images featured large sections of the ground, resulting in a reduced field of view being used to minimize the effect of the vehicles translation. Similarly it was assumed that the axis of the catadioptric camera system remained vertical in order to disregard the effect of roll and pitch on the images. The algorithm was also identified to suffer from integration error due to the use of data from consecutive frames. A similar image differencing technique with horizon detection in order to determine the heading of a ground-based vehicle, with the same sources of error as Labrosse [2006]. To provide all three degrees of freedom, a combination of the visual horizon and image differencing have been implemented as described below.

Mondragón et al. [2010] combine this image differencing technique with horizon detection in order to determine the attitude of the aircraft. Similarly Moore et al. [2011b] use the visual horizon to determine the roll and pitch angles, followed by an image differencing procedure to estimate the heading of the vehicle. In that study, the images used for the differencing process were stabilized using the roll and pitch data, obtained by using the horizon, in order to reduce the problem to detection of pure yaw. However, all of the methods that combine visual horizon detection for roll and pitch, and image differencing are influenced by the accuracy of the horizon estimates.

This paper describes a novel, vision-based method of estimating the attitude of a UAV during flight that uses an image matching algorithm applied to panoramic video footage of the sky. The advantage of using the sky image is that objects are far away, so that any changes in the sky image between successive frames are only due to the rotation of the aircraft, and not due to its translation. Individual frames are compared to a cumulative reference image by means of a sum of the absolute pixel-wise differences (SAD) intensity minimisation. This is done in order to utilize the rotational information encapsulated by the changes in the positions of the images of distant, stationary features in the sky, such as the sun, the clouds and intensity gradients across consecutive frames. The novelty of this approach is that it is a purely vision-based method of determining the 3DOF aircraft attitude, requiring only a view of the sky. This study illustrates the potential of such a system to perform as well as or better than both an industry standard IMU (as tested against a MicroStrain GX3), and other state of the art vision based methods in a variety of sky conditions.

2 Flight Platform

Within this study a custom built quad-rotor (see Figure 1) was used for flight testing. Rotor control for the aircraft was achieved using a MicroKopter flight controller, speed controller and motors. The quad-rotor featured an Intel NUC computer (with a 2.6GHz dual-core processor). A MicroStrain 3DM-GX3-25 AHRS IMU was installed, as well as the vision system (described in more detail by Thurrowgood et al. [2014]) shown in Figure 2.

The vision system uses two back to back 190° field of view fisheye cameras operating at a 25Hz frame rate (Point Grey Firefly MV cameras with Sunex DSL216 lenses). A stereo overlap of approximately 30° by 130°, with a baseline of 10cm, is achieved by tilting both cameras 10° towards the front. Calibration of the cameras was completed using a method similar to that described by Kannala and Brandt [2006]. A stitched panoramic image is created, with a field of view (FOV) of approximately 360° azimuth and 150° elevation, where the resulting stitched image has a resolution of approximately 360 by 220 pixels. Furthermore, a mask is generated
to conceal the regions of the aircraft, such as the rotor blades, that are visible within each frame.

3 Sky-based Attitude Estimation

Movement of the sky image (including the sun and any clouds, as well as the overall intensity and color gradients) provides pure rotational information about the aircraft, uncontaminated by any translation due to the sky features being at a large distance [Zeil et al., 2003].

The attitude computation proposed here takes advantage of this concept using difference images of the sky, captured in successive camera frames, to compute the roll, pitch and yaw of a UAV. This method of estimating UAV attitude can be broken into three stages:

1. Preprocessing of the panoramic images for the optimiser
2. Attitude estimation using an iterative optimisation on the sum of absolute differences of the current image and reference image
3. Updating the reference image to account for environmental changes

Every frame undergoes this process, enabling computation of the instantaneous absolute 3DOF attitude at any instant of time, without integration error. Each of these steps is explained in detail below.

3.1 Preprocessing

A Gaussian blur (using a $5\times5$ kernel) is applied to the panoramic images described in Section 2. The blurring operation is used to smooth the optimising function, which in turn reduces the number of local minima encountered during the optimisation process for attitude estimation. After the first frame is processed, it is stored as the base reference image.

3.2 Attitude Estimation

An accurate representation of an image in three dimensions can be obtained through the use of a “world sphere” as described in Figure 3. The world sphere provides a non-warped representation of the panoramic images generated by the vision system, mapped onto the unit sphere. As such, each pixel location represents a 3D vector on the world sphere. This concept is used here to define the roll, pitch and heading directions, while also allowing for the rotation of the image (through the application of a rotation matrix to the 3D vector representation of the pixel location). Note that projecting the panoramic images onto the world sphere does not provide any depth information, rather it helps to accommodate the earlier assumption that objects in the sky are infinitely far away (and therefore present no parallax).

The proposed attitude estimation procedure uses both the intensity gradient and the geometry of a scene to predict the rotated image in the next frame for any particular combination of roll, pitch and yaw. An iterative optimisation algorithm is applied to determine the specific combination of roll, pitch and yaw that best predict the image in the next frame. This provides the best estimate of the aircraft’s roll, pitch and yaw for any given frame, with respect to the aircraft’s orientation at the time of the reference image. This procedure is described in greater detail below.

![Figure 3: A pixel located in (a) the panoramic image space: $P(\phi, \theta)$, where $\phi$ represents the azimuth and $\theta$ the elevation; and (b) the world sphere: $P(r, p, h)$ along with respective coordinate definitions.](image)

![Figure 4: Illustration of the “sky cone” used to check the validity of a pixel.](image)

To estimate the rotation of the sky image in 3DOF, an initial arbitrary rotation in 3DOF is used to apply an
appropriate shift to each valid pixel in the current image. The validity of a pixel is determined by its location. Specifically, valid pixels are defined as pixels that do not lie within the mask (see Section 2) of both the current and reference images, but which also exist within the area encompassed by the intersection of a 150° globally vertically oriented cone (see Figure 4) and the surface of the world sphere.

The intensity of each pixel in the rotated image is calculated using bilinear interpolation of the neighbouring pixels at the estimated location within the reference image. After the intensities of all of the pixels in the rotated image have been computed, the pixel-wise intensity difference between the original and the rotated image is calculated by computing the sum of the absolute pixel-wise differences (SAD). In order to remove the effect of varying numbers of valid pixels on the estimated attitude, the SAD is normalized by the number of valid pixels that were used in the computation.

The SAD is minimised iteratively using an algorithm from the NLopt library [Johnson, 2014]. The algorithm used is BOBYQA [Powell, 2009], a derivative free gradient descent method. By using the absolute 3DOF rotation as the input variables, the NLopt optimization scheme is completed by minimizing the normalized SAD value as computed between the current frame and the reference image. This is done in order to maximize available pixel information (i.e. to increase the number of valid pixels) for future attitude estimations, while also compensating for changes in illumination, shape and position of any cloud formations. Furthermore, iteratively updating the reference image reduces the effect of objects (such as other aircraft or tall structures/trees) that are temporarily within the field of view.

The accumulating reference image used for this study was inspired by the work of Moore et al. [2011b]. The updating process in both that study and the present one is defined as

\[
d^i = \alpha_d c^i + (1 - \alpha_d)d^{i-1}
\]

where \(d_i\) and \(c_i\) are the \(i^{th}\) pixels in the accumulating reference image and the current image respectively, and \(\alpha_d\) is the accumulation rate of the reference image (similar to that study we use \(\alpha_d = 0.01\), which corresponds to an update time of \(\tau_d \sim 4s/0.25Hz\)). Only valid pixels (see Section 3.2) are used during this process to ensure that the reference frame is not contaminated. Note that the increase of information is due to the removal of sections of the image mask corresponding to the updated areas of the reference image. Figure 5 illustrates the effect of the accumulating reference image procedure as the quadrotor yaws through 360°.

As indicated in Section 3.2, the attitude estimation requires an initial input guess as well as search bounds, which are updated after the completion of the optimization process for each frame. These values are used in order to reduce the number of computations required, as
well as the search space and the chance of encountering a local minimum. Similarly the new location of the sky cone relative to the normal vector of the aircraft is computed using a three dimensional rotation matrix, further maximizing the number of valid pixels that are available for the subsequent optimisation process.

4 Results

This section provides a description of the processes undertaken, and the results obtained during the performance evaluation of the sky-based algorithm with respect to both a mechanical ground truth and the onboard IMU. Tests were completed in different sky conditions which are defined as: clear sky (less than 5% cloud), partial cloud (between 5 – 85% cloud) and full cloud (greater than 85% cloud).

4.1 Validation

In order to validate the algorithm outlined above, while also gauging the performance of the IMU, a tripod mount (see Figure 6) was designed for the quad-rotor to provide a mechanical ground truth. Three independent protractor-based measurement systems, accurate to ±0.5° for roll, pitch and heading rotations, with rotational axes parallel to those of both the vision system and the IMU were fixed to the 3-way pan-tilt head of the tripod. Each axis featured a locking mechanism to reduce operator error, while the mounting plate and counter-weight were designed to ensure that the tripod was balanced when loaded with the quad-rotor.

The secondary goal of the tripod validation procedure was to quantitatively determine the suitability of the IMU to act as a quasi-ground truth sensor during the analysis of the post-processed flight data, described in Section 4.2.

Each angle of rotation was tested individually, with the other (secondary) angles held constant at 0°. The primary angle was then increased from 0° through to either ±45° in 5° increments at ~15s intervals. Upon reaching 45°, the primary angle was returned directly to 0°, and held for a further ~15s. This test was performed a total of 27 times: three times for each rotation angle (roll, pitch and heading), in three different sky conditions (clear, partial cloud, and full cloud), resulting in nine complete data sets (with each set consisting of three tests, one for each rotation angle). One such data set is shown in Figure 7. Due to the mechanical nature of the ground truth system, only data contained within the shaded regions (corresponding to a period of 12s with the primary angle held constant) is deemed valid. The mean and standard deviation of the sky-based estimate and the IMU error relative to the mechanical ground truth during all shaded time periods in Figure 7 over each individual test was calculated and recorded in Table 1 for analysis.

It was identified from these results that the heading estimate produced by the IMU featured significant error, eliminating its potential use as a ground truth for heading during flight testing. In contrast, the roll and pitch evaluations of the IMU showed minimal error (on average close to or less than the accuracy of the tripod ground-truth, specifically ±0.5°). This indicated that these measurements were sufficiently accurate to be used as a reference for evaluating the performance of the proposed algorithm during the analysis of the post-processed flight data.

4.2 Flight Tests

The performance of the proposed attitude estimation scheme when subjected to a dynamic, real world environment was evaluated in 10 flight tests. Similar to the validation procedure described above, these tests were conducted in a variety of sky conditions (clear, partial cloud, and full cloud), while completing manoeuvres of varying intensity (ranging from placid to aggressive). Each flight was carried out under manual control, with both IMU and video data recorded for post-processing. After the application of the sky-based algorithm, each attitude estimate was compared to the corresponding IMU data, both directly and by computing probability distribution plots of the differences (note: for completeness, roll, pitch and heading were all considered, although the heading comparison does not provide valid quantitative results because the IMU readings were deemed to be erroneous for yaw, as shown in Figure 7 and Table 1).
Figure 7: The sky estimate (blue), the IMU data (red), and the tripod ground truth (black) for a roll test (a), a pitch test (b), and a heading test (c) all completed with full cloud sky conditions. Note that only data in the shaded regions is considered, as the ground truth is unknown outside these areas.

<table>
<thead>
<tr>
<th>Sky Condition</th>
<th>Sky-based ($\mu \pm \sigma^\circ$)</th>
<th>IMU ($\mu \pm \sigma^\circ$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>P</td>
</tr>
<tr>
<td>Clear 1</td>
<td>0.50 ± 0.03</td>
<td>-0.42 ± 0.05</td>
</tr>
<tr>
<td>Clear 2</td>
<td>-0.11 ± 0.08</td>
<td>-0.23 ± 0.04</td>
</tr>
<tr>
<td>Clear 3</td>
<td>-0.24 ± 0.02</td>
<td>-0.35 ± 0.07</td>
</tr>
<tr>
<td>Clear Average</td>
<td>0.05 ± 0.04</td>
<td>-0.33 ± 0.05</td>
</tr>
<tr>
<td>Partial Cloud 1</td>
<td>-0.10 ± 0.02</td>
<td>-0.05 ± 0.06</td>
</tr>
<tr>
<td>Partial Cloud 2</td>
<td>-1.50 ± 0.08</td>
<td>3.73 ± 0.13</td>
</tr>
<tr>
<td>(1.34 ± 0.10)</td>
<td>(0.75 ± 0.13)</td>
<td>(0.76 ± 0.05)</td>
</tr>
<tr>
<td>Partial Cloud 3</td>
<td>-0.70 ± 0.04</td>
<td>3.51 ± 0.13</td>
</tr>
<tr>
<td>Partial Cloud Average</td>
<td>-0.77 ± 0.05</td>
<td>2.40 ± 0.11</td>
</tr>
<tr>
<td>(0.18 ± 0.06)</td>
<td>(-0.08 ± 0.10)</td>
<td>(-0.88 ± 0.04)</td>
</tr>
<tr>
<td>Full Cloud 1</td>
<td>-3.56 ± 0.11</td>
<td>0.71 ± 0.08</td>
</tr>
<tr>
<td>(0.74 ± 0.11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Cloud 2</td>
<td>0.07 ± 0.02</td>
<td>-0.72 ± 0.07</td>
</tr>
<tr>
<td>Full Cloud 3</td>
<td>0.00 ± 0.01</td>
<td>-0.73 ± 0.04</td>
</tr>
<tr>
<td>Full Cloud Average</td>
<td>-1.16 ± 0.05</td>
<td>-0.25 ± 0.06</td>
</tr>
<tr>
<td>(0.27 ± 0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Average</td>
<td>-0.63 ± 0.05</td>
<td>0.61 ± 0.07</td>
</tr>
<tr>
<td>(0.17 ± 0.05)</td>
<td>(-0.22 ± 0.07)</td>
<td>(-0.53 ± 0.05)</td>
</tr>
<tr>
<td>Total Average</td>
<td>-0.10 ± 0.03</td>
<td>-0.26 ± 0.06</td>
</tr>
</tbody>
</table>

Table 1: Error of both the sky-based and IMU attitude estimates relative to the mechanically determined ground truth for a variety of sky conditions. The red highlights represent tests affected by lens flare, while the blue highlights represent tests affected by large amounts of rapid, uniform cloud movement. Note that the values in brackets are the results after an application of the correction method outlined in Section 5.4. To average over the same data sets, the greyed out IMU values were excluded.

An example of a typical flight data set is described in Figure 8 (a-c), with the sky-based estimate shown in blue and the IMU data in red. A graphical reconstruction of the approximate trajectory as recorded using visual odometry [Strydom et al., 2014; Thurrowgood et al., 2014] is illustrated in Figure 8 (d).

Figure 9 shows the probability distributions of the differences between the roll, pitch and heading estimates of the sky-based algorithm and the corresponding IMU readings. The data included in this evaluation encompass the attitude estimates of every frame for all 10 flight tests, corresponding to n=21,527 difference values. The mean and standard deviation of the difference values were calculated to be: for roll (0.28 ± 4.77º), pitch
Figure 8: Flight data taken while performing aggressive manoeuvres under a full cloud sky condition. (a)-(c) sky-estimate (blue) and IMU data (red) for roll, pitch, and heading respectively; (d) reconstruction of approximate flight path (take-off: ○, landing: ●) using visual odometry.

(1.84 ± 5.37°) and heading (2.29 ± 11.03°).

5 Discussion

This section summarises our analysis of the results described above, along with a comparison to alternative vision based attitude estimation techniques. Applications of the proposed algorithm are identified, while also touching upon the associated limitations.

5.1 Analysis of Results

The results of the tripod validation procedure (described in Section 4.1) suggest that the performance of the proposed sky-based attitude estimation is as good as or better than that of the on-board IMU for both roll and pitch, as the average angular error for each method is less than that of the ground truth system. Furthermore, the heading estimation of the sky-based algorithm was shown to be superior to that of the IMU, producing an average mean error of −0.16° compared to 2.31° respectively (see Table 1). Note that this is when considering the corrected values (blue highlights), and ignoring the effect of glare (red highlights) as explained in Section 5.4. Although the effect of coupling was not tested during the tripod validation procedure, it was indirectly investigated during the flight testing. There is no evidence of any error arising from coupling between pitch and roll. These findings are reinforced by the results of the post-processed flight tests, which show an average roll and pitch discrepancy of 0.28 ± 4.77° and 1.84 ± 5.37° when considering the mean difference between the pro-
posed algorithm and the IMU data for each frame over all 10 flights. As predicted by the results of the tripod evaluation, the average mean difference in heading estimation was substantially larger (2.29 ± 11.03°), however there was a distinct correlation between the shapes of the two plots (see Figure 8 (c)). The results presented here corroborate the high level of accuracy obtained by the proposed vision-based attitude estimation method in comparison with an IMU, even if the IMU is of a high industry standard (we used a MicroStrain 3DM-GX3-25 AHRS IMU).

5.2 Comparison with Existing Methods

Comparison between independent attitude estimation techniques is difficult due to the lack of a standard ground truth, however we shall attempt to do so here. Mondragón et al. [2010] also compare their vision based method to an IMU, reporting average root-mean-square errors of 1.7°, 3.3°, 5.8° for roll, pitch and heading respectively. The average angular error for roll and pitch reported by Moore et al. [2011b] was 1.49°, along with an error of 2.47° for their heading estimate. That study manually tracked the shadow of the aircraft as a method of determining the ground truth for yaw, while also manually computing the roll and pitch from the horizon for evaluating the accuracy of those angle estimates. Our algorithm performed substantially better than the method proposed by Mondragón et al. [2010] (−0.16° vs 1.7°, 1.84° vs 3.3°, 2.29° vs 5.8° for roll, pitch and heading), and slightly better than that of Moore et al. [2011b] (average angular error for roll and pitch of 1.06° vs 1.49°, and for heading: 2.29° vs 2.47°). Note that we have shown through the validation procedure (Section 4.1) that our method produces a heading estimation of higher accuracy than the IMU, so the heading comparison is only included above for completeness (the true heading error is expected to be significantly less).

Furthermore, the fundamental limitation of both the horizon based attitude estimation schemes (the requirement that the horizon is flat and visible) is overcome by the present method as it only requires that a portion of the sky be within the field of view. Additionally, the proposed algorithm generates absolute attitude estimates without the integration errors suffered by a number of the horizon-based techniques.

5.3 Applications

Vision based attitude estimation has a range of potential applications both as a corrective measure, or surrogate to, the standard IMU based navigation systems. For example, extra-terrestrial operations (such as the exploration of the surface of Mars where NASA’s Mars rover experienced up to 31° tilt in the first two years of exploration [Maimone et al., 2007], or flights within the atmosphere of Titan) face significant communication delays between the platform and the human operators on Earth. As this essentially removes the option of direct control, it is of utmost importance that navigational errors do not accumulate over time. This illustrates the need for a robust correction and validation scheme, such as a supplementary vision-based system - which would also provide a solution to IMU sensor failure. Similarly, one proposed commercial application of UAVs is the delivery of parcels within residential areas. A vision-based system would ensure highly accurate navigation over long flights where an IMU may drift. Furthermore, UAVs are being increasingly used in the aftermath of environmental disasters (such as following a cyclone or tsunami) to provide assistance in mapping areas where existing maps and often communication are no longer valid. In these situations, vision based attitude estimation could provide a cheap and reliable alternative to
an IMU, simply by further utilizing the camera system required for such applications.

5.4 Limitations

The proposed algorithm is reliant on a constant view of the sky, however this requirement is less restrictive than the corresponding limitation of the alternate vision based methods outlined in Section 1. Moreover, measures have been implemented to reduce the effects of objects temporarily obscuring the field of view through the use of the accumulative reference image as discussed in Section 3.3, further negating the effect of this constraint.

During the testing process, it was identified that the presence of lens flares dramatically reduced the performance of the algorithm. Flares unfortunately move in an unpredictable manner, resulting in substantial propagation of error as these artificial objects are incorrectly matched by the SAD minimization scheme. The entries of Table 1 highlighted in red are examples of this effect, with the error in the measurement of each angle ($-1.45 \pm 0.07^\circ$ and $-1.60 \pm 0.01^\circ$) being substantially higher than the average calculated when this data is excluded ($-0.16 \pm 0.01^\circ$). We have classified lens flares as an inevitable problem faced by vision based attitude estimation methods, and ultimately as an area requiring additional research.

Similarly, rapid uniform movement of clouds resulted in an artificial rotation, synonymous to a constant increase in estimation error. This is a by product of the image differencing process, however in most cases the accumulative reference image is able to negate the issue. Four occurrences of these rapid uniform changes which were not immediately resolved are listed in Table 1 (highlighted in blue). These data sets were obtained on particularly windy days, and all featured cloud coverage greater than $\sim 75\%$. A proposed solution to this issue is to estimate the angular velocity of the clouds by taking the mean angular step size between each frame when in a known static orientation. This correction method was applied to the post-processed tripod data using the first $\sim 65$ frames ($\sim 2.6s$) of footage to estimate the cloud angular velocity. The corrected results are listed in brackets below each blue highlighted entry in Table 1. Incorporating this solution directly into the proposed algorithm would completely account for this effect, only requiring a brief analysis of cloud velocity either when commencing a flight or while hovering using alternate methods of attitude stabilisation.

6 Conclusion

This study has described a pure vision-based method for robust determination of the absolute attitude of a UAV. The proposed algorithm utilizes the rotational information that is generated by the sky image (which comprises the sun, clouds, and other intensity gradients). This is achieved through the use of a SAD minimization applied to the pixel-wise intensity difference between successive video frames, producing a distinct minimum corresponding to the attitude of the aircraft.

We have shown through a rigorous validation process that the proposed method is capable of retaining accuracy when subject to a variety of sky conditions, producing results competitive with those of an industry standard IMU. Future work will include optimization of the algorithm in order to complete real time flight testing, as well as an extension of the image differencing method to incorporate translation and compute egomotion in 6DOF.

Acknowledgements

The authors of this paper would like to thank Mr. Jovian Delaforce for his invaluable assistance with field testing. We would also like to thank Mr. Dean Soccol for his ingenuity regarding the design and construction of the tripod ground truth system.

The research described here was supported partly by the ARC Centre of Excellence in Vision Science (CE0561903), by Boeing Defence Australia Grant SMP-BRT-11-044, ARC Linkage Grant LP130100483, ARC Discovery Grant DP140100896, a Queensland Premiers Fellowship, and an ARC Distinguished Outstanding Researcher Award (DP140100914).

References


