A 6 DoF Visual Tracking System for a Miniature Helicopter

Author:
Mak, Lin Chi; Furukawa, Tomonari

Event details:
2nd International Conference on Sensing Technology
Palmerston North, New Zealand

Publication Date:
2007

DOI:
https://doi.org/10.26190/unsworks/385

2nd International Conference on Sensing Technology
November 26-28, 2007 Palmerston North, New Zealand

A 6 DoF Visual Tracking System for a Miniature Helicopter

Lin Chi Mak, Tomonari Furukawa
Computational Mechanics and Robotics
School of Mechanical and Manufacturing Engineering
University of New South Wales, Sydney, Australia
{m.chi; t.furukawa}@cas.edu.au

Abstract

This paper presents a low-cost 6 Degrees-of-Freedom (6 DoF) visual tracking technique and its system for a small indoor rotary-wing Unmanned Aerial Vehicle (UAV) using three onboard LEDs and a single on-ground camera. The ellipse formed by the cyan LEDs on the blades and the red LED in the captured image are utilised for pose estimation. The developed system successfully implements the proposed technique with a sampling rate of 9-12Hz. The most notable contributions of this paper are the novel configuration of the LEDs for visual tracking using only three LEDs and one camera for 6 DoF motion tracking of a UAV and the wide range of observable flight. The experimental results show 1-3% percentage error in the developed system at various ranges, directions and vehicle velocities in an indoor environment, demonstrating the effectiveness of the proposed technique.

Keywords: Computer vision, unmanned aerial vehicle, orientation and position determination

1 Introduction

Over recent years, the use of small UAVs has gained significant interest in several research communities due to their wide robotics applications such as Urban Search And Rescues (USAR) and surveillance [1]. Robust pose estimation systems are crucial for UAVs to navigate indoors. The key challenges in such systems are the unavailability of GPS data and the limited sizes and payloads of the UAVs. Due to the high manoeuvrability of rotary-wing UAVs, they have distinct advantages over fixed-wing and flapping-wing UAVs in indoor environments with limited spaces [2], [3].

In the literature, there have been a variety of indoor pose estimation systems developed for aerial vehicles. The most popular techniques are the Inertia Measurement Unit (IMU) and the electromagnetic tracking device. The IMU measures the accelerations and integrates them over time to obtain the orientation and position. It provides low-noise motion measurements with high sampling rate [4], but the pose gradually drifts due to the accumulative errors caused by the integration. The electromagnetic tracking device measures the relative pose of the target very accurately. Castillo *et al.* [2] showed the successful pose estimation and control of a quad-rotor helicopter using an electromagnetic tracking sensor. However, this type of sensor is highly sensitive to electromagnetic noise, especially when the sensor is very close to the electric motors in a small UAV.

Another popular class of the approaches to the pose estimation of UAVs is the vision. Having cameras and landmarks on vehicles or on ground, the poses of the vehicles can be estimated. Existing vision-based pose estimation systems fall into three main categories. The first category is the outside-in, where cameras are on-ground to track onboard landmarks. In this category, Tisse *et al.* [4] utilised three onboard markers and on-ground stereo cameras to obtain the 6 DoF motion, consisting of the yaw, roll, pitch and 3-Dimensional position, of a rotary-wing UAV. Both Altuğ *et al.* [5] and Earl and Andrea [6] successfully controlled a quad-rotor helicopter with 4 and 5 markers/LEDs as onboard landmarks and an external camera under the helicopter. Wang *et al.* [7] presented an outside-in visual tracking system for a coaxial helicopter of no dimension larger than 140mm with an onboard marker. This type of the system is very suitable for small UAVs or Micro Aerial Vehicles (MAVs) with limited payload because only light markers or LEDs are required to be mounted on-board.

The second category is the inside-out, where cameras are onboard while landmarks are on-ground. In this category, Romero *et al.* [8] demonstrated stabilisation and localisation of a quad-rotor helicopter using four on-ground markers, whilst Tournier *et al.* [9] and Xu *et al.* [10] advantageously exploited onboard cameras and special patterns on ground to control and estimate the positions and the poses of aerial vehicles. These techniques provide highly accurate pose data, adequate for several autonomous tasks, such as way-point tracking and landing. Davidson *et al.* [11] presented a vision-based Simultaneous Localisation and Mapping (SLAM) technique which can determine the 3D trajectory of a monocular hand-held camera in an unknown scene by localising new landmarks. The last category, the hybrid system, localises the target by a combination of on-ground and onboard cameras. Altuğ *et al.* [12] improved the effectiveness of their estimation system

by having an additional onboard camera and an on-ground marker.

Although the existing techniques provide accurate pose estimation of flying vehicles, there are still some limitations in these systems. The existing inside-out systems [8]–[10], except 3D vision-based SLAM techniques, require artificial landmarks in the environments, or large ground vehicles to carry large patterns. Also, such systems suffer from the poor quality of received images from the onboard camera due to radio interference and vibrations in the UAVs. In the first and third categories of existing vision-based pose estimation systems, the ranges of observable flights are usually narrow [5], [6], [12], since the artificial landmarks are only visible from limited positions.

This paper presents a pose estimation technique and its low-cost system for a miniature indoor rotary-wing UAV using only three onboard LEDs and an on-ground camera. The circular orbit of the coloured LEDs on the blades forms an ellipse in the captured image which is easily recognised by colour thresholding and segmentation in indoor environments. Despite the shape of the ellipse varying with the orientation of the UAV, the width is only dependent on the distance between the camera and the centre of the orbit. Based on the width and the centre of the ellipse, the position of the UAV can be computed. The pose is estimated by the consideration of the shape of the ellipse and an additional onboard LED. The most significant contribution in this paper is the innovative configuration of the LEDs, which requires only three LEDs and one camera to track the 6 DoF motion of a UAV. The other strength of the proposed technique is that the wide range of observable flight since the orbit is visible in almost all directions.

This paper is organised as follows. The following section presents the visual tracking technique, while the visual tracking system is illustrated in section 3. Then, the efficacy of the developed system is demonstrated through the experimental results in section 4 and conclusions are summarised in the last section.

2 Pose Estimation

This section describes the proposed visual tracking technique for a miniature helicopter. In this paper, a coaxial helicopter is considered as the target, although the proposed technique can be applied to the other types of helicopters. The proposed technique consists of two major executive stages, which are objects segmentation and pose estimation. There are two cyan LEDs pointing outwards at the tips of the top blades, and a red LED at the end of the tail of the target.

Figure 1 shows the target coordinate system and positions of the LEDs on the target. The cyan LEDs are located at $[r\cos\mu, r\sin\mu, 0]_t^T$ and $[-r\cos\mu, -r\sin\mu, 0]_t^T$, $\mu \in (0, 2\pi]$, where r and

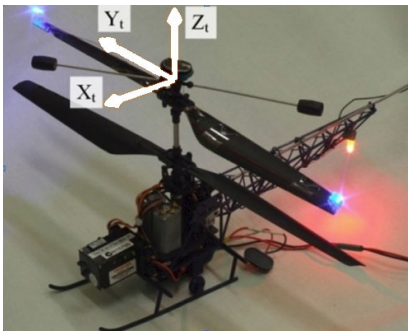


Figure 1: Target coordinate system and positions of the LEDs at the target.

$[\dagger]_t$ represent the radius of the circular orbit and a point, $\dagger$, in the target coordinate system respectively. The red LED is located at $[-x_t^r, 0, -z_t^r]_t^T$, where x_t^r and z_t^r are the distances between the red LED and the origin in X_t- and Z_t- components respectively.

The proposed technique and its system are based on the following assumptions:

Assumption 1: The LEDs on the helicopter are the major red and cyan light sources in the view of the camera.

Assumption 2: The exposure time per frame of the camera is slightly longer than the half period of the rotation of the blades with the cyan LEDs. Therefore, a full cyan ellipse is captured in the image.

Assumption 3: All of the LEDs are always in the view of the camera.

2.1 Object Segmentation

The first step of the object segmentation is colour recognition. Assume that $\mathbf{I}$ designates the entire image with a size of $w \times h$, where w and h represent the number of pixels of the width and the height. Subsequently, the image space is defined as

$$\mathbf{I} = \left\{ \mathbf{j} = [i, j]^T \mid \forall i \in (1, w), \forall j \in (1, h) \right\}.$$

The colour at $\mathbf{j}$ can be denoted by three components in a vector $[r_\mathbf{j}, g_\mathbf{j}, b_\mathbf{j}]^T$. Consequently, the cyan colour space can be represented by

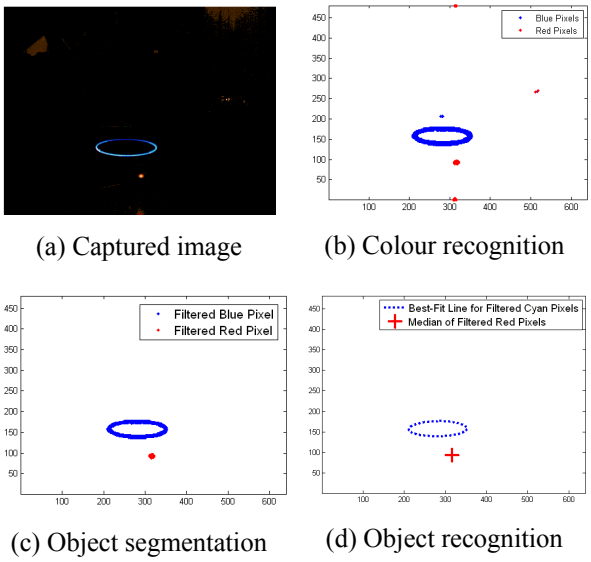
$$\mathbf{I}_c = \left\{ \mathbf{j} \in \mathbf{I} \mid 1.5b_\mathbf{j} + 0.5g_\mathbf{j} - 2r_\mathbf{j} > \tau_c \right\},$$

where τ_c and $\mathbf{I}_c$ represent the cyan threshold value and the set of cyan pixels respectively. Similarly, the red colour space can be designated by $\mathbf{I}_r$ as follows:

$$\mathbf{I}_r = \left\{ \mathbf{j} \in \mathbf{I} \mid 2r_\mathbf{j} - g_\mathbf{j} - b_\mathbf{j} > \tau_r \right\},$$

where τ_r stands for the red threshold.

Figure 2 depicts an example of the process of cyan ellipse and red LED recognition. Figure 2a illustrates the original image with a resolution of 640×480



(a) Captured image (b) Colour recognition

(c) Object segmentation (d) Object recognition

Figure 2: An example of the processes of the cyan ellipse and red LED recognition

pixels. Meanwhile, figure 2b visualizes the cyan and red colour spaces, $\mathbf{I}_c$ and $\mathbf{I}_r$. Usually, there is some noise in the image incorrectly recognized as the red or cyan pixels. Therefore, an object segmentation algorithm is employed to further filter the image.

The second step of the objects segmentation is to eliminate the noise based on the initially guess of the ellipse's position. According to assumption 1, the ellipse formed by the cyan LEDs dominates the cyan pixels in the captured image. Hence, the approximate ellipse's centre, $\hat{\mathbf{j}}_e$, can be obtained as the median centre of the cyan colour space:

$$\hat{\mathbf{j}}_e = \text{median}\{\mathbf{j} \in \mathbf{I}_c\} \quad . \tag{1}$$

All points in $\mathbf{I}_c$ and $\mathbf{I}_r$ which are far away from $\hat{\mathbf{j}}_e$ will be then discarded and two new spaces, $\mathbf{I}_e^0$ and $\mathbf{I}_a$ denoting the segmented ellipse space and the tail LED space respectively are created as

$$\mathbf{I}_e^0 = \left\{ \mathbf{j} \in \mathbf{I}_c \;\middle|\; \left(\left| \hat{\mathbf{j}}_e - \mathbf{j} \right| < \tau_d \right) \right\},$$

$$\mathbf{I}_a = \left\{ \mathbf{j} \in \mathbf{I}_r \;\middle|\; \left(\left| \hat{\mathbf{j}}_e - \mathbf{j} \right| < \tau_d \right) \right\},$$

where τ_d is the distance threshold value which is 1.5 times of the width of the ellipse in the prior captured image. If there is no prior information about the size of the ellipse, τ_d is set to be positively infinite.

Similar to the Saripalli's technique [13], $\mathbf{I}_e^0$ is then partitioned into m subregions $\mathbf{I}_e^1$, $\mathbf{I}_e^2$,..., $\mathbf{I}_e^m$ such that:

- $\mathbf{I}_e^1 \cup \mathbf{I}_e^2 \cup ... \cup \mathbf{I}_e^m = \mathbf{I}_e^0$,
- $\mathbf{I}_e^p$ is a connected region for $\forall p = (1,...,m)$,
- $\mathbf{I}_e^p \cup \mathbf{I}_e^q = \varnothing, \forall p, q = (1,...,m), q \neq p.$

All subregions $\mathbf{I}_e^p$ sized smaller than 10 pixels are discarded. Subsequently, the final segmented ellipse space $\mathbf{I}_e$ is yielded as

$$\mathbf{I}_e = \left\{ \mathbf{j} \in \mathbf{I}_e^p \mid \forall p = (1,...,m), \text{size}(\mathbf{I}_e^p) > 10 \right\} \subset \mathbf{I}_e^0 ,$$

where size(†) stands for the number of the elements in the set †. All pixels in $\mathbf{I}_e$ and $\mathbf{I}_a$ will be used for the pose estimation described in the next subsection.

2.2 Pose Estimation

Having determined the shape and size of the cyan ellipse in the image plane, the pose of the target excepting the yaw can be determined. Figure 3 shows a point, $[x_s, y_s, z_s]_s^T$, in the sensor coordinate system, its projection, $\mathbf{j} = [i, j]^T \in \mathbf{I}$, to the image plane and two vectors, $\mathbf{n}$ and $\mathbf{s}$, which will be exploited in equation (5). Let the image plane be perpendicular to the viewing axis at the focal length, f, from the origin of the sensor coordinate system. Consequently, the projection, $\mathbf{x}_i = [y_i, z_i]_i^T$, of a point at $[x_s, y_s, z_s]_s^T$ to the image plane can be denoted as

$$\mathbf{x}_i = \left[f y_s / x_s, f z_s / x_s \right]_i^T . \tag{2}$$

The relationship between a point in the image plane and the corresponding pixel in the captured image can be expressed as:

$$\mathbf{x}_i = G(\mathbf{j}) = \left[f \tan\left(\frac{i - w/2}{w} \varepsilon \right), f \tan\left(\frac{j - h/2}{w} \varepsilon \right) \right]_i^T , \tag{3}$$

where ε stands for the horizontal view angle of the camera. The ellipse in the image plane can be expressed as in [14] by

$$\mathrm{H}(\mathbf{x}_i) = a y_i^2 + 2 b y_i z_i + c z_i^2 + 2 f d y_i + 2 f e z_i + f^2 g = 0 , \tag{4}$$

where the coefficients, a, b, c, d, e and g are represented with r and two vectors in the sensor coordinate system. These are $\mathbf{s} = [s_1, s_2, s_3]_s^T$, the vector from the centre of the circle formed by the cyan LEDs to the camera, and $\mathbf{n} = [n_1, n_2, n_3]_s^T$, the normal unit vector of the surface, on which the circle lies. The relationship between the above coefficients, a, b, c, d, e, g, r, $\mathbf{s}$ and $\mathbf{n}$, is as follows:

$$\begin{aligned}
a &= s_3^2(1-n_1^2) + s_1^2(1-n_3^2) + 2 s_1 s_3 n_1 n_3 - r^2 n_2^2 , \\
b &= s_1 s_3 n_1 n_2 + s_1 s_2 n_1 n_3 + r^2 n_2 n_3 - s_1^2 n_2 n_3 - s_2 s_3 (1-n_1^2) , \\
c &= s_1^2(1-n_2^2) + s_2^2(1-n_1^2) + 2 s_1 s_2 n_2 n_3 - r^2 n_3^2 , \\
d &= s_2 s_3 n_1 n_3 + s_1 s_3 n_2 n_3 + r^2 n_1 n_2 - s_3^2 n_1 n_2 - s_1 s_2 (1-n_3^2) , \\
e &= s_2^2 n_1 n_3 + s_1 s_3 (1-n_2^2) - s_1 s_2 n_2 n_3 - s_2 s_3 n_1 n_2 - r^2 n_1 n_3 , \\
g &= s_2^2(1-n_3^2) + s_3^2(1-n_2^2) + 2 s_2 s_3 n_2 n_3 - r^2 n_1^2 .
\end{aligned} \tag{5}$$

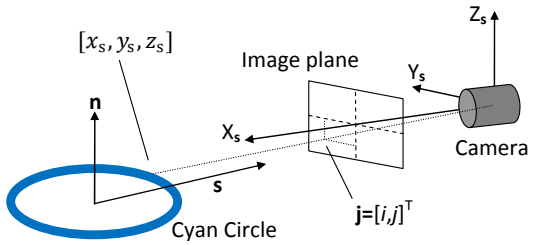


Figure 3: Image geometry and sensor coordinate system

The best-fit ellipse in the image plane can be determined via the optimization of the following objective function:

$$\sum_{p=1}^{\text{size}(\mathbf{I}_e)} \left| \mathrm{H}(\mathrm{G}(\mathbf{j}_p^e)) \right| \longrightarrow \min_{\mathbf{s},\mathbf{n}}, \tag{6}$$

where $\mathbf{j}_p^e \in \mathbf{I}_e$ stands for the p^{th} pixel in the final segmented ellipse space. In the optimization process, only one parameter is adjusted in each step until the error of the objective function converges to a small value, so the process is not computationally expensive. The pose of the target except the yaw in the sensor coordinate system can be estimated from the vectors, **s** and **n**, as follows:

$$[x_s^t, y_s^t, z_s^t]_s^T = -\mathbf{s} = [-s_1, -s_2, -s_3]_s^T, \tag{7}$$

$$[\alpha_s^t, \beta_s^t]_s^T = \left[\tan^{-1}(n_2 / n_3), \tan^{-1}(-n_1 / n_3) \right]_s^T, \tag{8}$$

where x_s^t, y_s^t and z_s^t are the x-, y- and z-coordinates of the origin of the target coordinate system in the sensor coordinate system, while α_s^t and β_s^t are the pitch and roll angle of the target.

The yaw of the target is estimated based on the red LED through four steps. Firstly, the centre of the red LED, $\hat{\mathbf{j}}_a$ in the captured image is determined contingent on all pixels in $\mathbf{I}_a$ as follows:

$$\hat{\mathbf{j}}_a = \text{median}\{\mathbf{j} \in \mathbf{I}_a\}. \tag{9}$$

Secondly, the projection $\hat{\mathbf{j}}_a'$ of $\hat{\mathbf{j}}_a$ to the plane on which the cyan ellipse lies can be computed as

$$\hat{\mathbf{j}}_a' = \hat{\mathbf{j}}_a + \begin{bmatrix} n_1 \\ n_3 \end{bmatrix} \cdot \left(\frac{w}{\varepsilon \sqrt{n_1^2 + n_3^2}} \tan^{-1}\left(\frac{z_t^r}{|\mathbf{s}|} \right) \right) \cdot \cos\left(\tan^{-1}(z_s^t / x_s^t) + \tan^{-1}(n_3 / n_1) \right). \tag{10}$$

Thirdly, $\hat{\mathbf{j}}_a'$ is transformed to $\hat{\mathbf{j}}_a^*$ in another coordinate system with the x-axis parallel to the main axis of the best-fit ellipse and the origin at the centre of the ellipse in the captured image as follows:

$$\hat{\mathbf{j}}_a^* = [\hat{i}_a^*, \hat{j}_a^*]^T = \frac{1}{\sqrt{n_1^2 + n_3^2}} \begin{pmatrix} n_3 & -n_1 \\ n_1 & n_3 \end{pmatrix} \left(\hat{\mathbf{j}}_a' - \tilde{\mathbf{j}}_e \right), \tag{11}$$

where $\tilde{\mathbf{j}}_e$ is the centre of the estimated best-fit ellipse in the captured image. Lastly, the yaw angle of the target in the sensor coordinate system γ_s can be expressed as:

$$\gamma_s = \begin{cases} \sin^{-1}\left(\hat{i}_a^* / \left((w/\varepsilon) \tan^{-1}(x_t^r / x_s^t) \right) \right), \text{for } h \le 0, \\ \pi - \sin^{-1}\left(\hat{i}_a^* / \left((w/\varepsilon) \tan^{-1}(x_t^r / x_s^t) \right) \right), \text{otherwise}, \end{cases} \tag{12}$$

where $h = \left([s_1, s_3] \cdot [n_1, n_3]^T \right) \hat{j}_a^*$. If $h > 0$, the target is pointed towards the camera.

3 Visual Tracking System

This section lists the properties of the developed visual tracking system. Figure 4 describes the developed system and the key parameters of each component in the system. The system consists of a USB webcam with a horizontal Field Of View (FOV), ε, of 60.8º±5%, a linear polarizing filter and a standard PC with 2 × 2.13GHz CPU and 2GB RAM. The webcam was chosen due to its low cost of AU$110 and large FOV. Noticeably, the total cost of the system, excluding the PC, is AU$150. The video of the ellipse recognition in the developed system is available at http://cmr.mech.unsw.edu.au/?q=node/27.

For the validity of assumption 2, the exposure time of the camera is constrained and hence a polarizing filter is used to reduce the average light level in the captured image. Since the gain of the selected webcam is not adjustable, a filter with a filter factor of 2 is located in front of the camera to halve the light level to prevent the captured image from over-exposing.

In the software, there are three techniques to enhance the accuracy and the sampling rate of the developed system. In order to maintain the sampling rate when the number of pixels in $\mathbf{I}_e$ is more than 200, only 200 points are randomly selected and considered for the estimation of the best-fit ellipse. With the current configuration of the PC, the average sampling rate is 9-12Hz. If no cyan pixel is detected, the system will notice that the helicopter is out of FOV and wait until it is in the FOV. To improve the accuracy of the developed system, any estimated pose will be discarded if it is notably different from the previous.

4 Experimental Results

This section discusses the accuracy of the developed system under different ranges, target speeds and sensor positions. Figure 5 shows the testing platform and the setup of the experiment. The Y_t-position of the target was controlled by a 0.8m-long motor-driven slider and monitored by an infrared (IR) range finder located, which gave 1cm positioning accuracy, at the end of the slider. The Y_s-axis was parallel to the Y_t-axis, whilst the angle between X_s-axis and X_t-axis was the true roll, β_s^0. The range between the origins of the

camera and the target coordinate systems was x_s. The experiment was undertaken in an indoor environment with fluorescent lighting.

There were three cases with different arrangements in the experiment. In case 1, the target was stationary at $[x_s, 0, 0]_s^T$ and $\beta_s^0 = 30^o$, where $1.5\text{m} \le x_s \le 7.6\text{m}$. In case 2, the target was moving at $[2\text{m}, y_s, 0]_s^T$, $-0.4\text{m} \le y_s \le 0.4\text{m}$, and $\beta_s^0 = 30^o$ at varying speeds between 0 and 0.42ms^{-1} along the Y_t-axis. In case 3, the target was moving at the average speed of 2.2ms^{-1} along the Y_t-axis at $[2\text{m}, y_s, 0]_s^T$ and β_s^0 varied between 0^o and 76^o, where $-0.4\text{m} \le y_s \le 0.4\text{m}$.

Figures 6a and 6b depict the root mean square (RMS) percentage errors in the position and orientation of the target at different ranges, x_s, in case 1. The percentage errors of the position and the orientation are obtained by the following equations:

$$\text{E}(*) = \left(\varepsilon(*) / x_s\right) \times 100\% \ , \tag{13}$$

$$\text{E}(\angle) = \left(\varepsilon(\angle) / 2\pi\right) \times 100\% \ , \tag{14}$$

where $\text{E}(\dagger)$ and $\varepsilon(\dagger)$ denote the percentage error and the RMS error of $\dagger$ respectively. Moreover, $*$ and $\angle$ stand for the x-, y- or z-position and the yaw, roll, or pitch of the target respectively in the sensor coordinate system.

According to the experimental results in case 1, similar to other single camera visual tracking systems, the percentage errors of both the target position and orientation rise with the range. This is because the size of the ellipse in the captured image reduces with the rise in the range. The x-positioning error is below 3% and both of the y- and z-positioning errors are under 1.2% when the range is shorter than 7m. The x-position error is larger than the others due to difficulties in estimating the ellipse width. Besides, the yaw error increases exponentially with the range up to 3% at 7m. Hence, long range visual tracking should be avoided when accurate yaw estimation is required. Within 7m range, the errors of the pitch and roll are smaller than 0.6%, indicating that the developed system is good at estimating them.

Figure 7 illustrates the RMS errors in case 2 at different target speeds within 0.42ms^{-1} along Y_t-axis, where 'ori.' represents the total orientation error, which is the root of the total sum of the squared yaw, pitch and roll. The accuracies of the pitch and the roll are not substantially affected by the target speed, but the yaw error increases from ±2º to ±5º with the rise in the target speed. In addition, the z-positioning error remains unchanged and the x-positioning error rises slightly from 0.17m to 0.26m with the increase in the target speed from 0 to 0.42ms^{-1}. Meanwhile, the RMS error of the y-position increases dramatically from 4mm to 50mm. Notably, the y-positioning error is highly dependent on the target speed in y-direction because of the measurement delay in the developed system, mainly due to the computational time required for image processing and image acquisition.

Figure 8 depicts the RMS errors in case 3 at varying true rolls, β_s^0, between 0^o and 76^o. The positioning errors in x-, y- and z-directions rise slightly by less than 0.01m with the increase in the true roll, pointing out that the positioning accuracies are not sensitive to the orientation of the target relative to the camera. Noticeably, the pitch error rises significantly from 1º to 7º with β_s^0 because the ellipse with large β_s^0 looks like a circle, making it very difficult to determine the pitch. Nonetheless, the decrease in the roll error with the increase in β_s^0 compensates the rise in the pitch error. The total orientation error of the target thus increases slightly from 6.5º to 8.3º. In conclusion, the performance of the developed system is satisfactory in most directions of the camera from the target.

5 Conclusion and Future Works

A novel visual tracking technique has been proposed and tested. The key advantage of the proposed technique is the wide range of observable flight. The experimental results first show the 3D positioning and the total orientation measurement errors of the static target at 2m away from the camera are 0.02m and 4º respectively. Meanwhile, the results also show positioning error below 0.05m and good orientation error under 8º in almost any helicopter's orientation, demonstrating the effectiveness of the proposed technique and the corresponding system. However, the yaw errors increase exponentially with the range, indicating the yaw estimation is not robust when the range is large.

In the future, the yaw accuracy of the developed system can be improved by having an additional on-board camera to determine the orientation using landmarks as in a hybrid system. The limited FOV of the camera can be solved by the exploitation of an active vision system. Moreover, this work can be extended to the autonomous control of a miniature helicopter in an indoor environment.

6 Acknowledgements

This work is partly supported by the US Department of Air Force (AOARD-0504026) and ARC Centre of Excellence programme, funded by the Australian Research Council (ARC) and the New South Wales State Government.

7 References

[1] S. Todorovic and M. C. Nechyba, "A Vision System for Intelligent Mission Profiles of Micro Air Vehicles", *IEEE Trans. on Vehicular Tech.*, 53(6), pp 1713-1725 (2004).

[2] P. Castillo, A. Dzul, and R. Lozano, "Real-Time Stabilization and Tracking of a Four-Rotor Mini Rotorcraft", *IEEE Trans. on Control Systems Technology*, 12(4), pp 510-516 (2004).

[3] S. Park, D. H. Won, M.S. Kang, T. J. Kim, H. G. Lee, and S. J. Kwon, "RIC (Robust Internal-loop Compensator) Based Flight Control of a Quad-Rotor Type UAV", Proceedings IEEE/RSJ Int. Conf. on Intell. Robots and Systems (IROS), pp 3542 - 3547 (2005).

[4] C. Tisse, T. Fauvel, and H. Durrant-Whyte, "MAV motion capture system", Proceedings Int. Conf. on Sensing Technologies (ICST 2005), Palmerston North, New Zealand, pp 533-538 (2005).

[5] E. Altuğ, J. P. Ostrowski, and R. Mahony, "Control of a Quadrotor Helicopter Using Visual Feedback", Proceedings IEEE Int. Conf. on Robotics and Automation, Washington, USA, pp 72-77 (2002).

[6] M. G. Earl, and R. D. Andrea, "Real-time Attitude Estimation Techniques Applied to a Four Rotor Helicopter", Proceedings 43rd IEEE Conf. on Decision and Control, Atlantis, Paradise Island, Bahamas, pp 3956-3961 (2004).

[7] W. Wang, G. Song, K. Nonami, M. Hirata, and O. Miyazawa, "Autonomous Control for Micro-Flying Robot and Small Wireless Helicopter X.R.B", Proceedings 2006 IEEE/RSJ Int. Conf. on Intell. Robots and Systems, Beijing, China, pp 2906-2911 (2006).

[8] H. Romero, R. Benosman, and R. Lozano, "Stabilization and location of a four rotor helicopter applying vision", Proceedings 2006 American Control Conf. Minneapolis, Minnesota, USA, pp 3930-3935 (2006).

[9] G. P. Tournier, M. Valentiy, and J. P. How, "Estimation and Control of a Quadrotor Vehicle Using Monocular Vision and Moirè Patterns", Proceedings AIAA Guidance, Navigation, and Control Conf. and Exhibit, Keystone, Colorado (2006).

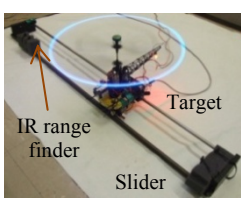
[10] C. Xu, L. Qiu, M. Liu, B. Kong, and Yun. Ge, "Stereo Vision based Relative Pose and Motion Estimation for Unmanned Helicopter Landing", Proceedings 2006 IEEE Int. Conf. On Information Acquisition, Weihai, Shandong, China, pp 31-36 (2006).

[11] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse, "MonoSLAM: Real-Time Single Camera SLAM", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 29(6), pp 1052-1067 (2007).

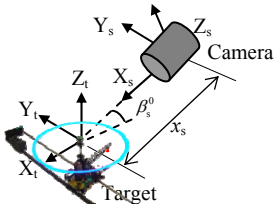
[12] E. Altuğ, J. P. Ostrowski, and C. J. Taylor, "Control of a Quadrotor Helicopter Using Dual Camera Visual Feedback", *Int. J. of Robotics Research*, 24, pp 329-341 (2005).

[13] S. Saripalli, J. F. Montgomery, and G. S. Sukhatme, "Visually guided landing of an unmanned aerial vehicle", *IEEE Trans. on Robotics and Automation*, 19(3), pp 371-380 (2003).

[14] M. Sato, and J. K. Aggarwal "Estimation of position and orientation from Image Sequence of a Circle", Proceedings IEEE Int. Conf. on Robotics and Automation, pp 2252-2257 (1997).

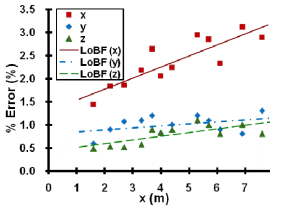


(a) Testing Platform

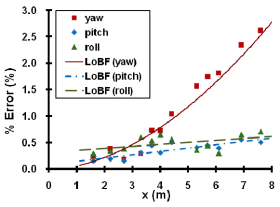


(b) Experimental Setup

Figure 5: Testing Platform and experimental Setup

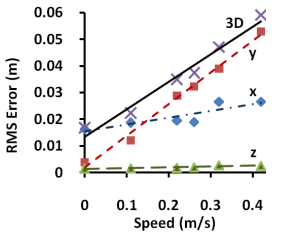


(a) Positioning Error

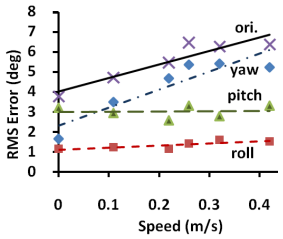


(b) Orientation Error

Figure 6: Experimental results in case 1

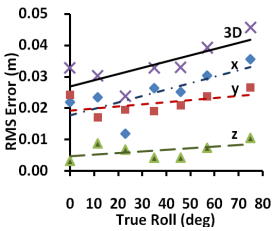


(a) Positioning Error

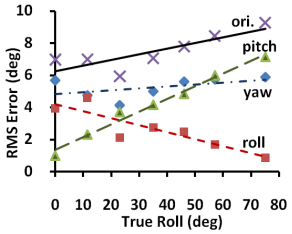


(b) Orientation Error

Figure 7: Experimental results in case 2



(a) Positioning Error



(b) Orientation Error

Figure 8: Experimental results in case 3