Journal of Intelligent & Robotic Systems (2022) 105: 72
https://doi.org/10.1007/s10846-022-01674-5

REGULAR PAPER

A Visual Predictive Control Framework for Robust and Constrained Multi-Agent Formation Control

Mostafa M. H. Fallah[1] · Farrokh Janabi-Sharifi[1] · Sina Sajjadi[2] · Mehran Mehrandezh[2]

Received: 30 September 2021 / Accepted: 3 June 2022 / Published online: 18 July 2022

Abstract
In this study, the problem of leader-follower position-based formation control is considered. Each agent in the multi-agent network is equipped with a perspective camera (assumed to be pinhole-type) and is servoed visually. A depth-based visual predictive controller is proposed. The framework optimizes the planned trajectory with a prediction horizon, while taking image- and physical-space constraints into account. Furthermore, the presented control scheme provides robustness against the camera occlusion, modeling errors and uncertainties. The performance of the proposed tracking algorithm is validated via numerous simulations.

Keywords Model predictive control · Visual servoing · Robotics · Discrete-time multi-agent systems

1 Introduction

In recent years, plenty of studies on multi-agent formation control has been conducted [1, 2], due to their numerous potential applications including those in the fields of search and exploration [3], cooperative surveillance [4, 5], mapping [6], etc. The problem of formation control of robots can be defined as coordinated control of a group of robots to maintain a desired spatial pattern, and as a group, to follow predefined trajectories. Typically, in formation control, finding a centralized controller would be a challenge, since a global information is not available to each robot in the team (i.e., limited cross-talk bandwidth among agents). Therefore, each robot relies on the local information, acquired by its own sensors or received from its neighbors, to generate the corresponding control signal. Based on the agents' interaction topology, and specifically, according to their sensing capabilities, various formation control problems have been introduced in the literature which could be classified into position-, displacement-, and distance-based formation control [1]. In this study, we focused on the position-based formation control (PBFC) approach, which integrates the advantage of achieving formation with reduced interactions needed among agents. In this approach, each agent senses its own position with respect to a predefined global coordinate system, and actively controls it to achieve and to maintain the final reference formation.

Among the existing approaches in this field, visual servoing (VS) methods have become popular due to their advantages [7–10]. Similar to the classic VS, the visual servo formation control approaches can be divided into two main categories: image-based (IBVS) [11–13], and position-based visual serving (PBVS) [14–24]. In PBVS methods, the predefined target with known geometry is utilized to estimate the relative pose of the leader and the followers, while in IBVS, unlike the previous one, followers achieve the tracking of the leader's pose by the direct control of the image errors. Due to the fact that the features are only being used in the pose estimation and not the control algorithm, the PBVS method may lead to violating

Category(7),(3)

✉ Farrokh Janabi-Sharifi
fsharifi@ryerson.ca

Mostafa M. H. Fallah
mostafa.mohammad@ryerson.ca

Sina Sajjadi
sinasajjadi@uregina.com

Mehran Mehrandezh
Mehran.Mehrandezh@uregina.ca

[1] Mechanical and Industrial Engineering Department, Ryerson University, 350 Victoria St, Toronto M5B 2K3, ON, Canada

[2] Faculty of Engineering and Applied Science, University of Regina, 3737 Wascana Pkwy, Regina S4S 0A2, SK, Canada

image constraints. Also, PBVS might demonstrate high sensitivity to the camera calibration errors. On the other hand, an IBVS may violateds physical constraints. In addition, it may lead to problems such as: reaching image singularities, and the camera-retreat problem. Generally, the problems in VS-based formation control would face challnges due to sensing, actuation, and control constraints. As far as imaging, cameras [16–19] or pinhole [11–15, 20–24] have limited field of view. On the other hand, actuation constraints are dictated by the structure of the utilized robots. Examples are: mobile robots [11–14, 16–24], unmanned aerial vehicles [15], etc. Finally, control limitations originate from the structure of the designed control system. Image Jacobian singularity and camera retreat/advance problem in the IBVS [11–13], and dependency on 3-D reconstruction and sensitivity to precise pose estimation and external camera calibration in PBVS [14–24] are examples of control constraints [25]. In the more recent papers, the problem of IBVS-based formation control under FOV constraints has been addressed [26, 27]. Moreover, the researchers have studied the problem under FOV constraints and uncertain system parameters and feature depth resulting in the development of adaptive and robust formation control schemes [28, 29]. However, the previous works on vision-based formation control fall short in simultaneously handling of aforementioned constraints and optimizing the control trajectory.

The contributions of the paper are as follows. (i) A novel depth-based visual predictive control (DVPC) is formulated to provide a hybrid control structure enabling control and constraint handling in both Cartesian and image spaces, (ii) A new VS-based framework is proposed for position-based formation control using DVPC in which a unique framework for information-exchange among agents is enabled to overcome potential occlusion of agents' camera(s). To the best of our knowledge, this is the first leader-follower formation control developed based on a VPC scheme. Incorporating control over the feature depth in the problem formulation would facilitate our algorithm to optimize the formation trajectories as well. In particular, it is shown via simulation that this will resolve the camera retreat problem [30–34].

In this approach, different agents' sense and actively control their own poses with respect to their corresponding coordinate systems to achieve the desired formation. In the proposed method, interactions among the agents are not necessary. However, it can be added to the tracking algorithm for the purposes of enhancing control performance [1], specifically in the case of issues such as unexpected occlusion (Fig. 1) and uncertain sensory data fairly easily. The proposed DVPC has been embedded in the well-known Internal Model Control (IMC) structure, where the model of the system is utilized to not only improve on the reference signals to track and the system robustness [35], but also

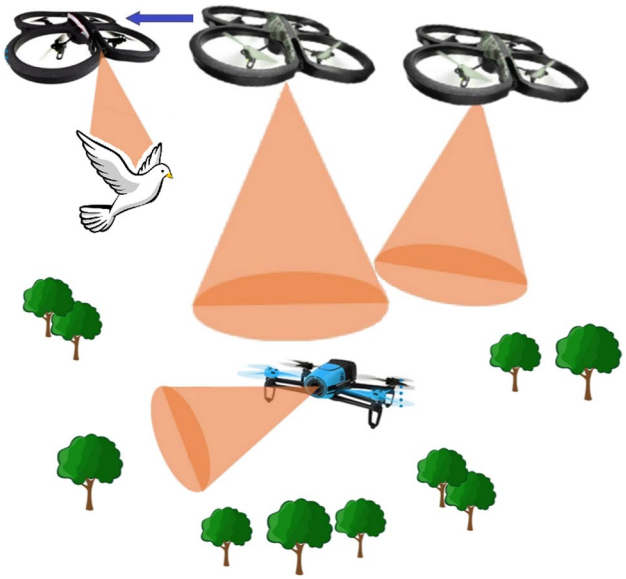


Fig. 1 Temporarily occluded follower in schematic multi UAV system

to take the role of an estimator for conveying the received measurement signals from the neighboring agents.

This paper is organized as follows. In Section 2, the details of the DVPC scheme is presented. The interaction method among agents is presented in Section 3. Section 4 is dedicated to the formal stability analysis of the proposed control scheme, while Section 5 presents some simulation results to demonstrate the performance of the developed control method under various conditions, and to show the advantages of this method compared to other control schemes. Finally, conclusions are given in the last section.

2 Control Structure

In this paper, each agent is equipped with a vision sensor to measure its pose with respect to the corresponding reference coordinate system and is controlled by a similar, but separate, depth-based visual predictive controller where the controllers can communicate in case of occlusion for sake of algorithm robustness. The control block diagram of the proposed DVPC scheme is presented in Fig. 2. As the block diagram shows, in the DVPC method, unlike the classic image-based visual predictive control, reference signal is comprised of two distinctive values, $\mathbf{s}^*(k) = [u^*(k), v^*(k), 1]^\top$ which is the homogeneous final coordinate vector of the visual feature(s), and $Z^*(k)$ or the depth of the corresponding feature(s) expressed in the reference camera coordinate frame.

In order to further increase the robustness of the controller in the presence of uncertainties, DVPC is encapsulated in well-known internal model control (IMC) structure. As depicted in Fig. 2, $\epsilon(k)$ represents the disturbances and

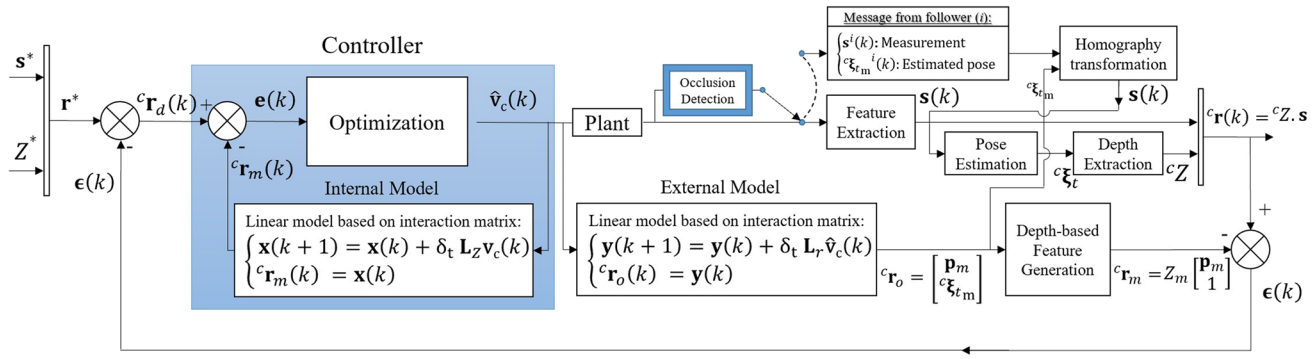


Fig. 2 Block diagram of the depth-based visual predictive control

modeling errors between current features and prediction. In fact, according to the IMC architecture, tracking the reference signal, $\mathbf{r}^*(k)$, by the system output $\mathbf{r}(k)$ is equal to the tracking of the desired trajectory, $\mathbf{r}_d(k)$, by the model output, ${}^c\mathbf{r}_m(k)$.

Similar to any other predictive controller, DVPC employs a model of the system to anticipate the output's behavior over the prediction horizon. In order to alleviate the issues due to the nonlinearity of the system, and to decrease the processing time (specifically when the prediction horizon is large), a linear model based on the depth-based Jacobian matrix, $\mathbf{L}_Z$, is employed. The corresponding Jacobian matrix is defined as,

$$\mathbf{L}_Z = [-\mathbf{M} \quad \mathbf{M}[\mathbf{M}^{-1}\mathbf{s}(k)]_\times], \tag{1}$$

where, $\mathbf{M}$, and $[\mathbf{M}^{-1}\mathbf{s}_i]_\times$ are the camera intrinsic matrix, and the corresponding skew-symmetric matrix of vector $\mathbf{M}^{-1}\mathbf{s}(k)$ [36]. The dynamic evolution of the hybrid depth-based feature is expressed by,

$${}^c\dot{\mathbf{r}}_m(k) = \mathbf{L}_Z\mathbf{v}_c(k), \tag{2}$$

and the linear state-space model of the system can be represented as,

$$\begin{cases} \mathbf{x}(k+1) = \mathbf{x}(k) + \delta_t\mathbf{L}_Z\mathbf{v}_c(k), \\ {}^c\mathbf{r}_m(k) = \mathbf{x}(k), \end{cases} \tag{3}$$

where, δ_t is the discrete time step and $\mathbf{v}_c(k)$, $\mathbf{x}(k)$, and ${}^c\mathbf{r}_m(k)$ are system input, states of the plant, and system output, respectively.

As it was mentioned, the algorithm in DVPC minimizes the difference between desired hybrid trajectory and predicted model output. Therefore, it can be formulated as a constrained optimization problem with the following objective function and constraints,

$$\min_{\hat{\mathbf{v}}_c(k)\in\mathbb{C}\subset\mathrm{R}^{6\times N_p}} \Gamma\big(\mathrm{v}_c(k)\big) = \sum_{j=k}^{k+N_p-1} \|\mathrm{r}_d(j) - {}^c\mathrm{r}_m(j)\|^{\mathrm{T}}\Omega\|\mathrm{r}_d(j) - {}^c\mathrm{r}_m(j)\| \tag{4}$$

$$\text{s.t.} \quad \mathbf{x}(k+1) = \mathbf{x}(k) + \delta_t\mathbf{L}_Z\mathbf{v}_c(k) \tag{5}$$

$$Z_{\min} \le Z_m(k) \le Z_{\max} \tag{6}$$

$$p_{\min} \le p(k) \le p_{\max} \tag{7}$$

$$\mathbf{v}_{\min} \le \mathbf{v}_c(k) \le \mathbf{v}_{\max} \tag{8}$$

$$\mathbf{r}_d(N_p+1) - {}^c\mathbf{r}_m(N_p+1) = \mathbf{0} \tag{9}$$

In the above formulas, N_p is the prediction horizon, $\mathbb{C}$ is the domain limited by all of the available constraints, and $\hat{\mathbf{v}}_c(k)$ is the optimal control signal, representing the velocity of camera for each individual agent or the leader. In addition, Ω is the weighting matrix which is considered to be the identity matrix. It is also necessary to mention that, due to saving the calculation time, the control horizon or N_c is equal one, and there is no change in the control signal inside the cost function.

As the last step, a quadratic program (QP) is employed to minimize the predefined cost function of sequentially linearised system and to handle the constraints. In this work, there are four different constraints to be handled during the optimization procedure which can be expressed as follows:

1. Image constraints such as camera's limited field of view. To avoid issues caused by these limitations, the visual features have to stay in the predefined area, or $p_{\min} \le p(k) \le p_{\max}$, where, $p_{\min} = [u_{\min}, v_{\min}]^\top$ and $p_{\max} = [u_{\max}, v_{\max}]^\top$ are respectively, the coordinates of the upper left and lower right corners of the camera's field of view.
2. Spatial constraints such as camera's limited depth of field. In this case, the camera has to avoid exceeding the predefined distance from the target points or, $Z_{\min} \le Z_m(k) \le Z_{\max}$, where, $Z_{\min}$ and $Z_{\max}$ are the minimum and the maximum acceptable depth of target points in the camera frame.
3. Control-related constraints such as actuators' saturation boundaries or, $\mathbf{v}_{\min} \le \mathbf{v}_c(k) \le \mathbf{v}_{\max}$, where, $\mathbf{v}_{\min}$ and $\mathbf{v}_{\max}$

are respectively, the minimum and maximum acceptable control signals.

4. Finally, the terminal equality constraint, (9), is set to guarantee the closed-loop stability of the control system.

3 Multi-Agent Position-Based Formation Control

In a typical position-based formation control (PBFC) problem, each agent needs to sense its absolute pose with respect to the preassigned global coordinate system [1], and find its own final reference pose by following the corresponding visual features and by approaching to the reference configuration. The desired formation will be accessible when all of the available agents reach to their own absolute predefined poses. It is important to mention that in PBFC, the dynamic interaction among agents inside a neighborhood is not usually required, and it is utilized to enhance the control performance and overcome some of the unexpected issues [1]. One of these problems is temporary and unexpected occlusion in which the data-sharing among occluded agent and its neighbors is beneficial. In such case, the necessary visual data can be transferred by homography transformation.

3.1 Homography transformation

homography is a method for the mapping of image coordinates between two planar projections of an image. Considering two cameras, α and β, looking at the same point $\mathbf{P}$, it is possible to pass from projection ${}^{\beta}\mathbf{s}(k) = [{}^{\beta}\mathbf{p}(k)^{\top},1]^{\top}$ of $\mathbf{P}$ in camera β to the projection ${}^{\alpha}\mathbf{s}(k) = [{}^{\alpha}\mathbf{p}(k)^{\top},1]^{\top}$ of $\mathbf{P}$ in camera α by the following formula

$$ {}^{\alpha}\mathbf{s}(k) = (\tfrac{{}^{\beta}Z(k)}{{}^{\alpha}Z(k)})\mathbf{M}_{\alpha}\mathbf{H}_{\alpha\beta}(k)\mathbf{M}_{\beta}^{-1}\mathbf{s}(k), \tag{10} $$

where, $\mathbf{M}_{\alpha}$ and $\mathbf{M}_{\beta}$ are the cameras' intrinsic parameter matrices, and the homography matrix $\mathbf{H}_{\alpha\beta}$ is given by

$$ \mathbf{H}_{\alpha\beta}(k) = {}^{\alpha}\mathbf{R}_{\beta}(k) - \frac{{}^{\alpha}\mathbf{t}_{\beta}(k)\mathbf{n}^{\top}}{d_{\beta}}. \tag{11} $$

In Eq. 11, ${}^{\alpha}\mathbf{R}_{\beta}(k)$, and ${}^{\alpha}\mathbf{t}_{\beta}(k)$ are respectively, the rotation matrix and translational vector from camera α to camera β, as depicted in Fig. 3. In addition, $\mathbf{n}$, and d_{β} are the normal vector of the plane and the distance from camera β to the plane, respectively.

Occlusion may occur either for the leader or the followers. Consequently, there is no measurement for the occluded camera and it is necessary to provide it with enough information to generate the proper control signal. This information includes the coordinates of the occluded visual features and the relative pose of the sender with respect to a common coordinate frame. In this step, each case is explained separately:

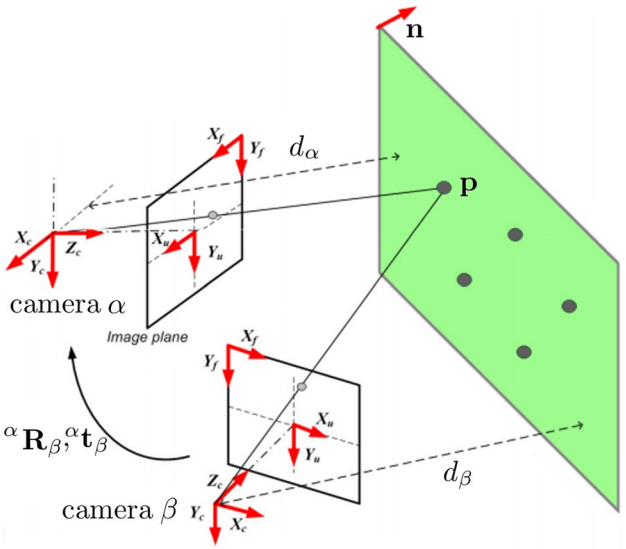


Fig. 3 homography transformation between two pinhole cameras

- **Occlusion of the follower's camera**: At each time instance (k), each follower inside the network receives packages of data sent by neighboring followers. Each package includes the measured coordinates of the target points attached to the leader at time instance (k) (or $\mathbf{s}(k)$) and the estimated pose of the sender with respect to the aforementioned target frame (or ${}^{c}\boldsymbol{\xi}_{tm}(k)$). It is necessary to mention that the external estimator model shown in Fig. 2 has two distinct responsibilities. As the estimator in the internal model control framework, it provides the system with estimated coordinate of visual feature (or $\mathbf{s}_{m}(k)$) and the relative estimated depth of the corresponding target point (or $Z_{m}(k)$). In addition, by estimating the relative pose with respect to the common target frame helps the occluded agent to translate the received data from neighboring agents to be utilized instead of the interrupted measurement. This linear model is formalized as follows [37]:

$$ \begin{cases} \mathbf{y}(k+1) = \mathbf{y}(k) + \delta_{t}\mathbf{L}_{r}\hat{\mathbf{v}}_{c}(k), \\ {}^{c}\mathbf{r}_{o}(k) = \mathbf{y}(k). \end{cases} \tag{12} $$

The external Jacobian matrix, $\mathbf{L}_{r}$ is defined as

$$ \mathbf{L}_{r}(k) = [\mathbf{L}_{p_1}(k)^{\top}, \ldots, \mathbf{L}_{p_g}(k)^{\top}, \mathbf{L}_{\xi}(k)^{\top}]^{\top}, \tag{13} $$

where, $\mathbf{L}_{p_i} \in \mathbb{R}^{2\times 6}$ and $\mathbf{L}_{\xi} \in \mathbb{R}^{6\times 6}$ are the image-based and position-based interaction matrices as proposed in [25]. By using the aforementioned interaction matrix, the

output of the external estimator model can be expressed as,

$$^{c}\mathbf{r}_{o}(k) = [\mathbf{p_m}(k), ^{\mathbf{c}}\boldsymbol{\xi}_{\mathbf{t_m(k)}}]$$

which is the concatenated coordinates of the estimated visual features, $\mathbf{p}_m(k) = (u_m(k), v_m(k))$, and the relative estimated target pose expressed in the camera frame, $^{c}\boldsymbol{\xi}_{tm}(k)$.

- **Occlusion of the leader's camera**: In this case, upon the leader's request, a package of data that includes two different items will be sent from its followers. The first item is the measured coordinates of the target points followed by the leader which may be seen by the followers, too. None of the followers use this set of target points to generate the control signal. The second item is the measured relative pose of the target attached to the leader with respect to the sender's camera or $^{c}\boldsymbol{\xi}_{t}(k)$. This relative pose is used by the leader to translate the corresponding measurement.

4 Stability Analysis

All of the agents in the proposed PBFC scheme utilize the DVPC as a successively linearized model predictive control method. The stability investigation of the proposed DVPC formation control system is similar to that of linear MPC. The stability of general model predictive controllers has been rigorously studied in [38], which yields the satisfaction of four axioms. As a matter of brevity, only the satisfaction of the above-mentioned axioms was discussed.

In the utilized predictive control approach, the state origin is shifted from $\mathbf{0}$ to $\mathbf{x}_d$ (where $\mathbf{r}_d = \mathbf{x}_d$), there is no terminal cost, and the control scheme is characterized by its terminal constraint (9). Given the predictive model of the system (3) and associated cost function (5), and based on the defined state origin, $\mathbf{x}_d$, the closed subset of feasible system states includes the origin, and consequently, the first axiom presented in [38] is satisfied. On the other hand, while in the origin we have $\mathbf{x} = \mathbf{x}_d$, the corresponding control signal vanishes, and the second axiom is fulfilled, too. According to the state-space model of the system, at the origin we will have,

$$\mathbf{A}\mathbf{x}(k) + \mathbf{B}\tau(\mathbf{x}(k)) := \mathbf{I}\mathbf{x}_d + \delta_t \mathbf{L}_Z.\mathbf{0} = \mathbf{x}_d, \tag{14}$$

where, $\mathbf{I}$ is identity matrix, and the third axiom is also satisfied. Finally, due to having the terminal equality (9), and since there is no terminal cost and $\tau(\mathbf{x}) \equiv \mathbf{0}$ we have,

$$\|\mathbf{r}_d - \mathbf{r}_d\|^2 = 0, \tag{15}$$

and satisfaction of the fourth axiom is also guaranteed, which implies the global asymptotic stability of the agents controlled by the proposed DVPC method.

5 Simulation Results

In this section, to compare the performance of the proposed control scheme with the state of art, and to demonstrate the capability of the proposed multi-agent formation control approach, under different conditions with significant uncertainty level and unexpected occlusion, some scenarios have been designed and simulated.

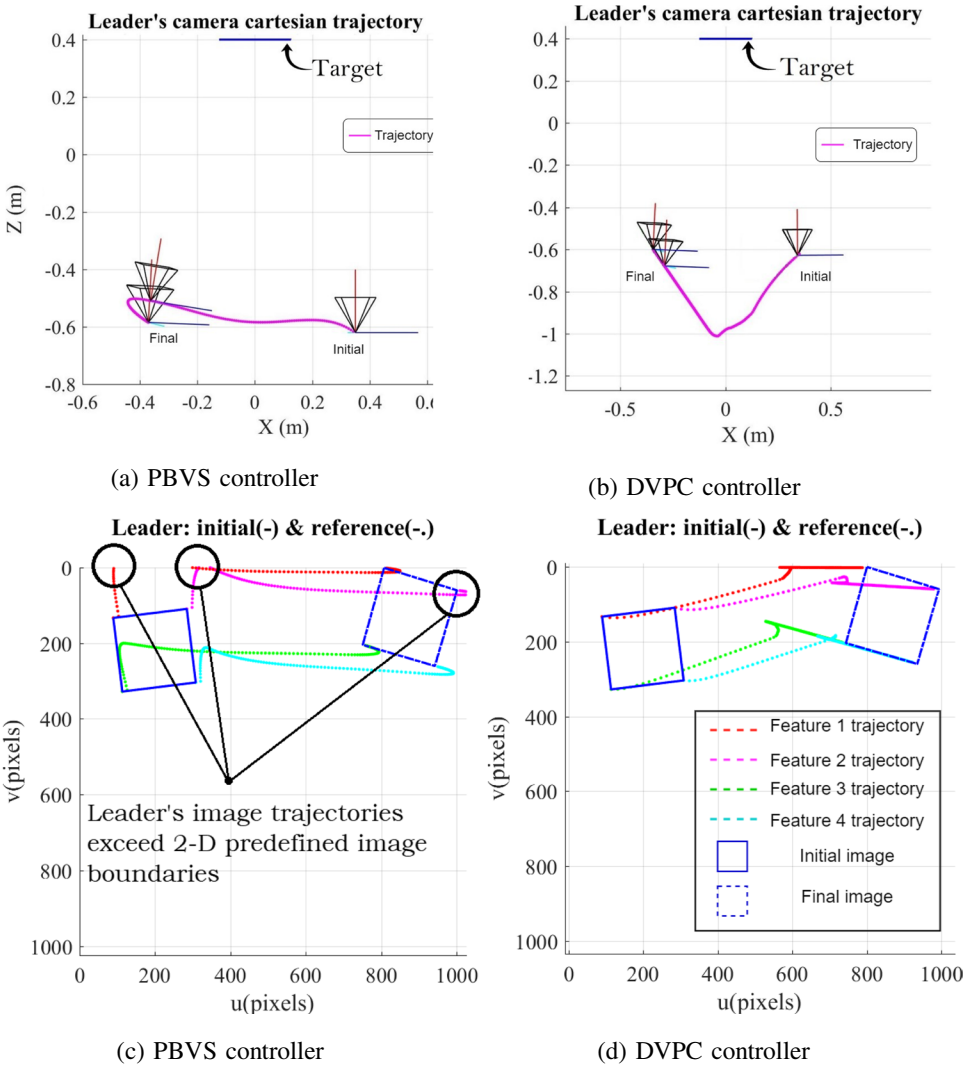
The sampling time step is set to $\delta_t = 0.03$ s, and all agents including leader and followers are utilizing the same camera type with focal length and depth of field of 8 mm, and 1.5 m, respectively, and the pixel size of $1e-05$ m by $1e-05$ m. The number of pixels is 1024×1024 and the principal point is located in the middle of image plane at (512,512). Therefore, in the following simulations, $\mathbf{p}_{\min} = (u_{\min}, v_{\min}) = (0, 0)$ and $\mathbf{p}_{\max} = (u_{\max}, v_{\max}) = (1024, 1024)$. In all of the following scenarios, each agent includes a pin-hole camera with 6 DOF, flying freely in the 3-D space, and we consider ± 0.25 m/s (rad/s) as the velocity limits to be fulfilled within the optimization procedure. The prediction horizon, $N_p = 15$, and the control horizon, $N_c = 1$ which means that the control signal is constant within the optimization procedure. In all of the following scenarios, at each iteration, first all of the available followers have to reach their reference relative pose with respect to the leader, and after that leader keeps on its motion toward its destination.

5.1 A comparison with PBVS approach

To validate our work in contrast to the state of the art, two different scenarios have been conducted. In the first scenario, the presented control approach has been compared with classic position-based visual servo (PBVS) technique [14–24]. The main target includes 4 coplanar points with absolute coordinates, $[-0.125, -0.125, 0.4]^{\top}$, $[-0.125, 0.125, 0.4]^{\top}$, $[0.125, 0.125, 0.4]^{\top}$, and $[0.125, -0.125, 0.4]^{\top}$, respectively, distinguished with a dark blue square as shown in Figs. 4 and 5.[1] The initial leader's camera frame is located at $[0.35, 0.42, -0.62]^{\top}$, and is rotated 7 deg around the z-axis. The initial pixel-wise coordinates of the main target points in the leader's image plane are, respectively, (90,133), (114,327), (309,304), and (285,109). Leader is gradually converging to its predefined reference pose where its frame center is located at $[-0.353, 0.513, -0.59]^{\top}$, and is rotated around x-, y-, and z-axes by -2.9 deg, 3 deg, and -14.9 deg, respectively. Based on the leader's reference configuration, the final pixel-wise coordinates of the main target points are, (800,1), (744,204), (935,260), and (993,60).

As it is clear in Fig. 4, although both of the controllers have a successful performance in controlling the

[1] For the sake of brevity, in the rest of this paper units are in meter, expressed in the global coordinate frame, unless specified otherwise.



(a) PBVS controller

(b) DVPC controller

(c) PBVS controller

(d) DVPC controller

Fig. 4 Leader's trajectories in the scenarios with PBVS and DVPC: Due to having no direct control over visual features in classic PBVS, visual features may leave the camera's field of view within the control procedure (Fig. 4c). On the other hand, in DVPC, the camera adjusts its Cartesian trajectory such that the visual features remain in its field of view (Fig. 4d)

camera 3-D trajectory, and all of the target points are kept within the predefined depth of field (Fig. 4a and b), but due to having no direct control over 2-D visual features, the leader controlled by PBVS misses two visual features within the control procedure which leads to the system interruption and failure (Fig. 4c). On the other hand, the proposed controller has direct control over 2-D image features and adjusts the camera trajectory such that all of the visual features are kept in the field of view while the leader converging to its reference configuration (Fig. 4d).

There is a secondary target attached firmly to the leader's coordinate frame, and indicated by a green square in Fig. 5a to c which consists of four corners of a 18.75×18.75 cm^2 square, and their relative coordinates expressed in the leader's camera frame are $[-0.2437,0,-0.0938]^\top$, $[-0.2437,0,0.0938]^\top$, $[-0.0562,0,0.0938]^\top$, and $[-0.0562,0,-0.0938]^\top$, respectively.

The leader is tracked by a follower whose coordinate system is located at points $[0.027,-1.011,-0.07]^\top$, expressed in the leader's camera frame, and is rotated by -75.7 deg, 20.2 deg, and 11.9 deg around x-, y-, and z-axes, respectively. In the initial configuration which is the follower's reference configuration, too, the measured pixel-wise coordinates of the leader's target points in the follower's image plane are, (500,851), (542,700), (688,743), and (650,907), respectively. The follower has to maintain its relative pose with respect to the leader's target set, and consequently, the leader's camera frame.

As it is clear in Fig. 5, by utilizing the proposed control method, the leader's camera converges to its preassigned final configuration successfully (Fig. 5d and f), while the follower maintains its relative pose with respect to the leader's camera frame (Fig. 5e and g).

5.2 A comparison with IBVS approach

In the second scenario, the presented control approach has been compared with formation control methods using IBVS

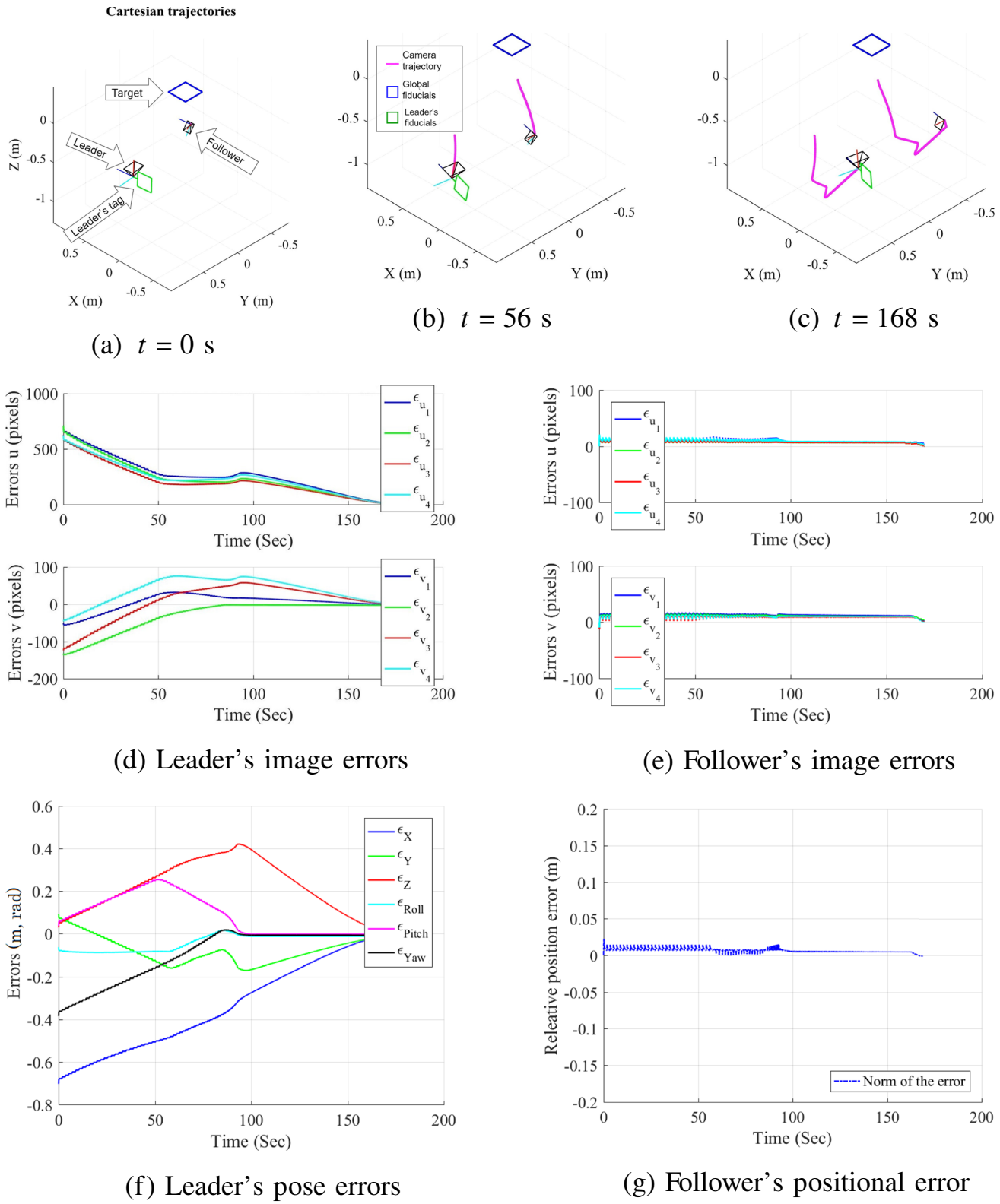
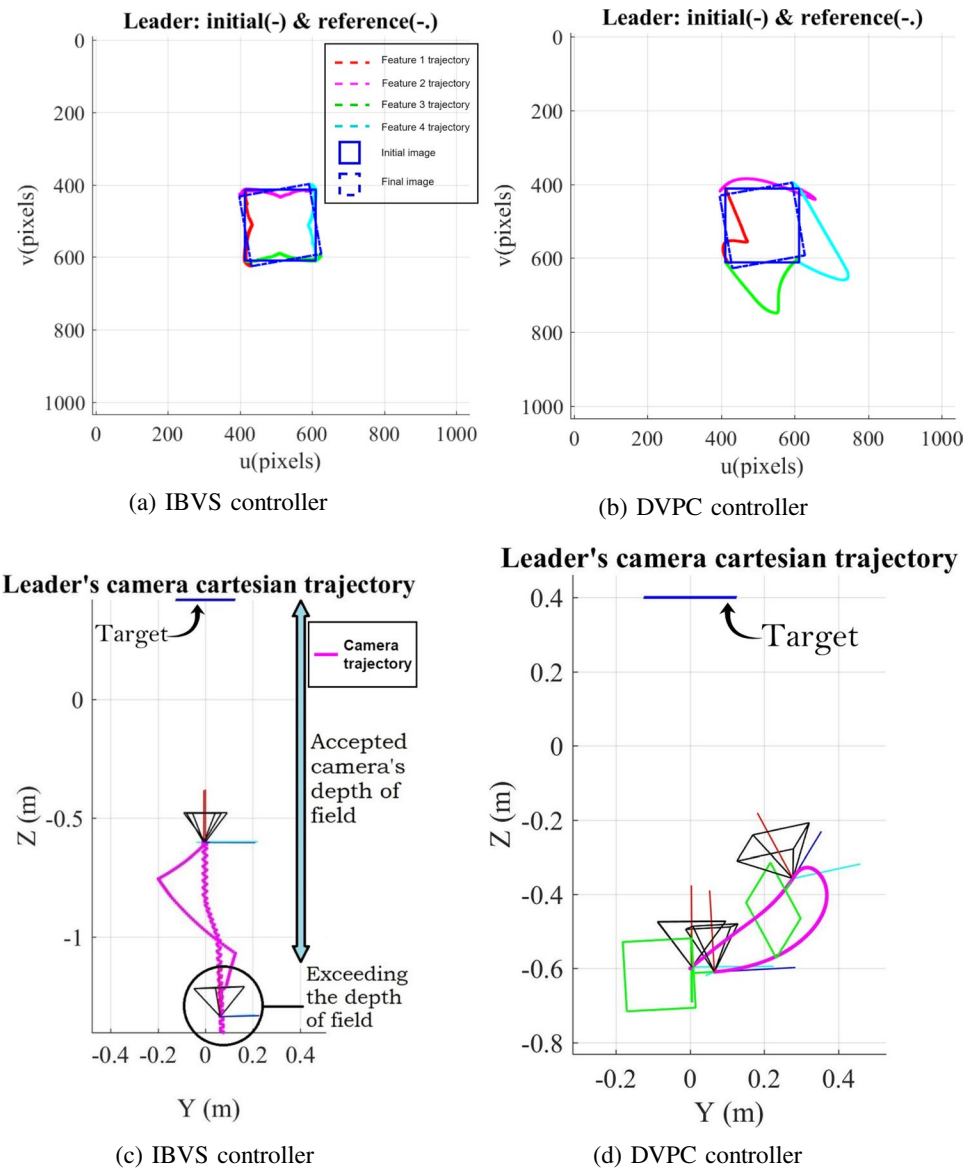


Fig. 5 Proposed formation control scheme: As it is demonstrated in Fig. 5a to c), the predefined form of the multi-agent system is kept constant, while the agents converge to their final configuration successfully, and the corresponding visual and spatial errors vanish (Fig. 5d to g)



(a) IBVS controller

(b) DVPC controller

(c) IBVS controller

(d) DVPC controller

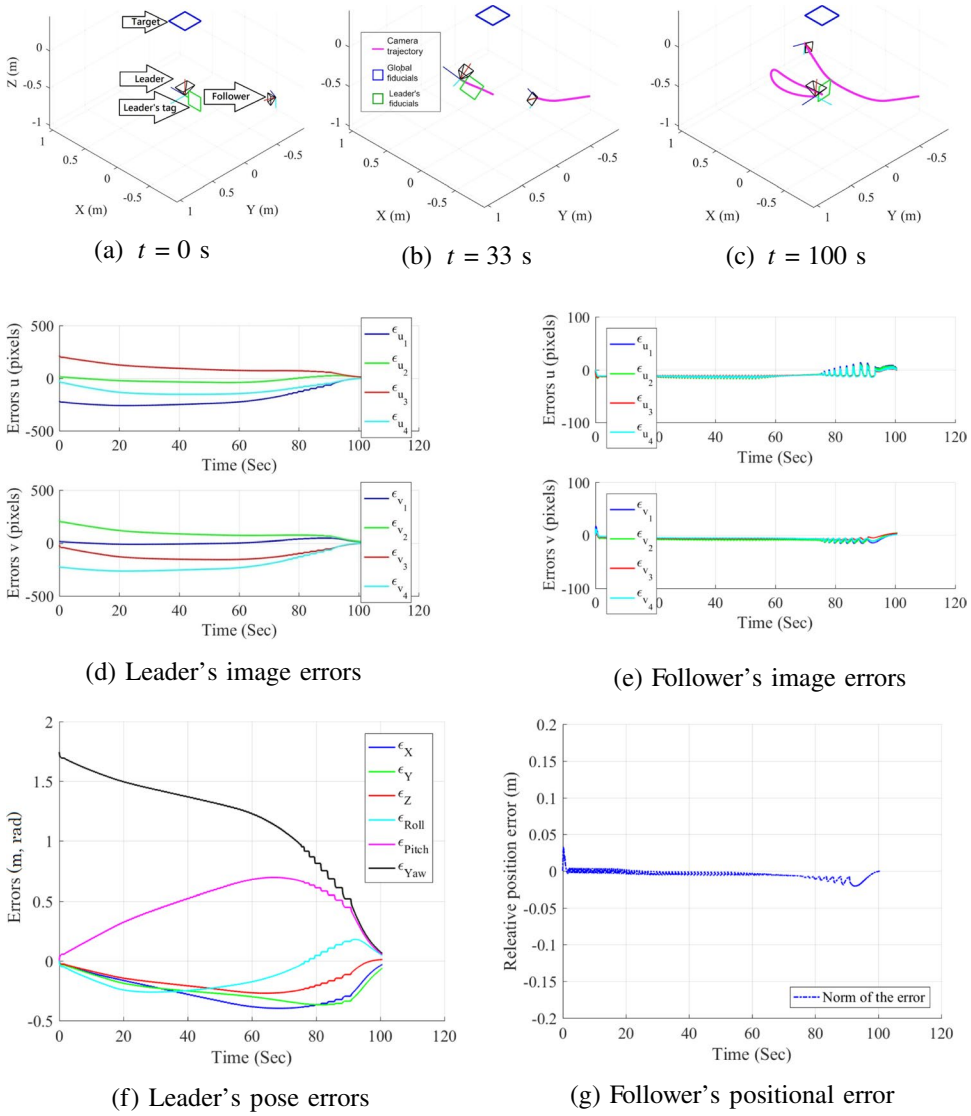
Fig. 6 Leader's trajectories in the scenarios with IBVS and DVPC: Due to having no direct control over the camera's Cartesian trajectory in classic IBVS, target points may exceed the boundaries of the accepted depth of field (Fig. 6d). On the other hand, the camera retreat problem could be handled easily in the system controlled by DVPC (Fig. 6c)

[11–13]. The similar target set as previous scenario - demonstrated with a dark blue square in Figs. 6 and 7 - is tracked by the leader's camera. Initially, the leader's camera frame is located at $[0,0,-0.6]^{\mathsf{T}}$, without any rotation. The initial pixel-wise coordinates of the main target points in the leader's image plane are (412,412), (412,612), (612,612), and (612,412), respectively. The reference pose of the leader is located at the same point as the initial one where its frame is rotated around z-axes by 100 deg. Based on this reference configuration, the final coordinates of the main target points are, (431,628), (628,593), (593,396), and (396,431), respectively.

As it is shown in Fig. 6, due to camera retreat problem, the leader's camera in the scenario with IBVS controller, exceeds the 1.5 m allowable depth of field (Fig. 6c), while the leader in the scenario with proposed controller converges to the final configuration, successfully (Fig. 6d).

Similar to the previous scenario, there is another target point set attached to the leader's camera frame, and demonstrated by a green square in Fig. 7a to c. This target is traced by a follower whose coordinate system is initially, located at the points $[-0.71,-0.7,-0.02]^{\mathsf{T}}$, expressed in the leader's camera frame, and is rotated by -90 deg, and 20 deg around x-, and y- axes, respectively. The measured pixel-wise coordinates of the leader's target points in the follower's image plane are (707,584), (707,401), (852,409), and (852,579), respectively, and follower has to maintain its relative pose with respect to the leader's camera frame, by tracking the mentioned secondary target set, similar to the Section 5.1.

As it is shown in Fig. 7, by employing the proposed control method, in contrast to the IBVS controller which suffers



(a) $t = 0$ s (b) $t = 33$ s (c) $t = 100$ s

(d) Leader's image errors (e) Follower's image errors

(f) Leader's pose errors (g) Follower's positional error

Fig. 7 Proposed formation control scheme: As it is demonstrated in Fig. 7a to c, the preassigned form of the multi-agent system is maintained, while each agent converges to its final configuration successfully, and the corresponding visual and spatial errors vanish (Fig. 7d to g)

from camera retreat problem, the leader's camera converges to its preassigned final configuration, successfully (Fig. 7d and f), while the follower maintains its relative pose with respect to the leader's camera frame, too (Fig. 7e and g).

5.2.1 Scenario with fully occluded leader[2]

In the following scenario, leader's camera is occluded within $t = 19$ s, and $t = 52$ s. This unexpected occlusion is indicated by black target points and black trajectories on the leader's image plane (Fig. 8), black spatial trajectory (Fig. 9), and by dashed lines on the corresponding diagrams (Figs. 10 to 12). In order to compensate this issue, there are two solutions. As the first solution, the leader's external model, temporarily, can estimate the necessary information until the issue is resolved. But, this method cannot be utilized when the main target itself is also time-varying. Therefore, it is assumed that, at least, one of the available followers has the main target in its field of view as shown in Fig. 8. The relative pose of two camera frames at each time instant, in addition to the coordinates of the main target points, could be measured and sent to the leader by the aforementioned assistant follower.

In this scenario, we not only consider that the leader is occluded, but also at the same time, all of the received measurement signals in addition to the generated control input are polluted by some level of noise to represent the real data transfer procedure issues (pollutant noises) between two agents and its undesired effects. As the first step, an

[2] The videos associated with leader's image plane and Cartesian trajectories of the networked agents are uploaded in the following links, respectively: https://www.youtube.com/watch?v=GlWUHPP2CFc. https://www.youtube.com/watch?v=dZvLYcinie4.

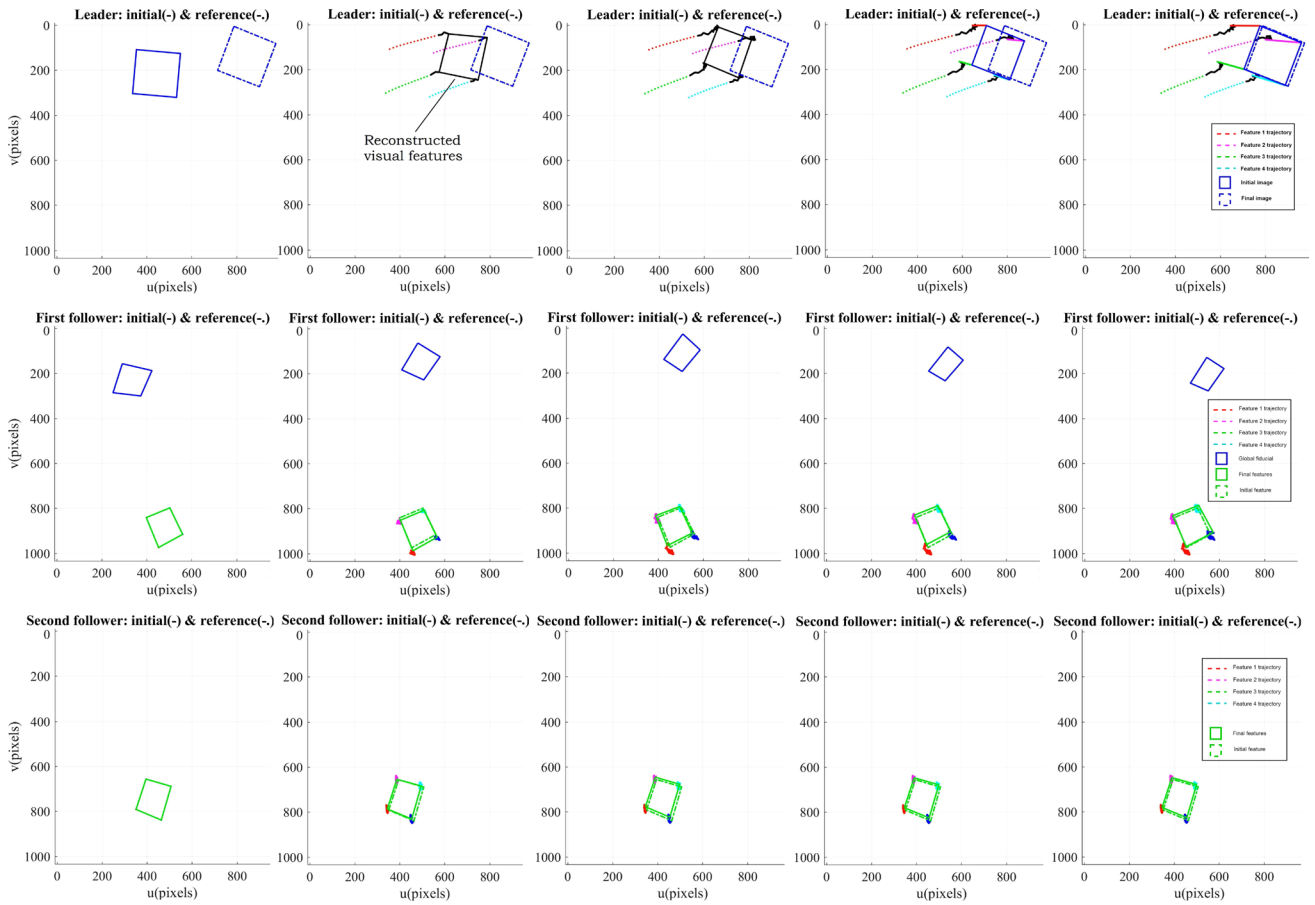


Fig. 8 Visual trajectories of agents, in the scenario with fully occluded leader: While the leader converges to its final configuration, all of the corresponding visual features remain in its field of view and the followers keep the predefined form

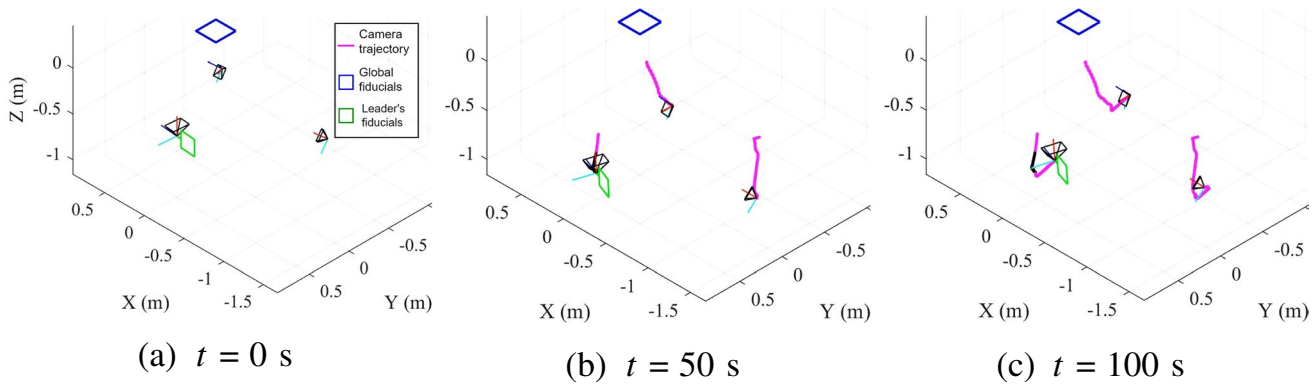


Fig. 9 Camera trajectories in the scenario with fully occluded leader: By utilizing the proposed data-sharing system among agents, the preassigned form of the multi-agents is maintained, while the system converges to its final configuration

additive white noise with an amplitude of 6 pixels, is added to the camera 2-D measurement which itself affects the final estimated depth, too. Additionally, to demonstrate the inaccuracy in depth estimation procedure, another multiplicative white noise with uniform distribution and amplitude equal to two percent of the estimated depth is added to it.

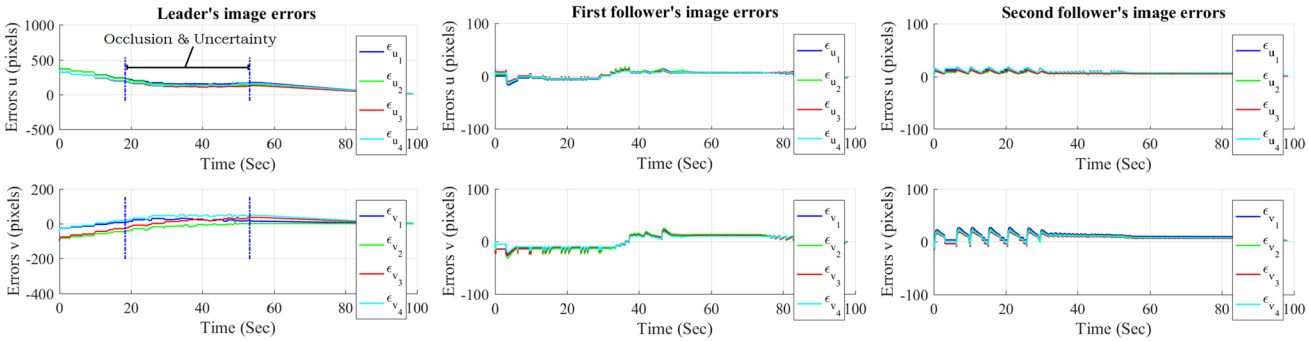


Fig. 10 Image errors of agents, in the scenario with fully occluded leader: As it is demonstrated, the corresponding visual errors vanish which means the successful convergence to the final configuration

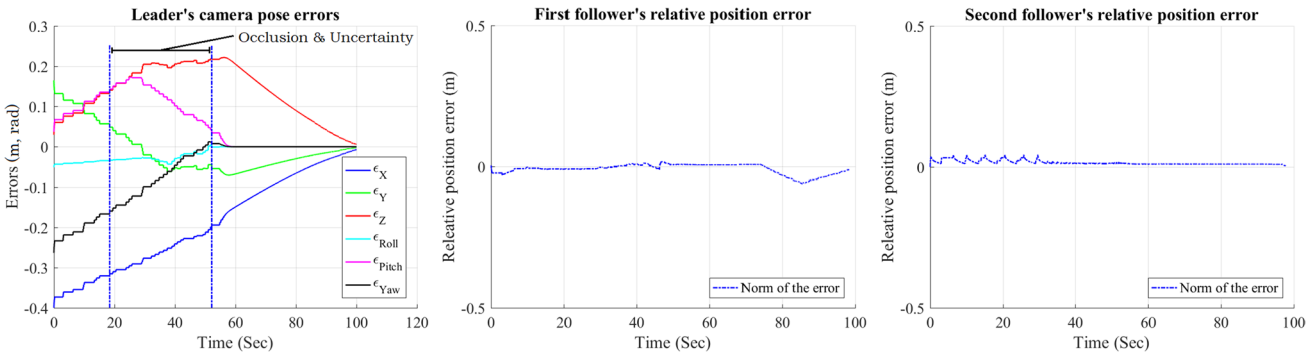


Fig. 11 Cartesian errors of agents, in the scenario with fully occluded leader: Similar to the previous figure, vanishing the Cartesian errors means the convergence to the reference configuration

Finally, control signal is polluted by another multiplicative white noise with an amplitude equal to three percent of its value.

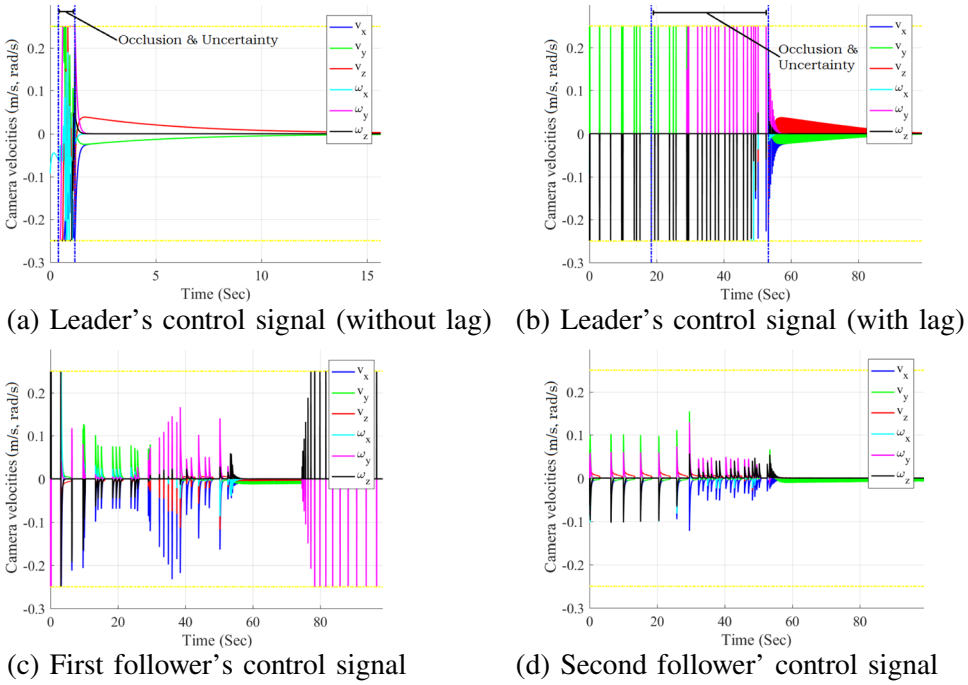
As it can be observed in Fig. 8, all the visual features stay inside the corresponding image planes. Specifically, in the leader's image plane and due to the closeness of the reference image to the boundaries of the plane, the camera motion is adjusted to meet the 2-D visual constraint. On the other hand, as it is demonstrated in the Fig. 9, the corresponding target points are located within the accepted range of cameras' depth of field, or ($0\,\text{m} < Z < 1.5\,\text{m}$), and the 3-D depth constraint is also met, accordingly. Finally, as shown in Fig. 12, all of the velocity parameters meet the predefined control constraint, and stay between two yellow dashed lines indicating the critical values ± 0.25 m/s (rad/s). As it is demonstrated in Figs. 10 and 11, by utilizing the proposed DVPC scheme and the second method to compensate the leader's occlusion, all of 2-D/3-D errors vanished gradually, and the corresponding agents converged to their final destinations successfully, while maintain their relative form.

5.2.2 Scenario with fully occluded follower[3]

In the next scenario, second follower's camera is occluded within $t = 36$ s, and $t = 50$ s, which is demonstrated by black target points on its image plane (Fig. 13), partially black 3-D trajectory (Fig. 14), and by dashed lines on its corresponding diagrams (Fig. 15).

Similar to the previous scenario, in order to resolve the issue, the missed coordinates of the secondary target points attached to the leader's camera are measured and sent to the occluded follower by its neighboring agent. In addition to the follower's occlusion, to demonstrate the undesirable effect of data transfer between agents, and to validate the robustness of the DVPC scheme against uncertainties, similar noises as Section 5.2.1 are added to the system. By

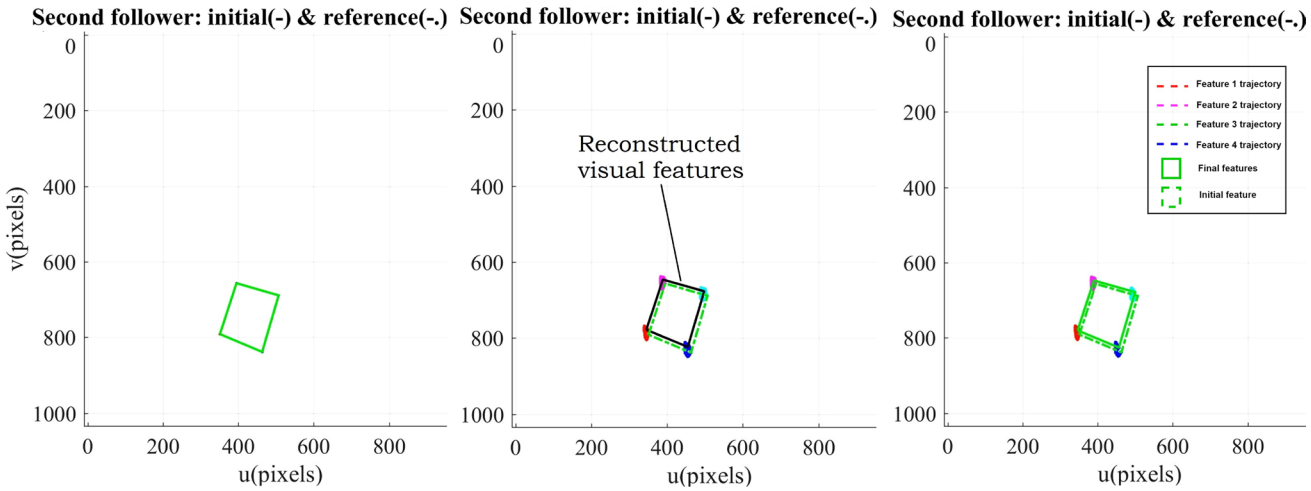
[3] The videos associated with second follower's image plane and Cartesian trajectories of the networked agents are uploaded in the following links, respectively: https://www.youtube.com/watch?v=th7Ni3aYvgk. https://www.youtube.com/watch?v=ENGUtgO80So.



(a) Leader's control signal (without lag) (b) Leader's control signal (with lag)

(c) First follower's control signal (d) Second follower' control signal

Fig. 12 Control signals of agents, in the scenario with fully occluded leader: As it is shown, all of the control inputs meet the predefined actuation constraints and remain within the minimum and maximum values which are depicted by yellow lines



(a) $t = 0$ s (b) $t = 50$ s (c) $t = 100$ s

Fig. 13 Visual trajectories of the second follower in the scenario with fully occluded follower: Although the follower becomes occluded, the data sent by its neighboring agents helps it to converge to its final configuration successfully

employing the proposed DVPC scheme, and providing the occluded agent by missed measurements, all of the active constraints are addressed, and all of the networked agents converge to their final configuration by maintaining preassigned form (Fig. 14). Same as the Section 5.2.1, the system is also robust against all of the injected noises (Fig. 15).

As shown in Fig. 13, despite the occlusion of the follower's view (observing the leader's features), the controller excels the VS, thanks to the homography-based communication system. As the results show, all the visual features stay inside the admissible image field-of-view constraint. Also, all of the manipulated variables (camera velocity) meet the predefined control constraint and are kept between two yellow dashed lines indicating the saturation values ± 0.25 m/s (rad/s). As it is demonstrated in Figs. 15 and 14, by utilizing the proposed DVPC scheme and homography-based communication method, all of the errors reduce gradually, and the agents are capable of maintaining their relative view (of the leader's features), despite occlusion.

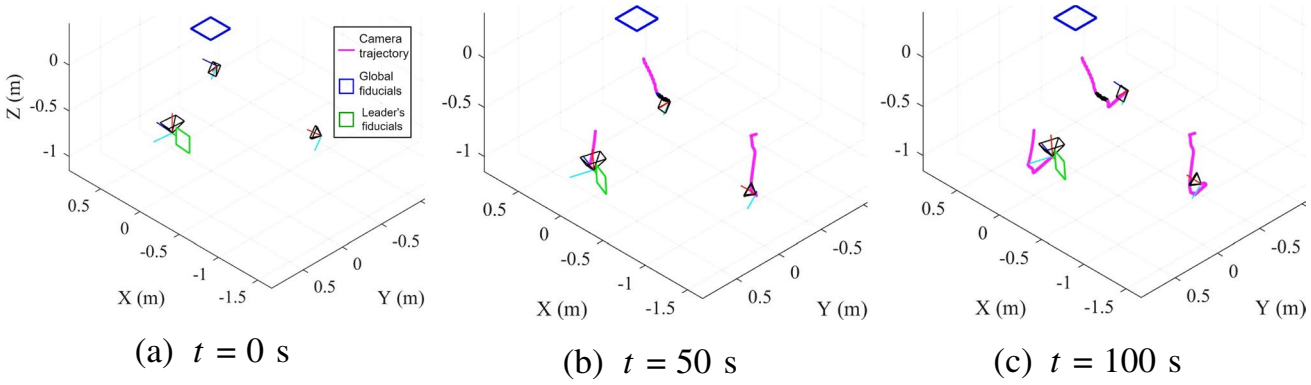


Fig. 14 Camera trajectories in the scenario with fully occluded follower: Similar to the previous figure, due to the embedded data sharing system, even the occluded follower can converge to the reference position successfully and maintain the predefined form

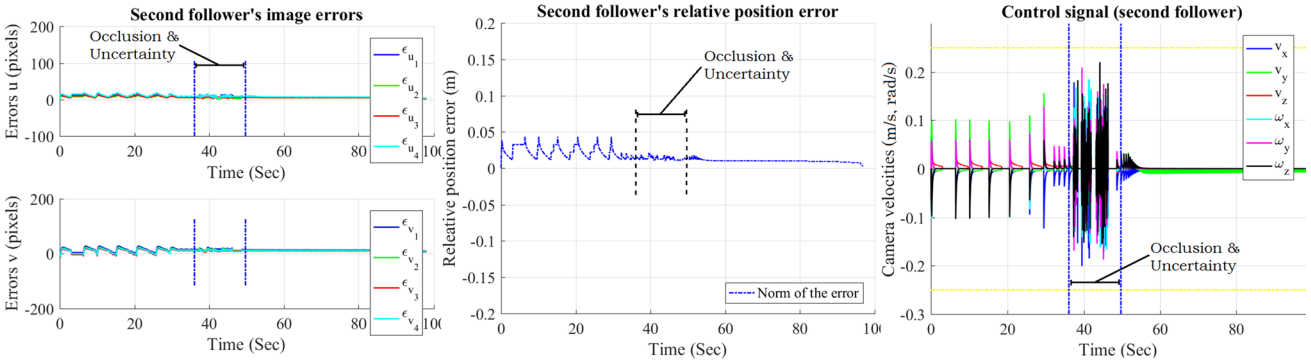


Fig. 15 Image errors, positional error, and control signal of the second follower, in the scenario with fully occluded follower. As it is depicted in the figure, while the follower converges to its reference position and corresponding errors vanish gradually, the control input meets the preassigned constraints

5.3 Incorporating kinematics/dynamics of the carrier platform

Up to this section, the paper has dealt only with free-flying cameras, ignoring the kinematics/dynamics of the carrier platform, following the trend in visual servoing literature which usually relies on free-flying cameras for the proof-of-the-concept [39]. In this section, the applicability of proposed DVPC approach in realworld scenarios is studied through implementation on multi-agent 6-DoF UAV systems.

The block diagram of the modified method is presented in Fig. 16. A two-tier cascaded controller similar to the idea proposed in [40] has been adopted to incorporate UAV dynamics in the simulations. In this diagram, the occlusion detection part is omitted for clarity. The external and internal model of the system has been updated to incorporate carrier robot Jacobian $\mathbf{J}_r$. This way, by including kinematics of the robot in the system, the camera velocity can be approximated by the term $\mathbf{J}_r\hat{\mathbf{v}}_r(k)$. Note that this kinematics based approximation neglects the dynamics of the robot for the favor of the real-time computational expense of the predictive controller. With the above mentioned adjustment, the modified internal model can be formulated as,

$$\begin{cases} \mathbf{x}(k+1) = \mathbf{x}(k) + \delta_t \mathbf{L}_Z \mathbf{J}_r \hat{\mathbf{v}}_r(k), \\ {}^c\mathbf{r}_m(k) = \mathbf{x}(k), \end{cases} \tag{16}$$

and the external model can be written in the form of,

$$\begin{cases} \mathbf{y}(k+1) = \mathbf{y}(k) + \delta_t \mathbf{L}_r \mathbf{J}_r \hat{\mathbf{v}}_r(k), \\ {}^c\mathbf{r}_o(k) = \mathbf{y}(k). \end{cases} \tag{17}$$

As the block diagram suggests, the DVPC with modified internal model calculates the optimised desired velocity and trajectory for the carrier robot, $\hat{\mathbf{v}}_r(k)$. The second stage controller, namely "Robot controller", is in charge of following the desired velocity trajectories generated by the previous stage.

A nonplanar 6 DoF multirotor vehicle proposed in [41] is considered as the camera carrier platform. The angled

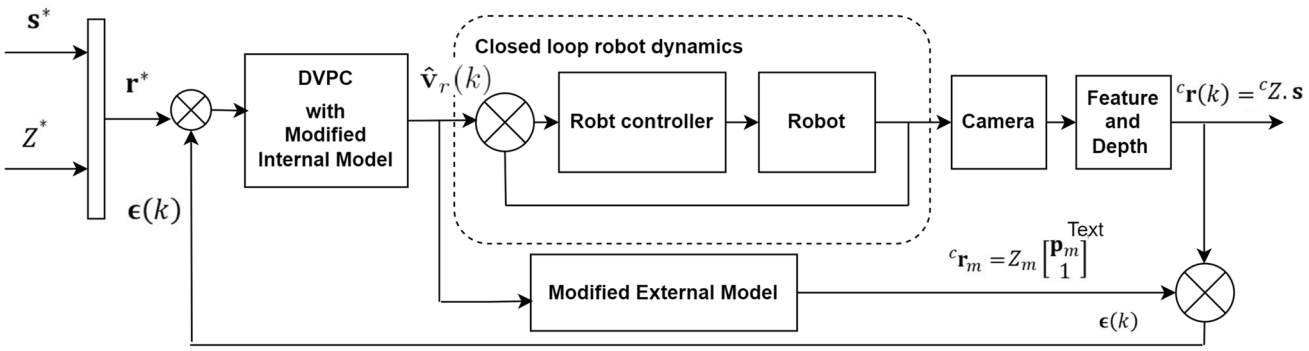


Fig. 16 Depth-based visual predictive control with robot dynamics

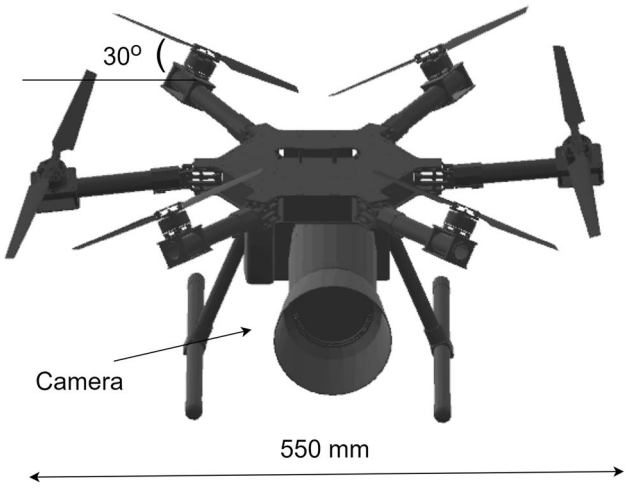


Fig. 17 6 DoF non-planar aerial vehicle equipped with camera

propeller arrangement of this class of hexacopters, shown in Fig. 17, provides the capability of independent control of position and attitude for both hover and flight scenarios, expanding the set of reachable motions from any instance of servoing. As a brevity measure, the dynamics of the UAV and the controller used in the simulation are presented in the Appendix. The results of a multi agent DVPC control scenario using 6 DoF drones is presented in the following.

The initial configurations of the leader and follower in a DVPC scenario are represented in Fig. 18. Also, the resulting image trajectories in the leader and followers' cameras are depicted in Fig. 19. The goal of the servoing for the leader is to reach the target image of global fiducial represented in Fig. 19a. On the other hand, the followers are keeping the initial image of the leaders fiducial intact by changing their relative pose to the leader as depicted in Fig. 19b and c. The motion trajectories of UAVs are also

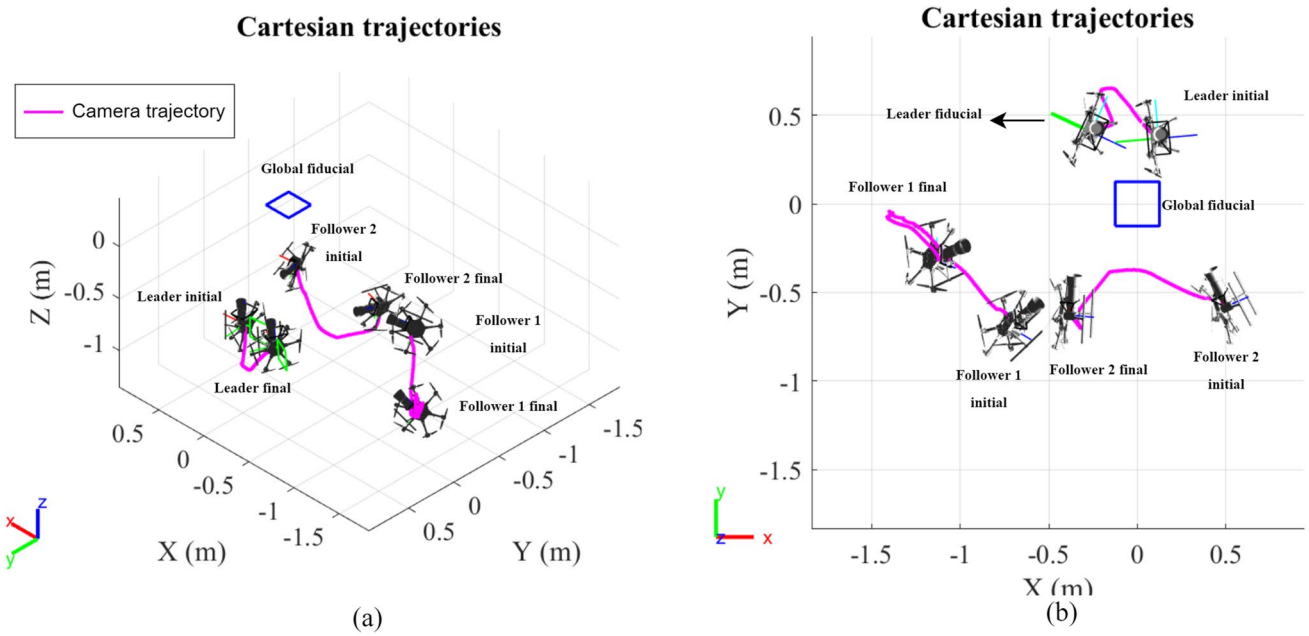


Fig. 18 Leader and follower UAV trajectories in a DVPC servoing scenario from different views

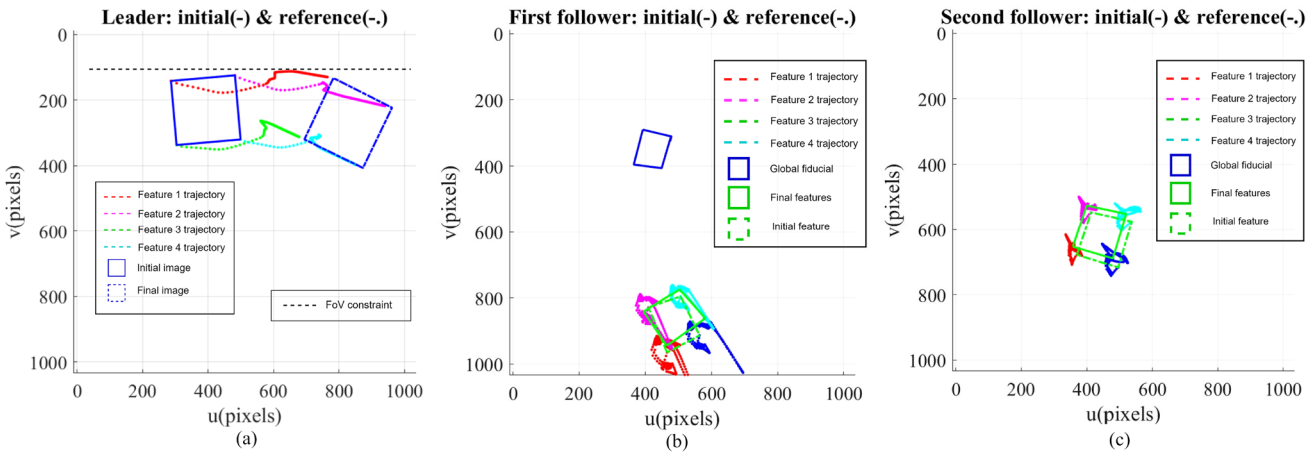


Fig. 19 Image trajectories: (**a**) leader's image trajectory, (**b**) the first follower's image trajectory and (**c**) the second follower's image trajectory

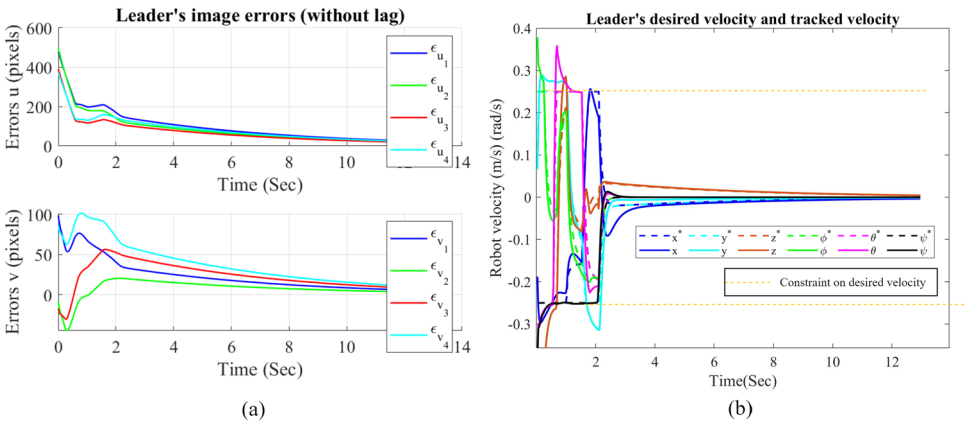


Fig. 20 (**a**) Leader's image error, and (**b**) desired robot velocity generated by visual controller and the tracked velocity by leader UAV

presented in Fig. 18 along with their final pose as result of servong. Figure 20a, shows that the leader excels at reaching the target view from the global fiducials, and as a result, image errors reduce with time. Figure 20b, demonstrates the constrained velocity generated by the visual controller and the velocity trajectories followed by the UAV. As shown in the graph, the DVPC takes the velocity constraint into account by generating a constrained reference velocity. However, a constraint aware controller such as MPC is required to function as the robot controller to fully comply the closed loop dynamics of the UAV with the constraints. Nevertheless, the generated reference velocities by the visual controller can be successfully limited to admissible levels.

6 Conclusion

An innovative multi-agent depth-based visual predictive control (DVPC) scheme is proposed in this paper to address some of the challenges in the position-based leader-follower formation control. The proposed framework shows promise in handling both images- and physical-space constraints while calculating the optimal trajectories for all the agents, within a prediction horizon. Furthermore, the superiority of the proposed controller in constraint handling was shown over that in the state-of-the-art image-based and position-based predictive control strategies via a myriad of simulated scenarios. A homography-transform, an image-based method for sharing data in scenarios with camera occlusion, was introduced as a means of further robustification of the proposed approach. The effectiveness of the proposed method in controlling a team of agents, including drones, was studied. The robustness to camera occlusion was explored in a variety of scenarios such as (1) partial occlusion of the leader, (2) complete occlusion of a follower, and (3) partial occlusion in neighbouring followers. The results show that the proposed DVPC approach successfully excels in the multi-agent servoing task while taking sensor and actuators constraints into account.

Appendix

The motion dynamics of the UAV [41], equipped by camera can be described by the following equations,

$$\begin{aligned}
\dot{x_1} &= x_2 \\
\dot{x_2} &= \begin{bmatrix} 0 \\ 0 \\ -g \end{bmatrix} + \frac{1}{m} \begin{bmatrix} c_\phi c_\psi - c_\theta s_\phi s_\psi & -c_\psi s_\phi - c_\phi c_\theta s_\psi & s_\theta s_\psi \\ c_\theta c_\psi s_\phi + c_\phi s_\psi & c_\phi c_\theta c_\psi - s_\phi s_\psi & -c_\psi s_\theta \\ s_\phi s_\theta & c_\phi s_\theta & c_\theta \end{bmatrix} \begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix} \\
\dot{x_3} &= \begin{bmatrix} 1 & 0 & -s_\theta \\ 0 & c_\phi & c_\theta s_\phi \\ 0 & -s_\phi & c_\theta c_\phi \end{bmatrix}^{-1} x_4 \\
\dot{x_4} &= \begin{bmatrix} \tau_\phi I_{xx} - 1 \\ \tau_\theta I_{yy} - 1 \\ \tau_\psi I_{zz} - 1 \end{bmatrix} - \begin{bmatrix} \frac{I_{yy} - I_{zz}}{I_{xx}} \omega_y \omega_z \\ \frac{I_{zz} - I_{xx}}{I_{yy}} \omega_x \omega_z \\ \frac{I_{xx} - I_{yy}}{I_{zz}} \omega_x \omega_y \end{bmatrix}
\end{aligned} \tag{18}$$

where, s and c refer to sin() and cos () operators and $x = (x_1,x_2,x_3,x_4)$ represents the state vector of the vehicle. Also, $x_1 = (x,y,z)$ and $x_2 = (\dot{x}, \dot{y}, \dot{z})$ are position and velocity of the UAV. The attitude of the UAV in the reference frame is denoted by $x_3 = (\theta,\phi,\psi)$, where θ, ϕ, ψ denote the roll, pitch and yaw in the global frame. In addition, the angular velocity of UAV in the body frame is described by $x_4 = (\omega_x,\omega_y,\omega_z)$. It worths mentioning that the angular velocity vector $\omega \neq \dot{\theta}$. In fact, the angular velocity is a vector pointing along the axis of rotation in the body frame, while $\dot{\theta}$ represents time derivative of pitch, yaw and roll. The manipulated variable for controlling the pose of the vehicle is the force, $\mathbf{F} = \begin{bmatrix} f_x & f_y & f_z \end{bmatrix}^{\mathrm{T}}$, and torque, $\mathbf{M} = \begin{bmatrix} \tau_\phi & \tau_\theta & \tau_\psi \end{bmatrix}^{\mathrm{T}}$ vector generated by the propellers. A linear model can be formulated to relate each propeller's angular velocity to generated force f and torque τ.

$$\begin{aligned} f &= K_f \cdot \omega^2 \\ \tau &= K_\tau \cdot \omega^2 \end{aligned} \tag{19}$$

where, K_f and K_τ are the force and torque coefficients, ω represent the propellers rotation velocity. The trasnfer function, T can be used to calculate the velocity of each propeller based on the required $\mathbf{F}$ and $\mathbf{M}$ vector calculated the controller.

$$\mathbf{\Omega} = T^{-1} \begin{bmatrix} \mathbf{F} \\ \mathbf{M} \end{bmatrix}$$

$$T = \mathrm{diag}\left(\frac{\sqrt{3}K_f s_\gamma}{2}, \frac{K_f s_\gamma}{2}, K_f c_\gamma, \frac{LK_f c_\gamma}{2}, \frac{\sqrt{3}LK_f s_\gamma}{2}, 1 \right) \begin{bmatrix} 1 & 0 & -1 & 1 & 0 & -1 \\ -1 & 2 & -1 & -1 & 2 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 \\ 1 & 2 & 1 & -1 & -2 & -1 \\ 1 & 0 & -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Table 1 Dynamics and control parameters of the UAV

Parameter	Value
Thrust constant (K_f)	6×10^{-3} N/(rad/s)^2
Torque constant (K_τ)	6×1^{-2} Nm/(rad/s)^2
Propeller distance (L)	0.3m
Mass (m)	1.3 kg
Inertia along x (I_{xx})	3×10^{-2} kgm^2
Inertia along y (I_{yy})	3.2×10^{-2} kgm^2
Inertia along z (I_{zz})	5×10^{-2} kgm^2
LQR gain (K)	diag([100, 100, 300, 10, 10, 10, 300, 300, 300, 10, 10, 10])
Propeller angle (γ)	30°

$$+ \mathrm{diag}\left(1, 1, 1, \frac{\sqrt{3}K_\tau s_\gamma}{2}, \frac{K_\tau s_\gamma}{2}, K_\tau c_\gamma \right) \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 1 & 0 & -1 \\ 1 & 2 & 1 & -1 & -2 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 \end{bmatrix}$$

where, the vector $\mathbf{\Omega} = \begin{bmatrix} \omega_1^2 & \omega_2^2 & \omega_3^2 & \omega_4^2 & \omega_5^2 & \omega_6^2 \end{bmatrix}^{\mathrm{T}}$ can be formed by square of spinning velocity of each propeller.

For simulations of the paper, an LQR controller was design by successively linearizing system dynamics using Jacobian in each control step. The successive linearization model-based LQR controller found to be effective in stabilizing and controlling the position and orientation of the proposed UAV. The rest of the dynamics and controller parameters used in the simulation can be found in Table 1.

Acknowledgments This work was sponsored by National Sciences and Engineering Research Council of Canada (NSERC) through Discovery Grant #2017 06930.

Author Contributions M. M. H. Fallah developed formulation and simulations, and authored the manuscript. F. Janabi-Sharifi supervised M. M. H. Fallah in problem formulation and control development, and participated in the authorship of the manuscript. S.Sajjadi contributed in formatting and editing of the paper and the simulations with the UAV section. M.Mehrandezh provided technical advise, reviewed and edited the paper.

Funding This work was sponsored by National Sciences and Engineering Research Council of Canada (NSERC) through Discovery Grant #2017 06930.

Data availability None

Code Availability None

Declarations

Ethics approval Not applicable

Consent to participate Not applicable

Consent for Publication Not applicable

Conflict of Interests The authors declare that they have no competing interests.

References

1. Oh, K.K., Park, M.C., Ahn, H.S.: A survey of multi-agent formation control. Automatica **53**, 424–440 (2015)
2. Zhang, Y., Mehrjerdi, H.: A survey on multiple unmanned vehicles formation control and coordination: normal and fault situations. In: Proc. 2013 International Conference on Unmanned Aircraft Systems (ICUAS), Atlanta, GA, USA, pp. 1087–1096 (2013)
3. Hu, J., Xu, J., Xie, L.: Cooperative search and exploration in robotic networks. Unmanned Systems **1**(01), 121–142 (2013)
4. Panagou, D., Kyriakopoulos, K.J.: Cooperative formation control of underactuated marine vehicles for target surveillance under sensing and communication constraints. In: Proc. 2013 IEEE International Conference on Robotics and Automation, Karlsruhe, Germany, pp. 1871–1876 (2013)
5. der Walle, D.V., Fidan, B., Sutton, A., Yu, C., Anderson, B.D.O.: Non-hierarchical Uav formation control for surveillance tasks. In: Proc. IEEE 2008 American Control Conference (ACC), Seattle, WA, USA, pp. 777–782 (2008)
6. Kim, J.H., Kwon, J.W., Seo, J.: Multi-uav-based stereo vision system without gps for ground obstacle mapping to assist path planning of ugv. Electron. Lett. **50**(20), 1431–1432 (2014)
7. Hutchinson, S., Hager, G.D., Corke, P.I.: A tutorial on visual servo control. IEEE Trans. Robot. Autom. **12**(5), 651–670 (1996)
8. Kragic, D., Christensen, H.I., et al.: Survey on visual servoing for manipulation. Computational Vision and Active Perception Laboratory, Fiskartorpsv **15**, 2002 (2002)
9. Cho, H.: Opto-mechatronic Systems handbook: Techniques and Applications. CRC Press, Massachusetts, US (2002)
10. Janabi-Sharifi, F., Deng, L., Wilson, W.J.: Comparison of basic visual servoing methods. IEEE/ASME Transactions on Mechatronics **16**(5), 967–983 (2010)
11. Chen, X., Jia, Y.: Adaptive leader-follower formation control of non-holonomic mobile robots using active vision. IET Control Theory & Applications **9**(8), 1302–1311 (2015)
12. Wang, H., Guo, D., Liang, X., Chen, W., Hu, G., Leang, K.K.: Adaptive vision-based leader–follower formation control of mobile robots. IEEE Trans. Ind. Electron. **64**(4), 2893–2902 (2016)
13. Liang, X., Wang, H., Liu, Y.H., Chen, W., Liu, T.: Formation control of nonholonomic mobile robots without position and velocity measurements. IEEE Trans. Robot. **34**(2), 434–446 (2017)
14. Chueh, M., Yeung, Y.L.W.A., Lei, K.P.C., Joshi, S.S.: Following controller for autonomous mobile robots using behavioral cues. IEEE Trans. Ind. Electron. **55**(8), 3124–3132 (2008)
15. Fathian, K., Doucette, E., Curtis, J.W., Gans, N.R.: Vision-based distributed formation control of unmanned aerial vehicles. arXiv:1809.00096 (2018)
16. Das, A.K., Fierro, R., Kumar, V., Ostrowski, J.P., Spletzer, J., Taylor, C.J.: A vision-based formation control framework. IEEE Trans. Robot. Autom. **18**(5), 813–825 (2002)
17. Mariottini, G.L., Morbidi, F., Prattichizzo, D., Pappas, G.J., Daniilidis, K.: Leader-follower formations: uncalibrated vision-based localization and control. In: Proc. 2007 IEEE International Conference on Robotics and Automation, Roma, Italy, pp. 2403–2408 (2007)
18. Mariottini, G.L., Morbidi, F., Prattichizzo, D., Valk, N.V., Michael, N., Pappas, G., Daniilidis, K.: Vision-based localization for leader–follower formation control. IEEE Trans. Robot. **25**(6), 1431–1438 (2009)
19. Mariottini, G.L., Pappas, G., Prattichizzo, D., Daniilidis, K.: Vision-based localization of leader-follower formations. In: Proc. 44Th IEEE Conference on Decision and Control, Seville, Spain, Pp. 635–640 (2005)
20. Dani, A.P., Gans, N., Dixon, W.E.: Position-based visual servo control of leader-follower formation using image-based relative pose and relative velocity estimation. In: Proc. 2009 American Control Conference (ACC), St. Louis, MO, USA, pp. 5271–5276 (2009)
21. Fidan, B., Gazi, V., Zhai, S., Cen, N., Karataş, E.: Single-view distance-estimation-based formation control of robotic swarms. IEEE Trans. Ind. Electron. **60**(12), 5781–5791 (2012)
22. Liang, X., Liu, Y.H., Wang, H., Chen, W., Xing, K., Liu, T.: Leader-following formation tracking control of mobile robots without direct position measurements. IEEE Trans. Autom. Control **61**(12), 4131–4137 (2016)
23. Orqueda, O.A.A., Fierro, R.: Robust Vision-Based Nonlinear Formation Control. In: Proc. 2006 American Control Conference (ACC), Minneapolis, MN, USA, pp. 1422–1427 (2006)
24. Poonawala, H., Satici, A.C., Gans, N., Spong, M.W.: Formation Control of Wheeled Robots with Vision-Based Position Measurement. In: Proc. 2012 American Control Conference (ACC), Montréal, Canada, pp. 3173–3178 (2012)
25. Chaumette, F., Hutchinson, S.: Visual servo control. i. basic approaches. IEEE Robotics & Automation Magazine **13**(4), 82–90 (2006)
26. Miao, Z., Zhong, H., Wang, Y., Zhang, H., Tan, H., Fierro, R.: Low complexity leader-following formation control of mobile robots using only fov-constrained visual feedback. IEEE Transactions on Industrial Informatics, pp. 1–1 (2021)
27. Miao, Z., Zhong, H., Lin, J., Wang, Y., Chen, Y., Fierro, R.: Vision-based formation control of mobile robots with fov constraints and unknown feature depth. IEEE Trans. Control Syst. Technol. **29**(5), 2231–2238 (2021)
28. Li, Z., Yuan, Y., Ke, F., He, W., Su, C.-Y.: Robust vision-based tube model predictive control of multiple mobile robots for leader–follower formation. IEEE Trans. Ind. Electron. **67**(4), 3096–3106 (2019)
29. Lin, J., Miao, Z., Zhong, H., Peng, W., Wang, Y., Fierro, R.: Adaptive image-based leader–follower formation control of mobile robots with visibility constraints. IEEE Trans. Ind. Electron. **68**(7), 6010–6019 (2020)
30. Allibert, G., Courtial, E., Chaumette, F.: Predictive control for constrained image-based visual servoing. IEEE Trans. Robot. **26**(5), 933–939 (2010)
31. Fallah, M.M.H., Ghazbi, S.N., Mehrkish, A., Janabi-Sharifi, F.: Depth-based visual predictive control of tendon-driven continuum robots. In: Proc. of IEEE/ASME International Conference on Advanced Intelligent Mechatronics, Boston, MA, USA, pp. 488–494 (2020)
32. Fallah, M.M.H., Janabi-Sharifi, F.: Linear position-based visual predictive control. In: Proc. of the Canadian Society for Mechanical Engineering International Congress, p 2020. Charlottetown, PEI, Canada (2020)
33. Sauvée, M., Poignet, P., Dombre, E.: Ultrasound image-based visual servoing of a surgical instrument through nonlinear model

predictive control. The International Journal of Robotics Research **27**(1), 25–40 (2008)
34. Sauvée, M., Poignet, P., Dombre, E., Courtial, E.: Image based visual servoing through nonlinear model predictive control. In: Proc. 45Th IEEE Conference on Decision and Control, San Diego, CA, USA, pp. 1776–1781 (2006)
35. Economou, C.G., Morari, M., Bernhard, B.O.: Internal model control: Extension to nonlinear system. Industrial & Engineering Chemistry Process Design and Development **25**(2), 403–411 (1986)
36. Cervera, E., Pobil, A.P.D., Berry, F., Martinet, P.: Improving image-based visual servoing with three-dimensional features. The International Journal of Robotics Research **22**(10-11), 821–839 (2003)
37. Fallah, M.M.H., Janabi-Sharifi, F.: Conjugated visual predictive control for constrained visual servoing. Journal of Intelligent & Robotic Systems **101**(26), 1–21 (2021)
38. Mayne, D.Q., Rawlings, J.B., Rao, C.V., Scokaert, P.O.: Constrained model predictive control: Stability and optimality. Automatica **36**(6), 789–814 (2000)
39. Chaumette, F., Hutchinson, S.: Visual servo control. ii. advanced approaches [tutorial]. IEEE Robotics & Automation Magazine **14**(1), 109–118 (2007)
40. Sajjadi, S., Mehrandezh, M., Janabi-Sharifi, F.: A nonlinear adaptive model-predictive approach for visual servoing of unmanned aerial vehicles. In: Martínez-García, A., Bhattacharya, I., Otani, Y., Tutsch, R (eds.) Progress in Optomechatronic Technologies, pp 153–164. Springer, Singapore (2019)
41. Crowther, B., Lanzon, A., Maya-Gonzalez, M., Langkamp, D.: Kinematic analysis and control design for a nonplanar multirotor vehicle. Journal of Guidance, Control, and Dynamics **34**(4), 1157–1171 (2011)

Mostafa Mohammad Hossein Fallah received the B.Sc. and the first M.Sc. degrees in mechanical engineering from Sharif University of Technology of Tehran, Iran in 2007 and 2009 respectively. After working for Hamoun Nyzeh Company (Hanyco)– manufacturer of ductile iron pipes in Iran for four years he started his second M.Sc. at the University of Ottawa, Canada in 2013. He was teaching and research assistant from 2013 to 2015 at the University of Ottawa and started his Ph.D. in the Department of Mechanical and Industrial Engineering of Ryerson University in 2016, where he is currently a graduate assistant in RMAL (Robotics, Mechatronics, and Automation Lab). His research interests include Robotics, Control Systems and Computer Vision.

Farrokh Janabi-Sharifi received the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 1995. He is currently a Professor of Mechanical and 1544 Industrial Engineering and the Director of Robotics, Mechatronics and Automation Laboratory (RMAL), Ryerson University, Toronto, ON, Canada. He was an NSERC Postdoctoral Fellow and Instructor with the Center for Intelligent Machines, Department of Electrical and Computer Engineering, McGill University, Montréal, QC, Canada, (1995–1997). His research interests include visual servoing with a focus on medical robotics, continuum robots, and unmanned systems. Dr. Janabi-Sharif has been an Associate Editor of the International Journal of Optomechatronics and a Technical Editor of the IEEE/ASME TRANSACTIONS ON MECHATRONICS. Dr. Janabi-Sharifi is a fellow of the Canadian Society of Mechanical Engineers, Engineering Institute of Canada, and Canadian Academy of Engineering.

Sina Sajjadi received his M.Sc. degree in mechatronic engineering from Tabriz University, Tabriz, Iran, in 2018. He is pursuing a Ph.D. degree in industrial systems engineering at the Faculty of Engineering and Applied Science, University of Regina, SK, Canada. He is currently a researcher at the Robotics, Mechatronics, and Automation Laboratory, Ryerson University. His research interests include robotics, control systems, machine vision and machine learning.

Mehran Mehrandezh did postdoctoral studies at Simon Fraser University in Canada in 2000-2001. He received his Ph.D. and M.A.Sc. in Mechanical Engineering from the University of Toronto, and the Queens University in 2000 and 1995, respectively. His research revolves around control and machine vision with applications in design, development, and utilization of unmanned systems. The pipe crawling robot co-invented by him was highlighted as one of the five high-tech fixes to infrastructure in the Popular Mechanics Magazine, with a worldwide readership, in 2009. He also holds a US patent on a self-reconfigurable and adaptable rope climbing vehicle. Dr. Mehrandezh was a recipient of a 3-year grant under the Strategic Grant Program (SPG) by NSERC to work on vision-based control of UAVs. Dr. Mehrandezh has published more than 60 papers in peer-reviewed and reputed journals, conference proceedings, and book chapters related to optomechatronics systems and vison-based control of Unmanned Aerial and Ground Vehicles.

Dr. Mehrandezh's team participated in the AgBOT challenge in 2016. His team won the competition. They stood before 12 other teams from top engineering schools such as Virginia Tech and Purdue. The theme of the challenge was on the design, development and field trail of a fully-autonomous seeder in a 12-acre test farm in Rockville, Indiana in May 2016. Dr. Mehrandezh has been a member of IEEE since 1995.