Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnline CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 52-30 (2019) 114–119

On Achieving Bounded Harvest Times in Robotic Fruit Harvesting: A Finite-Time Visual Servo Control Approach

S. S. Mehta * **C. Ton** ** **M. Rysz** *** **P. Ganesh** * **Z. Kan** **** **T. F. Burks** †

* *Department of Mechanical and Aersospace Engineering, University of Florida, Shalimar, FL-32579 (e-mail: {siddhart,prashant.ganesh}@ufl.edu).*
** *Southwest Research Institute, 6220 Culebra Rd, San Antonio, TX-78238 (e-mail: chau.t.ton@gmail.com)*
*** *Department of Information Systems & Analytics, Miami University, Oxford, OH-45056 (e-mail: ryszmw@miamioh.edu)*
**** *Department of Automation, University of Science and Technology of China, Hefei, Anhui, 230052, China (e-mail: zkan@ustc.edu.cn)*
† *Department of Agricultural and Biological Engineering, University of Florida, Gainesville, FL-32611 (e-mail: tburks@ufl.edu).*

Abstract: To improve commercial feasibility of robotic harvesters, it is utmost important to reduce and be able to guarantee harvesting times. A significant portion of this responsibility is on the control system. Current visual servo control methods can at best achieve exponential regulation of a robot (i.e., theoretically infinite convergence time), making it impossible to predict harvest time. The aim of this paper is to introduce a new finite-time visual servo control approach that guarantees finite (i.e., bounded) and computable harvest times. To this end, a continuous terminal sliding mode visual servo controller is developed, and Lyapunov-based stability analysis is presented to guarantee finite-time regulation of the robot to a target fruit. Further, we derive expressions for the bound on the harvesting time, which can aid post-harvest operations management. The developed controller is validated through numerical simulations.

Keywords: Harvesting time; Agricultural robotics; Terminal sliding mode control

1. INTRODUCTION

Many developed and rapidly developing countries around the world are exploring autonomous robotic solutions for harvesting specialty crop. The desire to invest in robotic solutions stems partly due to high labor prices, labor shortages, and low labor productivity. Robotic harvesting technology has the potential to overcome these challenges by providing economic advantages through increased productivity. However, the economic feasibility of the robotic harvesting operation depends at large on two factors: harvesting efficiency, which is a quantitative measure of successful harvest, and harvest time (i.e., the time required to harvest the crop). To maintain economic advantage, it is imperative that the harvest time for the robotic operation should be comparable to its manual counterpart. Picking time, which can be defined as the time for the robot to reach, detach and store a fruit or vegetable, accounts for most of the harvest time. The dynamic response of a robot and the performance of the underlying control system dictate the picking times. To this end, the objective of this paper is to design advanced control systems for robotic harvesting that can guarantee a desired bound on the picking time.

Vision-based control is the most popular control technique used in robotic harvesting. The objective of vision-based control in robotic harvesting is to autonomously position the robot in relation to a fruit for successful detachment using measurements provided by a vision system. For comprehensive overview of robotic systems and vision-based control in agriculture, readers are referred to Bac et al. (2014); Zhao et al. (2016). Vision-based control approaches in robotic harvesting can be broadly divided into open-loop control and closed-loop control. Although open-loop control systems are attractive due to their simplicity, these systems may suffer from excessive positioning errors in outdoor agricultural environments since continuous image feedback is not opted to verify and rectify the position of the robot with respect to the fruit. Closed-loop vision systems overcome these limitations by employing continuous image feedback. The first reported closed-loop visual servo control in robotic harvesting appears in Harrell et al. (1985), where a monochromatic vision system was implemented to track citrus fruits at 60Hz. Subsequently, the authors developed a direct visual servo controller in Harrell et al. (1989) to obtain

⋆ This research is supported in part by the USDA Small Business Innovation Research (SBIR) grant 2018-33610-28228 through GeoSpider Inc. and the USDA Capacity Building Grant (CBG) 2019-38821-29147. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funding agency.
**Corresponding author: S. S. Mehta

10.1016/j.ifacol.2019.12.507

joint inputs based on the image position of the fruit centroid with respect to the center of the image. However, the control structure resulted in geometric increase in control gains raising concerns for stability of the system and requiring an ad-hoc gain scheduling procedure. In Harrell et al. (1990) the authors developed a vision controller that computed joint velocity inputs to regulate a fruit at the principle point. The robot approached the fruit with predefined velocity until an ultrasonic transducer mounted in the end-effector detected the presence of fruit. Edan et al. (2000) used a cooperative sensing framework consisting of two monochromatic cameras serving as a far-vision sensor and a near-vision sensor. The far-vision camera detected oncoming melons, whereas the near-vision camera mounted on the gripper provided image feedback to reach centrally over a target melon until a proximity sensor detected the soil surface. In eggplant harvesting, Hayashi et al. (2002) developed a vision-based fuzzy controller to determine the forward, vertical, and angular motion of the end-effector based on the feedback from a monocular camera mounted in eye-in-hand configuration. Bulanon et al. (2005) considered an end-effector mounted camera and laser ranging sensor for apple harvesting robot. A proportional controller was developed to provide camera position inputs based on the error between the image position of an apple and the image center. Foglia and Reina (2006) also considered a monocular camera system in eye-in-hand configuration for radicchio harvesting, where the closed-loop system was aimed to compensate for positioning errors of the gripper due to unexpected speed variations of the carrier. In robotic apple harvesting, De-An et al. (2011) developed a proportional controller that related the image-space (pixel) errors to joint angle increments to position a fruit at the image center. A standard image-based visual servo (IBVS) controller was used in sweet pepper harvesting in Barth et al. (2016) using a single monocular camera in eye-in-hand configuration, where the center of gravity image coordinates of the largest segmented sweet-pepper served as the image feature. Although stability of the closed-loop control system is critical in achieving high harvesting efficiency, most of these results paid little or no attention to rigorous controller formulation and stability analysis. To guarantee and improve stability of robotic harvesters, Mehta and Burks (2014) presented a hybrid control framework that uses principles of IBVS and position-based visual servo (PBVS) to control the 3D translation of the camera. The image measurements of the fruit provide camera velocities along the image axes, whereas the camera velocity along the optical axis is obtained using the depth estimates of the fruit. The controller in Mehta and Burks (2014) guarantees exponential regulation of the robot to a target fruit. Subsequently, the hybrid control design was extended to develop robust and adaptive visual servo controllers in Mehta et al. (2016) and Mehta and Burks (2016), respectively, to compensate for the unknown fruit motion that may arise due to environmental disturbances. These advanced controllers ensure uniformly ultimately bounded stability, thereby regulating the robot in the vicinity of a target fruit. Recent result in Chen et al. (2019) developed a sliding mode controller for apple harvesting based on fuzzy neural networks to reduce chattering that is commonly associated with sliding mode control. The controller is proved to be asymptotically stable, which is a weaker notion than exponential stability.

The existing results can at best ensure exponential stability of the harvesting systems, and although exponential results provide the rate of error decay, a finite bound on the time required for the robot to reach a fruit cannot be established, i.e., the robot reaches a target fruit as time goes to infinity. As such, picking time, and hence the harvest time, cannot be guaranteed a priori. Motivated by the desire to improve harvest time, this paper presents a new visual servo controller that can provide a computable bound on the time required for the robot to reach a fruit. To this end, a continuous terminal sliding mode visual servo controller is developed, and Lyapunov-based stability analysis is provided to guarantee finite-time regulation of the robot to a target fruit. Additionally, expressions for the bound on the time required for the robot to reach a fruit are derived, and it is shown that this bound is a function of the control gains. Hence, a desired bound on the picking time can be obtained by appropriately selecting control gains. Finally, the performance of the developed controller is validated through numerical simulations.

2. EUCLIDEAN RECONSTRUCTION

The controller in this paper uses a cooperative camera configuration consisting of a fixed camera in the workspace and a camera-in-hand (CiH) attached to the robot end-effector. The fixed camera provides Euclidean position estimates of the fruits, which can be used to select a fruit to be harvested and align the robot end-effector such that the fruit is visible to the CiH. Subsequently, the robot is driven to the target fruit using image feedback from the CiH.

Consider three orthogonal coordinate frames $\mathcal{F}$, $\mathcal{F}_f$, and $\mathcal{F}_b$ as shown in Fig. 1. The coordinate frame $\mathcal{F}$ is attached to the CiH, and therefore, it is time-varying with respect to stationary frame $\mathcal{F}_b$. The coordinate frame $\mathcal{F}_f$ is attached to the fixed camera. Let the fruit to be harvested be O^* as shown in Fig. 1. The Euclidean position of the fruit with respect to the cameras is unknown as the depth information is lost in projective transformation. Let the unknown Euclidean coordinates of the fruit centroid, $\bar{m}(t), \bar{m}_f \in \mathbb{R}^3$, expressed in terms of $\mathcal{F}$ and $\mathcal{F}_f$, respectively, be written as

$$\bar{m}(t) = [\, x(t) \;\; y(t) \;\; z(t) \,]^T, \quad \bar{m}_f = [\, x_f \;\; y_f \;\; z_f \,]^T \tag{1}$$

where $z(t), z_f \in \mathbb{R}$ denote the unknown depth of the target fruit measured in $\mathcal{F}$ and $\mathcal{F}_f$, respectively. Since the Euclidean space is projected onto the image-space, let $m(t)$ and m_f denote the corresponding normalized Euclidean coordinates of the fruit centroid

$$m(t) = \left[\frac{x(t)}{z(t)} \;\; \frac{y(t)}{z(t)} \;\; 1 \right]^T, \quad m_f = \left[\frac{x_f}{z_f} \;\; \frac{y_f}{z_f} \;\; 1 \right]^T. \tag{2}$$

Assumption 1. In (2), it is assumed that the unknown depths $z(t), z_f \geq \varepsilon$, where $\varepsilon \in \mathbb{R}_{>0}$ is an arbitrary constant. This is a standard assumption in visual servo control, which physically means that the target is always in front of the camera.

In addition to having normalized Euclidean coordinates, the fruit will also have pixel coordinates acquired by the CiH and the fixed camera. Let $p(t), p_f \in \mathbb{R}^2$ denote the pixel coordinates of the fruit centroid expressed in $\mathcal{F}$ and $\mathcal{F}_f$, respectively, as

$$p(t) \triangleq [\, u(t) \;\; v(t) \,]^T, \qquad p_f \triangleq [\, u_f \;\; v_f \,]^T. \tag{3}$$

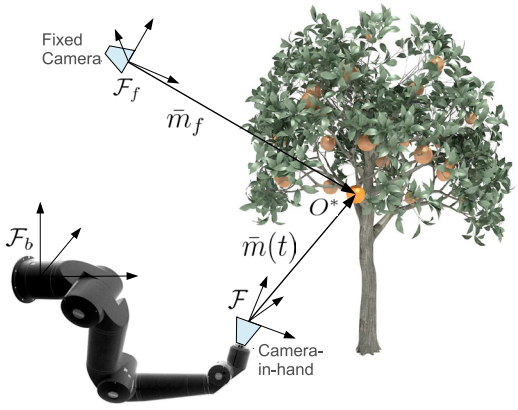

Fig. 1. Coordinate frame relationships, where the time-varying frame $\mathcal{F}$ is attached to the camera-in-hand, $\mathcal{F}_f$ corresponds to the fixed camera, and $\mathcal{F}_b$ is attached to the stationary base of the robot.

The normalized Euclidean coordinates in (2) cannot be measured directly, however, a global invertible transformation (i.e., the pinhole camera model) can be used to determine the normalized Euclidean coordinates from the corresponding pixel information as

$$\begin{bmatrix} p^T & 1 \end{bmatrix}^T = Am, \qquad \begin{bmatrix} p_f^T & 1 \end{bmatrix}^T = A_f m_f. \tag{4}$$

In (4), $A, A_f \in \mathbb{R}^{3\times3}$ are the known constant invertible intrinsic camera calibration matrices for the CiH and the fixed camera, respectively.

Remark 1. Prior to visual servo control, the estimated position of the fruit can be obtained in $\mathcal{F}_f$ using various depth estimation methods, for example, the statistical approach in Mehta and Burks (2014). The estimated fruit position in $\mathcal{F}_f$ be $\hat{\bar{m}}_f \in \mathbb{R}^3$. Using the extrinsic camera calibration parameters of the fixed camera and the robot forward kinematics, $\hat{\bar{m}}_f$ can be expressed in $\mathcal{F}_b$ and $\mathcal{F}$, and which we denote by $\hat{\bar{m}}_b, \hat{\bar{m}}'(t) \in \mathbb{R}^3$, respectively. These position estimates will be used to derive a rotation control law for the robot such that the fruit, if previously out of the field-of-view, becomes visible to the CiH.

Remark 2. When the fruit is visible to the CiH, the time-varying depth $\hat{z}(t)$ and position $\hat{\bar{m}}(t) \in \mathbb{R}^3$ of the fruit with respect to $\mathcal{F}$ can be obtained using the method in Mehta and Burks (2014). The depth $\hat{z}(t)$ depends on the area of the detected fruit in the image-space, hence $\hat{z}(t) \geq z(t), \forall t$.

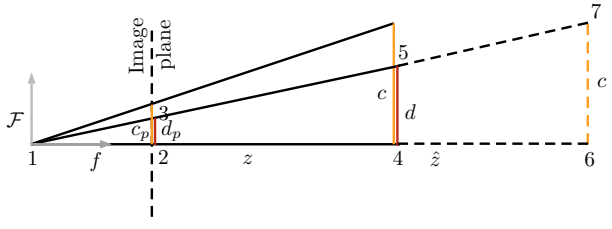

Fig. 2. Projective geometry of an object.

Proposition 1. The unknown depth ratio $\hat{z}/z$, denoted by $\xi \in \mathbb{R}_{>0}$, is constant under the given steady, but possibly uneven and imperfect, illumination conditions.

Proof 1. Consider a two-dimensional space where a camera is located at the origin, and $\mathcal{F}$ be the coordinate frame attached to the camera as shown in Fig. 2. Let an object (e.g., fruit) of size $c \in \mathbb{R}_{>0}$ be located at distance $z \in \mathbb{R}_{>0}$ from the origin of $\mathcal{F}$. The projection of this object in the image-plane, which is at a distance $f \in \mathbb{R}_{>0}$ from $\mathcal{F}$, is denoted by $c_p \in \mathbb{R}$. Further, image processing algorithms can be employeed to detect the object. Let $d_p \leq c_p$ be the size of the detected object. The inequality is due to the fact that image processing algorithms may not detect an entire object due to environmental uncertainties such as uneven illumination. The object of size d_p in the image-plane is equivalent to 1) a smaller object of size $d \in \mathbb{R}$ at distance z from the camera or 2) an object of the same size as the original, i.e., c, at a larger distance $z' \in \mathbb{R}_{>0}$ from the camera as shown in Fig. 2.

From the similar triangles $\Delta123$ and $\Delta145$, and $\Delta123$ and $\Delta167$, we get the following relationships:

$$\frac{f}{z} = \frac{d_p}{d} \qquad \frac{f}{\hat{z}} = \frac{d_p}{c}. \tag{5}$$

Using (5), the depth ratio can be obtained as $\hat{z}/z = c/d$. Given that c and d are the actual and perceived object sizes, respectively, the ratio $\hat{z}/z$ is constant $\forall t$ under the given illumination conditions. From Remark 2, the lower bound on ξ can be obtained as $\underline{\xi} = 1$. This lower bound corresponds to the case when an image processing algorithm perfectly detects a fruit such that $\hat{z} = z$.

3. CONTROL OBJECTIVE

The control objective is to regulate the robot end-effector to a target fruit for harvesting, which can be achieved by regulating the time-varying fruit pixel coordinates $p(t)$ to the image coordinates of the principal point (i.e., the intersection of an optical axis with the image plane), and regulating the end-effector to the desired fruit depth. Hence, mathematically, the control objective can be stated as

$$p(t) \to p_d, \; p_d = \begin{bmatrix} u_0 & v_0 \end{bmatrix}^T \text{ and } z(t) \to 0 \tag{6}$$

where $u_0, v_0 \in \mathbb{R}$ denote the pixel coordinates of the principal point of the CiH.

The actual kinematic control inputs for the robot can be obtained from the subsequently developed camera control velocities using the extrinsic calibration parameters as (Malis and Chaumette (2002))

$$\begin{bmatrix} v_c \\ \omega_c \end{bmatrix} = \begin{bmatrix} R_e & [t_e]_\times R_e \\ \mathbf{0} & R_e \end{bmatrix} \begin{bmatrix} v_r \\ \omega_r \end{bmatrix} \tag{7}$$

where $v_c(t)$, $\omega_c(t) \in \mathbb{R}^3$ denote the linear and angular velocity of the camera; $v_r(t)$, $\omega_r(t) \in \mathbb{R}^3$ denote the linear and angular velocity of the robot; $R_e \in \mathbb{R}^{3\times3}$ denotes the known constant rotation between the CiH and robot end-effector frames; and $[t_e]_\times \in \mathbb{R}^{3\times3}$ is the skew symmetric matrix of $t_e \in \mathbb{R}^3$, which represents the known constant translation vector between the CiH and robot frames.

4. CONTROLLER DEVELOPMENT

Based on the control objective, a decoupled rotation and a hybrid translation controller is developed. The fixed camera can view an entire tree canopy and select a fruit to be harvested. However, the target fruit selected by the fixed camera may not be visible to the CiH; therefore, the CiH needs to be oriented along the target fruit, which is achieved using the rotation controller. Following the orientation, the CiH is regulated to the desired fruit position using a hybrid translation controller.

4.1 Rotation Controller

The rotation error $e_\omega(t) \in \mathbb{R}^3$, defined as the orientation mismatch that brings the target fruit into the field-of-view of the CiH, can be expressed in terms of an angle-axis representation as

$$e_\omega \triangleq u\theta \tag{8}$$

where $u(t) \in \mathbb{R}^3$ represents a unit axis of rotation such that $u(t) = \overrightarrow{\hat{\hat{m}}}'(t) \wedge [\,0\ 0\ 1\,]^T$, and $\theta(t) \in \mathbb{R}$ is the angle of rotation about $u(t)$ defined as $\theta(t) = \cos^{-1}\langle \overrightarrow{\hat{\hat{m}}}'(t), [\,0\ 0\ 1\,]^T \rangle$, such that $0 \leq \theta(t) < 2\pi$. $\overrightarrow{\hat{\hat{m}}}'(t) \in \mathbb{R}^3$ represents a unit vector along $\hat{\hat{m}}'(t)$, which is defined in Remark 1. Taking the time derivative of (8), the open-loop error dynamics for $e_\omega(t)$ can be obtained as

$$\dot{e}_\omega = -L_\omega \omega_c \tag{9}$$

where $L_\omega\,(t) \in \mathbb{R}^{3\times 3}$ is defined as

$$L_\omega \triangleq I_3 - \frac{\theta}{2}[u]_\times + \left(1 - \frac{\text{sinc}(\theta)}{\text{sinc}^2(\frac{\theta}{2})}\right)[u]_\times^2. \tag{10}$$

In (10), the $\text{sinc}(\theta)$ term denotes the unnormalized sinc function. Given the open-loop rotation error dynamics in (9) and the subsequent stability analysis, the control input $\omega_c(t)$ can be designed as

$$\omega_c = \Lambda_\omega \text{diag}(|e_\omega|^{\alpha_1})\text{sgn}(e_\omega) \tag{11}$$

where $\Lambda_\omega \in \mathbb{R}_{>0}$ is the known constant control gain, $\text{diag}(\cdot)$ denotes a diagonal matrix, $\text{sgn}(\cdot)$ is a signum function, and $\alpha_1 \in \mathbb{R}$ is a known constant such that $\alpha_1 \in (0, 1)$. It must be pointed out that even in the presence of the discontinuous term $\text{sgn}(e_\omega)$ the angular velocity control input $\omega_c(t)$ in (11) is a continuous function of time. Substituting (11) into (9) gives the expression for the closed-loop error dynamics as

$$\dot{e}_\omega = -L_\omega \Lambda_\omega \text{diag}(|e_\omega|^{\alpha_1})\text{sgn}(e_\omega). \tag{12}$$

Theorem 1. The rotation controller in (11) guarantees that the orientation of the camera coordinate frame $\mathcal{F}$ is regulated to a desired orientation such that the origin is a practically globally finite-time-stable equilibrium for the closed-loop system in (12).

Proof 1. Consider a positive definite Lyapunov candidate function $V_\omega(e_\omega)$ as

$$V_\omega = \frac{1}{2} e_\omega^T e_\omega. \tag{13}$$

After taking the time-derivative of (13) and substituting (12) in the resulting expression, the Lyapunov derivative can be obtained as

$$\dot{V}_\omega = -e_\omega^T \Lambda_\omega \text{diag}(|e_\omega|^{\alpha_1})\text{sgn}(e_\omega) \tag{14}$$

where the fact that $e_\omega^T L_\omega = e_\omega^T$ is utilized. $e_\omega^T(t)$ in (14) can be written as $e_\omega^T = \text{sgn}(e_\omega^T)\text{diag}(|e_\omega|)$. Therefore, the Lyapunov derivative becomes

$$\dot{V}_\omega = -\text{sgn}(e_\omega^T)\Big[\Lambda_\omega \text{diag}(|e_\omega|)\text{diag}(|e_\omega|^{\alpha_1})\Big]\text{sgn}(e_\omega). \tag{15}$$

It can be observed that the bracketed quantity in (15) is a positive definite symmetric diagonal matrix. As a result, the Lyapunov derivative can be expressed as

$$\begin{aligned}\dot{V}_\omega &= -\text{sgn}(|e_\omega^T|)\Big[\Lambda_\omega \text{diag}(|e_\omega|)\text{diag}(|e_\omega|^{\alpha_1})\Big]\text{sgn}(|e_\omega|) \\ &= -\Lambda_\omega \|e_\omega\|^{\alpha_1+1} = -\Lambda_\omega (2V_\omega)^{\mu_1}\end{aligned} \tag{16}$$

where $\|\cdot\|^p$ denotes the p-norm, and $\mu_1 = (\alpha_1 + 1)/2$. Note that the Lyapunov derivative in (16) is negative definite, $\dot{V}_\omega(e_\omega) < 0$. Based on $V_\omega(e_\omega) > 0$ and $\dot{V}_\omega(e_\omega) < 0$, it can be concluded that $e_\omega(t) \in \mathcal{L}_\infty$ and $e_\omega(t) \in \mathcal{L}_2$. Based on the fact that $e_\omega(t) \in \mathcal{L}_\infty$, (8), (9), (11) and (12) can be used to prove that $u(t), \theta(t), L_\omega(t), \omega_c(t), \dot{e}_\omega(t) \in \mathcal{L}_\infty$. Using $\dot{e}_\omega(t) \in \mathcal{L}_\infty$ and (8), it is clear that $\dot{u}(t), \dot{\theta}(t) \in \mathcal{L}_\infty$. Additionally, with $\Lambda_\omega > 0$ and $\alpha_1 \in (0, 1)$, the rotation controller in (11) is continuous everywhere and locally Lipschitz everywhere except at the origin. Thus, according to Theorem 1 in Bhat and Bernstein (1995), the origin is a practically globally finite-time-stable equilibrium of the closed-loop system in (12). The term practically global is used in lieu of global since the result is not valid for the singular point $\theta = 2\pi$ associated with the angle of rotation (Fang et al., 2006). The upper bound on the convergence time for $e_\omega(0) \to 0$ can be obtained as (Bhat and Bernstein, 1995)

$$t_\omega(e_\omega(0)) \leq \frac{1}{\Lambda_\omega 2_1^\mu (1-\mu_1)} V_\omega(e_\omega(0))^{1-\mu_1} = t'_\omega \tag{17}$$

where $t'_\omega \in \mathbb{R}_{>0}$. Hence, the developed rotation controller guarantees that the orientation of the camera is regulated to the desired orientation such that the target fruit becomes visible to the CiH. ■

4.2 Translation Controller

After taking time derivative of the first expression in (4), the velocity $\dot{p}(t) \in \mathbb{R}^2$ of the fruit centroid in the image plane of the CiH can be obtained as

$$\dot{p} = \begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} = -\frac{1}{z} J_v v_c - J_\omega \omega_c \tag{18}$$

where $J_v(u, v), J_\omega(u, v) \in \mathbb{R}^{2\times 3}$ are the measurable image Jacobians that relate the linear and angular velocity, $v_c(t)$ and $\omega_c(t)$, respectively, of the CiH to the image plane velocity of the fruit centroid. Since no orientation change is required during translation control, the image dynamics for translation control can be obtained by substituting $\omega_c(t) = 0$ in (18) as

$$\dot{p} = -\frac{1}{z} J'_v v'_c - \frac{1}{z} J''_v v_{cz} \tag{19}$$

where $v'_c(t) = [\,v_{cx}\ v_{cy}\,]^T \in \mathbb{R}^2$ and $v_{cz}(t) \in \mathbb{R}$ are the velocities of the CiH along the image $x-$ and $y-$axis and along the optical axis, respectively. In (19), $J'_v \in \mathbb{R}^{2\times 2}$ and $J''_v(u, v) \in \mathbb{R}^2$ are the measurable Jacobians given by

$$J'_v = \begin{bmatrix} \lambda_x f & 0 \\ 0 & \lambda_y f \end{bmatrix} \quad J''_v = -\begin{bmatrix} u \\ v \end{bmatrix}. \tag{20}$$

Based on the control objective in (6), the translation error $e_v(t) = [e_{v1}(t)\ e_{v2}(t)]^T \in \mathbb{R}^3$ can be defined as

$$e_{v1} \triangleq p - p_d \tag{21}$$

$$e_{v2} \triangleq \hat{z}. \tag{22}$$

From Remark 2, the measurable error signal $e_{v2}(t)$ ensures that the depth for regulating the camera is greater than or equal to the depth of the target fruit in $\mathcal{F}$, i.e., $\hat{z}(t) \geq z(t)$. The choice of $e_{v2}(t)$ is appropriate in our case since the end-effector is equiped with an infrared proximity sensor that detects when the target fruit is reached to stop the motion of the robot.

Control input $\mathbf{v_z(t)}$: Taking time-derivative of (22), and using the depth ratio ξ defined in Proposition 1, the open-loop error dynamics for $e_{v2}(t)$ can be obtained as

$$\dot{e}_{v2} = -\xi v_{cz}. \tag{23}$$

Based on the above dynamics, the linear velocity $v_{cz}(t)$ along the optical axis of the CiH is designed as

$$v_{cz} = \Lambda_z \text{diag}(|e_{v2}|^{\alpha_2})\text{sgn}(e_{v2}) \tag{24}$$

where $\Lambda_z \in \mathbb{R}_{>0}$ is the constant control gain, and $\alpha_2 \in \mathbb{R}$ is a known constant such that $\alpha_2 \in (0,1)$. Substituting (24) into (23), the closed-loop error system for $e_{v2}(t)$ can be obtained as

$$\dot{e}_{v2} = -\xi \Lambda_z \text{diag}(|e_{v2}|^{\alpha_2})\text{sgn}(e_{v2}). \tag{25}$$

Control input $\mathbf{v_c'(t)}$: Taking time-derivative of (21) and substituting (19) into the resulting expression, the open-loop error system for $e_{v1}(t)$ can be obtained as

$$\dot{e}_{v1} = -\frac{1}{z} J_v' v_c' - \frac{1}{z} J_v'' v_{cz}. \tag{26}$$

Based on the open-loop error system in (26) and the subsequent stability analysis, the linear control velocity $v_c'(t) \in \mathbb{R}^2$ of the CiH, along the x- and y-axis of $\mathcal{F}$, can be designed as

$$v_c'(t) = J_v'^{-1}\Big(\hat{z}\Lambda_{xy}\text{diag}(|e_{v1}|^{\alpha_3})\text{sgn}(e_{v1}) - J_v'' v_{cz}\Big) \tag{27}$$

where $\Lambda_{xy} \in \mathbb{R}_{>0}$ is the constant control gain, and $\alpha_3 \in \mathbb{R}$ is a known constant such that $\alpha_3 \in (0,1)$. After substituting the control input in (27) into (26), and using $\hat{z}/z = \xi$, the closed-loop error system for $e_{v1}(t)$ can be obtained as

$$\dot{e}_{v1} = -\xi \Lambda_{xy} \text{diag}(|e_{v1}|^{\alpha_3})\text{sgn}(e_{v1}). \tag{28}$$

Theorem 2. The translation control inputs developed in (24) and (27) ensure that the robot is regulated to the target fruit in finite time.

Proof 2. To prove that the camera is regulated to the desired depth, consider $V_z(t)$ to be the following positive definite Lyapunov candidate function:

$$V_z = \frac{1}{2} e_{v2}^T e_{v2}. \tag{29}$$

Taking the time-derivative of $V_z(t)$, substituting (25), and writing $e_{v2}^T(t)$ as $\text{sgn}(e_{v2}^T)\text{diag}(|e_{v2}|)$, the Lyapunov derivative can be obtained as

$$\dot{V}_z = -\text{sgn}(e_{v2}^T)\Big[\text{diag}(|e_{v2}|)\xi\Lambda_z\text{diag}(|e_{v2}|^{\alpha_2})\Big]\text{sgn}(e_{v2}). \tag{30}$$

Since $\xi > 0$, it can be observed that the bracketed quantity in (15) is a positive definite symmetric diagonal matrix. As a result, the Lyapunov derivative can be expressed as

$$\begin{aligned}\dot{V}_z &= -\text{sgn}(|e_{v2}^T|)\Big[\xi\Lambda_z\text{diag}(|e_{v2}|)\text{diag}(|e_{v2}|^{\alpha_2})\Big]\text{sgn}(|e_{v2}|) \\ &= -\xi\Lambda_z\|e_{v2}\|^{\alpha_2+1} = -\xi\Lambda_z(2V_z)^{\mu_2}\end{aligned} \tag{31}$$

where $\mu_2 = \alpha_2 + 1$. Using similar signal chasing arguments as presented in Theorem 1, it can be shown that $e_{v2}(t) \in \mathcal{L}_\infty$ and $e_{v2}(t) \in \mathcal{L}_2$. Based on the fact that $e_{v2}(t) \in \mathcal{L}_\infty$, (22), (24), and (25) can be used to prove that $\hat{z}(t), v_{cz}(t), \dot{e}_{v2}(t) \in \mathcal{L}_\infty$. Further, with $\Lambda_z > 0$ and $\alpha_2 \in (0,1)$, the controller in (24) is continuous everywhere and locally Lipschitz everywhere except at the origin. Thus, according to Theorem 1 in Bhat and Bernstein (1995), the origin is a globally finite-time-stable equilibrium of (25). Using the bound $\underline{\xi} = 1$ defined in Proposition 1, the upper bound on the convergence time for $e_{v2}(0) \to 0$ can be obtained from Bhat and Bernstein (1995) as

$$t_z(e_{v2}(0)) \leq \frac{1}{\Lambda_z 2_2^{\mu}(1-\mu_2)} V_z(e_{v2}(0))^{1-\mu_2} = t_z' \tag{32}$$

where $t_z' \in \mathbb{R}_{>0}$.

To prove that the fruit is regulated to the principal point, consider a positive definite Lyapunov candidate function $V_{xy}(t)$ as

$$V_{xy} = \frac{1}{2} e_{v1}^T e_{v2}. \tag{33}$$

Taking the time-derivative of (33), substituting (28), and writing $e_{v1}^T(t)$ as $\text{sgn}(e_{v1}^T)\text{diag}(|e_{v1}|)$, the Lyapunov derivative can be obtained as

$$\begin{aligned}\dot{V}_{xy} &= -\text{sgn}(|e_{v1}^T|)\Big[\xi\Lambda_{xy}\text{diag}(|e_{v1}|)\text{diag}(|e_{v1}|^{\alpha_3})\Big]\text{sgn}(|e_{v1}|) \\ &= -\xi\Lambda_{xy}\|e_{v1}\|^{\alpha_3+1} = -\xi\Lambda_{xy}(2V_{xy})^{\mu_3}\end{aligned} \tag{34}$$

where the fact that $\xi, \hat{z}(t) > 0$ are used, and $\mu_3 = \alpha_3 + 1$. Using signal chasing arguments and the previously proved facts that $v_{cz}(t) \in \mathcal{L}_\infty$ and $z(t) \in \mathcal{L}_\infty$, it can be proved that the origin is a globally finite-time-stable equilibrium of the closed-loop system in (28). Further, the upper bound on the convergence time for $e_{v1}(0) \to 0$ can be obtained as

$$t_{xy}(e_{v1}(0)) \leq \frac{1}{\Lambda_{xy} 2_3^{\mu}(1-\mu_3)} V_{xy}(e_{v1}(0))^{1-\mu_3} = t_{xy}' \tag{35}$$

where $t_{xy}' \in \mathbb{R}_{>0}$. ■

From the stability analysis, it is proved that the translation controller in (24) and (27) regulates the robot to a target fruit in finite time. The worst-case upper bound on the convergence time for the largest possible errors $e_{v1}(0)$ and $e_{v2}(0)$ encountered in harvesting operation can be obtained as $\max(t_{xy}, t_z)$. Based on which, the convergence times in (32) and (35) can be adjusted to achieve a desired bound on the picking time by appropriately selecting Λ_z and Λ_{xy}, respectively.

5. SIMULATION RESULTS

A numerical simulation was performed to demonstrate the performance of the proposed robust controller for citrus harvesting. The initial position $t_r \in \mathbb{R}^3$ and orientation $R_r \in \mathbb{R}^{3\times3}$ of the CiH coordinate frame $\mathcal{F}$ with respect to $\mathcal{F}_b$ was considered to be

$$t_r = [\,20\ 60\ 1500\,]^T,\ R_r = \begin{bmatrix} 0.4698 & -0.1955 & 0.8608 \\ 0.1710 & 0.9769 & 0.1285 \\ -0.8660 & 0.0868 & 0.4924 \end{bmatrix}. \tag{36}$$

The position of the target fruit in $\mathcal{F}_b$ was assumed to be $O^* = [500\ 400\ 2500]^T$. The fruit pixel coordinates were assumed to be affected by a zero-mean Gaussian sensor noise of standard deviation of 1 pixel. Fig. 3 shows the time-varying image-space trajectory of the fruit in the CiH during rotation (dotted line) and translation control (continuous line). The fruit pixel coordinates are regulated to the principal point of the CiH, $pd = [320\ 240]^T$, so that the fruit can be harvested. The rotation and translation error plots are shown in Figs. 4 and 5, respectively, where it is to be noted that the rotation and translation errors go to zero in finite time. It can be seen that for the presented simulation scenario, the fruit can be reached in about 2s.

6. CONCLUSION

A terminal sliding mode based visual servo control approach is introduced for robotic harvesting. The controller ensures finite time regulation of the robotic manipulator to a target fruit thereby a bound on the picking time can be established.

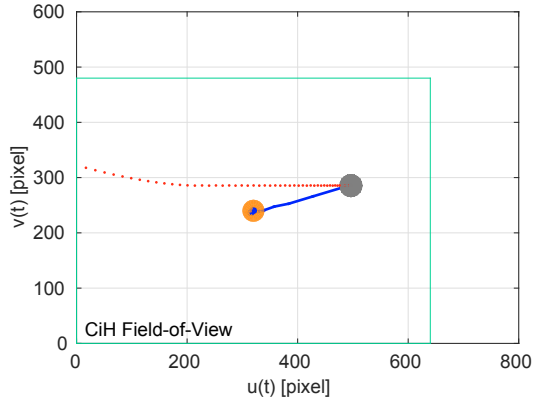

Fig. 3. Image-space trajectory of the target fruit in the CiH during rotation (dotted red line) and translation control (continuous blue line). The position of the fruit in the image plane after rotation control is shown in gray, and the final position of the fruit is shown in orange.

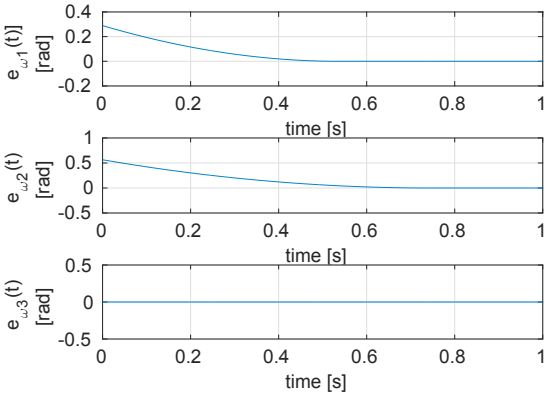

Fig. 4. Rotation error $e_\omega(t) = [e_{\omega 1}(t)\ e_{\omega 2}(t)\ e_{\omega 3}(t)]^T$.

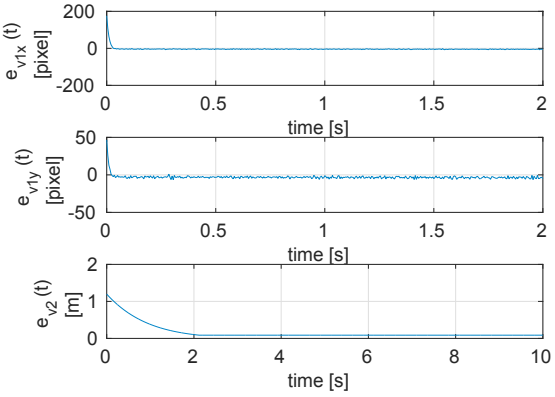

Fig. 5. Translation error $e_v(t) = [e_{v1}^T(t)\ e_{v2}(t)]^T$, where $e_{v1}(t) = [e_{v1x}(t)\ e_{v1y}(t)]^T$.

Future work will include designing a coupled translation controller, for motions along the image plane and optical axis, to provide a more informative bound on the convergence time.

REFERENCES

C Wouter Bac, Eldert J Henten, Jochen Hemming, and Yael Edan. Harvesting robots for high-value crops: State-of-the-art review and challenges ahead. *Journal of Field Robotics*, 31(6):888–911, 2014.

Ruud Barth, Jochen Hemming, and Eldert J van Henten. Design of an eye-in-hand sensing and servo control framework for harvesting robotics in dense vegetation. *Biosystems Engineering*, 146:71–84, 2016.

Sanjay P Bhat and Dennis S Bernstein. Lyapunov analysis of finite-time differential equations. In *American Control Conference, Proceedings of the 1995*, volume 3, pages 1831–1832. IEEE, 1995.

Duke M Bulanon, Hiroshi Okamoto, and Shun-Ichi Hata. Feedback control of manipulator using machine vision for robotic apple harvesting. In *2005 ASAE Annual Meeting*, page 1. American Society of Agricultural and Biological Engineers, 2005.

Wei Chen, Tongqing Xu, Junjie Liu, Mo Wang, and Dean Zhao. Picking robot visual servo control based on modified fuzzy neural network sliding mode algorithms. *Electronics*, 8(6):605, 2019.

Zhao De-An, Lv Jidong, Ji Wei, Zhang Ying, and Chen Yu. Design and control of an apple harvesting robot. *Biosystems engineering*, 110(2):112–122, 2011.

Yael Edan, Dima Rogozin, Tamar Flash, and Gaines E Miles. Robotic melon harvesting. *IEEE Transactions on Robotics and Automation*, 16(6):831–835, 2000.

Y. Fang, W.E. Dixon, D.M. Dawson, and J. Chen. An exponential class of model-free visual servoing controllers in the presence of uncertain camera calibration. *International Journal of Robotics and Automation*, 21(4):247–255, 2006.

Mario M Foglia and Giulio Reina. Agricultural robot for radicchio harvesting. *Journal of Field Robotics*, 23(6-7): 363–377, 2006.

R. C. Harrell, P. D. Adsit, and D. C. Slaughter. Real-time vision-servoing of a robotic tree fruit harvester. *ASAE paper*, pages 85–3550, 1985.

R. C. Harrell, D. C. Slaughter, and P. D. Adsit. A fruit-tracking system for robotic harvesting. *Machine Vision and Applications*, 2(2):69–80, 1989.

R. C. Harrell, P. D. Adsit, T. A. Pool, R. Hoffman, et al. The Florida robotic grove-lab. *Transactions of the ASAE*, 33 (2):391–399, 1990.

S. Hayashi, K. Ganno, Y. Ishii, and I. Tanaka. Robotic harvesting system for eggplants. *Japan Agricultural Research Quarterly*, 36:163–168, 2002.

E. Malis and F. Chaumette. Theoretical improvements in the stability analysis of a new class of model-free visual servoing methods. *Robotics and Automation, IEEE Transactions on*, 18(2):176–186, Apr. 2002.

S. S. Mehta and T. F. Burks. Vision-based control of robotic manipulator for citrus harvesting. *Computers and Electronics in Agriculture*, 102:146–158, 2014.

S. S. Mehta and T. F. Burks. Adaptive visual servo control of robotic harvesting systems. *IFAC-PapersOnLine*, 49 (16):287–292, 2016.

S. S. Mehta, W. MacKunis, and T. F. Burks. Robust visual servo control in the presence of fruit motion for robotic citrus harvesting. *Computers and Electronics in Agriculture*, 123:362–375, 2016.

Yuanshen Zhao, Liang Gong, Yixiang Huang, and Chengliang Liu. A review of key techniques of vision-based control for harvesting robot. *Computers and Electronics in Agriculture*, 127:311–323, 2016.