Image-Based Visual Servoing for Enhanced Cooperation of Dual-Arm Manipulation

Zizhe Zhang ⓘ, *Graduate Student Member, IEEE*, Yuan Yang ⓘ, *Member, IEEE*, Wenqiang Zuo, Guangming Song ⓘ, *Senior Member, IEEE*, Aiguo Song ⓘ, *Senior Member, IEEE*, Yang Shi ⓘ, *Fellow, IEEE*

Abstract—The cooperation of a pair of robot manipulators is required to manipulate a target object without any fixtures. The conventional control methods coordinate the end-effector pose of each manipulator with that of the other using their kinematics and joint coordinate measurements. Yet, the manipulators' inaccurate kinematics and joint coordinate measurements can cause significant pose synchronization errors in practice. This paper thus proposes an image-based visual servoing approach for enhancing the cooperation of a dual-arm manipulation system. On top of the classical control, the visual servoing controller lets each manipulator use its carried camera to measure the image features of the other's marker and adapt its end-effector pose with the counterpart on the move. Because visual measurements are robust to kinematic errors, the proposed control can reduce the end-effector pose synchronization errors and the fluctuations of the interaction forces of the pair of manipulators on the move. Theoretical analyses have rigorously proven the stability of the closed-loop system. Comparative experiments on real robots have substantiated the effectiveness of the proposed control.

Index Terms—Cooperative manipulation, image-based visual servoing, multirobot systems, passivity-based control.

I. INTRODUCTION

Dual-arm manipulation is a challenging research topic in the robotics society [1], [2]. Given that typical humanoid robots have bimanual capabilities, the study of dual-arm manipulation control is gaining more and more attention recently [3]. A typical application of dual-arm manipulation is to grasp and transport a rigid target object without extra fixtures [4]. In this context, the two arms need to keep their end effectors in frictional contact with the manipulated object throughout the task execution [5]. Thus, a fundamental control objective for the pair of two manipulators is to tightly coordinate their end-effector poses on the move [6].

The existing research on dual-arm manipulation control has synchronized the end-effector positions of both manipulators with event-triggered communications [7] and regulated their interaction forces with the object in the contact direction [8]. Accounting for the internal forces [9], constrained quadratic

Zizhe Zhang is with the Department of Electrical and System Engineering, University of Pennsylvania, Philadelphia, PA 19104, United States (e-mail: zizhez@seas.upenn.edu).

Yuan Yang, Guangming Song and Aiguo Song are with the School of Instrument Science and Engineering, Southeast University, Nanjing, Jiangsu 210096, China (e-mail: yuan_evan_yang@seu.edu.cn; mikesong@seu.edu.cn; a.g.song@seu.edu.cn).

Wenqiang Zuo is with the School of Materials Science and Engineering, Southeast University, Nanjing, Jiangsu 211189, China (e-mail: wenqiangzuo@seu.edu.cn).

Yang Shi is with the Department of Mechanical Engineering, University of Victoria, Victoria, BC V8W 2Y2, Canada (e-mail: yshi@uvic.ca).
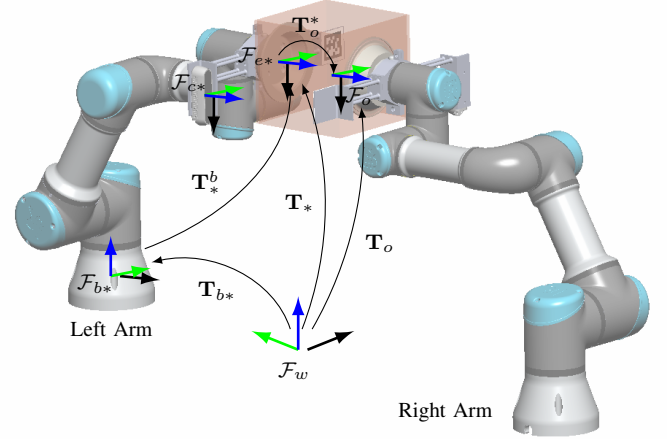
Figure 1. Two 6-DOF serial manipulators serve as the left and the right arms in a dual-arm manipulation system. The target object that they manipulate is a cubic box. $\mathcal{F}_w$ and $\mathcal{F}_o$ denote the world frame and the object frame, respectively. $\mathcal{F}_{b*}$, $\mathcal{F}_{e*}$ and $\mathcal{F}_{c*}$ are the base, end-effector and camera frames of the left $* = l$ and the right $* = r$ arms. $\mathbf{T}_o$, $\mathbf{T}_{b*}$, $\mathbf{T}_*$, $\mathbf{T}_*^b$ and $\mathbf{T}_o^*$ indicate the transformations between the frames.

programs have been formulated to minimize the contact forces while avoiding obstacles in the surrounding environment [10] and the separation or sliding of the frictional contact [11]. Under nonlinear model predictive control, a torque-controlled dual-arm manipulation system has recently been tested for a mirror grinding task [12]. However, torque-controlled manipulators with exactly known dynamic models are unavailable in many applications. Cooperation approaches for velocity- and position-controlled robot arms have thus been advocated. For velocity-controlled manipulators, the joint limit avoidance and Cartesian-space path following have been pursued with lower priority than their pose synchronization task [13]. A sparse kinematic control has minimized the number of the two arms' actuated joints using a hierarchical quadratic program [14]. For position-controlled manipulators, a linear elastic model for the contact of the follower arm with the object has been assumed to minimize the internal forces [15]. A communication-free coordination control has adopted the manipulators' interaction forces to transport deformable objects [16].

A position-based visual servoing (PBVS) approach has been deployed on a humanoid robot to control both arms/hands to grasp and manipulate its observed target objects in a kitchen environment [17]. An image-based visual servoing (IBVS) method has been applied to a dual-arm robot with both fixed and mobile cameras to align the poses of both arms when

performing a plug-socket insertion task [18]. Yet, the control design and stability analysis have not considered the closed kinematic chain. With IBVS, two simulated aerial manipulators have rigidly grasped an assembly bar and transported it to the target pose [19]. However, the control has not yet been verified on any real-robot platform.

A. Contributions

This paper develops an IBVS control algorithm for enhancing dual-arm cooperation. Its contributions to the coordination control and the visual servoing control are as follows.

The classical cooperative manipulation controls synchronize the end-effector poses of both arms [20]. However, previous studies [21], [22] have discovered that the kinematic errors of manipulators can induce significant pose biases between them and lead to possible task failures and hardware damage. Recursive least-squares estimators have estimated unknown kinematic parameters of space manipulators while suffering from joint measurement noises [23]. Because IBVS is robust to kinematic errors, the paper proposes an IBVS algorithm to reduce the pose errors and the accompanying interaction force fluctuations during dual-arm manipulation.

The basic [24] and advanced [25] visual servoing algorithms have been widely employed for a single manipulator to reach a stationary target or track a moving target. The interaction matrices depend only on the image features and their depths. However, in the context of dual-arm manipulation, the cameras and markers are mounted on two arms separately. As a result, the kinematic model of the system becomes entirely different, and the associated interaction matrix significantly differs from the classical one. The paper figures out the dependence of the customized interaction matrix on the poses of both arms explicitly. Moreover, it synthesizes the IBVS algorithm with pose synchronization control while guaranteeing the closed-loop system stability rigorously.

In the remainder of the paper, Section II builds the kinematic models of the manipulators and the vision system. Section III designs the control algorithm and analyses the system stability. Section IV discusses the state and input constraints of manipulators and some visual servoing issues. Section V validates the control via comparative experiments. Finally, Section VI concludes the design and future studies.

II. Modelling

This section introduces the closed kinematic chain formed by the pair of manipulators and the pinhole perspective projection model of their wrist-mounted cameras.

A. The Closed Kinematic Chain

The dual-arm manipulation system, as shown in Figure 1, contains a pair of 6-degrees-of-freedom (6-DOF) serial manipulators with only revolute joints. Let the world frame be $\mathcal{F}_w$, the base and the end-effector frames of the left $* = l$ and the right $* = r$ arms be $\mathcal{F}_{b*}$ and $\mathcal{F}_{e*}$, respectively. Let $\mathbf{q}_*$ be the joint coordinates and

$$\mathbf{T}_* = \begin{bmatrix} \mathbf{R}_* & \mathbf{x}_* \\ \mathbf{0} & 1 \end{bmatrix}$$

be the end-effector poses of the two arms $* = l, r$, where $\mathbf{R}_*$ and $\mathbf{x}_*$ are the rotation matrix representations of orientation and the positions respectively. The Jacobian matrices $\mathbf{J}_*$ of the two manipulators $* = l, r$ then link their joint velocity inputs $\dot{\mathbf{q}}_*$ to their end-effector velocities $\mathbf{v}_*$ by

$$\mathbf{v}_* = \begin{bmatrix} \boldsymbol{v}_* \\ \boldsymbol{\omega}_* \end{bmatrix} = \mathbf{J}_* \dot{\mathbf{q}}_*, \tag{1}$$

where the linear $\boldsymbol{v}_*$ and angular $\boldsymbol{\omega}_*$ velocities adapt their end-effector poses $\mathbf{T}_*$ by

$$\dot{\mathbf{x}}_* = \boldsymbol{v}_*, \tag{2a}$$
$$\dot{\mathbf{R}}_* = \boldsymbol{\omega}_*^\times \mathbf{R}_*. \tag{2b}$$

To manipulate the target object without fixtures, the two manipulators need to keep their end effectors rigidly attached to the object throughout the task execution. Let $\mathcal{F}_o$ be the object frame, and let $\mathbf{T}_o$ be the object pose. Then, the rigid attachment of the arms' end effectors to the object requires

$$\mathbf{T}_l \mathbf{T}_o^l = \mathbf{T}_o = \mathbf{T}_r \mathbf{T}_o^r, \tag{3}$$

where the poses of the object relative to the end effectors of the left $\mathbf{T}_o^l$ and the right $\mathbf{T}_o^r$ arms remain invariant. Simple algebraic operations derive that the transformations

$$\mathbf{T}_r^l = \mathbf{T}_l^{-1} \mathbf{T}_r = \mathbf{T}_o^l \mathbf{T}_o^{r-1}, \tag{4a}$$
$$\mathbf{T}_l^r = \mathbf{T}_r^{-1} \mathbf{T}_l = \mathbf{T}_o^r \mathbf{T}_o^{l-1} \tag{4b}$$

remain invariant. Namely, the two arms' end effectors and the object shall stay stationary with respect to each other. For the left arm $* = l$, the right arm $* = r$, and the object $* = o$, let

$$\overline{\mathbf{T}}_* = \begin{bmatrix} \overline{\mathbf{R}}_* & \overline{\mathbf{x}}_* \\ \mathbf{0} & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{T}_{*d} = \begin{bmatrix} \mathbf{R}_{*d} & \mathbf{x}_{*d} \\ \mathbf{0} & 1 \end{bmatrix}$$

denote their initial and desired poses, respectively. The pose synchronization control objective of the dual-arm manipulation system can thus be formulated as

$$\mathbf{T}_{od}^{-1} \mathbf{T}_o = \mathbf{I}, \tag{5a}$$
$$\overline{\mathbf{T}}_l^{-1} \mathbf{T}_l = \overline{\mathbf{T}}_r^{-1} \mathbf{T}_r = \overline{\mathbf{T}}_o^{-1} \mathbf{T}_o. \tag{5b}$$

In Equation (5a), $\mathbf{T}_{od}^{-1} \mathbf{T}_o$ is the displacement of the object from its desired pose $\mathbf{T}_{od}$. It thus enforces the object to move towards its desired pose. In Equation (5b), $\overline{\mathbf{T}}_*^{-1} \mathbf{T}_*$ denote the displacements of the left arm $* = l$, the right arm $* = r$, and the object $* = o$ from their initial poses $\overline{\mathbf{T}}_*$, respectively. As the two manipulators reliably hold the object at the initial poses $\overline{\mathbf{T}}_*$, the objective (5b) then secures the manipulation by enforcing them to keep their relative poses invariant.

B. The Pinhole Projection Model

The eye-in-hand camera mounted on the wrist of each manipulator, as shown in Figure 2, measures the image features of the fiducial marker fixed on the wrist of the other manipulator. In particular, the image features adopted in the control design are the four corner points of each AprilTag marker. Let the pose of the camera frame $\mathcal{F}_{c*}$ relative to the end-effector frame $\mathcal{F}_{e*}$ of the left $* = l$ or the right $* = r$ arm be

$$\mathbf{T}_c^* = \begin{bmatrix} \mathbf{R}_c^* & \mathbf{x}_c^* \\ \mathbf{0} & 1 \end{bmatrix}.$$
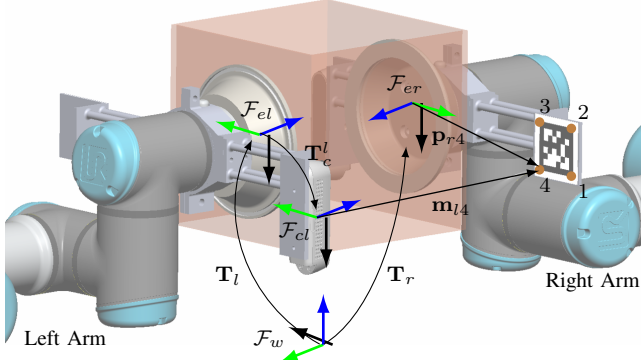
Figure 2. The positions of the corner points $i = 1, 2, 3, 4$ expressed of the right arm's marker expressed in the end-effector frame $\mathcal{F}_{er}$ and in the left camera frame $\mathcal{F}_{cl}$ are $\mathbf{p}_{ri}$ and $\mathbf{m}_{li}$, respectively. $\mathbf{T}_*$ and $\mathbf{T}_c^l$ denote the transformations between the world frame $\mathcal{F}_w$, the end-effector frames $\mathcal{F}_{e*}$ and the camera frame $\mathcal{F}_{cl}$ with $* = l, r$.

Let the position of the i-th corner point of the marker on the right $* = r$ or the left $* = l$ arm be expressed in the end-effector frame $\mathcal{F}_{e*}$ by $\mathbf{p}_{*i}$ with $i = 1, 2, 3, 4$.

Let the superscripts ^{x}, ^{y} and ^{z} index the components of position vectors in the associated axis. By the kinematic chain, the position $\mathbf{m}_{li} = (m_{li}^x, m_{li}^y, m_{li}^z)^\mathsf{T}$ of the i-th corner point of the right arm's marker expressed in the left arm's camera frame $\mathcal{F}_{cl}$ meets

$$\mathbf{T}_l \mathbf{T}_c^l \begin{bmatrix} \mathbf{m}_{li} \\ 1 \end{bmatrix} = \mathbf{T}_r \begin{bmatrix} \mathbf{p}_{ri} \\ 1 \end{bmatrix}.$$

The position $\mathbf{m}_{ri} = (m_{ri}^x, m_{ri}^y, m_{ri}^z)^\mathsf{T}$ of the i-th corner point of the left arm's marker expressed in the right arm's camera frame $\mathcal{F}_r$ meets

$$\mathbf{T}_r \mathbf{T}_c^r \begin{bmatrix} \mathbf{m}_{ri} \\ 1 \end{bmatrix} = \mathbf{T}_l \begin{bmatrix} \mathbf{p}_{li} \\ 1 \end{bmatrix}.$$

One can thus derive that

$$\mathbf{m}_{li} = \left(\mathbf{R}_l \mathbf{R}_c^l \right)^\mathsf{T} \left(\mathbf{R}_r \mathbf{p}_{ri} + \mathbf{x}_r - \mathbf{R}_l \mathbf{x}_c^l - \mathbf{x}_l \right), \quad (6a)$$

$$\mathbf{m}_{ri} = \left(\mathbf{R}_r \mathbf{R}_c^r \right)^\mathsf{T} \left(\mathbf{R}_l \mathbf{p}_{li} + \mathbf{x}_l - \mathbf{R}_r \mathbf{x}_c^r - \mathbf{x}_r \right), \quad (6b)$$

for the corner points $i = 1, 2, 3, 4$ of markers.

Let $\boldsymbol{\zeta}_{*i}$ normalize $\mathbf{m}_{*i}$ with $* = l, r$ by their z-coordinates

$$\boldsymbol{\zeta}_{*i} = \begin{bmatrix} \dfrac{m_{*i}^x}{m_{*i}^z} & \dfrac{m_{*i}^y}{m_{*i}^z} \end{bmatrix}^\mathsf{T}.$$

Let the intrinsic parameters matrices of the left $* = l$ and the right $* = r$ cameras be $\mathbf{K}_*$. The cameras then capture their image features $\mathbf{s}_{*i}$ with $i = 1, 2, 3, 4$ by

$$\mathbf{s}_{*i} = \mathbf{K}_* \boldsymbol{\zeta}_{*i}. \quad (7)$$

Taking the time derivative of Equation (7) yields

$$\dot{\mathbf{s}}_{*i} = \frac{1}{m_{*i}^z} \begin{bmatrix} \mathbf{K}_* & -\mathbf{s}_{*i} \end{bmatrix} \dot{\mathbf{m}}_{*i}. \quad (8)$$

Substituting Equation (6) into Equation (8) gives rise to

$$\begin{bmatrix} \dot{\mathbf{s}}_{li} \\ \dot{\mathbf{s}}_{ri} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_{li}^l & \mathbf{L}_{li}^r \\ \mathbf{L}_{ri}^l & \mathbf{L}_{ri}^r \end{bmatrix} \begin{bmatrix} \mathbf{v}_l \\ \mathbf{v}_r \end{bmatrix}, \quad (9)$$

where the interaction matrices $\mathbf{L}_{li}^l$, $\mathbf{L}_{li}^r$, $\mathbf{L}_{ri}^l$ and $\mathbf{L}_{ri}^r$ with $i = 1, 2, 3, 4$ are expressed as

$$\mathbf{L}_{li}^l = \frac{1}{m_{li}^z} \begin{bmatrix} \mathbf{K}_l & -\mathbf{s}_{li} \end{bmatrix} \left(\mathbf{R}_l \mathbf{R}_c^l \right)^\mathsf{T} \begin{bmatrix} -\mathbf{I} & (\mathbf{R}_r \mathbf{p}_{ri} + \mathbf{x}_r - \mathbf{x}_l)^\times \end{bmatrix},$$

$$\mathbf{L}_{li}^r = \frac{1}{m_{li}^z} \begin{bmatrix} \mathbf{K}_l & -\mathbf{s}_{li} \end{bmatrix} \left(\mathbf{R}_l \mathbf{R}_c^l \right)^\mathsf{T} \begin{bmatrix} \mathbf{I} & -(\mathbf{R}_r \mathbf{p}_{ri})^\times \end{bmatrix},$$

$$\mathbf{L}_{ri}^l = \frac{1}{m_{ri}^z} \begin{bmatrix} \mathbf{K}_r & -\mathbf{s}_{ri} \end{bmatrix} \left(\mathbf{R}_r \mathbf{R}_c^r \right)^\mathsf{T} \begin{bmatrix} \mathbf{I} & -(\mathbf{R}_l \mathbf{p}_{li})^\times \end{bmatrix},$$

$$\mathbf{L}_{ri}^r = \frac{1}{m_{ri}^z} \begin{bmatrix} \mathbf{K}_r & -\mathbf{s}_{ri} \end{bmatrix} \left(\mathbf{R}_r \mathbf{R}_c^r \right)^\mathsf{T} \begin{bmatrix} -\mathbf{I} & (\mathbf{R}_l \mathbf{p}_{li} + \mathbf{x}_l - \mathbf{x}_r)^\times \end{bmatrix}.$$

Let $\bar{\mathbf{s}}_{li}$ and $\bar{\mathbf{s}}_{ri}$ be the initial values of $\mathbf{s}_{li}$ and $\mathbf{s}_{ri}$, respectively. The enhanced cooperation control objective of the dual-arm system can thus be formulated as

$$\begin{bmatrix} \mathbf{s}_{li} \\ \mathbf{s}_{ri} \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{s}}_{li} \\ \bar{\mathbf{s}}_{ri} \end{bmatrix} \quad (10)$$

for the corner points $i = 1, 2, 3, 4$ of markers.

III. CONTROL DESIGN AND STABILITY ANALYSIS

This section designs the control inputs to the manipulators and analyses the stability of the closed-loop system.

A. Control Design

Equation (5b) implies that

$$\overline{\mathbf{T}}_l^{-1} \mathbf{T}_{ld} = \overline{\mathbf{T}}_r^{-1} \mathbf{T}_{rd} = \overline{\mathbf{T}}_o^{-1} \mathbf{T}_{od}.$$

Given that $\mathbf{T}_{od}$ is pre-planned, one can then distribute desired poses to the left and the right arms by

$$\mathbf{T}_{ld} = \overline{\mathbf{T}}_l \overline{\mathbf{T}}_o^{-1} \mathbf{T}_{od} \quad \text{and} \quad \mathbf{T}_{rd} = \overline{\mathbf{T}}_r \overline{\mathbf{T}}_o^{-1} \mathbf{T}_{od}.$$

As a result, Equation (5a) is equivalent to

$$\mathbf{T}_{*d}^{-1} \mathbf{T}_* = \begin{bmatrix} \mathbf{R}_{*d}^\mathsf{T} \mathbf{R}_* & \mathbf{R}_{*d}^\mathsf{T} (\mathbf{x}_* - \mathbf{x}_{*d}) \\ \mathbf{0} & 1 \end{bmatrix} = \mathbf{I}$$

for $* = l, r$ and can thus be reformulated as

$$\underset{\mathbf{q}_l, \mathbf{q}_r}{\text{minimize}} \quad V_d, \quad (11)$$

where the objective function is defined by

$$V_d = \frac{1}{2} \sum_{*=l,r} \left[\text{tr} \left(\mathbf{I} - \mathbf{R}_{*d}^\mathsf{T} \mathbf{R}_* \right) + (\mathbf{x}_* - \mathbf{x}_{*d})^\mathsf{T} (\mathbf{x}_* - \mathbf{x}_{*d}) \right]$$

with $\text{tr}(\mathbf{M})$ being the trace of any square matrix $\mathbf{M}$.

Equation (5b) implies also that

$$\overline{\mathbf{T}}_l^{-1} \mathbf{T}_l = \begin{bmatrix} \widetilde{\mathbf{R}}_l & \overline{\mathbf{R}}_l^\mathsf{T} \widetilde{\mathbf{x}}_l \\ \mathbf{0} & 1 \end{bmatrix} = \begin{bmatrix} \widetilde{\mathbf{R}}_r & \overline{\mathbf{R}}_r^\mathsf{T} \widetilde{\mathbf{x}}_r \\ \mathbf{0} & 1 \end{bmatrix} = \overline{\mathbf{T}}_r^{-1} \mathbf{T}_r,$$

where $\widetilde{\mathbf{R}}_* = \overline{\mathbf{R}}_*^\mathsf{T} \mathbf{R}_*$ and $\widetilde{\mathbf{x}}_* = \mathbf{x}_* - \overline{\mathbf{x}}_*$ for $* = l, r$. Namely,

$$\mathbf{I} = \widetilde{\mathbf{R}}_l \widetilde{\mathbf{R}}_r^\mathsf{T} \quad \text{and} \quad \mathbf{0} = \overline{\mathbf{R}}_l^\mathsf{T} \widetilde{\mathbf{x}}_l - \overline{\mathbf{R}}_r^\mathsf{T} \widetilde{\mathbf{x}}_r.$$

It can thus be reformulated as follows

$$\underset{\mathbf{q}_l, \mathbf{q}_r}{\text{minimize}} \quad V_t, \quad (12)$$

where the objective function is defined by

$$V_t = \frac{1}{2} \text{tr} \left(\mathbf{I} - \widetilde{\mathbf{R}}_l \widetilde{\mathbf{R}}_r^\mathsf{T} \right) + \frac{1}{2} \left(\overline{\mathbf{R}}_l^\mathsf{T} \widetilde{\mathbf{x}}_l - \overline{\mathbf{R}}_r^\mathsf{T} \widetilde{\mathbf{x}}_r \right)^\mathsf{T} \left(\overline{\mathbf{R}}_l^\mathsf{T} \widetilde{\mathbf{x}}_l - \overline{\mathbf{R}}_r^\mathsf{T} \widetilde{\mathbf{x}}_r \right).$$

The enhanced cooperation control objective (10) can be reformulated as the following quadratic program

$$\underset{\mathbf{q}_l, \mathbf{q}_r}{\text{minimize}} \quad V_s, \tag{13}$$

where the objective function is defined by

$$V_s = \frac{1}{2} \sum_{i=1}^{4} \left[(\mathbf{s}_{li} - \bar{\mathbf{s}}_{li})^{\mathsf{T}} (\mathbf{s}_{li} - \bar{\mathbf{s}}_{li}) + (\mathbf{s}_{ri} - \bar{\mathbf{s}}_{ri})^{\mathsf{T}} (\mathbf{s}_{ri} - \bar{\mathbf{s}}_{ri}) \right].$$

Therefore, IBVS for enhancing the cooperation of dual-arm manipulation is formulated as solving the optimization problems (11)-(13) simultaneously by

$$\underset{\mathbf{q}_l, \mathbf{q}_r}{\text{minimize}} \quad V = w_d V_d + w_t V_t + w_s V_s, \tag{14}$$

where the objective function V sums V_d, V_t and V_s by positive scalar weights w_d, w_t and w_s, respectively.

Following the steepest descent algorithm, the joint velocity control inputs to the left and the right arms are designed by (15), where $\dot{\mathbf{q}}_l$ and $\dot{\mathbf{q}}_r$ are respectively the negative gradients of the objective function V with respect to $\mathbf{q}_l$ and $\mathbf{q}_r$, and $\mathrm{sk}(\mathbf{M})^{\vee}$ returns the vector associated with the skew-symmetric part of any square matrix $\mathbf{M}$. The next subsection proves the closed-loop system stability.

B. Stability Analysis

One can investigate the stability of the closed-loop dual-arm manipulation system by the objective function V defined in Equation (14). Under the joint velocity control inputs (15), the kinematics (1) and (2) of the left and the right manipulators adapt the objective function V by

$$\dot{V} = w_d \dot{V}_d + w_t \dot{V}_t + w_s \dot{V}_s = -\dot{\mathbf{q}}_l^{\mathsf{T}} \dot{\mathbf{q}}_l - \dot{\mathbf{q}}_r^{\mathsf{T}} \dot{\mathbf{q}}_r \leq 0, \tag{16}$$

where the time derivatives of V_d, V_t and V_s are presented in detail in (17), (18) and (19), respectively. Because the objective function V is nonnegative and its time derivative is nonpositive, one can thus conclude that the states $\mathbf{q}_l$ and $\mathbf{q}_r$ of the system converge to the largest invariant set of

$$\left\{ (\mathbf{q}_l, \mathbf{q}_r) \mid \dot{\mathbf{q}}_l = \dot{\mathbf{q}}_r = \mathbf{0} \right\}$$

by LaSalle's invariance principle. That is, both arms become stationary asymptotically.

From Equations (17)-(19), one can find that the joint velocity control inputs $\dot{\mathbf{q}}_l$ and $\dot{\mathbf{q}}_r$ of the left and the right arms designed in (15) can be written as

$$\dot{\mathbf{q}}_l = -w_d \nabla_{\mathbf{q}_l} V_d - w_t \nabla_{\mathbf{q}_l} V_t - w_s \nabla_{\mathbf{q}_l} V_s, \tag{20a}$$
$$\dot{\mathbf{q}}_r = -w_d \nabla_{\mathbf{q}_r} V_d - w_t \nabla_{\mathbf{q}_r} V_t - w_s \nabla_{\mathbf{q}_r} V_s, \tag{20b}$$

where $\nabla_{\mathbf{q}_*} V_d$, $\nabla_{\mathbf{q}_*} V_t$ and $\nabla_{\mathbf{q}_*} V_s$ are the gradients of the objective functions V_d, V_t and V_s with respect to $\mathbf{q}_*$ for $* = l, r$. Then, the fact that $\dot{\mathbf{q}}_* \to \mathbf{0}$ together with (20) imply that the gradient of the objective function V tends to zero, i.e.,

$$\begin{bmatrix} \dot{\mathbf{q}}_l \\ \dot{\mathbf{q}}_r \end{bmatrix} = - \begin{bmatrix} \nabla_{\mathbf{q}_l} V \\ \nabla_{\mathbf{q}_r} V \end{bmatrix} = -\nabla V \to 0$$

as time goes to infinity, where $\nabla_{\mathbf{q}_*} V$ are the gradients of the objective function V with respect to $\mathbf{q}_*$. One can thus conclude that the objective function V converges to its minimum.

Note that the objective function V is not necessarily globally strictly convex with respect to $\mathbf{q}_l$ and $\mathbf{q}_r$. The steady-state configuration of the dual-arm system can possibly be a local minimum point of V. Nevertheless, $V = 0$ at the initial time is exactly its global minimum by its definition, which implies that the proposed controls (15) reduce V back to zero asymptotically once it diverges a bit from zero throughout the cooperative manipulation. Therefore, both the pose synchronization control (5) and the enhanced cooperation control (10) objectives are fulfilled.

IV. Discussion

This section discusses the state and input constraints of manipulators and visual servoing issues.

A. State and Input Constraints

Implementing the proposed control on a real-robot platform needs to cope with the system's state and input constraints. For the velocity-controlled manipulators $* = l, r$, their joint angles $\mathbf{q}_*$ and velocities $\dot{\mathbf{q}}_*$ are the states and inputs of the system, respectively. Typical state constraints of the velocity-control manipulators are imposed to avoid the mechanical limits of their joints, obstacles in the surrounding environment, and collisions between their bodies. Together, all these state constraints can be unified into an inequality of the manipulators' joint angles as follows

$$h(\mathbf{q}_l, \mathbf{q}_r) \leq 0.$$

The input constraints are imposed to accommodate the limits of the left $* = l$ and the right $* = r$ arms' joint velocities. Namely, the joint velocity control inputs $\dot{\mathbf{q}}_*$ cannot exceed their maximum $\dot{\mathbf{q}}_*^{\max}$ and the minimum $\dot{\mathbf{q}}_*^{\min}$ values as follows

$$\dot{\mathbf{q}}_*^{\min} \leq \dot{\mathbf{q}}_* \leq \dot{\mathbf{q}}_*^{\max}.$$

Under the state constraints, the joint velocity control inputs to the manipulators can be designed by solving the optimization problem (14) with the penalty function

$$\psi(\mathbf{q}_l, \mathbf{q}_r) = \exp\left[\frac{\bar{h} - h(\mathbf{q}_l, \mathbf{q}_r)}{h(\mathbf{q}_l, \mathbf{q}_r)} \right] - \frac{\bar{h} - h(\mathbf{q}_l, \mathbf{q}_r)}{h(\mathbf{q}_l, \mathbf{q}_r)} - 1$$

$$\dot{\mathbf{q}}_l = -w_d \mathbf{J}_l^{\mathsf{T}} \begin{bmatrix} \mathbf{x}_l - \mathbf{x}_{ld} \\ \mathrm{sk}\left(\mathbf{R}_l \mathbf{R}_{ld}^{\mathsf{T}}\right)^{\vee} \end{bmatrix} - w_t \mathbf{J}_l^{\mathsf{T}} \begin{bmatrix} \widetilde{\mathbf{x}}_l - \overline{\mathbf{R}}_l \overline{\mathbf{R}}_r^{\mathsf{T}} \widetilde{\mathbf{x}}_r \\ \mathrm{sk}\left(\mathbf{R}_l \widetilde{\mathbf{R}}_r^{\mathsf{T}} \overline{\mathbf{R}}_l^{\mathsf{T}}\right)^{\vee} \end{bmatrix} - w_s \mathbf{J}_l^{\mathsf{T}} \sum_{i=1}^{4} \left[\mathbf{L}_{li}^{l}{}^{\mathsf{T}} (\mathbf{s}_{li} - \bar{\mathbf{s}}_{li}) + \mathbf{L}_{ri}^{l}{}^{\mathsf{T}} (\mathbf{s}_{ri} - \bar{\mathbf{s}}_{ri}) \right] \tag{15a}$$

$$\dot{\mathbf{q}}_r = -w_d \mathbf{J}_r^{\mathsf{T}} \begin{bmatrix} \mathbf{x}_r - \mathbf{x}_{rd} \\ \mathrm{sk}\left(\mathbf{R}_r \mathbf{R}_{rd}^{\mathsf{T}}\right)^{\vee} \end{bmatrix} - w_t \mathbf{J}_r^{\mathsf{T}} \begin{bmatrix} \widetilde{\mathbf{x}}_r - \overline{\mathbf{R}}_r \overline{\mathbf{R}}_l^{\mathsf{T}} \widetilde{\mathbf{x}}_l \\ \mathrm{sk}\left(\mathbf{R}_r \widetilde{\mathbf{R}}_l^{\mathsf{T}} \overline{\mathbf{R}}_r^{\mathsf{T}}\right)^{\vee} \end{bmatrix} - w_s \mathbf{J}_r^{\mathsf{T}} \sum_{i=1}^{4} \left[\mathbf{L}_{li}^{r}{}^{\mathsf{T}} (\mathbf{s}_{li} - \bar{\mathbf{s}}_{li}) + \mathbf{L}_{ri}^{r}{}^{\mathsf{T}} (\mathbf{s}_{ri} - \bar{\mathbf{s}}_{ri}) \right] \tag{15b}$$

$$\dot{V}_d = -\frac{1}{2}\sum_{*=l,r}\mathrm{tr}\left(\mathbf{R}_{*d}^\mathsf{T}\dot{\mathbf{R}}_*\right) + \sum_{*=l,r}(\mathbf{x}_* - \mathbf{x}_{*d})^\mathsf{T}\dot{\mathbf{x}}_* = -\frac{1}{2}\sum_{*=l,r}\mathrm{tr}\left(\mathbf{R}_{*d}^\mathsf{T}\boldsymbol{\omega}_*^\times\mathbf{R}_*\right) + \sum_{*=l,r}(\mathbf{x}_* - \mathbf{x}_{*d})^\mathsf{T}\boldsymbol{v}_*$$

$$= -\frac{1}{2}\sum_{*=l,r}\mathrm{tr}\left(\boldsymbol{\omega}_*^\times\mathbf{R}_*\mathbf{R}_{*d}^\mathsf{T}\right) + \sum_{*=l,r}(\mathbf{x}_* - \mathbf{x}_{*d})^\mathsf{T}\boldsymbol{v}_* = \sum_{*=l,r}\boldsymbol{\omega}_*^\mathsf{T}\mathrm{sk}\left(\mathbf{R}_*\mathbf{R}_{*d}^\mathsf{T}\right)^\vee + (\mathbf{x}_* - \mathbf{x}_{*d})^\mathsf{T}\boldsymbol{v}_* \qquad (17)$$

$$= \sum_{*=l,r}\begin{bmatrix}\mathbf{x}_* - \mathbf{x}_{*d}\\\mathrm{sk}\left(\mathbf{R}_*\mathbf{R}_{*d}^\mathsf{T}\right)^\vee\end{bmatrix}^\mathsf{T}\mathbf{v}_* = \sum_{*=l,r}\left[\nabla_{\mathbf{q}_*}V_d\right]^\mathsf{T}\dot{\mathbf{q}}_* = \begin{bmatrix}\nabla_{\mathbf{q}_l}V_d\\\nabla_{\mathbf{q}_r}V_d\end{bmatrix}^\mathsf{T}\begin{bmatrix}\dot{\mathbf{q}}_l\\\dot{\mathbf{q}}_r\end{bmatrix}$$

$$\dot{V}_t = -\frac{1}{2}\mathrm{tr}\left(\dot{\widetilde{\mathbf{R}}}_l\widetilde{\mathbf{R}}_r^\mathsf{T} + \widetilde{\mathbf{R}}_l\dot{\widetilde{\mathbf{R}}}_r^\mathsf{T}\right) + \left(\overline{\mathbf{R}}_l^\mathsf{T}\widetilde{\mathbf{x}}_l - \overline{\mathbf{R}}_r^\mathsf{T}\widetilde{\mathbf{x}}_r\right)^\mathsf{T}\left(\overline{\mathbf{R}}_l^\mathsf{T}\dot{\widetilde{\mathbf{x}}}_l - \overline{\mathbf{R}}_r^\mathsf{T}\dot{\widetilde{\mathbf{x}}}_r\right)$$

$$= -\frac{1}{2}\mathrm{tr}\left(\overline{\mathbf{R}}_l^\mathsf{T}\dot{\mathbf{R}}_l\widetilde{\mathbf{R}}_r^\mathsf{T} + \widetilde{\mathbf{R}}_l\dot{\mathbf{R}}_r^\mathsf{T}\overline{\mathbf{R}}_r\right) + \left(\overline{\mathbf{R}}_l^\mathsf{T}\widetilde{\mathbf{x}}_l - \overline{\mathbf{R}}_r^\mathsf{T}\widetilde{\mathbf{x}}_r\right)^\mathsf{T}\left(\overline{\mathbf{R}}_l^\mathsf{T}\dot{\mathbf{x}}_l - \overline{\mathbf{R}}_r^\mathsf{T}\dot{\mathbf{x}}_r\right)$$

$$= -\frac{1}{2}\mathrm{tr}\left(\dot{\mathbf{R}}_l\widetilde{\mathbf{R}}_r^\mathsf{T}\overline{\mathbf{R}}_l^\mathsf{T}\right) - \frac{1}{2}\mathrm{tr}\left(\overline{\mathbf{R}}_r\widetilde{\mathbf{R}}_l\dot{\mathbf{R}}_r^\mathsf{T}\right) + \left(\overline{\mathbf{R}}_l^\mathsf{T}\widetilde{\mathbf{x}}_l - \overline{\mathbf{R}}_r^\mathsf{T}\widetilde{\mathbf{x}}_r\right)^\mathsf{T}\overline{\mathbf{R}}_l^\mathsf{T}\dot{\mathbf{x}}_l + \left(\overline{\mathbf{R}}_r^\mathsf{T}\widetilde{\mathbf{x}}_r - \overline{\mathbf{R}}_l^\mathsf{T}\widetilde{\mathbf{x}}_l\right)^\mathsf{T}\overline{\mathbf{R}}_r^\mathsf{T}\dot{\mathbf{x}}_r$$

$$= -\frac{1}{2}\mathrm{tr}\left(\boldsymbol{\omega}_l^\times\mathbf{R}_l\widetilde{\mathbf{R}}_r^\mathsf{T}\overline{\mathbf{R}}_l^\mathsf{T}\right) + \frac{1}{2}\mathrm{tr}\left(\overline{\mathbf{R}}_r\widetilde{\mathbf{R}}_l\mathbf{R}_r^\mathsf{T}\boldsymbol{\omega}_r^\times\right) + \left(\overline{\mathbf{R}}_l^\mathsf{T}\widetilde{\mathbf{x}}_l - \overline{\mathbf{R}}_r^\mathsf{T}\widetilde{\mathbf{x}}_r\right)^\mathsf{T}\overline{\mathbf{R}}_l^\mathsf{T}\boldsymbol{v}_l + \left(\overline{\mathbf{R}}_r^\mathsf{T}\widetilde{\mathbf{x}}_r - \overline{\mathbf{R}}_l^\mathsf{T}\widetilde{\mathbf{x}}_l\right)^\mathsf{T}\overline{\mathbf{R}}_r^\mathsf{T}\boldsymbol{v}_r \qquad (18)$$

$$= \boldsymbol{\omega}_l^\mathsf{T}\mathrm{sk}\left(\mathbf{R}_l\widetilde{\mathbf{R}}_r^\mathsf{T}\overline{\mathbf{R}}_l^\mathsf{T}\right)^\vee + \boldsymbol{\omega}_r^\mathsf{T}\mathrm{sk}\left(\mathbf{R}_r\widetilde{\mathbf{R}}_l^\mathsf{T}\overline{\mathbf{R}}_r^\mathsf{T}\right)^\vee + \left(\widetilde{\mathbf{x}}_l - \overline{\mathbf{R}}_l\overline{\mathbf{R}}_r^\mathsf{T}\widetilde{\mathbf{x}}_r\right)^\mathsf{T}\boldsymbol{v}_l + \left(\widetilde{\mathbf{x}}_r - \overline{\mathbf{R}}_r\overline{\mathbf{R}}_l^\mathsf{T}\widetilde{\mathbf{x}}_l\right)^\mathsf{T}\boldsymbol{v}_r$$

$$= \begin{bmatrix}\widetilde{\mathbf{x}}_l - \overline{\mathbf{R}}_l\overline{\mathbf{R}}_r^\mathsf{T}\widetilde{\mathbf{x}}_r\\\mathrm{sk}\left(\mathbf{R}_l\widetilde{\mathbf{R}}_r^\mathsf{T}\overline{\mathbf{R}}_l^\mathsf{T}\right)^\vee\end{bmatrix}^\mathsf{T}\mathbf{v}_l + \begin{bmatrix}\widetilde{\mathbf{x}}_r - \overline{\mathbf{R}}_r\overline{\mathbf{R}}_l^\mathsf{T}\widetilde{\mathbf{x}}_l\\\mathrm{sk}\left(\mathbf{R}_r\widetilde{\mathbf{R}}_l^\mathsf{T}\overline{\mathbf{R}}_r^\mathsf{T}\right)^\vee\end{bmatrix}^\mathsf{T}\mathbf{v}_r = \begin{bmatrix}\nabla_{\mathbf{q}_l}V_t\\\nabla_{\mathbf{q}_r}V_t\end{bmatrix}^\mathsf{T}\begin{bmatrix}\dot{\mathbf{q}}_l\\\dot{\mathbf{q}}_r\end{bmatrix}$$

$$\dot{V}_s = \sum_{i=1}^{4}\begin{bmatrix}\mathbf{s}_{li} - \bar{\mathbf{s}}_{li}\\\mathbf{s}_{ri} - \bar{\mathbf{s}}_{ri}\end{bmatrix}^\mathsf{T}\begin{bmatrix}\dot{\mathbf{s}}_{li}\\\dot{\mathbf{s}}_{ri}\end{bmatrix} = \sum_{i=1}^{4}\begin{bmatrix}\mathbf{s}_{li} - \bar{\mathbf{s}}_{li}\\\mathbf{s}_{ri} - \bar{\mathbf{s}}_{ri}\end{bmatrix}^\mathsf{T}\begin{bmatrix}\mathbf{L}_{li}^l & \mathbf{L}_{li}^r\\\mathbf{L}_{ri}^l & \mathbf{L}_{ri}^r\end{bmatrix}\begin{bmatrix}\mathbf{v}_l\\\mathbf{v}_r\end{bmatrix} = \sum_{i=1}^{4}\left(\begin{bmatrix}\mathbf{L}_{li}^l & \mathbf{L}_{li}^r\\\mathbf{L}_{ri}^l & \mathbf{L}_{ri}^r\end{bmatrix}^\mathsf{T}\begin{bmatrix}\mathbf{s}_{li} - \bar{\mathbf{s}}_{li}\\\mathbf{s}_{ri} - \bar{\mathbf{s}}_{ri}\end{bmatrix}\right)^\mathsf{T}\begin{bmatrix}\mathbf{v}_l\\\mathbf{v}_r\end{bmatrix}$$

$$= \sum_{i=1}^{4}\begin{bmatrix}\mathbf{L}_{li}^{l\,\mathsf{T}}(\mathbf{s}_{li} - \bar{\mathbf{s}}_{li}) + \mathbf{L}_{ri}^{l\,\mathsf{T}}(\mathbf{s}_{ri} - \bar{\mathbf{s}}_{ri})\\\mathbf{L}_{li}^{r\,\mathsf{T}}(\mathbf{s}_{li} - \bar{\mathbf{s}}_{li}) + \mathbf{L}_{ri}^{r\,\mathsf{T}}(\mathbf{s}_{ri} - \bar{\mathbf{s}}_{ri})\end{bmatrix}^\mathsf{T}\begin{bmatrix}\mathbf{v}_l\\\mathbf{v}_r\end{bmatrix} = \begin{bmatrix}\nabla_{\mathbf{q}_l}V_s\\\nabla_{\mathbf{q}_r}V_s\end{bmatrix}^\mathsf{T}\begin{bmatrix}\dot{\mathbf{q}}_l\\\dot{\mathbf{q}}_r\end{bmatrix} \qquad (19)$$

---

for $\overline{h} < h(\mathbf{q}_l, \mathbf{q}_r) < 0$ and $\psi(\mathbf{q}_l, \mathbf{q}_r) = 0$ for $h(\mathbf{q}_l, \mathbf{q}_r) \leq \overline{h}$ with $\overline{h} < 0$. Then, the input constraints are guaranteed by scaling the velocity inputs $\dot{\mathbf{q}}_*$ with factors

$$\alpha_* = \max_{i=1,\cdots,6}\left(\frac{\dot{q}_{*i}}{\dot{q}_{*i}^{\max}}, \frac{\dot{q}_{*i}}{\dot{q}_{*i}^{\min}}, 1\right),$$

where $\dot{q}_{*i}$, $\dot{q}_{*i}^{\max}$ and $\dot{q}_{*i}^{\min}$ are the i-th elements of $\dot{\mathbf{q}}_*$, $\dot{\mathbf{q}}_*^{\max}$ and $\dot{\mathbf{q}}_*^{\min}$, respectively. Together, the left $* = l$ and the right $* = r$ arms cooperatively solve the optimization problem (14) under the state and input constraints by their joint velocity inputs

$$\dot{\mathbf{q}}_* = -\frac{1}{\alpha_*}\left[\nabla_{\mathbf{q}_*}V + \nabla_{\mathbf{q}_*}\psi(\mathbf{q}_l, \mathbf{q}_r)\right],$$

where $\nabla_{\mathbf{q}_*}\psi(\mathbf{q}_l, \mathbf{q}_r)$ are the gradients of the penalty function $\psi(\mathbf{q}_l, \mathbf{q}_r)$ with respect to $\mathbf{q}_*$.

B. Visual Servoing Issues

The system performance is limited by the low data rate of vision sensors and the latency of feature extracting algorithms. Thus, extended Kalman filters have been integrated to predict and track moving targets [26], and sampled-data controls have been developed to stabilize the pick and place operations [27]. The lighting conditions are also unignorable for deploying vision-based controls in complex environments. Poor illumination conditions significantly degrade the performance of vision sensors. Thus, the mutual information [28] and the dense depth map [29] have robustified visual servoing against illumination variations at the cost of positioning accuracy. Event-based cameras are promising for overcoming the low data rate and the sensitivity to illumination of traditional vision sensors.

Large objects can possibly occlude markers mounted on the two manipulators. The visual servoing control then becomes invalid. In this situation, one can attach fiducial markers to the object and design visual servoing algorithms for the two manipulators with wrist-mounted cameras to precisely regulate their poses relative to the markers, thereby securely holding and manipulating the objects. However, the kinematic and dynamic models of the closed-loop system would become entirely different due to the possible relative motions between robots and objects. Future research will design new control algorithms to handle the manipulation of large objects.

PBVS and 2-1/2-D visual servoing [30] both possess larger domains of attraction than IBVS by adding some 3D information to the visual features. Yet, estimating the full or partial poses of markers/cameras requires more complicated image processing algorithms, which may induce extra latencies and degrade the system's responsiveness. Also, pose estimation algorithms are vulnerable to sensing noises. Even small errors in the image measurements can cause significant errors in the estimated pose, thereby undermining the system's positioning accuracy. Developing fast and robust pose estimation algorithms remains a future work.

V. Experimental Results

This section evaluates the proposed control via comparative experiments on a real-robot platform. The experimental setup includes a pair of Universal Robots UR3/UR3e manipulators, as shown in Figure 3. The task of the dual-arm system is to hold the target object, a 15 cm×15 cm×15 cm cubic orange box of weight 2.28 kg, without fixtures and transport it from the initial pose to the target pose while passing 3 waypoints. Two customized bowl-shaped grippers with silicone covers of shore 0A hardness and 25 mm thickness are mounted to the flanges of both arms to prevent them from over-squeezing the manipulated object. The interaction wrenches between the end effectors of the two arms and the box are measured by the built-in force/torque sensor of the UR3e arm. On the wrist of each arm, an Intel RealSense D435 camera captures the four corner points of the AprilTag marker from the 36h11 family mounted on the wrist of the other arm at about 30 Hz. The coordinates of the corner points on the image planes of the cameras serve as the image features for IBVS. The controls of the UR3/UR3e manipulators at 125 Hz/500 Hz are implemented in C++ programs that run on a single Ubuntu machine. The Real-Time Data Exchange (RTDE) synchronizes control programs with the UR robot hardware over a standard TCP/IP connection without breaking real-time properties.
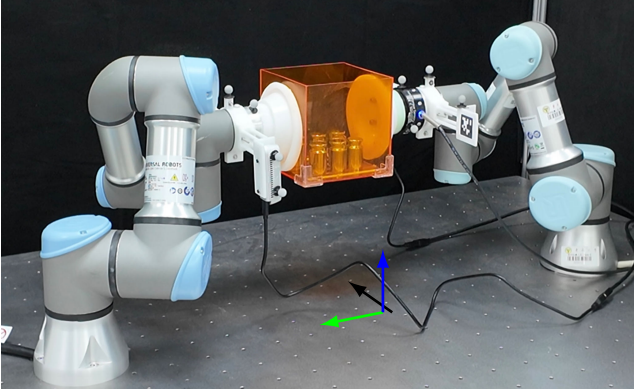

Figure 3. The experimental setup includes two Universal Robots UR3/UR3e manipulators. Both arms equip elastic tools on their flanges to transport a cubic box via frictional contacts. The camera mounted on the wrist of each arm captures the marker mounted on the wrist of the other arm.

In all experiments, the left and right manipulators locate their bases at $(0, -0.55, 0)$ m and $(0, 0.55, 0)$ m and start their joint coordinates with $(-30°, -140°, -85°, 45°, 75°, 0°)$. By the Monte Carlo algorithm, Table I lists the maximum and minimum coordinates that the end effectors of the two arms can reach. Accounting for the limited workspaces, the two arms transport the object to move along 3 paths. Each path has two waypoints B_o and C_o other than the initial/final pose A_o, of which the positions and orientations are listed in Table II.

Figure 4 plots the positions and the orientations of the two arms' end effectors when the object moves along path1, path2 and path3, respectively. The colors of the paths change with time as indicated in the color bar, and the coordinate frames at the waypoints mark the end-effector orientations of both arms. The left $* = l$ and the right $* = r$ arms start with their initial poses at the waypoints A_*, move towards the waypoints B_*

Table I
THE LEFT AND THE RIGHT ARMS' WORKSPACES IN METERS.

arms		x-coordinate	y-coordinate	z-coordinate
left	min	-0.4375	-0.0083	-0.3497
	max	$+0.4394$	$+0.5788$	$+0.6389$
right	min	-0.4394	$+0.1289$	-0.3497
	max	$+0.4375$	$+0.7408$	$+0.6389$

and C_*, and back to the waypoints A_*. Here, the poses of the waypoints are exactly the desired poses $(\mathbf{R}_{*d}, \mathbf{x}_{*d}) \in SE(3)$ in the controls (15) of the two arms.

Two groups of experiments examine the performances of the dual-arm system under the proposed control by picking different sets of gains w_d, w_t and w_s. With $w_d = 1.0$ and $w_d = 1.5$, the first and the second groups transport the object at a relatively lower and higher speed, respectively. In each group, the dual-arm system performs four experimental trials via setting $w_s = 0.0$, $w_s = 0.3$, $w_s = 0.5$ and $w_s = 1.0$ in ascending order while letting $w_t = 1 - w_s$. Basically, w_s weighs the impact of the IBVS controls on the coordination of the two arms. When $w_s = 0.0$ and $w_s = 1.0$, the proposed control becomes a purely pose synchronization control [20] and a purely IBVS control, respectively.

At a relatively lower speed, the first group of experiments all successfully transport the object along the planned three paths. Nevertheless, the performances of the dual-arm manipulation system change with the control gains w_s and $w_t = 1 - w_s$. Figure 5 depicts the errors of the image features from their desired values when the two arms transport the object to move along path1. The gain w_s grows in the order of 0.0, 0.3, 0.5 and 1.0. For most of the time, the purely pose synchronization control ($w_s = 0.0$) exhibits the greatest error of image features. Its maximum error is 2 times larger than other trials. As w_s raises to 0.3 and 0.5 sequentially, the errors of image features reduce accordingly. When the control becomes purely IBVS ($w_s = 1.0$), the error grows again, staying between the errors of the trials $w_s = 0.3$ and $w_s = 0.5$. Overall, setting equal weights $w_s = w_t = 0.5$ for the pose synchronization and IBVS controls obtains the smallest error. The error profiles display a similar trend in all trials, which suggests that they are related to the manipulators' postures as well.

Figure 6 presents the deviations of the interaction forces and torques between the end effector of the UR3e arm and the object. Corresponding to Figure 5, it delivers that the interaction forces and torques vary with the image feature errors throughout the object transportation. Roughly speaking, the interaction forces and torques increase/decrease with the errors of image features. From the 7 s to the 27 s, however, the interaction forces deviate less when the pose synchronization control functions alone than together with the IBVS control using $w_s = 0.3$. One can also notice that, at the 30 s, the torque deviations of the purely pose synchronization control and the purely IBVS control both reach their peaks. By the contrast, weighing the pose synchronization control and the IBVS control equally with $w_s = w_t = 0.5$ maintains both the

Table II
THE POSITION COORDINATES AND THE ROLL-PITCH-YAW ANGLES OF THE OBJECT'S WAYPOINTS ON 3 PATHS.

paths		path1	path2	path3
waypoint B_o	position [m]	$(0.13, -0.03, -0.02)$	$(0, -0.1, -0.05)$	$(-0.1, -0.05, 0)$
	orientation [rad]	$(\pi/10, -\pi/18, 0)$	$(0, -\pi/18, 0)$	$(-\pi/12, -\pi/12, \pi/12)$
waypoint C_o	position [m]	$(-0.03, 0.01, 0)$	$(-0.1, 0.05, 0.05)$	$(0.1, -0.05, 0)$
	orientation [rad]	$(-\pi/12, \pi/12, 0)$	$(\pi/18, 0, 0)$	$(-\pi/12, \pi/12, -\pi/10)$

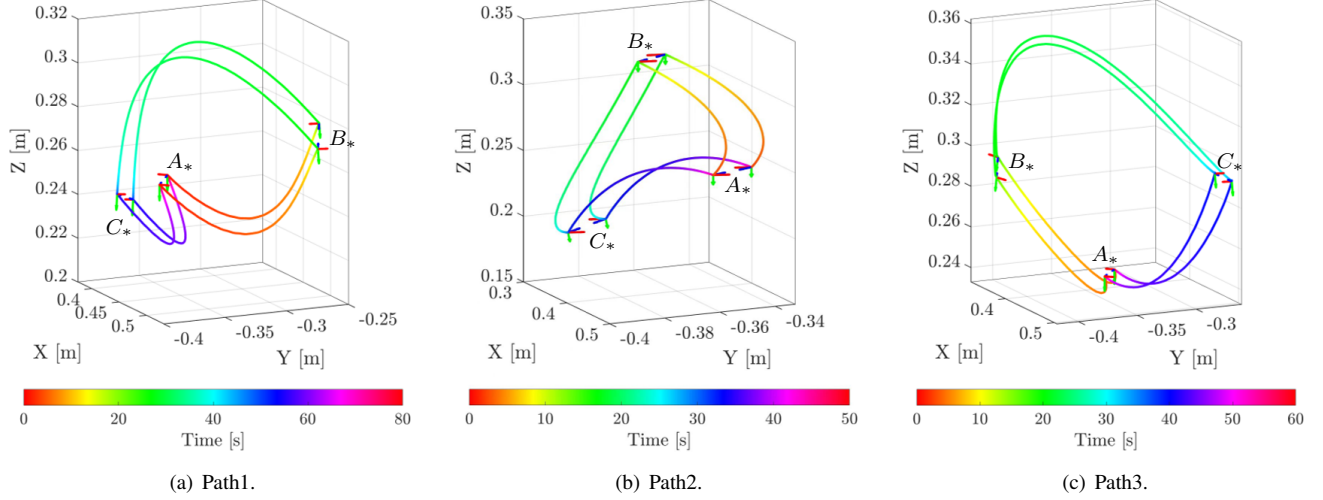
(a) Path1.
(b) Path2.
(c) Path3.

Figure 4. The Cartesian paths and the waypoints' orientations of the two arms' end effectors when the object moves along path1, path2 and path3, respectively. The final poses of both arms coincide with their initial poses.
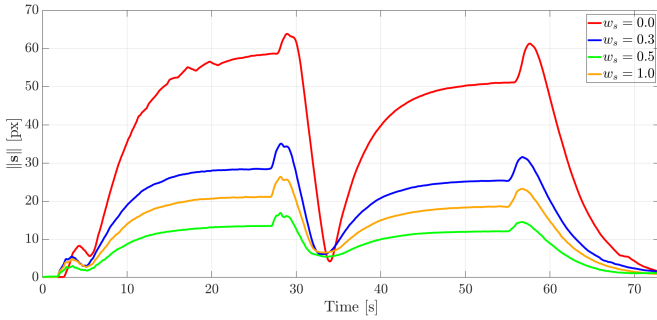

Figure 5. The errors of the image features $\mathbf{s} = (\mathbf{s}_l^\mathsf{T}, \mathbf{s}_r^\mathsf{T})^\mathsf{T}$ from their desired (initial) values when transporting the object along path1 during the first group of experiments with $w_d = 1.0$ and $w_t = 1 - w_s$.
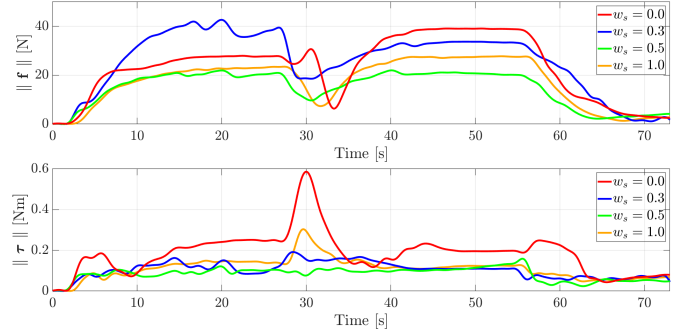

Figure 6. The deviations of the interaction forces (the top subplot) and torques (the bottom subplot) between the end effector of the UR3e manipulator and the object when transporting the object along path1 during the first group of experiments with $w_d = 1.0$ and $w_t = 1 - w_s$.

force and torque deviations flattest and smallest among all the four trials.

At a relatively higher speed, the second group of experiments encounter some failures. In particular, both the purely pose synchronization control ($w_s = 0.0$) and the purely IBVS control ($w_s = 1.0$) cause the object to fall down or slide when transporting it along path1 and path3. In contrast, weighing the pose synchronization control and the IBVS control equally with $w_s = w_t = 0.5$ succeeds in transporting the object along all three paths. Besides the accompanying videos, Figure 7 compares the image trajectories of the markers' corners when the two arms transport the object along path3 with $w_s = 0.0$,

$w_s = 1.0$ and $w_s = 0.5$, respectively. One can find that the markers' corners exhibit the greatest deviations with $w_s = 0.0$, the moderate deviations with $w_s = 1.0$, and the least deviations with $w_s = 0.5$. Correspondingly, the object falls down, slides, and holds on in the respective experimental trials.

VI. CONCLUSION

This paper has developed an IBVS approach for enhancing the cooperation of dual-arm manipulation. By coordinating the two arms with their kinematics and joint angle measurements, the conventional synchronization control leads to significant end-effector pose errors of the dual-arm manipulation system
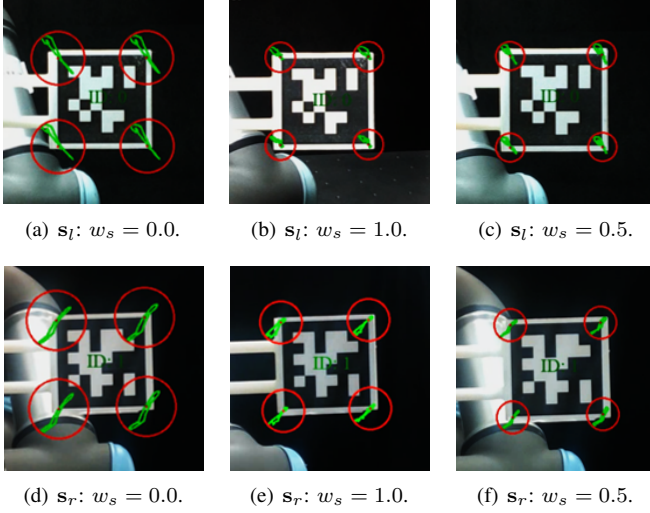
(a) $\mathbf{s}_l$: $w_s = 0.0$.   (b) $\mathbf{s}_l$: $w_s = 1.0$.   (c) $\mathbf{s}_l$: $w_s = 0.5$.

(d) $\mathbf{s}_r$: $w_s = 0.0$.   (e) $\mathbf{s}_r$: $w_s = 1.0$.   (f) $\mathbf{s}_r$: $w_s = 0.5$.

Figure 7. The trajectories $\mathbf{s}_*$ of the markers' corners on the images of the left $* = l$ and the right $* = r$ cameras when transporting the object along path3 in the second group of experiments with $w_d = 1.5$ and $w_t = 1 - w_s$. The green curves are the image trajectories. The centers and radii of red circles indicate the original locations and the maximum deviations of the markers' corners on the images, respectively.

when systematic errors exist. In contrast, the proposed control can mitigate the pose synchronization errors because the IBVS enhancement is robust to the kinematic errors of manipulators. Technically, the paper has modelled the kinematics of a dual-arm manipulation system, has tailored the interaction matrix to the dual-arm system, has designed the joint velocity control inputs with rigorous stability proof, and has discussed the state and input constraints of manipulators and visual servoing issues. Future studies will control the interaction forces between manipulators and objects and will coordinate them to assemble parts and components on floating bases.

REFERENCES

[1] A. Billard and D. Kragic, "Trends and challenges in robot manipulation," *Science*, vol. 364, no. 6446, p. eaat8414, 2019.

[2] M. Abbas, J. Narayan, and S. K. Dwivedy, "A systematic review on cooperative dual-arm manipulators: modeling, planning, control, and vision strategies," *International Journal of Intelligent Robotics and Applications*, vol. 7, pp. 683–707, 2023.

[3] D. Rakita, B. Mutlu, M. Gleicher, and L. M. Hiatt, "Shared control–based bimanual robot manipulation," *Science Robotics*, vol. 4, no. 30, p. eaaw0955, 2019.

[4] F. Krebs and T. Asfour, "A bimanual manipulation taxonomy," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 11031–11038, 2022.

[5] F. Caccavale and M. Uchiyama, "Cooperative manipulation," in *Springer Handbook of Robotics*, B. Siciliano and O. Khatib, Eds. Berlin, Germany: Springer, 2016, pp. 989–1006.

[6] P. Chiacchio, S. Chiaverini, and B. Siciliano, "Direct and inverse kinematics for coordinated motion tasks of a two-manipulator system," *Journal of Dynamic Systems, Measurement, and Control*, vol. 118, no. 4, pp. 691–697, 12 1996.

[7] P. B. g. Dohmann and S. Hirche, "Distributed control for cooperative manipulation with event-triggered communication," *IEEE Transactions on Robotics*, vol. 36, no. 4, pp. 1038–1052, 2020.

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

[9] I. D. Walker, R. A. Freeman, and S. I. Marcus, "Analysis of motion and internal loading of objects grasped by multiple cooperating manipulators," *The International Journal of Robotics Research*, vol. 10, no. 4, pp. 396–409, 1991.

[10] Y. He, M. Wu, and S. Liu, "A cooperative optimization strategy for distributed multi-robot manipulation with obstacle avoidance and internal performance maximization," *Mechatronics*, vol. 76, p. 102560, 2021.

[11] N. Dehio, J. Smith, D. L. Wigand, P. Mohammadi, M. Mistry, and J. J. Steil, "Enabling impedance-based physical human–multi–robot collaboration: Experiments with four torque-controlled manipulators," *The International Journal of Robotics Research*, vol. 41, no. 1, pp. 68–84, 2022.

[12] Y. Zhang, X. Zhao, B. Tao, and H. Ding, "Multi-objective synchronization control for dual-robot interactive cooperation using nonlinear model predictive policy," *IEEE Transactions on Industrial Electronics*, vol. 70, no. 1, pp. 582–593, 2023.

[13] D. Ortenzi, R. Muthusamy, A. Freddi, A. Monteriù, and V. Kyrki, "Dual-arm cooperative manipulation under joint limit constraints," *Robotics and Autonomous Systems*, vol. 99, pp. 110–120, 2018.

[14] S. Tarbouriech, B. Navarro, P. Fraisse, A. Crosnier, A. Cherubini, and D. Sallé, "Dual-arm relative tasks performance using sparse kinematic control," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 6003–6009.

[15] M. Costanzo, G. De Maria, and C. Natale, "Tactile feedback enabling in-hand pivoting and internal force control for dual-arm cooperative object carrying," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 11466–11473, 2022.

[16] Y. Gombo, A. Tiwari, M. Safwat, H. Chang, and S. Devasia, "Delayed self-reinforcement to reduce deformation during decentralized flexible-object transport," *IEEE Transactions on Robotics*, vol. 40, pp. 999–1018, 2024.

[17] N. Vahrenkamp, C. Böge, K. Welke, T. Asfour, J. Walter, and R. Dillmann, "Visual servoing for dual arm motions on a humanoid robot," in *2009 9th IEEE-RAS International Conference on Humanoid Robots*, 2009, pp. 208–214.

[18] J. Qu, F. Zhang, Y. Fu, and S. Guo, "Multi-cameras visual servoing for dual-arm coordinated manipulation," *Robotica*, vol. 35, no. 11, p. 2218–2237, 2017.

[19] R. Mebarki, V. Lippiello, and B. Siciliano, "Toward image-based visual servoing for cooperative aerial manipulation," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 6074–6080.

[20] T. Hatanaka, Y. Igarashi, M. Fujita, and M. W. Spong, "Passivity-based pose synchronization in three dimensions," *IEEE Transactions on Automatic Control*, vol. 57, no. 2, pp. 360–375, 2012.

[21] Y. Yang, D. Constantinescu, and Y. Shi, "Passive multiuser teleoperation of a multirobot system with connectivity-preserving containment," *IEEE Transactions on Robotics*, vol. 38, no. 1, pp. 209–228, 2022.

[22] Y. Yang, A. Song, L. Zhu, B. Xu, G. Song, and Y. Shi, "Passivity-based control of distributed teleoperation with velocity/force manipulability optimization," *IEEE Transactions on Robotics*, vol. 41, pp. 647–665, 2025.

[23] F. Aghili, "Adaptive control of manipulators forming closed kinematic chain with inaccurate kinematic model," *IEEE/ASME Transactions on Mechatronics*, vol. 18, no. 5, pp. 1544–1554, 2013.

[24] F. Chaumette and S. Hutchinson, "Visual servo control. i. basic approaches," *IEEE Robotics & Automation Magazine*, vol. 13, no. 4, pp. 82–90, 2006.

[25] ——, "Visual servo control. ii. advanced approaches [tutorial]," *IEEE Robotics & Automation Magazine*, vol. 14, no. 1, pp. 109–118, 2007.

[26] A. A. Oliva, E. Aertbeliën, J. De Schutter, P. Robuffo Giordano, and F. Chaumette, "Towards dynamic visual servoing for interaction control and moving targets," in *2022 International Conference on Robotics and Automation (ICRA)*, 2022, pp. 150–156.

[27] M. Costanzo, G. De Maria, C. Natale, and A. Russo, "Modeling and control of sampled-data image-based visual servoing with three-dimensional features," *IEEE Transactions on Control Systems Technology*, vol. 32, no. 1, pp. 31–46, 2024.

[28] A. Dame and E. Marchand, "Mutual information-based visual servoing," *IEEE Transactions on Robotics*, vol. 27, no. 5, pp. 958–969, 2011.

[29] C. Teulière and E. Marchand, "A dense and direct approach to visual servoing using depth maps," *IEEE Transactions on Robotics*, vol. 30, no. 5, pp. 1242–1249, 2014.

[30] E. Malis, F. Chaumette, and S. Boudet, "2-1/2-d visual servoing," *IEEE Transactions on Robotics and Automation*, vol. 15, no. 2, pp. 238–250, 1999.