Vision-encoder-based Payload State Estimation for Autonomous MAV With a Suspended Payload

Yunfan Ren[§], Jianheng Liu[§], Haoyao Chen and Yunhui Liu

***Abstract*— Autonomous delivery of suspended payloads with MAVs has many applications in rescue and logistics transportation. Robust and online estimation of the payload status is important but challenging especially in outdoor environments. The paper develops a novel real-time system for estimating the payload position; the system consists of a monocular fisheye camera and a novel encoder-based device. A Gaussian fusion-based estimation algorithm is developed to obtain the payload state estimation. Based on the robust payload position estimation, a payload controller is presented to ensure the reliable tracking performance on aggressive trajectories. Several experiments are performed to validate the high performance of the proposed method.**

I. Introduction

Cargo transportation with unmanned aircraft has attracted increasing attention in robotic field. By using a cable, the aircraft can carry a payload larger than itself without being disturbed by the airflow. Also, in complex environments, such as jungles and other difficult-to-land scenes, the aircraft can use a cable to transport cargo [1]. As shown in Fig. 1 (left), fast cargo transportation by deliberately swinging the payload can greatly increase the efficiency and reduce energy consumption. However, the practical cargo transportation generally needs experienced helicopter pilots. The autonomous cargo transportation with unmanned aircraft is still a challenging research topic, especially due to the difficult estimation of payload's state.

Many researchers have developed approaches to the problem of flight stability for MAVs with a suspended payload [2]–[4]. Tang et al. [5] developed a controller to realize high MAV maneuvering around obstacles such as windows shorter than the suspension-cable's length. Crousaz et al. [6] applied SLQ control on a quadrotor with a slung load to perform a go-to-goal task. Foehn et al. [7] used an optimized trajectory to drop the payload to a target area. Palunko et al. [8] implemented a method based on reinforcement learning to track the reference trajectory. However, the existing methods completely relied on an external motion capture system to estimate the payload position, and cannot be applied to outdoor scenes.

Fig. 1. Experienced helicopter pilot transporting Christmas tree (left). Autonomous MAV throwing payload precisely to a moving vehicle, using our estimation system (right).

There exist some approaches implementing onboard payload estimation. For example, Tang et al. [9] used a visual method to estimate the position of the load. This method requires the payload to have a specific artificial pattern on the load (like a white circular tag), and the low frequency and high latency of the video stream make it difficult to accurately estimate the states of the load. Lee and Kim [10] used the inertial measurement unit (IMU) to estimate the position of the load; however, the IMU as an integral-type sensor may cause accumulated errors after working for a long time.

In this paper, we aim to develop a high frequency and accuracy method to achieve onboard position estimation of the suspended payload. The MAV-payload system is a hybrid system, when the cable is not taut, the system degenerates to a normal MAV system, and the suspended payload will not influence the MAV. The MAV controller needs autonomous switch for taut and slack situations. The paper contributions are three-fold.

First, a novel onboard payload estimation system by combining encoder and fish-eye camera is designed for practical scenarios without external observations. A Gaussian fusion-based estimation algorithm is further developed by fusing encoder and vision information to obtain accurate and robust performance.

Second, a payload controller which can ensure the reliable tracking performance on an aggressive trajectory is proposed. The vision-encoder estimation system detects whether the cable is taut and the control strategy is switched automatically.

[§] Y.F. Ren and J.H. Liu contributed equally.

This work was supported in part by the National Natural Science Foundation of China under Grant U1713206 and in part by the Shenzhen Science and Innovation Committee under JCYJ20200109113412326, JCYJ20180507183837726, and JCYJ20180507183456108. (Corresponding author: Haoyao Chen.)

Y.F. Ren, J.H. Liu, and H.Y. Chen* are with the School of Mechanical Engineering and Automation, Harbin Institute of Technology Shenzhen, P.R. China, e-mail: hychen5@hit.edu.cn.

Y.H. Liu is with the Department of Mechanical and Automation Engineering, the Chinese University of Hong Kong, P.R. China, e-mail: yhliu@mae.cuhk.edu.hk.

Finally, validation experiments are performed to show the robustness of the whole estimation system to complex flight environment and extreme payload changes.

II. Estimator Design and Modeling

In this section, the design of the vision-encoder-based estimator and the dynamic modeling of the MAV-payload system are presented, which is the basis of our payload estimation and payload control algorithm in Section III and Section IV.

TABLE I
VARIABLES OF THE SYSTEM

$\mathcal{W}, \mathcal{B}, \mathcal{E}$	World, body and estimator frame
$l, g \in \mathbb{R}$	Cable length and gravity constant
$m_Q, m_L \in \mathbb{R}$	Mass of MAV and payload
$\mathbf{x}_Q^{\mathcal{W}}, \mathbf{x}_L^{\mathcal{W}} \in \mathbb{R}^3$	Position vectors of MAV and payload in $\mathcal{W}$
$\mathbf{x}_L^{\mathcal{E}} \in \mathbb{R}^3$	Position vectors of payload in $\mathcal{E}$
$\mathbf{q} \in S^2 \subset \mathbb{R}^3$	Unit vector from MAV to the payload
$\omega \in \mathbb{R}^3$	Anguluar velocity of payload in $\mathcal{W}$
$\mathbf{R}_{\mathcal{B}}^{\mathcal{W}}, \mathbf{R}_{\mathcal{E}}^{\mathcal{B}}, \mathbf{R}_{\mathcal{C}}^{\mathcal{E}} \in SO(3)$	Rotation from $\mathcal{B}$ to $\mathcal{W}$, $\mathcal{E}$ to $\mathcal{B}$ and $\mathcal{C}$ to $\mathcal{E}$
$\mathbf{\Omega} \in \mathbb{R}^3$	Anguluar velocity of MAV in $\mathcal{B}$
$k_{(\cdot)}$	Diagonal gain matrix

A. Frame and variable definitions

Table I illustrates the frame and variable definitions. The right uppercase superscription is used to indicate the coordinate systems. Three coordinate systems are defined as shown in Fig. 2. The body frame $\mathcal{B}$ is defined with its origin at the geometric center of the MAV. The x-axis is pointing forward, the y-axis is pointing to the left, and the z-axis is pointing upward. The vision-encoder-based estimator frame $\mathcal{E}$ is defined with its origin at the geometric center of the encoder device. The x-axis is pointing forward, the y-axis is pointing to the right, and the z-axis is pointing downward. The world frame $\mathcal{W}$ coincides with $\mathcal{B}$ at the initial position. In each frame, e_1, e_2, e_3 are in red, green and blue respectively.

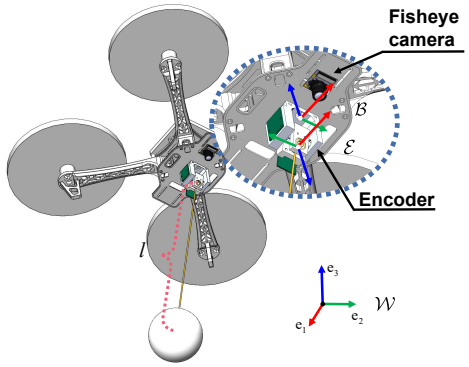

Fig. 2. Definition of coordinate frames.

B. Encoder design and angle definition

To make the payload controllable, the cable between the payload and the MAV should be kept taut. If the cable is not taut, the MAV-payload system degenerates into an ordinary MAV system. Therefore, in general, the workspace of the payload is a spherical shell which is controllable with only two degrees of freedom.

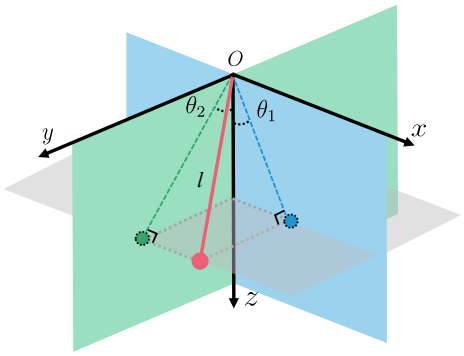

Fig. 3. Polar coordinate definition of the frame $\mathcal{E}$.

By considering the Cartesian coordinates in the estimator frame, the payload vector is projected to plane xOz, yOz. Two angles θ_1 and θ_2 are defined respectively, as shown in Fig. 3. And then, the payload position in the estimator's Cartesian coordinate frame is given as

$$\begin{cases} x^2 + y^2 + z^2 = l^2 \\ \frac{x}{z} = \tan\theta_1 \\ \frac{y}{z} = \tan\theta_2, \end{cases} \tag{1}$$

where l is a constant when the cable is taut.

To measure the above θ_1 and θ_2 values, a novel sensing system is designed. Fig. 4 shows the designed mechanism of the system. AS5047P, a magnetic rotary position sensor, is chosen as our angle sensor. The AS5047P is a high-resolution rotary position sensor for high speed (up to 28krpm) angle measurement over a full 360-degree range with almost 0 latency, and it offers a robust design that suppresses the influence of any homogenous external stray magnetic field. A standard 4-wire SPI serial interface allows a host microcontroller to read 14-bit absolute angle position data from the AS5047P and to program non-volatile settings without a dedicated programmer.

As shown in Fig. 4, a two-axis pan-tilt mechanism is designed to measure the two angles mentioned in Fig. 3. To reduce the friction between the axles, high-speed bearings are adopted and embedded in the 3D-printed support. The support can be further replaced by the metal material, and the equipment can be used in scenes with greater load-bearing requirements. The details of the sensor models are provided as follows.

C. Camera model

A downward-facing fish-eye camera is equipped to estimate the state of the payload, as shown in Fig. 2. Let us

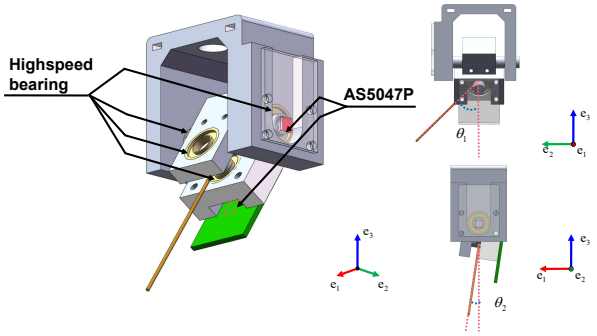

Fig. 4. Mechanism of the encoder-based estimator. The encoder angles θ_1 and θ_2 are illustrated.

assume that the end of the cable coincides with the origin of the encoder-based estimator. The origin of encoder frame is $\mathbf{p}_{\mathcal{E}}^{\mathcal{C}}$ away from the optic center of the camera. According to the pin-hole camera model, the unit vector from the camera to payload is defined as

$$\mathbf{q}^{\mathcal{C}} = \frac{(\frac{\Delta u}{f_x}, \frac{\Delta v}{f_y}, 1)^T}{\|(\frac{\Delta u}{f_x}, \frac{\Delta v}{f_y}, 1)\|}, \tag{2}$$

where Δu and Δv are pixel coordinates of the payload from the image center; f_x and f_y are the intrinsic parameters of the camera model acquired by camera calibration [11]. The payload position in the world frame is obtained by

$$\theta = \arccos(\mathbf{q}^{\mathcal{C}} \cdot \mathbf{p}_{\mathcal{E}}^{\mathcal{C}}), \tag{3}$$

$$\bar{l} = \|\mathbf{p}_{\mathcal{E}}^{\mathcal{C}}\| \cos\theta + \sqrt{\|\mathbf{p}_{\mathcal{E}}^{\mathcal{C}}\|^2 \cos^2\theta - \|\mathbf{p}_{\mathcal{E}}^{\mathcal{C}}\|^2 + l^2}, \tag{4}$$

$$\mathbf{x}_L^{\mathcal{E}} = \mathbf{R}_{\mathcal{C}}^{\mathcal{E}} \bar{l} \mathbf{q}^{\mathcal{C}} + \mathbf{p}_{\mathcal{C}}^{\mathcal{E}}, \tag{5}$$

$$\mathbf{x}_L^{\mathcal{W}} = \mathbf{x}_Q^{\mathcal{W}} + \mathbf{R}_{\mathcal{B}}^{\mathcal{W}} \mathbf{R}_{\mathcal{E}}^{\mathcal{B}} \mathbf{x}_L^{\mathcal{E}}, \ \mathbf{q}^{\mathcal{W}} = \mathbf{x}_L^{\mathcal{W}} / \|\mathbf{x}_L^{\mathcal{W}}\|, \tag{6}$$

where θ is the angle between $\mathbf{q}^{\mathcal{C}}$ and $\mathbf{p}_{\mathcal{E}}^{\mathcal{C}}$; $\bar{l}$ is the estimate distance from the camera to the payload.

For simplicity, we omit the superscription of $\mathcal{W}$ in the following sections, i.e. $\mathbf{q}$ denotes $\mathbf{q}^{\mathcal{W}}$, $\mathbf{x}$ to represent $\mathbf{x}^{\mathcal{W}}$, and $\mathbf{R}$ to represent $\mathbf{R}_{\mathcal{B}}^{\mathcal{W}}$.

D. MAV-payload system dynamics

To control the payload efficiently, we model the payload as a point-mass and the cable as a massless rod. According to the Euler-Lagrange equation, the system's kinetic and potential energy functions are respectively given as follows.

$$\mathcal{T} = \frac{1}{2} m_Q \dot{\mathbf{x}}_Q \cdot \dot{\mathbf{x}}_Q + \frac{1}{2} m_L \dot{\mathbf{x}}_L \cdot \dot{\mathbf{x}}_L + \frac{1}{2} \mathbf{\Omega} \mathbb{I} \mathbf{\Omega}, \tag{7}$$

$$\mathcal{U} = m_Q g \mathbf{x}_Q \cdot \mathbf{e}_3 + m_L g \mathbf{x}_L \cdot \mathbf{e}_3. \tag{8}$$

The Lagrangian for the system, $\mathcal{L} : TQ \to \mathbb{R}$, is defined by $\mathcal{L} = \mathcal{T} - \mathcal{U}$. The dynamics of the system satisfies the Lagranged'Alembert principle:

$$\delta \int_0^\tau \mathcal{L}\, dt + \int_0^\tau \left(\langle W_1, \hat{M} \rangle + W_2 \cdot f\mathbf{R}\mathbf{e}_3 \right) dt = 0, \tag{9}$$

where f is the thrust magnitude, $\mathbf{M}$ is the moment vector; $W_1 = \mathbf{R}^T \delta \mathbf{R}$ and $W_2 = \delta \mathbf{x}_Q = \delta \mathbf{x}_L - l\delta \mathbf{q}$ are the virtual work of the system. To sum up, the dynamics equation for the quadrotor with a suspended payload [12] is given as

$$(m_Q + m_L)(\ddot{\mathbf{x}}_L + g\mathbf{e}_3) = (\mathbf{q} \cdot f\mathbf{R}\mathbf{e}_3 - m_Q l(\dot{\mathbf{q}} \cdot \dot{\mathbf{q}}))\mathbf{q}, \tag{10}$$

$$m_Q l \dot{\omega} = -\mathbf{q} \times f\mathbf{R}\mathbf{e}_3, \tag{11}$$

$$\dot{\mathbf{R}} = \mathbf{R}\hat{\mathbf{\Omega}}, \tag{12}$$

$$J_Q \dot{\mathbf{\Omega}} + \mathbf{\Omega} \times J_Q \mathbf{\Omega} = \mathbf{M}. \tag{13}$$

III. Estimation of Payload Position and Taut State

The autonomous control of the MAV-payload system requires accurate payload state estimation. In this section, the algorithm framework of the vision-encoder-based payload estimator is introduced as shown in Fig. 5. The estimation results from the encoder and fish-eye camera are fused with a Gaussian filter, yielding high-frequency, real-time and robust estimate of the payload's position and cable's taut state.

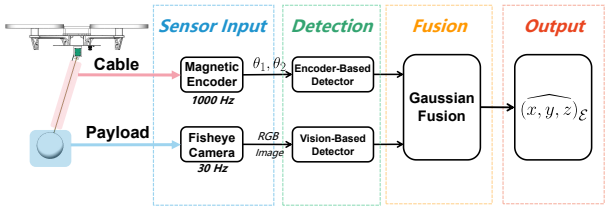

Fig. 5. The framework of the estimation system.

A. Encoder-based estimator

For each magnet encoder, the angle data is obtained through analog-to-digital conversion, with the total range $n_{total} = 2^{14}$. By assuming that the measurement of the whole encoder system is disturbed by white noise, the measured angle can be expressed as

$$\hat{\theta} = 2\pi \frac{n_{adc}}{n_{total}} + w_n, \tag{14}$$

where $n_{adc} \in [0, n_{total}]$ is the raw measured angle, and $w_n \sim \mathcal{N}(0, \sigma_e^2)$.

Given the mechanical design of the encoder system in Section II, the measurements can be expressed as

$$\begin{aligned} \mathbf{x}_L^{\mathcal{E}} &= h(\hat{\theta}) \\ &= h(\theta) + \eta, \end{aligned} \tag{15}$$

where the function $h(\cdot)$ is the measurement function shown in equation (1), $\eta \in \mathbb{R}^3$ is the transformed noise variance. The noise is described by zero-mean Gaussian noise with a covariance matrix of R. For the measurements θ, the conditional probability density is as follows:

$$\begin{aligned} p(\mathbf{x}_L^{\mathcal{E}}|\theta) &= \mathcal{N}(\mathbf{x}_L^{\mathcal{E}}; h(\theta), R) \\ &= \frac{1}{\sqrt{|2\pi R|}} \exp\left(-\frac{1}{2}\|h(\theta) - \mathbf{x}_L^{\mathcal{E}}\|_R^2\right). \end{aligned} \tag{16}$$

B. Vision-based estimator

It is noting that the encoder-based estimator is not able to provide robust estimation of the payload position if the cable is in slack state; this will lead to failure of MAV flight control. To address the problem, we further leverage a monocular fish-eye camera; the camera is not only used to detect the taut state of the cable but also provide the position information of the payload.

To realize robust detection, STAPLE [13] is applied to track the payload, and it utilizes a linear combination of a Correlation Filter (using HOG features) and a global color histogram to construct a score function as follows

$$f(u,v) = \gamma_{tp} f_{tp}(u,v) + \gamma_{ht} f_{ht}(u,v), \tag{17}$$

where $f_{tp}(u,v)$ is the template score computed by the correlation filter; $f_{ht}(u,v)$ is the histogram score; γ_{tp} and γ_{ht} are the weights of score functions. STAPLE has great robustness to challenging situations like motion blur, illumination changes, object deformation and complex background. With those features, our vision system shows great robustness in indoor and outdoor experiments as shown in Fig. 6.

Moreover, benefit from the encoder-based estimator, we give STAPLE the ability to retrieve the lost tracking object. The original STAPLE has a low probability to re-track the object when it is out of the field of view. In our system, if the object tracking failed in the vision-based estimator, it is re-initialized by the estimated position from the encoder-based estimator.

Fig. 6. Tracking experiments in different challenging situations

To judge the taut state of the cable, the estimated position difference $\Delta \mathbf{x}_L = \|\mathbf{x}_{L_{vision}} - \mathbf{x}_{L_{encoder}}\|_2$ between the encoder-based estimator and the vision-based estimator is computed. Define the circumscribed radius of payload is r. When $\Delta \mathbf{x}_L < r$, the rope is taut, otherwise, is slack. This process is necessary for our automatical controller switch strategy detailed in IV-B.

C. Vision-encoder-based sensor fusion

As discussed previously, both encoder and vision-based estimators provides the position estimation of the payload. Because the two estimations suffers from noise interference, it is necessary to fuse both sensor measurements to provide accurate results. Without loss of generality, the position estimation $\mathbf{x}^{\mathcal{B}}_{L_{est}}$ obtained by the entire vision-encoder system is assumed to obey Gaussian distribution

$$P(\mathbf{x}^{\mathcal{B}}_{L_{est}}) = \mathcal{N}(\mu, \sigma^2). \tag{18}$$

And the observations from vision and encoder also obey Gaussian distribution

$$P(\mathbf{x}^{\mathcal{B}}_{L_{obs}}) = \mathcal{N}(\mu_{obs}, \sigma^2_{obs}). \tag{19}$$

According to the product of the Gaussian distribution [14], the parameters in (18) are calculated as

$$\mu_{est} = \frac{\sigma^2_{vision}\mu_{encoder} + \sigma^2_{vision}\mu_{encoder}}{\sigma^2_{vision} + \sigma^2_{encoder}}, \tag{20}$$

$$\sigma^2_{est} = \frac{\sigma^2_{encoder}\sigma^2_{vision}}{\sigma^2_{vision} + \sigma^2_{encoder}}. \tag{21}$$

As estimator frame $\mathcal{E}$ is a polar coordinate defined in II-B, using the equations above in each angle, a fused position estimation is obtained in real-time.

IV. Trajectory Generation and Control

A. Trajectory generation

To ensure the payload move smoothly and efficiently, a minimum snap trajectory [15] is planned as a differentially-flat reference input for the payload. The trajectory is given by piecewise-polynomials:

$$f(t) = \begin{cases} f_1(t) = \sum_{i=0}^{N} p_{1,i} t^i, \ T_0 \le t \le T_1 \\ \cdots \\ f_M(t) = \sum_{i=0}^{N} p_{M,i} t^i, \ T_{M-1} \le t \le T_M, \end{cases} \tag{22}$$

where $T_0, \cdots, T_M$ are the duration for each segment. For dynamic feasibility, continuity constraints and derivative constraints are imposed as

$$f^{(k)}_j(T_j) = f^{(k)}_{j+1}(T_j), \ \begin{cases} f^{(k)}_j(T_{j-1}) = x^{(k)}_{0,j} \\ f^{(k)}_j(T_j) = x^{(k)}_{T,j}, \end{cases} \tag{23}$$

where x_j is the desired waypoint given. To minimize differential thrust, a optimal parameter vector of piecewise-polynomials $\mathbf{p}^* = [p_{1,0}, p_{2,0}, \cdots, p_{M,N}]^T$ can be found by

$$\mathbf{p}^* = \arg\min_{\mathbf{p}} \int_{T_0}^{T_M} (f^{(4)}(t))^2 dt, \tag{24}$$

The optimization problem described above can be optimized by solver [16] to generate the minimum snap trajectory, and obtain the desired payload trajectory $\mathbf{x}^{(k)}_{Ld} = f^{(k)}(t)$.

B. Control design and controller switch strategy

The cable-suspended load is a hybrid system in which two states are determined by the tension in the cable, being either

taut or slack [17]. A payload controller [12] that can stabilize the payload oscillation is utilized in the paper.

1) The cable is taut: Consider the load position dynamics in (10), and consider the computed load attitude defined as

$$\mathbf{q}_c = \frac{A}{\|A\|}, \tag{25}$$

$$\begin{aligned} A = &- k_x \mathbf{e}_{\mathbf{x}_L} - k_v \mathbf{e}_{v_L} \\ &+ (m_Q + m_L)(\ddot{\mathbf{x}}_L^d + g\mathbf{e}_3) + m_Q l(\dot{\mathbf{q}} \cdot \dot{\mathbf{q}})\mathbf{q}, \end{aligned} \tag{26}$$

with tracking error functions in $\mathbb{R}^3$ defined as $\mathbf{e}_{\mathbf{x}_L} = \mathbf{x}_L - \mathbf{x}_{L_d}$, $\mathbf{e}_{v_L} = \dot{\mathbf{x}}_L - \dot{\mathbf{x}}_{L_d}$. By combining the dynamic model and the desired yaw ϕ of the MAV, the computed MAV attitude is acquired by

$$\mathbf{R}_c = [\mathbf{b}_{1_c}, \mathbf{b}_{3_c} \times \mathbf{b}_{1_c}, \mathbf{b}_{3_c}],\ \hat{\mathbf{\Omega}}_c = \mathbf{R}_c^T \dot{\mathbf{R}}_c, \tag{27}$$

where $\mathbf{b}_{1_c}$, $\mathbf{b}_{3_c} \in S^2$ is defined by

$$\mathbf{b}_{1_d} = [\cos\phi, \sin\phi, 0]^T, \tag{28}$$

$$\mathbf{b}_{3_c} = \frac{F}{||F||},\ \mathbf{b}_{1_c} = \frac{(\mathbf{b}_{3_c} \times \mathbf{b}_{1_d}) \times \mathbf{b}_{3_c}}{||\mathbf{b}_{3_c} \times \mathbf{b}_{1_d}||}, \tag{29}$$

where,

$$F = F_n - F_{pd} - F_{ff}, \tag{30}$$

$$F_n = (A \cdot \mathbf{q})\mathbf{q}, \tag{31}$$

$$F_{pd} = -k_q \mathbf{e}_q - k_\omega \mathbf{e}_{\dot{q}}, \tag{32}$$

$$F_{ff} = m_Q l\big(\langle \mathbf{q}, \mathbf{q}_d \times \dot{\mathbf{q}}_d \rangle (\mathbf{q} \times \dot{\mathbf{q}}) + (\mathbf{q}_d \times \ddot{\mathbf{q}}_d) \times \mathbf{q}\big), \tag{33}$$

where orientation error functions on S^2 are defined as $\mathbf{e}_q = \hat{\mathbf{q}}^2 \mathbf{q}_d, \mathbf{e}_{\dot{q}} = \dot{\mathbf{q}} - (\mathbf{q}_d \times \dot{\mathbf{q}}_d) \times \mathbf{q}$. Then replace the desired quadrotor and load attitude by their computed values, $\mathbf{R}_c$, $\mathbf{\Omega}_c$ and $\mathbf{q}_c$ respectively [12]. Define quadrotor thrust f as:

$$f = F \cdot \mathbf{R}\mathbf{e}_3. \tag{34}$$

According to the quadrotor's attitude, the configuration error functions [18] are given by $\mathbf{e}_R = \frac{1}{2}(\mathbf{R}_d^T \mathbf{R} - \mathbf{R}^T \mathbf{R}_d)^\vee$, $\mathbf{e}_{\mathbf{\Omega}} = \mathbf{\Omega} - \mathbf{R}^T \mathbf{R}_d \mathbf{\Omega}_d$. The moment defined as

$$\begin{aligned} \mathbf{M} = &- k_R \mathbf{e}_R - k_\Omega \mathbf{e}_\Omega + \mathbf{\Omega} \times J_Q \mathbf{\Omega} \\ &- J_Q(\hat{\mathbf{\Omega}} \mathbf{R}^T \mathbf{R}_d \mathbf{\Omega}_d - \mathbf{R}^T \mathbf{R}_d \dot{\mathbf{\Omega}}_d). \end{aligned} \tag{35}$$

The closed-loop payload controller is exponentially stable and enables the quadrotor with a suspended load to track an aggressive trajectory.

2) The cable is slack: Both of the estimators will bring wrong feedback as the taut-rope assumption is not meet. And the payload has no force to the MAV. The MAV-payload controller above will generate wrong feedforward force (26) as shown in Fig. 7(b), which ultimately leads to the crash of the aircraft. To avoid this, our system will automatically switch to a normal quadrotor controller [19] which is proofed to be exponential asymptotic stable even in aggressive pitch and roll angle. The MAV-payload system degenerates into an ordinary MAV system as shown in Fig. 7(c)(d).

V. Experiments

Several experiments were performed to verify the proposed system. The experimental MAV is a quadrotor with an onboard computer running robot operation system (ROS) in Linux and some basic parameters are given in Table II.

TABLE II

MAV Diameter	450mm
MAV weight	1.96kg
Rope length	570mm
Payload weight	0.168kg
Onboard computer	intel i7-5557U, no GPU
Fish-eye camera	HBV-1466, 640*480 30FPS 160°

The static noise of the vision-based and encoder-based estimators are obtained with experimental tests, separately, and the results are shown in Table III. θ_1, θ_2 are defined in II-B.

TABLE III

	σ_{θ_1}	σ_{θ_2}
Encoder-based	8.55×10^{-4}	1.20×10^{-3}
Vision-based	3.66×10^{-2}	4.04×10^{-2}

All the algorithms were implemented in the onboard computer. The pose of the MAV ($\mathbf{R}_{\mathcal{B}}^{\mathcal{W}}, \mathbf{x}_Q^{\mathcal{W}}$) was obtained from the VICON motion capture system and was sent to the onboard computer by Wi-Fi using TCP/IP at 100Hz. And the position of the payload was fully estimated by the vision-encoder system at 100Hz; it can be higher frequency up to 1000Hz. The ground truth of payload position ($\mathbf{x}_{L_{gt}}^{\mathcal{B}}$) in $\mathcal{B}$ frame and the estimated payload position ($\mathbf{x}_{L_{est}}^{\mathcal{W}}$) in frame $\mathcal{W}$ are calculated by

$$\begin{cases} \mathbf{x}_{L_{gt}}^{\mathcal{B}} = \mathbf{R}_{\mathcal{E}}^{\mathcal{B}} \mathbf{R}_{\mathcal{W}}^{\mathcal{E}} (\mathbf{x}_{L_{gt}}^{\mathcal{W}} - \mathbf{x}_{Q_{gt}}^{\mathcal{W}}) \\ \mathbf{x}_{L_{est}}^{\mathcal{W}} = \mathbf{x}_Q^{\mathcal{W}} + \mathbf{R}_{\mathcal{B}}^{\mathcal{W}} \mathbf{R}_{\mathcal{E}}^{\mathcal{B}} \mathbf{x}_{L_{est}}^{\mathcal{E}}. \end{cases} \tag{36}$$

A. Estimation results

The first experiment was performed to verify the accuracy, rapidity, and low latency of our payload position estimator. We make the vision-encoder-based suspended payload position estimator to move a path of our lab name NRSL as shown in Fig. 8. The results of $\mathbf{x}_L^{\mathcal{B}}$ is illustrated in Fig. 9, where the ground truth is obtained from the VICON motion capture system. The estimated data is the fused output of the proposed vision-encoder system.

The root mean square error is used to evaluate the estimation accuracy of the estimator during the entire process.

$$RMSE = \sqrt{\frac{\sum_{i=0}^{n} ||\mathbf{x}_i^{gt} - \mathbf{x}_i^{est}||}{n}},$$

where $\mathbf{x}_i^{(\cdot)}$ is the payload position in frame $\mathcal{B}$ at time i. The results are given in Table IV.

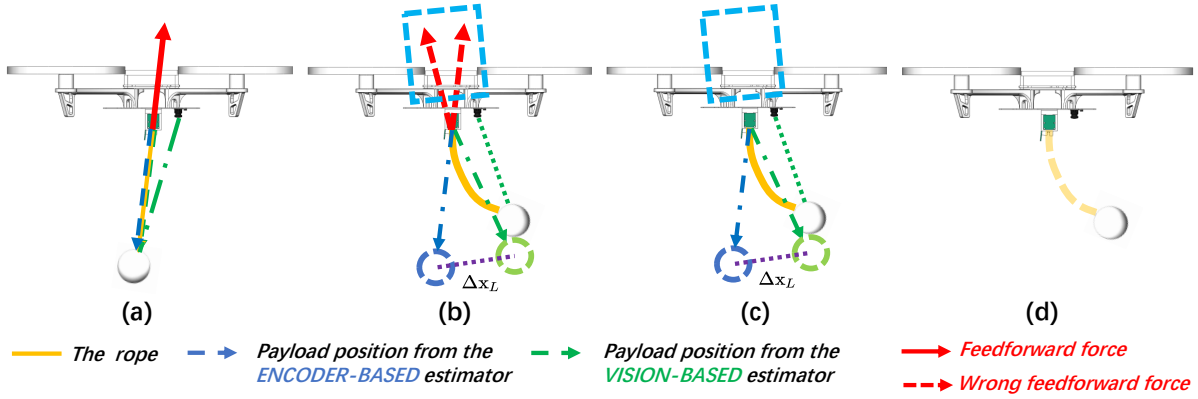

Fig. 7. (a) The cable is taut; (b) The cable is slack and the MAV-payload controller generates wrong force; (c) The automatical controller switch strategy works; (d) The MAV-payload system degerates to a normal quadrotor system.

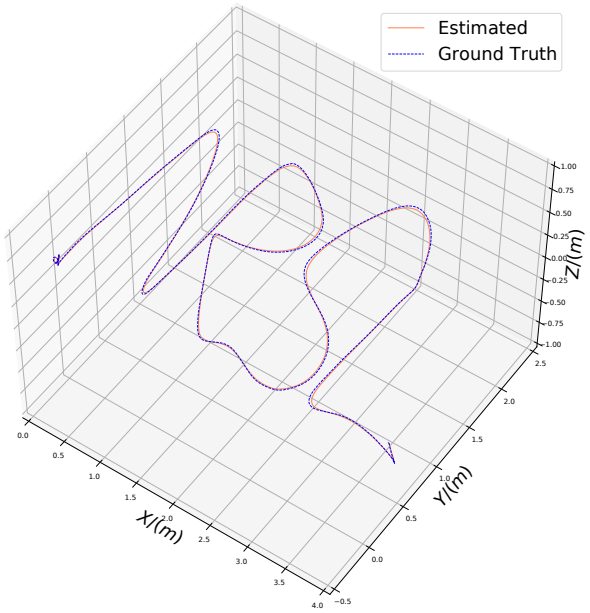

Fig. 8. Trajectory of the payload in the frame $\mathcal{W}$ in NRSL flight experiment.

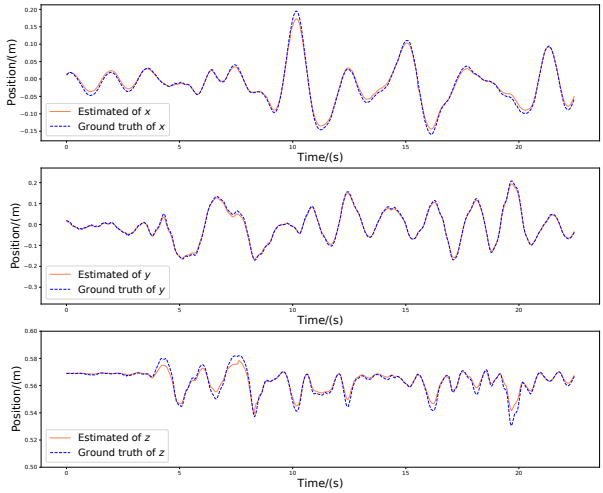

Fig. 9. Position of the payload in the frame $\mathcal{B}$ on each axis in NRSL flight experiment.

TABLE IV

TRACKING ERROR ON EACH AXIS

	x/m	y/m	z/m
Tracking Error	6.718e-03	5.593e-03	2.485e-03

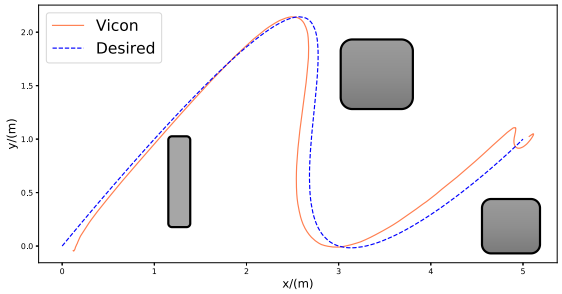

Fig. 10. Flight result I: Desired trajectory and actual trajectory of the payload in the coordinate $\mathcal{W}$. The gray boxes are obstacles in configuration space.

B. Flight result I: payload trajectory tracking

The proposed system was further tested by tracking a minimum snap trajectory; the experiment aims to verify the ability of our system about driving a suspended payload to stably track a desired trajectory. The trajectory is constrained to pass through the given waypoints at a constant height, $0.58\,m$. Each duration between waypoints is allocated by trapezoidal velocity time profile and the total executing time is $11\,s$. In the whole process, the cable remains taut.

Fig. 10 compares the desired payload trajectory with the actual payload trajectory. The tracking position error computed by $\|\mathbf{x}_L - \mathbf{x}_L^d\|_2$ is $0.114\,m$ in average. It is seen that the actual payload trajectory is approximate to the desired payload trajectory with a low tracking position error. The tracking error increases when the system tries to stabilize the payload at a high velocity of $2.55\,m/s$.

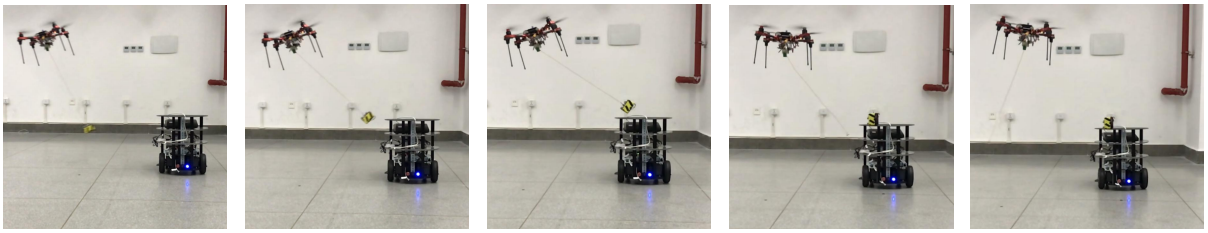

Fig. 11. Flight result II: Snapshots of the autonomous MAV dropping the payload to a moving vehicle with a maximum swing angle over $59°$.

C. Flight result II: drop the payload to a moving vehicle

To verify the practicability of our estimation system and control algorithm, a smooth trajectory with an aggressive swing angle was generated to enable the MAV to drop the payload to a continuously moving vehicle. Fig. 11 illustrates the snapshots of payload dropping. When the suspended payload was released, the visual sensor detected the abnormal movement of the payload and judged that the cable is not tensioned. And then the MAV switched the controller to ensure that the aircraft was not disturbed by incorrect payload state estimation. The maximum speed of the MAV is $2.84\,m/s$ with the acceleration over $13.5\,m/s^2$, and the swing angle when dropping the payload is over $59°$. Our planned trajectory made the payload have zero acceleration and zero speed when disconnected; the cargo delivery was safe and stable.

VI. Conclusion

The paper presents an efficient and robust vision-encoder-based estimator for the estimation of the payload state. To enhance the maneuverability of the MAV, a practical evaluation of the rope's taut state by using a vision-based estimator is applied. The entire systems with state estimation, trajectory planning and flight control for a MAV with a suspended payload is also provided in the paper. Experiments are performed for validation of the proposed approach. With the newly developed approach, a more agile movement for a MAV with suspended payload is able to perform in the future. Furthermore, a more straightforward method using extra sensors like force sensor to judge if the rope is taut can be applied to reduce dependence on the visual system.

References

[1] Markus Bernard and Konstantin Kondak. Generic slung load transportation system using small size helicopters. In *2009 IEEE International Conference on Robotics and Automation*, pages 3258–3264. IEEE, 2009.

[2] Paul EI Pounds, Daniel R Bersak, and Aaron M Dollar. Stability of small-scale uav helicopters and quadrotors with added payload mass under pid control. *Autonomous Robots*, 33(1):129–142, 2012.

[3] Ivana Palunko, Patricio Cruz, and Rafael Fierro. Agile load transportation: Safe and efficient load manipulation with aerial robots. *IEEE robotics & automation magazine*, 19(3):69–79, 2012.

[4] Byung-Yoon Lee, Hae-In Lee, Dong-Wan Yoo, Gun-Hee Moon, Dong-Yeon Lee, Yun young Kim, and Min-Jea Tahk. Study on payload stabilization method with the slung-load transportation system using a quad-rotor. In *2015 European Control Conference (ECC)*, pages 2097–2102. IEEE, 2015.

[5] Sarah Tang and Vijay Kumar. Mixed integer quadratic program trajectory generation for a quadrotor with a cable-suspended payload. In *2015 IEEE international conference on robotics and automation (ICRA)*, pages 2216–2222. IEEE, 2015.

[6] Cedric De Crousaz, Farbod Farshidian, Michael Neunert, and Jonas Buchli. Unified motion control for dynamic quadrotor maneuvers demonstrated on slung load and rotor failure tasks. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2223–2229. IEEE, 2015.

[7] Philipp Foehn, Davide Falanga, Naveen Kuppuswamy, Russ Tedrake, and Davide Scaramuzza. Fast trajectory optimization for agile quadrotor maneuvers with a cable-suspended payload. In *Robotics: Science and Systems (RSS) 2017*, 2017.

[8] Ivana Palunko, Aleksandra Faust, Patricio Cruz, Lydia Tapia, and Rafael Fierro. A reinforcement learning approach towards autonomous suspended load manipulation using aerial robots. In *2013 IEEE international conference on robotics and automation*, pages 4896–4901. IEEE, 2013.

[9] Sarah Tang, Valentin Wüest, and Vijay Kumar. Aggressive flight with suspended payloads using vision-based control. *IEEE Robotics and Automation Letters*, 3(2):1152–1159, 2018.

[10] Seung Jae Lee and H Jin Kim. Autonomous swing-angle estimation for stable slung-load flight of multi-rotor uavs. In *2017 IEEE international conference on robotics and automation (ICRA)*, pages 4576–4581. IEEE, 2017.

[11] Juho Kannala and Sami S Brandt. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. *IEEE transactions on pattern analysis and machine intelligence*, 28(8):1335–1340, 2006.

[12] Koushil Sreenath, Taeyoung Lee, and Vijay Kumar. Geometric control and differential flatness of a quadrotor uav with a cable-suspended load. In *52nd IEEE Conference on Decision and Control*, pages 2269–2274. IEEE, 2013.

[13] Luca Bertinetto, Jack Valmadre, Stuart Golodetz, Ondrej Miksik, and Philip HS Torr. Staple: Complementary learners for real-time tracking. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1401–1409, 2016.

[14] Norbert Freier and TU ITVS. Kalman filter: Multiplying normal distributions. *TU Dresden*, 2013.

[15] Daniel Mellinger and Vijay Kumar. Minimum snap trajectory generation and control for quadrotors. In *2011 IEEE international conference on robotics and automation*, pages 2520–2525. IEEE, 2011.

[16] E Michael Gertz and Stephen J Wright. Object-oriented software for quadratic programming. *ACM Transactions on Mathematical Software (TOMS)*, 29(1):58–81, 2003.

[17] Patricio J Cruz, Meeko Oishi, and Rafael Fierro. Lift of a cable-suspended load by a quadrotor: A hybrid system approach. In *2015 American control conference (ACC)*, pages 1887–1892. IEEE, 2015.

[18] Francesco Bullo and Andrew D Lewis. *Geometric control of mechanical systems: modeling, analysis, and design for simple mechanical control systems*, volume 49. Springer, 2019.

[19] Taeyoung Lee, Melvin Leok, and N Harris McClamroch. Geometric tracking control of a quadrotor uav on se (3). In *49th IEEE conference on decision and control (CDC)*, pages 5420–5425. IEEE, 2010.