Using Stereo Vision to Acquire the Flight Information of Flapping-Wing MAVs*

F. Y. Hsiao[1]*, H. K. Hsu[1], C. L. Chen[1], L. J. Yang[2] and J. F. Shen[3]

[1]*Department of Aerospace Engineering, Tamkang University, Tamsui, Taiwan 251, R.O.C.*
[2]*Department of Mechanical and Electro-Mechanical Engineering, Tamkang University, Tamsui, Taiwan 251, R.O.C.*
[3]*Primax Electronics Ltd, Neihu, Taipei, Taiwan, R.O.C.*

Abstract

This paper investigates the potential to acquire flight information, including the spatial position and attitude, of a flapping-wing micro-aerial vehicle (MAV) utilizing a stereo-vision system. The flapping-wing MAV used in this paper is the *Golden Snitch* developed by the MEMS Laboratory in the Tamkang University. The *Golden Snitch* has wing span of 20 cm and weight of 8 g. Due to limited loading capacity, a conventional inertia measurement unit cannot be installed onboard. As a result, an external stereo-vision system is a potential solution to navigate a flapping-wing MAV. At beginning, techniques of image processing and stereo vision are briefly reviewed. Then, formulae to obtain flight information through the measurements of the stereo-vision system are derived. Four types of experiments are designed and accomplished to evaluate the performance of the system. Experiment results are provided to demonstrate the applicability and constraints of our algorithms.

Key Words: Flapping-Wing MAV, Stereo Vision, Image Processing, Attitude Determination, Autonomous Flight

1. Introduction

This paper investigates the acquisition of flight information of flapping-wing micro-aerial vehicles (MAVs) using a stereo-vision system. The flight information includes aircraft attitude and spatial position. The knowledge of flight information is crucial to the control of an unmanned aerial vehicle (UAV). For a regular UAV, it is not difficult to measure its flight information. An initial measurement unit (IMU) is usually installed onboard to collected required flight information [1].

For an MAV, however, it is too small to install these measurement instruments. As a result, an alternative method should be developed. In this paper we propose to acquire the flight information of MAVs with a stereo-vision system. Machine vision has been utilized in target tracking for long time. Many researchers have committed to the development and applications of machine vision [2,3]. In most researches, only one single camera is employed. The disadvantage of using a single camera is that only the relative position of the target in the image plane can be detected. To obtain the three dimensional location of an object one needs to employ a stereo-vision system, where two cameras are employed. Some examples of researches on the applications of stereo-vision systems can be found in [4] and [5].

One the other hand, a flapping-wing MAV draws more attention among researchers. Flapping flight is an efficient way to transport a unit of mass over a unit of distance, even though it requires high power output [6]. Many laboratories are committed to the development of

*The previous version of this paper was presented as a candidate of the Student Best Paper Award in the 2011 AASRC Conference, Feng-Chia University, Taichung, Nov. 5, 2011.

*Corresponding author. E-mail: fyhsiao@mail.tku.edu.tw

flapping-wing MAVs [7–11]. Among the developed MAVs, the *Golden Snitch*, named after the device referenced in the popular Harry Potter series, is developed by the MEMS Laboratory in the Tamkang University (TKU). It has wingspan of 20 cm, and weight of 8 g, including the fuselage, flapping wings, tail wings, a battery, a motor, and a set of gear system [9]. A similar robotic bird is the Delfly II [10]. Due to the constraint of the bird size and loading capability, both of the *Golden Snitch* and the Delfly II are navigated and guided by external machine-vision systems [10,12]. The *Golden Snitch* is navigated and guided using a stereo-vision system, while the Delfly II uses a mono-camera system.

Based on the results of [5], this paper proposes to measure the positioning and attitude information of a flapping-wing MAV with a stereo-vision system. At beginning, image processing techniques used in the experiments are briefly introduced. Formulae to compute the three dimensional location of a point target are derived. With the knowledge of three to four points on an aircraft, the attitude of the aircraft can be obtained. Accordingly, given several marked anchor points on the fuselage and wings, we are able to compute the attitude of an aircraft. Measurements from an IMU are introduced to verify the correctness of our experiments. Experiment results are provided to demonstrate the applicability and constraints of our algorithms.

2. Stereo Vision

2.1 Image Processing

Before the discussion of navigation laws, we would like to briefly review image processing procedures. Figure 1 gives an example on the standard procedure of image processing. When a camera shoots a video, images are captured from the frames of the video, and undergone image processing for further analysis. In general, images from a webcam are in the Red-Green-Blue (*RGB*) format. This format is difficult to analyze due to the sensitive nature of *RGB* to the variation of illumination.

Consequently, color space transformation is usually necessary. Two types of color space are selected and compared in this paper. They are the Hue-Saturation-Intensity (*HSI*) and Illuminance-Blue-Red (YC_bC_r). The details of those definitions and transformations are available in [13].

After finishing color space transformation, the image is binarized with a selected threshold value.

$$g_b(i,j) = \begin{cases} 1 & T_{\min} \le g(i,j) \le T_{\max} \\ 0 & otherwise \end{cases} \tag{1}$$

where $g(i,j)$ is the original pixel value, $g_b(i,j)$ is the assigned pixel value in the binarized image, $T_{\min}$ and $T_{\max}$ are the thresholds. In this research, there are two ways to determine the threshold values. In some cases, the *hue* value is more sensitive, while in certain circumstances the C_b and C_r are more sensitive. Therefore, $g(i, j) = H(i,j)$ if the *HSI* transformation is selected, and $g(i, j) = C_b(i,j)$ or $g(i,j) = C_r(i,j)$ if the YC_bC_r transformation is selected. After the binarization the standard procedures in *filtering*, *erosion and dilation* are introduced to eliminate noises, and the *edge detection* is applied to identify the object.

2.2 Centroid Location

When a target is identified with image processing techniques, we then compute the centroid location. The projected area of an object on the image plane is usually not a point. As a result, a natural representation of the body is the centroid of the object image. Suppose the pixel of coordinate (i, j) has the value $g_b(i, j)$ obtained from Eq. (1). Given an image of size $m \times n$, we define the total pixel value as

$$M = \sum_{i=1}^{m}\sum_{j=1}^{n} g_b(i,j) \tag{2}$$

Moreover, define

$$G_b(i) = \sum_{j=1}^{n} g_b(i,j) \tag{3}$$

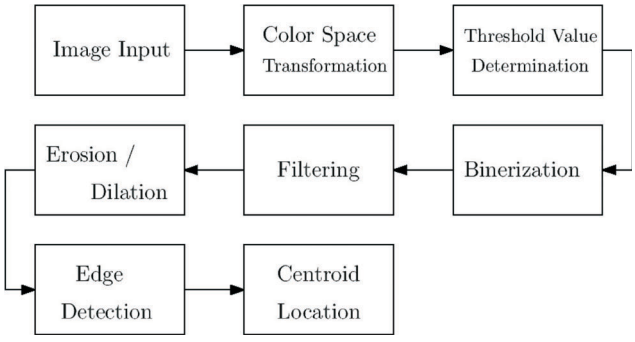

Figure 1. Flow chart of standard procedure on image processing.

$$G_b(j) = \sum_{i=1}^{m} g_b(i,j) \tag{4}$$

Then, the centroid location of an object on the image plane can be found by

$$P_x = \frac{\sum_{j=1}^{n} jG_b(j)}{M} \tag{5}$$

$$P_y = \frac{\sum_{i=1}^{m} iG_b(i)}{M} \tag{6}$$

2.3 Position Acquisition Using Stereo Vision

There are two types of geometry to obtain stereo vision: the crossing method and the parallel method, as shown in Figure 2. The cross method resembles human eyes more. In this paper, however, we select the parallel method. To implement the crossing method one requires the target around the crossing spot for less navigation error, implying that the camera has to track the object all the time when the object is moving. As a result, controllers and actuators should be installed for the cameras to track the target, and this increases the complexity of the whole system. Hence, the parallel method is an easier choice.

Consider a stereo-vision system with the parallel method as described in the preceding paragraph with the geometric parameters defined in Figure 3. The origin of the "camera coordinate system" locates at the center of two cameras. The *x*-axis points rightward, the *y*-axis points from the camera to the image plane, and the *z*-axis points upward. Define the disparity of two cameras to be *C*. Assume that the point × is the target to observe. The angles between the line of sight (LOS) and the central line of a camera are θ_1 and θ_2 for the right and the left cameras, respectively. Assume the depth of the image plane is *L*, and the points locate at x away from the y-z plane.

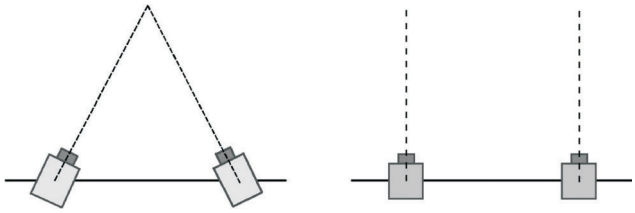

Figure 2. Geometries of the stereo camera. (*Left*) Crossing method; (*Right*) Parallel method.

Then Figure 3 shows the view of the left and the right camera, respectively.

Define $P_{\max}$ as the largest pixel numbers counted from the central line, and $\theta_{\max}$ as the half field of horizontal view, as defined in Figure 3.

The *x* coordinate of the target satisfies

$$\tan\theta_1 = \frac{x + C/2}{L} \tag{7}$$

$$\tan\theta_2 = \frac{x - C/2}{L} \tag{8}$$

In addition, we can map the two view angles, as shown in Figure 3, to the pixel numbers on the camera screen, given by

$$\frac{P_1}{P_{\max}} = \frac{\tan\theta_1}{\tan\theta_{\max}} \tag{9}$$

$$\frac{P_2}{P_{\max}} = \frac{\tan\theta_2}{\tan\theta_{\max}} \tag{10}$$

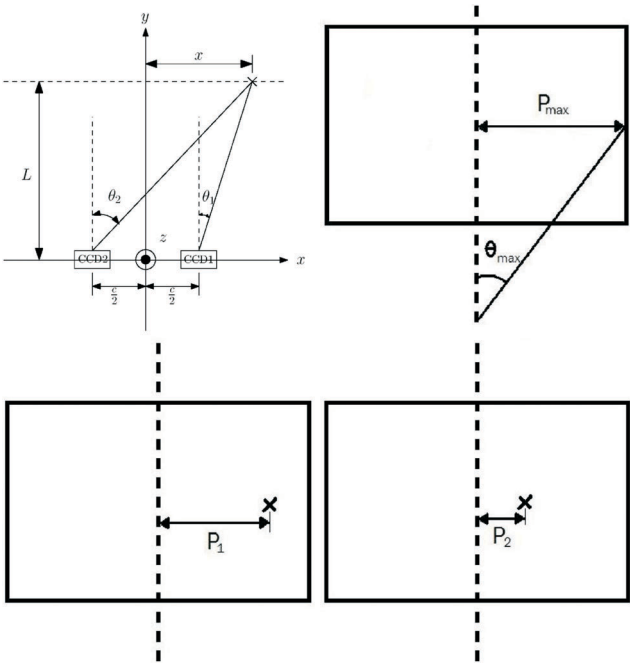

Figure 3. (*Upper Left*) The Camera coordinate system. The origin of the system locates at the center of two cameras. The *x*-axis points rightward, the *y*-axis points from the camera to the image plane, and the *z*-axis points upward. (*Upper Right*) The Largest pixel numbers in the horizontal direction and the half field of view. (*Lower Left*) The view of the left camera; (*Lower Right*) The view of the right camera.

Then $\tan\theta_1$ and $\tan\theta_2$ can be expressed in terms of pixel numbers and half field of view, given by

$$\frac{P_1}{P_{\max}}\tan\theta_{\max} = \frac{x + C/2}{L} \tag{11}$$

$$\frac{P_2}{P_{\max}}\tan\theta_{\max} = \frac{x - C/2}{L} \tag{12}$$

Solving the above two equations by canceling out L yields

$$\frac{P_1}{P_2} = \frac{x + C/2}{x - C/2} \tag{13}$$

Define $\rho = P_1/P_2$. x can be expressed in terms of C and ρ by

$$x = C\left(\frac{1}{2} + \frac{1}{\rho - 1}\right) \tag{14}$$

Equation (14) implies that the x coordinate of the target can be found once we are aware of the disparity distance, and identify the object on image planes of the two cameras.

As for the depth, L, it is attainable through solving Eq. (12). We first define $\gamma = P_2/P_{\max}$. Then L can be obtained by

$$L = \frac{C}{\gamma(\rho - 1)\tan\theta_{\max}} \tag{15}$$

The height of the target can be computed in the similar way. Having obtained L, we use it to calculate the height z by

$$\tan\phi = \frac{z}{L} \tag{16}$$

where ϕ is the vertical view angle as defined in Figure 4. By taking account the pixel relations we have

$$\frac{q}{q_{\max}} = \frac{\tan\phi}{\tan\phi_{\max}} \tag{17}$$

Similarly, $\tan\phi$ can be replaced in terms of known geometric parameters, given by

$$\frac{q}{q_{\max}}\tan\phi_{\max} = \frac{z}{L} \tag{18}$$

Equation (18) then yields

$$z = \mu\tan\phi_{\max} \tag{19}$$

where $\mu = q/q_{\max}$.

3. Flight Information Using Stereo Vision

3.1 Velocity and Acceleration Acquisition

With the algorithm described in the previous section, we are able to obtain the spatial position of a target at every moment. The algorithm of identifying the spatial position of a target remains the same even if the target is moving, because the target is temporary "frozen" on the image at every moment.

Assume that the position at $t = t_i$ is given by $r_i = (x_i, y_i, z_i)$, where $i = 1\ldots n$. The velocity $\mathbf{v}_i = (v_{x_i}, v_{y_i}, v_{z_i})$ can then be computed by

$$v_{x_i} = \frac{x_{i+1} - x_i}{t_{i+1} - t_i} \tag{20}$$

$$v_{y_i} = \frac{y_{i+1} - y_i}{t_{i+1} - t_i} \tag{21}$$

$$v_{z_i} = \frac{z_{i+1} - z_i}{t_{i+1} - t_i} \tag{22}$$

The acceleration can be estimated by fitting the observed data into a parabolic curve. Given time history of three positions, $(t_i, \mathbf{r}_i)$, $(t_{i+1}, \mathbf{r}_{i+1})$, and $(t_{i+2}, \mathbf{r}_{i+2})$, we must be able to find a parabola $\mathbf{r}_i = \mathbf{c}_{2_i} t^2 + \mathbf{c}_{1_i} t + \mathbf{c}_{0_i}$ that passes through these points. We note that $\mathbf{c}_{0_i}$, $\mathbf{c}_{1_i}$ and $\mathbf{c}_{2_i}$ are coefficient vectors of three components at $t = t_i$. If the time difference between consecutive points is very small, i.e., $\Delta t = t_{i+1} - t_i < 1$, the acceleration can be treated con-

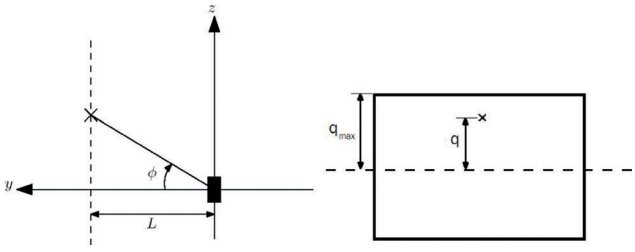

Figure 4. Definition of vertical parameters.

stant. Accordingly, the x coordinate must satisfy

$$\mathbf{r}_{i+1} = \mathbf{r}_i + \mathbf{v}_i t + \frac{1}{2}\mathbf{a}_i t^2 \tag{23}$$

Comparing the coefficients we conclude that $\mathbf{a}_i = 2\mathbf{c}_{2_i}$, where $\mathbf{a}_i = (a_{xi}, a_{yi}, a_{zi})$.

3.2 Attitude Acquisition

Assume that we are able to identify any point and calculate its spatial position relative to the camera coordinate system defined in Figures 3 and 4. Suppose four points on the plane, the head, the tail, and two wings, are marked and their coordinates are identified. By subtracting the coordinates of tail from the head, and left wing from the right wing, we can obtain the vectors along fuselage and wings, given by $\mathbf{r}_f$ and $\mathbf{r}_w$, respectively.

In our research, the model should be slightly modified. Two modifications are made. Different from a fixed-wing aircraft, whose main wing is static, a flapping wing vehicle vibrates its wings all the time. Consequently, to mark on the main wings of a flapping-wing vehicle as shown in Figure 5 is unrealistic. Instead, we mark on the stabilizers, as shown in Figure 6. Two tips of the stabilizers also form the vector $\mathbf{r}_w$.

On the other hand, three points are marked instead of four points as in Figure 5. In a fixed-wing vehicle, the fuselage vector can be found by connecting the cockpit and the tail. In our case, however, it is not easy to mark the cockpit. The mark in the cockpit will be shadowed by the flapping wings, and this might influence the observation. An alternative method is proposed. In order to have better performance, the stabilizers are designed backswept. As a result, by finding the middle point of the two tips of the stabilizers, we are able to locate the tail of the aircraft. We also mark the root of the stabilizers.

Suppose $\mathbf{r}_{ls}$ and $\mathbf{r}_{rs}$ represent the locations of the two tips of the left and right stabilizer, respectively. Then,

$$\mathbf{r}_w = \mathbf{r}_{rs} - \mathbf{r}_{ls} \tag{24}$$

Let

$$\mathbf{r}_t = \frac{1}{2}(\mathbf{r}_{rs} + \mathbf{r}_{ls}) \tag{25}$$

and $\mathbf{r}_r$ be the location of the root of the stabilizers. From the geometry, we conclude that

$$\mathbf{r}_f = \mathbf{r}_r - \mathbf{r}_t \tag{26}$$

The attitude of an aircraft is usually described in terms of Euler angles: the pitch angle θ, the yaw angle ψ, and the roll angle ϕ. It is well known that the rotation sequence is important in finite rotation. As a result, we select the sequence yaw-pitch-roll throughout this paper, as adopted in conventional flight dynamics.

Note that in the camera coordinate system the z-axis points upward while the conventional flight dynamics requires z-axis pointing toward the Earth. Accordingly, to get vectors in the camera coordinate system, we have to pre-multiply $\mathbf{r}_f$ and $\mathbf{r}_w$ by a transformation matrix $\mathbf{T}$. That is,

$$\mathbf{r}_{f,c} = \mathbf{T}\mathbf{r}_f \tag{27}$$

$$\mathbf{r}_{w,c} = \mathbf{T}\mathbf{r}_w \tag{28}$$

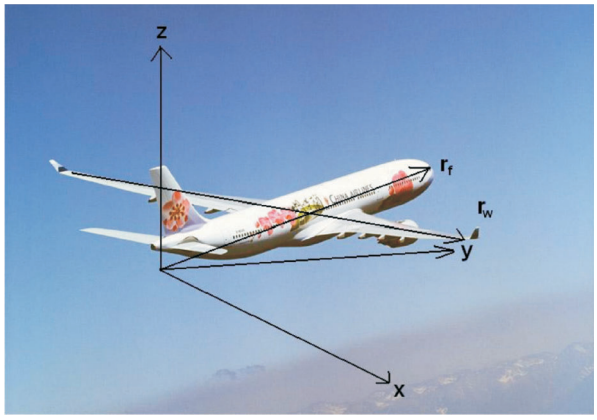

Figure 5. How attitude is determined using vectors along fuselage and wing.

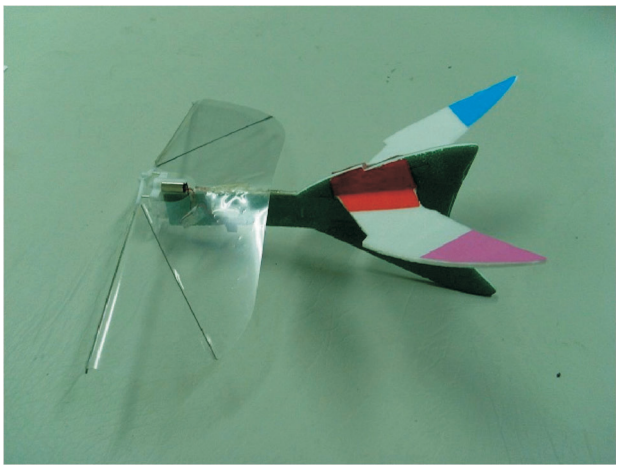

Figure 6. A marked flapping-wing MAV used in the experiments. Three points on the stabilizers are marked, instead of four for a fixed-wing aircraft.

where

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \tag{29}$$

$$\mathbf{T}^{-1} = \mathbf{T} \tag{30}$$

and $\mathbf{r}_{f,c}$ and $\mathbf{r}_{f,c}$ denote the fuselage and wing vectors relative to the camera coordinate system, respectively.

The transformation of a vector in the body-fixed frame to that in the inertia frame is given by [14]

$$\mathbf{q}_b = R_x(\phi)R_y(\theta)R_z(\psi)\mathbf{v} \tag{31}$$

$$= R_x(\phi)R_y(\theta)R_z(\psi)\mathbf{T}^{-1}\mathbf{q}_c \tag{32}$$

$$= R_x(\phi)R_y(\theta)R_z(\psi)\mathbf{T}\mathbf{q}_c \tag{33}$$

where $\mathbf{q}_b$ denotes a vector in body-fixed frame while $\mathbf{q}_c$ denotes a vector in the inertial frame. Moreover, the matrix representation of $\hat{\mathbf{r}}_f$ in body-fixed frame is (1,0,0) and $\hat{\mathbf{r}}_w$ is (0,1,0). $R_x(\cdot)$, $R_y(\cdot)$, and $R_z(\cdot)$ are the rotation matrices about the local x, y, z axis, respectively, given in [14].

One of the advantages to adopt the *yaw-pitch-roll* sequence is that the pitch and yaw angles are intuitively defined. The pitch is always defined as the angle between the local x-axis and the inertial x-y plane. The yaw angle is attainable via considering angle between the projection of the local x-axis in the inertial x-y plane and the inertial x-axis, as depicted in Figure 7.

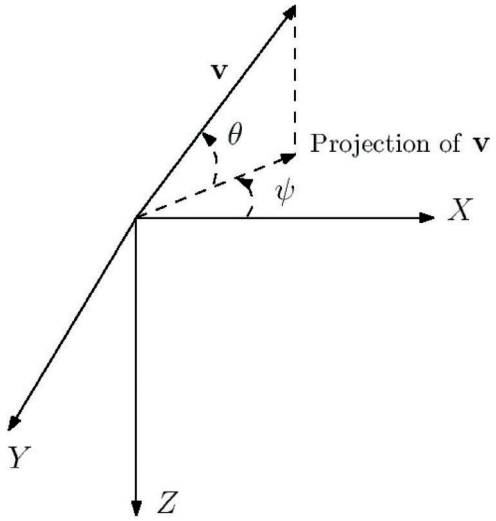

Figure 7. A cartoon depicting the definition of pitch and yaw angles.

Accordingly, the pitch angle can be found by

$$\theta = \frac{\pi}{2} - \cos^{-1}\left(\frac{-\mathbf{z}\cdot\mathbf{r}_f}{|\mathbf{r}_f|}\right) \tag{34}$$

The yaw angle can be found by

$$\psi = \tan^{-1}\left(\frac{-\mathbf{y}\cdot\mathbf{r}_f}{\mathbf{x}\cdot\mathbf{r}_f}\right) \tag{35}$$

Having obtained θ and ψ, the roll angle ϕ must satisfy Eq. (33) by

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = R_x(\phi)R_y(\theta)R_z(\psi)\hat{\mathbf{r}}_w \tag{36}$$

Solving Eq. (36) we can obtain the roll angle ϕ.

4. Hardware and Test Bed

4.1 Hardware

4.1.1 CCD Camera

Two commercial, automatically focusing CCD cameras are installed in the experiment, showing in the left of Figure 8. In order to lower down the development expenses, we don't select a very expensive but high performance camera. Instead, an inexpensive, commercial webcam is selected. The specifications are provided in the following: The specifications of the CCD are listed in the following:

- Product: JX-A7428
- Interface: USB2.0
- Pixel: Dynamic 1.3 million, photo 5 million

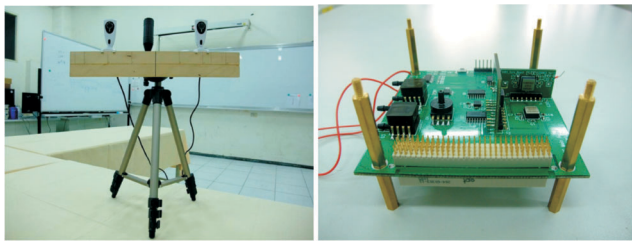

Figure 8. (*Left*) The CCD cameras used in the experiments; (*Right*) The inertial measurement unit used in this research for comparison of verification of stereo-vision observation.

- Dynamic dpi: 1280 × 1024
- Static dpi: 1280 × 1024
- Sensor: CMOS lens

4.1.2 IMU

In order to understand the performance of the stereo-vision system more qualitative, an IMU is employed to compare the measurements. The IMU, shown in the right of Figure 8 is developed by the Avionics and Flight Simulation Laboratory in the Tamkang University. The accelerometer is a two-axis, MEMS accelerometer of series number ADXL320 developed by the Analog Devices. The output signal is analog, and the range of measurement is ±5 g. It measures both dynamic and static acceleration [15].

Here list some specifications:

- Input voltage: 5 V
- Sensitivity: 312 mV/g
- Output voltage at 0 g: 2.5 V
- Noise (RMS): 150 $\mu g/\sqrt{Hz}$

4.1.3 Four-Axis Platform

A four-axis platform is introduced to verify the attitude measurements by the stereo-vision system. A four-axis platform is introduced to cross-verify our experimental results. In the platform system, the z-axis points along the axial direction, while the x- and y-direction point along the transverse directions. As a result, the heading of the MAV can be changed by the rotation about the z-axis. The pitch and yaw angles can be changed by the rotation about the x- or y-axis. The designated load is 25 kg. Here list the rotation range and maximum rotation rate of every axis:

- x-axis – Range: ±45°. Maximum rotation rate: 50 rpm.
- y-axis – Range: ±60°. Maximum rotation rate: 50 rpm.
- z-axis – Maximum rotation rate: 75 rpm.

4.2 User Interface

For the purpose of easy manipulation, we develop a GUI panel using Matlab [16]. This panel controls image grabbing, image processing, coordinate and attitude computation, and display the final results. All the processes can be finished within 0.1 to 0.3 seconds, depending on the complexity of the incoming figures. On the other word, we can acquire attitude at the frequency of around 3 to 10 Hz.

4.3 Calibration

4.3.1 Image Distortion

In order to lower down the experimental expenses, a commercial webcam is selected in this research. However, the low-cost equipments do not have the function of distortion correction. Due to the imperfect of lens, there must exits certain distortion in the taken image. The distorted image will then influence the calculation of coordinates and attitude, and needs calibrating. A simple experiment is conducted to calibrate those errors.

As shown in Figure 9, a grid, dotted array is displayed in front of our cameras. By comparing the measured position and the actual position of each dot, we are able to establish an error lookup table. When experiments are performed, a corresponded correction is introduced.

4.3.2 Image Plane Alignment

In addition to the error induced from distorted image, the alignment of camera coordinate system with the inertial frame is critical, too. Practically it is difficult to achieve the alignment. The image plane, generated by the aiming of two cameras, may be slightly away from the vertical plane. The misalignment of image plane will introduce an offset to the measurement of the target position.

A calibration experiment is designed in Figure 10. The left picture shows the actual equipment while the right shows the numbering of points. The stereo vision is able to measure the position of every point, denoted by

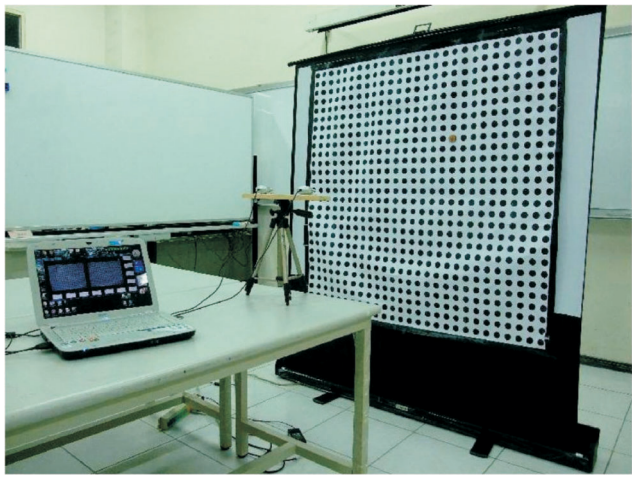

Figure 9. The calibration grid.

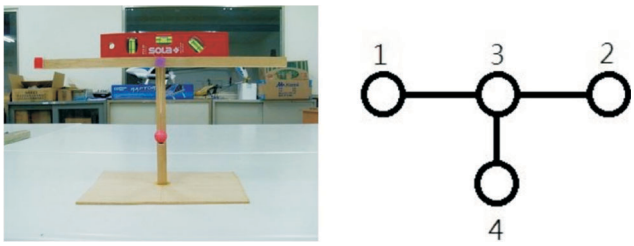

Figure 10. A setup for calibration of image misalignment. (*Left*) The actual setup. (*Right*) The numbering of points.

$\mathbf{r}_{sv,i}$, where $i = 1, 2, 3, 4$, relative to the camera coordinate system defined in Figure 3. Define the horizontal and vertical vectors in the camera coordinate system, respectively, by

$$\mathbf{r}_{sv,h} = \mathbf{r}_{sv,3} - \mathbf{r}_{sv,1} \tag{37}$$

$$\mathbf{r}_{sv,v} = \mathbf{r}_{sv,3} - \mathbf{r}_{sv,4} \tag{38}$$

On the other hand, we can physically measure the positions of these points in the local-vertical-local-horizon (LVLH) system, denoted by $\mathbf{r}_{LVLH,i}$, where $i = 1, 2, 3, 4$, with the assumption that $\mathbf{r}_{LVLH,3} = (0,0,0)$. Then, the horizontal and vertical vectors in the LVLH system can be respectively defined by

$$\mathbf{r}_{LVLH,h} = \mathbf{r}_{LVLH,3} - \mathbf{r}_{LVLH,1} \tag{39}$$

$$\mathbf{r}_{LVLH,v} = \mathbf{r}_{LVLH,3} - \mathbf{r}_{LVLH,4} \tag{40}$$

Assume that the distortion of image plane has been corrected with the method proposed in the previous section. If no distortion of image plane is considered, the LVLH and the camera coordinate system can be transformed by

$$\mathbf{r}_{LVLH} = \mathbf{r}_{sv,0} + \mathbf{R}\mathbf{r}_{sv} \tag{41}$$

where $\mathbf{r}_{sv,0}$ is the offset of these two systems and $\mathbf{R}$ is the rotation matrix that considers the difference in the orientation of the two systems.

There are six unknowns in Eq. (41), three in $\mathbf{r}_{sv,0}$ and three orientation angles in the matrix R. The transformations of horizontal and vertical vectors offer six equations. Thus, this system would be consistent. Having obtained $\mathbf{r}_{sv,0}$ and $\mathbf{R}$, we are able to transform every measurement in the camera coordinate system to the absolute position in the LVLH system with Eq. (41)

5. Experiments

5.1 Translational Information – Ground Test

The translation information includes position, velocity and acceleration of the target. The stereo-vision measures the position of the target at every moment. Then the velocity and acceleration are obtained from equations provided in section 3.1.

In order to understand the precision of the proposed, an IMU as well as a wireless transmitter are equipped on a remote-control car, as shown in Figure 11. In the experiment, the car repeats stop and run for several times. The positions of the car are recorded by the stereo vision and the accelerations are sensed by the IMU.

5.2 Translational Information – Flight Test

The first type of flight tests is designed to measure the flight altitude. Examples of these experiments are shown in Figure 12, and a video showing the experiments is provided in [17].

Due to space constraints in the laboratory, the MAV is attached to the ceiling with a string of negligible mass, which forces the MAV to fly within a certain range. Moreover, the length of the string is long enough not to influence its vertical motion. Although the string becomes tight to provide centripetal force when the MAV is flying, the fact that any deviation from the nominal altitude still results in the trigger of control. By flying faster or slower, the robotic bird can adjust itself to the designated nominal altitude. We place several blue stripes on the wall, as shown in Figure 12, denoting 1.8 m, 1.6 m,

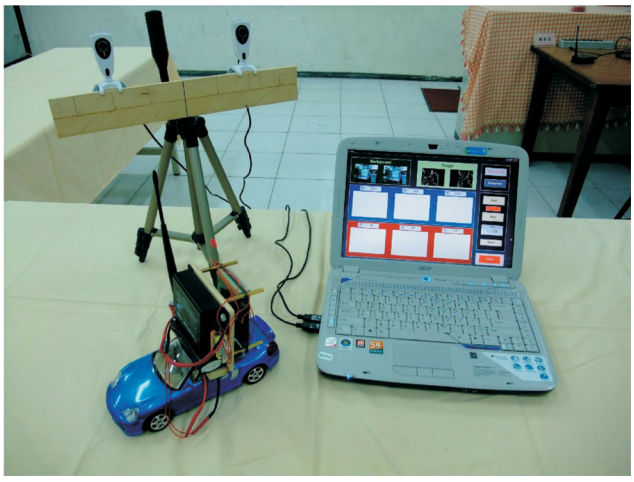

Figure 11. A remote-control car equipped with an inertia measurement unit (IMU) and a wireless transmitter.

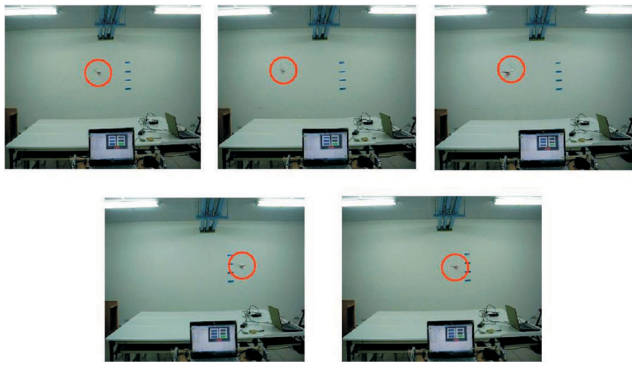

Figure 12. Flight tests in the laboratory. The MAV is hung to the ceiling with a string of negligible mass, in order to force the MAV to fly within a certain range. The blue stripes on the wall, denoting 1.8 m, 1.6 m, 1.4 m, and 1.2 m, respectively, from top to bottom, are set to provide visual references for human operators. The MAV is highlighted by a red circle. It is obvious that the MAV flies around the pre-assigned cruise altitude $h = 1.5$ m.

1.4 m, and 1.2 m, respectively, from top to bottom, to provide visual references for human operators. The measured results are provided in Figures 16 and 17, where the nominal altitudes are controlled at 1.5 m and 1.2 m, respectively.

5.3 Attitude Acquisition – Static Tests

In this part, the flapping-wing MAV is fixed on the four-axis platform, but the platform does not rotate. Two types of experiments are performed, the wings flap and the wings don't flap. In the experiments, we rotate the four-axis platform with designated angles, and place the MAV in different orientation so that various static attitudes can be generated. The main purpose of these tests is to verify the accuracy of the stereo-vision system, and to understand whether or not flapping is going to affect observations.

Three cases are run for the static tests of attitude determination. In each case, two Euler angles are fixed and one angle varies. It only requires the x-axis to be manipulated if the experiments are presented in this way. This simplifies the procedure of experiments. Suppose that the platform rotates about the x-axis with an angle. The MAV still attains different attitudes by placing it toward different directions. In our experiments, the MAV is placed toward four directions, the rightward, the forward, the leftward, and the backward, corresponded to four orientations. As a result, for every provided angle about the x-axis, we are able to obtain four sets of Euler angles. In the whole experiments, three angles are about the x-axis are provided, leading to 12 situations in total.

5.4 Attitude Acquisition – Dynamic Tests

In this part, the flapping-wing MAV is fixed on the four-axis platform, and the platform rotates about the z-axis. One arm is fixed along the y-axis, and the MAV is fixed at the tip of the arm. By rotating the y-axis, we are able to simulate different pitch angles. By tilting the arm up or down, i.e., to rotate the arm about the x-axis, we are able to simulate different roll angles. By rotating the arm about the z-axis, we can simulate the MAV circling around and the yaw angle varies with time.

In reality, the flight speed of the flapping-wing MAV developed in the TKU is around 4 m/s. As a result, the corresponded rotation rate of the four-axis platform is set to be 230 deg/s in this experiment.

6. Result and Discussion

6.1 Translational Information – Ground Test

The experiment results are provided in Figures 13 to 15. We can see that the information obtained via from the stereo vision is quite similar to that from IMU. However, at the instants when the car "starts to run" and "starts to stop", the data from IMU is more accurate since the sampling rate is higher. The curves obtained from the stereo vision at those moments are smoother because the sampling rate is too slow, implying that the application of this navigation system restricts to slower motion dynamics.

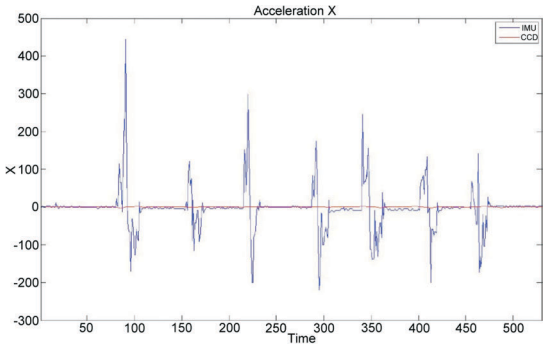

Figure 13. The comparison of acceleration history along the x-axis obtained from stereo vision and IMU. The blue line denotes the measurements from the IMU while the red line demotes the measurements from the stereo vision. The unit for the acceleration is cm/s^2.

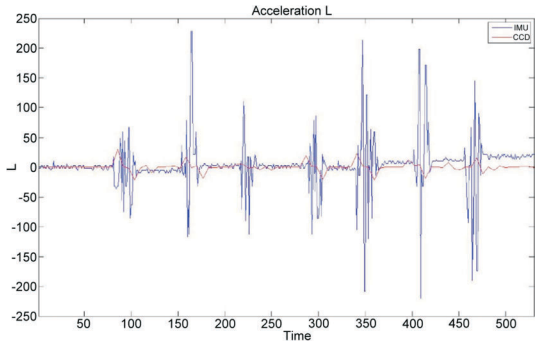

Figure 14. The comparison of acceleration history along the y-axis obtained from stereo vision and IMU. The blue line denotes the measurements from the IMU while the red line demotes the measurements from the stereo vision. The unit for the acceleration is cm/s^2.

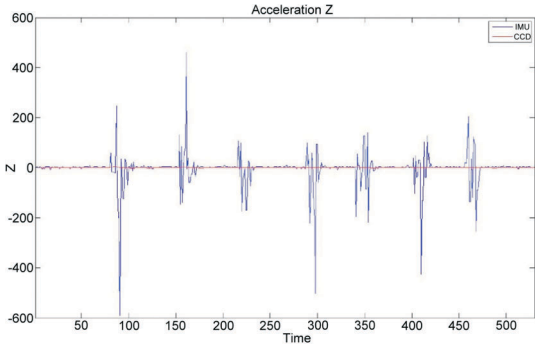

Figure 15. The comparison of acceleration history along the z-axis obtained from stereo vision and IMU. The blue line denotes the measurements from the IMU while the red line demotes the measurements from the stereo vision. The unit for the acceleration is cm/s^2.

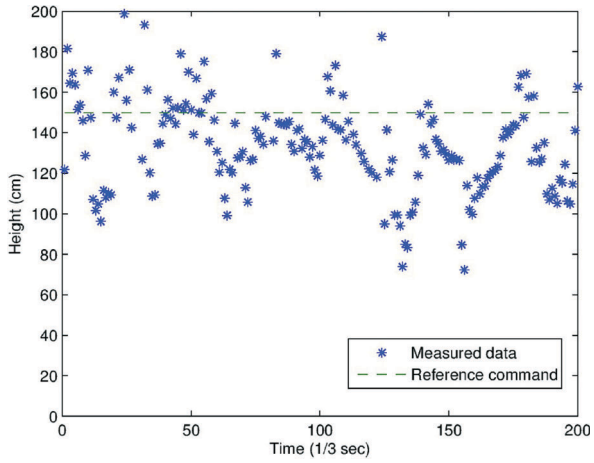

Figure 16. Position data acquired by the stereo-vision system. The nominal altitude is set as 1.5 m.

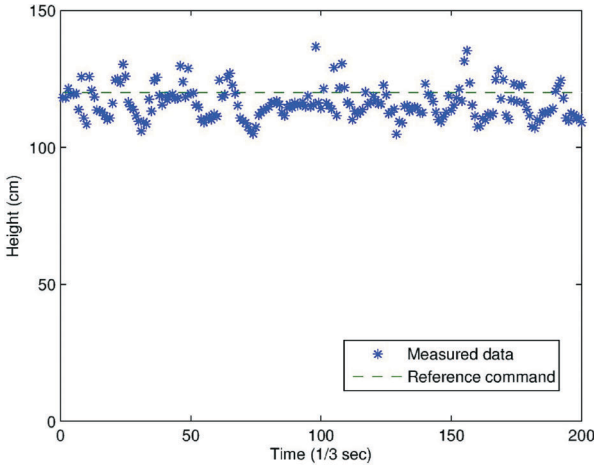

Figure 17. Position data acquired by the stereo-vision system. The nominal altitude is set as 1.2 m.

6.2. Translational Information – Flight Test

Two flight tests result are shown in Figures 16 and 17. In Figure 16 the nominal altitude is set at h = 1.5 m, and in Figure 17 the nominal altitude is set at h = 1.2 m. The observations are noisy for both cases, but the results are close to what is designed.

In Figure 16, the average measured height is about h = 1.27 m with standard deviation (STD) of 24 cm. As a result, the average error of this observation is 13 cm. In Figure 17, the average measured height is about h = 1.13 m with STD of 7 cm. As a result, the average error of this observation is 7 cm. Define the error percentage by

$$e = \frac{|h_{nom} - h_{sv}|}{h_{nom}} \times 100\% \tag{42}$$

where h_{nom} is the designated altitude and h_{sv} is the average measured altitude by the stereo vision. According to the definition $e_{1.5\text{m}}$ = 15.3% and $e_{1.2\text{m}}$ = 5.8%.

6.3 Attitude Acquisition – Static Tests

Four cases are run for the static tests of attitude determination. The results are provided in Tables 1 to 4. In each case, two Euler angles are fixed and one angle varies. 100 measurements are collected for the statistic purpose for every angle set in each case. In this experiment, we define $\psi = 0°$ if it aligns with the x-axis of the camera coordinate system.

Note that the definition of Eq. (42) may not applicable in this case, because some of the nominal angles are 0° and this will cause singularity. Here two errors are defined. The individual absolute error is given by

$$e_\gamma = |\gamma_{nom} - \gamma_{sv}| \tag{43}$$

where γ_{nom} is the nominal Euler angle and γ_{sv} denotes the measurements with the stereo vision. $\gamma = \psi$ if the yaw angle is considered. The same rule applies to $\gamma = \phi$ and $\gamma = \theta$. We also define the gross error by

$$e_g = \sqrt{e_\psi^2 + e_\phi^2 + e_\theta^2} \tag{44}$$

From Tables 1 to 3, we realize that the flapping or not won't affect the measurements of the MAV attitude. The error is very similar given the MAV flapping or not flapping. However, the STD is higher, implying the measurements are noisier, if the wings are flapping. This is reasonable. The fuselage vibrates when the wings are flapping. Vibration will result in blurs to the image, and the blur in the image will cause errors in the centroid and attitude determination. On the other hand, Table 4 gives the gross error for each combination. We can see that the maximum gross error for each set of Euler angles is around 10°, regardless of the attitude.

6.4 Attitude Acquisition – Dynamic Tests

In this part, the flapping-wing MAV is fixed on the four-axis platform, and the platform rotates about the z-axis. The rotation rate of the four-axis platform is set to be 230 degs/s. 150 data is collected and analyzed. The results are presented in Figures 18 to 20.

Table 1. Case 1: $\psi = 0°$, $\theta = 0°$, $\phi = (20°, 25°, 30°)$. The unit of this table is *degree*

Not Flapping									
	Yaw Angle (ψ)			Roll Angle (ϕ)			Pitch Angle (θ)		
Item	Angle 1	Angle 2	Angle 3	Angle 1	Angle 2	Angle 3	Angle 1	Angle 2	Angle 3
Set Value	0	0	0	20	25	30	0	0	0
Measured Value	0.44	1.89	-4.38	21.63	26.07	28.42	-3.44	-5.38	-2.69
Error	0.44	1.89	4.38	1.63	1.07	1.58	3.44	5.38	2.69
STD	0.56	0.71	1.09	1.61	1.64	2.64	3.80	1.58	2.58
Flapping									
	Yaw Angle (ψ)			Roll Angle (ϕ)			Pitch Angle (θ)		
Item	Angle 1	Angle 2	Angle 3	Angle 1	Angle 2	Angle 3	Angle 1	Angle 2	Angle 3
Set Value	0	0	0	20	25	30	0	0	0
Measured Value	1.35	1.49	-3.98	22.15	25.30	26.07	-3.04	-6.05	-2.40
Error	1.35	1.49	3.98	2.15	0.3	3.93	3.04	6.05	2.40
STD	1.62	0.77	0.94	2.54	1.83	2.13	4.66	1.76	2.46

Table 2. Case 2: $\psi = -90°$, $\phi = 0°$, $\theta = (-20°, -25°, -30°)$. The unit of this table is *degree*

Not Flapping									
	Yaw Angle (ψ)			Roll Angle (ϕ)			Pitch Angle (θ)		
Item	Angle 1	Angle 2	Angle 3	Angle 1	Angle 2	Angle 3	Angle 1	Angle 2	Angle 3
Set Value	-90	-90	-90	0	0	0	-20	-25	-30
Measured Value	-79.59	-82.10	-80.92	-0.74	-0.74	-0.65	-17.17	-20.10	-24.63
Error	10.41	7.90	9.08	0.74	0.74	0.65	2.83	4.90	5.37
STD	3.04	3.78	1.87	0.67	0.80	0.85	2.73	2.61	4.15
Flapping									
	Yaw Angle (ψ)			Roll Angle (ϕ)			Pitch Angle (θ)		
Item	Angle 1	Angle 2	Angle 3	Angle 1	Angle 2	Angle 3	Angle 1	Angle 2	Angle 3
Set Value	-90	-90	-90	0	0	0	-20	-25	-30
Measured Value	-81.11	-81.12	-83.59	-0.67	-1.07	-4.62	-16.82	-21.28	-26.59
Error	8.89	8.88	6.41	0.67	1.07	4.62	3.18	3.72	3.41
STD	3.36	4.03	3.99	0.60	0.33	0.91	4.69	3.64	5.43

Table 3. Case 4: ψ = 90°, ϕ = 0°, θ = (20°, 25°, 30°). The unit of this table is *degree*

Not Flapping									
	Yaw Angle (ψ)			Roll Angle (ϕ)			Pitch Angle (θ)		
Item	Angle 1	Angle 2	Angle 3	Angle 1	Angle 2	Angle 3	Angle 1	Angle 2	Angle 3
Set Value	90	90	90	0	0	0	20	25	30
Measured Value	85.47	82.85	87.06	1.56	2.18	1.67	22.96	29.17	41.05
Error	4.53	7.15	2.94	1.56	2.18	1.67	2.96	4.17	11.05
STD	2.12	1.95	10.06	1.50	0.39	2.12	6.24	4.22	4.82
Flapping									
	Yaw Angle (ψ)			Roll Angle (ϕ)			Pitch Angle (θ)		
Item	Angle 1	Angle 2	Angle 3	Angle 1	Angle 2	Angle 3	Angle 1	Angle 2	Angle 3
Set Value	90	90	90	0	0	0	20	25	30
Measured Value	84.94	82.07	87.55	1.66	1.83	1.62	25.05	29.82	41.29
Error	5.06	7.93	2.45	1.66	1.83	1.62	5.05	4.82	11.29
STD	1.89	2.32	11.95	0.93	0.46	1.75	5.85	4.46	3.80

Table 4. The gross error for various combinations of Euler angles listed in Table 1 to 3. The angle sequence is given by (ψ, ϕ, θ). The unit of this table is *degree*

	Flapping			Not Flapping		
Items	(0,20,0)	(0,25,0)	(0,30,0)	(0,20,0)	(0,25,0)	(0,30,0)
e_g	3.8320	5.8018	5.3774	3.9606	6.2380	6.0865
Items	(-90,0,-20)	(-90,0,-25)	(-90,0,-30)	(-90,0,-20)	(-90,0,-25)	(-90,0,-30)
e_g	10.8132	9.3256	10.5691	9.4654	9.6870	8.6058
Items	(90,0,20)	(90,0,25)	(90,0,30)	(90,0,20)	(90,0,25)	(90,0,30)
e_g	5.6317	8.5594	11.5557	7.3391	9.4587	11.6658

Define the average error, e_γ, for parameter γ, where γ functions the same as in Eq. (43).

$$e_\gamma = \frac{\sum|\gamma_p - \gamma_m|}{N_v} \tag{45}$$

where γ_p and γ_m denote the platform angle and the measurements, respectively, and N is the total amount of valid data. In addition to the average error, we also define successful rate by

$$\text{successful rate} = \frac{N_s}{N} \times 100\% \tag{46}$$

where N_s is the "successful data" and N is the total amount of data.

The Euler angles of the MAV can be computed only if the three highlighted points on the fuselage are identified successfully. However, some points may suffer from disturbances. For example, one or more points are sheltered by the fuselage, or the illumination varies so much that the image processing cannot recognize the target with pre-assigned threshold. When the stereo vision cannot compute the Euler angles successfully, it outputs a default value. Therefore, the index of *successful rate* is also defined to evaluate the performance of the stereo vision system.

From Figures 18 and 19 we can see that the average error is around 10° for pitch and roll. This is very consistent with the results from static experiments. For the yaw angle, the first glance of the error is as high as 107°. Actually, the processing rate is around 3 Hz in this experiment. We can see that the trend of the platform motion is similar to that of the measurements, and an time delay of around 0.3 seconds exists between the IMU and stereo-vision data, as shown in Figure 20. Since the four-axis platform is set to rotate with the rate of 230 degs/s, we conclude the time delay may generate an error of as high as 70 degrees. Taking the measurement noise into con-

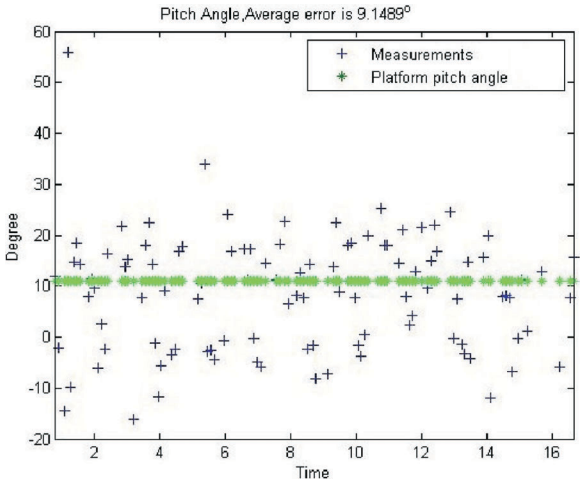

Figure 18. The pitch angle in dynamic test. The dots denote the pitch angle offered by the platform, whereas the solid line denotes the pitch angle measured by the stereo-vision system.

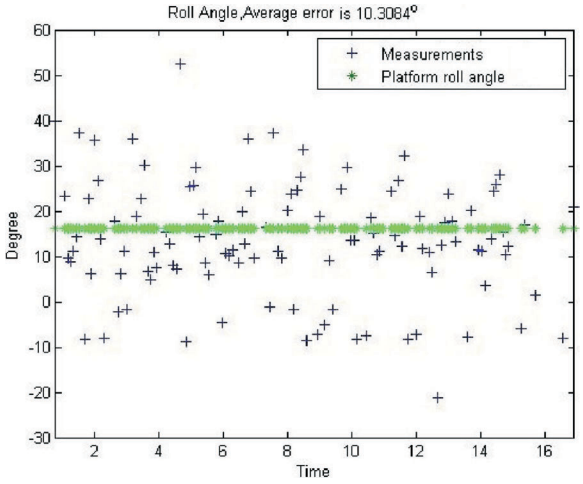

Figure 19. The roll angle in dynamic test. The dots denote the roll angle offered by the platform, whereas the solid line denotes the roll angle measured by the stereo-vision system.

sideration, an error of around 107° is possible in this system. This fact implies that the measurements should be corrected by the phase lag in practical implementation. The successful rates of the three experiments are provided in Table 5. The successful rates for all experiments are more than 70%.

6.5 Overall Discussion

Four types of experiments have been done to verify the potential on applying the stereo vision to flight information acquisition. The results are presented in the previous sections. Although the raw data looks very

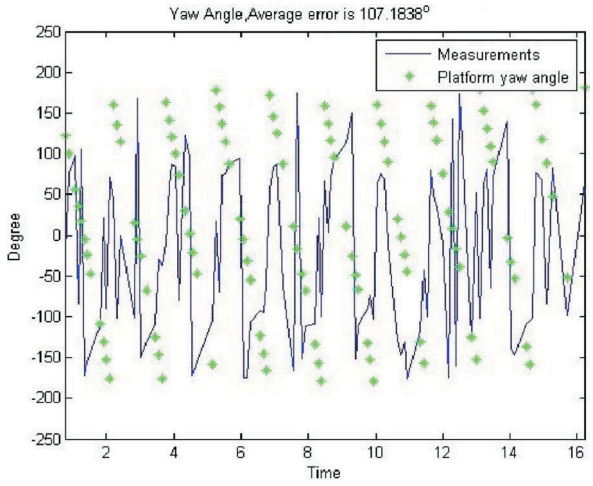

Figure 20. The yaw angle in dynamic test. The dots denote the yaw angle offered by the platform, whereas the solid line denotes the yaw angle measured by the stereo-vision system.

Table 5. Performance comparison for the attitude determination in dynamic tests

Item	Yaw	Pitch	Roll
Successful Rate	77%	71%	77%
Average Error	107.1836°	9.1489°	10.3084°

noisy, the average error is acceptable. Consequently, the measurement should be useful if it goes through a filter, such as a moving-average filter.

From the experiments of dynamic attitude determination, we realize that phase lag will occur if the motion of the MAV is faster than the processing speed of the flapping-wing MAV. Therefore, this result suggests that the results from the stereo vision should be corrected by the time-delay effect.

7. Conclusion

This paper investigates the potential to acquire flight information, including the spatial position and attitude, of a flapping-wing micro-aerial vehicle (MAV) utilizing a stereo-vision system. The flapping-wing MAV used in this paper is the *Golden Snitch* developed by the MEMS Laboratory in the Tamkang University. The *Golden Snitch* has wing span of 20 cm and weight of 8 g. Due to limited loading capacity, a conventional inertia measurement unit cannot be installed onboard. As a result, an external stereo-vision system is a potential solution to the autonomous flight of a flapping-wing MAV. In the paper,

formulae to obtain flight information through the measurement of the stereo-vision system are derived, and indices to evaluate the performance of the stereo vision-system are defined. Four types of experiments are accomplished to evaluate the performance. Experiments results suggest that the error percentage in the measurement of flight altitude vary from 5% to 15%. The absolute errors of Euler angles remain around 10° if the MAV is static or undergoes slow motion. However, the measurements are very noisy, and phase lag will occur if the motion of the MAV is faster than the processing speed of the flapping-wing MAV. As a result, with the introduction of a filter, the stereo-vision system should be applicable to acquiring the information of slow dynamics of flapping-wing MAVs in the future.

Acknowledgement

The work described here is partly funded by the Tamkang University through project 99-Tactical Project-Engineering (08), and partly by the Taiwan National Science Council through project NSC 100-2221-E-032-022.

References

[1] Chen, P. C., *The Study, Design and Realization of the Portable Unmanned Aerial Vehicle*, Master thesis, Department of Aerospace Engineering, Tamkang University (2009).

[2] Liu, Y. N., *Vision-Based Moving Target Tracking and Estimation for Unmanned Helicopter*, Master thesis, Tainan, Natioanl Cheng-Kung University (2009).

[3] Hsu, R. L., "Face Detection in Color Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, pp. 696–706 (2002)

[4] Chen, K. Y., Cheng-Chin Chien, C. C., Chang, W. L. and Hsieh, C. C., "Improving the Accuracy of Depth Estimation Using a Modified Stereo Vision Model in Binocular Vision," *The 10th International Symposium on Measurement Technology and Intelligent Instruments*, Daejeon, Korea (2011)

[5] Chen, C. L. and Hsiao, F. Y., "Attitude Acquisition Using Stereo-Vision Methodology," presented as Paper VIIP 652-108 at the 2009 IASTED Conference, Cambridge, UK, Jul. 13-15 (2009)

[6] Norberg, U. M., *Vertebrate Flight: Mechanics, Physiology, Morphology, Ecology and Evolution*, New York, Springer-Verlag (1990)

[7] Pornsin-sirirak, T. N., Tai, Y. C., Nassef, H. and Ho, C. M., "Titanium-Alloy MEMS Wing Technology for a Micro Aerial Vehicle Application," *Sensors and Actuators A: Physical*, Vol. 89, pp. 95–103 (2001)

[8] Yang, L. J., Hsu, C. K., Ho, J. Y. and Feng, C. K, "Flapping Wings with PVDF Sensors to Modify the Aerodynamic Forces of a Micro Aerial Vehicle," *Sensors and Actuators A: Physical*, Vol. 139, pp. 95–103 (2007).

[9] Yang, L. J., Hsu, C. K., Kao, C. Y., Hsiao, F. Y. and Feng, C. K, "Weight Reduction of Flapping Micro Aerial Vehicles by Electrical Discharge Wire Machining," *Journal of Aeroautics, Astronautics and Aviation, Series A*, Vol. 41, pp. 165–172 (2009).

[10] de Croon, G. C. H. E., de Clercq, K. M. E., Ruijsink, R., Remes, B. and de Wagter, C., "Design, Aerodynamics, and Vision-Based Control of the DelFly," *International Journal of Micro Air Vehicles*, Vol. 1, pp. 71–97 (2009)

[11] Bermudez, F. G. and Fearing, R., "Optical Flow on a Flapping Wing Robot," *Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, IEEE Press Piscataway, NJ, USA (2009).

[12] Lin, S. H., Hsiao, F. Y., Chen, C. L. and Shen, J. F., "Altitude Control of Flapping-Wing MAV Using Vision-Based Navigation," *Paper #WeA01.1 at the 2010 American Control Conference at an Invited Session*, Baltimore, MD, USA (2010).

[13] Cheng, Y. H., *Vision-Based Track Following for Unmanned Aerial Vehicle*. Master's thesis, Tainan, National Cheng Kung University (2008).

[14] Phillips, W. F., *Mechanics of Flight*, John Wiley and Sonic, pp. 623–625 (2004).

[15] Wu, T. H., *Design and Implementation of a MEMS-Based Flight Information Unit*, Master thesis, Department of Aerospace Engineering, Tankang University (2008).

[16] Gonzalez, R. C., Woods, R. E. and Eddins, S. L., *Digital Image Processing Using MATLAB*, 2ed, Gatesmark Publishing (2009).

[17] Flight Test of *Golden Snitch*, Available at http://www.youtube.com/watch?v=bexOl4YNnd0.

Manuscript Received: Mar. 9, 2012
Accepted: May 10, 2012