IEEE Robotics and Automation Letters (RAL) paper presented at
2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 14-18, 2024. Abu Dhabi, UAE

Image-Based Time-Varying Contact Force Control of Aerial Manipulator Using Robust Impedance Filter

Jeonghyun Byun, *Graduate Student Member, IEEE*, Junha Kim, Dohyun Eom, Dongjae Lee, Changhyeon Kim, and H. Jin Kim, *Member, IEEE*

***Abstract*—The use of aerial manipulators for safe and efficient physical interaction with their surrounding environments has been gaining attention within the aerial robotics research community. In this letter, we present an image-based time-varying force tracking controller for an aerial manipulator conducting forceful interaction with a static surface. To this end, we first extract visual features from the surface using a monocular camera and calculate image feature vectors for the rotational and translational movements of the camera. Then, the RISATE (Robust Integral of SATuration Error)-based impedance filter continuously updates the desired values of the image features based on the previously designated force profile. Our stability analysis verifies that the error between the desired and actual contact force is uniformly ultimately bounded in an arbitrarily small bound with the proposed strategy even with the switching between free-flight and contact modes. Moreover, through time-varying force-tracking experiments with a quadrotor-based aerial manipulator, we validate the reproducibility and improved force-tracking performance of the proposed method.**

***Index Terms*—Aerial systems: applications, aerial systems: mechanics and control, force control.**

I. Introduction

An AERIAL manipulator that can interact with structures located in hard-to-reach areas such as walls or windows installed on tall structures and high ceilings has been gaining significant attention in the aerial robotics research community. There have been preliminary works involving aerial physical interaction, for example, push-and-slide operation [1], window-cleaning [2], and aerial drilling [3]. To guarantee the safe operation of the aerial physical interaction, it is necessary to regulate the contact force exerted on the aerial manipulator's end-effector. Specifically, there needs the precise tracking of time-varying desired force profiles in situations such as tracking of a force reference generated by a human operator during teleoperation [4], exerting a variable force to carry an object during cooperative payload transportation [5], and extracting a wedged object from a static structure by gradually increasing a pulling force [6].

To conduct such tasks, the pose of a contact surface needs to be given before the operation. Since there might be a lack of prior information on the surface in real applications, we need to measure or estimate the surface pose using onboard sensors such as cameras and LiDAR (Light Detection And Ranging). However, only few studies consider a time-varying motion/force tracking control using the feedback from the onboard sensors during aerial physical interaction.

There had been several works on the force-tracking control during aerial physical interaction. [7] and [8] presented methods to track the constant force reference trajectory for an aerial manipulator. Meanwhile, in [9], [10], [11], time-varying force tracking controllers were introduced. However, those works require prior information on the pose of the surface where the aerial manipulator exerts force.

There have been several studies on the design of a force-tracking controller for an aerial manipulator using a camera sensor. [12] and [13] proposed image-based motion/force controllers using onboard RGB-D and Time-of-Flight (ToF) cameras, respectively. In [14], the authors employed a monocular camera to conduct the force-tracking using the aerial manipulator. However, the above works were still limited to the constant desired force. Even though [15] treated the tracking of *non*-constant desired force with a monocular camera, it focused only on maintaining flight stability during the contact, not on enhancing the force-tracking performance.

In this paper, we propose an image-based visual servoing (IBVS) method to make the force exerted on the end-effector track a time-varying force reference trajectory with no prior information on the contact surface's pose. We use a monocular camera due to its popularity with low computational burden, cost-effectiveness, and lightweight. Vision-based control designed for this camera can be expanded to other sensors that can easily gather depth information such as stereo or depth cameras. Assuming that four point features can be extracted and tracked from the image of a visual object on a planar surface by existing computer-vision methods [16], we calculate image feature vectors that represent the translational and rotational movements of the camera mounted on the aerial manipulator and formulate an IBVS algorithm that enables the tracking of time-varying force profiles. Particularly, our main contributions are arranged as follows:

Manuscript received 3 December 2023; accepted 22 March 2024. Date of publication 2 April 2024; date of current version 16 April 2024. This letter was recommended for publication by Associate Editor C. Gabellieri and Editor G. Loianno upon evaluation of the reviewers' comments. This work was supported in part by the National Research Foundation of Korea (NRF) through Unmanned Vehicles Core Technology Research and Development Program and in part by the Unmanned Vehicle Advanced Research Center (UVARC) funded by the Ministry of Science and ICT, Republic of Korea, under Grant NRF-2020M3C1C1A010864. *(Corresponding author: H. Jin Kim.)*

Jeonghyun Byun, Junha Kim, Dohyun Eom, Dongjae Lee, and H. Jin Kim are with the Department of Aerospace Engineering, Automation and System Research Institute(ASRI) and Institute of Advanced Aerospace Technology(IAAT), Seoul National University, Seoul 08826, South Korea (e-mail: quswjdgus97@snu.ac.kr; wnsgk02@snu.ac.kr; djaehgus@snu.ac.kr; ehdwo713@snu.ac.kr; hjinkim@snu.ac.kr).

Changhyeon Kim is with Samsung Research, Seoul 06765, South Korea (e-mail: rlackd93@snu.ac.kr).

This letter has supplementary downloadable material available at https://doi.org/10.1109/LRA.2024.3382963, provided by the authors.

Digital Object Identifier 10.1109/LRA.2024.3382963

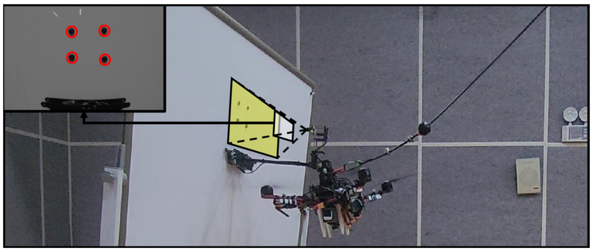

Fig. 1. Aerial manipulator conducting an image-based force tracking control on a whiteboard.

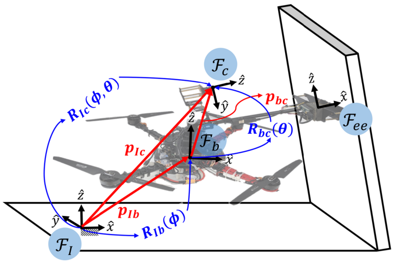

Fig. 2. Configuration of an aerial manipulator with the key coordinate frames, displacement vectors, and rotational matrices.

- We design a monocular-camera-based IBVS method that makes the force exerted on the surface track the *non*-constant desired force profile.
- We theoretically prove that the error between the desired and measured force is uniformly ultimately bounded within an arbitrarily small bound despite the errors in the image feature vectors while considering the switching behavior between free-flight and contact modes, not just the steady-state behavior.
- We validate the reproducibility and improved force-tracking performance of our controller through comparative experiments using a quadrotor-based aerial manipulator.

This paper is outlined as follows; In Section II, we describe an overview of an aerial manipulator system, and Section III depicts the design of two image feature vectors that represent the translation and rotational parts of the monocular camera's movement. In Section IV, we explain the overall framework of our image-based force-tracking control strategy. We conduct the stability analysis on the given system in Section V, and Section VI presents the experimental validation of the proposed control law, followed by conclusion in Section VII.

Notations: For a vector $\boldsymbol{\alpha}$, we let α_i denote the i-th element of $\boldsymbol{\alpha}$. Also, for $\boldsymbol{\alpha_1}$ and $\boldsymbol{\alpha_2} \in \mathbb{R}^3$, we let $[\boldsymbol{\alpha_1}]_\times \in \mathbb{R}^{3\times 3}$ denote the operator which maps $\boldsymbol{\alpha_1}$ into a skew-symmetric matrix such as $[\boldsymbol{\alpha_1}]_\times \boldsymbol{\alpha_2} = \boldsymbol{\alpha_1} \times \boldsymbol{\alpha_2}$. For scalars a_0, a_1, …, a_n, we use a semicolon as $[a_0\ a_1 \cdots a_n]^\top = [a_0; a_1; \cdots; a_n]$ to concatenate them into a column matrix, and we define $\mathrm{diag}\{a_0, a_1, \ldots, a_n\}$ as a diagonal matrix where the i-th diagonal element is a_i. We let $\boldsymbol{0_{ij}}$ and $\boldsymbol{I_n}$ denote $i \times j$ zero matrix and $n \times n$ identity matrix, and abbreviate the phrase "with respect to" to w.r.t..

II. AERIAL MANIPULATOR SYSTEM OVERVIEW

This section presents the configuration of an aerial manipulator system shown in Fig. 2 and the forward kinematics from the twist of the multirotor and the joint velocities to the twist of the camera.

A. Aerial Manipulator Configuration

The aerial manipulator utilized in this work consists of a multirotor and a n_a-degree-of-freedom (DOF) robotic arm where $n_a (\geq 2)$ represents the number of the arm's actuators. The generalized coordinate variables of the aerial manipulator consist of the position of the multirotor w.r.t. $\mathcal{F}_I$, $\boldsymbol{p_{Ib}} \in \mathbb{R}^3$, the Euler angles of the multirotor, $\boldsymbol{\phi} \in \mathbb{R}^3$, and the joint angles of the robotic arm, $\boldsymbol{\theta} \in \mathbb{R}^{n_a}$.

As depicted in Fig. 2, we let $\mathcal{F}_I$, $\mathcal{F}_b$, $\mathcal{F}_c$ and $\mathcal{F}_{ee}$ denote the inertial, multirotor body, monocular camera, and end-effector coordinate frames, respectively. For two frames $\mathcal{F}_F$ and $\mathcal{F}_{F'}$ $(F, F' \in \{I, b, c, ee\})$, we let $\boldsymbol{p_{FF'}}$ denote the position of $\mathcal{F}_{F'}$ w.r.t. $\mathcal{F}_F$, and the linear and angular velocities of $\mathcal{F}_{F'}$ w.r.t. $\mathcal{F}_F$ expressed in $\mathcal{F}_{F'}$ are defined as $\boldsymbol{\dot{p}^{F'}_{FF'}}$ and $\boldsymbol{\omega^{F'}_{FF'}}$, respectively. Moreover, we define $\boldsymbol{R_{FF'}} \in SO(3)$ as the rotation matrix from $\mathcal{F}_{F'}$ to $\mathcal{F}_F$.

B. Forward Kinematics

As shown in Fig. 2, the attitude and position of the camera, $\boldsymbol{R_{Ic}}$ and $\boldsymbol{p_{Ic}}$, are expressed w.r.t. $\boldsymbol{p_{Ib}}$, $\boldsymbol{p_{bc}}$, $\boldsymbol{R_{Ib}}$ and $\boldsymbol{R_{bc}}$ as follows:

$$\boldsymbol{R_{Ic}} = \boldsymbol{R_{Ib}}\boldsymbol{R_{bc}},\ \boldsymbol{p_{Ic}} = \boldsymbol{p_{Ib}} + \boldsymbol{R_{Ib}}\boldsymbol{p_{bc}}. \tag{1}$$

Since $\boldsymbol{\dot{R}_{FF'}} = \boldsymbol{R_{FF'}}[\boldsymbol{\omega^{F'}_{FF'}}]_\times$, we obtain $[\boldsymbol{\omega^{c}_{Ic}}]_\times = [\boldsymbol{R^\top_{bc}}\boldsymbol{\omega^{b}_{Ib}}]_\times + [\boldsymbol{\omega^{c}_{bc}}]_\times$. Hence, $\boldsymbol{\omega^{c}_{Ic}}$ is expressed as follows:

$$\boldsymbol{\omega^{c}_{Ic}} = \boldsymbol{R^\top_{bc}}\boldsymbol{\omega^{b}_{Ib}} + \boldsymbol{\omega^{c}_{bc}}. \tag{2}$$

because the operator $[\cdot]_\times$ is a closed set w.r.t. the linear combination. Since there exist functions $\boldsymbol{Q_{rp}}(\phi_1, \phi_2) \in \mathbb{R}^{3\times 2}$, $\boldsymbol{Q_y} \in \mathbb{R}^3$, $\boldsymbol{J^c_{bc}}(\boldsymbol{\theta}) \in \mathbb{R}^{3\times n_a}$ such that $\boldsymbol{\omega^{b}_{Ib}} = \boldsymbol{Q_{rp}}[\dot{\phi}_1; \dot{\phi}_2] + \boldsymbol{Q_y}\dot{\phi}_3$ and $\boldsymbol{\omega^{c}_{bc}} = \boldsymbol{J^c_{bc}}\boldsymbol{\dot{\theta}}$, (2) is rearranged as follows:

$$\boldsymbol{\omega^{c}_{Ic}} = \boldsymbol{R^\top_{bc}}\boldsymbol{Q_{rp}}[\dot{\phi}_1; \dot{\phi}_2] + \boldsymbol{R^\top_{bc}}\boldsymbol{Q_y}\dot{\phi}_3 + \boldsymbol{J^c_{bc}}\boldsymbol{\dot{\theta}}. \tag{3}$$

Also, since there exist a function $\boldsymbol{J_{t,bc}}(\boldsymbol{\theta}) \in \mathbb{R}^{3\times n_a}$ such that $\boldsymbol{\dot{p}^{b}_{bc}} = \boldsymbol{J_{t,bc}}(\boldsymbol{\theta})\boldsymbol{\dot{\theta}}$, the calculation of $\boldsymbol{\dot{p}^{c}_{Ic}} = \boldsymbol{R^\top_{Ic}}\boldsymbol{\dot{p}^{I}_{Ic}} = \boldsymbol{R^\top_{bc}}\boldsymbol{R^\top_{Ib}}\boldsymbol{\dot{p}^{I}_{Ic}}$ is arranged as follows:

$$\begin{aligned}\boldsymbol{\dot{p}^{c}_{Ic}} &= \boldsymbol{R^\top_{bc}}\boldsymbol{R^\top_{Ib}}\boldsymbol{\dot{p}^{I}_{Ib}} - \boldsymbol{R^\top_{bc}}[\boldsymbol{p_{bc}}]_\times\boldsymbol{Q_{rp}}[\dot{\phi}_1; \dot{\phi}_2] \\ &\quad - \boldsymbol{R^\top_{bc}}[\boldsymbol{p_{bc}}]_\times\boldsymbol{Q_y}\dot{\phi}_3 + \boldsymbol{R^\top_{bc}}\boldsymbol{J_{t,bc}}\boldsymbol{\dot{\theta}}.\end{aligned} \tag{4}$$

III. DESIGN OF IMAGE FEATURE VECTORS

In Fig. 3, the coordinates of four point features projected on the normalized image plane are calculated as follows:

$$x_i = \frac{p_{ci,1}}{p_{ci,3}}, \quad y_i = \frac{p_{ci,2}}{p_{ci,3}}, \quad i = 1, \ldots, 4 \tag{5}$$

where we let $\boldsymbol{p_{ci}} \in \mathbb{R}^3$ denote the position of the i^{th} point w.r.t. the camera frame's origin.

The image obtained through the lens of the camera can be viewed as an image feature that has information on the relative pose of the camera w.r.t. the surface. We can derive a vector $\boldsymbol{s_p} \in \mathbb{R}^6$ that represents the 6-DOF movement of the camera utilizing the coordinates of the four projected points. Also, since the camera's 6-DOF motion can be divided into the rotational and translational components, $\boldsymbol{s_p}$ can also be decoupled into two three-dimensional vectors: one for the rotational part and another for the translational part.

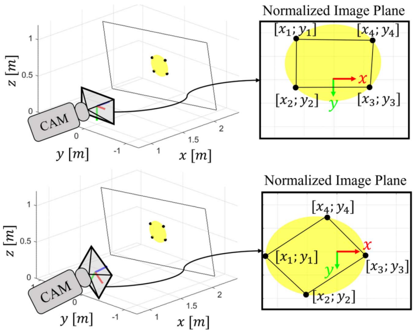

Fig. 3. Illustrations of four point features projected on the normalized image plane obtained from two camera orientations.

In the following subsections, the definitions of the *rotational image feature vector*, $\boldsymbol{s_r} \in \mathbb{R}^3$, and *translational image feature vector*, $\boldsymbol{s_t} \in \mathbb{R}^3$, will be presented.

A. Rotational Image Feature Vector

As shown in Fig. 3, if we change the attitude of the camera w.r.t. the surface, the angles formed between the x or y axis of the image plane and the lines connecting the points also change. Hence, $\boldsymbol{s_r}$ is defined as follows to include the information on the aforementioned angles [16]:

$$\begin{aligned} s_{r,1} &\triangleq \frac{1}{2}\left(\tan^{-1}\left(\frac{x_2-x_1}{y_1-y_2}\right) - \tan^{-1}\left(\frac{x_3-x_4}{y_4-y_3}\right)\right) \\ s_{r,2} &\triangleq \frac{1}{2}\left(\tan^{-1}\left(\frac{y_3-y_2}{x_3-x_2}\right) - \tan^{-1}\left(\frac{y_4-y_1}{x_4-x_1}\right)\right) \\ s_{r,3} &\triangleq \frac{1}{2}\left(\tan^{-1}\left(\frac{y_1-y_4}{x_4-x_1}\right) + \tan^{-1}\left(\frac{y_2-y_3}{x_3-x_2}\right)\right). \end{aligned} \tag{6}$$

Also, $\dot{\boldsymbol{s}}_{\boldsymbol{r}}$ is calculated as follows:

$$\dot{\boldsymbol{s}}_{\boldsymbol{r}} = \boldsymbol{L_{rr}}\boldsymbol{\omega}_{\boldsymbol{Ic}}^{\boldsymbol{c}} \tag{7}$$

where the detailed expression of $\boldsymbol{L_{rr}} \in \mathbb{R}^{3\times 3}$ can be found in [16]. Note that the vector $\boldsymbol{s_r}$ is solely related to the rotational movement of the camera.

B. Translational Image Feature Vector

According to [17], image moments are recognized as practical visual features that demonstrate the translational movement of a planar object. Using this approach, we formulate $\boldsymbol{s_t}$ as follows:

$$\boldsymbol{s_t} = \frac{C}{2\sqrt{n_{20}+n_{02}}}[x_g; y_g; 1] \tag{8}$$

where $x_g \triangleq \Sigma_{i=1}^4 x_i$, $y_g \triangleq \Sigma_{i=1}^4 y_i$ and $n_{lm} \triangleq \frac{1}{4}\Sigma_{i=1}^4 (x_i - x_g)^l (y_i - y_g)^m$. Also, C equals $2l_{cee}\sqrt{n_{20}(t_0) + n_{02}(t_0)}$ where t_0 and l_{cee} represent the time instant of the first contact and the shortest distance from $\mathcal{F}_c$ to the end-effector, respectively.

According to [17], the relation between $\dot{\boldsymbol{s}}_{\boldsymbol{t}}$ and the camera twist $[\dot{\boldsymbol{p}}_{\boldsymbol{Ic}}^{\boldsymbol{c}}; \boldsymbol{\omega}_{\boldsymbol{Ic}}^{\boldsymbol{c}}]$ is calculated as follows:

$$\dot{\boldsymbol{s}}_{\boldsymbol{t}} = -\dot{\boldsymbol{p}}_{\boldsymbol{Ic}}^{\boldsymbol{c}} + \boldsymbol{L_{tr}}\boldsymbol{\omega}_{\boldsymbol{Ic}}^{\boldsymbol{c}} \tag{9}$$

where $\boldsymbol{L_{tr}} \in \mathbb{R}^{3\times 3}$ can be found in [14].

IV. CONTROL FRAMEWORK

In this section, a control framework shown in Fig. 4 will be explained. At first, we design the strategy for reference trajectory generation of $\boldsymbol{s_r}$ and $\boldsymbol{s_t}$. Then, the IBVS algorithm is formulated to calculate the desired twist of the camera. Finally, we derive the inverse kinematics from the desired twist of the camera to the desired velocity and yaw rate of the multirotor and joint velocities based on the forward kinematics introduced in Section II-B.

A. Reference Trajectory Generation for $\boldsymbol{s_r}$ and $\boldsymbol{s_t}$

1) Rotational Part: $\boldsymbol{s_{r,d}}$ is differently calculated depending on whether the configuration of four point features is known in advance or not.

- *Known:* $\boldsymbol{s_{r,d}}$ is determined based on the formation of the four point features. For instance, if four markers on the surface form a rectangle, then $\boldsymbol{s_{r,d}} = \boldsymbol{0_{31}}$.
- *Unknown:* Assuming that the localization of the camera's pose is feasible, the attitude of the surface can be estimated by the triangulation method introduced in [16]. Since we can align the end-effector with the surface based on the attitude of the surface before contact, $\boldsymbol{s_{r,d}}$ is set as the value of $\boldsymbol{s_r}$ when the camera's image plane becomes parallel to the contact surface.

2) Translational Part: When the camera plane is parallel to the object plane, (9) becomes $\dot{\boldsymbol{s}}_{\boldsymbol{t}} = -\dot{\boldsymbol{p}}_{\boldsymbol{Ic}}^{\boldsymbol{c}}$. If we let $\boldsymbol{s_{t,d}}$ as the desired value of $\boldsymbol{s_t}$, since $\dot{p}_{Ic,3}^c$ is related to the camera's movement normal to the contact surface, $s_{t,d,3}$ needs to be updated by an additional impedance filter to make the contact force track the desired force profile. On the other hand, because the camera can freely move in the tangential direction of the surface, we can arbitrarily set $[s_{t,d,1}; s_{t,d,2}]$.

Let $f_d(t)$ denote the reference trajectory of the contact force which is $\mathcal{C}^4$ w.r.t. time. Also, a saturation function $\text{sat}_\varepsilon(\cdot)$ is defined as follows:

$$\text{sat}_\varepsilon(\delta) = \begin{cases} \frac{\delta}{\varepsilon}, & |\delta| \le \varepsilon \\ \frac{\delta}{|\delta|}, & |\delta| > \varepsilon \end{cases} \tag{10}$$

where ε is a small positive constant. Then, the *RISATE (Robust Integral SATuration of the Error)-based impedance filter* updates $s_{z,d} \triangleq s_{t,d,3}$ to make f track f_d as follows:

$$\begin{aligned} &m_d\ddot{s}_{z,d} + b_d\dot{s}_{z,d} - k_1(b_d + k_1 m_d)(s_{z,d} - s_z) \\ &= -k(z_f(t) - z_f(t_0)) - \int_{t_0}^{t}(k\lambda_2 z_f(\tau) + 2\lambda_3 \text{sat}_\varepsilon(z_f(\tau)))d\tau, \\ &\qquad s_{z,d}(t_0) = s_z(t_0) \end{aligned} \tag{11}$$

where $z_f(t) \triangleq \dot{e}_f + \lambda_1 e_f$ and $e_f(t) \triangleq f_d - f$ with the desired mass m_d, desired damping b_d, desired stiffness $k_1(b_d + k_1 m_d)$ and filter gains [18] k, λ_1, λ_2 and λ_3. Particularly, λ_3 is related to the force-tracking speed.

B. Image-Based Visual Servoing (IBVS)

The desired camera twist $[\dot{\boldsymbol{p}}_{\boldsymbol{Ic,d}}^{\boldsymbol{c}}; \boldsymbol{\omega}_{\boldsymbol{Ic,d}}^{\boldsymbol{c}}]$ is generated as follows:

$$\begin{aligned} \boldsymbol{\omega}_{\boldsymbol{Ic,d}}^{\boldsymbol{c}} &= \boldsymbol{L}_{\boldsymbol{rr}}^{-1}(\dot{\boldsymbol{s}}_{\boldsymbol{r,d}} + \boldsymbol{K_r}(\boldsymbol{s_{r,d}} - \boldsymbol{s_r})) \\ \dot{\boldsymbol{p}}_{\boldsymbol{Ic,d}}^{\boldsymbol{c}} &= -(\dot{\boldsymbol{s}}_{\boldsymbol{t,d}} + \boldsymbol{K_t}(\boldsymbol{s_{t,d}} - \boldsymbol{s_t})) + \boldsymbol{L_{tr}}\boldsymbol{\omega}_{\boldsymbol{Ic,d}}^{\boldsymbol{c}} \end{aligned} \tag{12}$$

where the controller gains $\boldsymbol{K_r} \in \mathbb{R}^{3\times 3}$ and $\boldsymbol{K_t} \in \mathbb{R}^{3\times 3}$ are defined as $\text{diag}\{k_{r,1}, k_{r,2}, k_{r,3}\}$ and $\text{diag}\{k_{t,1}, k_{t,2}, k_1\}$ with positive scalars $k_{r,1}$, $k_{r,2}$, $k_{r,3}$, $k_{t,1}$ and $k_{t,2}$. In case that $\boldsymbol{L_{rr}}$ is singular, the value of $\boldsymbol{\omega}_{\boldsymbol{Ic,d}}^{\boldsymbol{c}}$ is not updated.

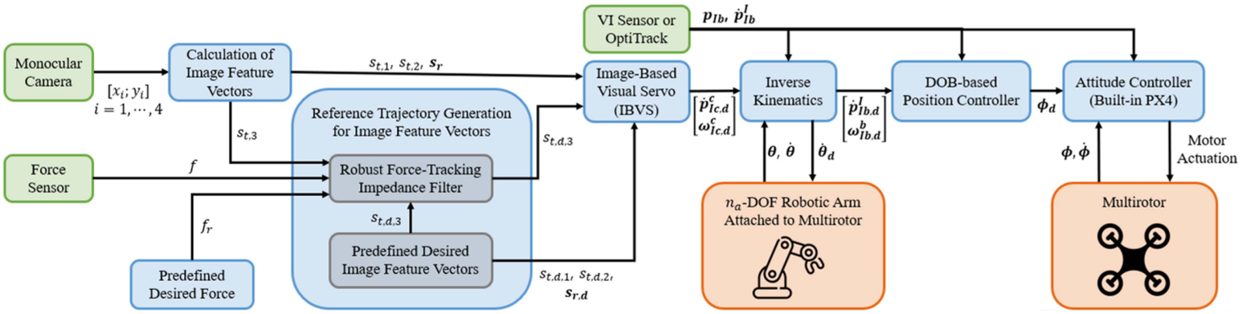

Fig. 4. Block diagram of the image-based time-varying contact force control for an aerial manipulator using robust impedance filter.

C. Inverse Kinematics: Reference Trajectory Generation for Flat Outputs of Aerial Manipulator

In [19], the aerial manipulator dynamics was shown to be differentially flat with flat outputs including the multirotor's position and yaw angle, $[\boldsymbol{p_{Ib}}; \phi_3]$, and the joint angles, $\boldsymbol{\theta}$. Here, we derive the desired profiles of the flat outputs from the desired camera twist obtained from (12).

In (3) and (4), the result of forward kinematics from $[\dot{\boldsymbol{p}}^{\boldsymbol{I}}_{\boldsymbol{Ib}}; \dot{\phi}_3; \dot{\boldsymbol{\theta}}_{\boldsymbol{d}}]$ to $[\dot{\boldsymbol{p}}^{\boldsymbol{c}}_{\boldsymbol{Ic}}; \boldsymbol{\omega}^{\boldsymbol{c}}_{\boldsymbol{Ic}}]$ is introduced. From those equations, $\dot{\boldsymbol{p}}^{\boldsymbol{I}}_{\boldsymbol{Ib,d}}$ and $[\dot{\phi}_{d,3}; \dot{\boldsymbol{\theta}}_{\boldsymbol{d}}]$ are obtained as follows:

$$\begin{aligned}
[\dot{\phi}_{d,3}; \dot{\boldsymbol{\theta}}_{\boldsymbol{d}}] &= [\boldsymbol{R}^{\top}_{\boldsymbol{bc}} \boldsymbol{Q}_{\boldsymbol{y}} \vdots \boldsymbol{J}^{\boldsymbol{c}}_{\boldsymbol{bc}}]^{\dagger} (\boldsymbol{\omega}^{\boldsymbol{c}}_{\boldsymbol{Ic,d}} - \boldsymbol{R}^{\top}_{\boldsymbol{bc}} \boldsymbol{Q}_{\boldsymbol{rp}} [\dot{\phi}_1; \dot{\phi}_2]) \\
\dot{\boldsymbol{p}}^{\boldsymbol{I}}_{\boldsymbol{Ib,d}} &= \boldsymbol{R}_{\boldsymbol{Ib}} (\boldsymbol{R}_{\boldsymbol{bc}} \dot{\boldsymbol{p}}^{\boldsymbol{c}}_{\boldsymbol{Ic,d}} + [\boldsymbol{p}_{\boldsymbol{bc}}]_{\times} \boldsymbol{Q}_{\boldsymbol{y}} \dot{\phi}_3 \\
&\quad - \boldsymbol{J}_{\boldsymbol{t,bc}} \dot{\boldsymbol{\theta}} + [\boldsymbol{p}_{\boldsymbol{bc}}]_{\times} \boldsymbol{Q}_{\boldsymbol{rp}} [\dot{\phi}_1; \dot{\phi}_2]).
\end{aligned} \tag{13}$$

where † means the pseudo-inverse.

V. STABILITY ANALYSIS

In this section, we present the theoretical results of stability analysis based on the assumption shown below:

Assumption 1: Four point features in Fig. 3 are well extracted and tracked by existing computer vision-based methods.

While Theorem 1 states that the stable contact between the end-effector and the surface is guaranteed, Theorem 2 illustrates the convergence of e_f to an arbitrarily small bound thanks to the effect of the saturation function term in (11).

Theorem 1: Assuming that $z_{f,d} = \dot{f}_d + \lambda_1 f_d > 0$, the end-effector makes contact with the surface again even though the detachment between the end-effector and the contact surface occurs.

Proof: Refer to Appendix A. ■

Theorem 2: There exist class-$\mathcal{K}_\infty$ functions $\alpha_1(\cdot)$ and $\alpha_2(\cdot)$ such that $|e_f|$ becomes uniformly ultimately bounded (UUB) with an arbitrarily small bound $\alpha_1^{-1}(\alpha_2(\varepsilon))$.

Proof: Refer to Appendix B. ■

VI. EXPERIMENTAL RESULTS

This section reports the experimental validation of the proposed image-based force control strategy.

A. Experimental Setups

The aerial manipulator used to validate the proposed controller consists of four parts: an underactuated quadrotor, a robotic arm, a monocular camera, and a 1-axis force sensor. The quadrotor was assembled with the off-the-shelf frame DJI S500, four 11-inch T-Motor carbon fiber propellers, four U3-KV700 motors with corresponding ALPHA-40A-LV electronic speed controllers (ESCs), Intel NUC for computing, a 4S Turnigy Lipo battery to power up Intel NUC, and a 6S Polytronics Lipo battery for the power supplement of the four motors and robotic arm. On Intel NUC, Robot Operating System (ROS) is installed in Ubuntu 18.04. An algorithm to obtain the 2-dimensional points of four point features projected on the camera's image plane, a position controller for the quadrotor, a servo-angle controller for the robotic arm, and a navigation algorithm are executed on Intel NUC. Based on the outputs from the position controller, the attitude controller embedded in Pixhawk 4 generates the motor command which is sent to the four ESCs. The robotic arm is comprised of ROBOTIS dynamixel XH540 and XM430 servo motors and is controlled by U2D2. For the monocular camera, we mount an RGB-monocular camera, MLC2000wG, parallel to the end-effector where the camera axis heads to the direction of force exertion, and set the camera to 752x480 resolution and 20fps. On the end-effector, Honeywell FSS2000NSB 1-axis force sensor is equipped. As a force-exerting surface, we use a whiteboard.

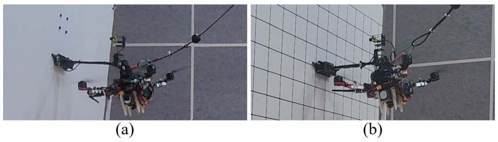

Fig. 5. Aerial manipulators in front of the whiteboard filled with (a) four round markers and (b) grid patterns.

In principle, the proposed control framework can be operated without an external positioning system by estimating the vehicle's velocity using camera-IMU-based localization algorithms [20], [21]. However, since this work aims to verify the tracking of a time-varying force reference trajectory based on the well-extracted and -tracked four point features, we use OptiTrack for the navigation to focus on the validation of the proposed control strategy.

B. Control Strategy

According to the robotic arm's configuration presented in Fig. 6, a unit vector expressing the x axis of $\mathcal{F}_{ee}$ w.r.t. $\mathcal{F}_I$, $\boldsymbol{n_f} \in \mathbb{R}^3$, is derived as follows:

$$\boldsymbol{n_f} = \begin{bmatrix} c_{\theta_2} & s_{\theta_2} s_{(\theta_1+\phi_1)} & -s_{\theta_2} c_{(\theta_1+\phi_1)} \\ 0 & c_{(\theta_1+\phi_1)} & s_{(\theta_1+\phi_1)} \\ s_{\theta_2} & -c_{\theta_2} s_{(\theta_1+\phi_1)} & c_{\theta_2} c_{(\theta_1+\phi_1)} \end{bmatrix} \begin{bmatrix} c_{\phi_2} c_{\phi_3} \\ -s_{\phi_2} \\ s_{\phi_2} c_{\phi_3} \end{bmatrix}$$

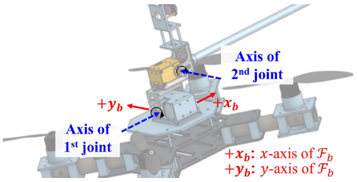

Fig. 6. Configuration of the robotic arm attached to the multirotor.

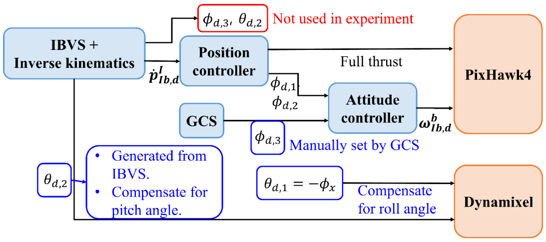

Fig. 7. Controller diagram for the actual experiments.

TABLE I
PARAMETERS AND GAINS USED FOR REPRODUCIBILITY TEST

Variable	m_d	b_d	k	k_1	$k_{t,1}$	$k_{t,2}$
Value	2.0	50	0.20	0.20	0.20	0.20
$k_{r,1}$	$k_{r,2}$	$k_{r,3}$	λ_1	λ_2	λ_3	ε
0.20	0.20	0.20	60	0.16	1.0×10^{-3}	0.30

where c_* and s_* represents $\cos *$ and $\sin *$, respectively. If we set θ_1 to follow $-\phi_1$, we can rearrange $\boldsymbol{n_f}$ as follows:

$$\boldsymbol{n_f} = \begin{bmatrix} c_{\theta_2} & 0 & -s_{\theta_2} \\ 0 & 1 & 0 \\ s_{\theta_2} & 0 & c_{\theta_2} \end{bmatrix} \begin{bmatrix} c_{\phi_2} c_{\phi_3} \\ -s_{\phi_2} \\ s_{\phi_2} c_{\phi_3} \end{bmatrix}. \tag{14}$$

Hence, we can align $\boldsymbol{n_f}$ normal to the contact surface by regulating ϕ_3 and θ_2 where ϕ_2 is measured by an IMU sensor mounted on a multirotor. In our experimental setting, we only extract $\dot{\theta}_{d,2}$ from $[\dot{\phi}_{d,3}; \dot{\boldsymbol{\theta}}_{\boldsymbol{d}}]$ obtained by (12) and (13) to focus on checking whether $\dot{\theta}_{d,2}$ can compensate for the tilting angles of the multirotor and the contact surface. As depicted in Fig. 7, $\phi_{d,3}$ is controlled by a ground control system (GCS) and $\theta_{d,1}$ is set to $-\phi_1$ to compensate for the roll angle of the multirotor. Finally, based on $\dot{\boldsymbol{p}}^{\boldsymbol{I}}_{\boldsymbol{Ib,d}}$ and $\dot{\theta}_{d,2}$ calculated in (13), we integrate them w.r.t. time to obtain $\boldsymbol{p_{Ib,d}}$, $\theta_{d,2}$.

For the position control, we use the disturbance-observer-based controller introduced in [22], and for the attitude control, we rely on an internal controller operated in the PixHawk4 flight controller unit. Also, for the joint angle control, we utilize the built-in controller made by Dynamixel.

C. Experimental Scenarios

1) Reproducibility Test: To validate the reproducibility of the proposed algorithm, we conducted ten experiments on the whiteboard with four round markers tilted by -6 [deg] (Fig. 5(a)) and two experiments on the vertical whiteboard filled with grid patterns (Fig. 5(b)). For all experiments, the force profile $f_d(t) = f_{sine}(t) \triangleq 2 - \cos\left(\frac{\pi(t-t_0)}{5}\right)$ [N] was utilized with the parameters and gains shown in Table I.

TABLE II
PARAMETERS AND GAINS USED FOR FORCE-TRACKING PERFORMANCE COMPARISON

Variable	m_d	b_d	k	k_1	$k_{t,1}$	$k_{t,2}$
Value	2.0	50	0.30	0.20	0.20	0.20
$k_{r,1}$	$k_{r,2}$	$k_{r,3}$	λ_1	λ_2	λ_3	ε
0.20	0.20	0.20	10	Vary	Vary	0.30

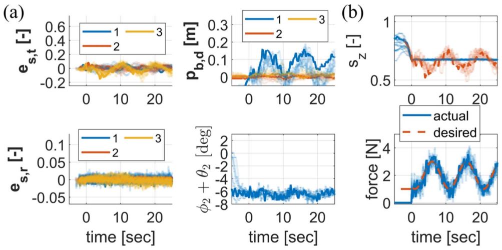

Fig. 8. (a) Histories of the errors in image feature vectors ($\boldsymbol{e_{s,t}}$&$\boldsymbol{e_{s,r}}$), desired position of the multirotor ($\boldsymbol{p_{Ib,d}}$) and pitch angle of the end-effector ($\phi_2 + \theta_2$), and (b) histories of the contact and desired force, and the actual and desired values of $s_{t,3} = s_z$ from the ten experiments on the whiteboard with four point features. The best results are expressed in bold lines.

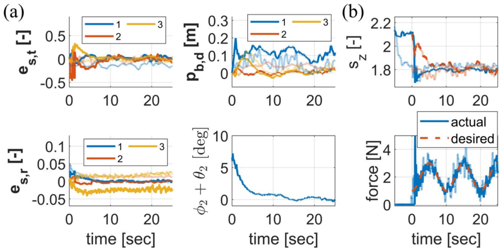

Fig. 9. (a) Histories of the errors in image feature vectors ($\boldsymbol{e_{s,t}}$&$\boldsymbol{e_{s,r}}$), desired position of the multirotor ($\boldsymbol{p_{Ib,d}}$) and pitch angle of the end-effector ($\phi_2 + \theta_2$), and (b) histories of the contact and desired force, and the actual and desired values of $s_{t,3} = s_z$ from the two experiments on the whiteboard filled with grid patterns. The best results are expressed in bold lines.

2) Force-Tracking Performance Comparison: To show the enhanced force-tracking performance of our IBVS method, we conducted force-tracking experiments with two different force profiles shown below:

1) Sinusoidal: $f_d(t) = f_{sine}(t)$ [N] $\forall t \in [t_0, t_0 + 10]$
2) Smoothed trapezoid:

$$f_d(t) = \begin{cases} 3, & t \in [t_0 + 5, t_0 + 15) \\ f_{sine}(t), & \text{otherwise} \end{cases} \text{[N]}$$

These experiments were conducted on the tilted whiteboard (Fig. 5(a)) with the gains and parameters shown in Table II .

D. Results

1) Reproducibility Test: In Fig. 8, we plot the measured values of $\boldsymbol{e_{s,t}} \triangleq \boldsymbol{s_{t,d}} - \boldsymbol{s_t}$, $\boldsymbol{e_{s,r}} \triangleq \boldsymbol{s_{r,d}} - \boldsymbol{s_r}$, $\phi_2 + \theta_2$, s_z, $s_{z,r}$, f, and f_d for the ten repetitive experiments on the tilted whiteboard with four point features. Also, for two repetitive experiments on the vertical whiteboard filled with grid patterns, we illustrate the aforementioned values in Fig. 9. Since the value of $\phi_2 + \theta_2$ converges to the tilted angle of the whiteboard when the optical axis of the camera aligns with the normal direction of the surface,

TABLE III
QUANTITATIVE RESULTS ON REPRODUCIBILITY TEST. AVG. MEANS THE AVERAGE VALUE

	Four round markers (Tilted)		Grid patterns (Vertical)	
Trial	1-5 (Avg.)	6-10 (Avg.)	1	2
$M_{\lvert e_f \rvert}$ [N]	0.892	1.11	1.44	0.721
$\mu_{f,\text{meas}}$ [N]	0.229	0.280	0.358	0.179
$\sigma_{f,\text{meas}}$ [N]	0.169	0.226	0.268	0.133

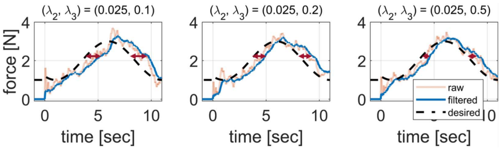

Fig. 10. Histories of the raw, estimated, and desired force from the three experiments using different values of λ_3 with the sinusoidal force profile.

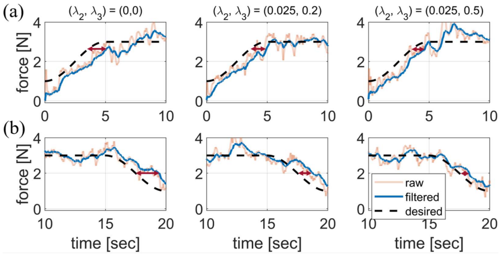

Fig. 11. Histories of the raw, estimated, and desired force from the three experiments using different values of λ_3 with the smoothed trapezoid force profile. (a) Increasing phase and (b) decreasing phase.

we plot this value to evaluate the performance of IBVS in the rotational part.

Fig. 8(a) shows that both $\boldsymbol{e_{s,t}}$ and $\boldsymbol{e_{s,r}}$ are well regulated close to zero, and $\phi_2 + \theta_2$ also converges to -6 [deg]. Also, Fig. 8 b reports that the measured force adequately tracks its sinusoidal desired profile by updating $s_{z,d}$ along with (11).

Similarly, Fig. 9(a) shows that $\boldsymbol{e_{s,t}}$ and $\boldsymbol{e_{s,r}}$, and $\phi_2 + \theta_{d,2}$ converge to zero, and as shown in Fig. 9(b), the force-tracking results of the two experiments are similar to those of the previous ten experiments.

To quantitatively validate the reproducibility of our method, we calculate $|e_f|$'s maximum value $M_{|e_f|} \triangleq \max_{t_i \le t \le t_f} |e_f|$, time-average $\mu_{f,\text{meas}} \triangleq \frac{1}{20}\int_{t_i}^{t_f} |e_f(\tau)| d\tau$, and time-standard-deviation $\sigma_{f,\text{meas}} \triangleq \sqrt{\frac{1}{20}\int_{t_i}^{t_f} (|e_f(\tau)| - \mu_{f,\text{meas}})^2 d\tau}$ between $t_i = t_0 + 5$ [sec] and $t_f = t_0 + 25$ [sec]. From Table III, we can notice that all quantitative metrics show a similar degree of force-tracking performance. This validates the reproducibility of the proposed impedance filter.

2) Force-Tracking Performance Comparison: In Figs. 10 and 11, to observe the tendency of change in f, we plot $\hat{f}$, which is the low-pass filtered value of f with the cutoff frequency of 0.016 Hz.

Fig. 10 illustrates the results of three experiments on following the sinusoidal force reference trajectory conducted with three different values of λ_3 in (11)y. From the results of $\hat{f}$, we can notice that $\hat{f}$ tracks f_d faster as the value of λ_3 increases.

In Fig. 11, the histories of f, $\hat{f}$ and f_d are partitioned into an increasing phase ($t \in [t_0, t_0 + 10)$) and a decreasing phase ($t \in [t_0 + 10, t_0 + 20)$). Also, the force-tracking results of the experiments with three different combinations of (λ_2, λ_3) are depicted. From the histories of $\hat{f}$ and f_d in Fig. 11, we can confirm a tendency that the speed of force-tracking gets faster with the increasing value of λ_3.

From the results of experiments with two different time-varying force profiles, we can notice that the transient force-tracking performance gets better with increasing value of λ_3.

3) Discussion: When the desired force is *constant*, the baseline impedance filter introduced in [[14] Section IV] (i.e., (11) substituted by $\lambda_2 = 0$ and $\lambda_3 = 0$) also guarantees the exponential stability of the error between the desired and actual force. Otherwise, according to our stability analysis, the integral term in (11) plays an important role in making $|e_f| \le \alpha_1^{-1}(\alpha_2(\varepsilon))$ even when f_d is *not* constant. The improved transient force-tracking performance depicted in Figs. 10 and 11 indeed confirms the advantage of the integral term, especially the term related to λ_3.

VII. CONCLUSION

This letter presents an image-based time-varying force-tracking control law for an aerial manipulator interacting with a structure. The two image feature vectors corresponding to the translational and rotational movements of the camera are calculated from the four point features. Then, the RISATE-based impedance filter utilizes those image feature vectors to generate their reference trajectories. The IBVS method determines the desired camera twist, and the desired velocities of the multirotor, its yaw rate, and the joint velocities are obtained by employing inverse kinematics. Additionally, we conduct stability analysis to prove that the contact force error between the actual and desired values is UUB with an arbitrary small bound. To evaluate the efficacy of the control approach, aerial force-tracking experiments were conducted. The results validate that the proposed controller has reproducibility and outperforms the existing vision-based force controller. For future works, we can extend the research to enhance the transient performance during the flight mode transition.

APPENDIX
PROOF OF THEOREM 1

According to [14], the contact force is modeled as follows:

$$f = \begin{cases} -k_e(s_z - s_e), & s_z \le s_e \\ 0, & s_z > s_e \end{cases} \tag{15}$$

where k_e and s_e represent the environmental stiffness and the value of s_z when the end-effector makes contact with the surface, respectively. By substituting (12) for (9), the closed-loop error dynamics for $e_s \triangleq s_{z,d} - s_z$ is derived as follows:

$$\dot{e}_s = k_1 e_s. \tag{16}$$

Then, by putting (15) and (16) into (11), the following equation is derived:

$$m_d(\ddot{s}_z + \ddot{e}_s) + b_d(\dot{s}_z + \dot{e}_s) - k_1(b_d + k_1 m_d)e_s = -k(z_f(t)$$

$$- z_f(t_0)) - \int_{t_0}^{t} (k\lambda_2 z_f(\tau) + 2\lambda_3 \text{sat}_\varepsilon(z_f(\tau)))d\tau. \tag{17}$$

When the end-effector detaches from the surface, (17) is reformulated as follows:

$$m_d\ddot{s}_z + b_d\dot{s}_z = -u$$

where $u \triangleq k(z_{f,d}(t) - z_{f,d}(t_0)) + \int_{t_0}^{t}(k\lambda_2 z_{f,d}(\tau) + 2\lambda_3\text{sat}(z_{f,d}(\tau)))d\tau$ and $z_{f,d} \triangleq \dot{f}_d + \lambda_1 f_d$. Thus, $\dot{s}_z$ is derived as follows:

$$\dot{s}_z = \dot{s}_z(t_b) - \tfrac{b_d}{m_d}(s_z - s_e) - \tfrac{1}{m_d}\int_{t_b}^{t} u(\tau)d\tau \tag{18}$$

where t_b means the time instant of breaking contact with the surface. With the assumption that $z_{f,d} = \dot{f}_d + \lambda_1 f_d > 0$, there exists a positive number c_u such that $c_u \leq \dot{u}$. Thus, u ultimately becomes positive despite $z_{f,d}(t_d) - z_{f,d}(t_0) < 0$ unless the end-effector makes contact with the surface again. According to (18), $\dot{s}_z$ becomes strictly decreasing since $u > 0$ and $s_z > s_e$. Hence, after the detachment, s_z again converges to $s_z = s_e$ so that we can guarantee that the end-effector makes contact with the surface again.

A2 Proof of Theorem 2

To prove that e_f is UUB with an arbitrarily small bound, we first derive the closed-loop error dynamics w.r.t. e_f. Then, we define a Lyapunov candidate function V to investigate the behavior of e_f, and show that V is positive. Then, we finish the proof of Theorem 2 by showing the negativeness of $\dot{V}$.

1) Closed-Loop Error Dynamics W.r.t. e_f: When the end-effector is in contact with the surface, since $\dot{s}_z = \frac{\dot{e}_f - \dot{f}_d}{k_e}$ and $\ddot{s}_z = \frac{\ddot{e}_f - \ddot{f}_d}{k_e}$ from (15), we can rearrange (17) as follows:

$$\begin{aligned} &m_d(\ddot{e}_f - \ddot{f}_d) + b_d(\dot{e}_f - \dot{f}_d) + k_e m_d(\ddot{e}_s - k_1\dot{e}_s) \\ &+ k_e b_d(\dot{e}_s - k_1 e_s) = -k_e k(z_f - z_f(t_0)) \\ &- k_e\int_{t_0}^{t}(k\lambda_2 z_f(\tau) + 2\lambda_3\text{sat}_\varepsilon(z_f(\tau)))d\tau. \end{aligned} \tag{19}$$

By substituting (16) and its time-derivative for (19), the third and fourth terms of the left-hand side of (19) become zero. Then, by differentiating (19) with time, $\dot{\zeta}_f \triangleq \dot{z}_f + \lambda_2 z_f$ is calculated as follows:

$$m_d\dot{\zeta}_f = N - k_e k\zeta_f - 2k_e\lambda_3\text{sat}_\varepsilon(z_f) - k_e z_f \tag{20}$$

where $N \triangleq (m_d(\lambda_1 + \lambda_2) - b_d)\ddot{e}_f + (m_d\lambda_1\lambda_2 + k_e)\dot{e}_f + k_e\lambda_1 e_f + m_d\dddot{f}_d + b_d\ddot{f}_d$.

2) Definition of a Lyapunov Candidate Function V: We define V as $\frac{k_e}{2}e_f^2 + \frac{k_e}{2}z_f^2 + \frac{m_d}{2}\zeta_f^2 + Q(t)$ where

$$\begin{aligned} &Q(t) \triangleq 2k_e\lambda_3\left(\int_{t_0}^{t}\dot{z}_f\text{sat}_\varepsilon(z_f)d\tau + |z_f(t_0)|\right) - N_d z_f, \\ &N_d \triangleq m_d\dddot{f}_d + b_d\ddot{f}_d. \end{aligned} \tag{21}$$

To investigate the stability of $\boldsymbol{e_f}$, we need to examine the properties of V and $\dot{V}$.

3) Positiveness of V: In this part, we will prove that V is positive and has its lower and upper bounds which are all increasing functions w.r.t. $\|\boldsymbol{e_f}\|$. Since the part including the first to third terms of V is positive and an increasing function w.r.t. $\|\boldsymbol{e_f}\|$, we only need to investigate the sign of $Q(t)$.

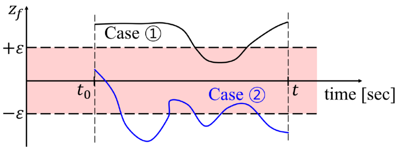

Fig. 12. Two cases of the behavior of $z_f(t)$.

Because $\int_{t_1}^{t_2}\dot{z}_f\text{sat}_\varepsilon(z_f)d\tau$ in (21) becomes zero when $|z_f(t_1)| = |z_f(t_2)| = \varepsilon,\quad \forall t_1 \in (t_0, t_2]$, the value of $\int_{t_0}^{t}\dot{z}_f\text{sat}_\varepsilon(z_f)d\tau$ only depends on whether $|z_f(t_0)|$ and $|z_f(t)|$ are smaller or larger than ε. Since we focus on the case where $\varepsilon \leq z_f(t)$, we calculate $Q(t)$ for two cases shown in Fig. 12.

i) For case ① ($\varepsilon \leq |z_f(t_0)|, |z_f|$), we obtain:

$$\int_{t_0}^{t}\dot{z}_f\text{sat}_\varepsilon(z_f)d\tau = \int_{t_0}^{t}\dot{z}_f\tfrac{z_f}{|z_f|}d\tau = |z_f| - |z_f(t_0)|.$$

Thus, $Q(t) = 2k_e\lambda_3|z_f| - N_d z_f$. Since $f_d \in \mathcal{C}^4$, N_d and $\dot{N}_d$ are bounded so that there exist $\lambda_2, \lambda_3 \in (0, \infty)$ which satisfy the following inequality:

$$|N_d| + \tfrac{|\dot{N}_d|}{\lambda_2} \leq k_e\lambda_3. \tag{22}$$

If we substitute (22) for $Q(t)$, the lower bound of $Q(t)$ is derived as follows:

$$Q(t) \geq (2k_e\lambda_3 - |N_d|)|z_f| \geq (|N_d| + 2\tfrac{|\dot{N}_d|}{\lambda_2})|z_f| \geq 0. \tag{23}$$

ii) For case ② ($|z_f(t_0)| < \varepsilon \leq |z_f|$), $\int_{t_0}^{t}\dot{z}_f\text{sat}_\varepsilon(z_f)d\tau$ is calculated as follows:

$$\begin{aligned} \int_{t_0}^{t}\dot{z}_f\text{sat}_\varepsilon(z_f)d\tau &= \int_{t_2^{**}}^{t}\dot{z}_f\tfrac{z_f}{|z_f|}d\tau + \int_{t_0}^{t_2^*}\tfrac{d}{d\tau}\left(\tfrac{z_f^2}{2\varepsilon}\right)d\tau \\ &= |z_f| - \varepsilon + \tfrac{\varepsilon}{2} - \tfrac{z_f^2(t_0)}{2\varepsilon} = |z_f| - \tfrac{z_f^2(t_0)+\varepsilon^2}{2\varepsilon} \end{aligned}$$

where $|z_f(t_2^{**})| = |z_f(t_2^*)| = \varepsilon$ for $t_0 < t_2^* < t_2^{**} < t$. Therefore, the following inequality holds:

$$\begin{aligned} Q(t) &= 2k_e\lambda_3\left(|z_f| - \tfrac{(|z_f(t_0)-\varepsilon|)^2}{2\varepsilon}\right) - N_d z_f \\ &\geq |N_d|\left(|z_f| - \tfrac{(|z_f(t_0)|-\varepsilon)^2}{\varepsilon}\right) + \tfrac{2|\dot{N}_d|}{\lambda_2}\left(|z_f| - \tfrac{(|z_f(t_0)|-\varepsilon)^2}{2\varepsilon}\right) \\ &\geq \tfrac{|\dot{N}_d||z_f|}{\lambda_2} + \left(|N_d| + \tfrac{|\dot{N}_d|}{\lambda_2}\right)\left(|z_f| - \tfrac{(|z_f(t_0)|-\varepsilon)^2}{\varepsilon}\right) \geq 0. \end{aligned}$$

Meanwhile, according to (21), $Q(t)$ also satisfies:

$$Q(t) \leq 2k_e\lambda_3|z_f| + |N_d||z_f| \leq (2k_e\lambda_3 + |N_d|)|z_f|. \tag{24}$$

Thus, $Q(t)$ satisfies the following inequality:

$$0 \leq Q(t) \leq (2k_e\lambda_3 + |N_d|)|z_f|. \tag{25}$$

Hence, there exist positive constants γ_1 and γ_2 satisfying:

$$\gamma_1\|\boldsymbol{e_f}\|^2 \leq V \leq \gamma_2\|\boldsymbol{e_f}\|^2 + (2\lambda_3 + |N_d|)|z_f|. \tag{26}$$

4) Negativeness of $\dot{V}$: In this procedure, we develop equations and inequalities with the condition $|z_f| \geq \varepsilon$. To prove the performance of force tracking, we first calculate $\dot{V}$ as follows:

$$\begin{aligned} \dot{V} &= k_e e_f\dot{e}_f + k_e z_f\dot{z}_f + \zeta_f(m_d\dot{\zeta}_f) \\ &+ 2k_e\lambda_3\dot{z}_f\text{sat}_\varepsilon(z_f) - N_d\dot{z}_f - \dot{N}_d z_f \end{aligned} \tag{27}$$

By substituting (20) for (27), $\dot{V}$ is reformulated as follows:

$$\begin{aligned}\dot{V} = & -k_e\lambda_1 e_f^2 + k_e e_f z_f - k_e\lambda_2 z_f^2 - k_e k\zeta_f^2 \\ & + \tilde{N}\zeta_f - k_e\lambda_2\lambda_3 z_f \mathrm{sat}_\varepsilon(z_f) + R_f(t) \end{aligned} \tag{28}$$

where $R_f(t) \triangleq -(\dot{N}_d - \lambda_2 N_d)z_f - k_e\lambda_2\lambda_3|z_f|$. From (22), we can derive the inequality below:

$$-(\dot{N}_d - \lambda_2 N_d)z_f \leq (|\dot{N}_d| + \lambda_2|N_d|)|z_f| \leq k_e\lambda_2\lambda_3|z_f|$$

so that $R_f(t) \leq 0$ holds. Because $e_f z_f \leq \frac{1}{2}(c_e^2 e_f^2 + \frac{z_f^2}{c_e^2})$ and $R_f(t) \leq 0$ hold with an arbitrary constant $c_e > 0$, $\dot{V}$ in (28) satisfies the following inequality:

$$\begin{aligned}\dot{V} \leq & -k_e(\lambda_1 - \tfrac{c_e^2}{2})e_f^2 - k_e(\lambda_2 - \tfrac{1}{2c_e^2})z_f^2 - \tfrac{1}{2}k_e k\zeta_f^2 \\ & - \tfrac{1}{2}k_e k\zeta_f^2 + |\tilde{N}||\zeta_f| - k_e\lambda_2\lambda_3 z_f \mathrm{sat}_\varepsilon(z_f) \\ \leq & -k_e(\lambda_1 - \tfrac{c_e^2}{2})e_f^2 - k_e(\lambda_2 - \tfrac{1}{2c_e^2})z_f^2 - \tfrac{1}{2}k_e k\zeta_f^2 \\ & - \tfrac{k_e k}{2}\left(\zeta_f - \tfrac{1}{k_e k}|\tilde{N}|\right)^2 + \tfrac{1}{2k_e k}|\tilde{N}|^2 - k_e\lambda_2\lambda_3 z_f \mathrm{sat}_\varepsilon(z_f) \end{aligned} \tag{29}$$

where $\tilde{N} \triangleq N - N_d$

Since $\tilde{N} = (m_d(\lambda_1 + \lambda_2) - b_d)\ddot{e}_f + (m_d\lambda_1\lambda_2 + k_e)\dot{e}_f + k_e\lambda_1 e_f$, it becomes a linear combination of e_f, $\dot{e}_f$ and $\ddot{e}_f$ so that there exist constants c_1, c_2 and c_3 such that $\tilde{N} = c_1 e_f + c_2 z_f + c_3\zeta_f$. Thus, we can find $\mu > 0$ satisfying

$$|\tilde{N}| \leq \mu\|\boldsymbol{e_f}\|. \tag{30}$$

If $\lambda_1 > \frac{c_e^2}{2}$ and $\lambda_2 > \frac{1}{2c_e^2}$, then (29) is rearranged as follows:

$$\dot{V} \leq -\left(\lambda^* - \frac{\mu^2}{2k_e k}\right)\|\boldsymbol{e_f}\|^2 - k_e\lambda_2\lambda_3 z_f \mathrm{sat}_\varepsilon(z_f) \tag{31}$$

where $\lambda^* \triangleq k_e \min\{\lambda_1 - \frac{c_e^2}{2}, \lambda_2 - \frac{1}{2c_e^2}, \frac{k}{2}\}$.

5) e_f is UUB in a Arbitrarily Small Bound: If $\mu^2/2k_e k < \lambda^*$, there exists $c^* > 0$ which satisfies $\dot{V} \leq -c^* V\ \forall |z_f| > \varepsilon$.

Owing to (26) and $\dot{V} \leq -c^* V$, there exist class $\mathcal{K}_\infty$ functions α_1 and α_2 and a continuous positive definite function $W : \mathbb{R}^3 \to \mathbb{R}$ such that

$$\alpha_1(\|\boldsymbol{e_f}\|) \leq V(\|\boldsymbol{e_f}\|) \leq \alpha_2(\|\boldsymbol{e_f}\|)$$

$$\dot{V} \leq -W(\boldsymbol{e_f}), \quad \forall \|\boldsymbol{e_f}\| \geq |z_f| \geq \varepsilon > 0$$

$\forall\, t \geq t_0$. According to [23, Theorem 4.18], there exists a class $\mathcal{KL}$ function η and a constant $T_f \geq 0$ such that

$$\|\boldsymbol{e_f}\| \leq \eta(\|\boldsymbol{e_f}(t_0)\|, t - t_0), \quad \forall\, t_0 \leq t \leq t_0 + T_f$$

$$\|\boldsymbol{e_f}\| \leq \alpha_1^{-1}(\alpha_2(\varepsilon)), \quad \forall\, t \geq t_0 + T_f$$

for any initial state $\|\boldsymbol{e_f}(t_0)\|$. Therefore, $|e_f| \leq \|\boldsymbol{e_f}\|$ becomes UUB with an arbitrarily small bound $\alpha_1^{-1}(\alpha_2(\varepsilon))$.

ACKNOWLEDGMENT

The authors would like to thank Yeonjoon Kim (Seoul National University) for his help conducting the experiments.

REFERENCES

[1] G. Nava, Q. Sablé, M. Tognon, D. Pucci, and A. Franchi, "Direct force feedback control and online multi-task optimization for aerial manipulators," *IEEE Robot. Automat. Lett.*, vol. 5, no. 2, pp. 331–338, Apr. 2019.

[2] Y. Sun, Z. Jing, P. Dong, J. Huang, W. Chen, and H. Leung, "A switchable unmanned aerial manipulator system for window-cleaning robot installation," *IEEE Robot. Automat. Lett.*, vol. 6, no. 2, pp. 3483–3490, Apr. 2021.

[3] C. Ding, L. Lu, C. Wang, and C. Ding, "Design, sensing, and control of a novel UAV platform for aerial drilling and screwing," *IEEE Robot. Automat. Lett.*, vol. 6, no. 2, pp. 3176–3183, Apr. 2021.

[4] M. Allenspach, N. Lawrance, M. Tognon, and R. Siegwart, "Towards 6dof bilateral teleoperation of an omnidirectional aerial vehicle for aerial physical interaction," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2022, pp. 9302–9308.

[5] E. Shahriari, S. A. B. Birjandi, and S. Haddadin, "Passivity-based adaptive force-impedance control for modular multi-manual object manipulation," *IEEE Robot. Automat. Lett.*, vol. 7, no. 2, pp. 2194–2201, Apr. 2022.

[6] J. Byun, I. Jang, D. Lee, and H. J. Kim, "A hybrid controller enhancing transient performance for an aerial manipulator extracting a wedged object," *IEEE Trans. Automat. Sci. Eng.*, early access, doi: 10.1109/TASE.2023.3277508.

[7] L. Marković, M. Car, M. Orsag, and S. Bogdan, "Adaptive stiffness estimation impedance control for achieving sustained contact in aerial manipulation," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2021, pp. 117–123.

[8] L. Peric, M. Brunner, K. Bodie, M. Tognon, and R. Siegwart, "Direct force and pose NMPC with multiple interaction modes for aerial push-and-slide operations," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2021, pp. 131–137.

[9] C. Izaguirre-Espinosa, A.-J. Muñoz-Vázquez, A. Sanchez-Orta, V. Parra-Vega, and P. Castillo, "Contact force tracking of quadrotors based on robust attitude control," *Control Eng. Pract.*, vol. 78, pp. 89–96, 2018.

[10] Y. Rong, W. Chou, and R. Jiao, "Robust fault-tolerant motion/force control of a fully-actuated hexarotor using adaptive sliding mode impedance control," *Int. J. Robust Nonlinear Control*, vol. 32, no. 7, pp. 4149–4172, 2022.

[11] J. Byun, B. Kim, C. Kim, D. D. Oh, and H. J. Kim, "Stable contact guaranteeing motion/force control for an aerial manipulator on an arbitrarily tilted surface," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2023, pp. 5345–5351.

[12] G. He, Y. Jangir, J. Geng, M. Mousaei, D. Bai, and S. Scherer, "Image-based visual servo control for aerial manipulation using a fully-actuated UAV," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2023, pp. 5042–5049.

[13] K. Bodie et al., "Active interaction force control for contact-based inspection with a fully actuated aerial vehicle," *IEEE Trans. Robot.*, vol. 37, no. 3, pp. 709–722, Jun. 2020.

[14] M. Xu, A. Hu, and H. Wang, "Image-based visual impedance force control for contact aerial manipulation," *IEEE Trans. Automat. Sci. Eng.*, vol. 20, no. 1, pp. 518–527, Jan. 2022.

[15] A. Hu, Q. De, and H. Wang, "Purely image-based dynamic impedance control of a fully actuated aerial vehicle," *IEEE/ASME Trans. Mechatron.*, early access, doi: 10.1109/TMECH.2023.3293102.

[16] M. Keshmiri and W.-F. Xie, "Image-based visual servoing using an optimized trajectory planning technique," *IEEE/ASME Trans. Mechatron.*, vol. 22, no. 1, pp. 359–370, Feb. 2017.

[17] O. Tahri and F. Chaumette, "Point-based and region-based image moments for visual servoing of planar objects," *IEEE Trans. Robot.*, vol. 21, no. 6, pp. 1116–1127, Dec. 2005.

[18] D. Lee, J. Byun, and H. J. Kim, "Rise-based trajectory tracking control of an aerial manipulator under uncertainty," *IEEE Contr. Syst. Lett.*, vol. 6, pp. 3379–3384, 2022.

[19] H. Seo, S. Kim, and H. J. Kim, "Locally optimal trajectory planning for aerial manipulation in constrained environments," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2017, pp. 1719–1724.

[20] M. Achtelik, M. Achtelik, S. Weiss, and R. Siegwart, "Onboard IMU and monocular vision based control for mavs in unknown in-and outdoor environments," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2011, pp. 3056–3063.

[21] J. Engel, J. Sturm, and D. Cremers, "Camera-based navigation of a low-cost quadrocopter," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2012, pp. 2815–2821.

[22] D. Lee, H. Seo, I. Jang, S. J. Lee, and H. J. Kim, "Aerial manipulator pushing a movable structure using a DOB-based robust controller," *IEEE Robot. Automat. Lett.*, vol. 6, no. 2, pp. 723–730, Apr. 2020.

[23] H. K. Khalil, *Nonlinear Syst.*, 3rd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2002.