TCM: A Fast Technique to Determine if an Object is Moving or Stationary from a UAV

Reuben Strydom, Saul Thurrowgood and Mandyam V. Srinivasan
The Queensland Brain Institute and the School of Information Technology
and Electrical Engineering at The University of Queensland, Brisbane, Australia
r.strydom@uq.edu.au

Abstract

This paper demonstrates a novel technique to distinguish between moving or stationary objects from an Unmanned Aerial Vehicle (UAV). Our Triangle Closure Method (TCM) is able to robustly classify – with an average accuracy of 91.9% as computed in outdoor field tests – the motion of an object using information on the egomotion of the UAV (derived from optic flow) and on the expansion/contraction of the object's image. These parameters are used to compare the predicted and measured changes in the distance to the object, which is then used to classify the object's motion. The main advantage of this method compared to current solutions is that it can be applied in environments with strong motion parallax.

1 Introduction

Unmanned Aerial Vehicles (UAVs) are becoming an integral tool in military and civilian applications. Safety is of growing concern as the numbers of autonomous or semi-autonomous vehicles increase at unprecedented rates [Valavanis, 2008]. Significant research has been conducted on navigation and guidance of autonomous robots, including object detection, collision avoidance and interception. Before an object can be intercepted or avoided it must first be detected, and it is paramount to differentiate between static and moving objects, either for the purpose of pursuing them or avoiding collisions with them [Lai *et al.*, 2013; Nussberger *et al.*, 2014].

Radar is a popular technique to detect objects and determine their motion. Ground-based (e.g. [Clothier *et al.*, 2012; Wilson, 2012]) or airborne (e.g. [Korn and Edinger, 2008; Viquerat *et al.*, 2008]) radar have been employed to detect the motion of objects. However, with the growing trend towards UAVs and micro UAVs, these systems are prohibitively large and expensive. To meet the cost and size constraints for UAVs, effective vision-based bio-inspired algorithms are utilised for navigation and guidance [Green *et al.*, 2004; Zufferey and Floreano, 2006; Denuelle *et al.*, 2015].

For informed UAV guidance decisions in urban environments, knowledge of obstacles and their motion is required. First, an object must be detected, then classified as moving or stationary, and finally it can be avoided, intercepted or followed. As the field of object detection is well studied, this paper will focus on the second part, which is to reliably classify if a detected object is moving or stationary.

It is a relatively simple task to determine whether an object is moving or stationary from a static vision system. One method that is commonly employed is static background subtraction [Elhabian *et al.*, 2008] – in this paper, background subtraction in the context of a stationary observer will be referred to as static background subtraction. The first step of static background subtraction is to initialise a model of the background. As the vision system is static, it is possible to assume that the background is static, and thus only object motion will change the image. The background model must be updated periodically to adapt to dynamic scenes (e.g. illumination changes) [Zhu and Yuille, 1996]. Using the static background assumption, the next step is to detect foreground pixels that differ from the background model. Once the foreground pixels are detected, they must be post-processed to increase the signal to noise ratio. Following post-processing, the foreground pixels are then clustered utilising a connected component algorithm in order to detect the moving object [Wu and Leahy, 1993]. The benefit of this method is that it can detect very slow moving objects, however, background subtraction can produce false positives under dynamic scenes. Another approach is to use temporal differencing (known as frame subtraction), which encompasses techniques that compute a pixel-wise difference between consecutive frames [Spagnolo *et al.*, 2006]. Although frame differencing is a simple, fast technique that is well suited to dynamic scenes (e.g. lighting changes), it can overlook slow moving objects [Paragios and Deriche, 2000]

and is sensitive to thresholds.

The main limitation to the methods explained above is that they are constrained to a static camera. However, when the vision system itself is in motion, the task becomes more challenging. This is due to the fact that, in general, the images of the object as well as the background will be in motion, even when the object is stationary. Thus, a moving object cannot be detected simply in terms of motion parallax. Although humans can do this effortlessly, it is computationally non-trivial.

There are three main approaches for determining whether an object is moving while the observer is in motion, (as discussed in [Pinto *et al.*, 2015]): (1) organising the background into mosaics, (2) background subtraction and (3) optical flow and geometrical models.

To detect if an object is moving by organising the background into mosaics, a spatial image registration technique is applied [Ibrahim *et al.*, 2010; Li *et al.*, 2013]. Simply registering two frames by mosaicing often induces a large amount of noise – this in turn could increase the number of false positives.

The most popular methods by which a moving vision system detects a moving object use a cue based on motion contrast (encompassing (2) and (3)), i.e., the presence of a difference in the direction and/or magnitude induced by the motion in the image of the object, relative to its immediate background. The use of background subtraction to detect moving objects from a freely moving vision system is well studied [Hayman and Eklundh, 2003; Sheikh *et al.*, 2009; Kim *et al.*, 2010]. This method is similar to that of static background subtraction, however, the motion of the vision system is first taken into account. By utilising the known camera motion, the spatial information is less noisy than that of organising the background into mosaics. Optic flow is another method utilised to detect motion contrast [Thakoor *et al.*, 2004; Pinto *et al.*, 2014; Choi *et al.*, 2011]. These methods, however, usually require computation of dense optic flow, as sparse optic flow could miss small or slow moving objects [Sundaram *et al.*, 2010]. Although dense optic flow provides the ability to utilise clustering techniques, it drastically increases the computational time.

Motion contrast is a reliable method that detects a moving object if it is approximately at the same distance from the vision system as the background. However, when the object is at a different distance, such as an elevated ball, motion contrast will detect motion even when the object is stationary. This is currently a major limitation of many current techniques that rely solely on motion contrast to detect if an object is moving.

The Triangle Closure Method (TCM), presented in this paper, overcomes this limitation by combining the expansion of an object and the direction of motion of the UAV to classify the motion of an object regardless of the object's position with respect to the background. This method is well suited for the visual guidance of robots in cluttered environments. The primary goal of this study is not object detection *per se*, but rather determining whether a given object is stationary or moving in absolute 3D space. While this may seem like a trivial task (from everyday human experience), it is algorithmically challenging for a machine vision system.

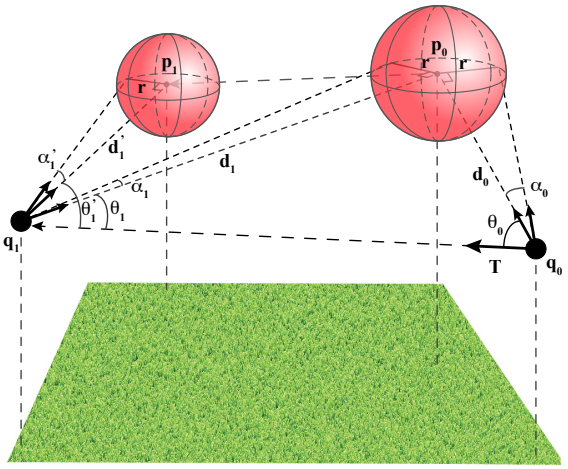

Figure 1: Determination of whether an object is moving or static, from information about the UAV's translation unit direction (T), the unit direction vectors (d_0 and d_1) connecting the UAV and the object, and its expansion between two image frames. α_0 and α_1 represent the angular sizes of the images of the object in the two frames. Primed symbols ($d_1^{'}$ and $\alpha_1^{'}$) represent the angles and vectors, had the object moved between two frames. When the object is stationary a closed triangle is formed between the motion of the UAV and a static object (vertices q_0, q_1 and p_0).

2 TCM: The Triangle Closure Method

The Triangle Closure Method, coined TCM, is a novel technique to determine, once detected, if an object is moving or stationary. The major advantage of TCM, compared with current techniques that use motion contrast, is that it robustly distinguishes between a static or moving object even if it is at a different distance to the background. It can also easily be implemented on a vision-based platform, as described later.

A typical scenario would be a quadrotor translating from position q_0 at t_0 to q_1 at t_1, while there are two possibilities for the object: (1) the object is static at p_0 between t_0 and t_1 or (2) the object moves from p_0 to p_1. Between two frames (t_0 to t_1) the Triangle Closure Method (TCM) classifies the motion of an object by computing the following information:

1. The translation direction of the UAV
2. The direction to the centre of the object and its radius
3. Predicted change in distance of the object
4. Measured change in distance of the object

To compute the predicted change in distance between two frames, the translation direction and the direction to the centre of the object are used. The vectors T, d_0 and d_1 are used to compute the angles θ_0 and θ_1 in (1). Exploiting the law of sines, the angles in (1) are utilised to compute the ratio between the distances d_0 and d_1 (2).

$$\begin{aligned} \theta_0 &= cos^{-1}(\hat{d_0} \cdot \hat{T}) \\ \theta_1 &= cos^{-1}(\hat{d_1} \cdot \hat{T}) \end{aligned} \tag{1}$$

$$\frac{\|d_0\|}{\|d_1\|} = \frac{sin(\theta_1)}{sin(\theta_0)} \tag{2}$$

Once the predicted distance ratio is computed, it is compared to the measured change in distance. The change in distance in this case is computed by measuring the expansion/contraction of the object's image (3); i.e. if the object is approaching, then α_1 would be greater than α_0 and the opposite is true for a receding object.

$$\frac{\|d_0\|}{\|d_1\|} = \frac{tan(\alpha_1)}{tan(\alpha_0)} \tag{3}$$

This comparison between the predicted and the actual ratio of distances can be used to determine if the object is moving or stationary. For example, in theory if the object were static, the predicted and measured distance changes would be equal (4). On the other hand, if the object were moving, there would be a discrepancy between the predicted and measured values (5), which we will refer to as disparity (δ).

$$\frac{sin(\theta_1)}{sin(\theta_0)} = \frac{tan(\alpha_1)}{tan(\alpha_0)} \tag{4}$$

$$\delta = \left|\left| \frac{sin(\theta_1)}{sin(\theta_0)} - \frac{tan(\alpha_1)}{tan(\alpha_0)} \right|\right| \tag{5}$$

Note: The derivation provided above uses the values in Figure 1 had the object been stationary (e.g. d_0, d_1, α_0 and α_1). However, had the ball been moving, that is, $\delta > \delta_{threshold}$ from t_0 to t_1, these values would be represented by the primed symbols ($d_1^{'}$ and $\alpha_1^{'}$). Although the disparity should equal zero if the object is stationary, in reality there is noise in the measurements of the UAV's translation direction, and the directions and sizes of the images of the object. Thus, a threshold is used to account for this noise, as shown in (6). A full sensitivity and noise analysis is provided in the next section.

$$Classifier : \begin{cases} Moving\ if\ \delta > \delta_{threshold} \\ Static\ if\ \delta < \delta_{threshold} \end{cases} \tag{6}$$

2.1 Sensitivity and noise analysis

The signal to be detected is the disparity, δ. It is evident from (5) that δ will have the largest magnitude when $\theta_0 \cong 0°$, $\theta_1 \cong 90°$, α_0 is large, and α_1 is small. This would correspond to the object being viewed frontally and presenting a large image in the first frame, and laterally, presenting a small image in the second frame. This condition would occur when the object is moving away from the aircraft, along a perpendicular trajectory. This is confirmed by the results of the simulation shown in Figure 5, where the moving object generates maximum disparity when it is moving in a direction perpendicular to the rotorcraft's trajectory.

Another condition that would produce a large magnitude of δ would be when $\theta_0 \cong 90°$, $\theta_1 \cong 0°$, α_0 is small, and α_1 is large. This would correspond to the object being viewed laterally and presenting a small image in the first frame, and frontally, presenting a large image in the second frame. Here, the object is moving toward the aircraft, along a perpendicular trajectory. For purposes of illustration, we consider only the first condition in the further discussion below.

Let us now consider the influence of noise in the measurements of the various angles. We rewrite (5) as:

$$\delta = |A - B| \tag{7}$$

where $A = \frac{sin(\theta_0)}{sin(\theta_1)}$, and $B = \frac{tan(\alpha_0)}{tan(\alpha_1)}$

The perturbation of term A due to measurement noise can be expressed as:

$$\begin{aligned} \Delta A &= \frac{sin(\theta_0) + \Delta\theta_0}{sin(\theta_1) + \Delta\theta_1} - A \\ &= \frac{sin(\theta_0) + \Delta\theta_0}{sin(\theta_1)[1 + (\Delta\theta_1 / sin(\theta_1))]} - A \end{aligned} \tag{8}$$

where the Δ symbols represent small perturbations in the measurements of the corresponding angles.

For small $\Delta\theta_1$, (8) can be rewritten as:

$$\begin{aligned} \Delta A &\cong \frac{[sin(\theta_0) + \Delta\theta_0]}{sin(\theta_1)} \left[1 - \frac{\Delta\theta_1}{sin(\theta_1)}\right] - A \\ &= \left[\frac{sin(\theta_0)}{sin(\theta_1)} + \frac{\Delta\theta_0}{sin(\theta_1)}\right] \left[1 - \frac{\Delta\theta_1}{sin(\theta_1)}\right] - A \end{aligned} \tag{9}$$

Noting that $\frac{sin(\theta_0)}{sin(\theta_1)} = A$, and dropping second-order Δ terms, (9) can be expressed, after some algebraic manipulation, as:

$$\begin{aligned} \Delta A &\cong \frac{\Delta\theta_0 \cdot sin(\theta_1) - \Delta\theta_1 \cdot sin(\theta_0)}{sin^2(\theta_1)} \\ &= \frac{\Delta\theta_0}{sin(\theta_1)} - \frac{\Delta\theta_1 \cdot sin(\theta_0)}{sin^2(\theta_1)} \end{aligned} \tag{10}$$

It is evident from (10) that, for ΔA to be a minimum, irrespective of the polarities of $\Delta\theta_0$ and $\Delta\theta_1$, we require $\theta_0 \cong 0°$, and $\theta_1 \cong 90°$. We note that the condition for maximum signal amplitude ($\theta_0 \cong 90°$, $\theta_1 \cong 0°$) is precisely the opposite of the condition for minimum noise ($\theta_0 \cong 0°$, and $\theta_1 \cong 90°$). Thus, there is a trade-off between the two requirements.

Let us now consider the noise-to-signal ratio in the measurement of A. This ratio, which we denote by NS_A, can be written:

$$NS_A = \frac{\Delta A}{A} = \left[\frac{sin(\theta_1)}{sin(\theta_0)}\right]\left[\frac{\Delta\theta_0 \cdot sin(\theta_1) - \Delta\theta_1 \cdot sin(\theta_0)}{sin^2(\theta_1)}\right] = \frac{\Delta\theta_0}{sin(\theta_0)} - \frac{\Delta\theta_1}{sin(\theta_1)} \quad (11)$$

It is evident from (11) that, for NS_A to be a minimum, irrespective of the polarities of $\Delta\theta_0$ and $\Delta\theta_1$, we require $\theta_0 \cong 90°$, and $\theta_1 \cong 90°$. Thus, the noise-to-signal ratio in the measurement of A is a minimum (or the signal-to-noise ratio is a maximum) when the target is viewed approximately laterally in both video frames.

A similar perturbation analysis of the parameter $B = \frac{tan(\alpha_0)}{tan(\alpha_1)}$ reveals that:

$$\Delta B \cong \frac{\Delta\alpha_0 \cdot tan(\alpha_1) - \Delta\alpha_1 \cdot tan(\alpha_0)}{tan^2(\alpha_1)} = \frac{\Delta\alpha_0}{tan(\alpha_1)} - \frac{\Delta\alpha_1 \cdot tan(\alpha_0)}{tan^2(\alpha_1)} \quad (12)$$

For ΔB to be a minimum irrespective of the polarities of $\Delta\alpha_0$ and $\Delta\alpha_1$, we require $\alpha_0 \cong 0°$, and $\alpha_1 \cong 90°$.

It can be shown that the noise-to-signal ratio of the measurement of B, denoted by NS_B, is given by:

$$NS_B = \frac{\Delta\alpha_0}{tan(\alpha_0)} - \frac{\Delta\alpha_1}{tan(\alpha_1)} \quad (13)$$

For NS_B to be a minimum irrespective of the polarities of $\Delta\alpha_0$ and $\Delta\alpha_1$, we require both α_0 and α_1 to be large. Thus, the noise-to-signal ratio in the measurement of B is a minimum (or the signal-to-noise ratio is a maximum) when the target produces a large image in both video frames. This is intuitively reasonable, as the percentage errors in measuring the size of the target image will be small if the image is large. In summary, the signal-to-noise ratio of the measurement of the disparity δ will be highest when the target is viewed laterally, and its image is large in both frames.

3 TCM Implementation

The TCM approach can be implemented utilising any sensors that provide the required information. For example, on a terrestrial robot, the distance to an object can be determined with a LiDAR sensor and the translation vector could be estimated with wheel encoder. These sensors, such as LiDAR, although common on ground robots, cannot be easily implemented on UAVs. Therefore, in this paper a vision-based approached is implemented.

3.1 Object Detection

To compute the expansion of the object (2), the object is detected using a colour detector. Although there are a number of methods that could be employed to detect the object before classifying its motion, it was decided that colour would be used for simplicity. Colour information allows the detection of both a stationary and moving object without a priori knowledge of its shape or size. The technique utilised in this paper is based on seeded region growing, as described in [Strydom *et al.*, 2015]. The method in [Strydom *et al.*, 2015] employs an additional refinement that is used for a ground target, which is not implemented for this study. This method was chosen over basic colour channel separation methods in order to increase robustness to lighting changes and colour gradients caused by specular reflection.

3.2 Egomotion Computation

To compute the translation direction used by the TCM algorithm, an optic flow and stereo technique, as described in [Thurrowgood *et al.*, 2014; Strydom *et al.*, 2014], is implemented. The pattern of flow vectors is computed using an iterative 400-point pyramidal block-matching algorithm on the panoramic image, as discussed in Section 4. The motion of the aircraft is then computed from the optic flow by using an algorithm that characterises the 3D motion of the aircraft in terms of a 3D translational vector (T) and a 3D rotational vector (R). To compute the translation and rotation, the normal vector to the ground plane is required, which in this case is provided by an Inertial Measurement Unit (IMU).

4 Vision System and Flight Platform

The custom built quadrotor used in this study (see Figure 2) incorporates a MicroKopter flight controller. MicroKopter also supplies the motors and speed controllers.

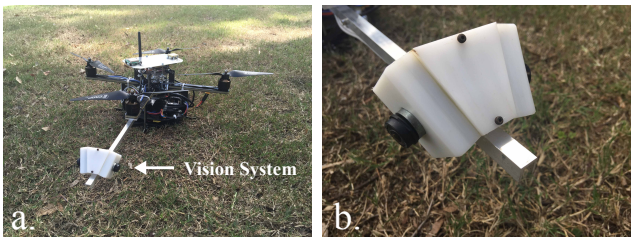

Figure 2: Flight platform (a) and vision system (b).

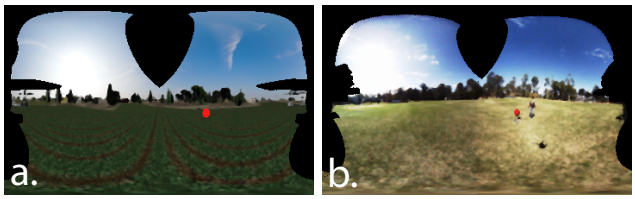

Figure 3: An example of an input image to TCM as perceived by the vision system in the Open Scene Graph simulator (a) and real flight (b). Note that the images are blurred to decrease noise.

An Intel NUC – featuring an Intel i5 core 2.6GHz dual core processor, 8GB of RAM and a 120GB SSD – is used to process the data obtained by the on-board sensors and the vision system in real time. Additional sensors include: a MicroStrain 3DM-GX3-25 IMU, and a vision system, as illustrated in Figure 2(b).

The vision system consists of two fisheye cameras (Point Grey Firefly MV cameras, Sunex DSL216 lenses). The cameras have a 25Hz frame rate, are software synchronised, and are calibrated as described in [Kannala and Brandt, 2006]. The calibration is then used to create a stitched panoramic image with a 360° by 150° field of view (360 by 220 pixel resolution). The large field of view (FOV) sensing is useful for (a) detecting moving objects anywhere in the environment and (b) accurate estimation of egomotion, which is an important intermediate step in the TCM algorithm, as we shall see below.

5 Method Validation

This section describes the optimisations to TCM, such as enhancing the signal-to-noise ratio for the disparity computed by TCM and determining a robust disparity threshold. This section will also verify, through simulated experiments, that TCM is able to determine whether an object is moving or stationary. This validation process was tested in a simulated Open Scene Graph (OSG) environment before real flight tests were performed. The OSG environment simulates the quadrotor motion and the environment as perceived by the quadrotor's vision system, which includes the moving object – in this case a red sphere. Figure 3 illustrates the similarities between the flight simulator and the real vision system.

5.1 Optimising classification

To determine an optimal threshold, a target was simulated in the OSG environment. The OSG environment was also useful to verify the sensitivity of TCM (theory described in Section 2.1), that is, how various object distances, trajectories and velocities affect the TCM disparity. As the disparity depends on the object's distance, velocity and trajectory, it is more appropriate to compare the disparity with the angular velocity of the object's motion. Figure 4 demonstrates that as the angular velocity of the object increases, so does the TCM disparity value. The threshold selected for classifying whether an object is moving or stationary is a compromise between the noise in the measurements and the minimum angular velocity. As well as the angular velocity, the frame interval (FI) between the frames used to compute the disparity affects the sensitivity to motion, as seen in Figure 4. As the frame interval increases, so does the sensitivity of the TCM algorithm.

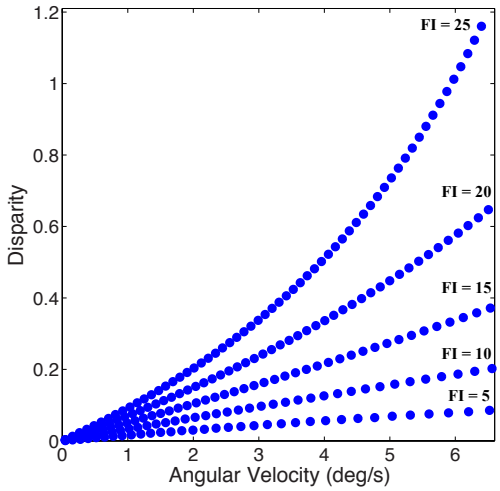

Figure 4: Disparity computed by TCM for various angular velocities and Frame Intervals (FI) as perceived by the vision system.

To improve the reliability of movement detection, the TCM algorithm is computed from two frames that are separated by 25 frames (a FI of 25). This increases the signal-to-noise ratios of the various angle measurements and produces more reliable results, provided the direction vectors and egomotion estimate remain reliable over this larger inter-frame interval (which turns out to be the case in our experiments). A frame interval greater than 25 frames, however, was found to amplify the noise in the TCM inputs. Once the optimal frame interval was determined, the disparity threshold could be optimised.

As shown in Figure 5, the mean noise in the disparity measurements produced by a static target in the OSG simulator was approximately 0.02 units. This noise is largely due to errors in the egomotion computation and the accuracy of the object detection. As the noise observed for a static object in real experiments increases significantly compared to the OSG simulation (see Figure 8), a more conservative disparity threshold of 0.09 was chosen through simulation and experimentation.

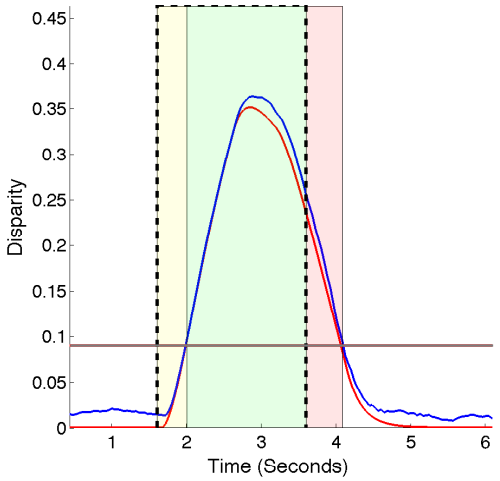

Figure 5: Disparity (TCM: blue line and ground truth: red line) produced in the simulator by an object that is initially static, then moves, which is indicated by a black dashed line, and finally comes to rest. The quadrotor trajectory in the simulator is perpendicular to the target trajectory.

5.2 Verification Results

The first scenario tested in the OSG environment was to move the target perpendicular to the quadrotor trajectory. The target was initially static at position (0.5, 3, 2)m, then moved at a velocity of $1m \cdot s^{-1}$ perpendicular to the quadrotor's trajectory to the position (3.5, 3, 2)m (where it was static once again), while the quadrotor translated from (0, 0, 3)m to (0, 3, 3)m at $0.5m \cdot s^{-1}$. The results in Figure 5 illustrate the ability of TCM to detect the motion of the object in simulation.

To demonstrate how various trajectories can influence the angular velocity of the object and thus, the disparity, the object was moved in a circle, as illustrated in Figure 6(a). As the angular velocity is maximum at $\theta = 90°$ or $270°$, the maximum disparity should occur at $\theta = 90°$ or $270°$ (See also Section 2.1 on sensitivity and noise analysis). It is possible to see a distinct peak value for disparity in Figure 6(b), however it occurs at approximately 330°. This delay is due to the 25 frame interval that is used to increase the signal-to-noise ratio (discussed in Section 5.1).

6 Field Test Results

To demonstrate the robustness of TCM in an outdoor environment, three outdoor field tests were performed: (1) a static target at position T_a (in Figure 7), (2) a moving object test where the object is initially static and comes to rest after moving from T_a to T_b, and finally (3) an oscillating target test. Table 1 summarises the results of tests in the three scenarios, showing the number of classified and misclassified frames for each of the flight tests.

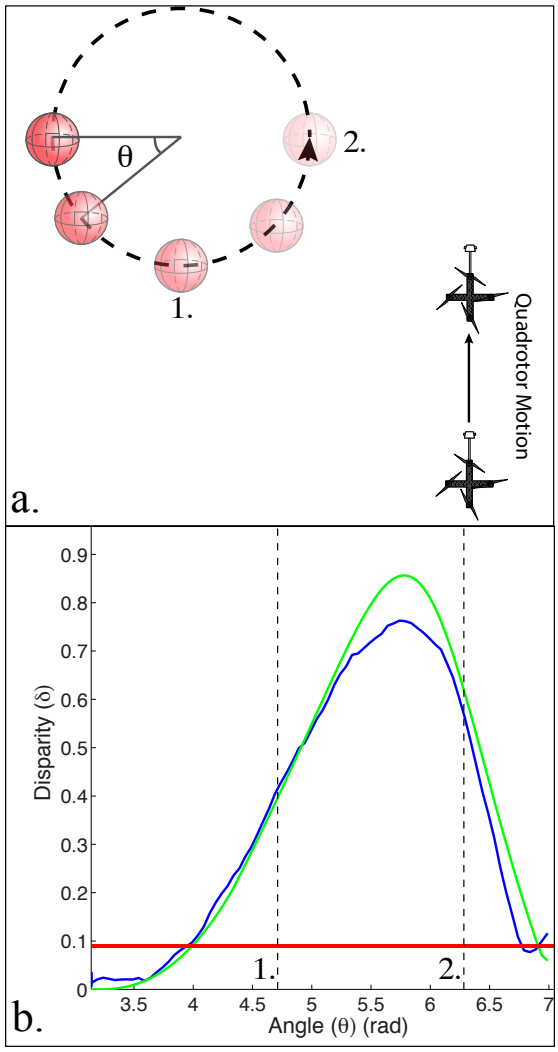

Figure 6: Disparity produced by an object moving along a circular trajectory. (a) shows the ball and quadrotor trajectories and (b) illustrates the results for the disparity produced by the moving object.

Figure 7: Motion of quadrotor and target throughout the various flight tests.

Table 1: Flight statistics for motion classification.

Flight test	# Classified	# Misclassified	Accuracy (%)
1	784	0	100.00
2	317	27	91.48
3	360	52	85.56
Average	487	39.5	91.89

It also details the accuracy of TCM for each test. The average accuracy was deemed to be 91.89% (a detailed performance evaluation is discussed in Section 7.1).

6.1 Scenario 1: stationary target

The first scenario was an elevated stationary target. The target position was (0.5,3,1)m (T_a) and the quadrotor started at a position of (0,0,2)m (Q_a) and moved to (0,3,2)m (Q_b). The purpose of this test was to demonstrate that TCM does not misclassify the object to be moving when in fact it is static – even if the object is at a different distance to the background, which is a common issue for methods that use motion contrast.

As seen in Figure 8, TCM correctly classified that the object is stationary for the entire flight, which consisted of 784 frames (approximately 31 seconds), that is, in this test there were zero false positives (FP). This test also provides a verification of a real outdoor test, using the threshold chosen from simulation.

6.2 Scenario 2: stationary-moving-stationary target

In Scenario 2, the object was static at a position of (0.5,3,1)m (T_a) for approximately 4.5 seconds, then moved 2m perpendicular to the quadrotor's translation (T_a to T_b), and finally came to rest at (3.5,3,1)m (T_b) for a further 4.5 seconds. The quadrotor started at a position of (0,0,2)m (Q_a) and moved to (0,3,2)m (Q_b).

This test provides an insight into the time to detect the change in the object's motion status (e.g. stationary to moving or visa versa), as shown in Figure 9.

6.3 Scenario 3: oscillating target

The third scenario is similar to scenario 2. However, instead of only moving the ball from T_a to T_b, the ball is also moved back from T_b to T_a. This test was again used to demonstrate that TCM correctly detects when the object is moved but it also demonstrates that the object can be detected at varying velocities and distances.

The peak disparity, as shown in Figure 10, for the first movement from T_a to T_b is smaller than the second movement from T_b to T_a. This is because the object is closer to the quadrotor and is moving 25% faster ($1m{\cdot}s^{-1}$ instead of $0.8m \cdot s^{-1}$) in the second movement. This increased the angular velocity, which in turn increased the disparity signal.

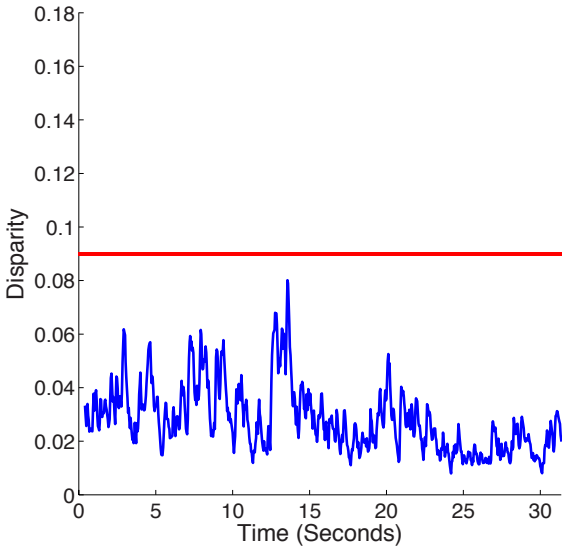

Figure 8: Scenario 1: disparity produced by a static object.

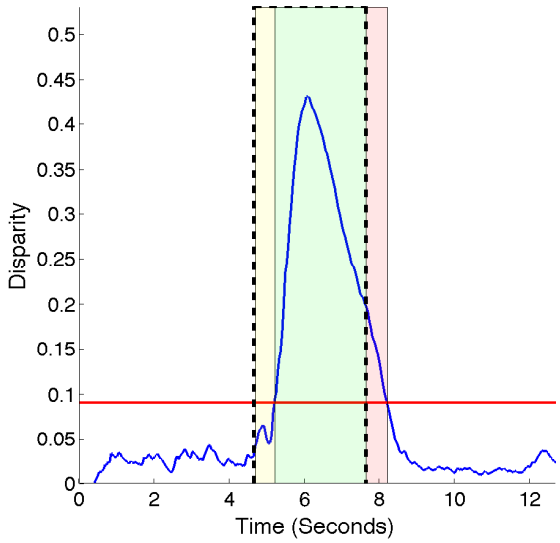

Figure 9: Scenario 2: disparity (blue line) produced by an object that is initially static at T_a, then moves from T_a to T_b and finally remains at T_b. The quadrotor moves from Q_a to Q_b. The red line represents the disparity threshold for classification and the black dashed line signifies when the object was truly in motion. The yellow shading depicts results that produce false negatives, the green true positives, the red false positives and no shading represents true negatives.

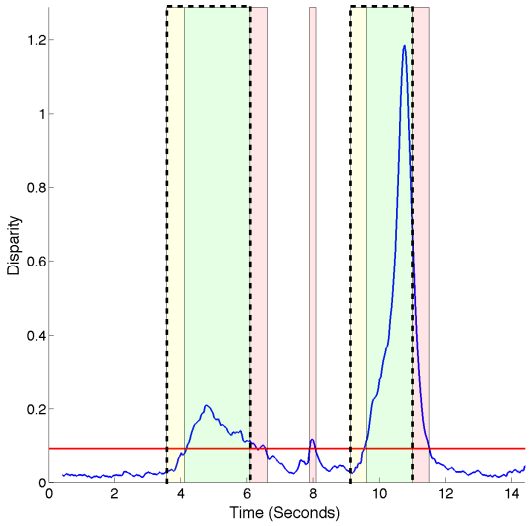

Figure 10: Scenario 3: disparity (blue line) produced by an object that is initially static at T_a, then moves from T_a to T_b and back again, and finally remains at T_a. The quadrotor moves from Q_a to Q_b. The red line represents the disparity threshold for classification and the black dashed line signifies when the object was truly in motion. The yellow shading depicts results that produce false negatives, the green true positives, the red false positives and no shading represents true negatives.

7 Discussion

To distinguish between moving and stationary objects, we performed experiments in a controlled Open Scene Graph (OSG) simulator, as well as outdoor flight tests. In both instances it was shown that the Triangle Closure Method (TCM) was able to detect the change in motion of an object, whether it is from moving to static or vice versa. The major contribution of this paper is that TCM provides a robust method for classifying the motion of an object and is unaffected by the presence of motion parallax, as demonstrated in the static scenario.

7.1 Performance Evaluation

To provide a quantitative measure of the performance of TCM, the number of True Positives (TP), False Positives (FP), True Negatives (TN) and False Negatives (FN) were computed for each flight. The on-board video footage was used to manually determine when the ball was in motion. As the camera system runs at 25Hz it is possible, in theory, to determine the object's motion with a temporal resolution of 40ms. Another metric that is crucial for real time applications is the computation time of the algorithm.

As seen in Figure 8, TCM correctly classified that the object was stationary for the entirety of the first scenario, which consisted of 784 frames (approximately 31 seconds). In this test there were no false positives (FP). This is largely due to the conservative disparity threshold of 0.09.

The accuracy of scenario 2 was computed to be 91.48%, where 317 frames were correctly classified out of a total of 344. Of the misclassified frames, 13 were FN and 14 were FP. These frames were misclassified due to the 25-frame interval; the object must be moving in more than half of the 25 frames.

As the object was moved twice in scenario 3, the number of misclassified frames roughly doubled due to the 25-frame interval but the total number of frames was only slightly more than in scenario 2. Therefore, there is a reduction in accuracy from 91.48% to 85.56%. Although the total accuracy of this test was less than that of the second scenario, TCM still correctly detected the object's motion in both instances.

The computation time for TCM was established on the NUC that processed the data on the flight platform. The NUC has an Intel i5 2.6GHz processor with 8GB of RAM. The vision system on board has a frame rate of 25Hz (40ms), thus all processing must be completed within this time frame. The image pre-processing and egomotion computation required 13ms and the object detection and the TCM method ran in a mere 6ms. Thus, the entire system runs in less than 20ms, which is within the 40ms window.

7.2 Applications

A moving platform with the capacity to classify objects as moving or stationary can significantly improve the safety of robots working in urban environments – in particular UAVs that fly at low altitudes. As the number of UAVs increase, awareness of surrounding objects is paramount for executing maneuvers that are appropriate, timely and effective. TCM would also be useful in interception tasks, in particular where the goal is to intercept only moving objects. It is a versatile method that could be implemented on both ground and air-based robots.

7.3 Limitations and Future Work

One limitation that would compromise the ability of TCM is that it is unable to detect a moving object when the object moves along a trajectory parallel to that of the UAV. This will generate the same disparity (zero) as a stationary object at a greater distance (as illustrated in Figure 11(a)) and will therefore be deemed to be stationary: $d_1/d_2 = D_1/D_2$. On the other hand, an object that moves in a direction that is non-parallel to the quadrotor trajectory (Figure 11(b)) will always generate a non-zero disparity ($d_1/d_2 \neq D_1/D_2$) and will thus be classified as moving. In the special case of parallel motion, the object moves in such a way as to conceal its motion ('motion camouflage', as described by [Srinivasan and Davey, 1995]). This type of motion, to the

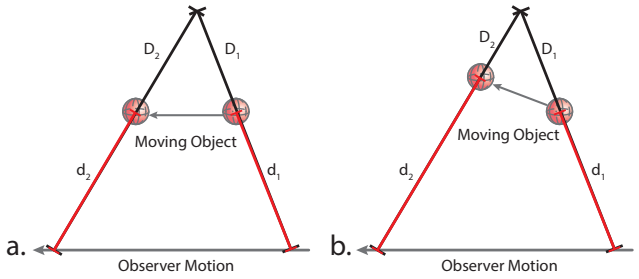

Figure 11: Object moving parallel (a) and non-parallel (b) to aircraft.

best of our knowledge, is undetectable by any monocular vision-based algorithm, including ours. However, such an object is not a potential threat, as it would always be at a safe distance away from the UAV.

In this study we have assumed that the object is spherical. This ensures that the object is always viewed as a circle, which simplifies the expansion/contraction computation. An object that is shaped very differently from a sphere would increase the noise in the TCM disparity signal, as the area of the shape in the image would then vary depending upon the viewing direction. Thus, the image size would then not be directly related to object distance. This problem could be solved by tracking the distance between two distinct features on the object, or by evaluating the change in image size along an axis perpendicular to the UAVs translation direction.

8 Conclusions

An efficient strategy, coined TCM, is proposed to distinguish whether an object is moving or stationary. We have demonstrated that using a vision-based approach, the TCM method can accurately and robustly classify if an object is moving or static even if it is at a different distance to the background. The method was validated through the use of a virtual environment (Open Scene Graph (OSG)). The OSG environment was used to demonstrate the ability of TCM when the target is constantly varying its trajectory. Finally, three scenarios were conducted in outdoor field tests, which provide quantitative measures of the performance of the Triangle Closure Method (TCM). The overall accuracy was measured to be 91.89%.

9 Acknowledgments

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 Premier's Fellowship, and an ARC Distinguished Outstanding Researcher Award (DP140100914). We thank Dean Soccol for assistance with the mechanical aspects of this work and Aymeric Denuelle for his assistance with the colour detection algorithm.

References

[Choi *et al.*, 2011] Jay Hyuk Choi, Dongjin Lee, and Hyochoong Bang. Tracking an unknown moving target from uav: Extracting and localizing an moving target with vision sensor based on optical flow. In *Automation, Robotics and Applications (ICARA), 2011 5th International Conference on*, pages 384–389. IEEE, 2011.

[Clothier *et al.*, 2012] Reece Clothier, Dennis Frousheger, Michael Wilson, and I Grant. The smart skies project: enabling technologies for future airspace environments. In *28th Congress of the International Council of the Aeronautical Sciences*. ICAS, 2012.

[Denuelle *et al.*, 2015] Aymeric Denuelle, Saul Thurrowgood, Reuben Strydom, Farid Kendoul, and Mandyam V Srinivasan. *Biologically-inspired visual stabilization of a rotorcraft UAV in unknown outdoor environments*. IEEE, 2015.

[Elhabian *et al.*, 2008] Shireen Y Elhabian, Khaled M El-Sayed, and Sumaya H Ahmed. Moving object detection in spatial domain using background removal techniques-state-of-art. *Recent patents on computer science*, 1(1):32–54 2008.

[Green *et al.*, 2004] William E Green, Paul Y Oh, and Geoffrey Barrows. *Flying insect inspired vision for autonomous aerial robot maneuvers in near-earth environments*, volume 3. IEEE, 2004.

[Hayman and Eklundh, 2003] Eric Hayman and Jan-Olof Eklundh. *Statistical background subtraction for a mobile observer*. IEEE, 2003.

[Ibrahim *et al.*, 2010] Aryo Wiman Nur Ibrahim, Pang Wee Ching, GL Gerald Seet, WS Michael Lau, and Witold Czajewski. *Moving objects detection and tracking framework for UAV-based surveillance*. IEEE, 2010.

[Kannala and Brandt, 2006] Juho Kannala and Sami S Brandt. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(8):1335–1340

[Kim *et al.*, 2010] In Su Kim, Hong Seok Choi, Kwang Moo Yi, Jin Young Choi, and Seong G Kong. Intelligent visual surveillance—a survey. *International Journal of Control, Automation and Systems*, 8(5):926–939

[Korn and Edinger, 2008] Bernd Korn and Christiane Edinger. *UAS in civil airspace: Demonstrating "sense and avoid" capabilities in flight trials*. IEEE, 2008.

[Lai *et al.*, 2013] 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

[Li *et al.*, 2013] Jing Li, Hongchun Yuan, and Ai Hong. *Moving Vehicle Detection in Dynamic Background from Airborne Monocular Camera*, volume 1. IEEE, 2013.

[Nussberger *et al.*, 2014] Andreas Nussberger, Helmut Grabner, and Luc Van Gool. Aerial object tracking from an airborne platform. In *Unmanned Aircraft Systems (ICUAS), 2014 International Conference on*, pages 1284–1293. IEEE, 2014.

[Paragios and Deriche, 2000] Nikos Paragios and Rachid Deriche. Geodesic active contours and level sets for the detection and tracking of moving objects. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(3):266–280

[Pinto *et al.*, 2014] Andry Maykol Pinto, A Paulo Moreira, Miguel V Correia, and Paulo G Costa. A flow-based motion perception technique for an autonomous robot system. *Journal of Intelligent Robotic Systems*, 75(3-4):475–492 0921–0296, 2014.

[Pinto *et al.*, 2015] Andry Maykol Pinto, Paulo Gomes Costa, and António Paulo Moreira. *Introduction to Visual Motion Analysis for Mobile Robots*. Springer, 2015.

[Sheikh *et al.*, 2009] Yaser Sheikh, Omar Javed, and Takeo Kanade. *Background subtraction for freely moving cameras*. IEEE, 2009.

[Spagnolo *et al.*, 2006] Paolo Spagnolo, M Leo, and A Distante. Moving object segmentation by background subtraction and temporal analysis. *Image and Vision Computing*, 24(5):411–423

[Srinivasan and Davey, 1995] Mandyam V Srinivasan and Matthew Davey. Strategies for active camouflage of motion. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 259(1354):19–25

[Strydom *et al.*, 2014] Reuben Strydom, Saul Thurrowgood, and Mandyam V Srinivasan. Visual odometry: Autonomous uav navigation using optic flow and stereo. In *Australasian Conference on Robotics and Automation (ACRA)*, pages 1–10. Australian Robotics and Automation Association, 2014.

[Strydom *et al.*, 2015] Reuben Strydom, Saul Thurrowgood, Aymeric Denuelle, and Mandyam V Srinivasan. *UAV Guidance: A Stereo-Based Technique for Interception of Stationary or Moving Targets*, pages 258–269 Springer, 2015.

[Sundaram *et al.*, 2010] Narayanan Sundaram, Thomas Brox, and Kurt Keutzer. *Dense point trajectories by GPU-accelerated large displacement optical flow*, pages 438–451 Springer, 2010.

[Thakoor *et al.*, 2004] Ninad Thakoor, Jean Gao, and Huamei Chen. *Automatic object detection in video sequences with camera in motion*. Citeseer, 2004.

[Thurrowgood *et al.*, 2014] Saul Thurrowgood, Richard JD Moore, Dean Soccol, Michael Knight, and Mandyam V Srinivasan. A biologically inspired, vision-based guidance system for automatic landing of a fixed-wing aircraft. *Journal of Field Robotics*, 31(4):699–727, 2014.

[Valavanis, 2008] Kimon P Valavanis. *Advances in unmanned aerial vehicles: state of the art and the road to autonomy*, volume 33 Springer Science Business Media, 2008.

[Viquerat *et al.*, 2008] Andrew Viquerat, Lachlan Blackhall, Alistair Reid, Salah Sukkarieh, and Graham Brooker. *Reactive collision avoidance for unmanned aerial vehicles using doppler radar*. Springer, 2008.

[Wilson, 2012] Michael Wilson. Ground-based sense and avoid support for unmanned aircraft systems. *Congress of the International Council of the Aeronautical Sciences (ICAS)*, 2012.

[Wu and Leahy, 1993] Zhenyu Wu and Richard Leahy. An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 15(11):1101–1113

[Zhu and Yuille, 1996] Song Chun Zhu and Alan Yuille. Region competition: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 18(9):884–900

[Zufferey and Floreano, 2006] Jean-Christophe Zufferey and Dario Floreano. Fly-inspired visual steering of an ultralight indoor aircraft. *Robotics, IEEE Transactions on*, 22(1):137–146 2006.