Image-Based Visual Relative Information for Distributed Rigid Formation Control in 3D Space

Rosa, M.R.; Berkel, A.; Jayawardhana, B.

Published in:
IEEE Control Systems Letters

DOI:
10.1109/LCSYS.2024.3401025

Publication date:
2024

Image-Based Visual Relative Information for Distributed Rigid Formation Control in 3D Space

M. R. Rosa, *Graduate Student Member, IEEE*, A. Berkel, and B. Jayawardhana, *Senior Member, IEEE*

***Abstract*—Image-based visual relative information (IBVR) for distributed rigid formation control of agents moving in 3D space is proposed for single-integrator agents. The IBVR approach is based solely on the local information of the neighbors' visible area, of neighbors' local coordinates in the image plane, and of the camera parameters in order to achieve and maintain rigid formation distributedly. Each agent is represented as a spherical agent and modeled as a single integrator. We introduce a rigid formation framework based on the image-based visual information and subsequently use it to design the gradient-based distributed formation control in 3D space.**

***Index Terms*—Distributed control, formation control, image-based, visual relative information, 3D workspace.**

I. Introduction

FORMATION-CONTROL theories and methods have attracted many researchers in various research fields, especially in robotics and in autonomous systems. Formation control theory is applicable to a multitude of robotic applications. Some popular applications of formation control methods are autonomous underwater vehicles (AUVs), unmanned aerial vehicles (UAVs), and mobile robots/unmanned ground vehicles (UGV). In these applications, the (distributed) formation control is used to reach and maintain desired formation shapes based on relative information that is measured by the on-board local sensor systems. In literature, various formation control methods have been proposed that depend on the form of relative information and on the location of the sensor systems. We refer interested readers to the review papers in [1], [2], [3].

In contrast to various formation control methods, the distance-based approach can be fully implemented in a distributed way and in a coordinate-free fashion, which has made it popular for the past decade. In the distance-based formation control framework, the desired shape is defined based on inter-agent distance. Correspondingly, each agent uses only relative position defined on its local coordinate systems so that all agents can reach and maintain the desired shape. Due to its distributed nature, it has an inherent robustness such that it can be combined with other complex algorithms, such as, the fully-homomorphic encryption algorithm as in [4].

As an alternative to the distance-based method, bearing-based approach has been studied where the relative bearing of neighboring agents is used to form a 2D/3D formation. However, it requires a common global reference framework or additional measurements such as relative orientation, azimuth angle and altitude/elevation angle to reach and maintain 3D formation [5], [6], [7]. Recently, an internal angle approach was proposed in [8] for 2D formation that uses only local information and local frame of reference. It assumes that all agents are circular agents with uniform radius so that the internal angle of neighboring agents can be used to define rigid formation.

Inspired by the work in [8], we study in this letter the generalization of internal angle approach to the 3D space. It was recently demonstrated that distance measurement of three-dimensional objects (3D-objects) from an image snapshot using a single camera is feasible [9]. Motivated by the approach in [9], we study in this letter the problem of distributed formation control that uses directly image information. It is based on the same mechanism as found in nature, where bees use the image information of visible landmarks that is compared to the ones in their memory [10]. We note that the use of a single camera for distance measurement in a 3D environment is particularly attractive due to its reduced complexity and cost compared to alternative sensors like 3D LIDAR or stereo cameras [11].

The formation control framework using image information has recently been presented in the literature, which is primarily driven by the use of image-based visual servoing (IBVS) approach. In the IBVS approaches, the distributed control uses directly position coordinate of the neighbors via image information to steer the agents towards the desired shape via leader-follower scheme [12], [13], [14]. In this case, the followers ensure that their leaders are steered toward the right image coordinate by adjusting their own position in the physical space. The uncertainty problem arising from unknown leaders' parameters such as velocities

Manuscript received 20 February 2024; revised 17 April 2024; accepted 2 May 2024. Date of publication 15 May 2024; date of current version 29 May 2024. This work was supported by the Indonesia Endowment Fund for Education (LPDP) under Grant 0000201/AUT/D/9/lpdp2022. Recommended by Senior Editor S. Tarbouriech. *(Corresponding author: M. R. Rosa.)*

M. R. Rosa is with the Faculty of Science and Engineering, Engineering and Technology Institute Groningen, University of Groningen, 9747 AG Groningen, The Netherlands, and also with the School of Electrical Engineering, Telkom University, Bandung 40257, Indonesia (e-mail: m.r.r.rosa@rug.nl).

A. Berkel and B. Jayawardhana are with the Faculty of Science and Engineering, Engineering and Technology Institute Groningen, University of Groningen, 9747 AG Groningen, The Netherlands (e-mail: a.m.a.berkel@student.rug.nl; b.jayawardhana@rug.nl).

Digital Object Identifier 10.1109/LCSYS.2024.3401025

or vision-related parameter uncertainties can be solved by introducing adaptive algorithm as discussed in [15], [16]. However, in the IBVS formation control methods, the use of a global position system is still required by the leader and the formation is kept via leader-follower configuration, which does not guarantee the rigidity of the formation shape. To the best of authors' knowledge, IBVS rigid distributed formation control has not been presented in the literature of formation control. In this sense, we do not use the vision systems for servoing purposes as commonly used in the IBVS literature. Correspondingly, in order to emphasize the use of vision systems for obtaining relative information in the distributed control, we introduce the notion of Image-Based Visual Relative (IBVR) information.

Correspondingly, the main contribution of this letter is on the design of IBVR distributed formation control that does not rely on the use of a global positioning system. Firstly, we propose an IBVR distributed control for reaching and maintaining a rigid formation shape in a 3D space. This letter extends naturally the work in [8] to 3D space. In standard geometry, one possible generalization of internal angle from 2D to the 3D case can be based on the use of solid angle. Inspired by the notion of solid angle, we use camera image to obtain the visible area that can be linked to the solid angle. The proposed IBVR approach relies only on the visible area from the image and the direction to the centroid of neighboring agents to form a 3D formation, without requiring a common global reference frame or additional angle measurements. The use of IBVR eliminates the need for direct distance and relative position measurements or inferring this information from the image. By assuming that the neighboring agents are uniform spherical objects and defined as single-integrator agent, our approach hinges on measuring the visible area of agents along with image data and camera parameters and uses this IBVR information to define rigid formation shape and the associated potential function suitable for the gradient-based formation control law. In contrast to the approach presented in [17], which focuses on *distance-based* formation control, here, we introduce a novel *image-based* relative information method that can be obtained via simple camera systems.

This letter is structured as follows. In Section II, we present the preliminaries on rigid formation framework that is commonly adopted in distance-based formation control. In Section III, we discuss the camera model and present the relationship between the visible area of neighboring agents to the distance. Accordingly, we present an image-based rigid formation framework that can be used to construct IBVR distributed formation control in Section IV. For showing the efficacy of the approach, we present the numerical simulations in Section V, and the conclusions in Section VI.

II. Preliminaries

In this letter we use the following standard notations. The identity matrix of dimension n is denoted as $\boldsymbol{I_n}$. The Kronecker product is denoted by $\otimes$. Given a square matrix $\boldsymbol{R}$, we define $\bar{\boldsymbol{R}} \coloneqq \boldsymbol{R} \otimes \boldsymbol{I_n}$. For a set of vectors $\boldsymbol{x_i} \in \mathbb{R}^n$, where $i \in \{1, \dots, k\}$, we define the concatenated vector $\boldsymbol{x} \in \mathbb{R}^{kn}$ by $\boldsymbol{x} \triangleq \begin{bmatrix} \boldsymbol{x_1}^\top & \boldsymbol{x_2}^\top & \dots & \boldsymbol{x_k}^\top \end{bmatrix}^\top$, the corresponding block diagonal matrix by $\boldsymbol{D_x} = \operatorname{diag}(\boldsymbol{x_i})_{i\in\{1,\dots,k\}} \in \mathbb{R}^{km\times kn}$ and its 2-norm by $\|x_i\|$. We also define $\operatorname{col}(\cdot)$ as a concatenated column vector operator. The space $\mathscr{L}_2(\mathbb{R}_+)$ is the space of continuous-time functions/signals $\boldsymbol{x} : \mathbb{R}_+ \to \mathbb{R}^n$, which are square-integrable, i.e., $\int_0^\infty \|\boldsymbol{x}(t)\|^2 \mathrm{d}t < \infty$. Similarly, the space $\mathscr{L}_\infty(\mathbb{R}_+)$ denotes the space of functions $\boldsymbol{x} : \mathbb{R}_+ \to \mathbb{R}^n$ that are essentially bounded, i.e., $\operatorname{ess-sup}_{t\in\mathbb{R}_+} \|\boldsymbol{x}(t)\| < \infty$.

Consider a graph of agents given by $\mathbb{G} = (\mathscr{V}, \mathscr{E})$, where $\mathscr{V} = \{1, \dots, N\}$ is the set of N agents, and $\mathscr{E} \subseteq \mathscr{V} \times \mathscr{V}$ is set of edges describing pairs of interacting agents. We define the set of neighbors of agent i by $\mathscr{N}_i \coloneqq \{j \in \mathscr{V} | i, j \in \mathscr{E}\}$. For every pair of interacting agents in the edge set $\mathscr{E}$, the relative information between them can be described by an incidence matrix $\boldsymbol{B} \in \mathbb{R}^{|\mathscr{V}|\times|\mathscr{E}|}$, where the k-th column vector is associated to the k-th edge $_k$ and its elements are given by

$$b_{ik} = \begin{cases} -1, & \text{if } i = \mathscr{E}_k^{\text{head}} \\ +1, & \text{if } i = \mathscr{E}_k^{\text{tail}} \\ 0, & \text{otherwise.} \end{cases}$$

Associated to each node in the graph $\mathbb{G}$, we can assign the agent's position $\boldsymbol{q_i} \in \mathbb{R}^3$. The tuple of positions $\boldsymbol{q} = \begin{bmatrix} \boldsymbol{q_1}^\top & \boldsymbol{q_2}^\top & \cdots & \boldsymbol{q_N}^\top \end{bmatrix}^\top$ and the graph $\mathbb{G}$ defines the *framework*. Using the incidence matrix $\boldsymbol{B}$, we can define the relative position $\boldsymbol{z}$ between interacting agents by $\boldsymbol{z} = \bar{\boldsymbol{B}}^\top \boldsymbol{q}$, where $\bar{\boldsymbol{B}} = \boldsymbol{B} \otimes \boldsymbol{I_n}$. For defining a geometrical shape and its rigidity property, a scalar positive-definite function $\boldsymbol{g}$ can be defined for every edge in the framework that can be related to the geometry. For instance, one can consider the use of Euclidean distance based on the position of interacting agents $g(\boldsymbol{q_i}, \boldsymbol{q_j}) = \|\boldsymbol{q_i} - \boldsymbol{q_j}\| = \|\boldsymbol{z_k}\|$, where $\mathscr{E}_k = (i, j)$, as commonly used in the distance-based formation control. In the literature of distance-based formation control, it is common to consider

$$g(\boldsymbol{q_i}, \boldsymbol{q_j}) = \|\boldsymbol{z_k}\|^\ell, \tag{1}$$

with $\ell \geq 1$. Another commonly used function is the relative bearing between interacting agents $g(\boldsymbol{q_i}, \boldsymbol{q_j}) = \angle \boldsymbol{z_k}$.

Correspondingly, an edge function $\boldsymbol{q} \mapsto \boldsymbol{f}(\boldsymbol{q}) \in \mathbb{R}_+^{|\mathscr{E}|}$ can be defined that collects all scalar measures in each edge. In our main results, we will later propose an image-based edge function where g will be defined based on the apparent image size of the neighboring agent.

Let us revisit the concept of an infinitesimally rigid framework, which has been studied extensively in prior works, such as [18], [19], [20]. The infinitesimal rigidity of a given geometrical shape defined by $\boldsymbol{f}(\boldsymbol{q})^* = \boldsymbol{d}^*$, where $\boldsymbol{d}^*$ is the desired edge quantities, which can be characterized by the null space of the rigidity matrix defined by $\boldsymbol{R}(\boldsymbol{q}) = \frac{\partial \boldsymbol{f}(\boldsymbol{q})}{\partial \boldsymbol{q}}$. For instance, using the distance-based edge function as in (1), which can be rewritten as

$$\boldsymbol{f}_{\text{dist}}(\boldsymbol{q}) = \begin{bmatrix} \|\boldsymbol{z_1}\|^\ell & \dots & \|\boldsymbol{z}_{|\mathscr{E}|}\|^\ell \end{bmatrix}^\top = \bar{\boldsymbol{D}}_{\tilde{\boldsymbol{z}}} \boldsymbol{D}_{\boldsymbol{z}}^\top \boldsymbol{z}, \tag{2}$$

where $\tilde{\boldsymbol{z}} = \begin{bmatrix} \|\boldsymbol{z_1}\|^{\ell-2} & \dots & \|\boldsymbol{z}_{|\mathscr{E}|}\|^{\ell-2} \end{bmatrix}^\top$, the distance-based rigidity matrix is given by

$$\boldsymbol{R}_{\text{dist}}(\boldsymbol{q}) = \frac{\partial \boldsymbol{f}_{\text{dist}}(\boldsymbol{q})}{\partial \boldsymbol{q}} = \bar{\boldsymbol{D}}_{\tilde{\boldsymbol{z}}} \boldsymbol{D}_{\boldsymbol{z}}^\top \bar{\boldsymbol{B}}^\top.$$

In this case, by denoting $\boldsymbol{d}^*$ as the desired distance vector in every edge, the set of all positions with the desired shape is given by

$$\mathscr{D} \coloneqq \{\boldsymbol{q}^* \,|\, \boldsymbol{f}_{\text{dist}}(\boldsymbol{q}^*) = \boldsymbol{d}^*\}.$$

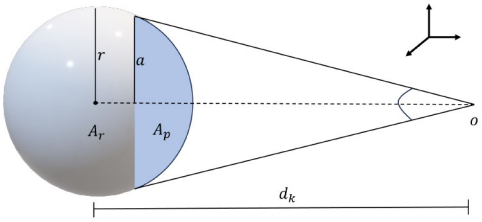

Fig. 1. The black area illustrates the visible area of the spherical neighboring agent as seen by an observer at o. The parameters r, d_k, a, A_r and A_p are the radius, distance to the centroid, apparent radius, total surface area and visible area, respectively.

The corresponding group motion that makes $\mathscr{D}$ invariant can be established by studying the infinitesimal displacement $\delta \boldsymbol{q}$ such that $\boldsymbol{R}(\boldsymbol{q})\delta \boldsymbol{q} = 0$. For the distance-based edge function, the admissible infinitesimal displacement includes translations and rotations of the entire framework and it is related also to the rank of the distance rigidity matrix by $\text{rank}(\boldsymbol{R}_{\text{dist}}(\boldsymbol{q})) = n|\mathscr{V}| - \frac{n(n+1)}{2}$, where n is the dimension of the working space.

III. IBVR Rigid Framework

In this section, we present the use of IBVR to define a rigid formation framework. As described in the Introduction, IBVR can be used as a proxy to measure the distance and relative position of the neighboring agents. The proposed technique relies on camera parameters and only measures objects with a specific shape, which will be a uniform sphere in this letter. The main motivation of using sphere is due to its simple mathematical representation, which facilitates the design of IBVR-based methods and the corresponding analysis. We note that this assumption is satisfied by a number of recent design of unmanned aerial vehicles [21], [22], [23] or those encapsulated in a sphere cage [24], [25], [26].

Let us consider a tetrahedron-shaped formation consisting of four spherical agents working in 3D-space, Suppose that each agent is capable of visual observation of other neighboring agents in 3D space via camera. In this case, the image-based visual information of the neighboring agents is given by the area of spherical caps as shown in Figure 1. The plane curve (conic) of the sphere is a circle, obtained by intersecting the line with the image plane, and its apparent sizes can be related to distance information [27], [28].

Let us now derive the relation between the visible area and the distance to the neighboring agents. Figure 1 shows an illustration of these variables where a neighboring spherical agent j is observed from the centroid of an agent i (indicated by o) with $(i, j) \in \mathcal{E}_k$. In this figure, the spherical cap visible area A_p is part of the total surface area A_r [29]. The visible circle that encapsulates A_p has a radius of a, which is smaller than the sphere's radius of r. As before, the distance between two agents is given by $d_k = \|\boldsymbol{z}_k\|$. The ratio between the spherical cap visible area and the sphere's surface area satisfies

$$A_{\text{ratio}} = \frac{A_p}{A_r} = \frac{d_k - r}{2d_k}. \tag{3}$$

Now, let us discuss the relationship between the size of the visible area of an agent and the distance at which that area can be seen. To illustrate this relationship, we will employ

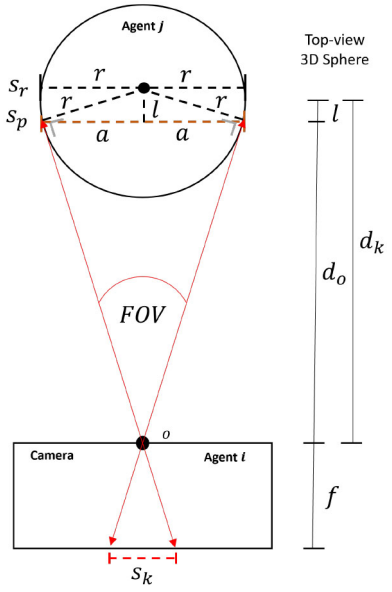

Fig. 2. The top-view plot of a pinhole camera system that projects a spherical object to the image plane.

principles of geometry and the characteristics of a pinhole camera. Figure 2 shows the top view of pinhole camera system, where o is the optical center (pinhole), f is focal length measured in pixels, d_o is the distance from pinhole to the surface of the object, l is the distance between the visible part of the sphere and its center, s_r is an area of a circle with radius r, s_p is the area of a circle with radius a and s_k is the circular area of the neighbors j on the image plane of agent i. Based on this pinhole camera model, we can estimate object distances by using the pixel area measurements as proposed in [9], [30]. The relationship between the distance of an agent to its neighbors and the measured visible area in 3D working space is given by

$$\frac{f^2}{d_o^2} = \frac{s_k}{s_p}. \tag{4}$$

For a sufficiently large inter-agent distance $d_k \gg r$, we have that $a \approx r$, $A_p \approx A_r$ and $s_p \approx s_r$. For instance, when $d_k > 10r$, and the real distance defined as $\hat{d}_k$, it holds that $A_p > 0.45A_r$, $s_p > 0.90s_r$, and $d_k > 0.948\hat{d}_k$. For the rest of this letter, we will assume that $d_k \gg r$ so that the following approximation holds. In this case, the relationship between the visible area and its projection satisfies the pinhole camera equation given by $\frac{f^2}{d_o^2} \approx \frac{f^2}{d_k^2}$ and $\frac{s_k}{s_p} \approx \frac{s_k}{s_r}$, or in other words, following (4),

$$s_k \approx \frac{s_r f^2}{d_k^2}. \tag{5}$$

This approximation allows us to measure the distance d_k using the image-based visual relative information (IBVR) of s_k. The large distance condition of $d_k \gg r$ means also that the spherical agents are not in the collision range.

In the following, we will establish the relation between the IBVR variable s_k and the relative position z_k as illustrated in Figure 3. First, note that the relative position between agents i and j on the camera coordinate plane can be obtained using

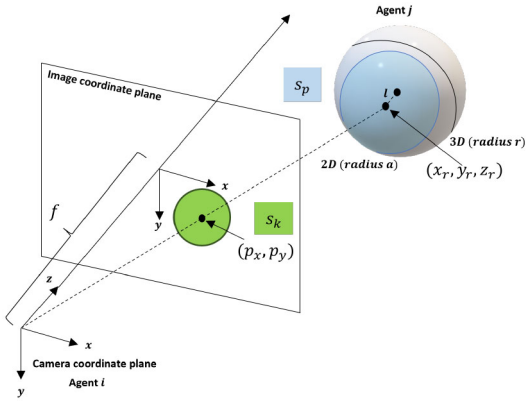

Fig. 3. An illustration of a pinhole camera model depicting the relative position of agent j compared to the camera coordinate plane of agent i.

the pinhole camera model [31]. We can obtain the relationship between $\boldsymbol{z_k}$ and the camera coordinate plane p_x, p_y by using the following equations

$$\boldsymbol{z_k} = \begin{bmatrix} \frac{z_r}{f} p_x & \frac{z_r}{f} p_y & z_r \end{bmatrix}^\top, \tag{6}$$

where $z_r = \sqrt{\frac{s_r}{s_k}} f$. Let us discuss the use of s_k to define rigid formation shape. Using s_k, we define a *visible-area-based edge function* $\boldsymbol{f}_{\mathbf{area}}(s)$ by

$$\boldsymbol{f}_{\mathbf{area}}(\boldsymbol{q}) = \begin{bmatrix} \dots & s_k & \dots \end{bmatrix}^\top, \ \forall k \in \{1, \dots, |\mathscr{E}|\}.$$

Based on (5), it can also be expressed as

$$\boldsymbol{f}_{\mathbf{area}}(\boldsymbol{q}) = \begin{bmatrix} \dots & \frac{s_r f^2}{d_k^2} & \dots \end{bmatrix}^\top, \ \forall k \in \{1, \dots, |\mathscr{E}|\}. \tag{7}$$

For describing the *visible area rigidity* matrix $\boldsymbol{R}_{\mathbf{area}}(\boldsymbol{q})$, one can take the Jacobian of (7) as follows

$$\begin{aligned} \boldsymbol{R}_{\mathbf{area}}(\boldsymbol{q}) &= \frac{\partial \boldsymbol{f}_{\mathbf{area}}(\boldsymbol{q})}{\partial \boldsymbol{q}} = \frac{\partial \boldsymbol{f}_{\mathbf{area}}(\boldsymbol{d_k})}{\partial \boldsymbol{Q}} \frac{\partial \boldsymbol{Q}}{\partial \boldsymbol{q}} \\ &= -\boldsymbol{T}(\boldsymbol{d}) \boldsymbol{R}_{\mathbf{dist}}(\boldsymbol{q}) = -\boldsymbol{T}(\boldsymbol{d}) \bar{\boldsymbol{D}}_{\bar{\boldsymbol{z}}} \boldsymbol{D}_z^\top \bar{\boldsymbol{B}}^\top, \end{aligned} \tag{8}$$

where $\boldsymbol{Q} = \boldsymbol{D_d} \in \mathbb{R}^{|\mathscr{E}|}$, and $\boldsymbol{T}(\boldsymbol{d}) = \operatorname{diag}(s_r f^2 (d_k)^{-4}_{k \in \mathscr{E}_k})$. The matrix $\boldsymbol{T}(\boldsymbol{d})$ is positive definite as each $d_k \gg 0$. Thus, we have rank $(\boldsymbol{R}_{\mathbf{area}})$ = rank $(\boldsymbol{R}_{\mathbf{dist}})$. It shows a one-to-one relationship between infinitesimal-rigid formation shape that is defined using the distance-based edge function $\boldsymbol{f}_{\mathbf{dist}}$ as in (2) with $\ell = 1$ and using the visible-area-based edge function $\boldsymbol{f}_{\mathbf{area}}$. In particular, if we use visible-area vector $\boldsymbol{s}^* = \operatorname{col}_{k \in \mathscr{E}_k}(s_k^*)$ to define a formation shape of $(\mathbb{G}, \boldsymbol{q}^*)$, the formation is *infinitesimally rigid* if $\operatorname{rank}(\boldsymbol{R}_{\mathbf{area}}(\boldsymbol{q}^*)) = 3|\mathscr{V}| - 6$. Similar as before, we define the set of all positions with the desired shape using $\boldsymbol{f}_{\mathbf{area}}$ by

$$\mathscr{D}_{\text{area}} := \{\boldsymbol{q}^* | \boldsymbol{f}_{\mathbf{area}}(\boldsymbol{q}^*) = \boldsymbol{s}^*\}.$$

We can now specify the desired target formation shape using a framework denoted as $(\mathbb{G}, \boldsymbol{q}^*)$. Here, the vector $\boldsymbol{q}^* \in \mathbb{R}^3$ adheres to a set of desired visible areas $\boldsymbol{s}^*$. One approach to determine these distances is by employing (5) when the visible area constraints are measured. This approach utilizes directly the image measurement to reach and maintain a rigid formation. Throughout the rest of this letter, we consider a set of desired visible area $\{s_k^*\}$, denoted by $\boldsymbol{s}^*$, such that the corresponding visible area rigidity matrix $\boldsymbol{R}_{\mathbf{area}}$ has rank $3|\mathscr{V}| - 6$, i.e., the formation shape $(\mathbb{G}, \boldsymbol{q}^*)$ is infinitesimally rigid.

IV. IBVR Distributed Formation Control

Consider single-integrator agents given by

$$\dot{\boldsymbol{q}}_{\boldsymbol{i}} = \boldsymbol{u}_{\boldsymbol{i}}, \qquad \forall i = 1, \dots, N, \tag{9}$$

where $\boldsymbol{q_i} \in \mathbb{R}^3$ is the position and $\boldsymbol{u_i} \in \mathbb{R}^3$ is the velocity input. In the following, we formulate a gradient-based controller utilizing a *visible-area-based* potential function

$$e_k = s_k^* - s_k, \tag{10}$$

where s_k^* is the desired visible-area of the neighbor j on the image plane of agent i. Recall that the set of desired visible-area $\{s_k^*\}$ is such that the corresponding formation shape is infinitesimally rigid. Let us consider the following potential function

$$V(\boldsymbol{e}) = \frac{1}{2} \sum_{k=1}^{|\mathscr{E}|} \kappa_k \left(s_k^* - s_k\right)^2, \tag{11}$$

where $\boldsymbol{e} = \operatorname{col}_{k \in \mathscr{E}_k}(e_k)$, κ_k is a positive constant for every $k \in \{1, \dots, |\mathscr{E}|\}$. Following the standard gradient-based control law, the distributed control law for every agents i is given by

$$\boldsymbol{u_i} = -\sum_{k=1}^{|\mathscr{E}_k|} \frac{\partial}{\partial q_k} V(\boldsymbol{e}).$$

By direct substitution of (11) into the above gradient-based control law, we have

$$\boldsymbol{u_i} = -\sum_{j \in \mathscr{N}_i} \frac{\boldsymbol{z_{ij}}}{\|\boldsymbol{z_{ij}}\| s_r f^2} s_{ij}^2 \kappa_{ij} e_{ij}, \tag{12}$$

where $e_{ij} = e_k$, $\kappa_{ij} = \kappa_k$, $\boldsymbol{z_{ij}} = \boldsymbol{z_k}$, and $s_{ij} = s_k$ with $(i, j) = \mathscr{E}_k$. In addition to the directional unit vector towards the centroid of neighboring agent j given by $\frac{\boldsymbol{z_{ij}}}{\bar{s}_{ij}}$, the above distributed control law uses only visible-area information of the neighbor to reach the desired formation shape.

Proposition 1: Consider a set of single-integrator agents (9) and assume an infinitesimally rigid framework $(\mathbb{G}, \boldsymbol{q}^*)$ with the set of desired visible area $\boldsymbol{s}^*$ describing the desired formation shape. Then the following distributed control law

$$\boldsymbol{u} = -\bar{\boldsymbol{B}} \boldsymbol{D}_{\boldsymbol{z}} \bar{\boldsymbol{D}}_{\boldsymbol{s}}^\top \boldsymbol{T}(\boldsymbol{s}) \boldsymbol{D}_{\boldsymbol{\kappa}} \boldsymbol{e}, \tag{13}$$

solves the problem of image-based visual relative information distributed formation control of single-integrator agents locally and exponentially. Particularly, for all initial conditions $q(0)$ in the neighborhood of $\mathscr{D}_{\text{area}}$, the visible area error e_k exponentially converges to zero for all $k \in 1, \dots, |\mathscr{E}|$, ensuring that all agents' positions $\boldsymbol{q_i}(t)$ are bounded and exponentially converge to the desired formation shape.

Proof: First, we will show the asymptotic convergence of error e_k for all $k \in \{1, \dots, |\mathscr{E}|\}$ to zero using the potential function

$$V(\boldsymbol{e}) = \frac{1}{2} \sum_{k=1}^{|\mathscr{E}|} \kappa_k e_k^2. \tag{14}$$

It is known that the gradient of the potential function along $\boldsymbol{q}$ satisfies

$$\frac{\partial V(\boldsymbol{e})}{\partial \boldsymbol{q}} = \sum_{k=1}^{|\mathscr{E}|} \frac{\partial \boldsymbol{P}(e_k)}{\partial e_k} \frac{\partial e_k}{\partial \boldsymbol{Q}} \frac{\partial \boldsymbol{Q}}{\partial \boldsymbol{q}}, \tag{15}$$

where $\boldsymbol{P}(e_k) = \frac{1}{2}\kappa_k(s_k^* - s_k)^2$ is the potential function for each edge. For each term in (15), we have the following computation

$$\begin{aligned} \frac{\partial \boldsymbol{e_k}}{\partial \boldsymbol{Q}} &= \operatorname{diag}\left(s_r f^2 (d_k)^{-4}_{k\in\mathscr{E}_k}\right) := \boldsymbol{T}(\boldsymbol{d}), \\ \frac{\partial \boldsymbol{Q}}{\partial \boldsymbol{q}} &= \bar{\boldsymbol{D}}_{\tilde{\boldsymbol{z}}} \boldsymbol{D}_z^\top \bar{\boldsymbol{B}}^\top := \boldsymbol{R}_{\mathbf{dist}}, \quad \text{and} \quad \frac{\partial \boldsymbol{P}(e_k)}{\partial e_k} = \kappa_k e_k. \end{aligned} \tag{16}$$

Correspondingly, (15) is given by $\frac{\partial V(\boldsymbol{e})}{\partial \boldsymbol{q}} = \bar{\boldsymbol{B}}\boldsymbol{D}_z\bar{\boldsymbol{D}}_{\tilde{\boldsymbol{z}}}^\top \boldsymbol{T}(\boldsymbol{d})\boldsymbol{D}_\kappa \boldsymbol{e}$. Note that as we only have image information of the neighbors, we need to rewrite $\boldsymbol{d}$ in this equation as a function of image information using (5). By substituting (5) into $\boldsymbol{T}(\boldsymbol{d})$, we can define $\boldsymbol{T}(\boldsymbol{d})$ as a function of visual relative information $\boldsymbol{T}(\boldsymbol{s}) = \operatorname{diag}(\{s_k^2(s_r f^2)^{-1}\}_{k\in\mathscr{E}_k})$ which is a positive definite matrix. In addition, using (5) one can define $\tilde{\boldsymbol{z}}$, given after (2) and using image measurement in (6), as a function of visual relative information, $\tilde{\boldsymbol{s}} = \operatorname{col}_{k\in\mathscr{E}_k}(\frac{1}{\|\boldsymbol{z_k}\|})$. Then one can use (8) to get $\boldsymbol{R}_{\mathbf{area}}$ from $\boldsymbol{R}_{\mathbf{dist}}$, and the expression of $\frac{\partial V(\boldsymbol{e})}{\partial \boldsymbol{q}}$ becomes

$$\frac{\partial V(\boldsymbol{e})}{\partial \boldsymbol{q}} = \bar{\boldsymbol{B}}\boldsymbol{D}_z\bar{\boldsymbol{D}}_{\tilde{\boldsymbol{s}}}^\top \boldsymbol{T}(\boldsymbol{s})\boldsymbol{D}_\kappa \boldsymbol{e} = -\boldsymbol{R}_{\mathbf{area}}^\top \boldsymbol{D}_\kappa \boldsymbol{e}. \tag{17}$$

By substituting the distributed control laws in its compact form as in (17) into the dynamics (9), we have $\dot{\boldsymbol{q}} = \boldsymbol{R}_{\mathbf{area}}^\top \boldsymbol{D}_\kappa \boldsymbol{e}$. The time derivative of visible area error satisfies

$$\dot{\boldsymbol{e}} = \frac{\partial \boldsymbol{e}}{\partial \boldsymbol{Q}} \frac{\partial \boldsymbol{Q}}{\partial \boldsymbol{q}} \dot{\boldsymbol{q}} = \boldsymbol{T}(\boldsymbol{d})\boldsymbol{R}_{\mathbf{dist}}\dot{\boldsymbol{q}} = -\boldsymbol{R}_{\mathbf{area}}\boldsymbol{R}_{\mathbf{area}}^\top \boldsymbol{D}_\kappa \boldsymbol{e}. \tag{18}$$

By calculating the time derivative of (14) along the trajectories of visible area error equation (18), we have

$$\dot{V}(\boldsymbol{e}) = \boldsymbol{e}^\top \boldsymbol{D}_\kappa \dot{\boldsymbol{e}} = -\boldsymbol{e}^\top \boldsymbol{D}_\kappa \boldsymbol{R}_{\mathbf{area}}\boldsymbol{R}_{\mathbf{area}}^\top \boldsymbol{D}_\kappa \boldsymbol{e}. \tag{19}$$

As $\boldsymbol{R}_{\mathbf{dist}}$ has full row rank at $\boldsymbol{z} \in \mathscr{D}$, it follows that $\boldsymbol{R}_{\mathbf{dist}}(\boldsymbol{q}^*)\boldsymbol{R}_{\mathbf{dist}}(\boldsymbol{q}^*)^\top$ is a positive definite matrix. As consequence, the matrix $\boldsymbol{R}_{\mathbf{area}}(\boldsymbol{q}^*)\boldsymbol{R}_{\mathbf{area}}(\boldsymbol{q}^*)^\top$ is a positive definite matrix. Correspondingly we can rewrite (19) into $\dot{V}(\boldsymbol{e}) = -\boldsymbol{e}^\top \mathbb{Q}\boldsymbol{e} \leq -\lambda_{\min}\|\boldsymbol{e}\|^2$, where $\mathbb{Q} = \boldsymbol{D}_\kappa \boldsymbol{R}_{\mathbf{area}}\boldsymbol{R}_{\mathbf{area}}^\top \boldsymbol{D}_\kappa$ and $\lambda_{\min}$ represents the smallest eigenvalue of matrix $\mathbb{Q}$. Using the standard comparison lemma and considering that V is quadratic in $\boldsymbol{e}$ as shown in (14), we can conclude that $\boldsymbol{e}(t)$ is bounded and converges exponentially to zero. It follows that $|\boldsymbol{e}(t)| < \frac{1}{\kappa_{k,\min}}\sqrt{V(0)}$ for all t, i.e., $\boldsymbol{e} \in \mathscr{L}_\infty$ and $s_k \in \mathscr{L}_\infty$ from (10). Correspondingly, the distributed input $\boldsymbol{u_i}$ as in (12) is bounded and converges to zero exponentially. Using (9), it follows that the agents' position is bounded for all time. This concludes the proof. ■

V. Numerical Simulations

We consider the distributed formation of tetrahedron via our proposed method. For the simulation setup, we define initial positions for each agent as follows: Agent 1 $\{0, 0, 0\}$, Agent 2 $\{3, 5, -1\}$, Agent 3 $\{6, 2, 1\}$, and Agent 4 $\{3, 1, 5\}$. The radius of each agent is set to $r = 0.5$ meters. We consider a camera with sensor resolution 4608×2592 pixels, pixel size $1.4\mu\text{m} \times 1.4\mu\text{m}$, and the focal length is set to $f = 2.75$ mm. The incidence matrix $\boldsymbol{B}$ is set as follows

$$\boldsymbol{B} = \begin{bmatrix} 1 & -1 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}, \tag{20}$$

which corresponds to all-to-all connectivity. We consider a tetrahedron formation shape where for each edge k in the set $1, \ldots, |\mathscr{E}|$, we set the desired visible area $s_k^* = 1.2121 \times 10^5$ pixel^2 $= 2.3758 \times 10^{-7}$ m^2. This desired area corresponds to the desired distance of $d_k^* = 5$m according to the relation in (5). For the numerical simulation of the virtual spring constant κ_k, we use the following values

$$\begin{aligned} &\kappa_1 = 9 \times 10^{13}, \kappa_2 = 8 \times 10^{13}, \kappa_3 = 7 \times 10^{13}, \\ &\kappa_4 = 6 \times 10^{13}, \kappa_5 = 5 \times 10^{13}, \kappa_6 = 4 \times 10^{13}. \end{aligned} \tag{21}$$

The varying gains indicate that the proposed approach can be implemented for heterogeneous agents, given the well-posed nature of the problem. It can be seen in Figure 4 that every single-integrator agent converges to the desired tetrahedron shape in the 3D space. The initial positions of the agents are represented by diamonds, while their final positions are represented by pentagrams, as depicted by the markers in Figure 4. Figure 5 shows the convergence of the visible area error to zero for each edge.

To analyze the region-of-attraction and its robustness, we run a Monte Carlo simulation where 1000 samples of initial position are taken in the neighborhood of the desired shape $\mathscr{D}_{\text{area}}$, as illustrated in Figure 6. Figure 6 reveals that when initial positions are close to the desired shape (within 95% of the normal distribution), the visible area error converges exponentially, as depicted in cyan. However, when the initial positions nearly cause the agents to align, the visible area

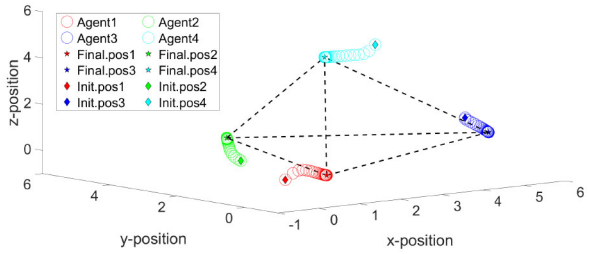

Fig. 4. The simulation result of an IBVR distributed formation control where four agents must form a tetrahedron formation shape. The circles show the trajectories of the agents towards the desired 3D shape.

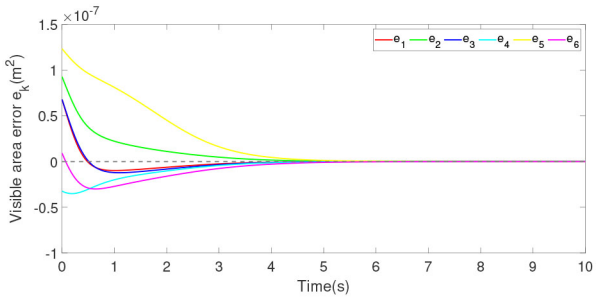

Fig. 5. The plot of visible area error e_k for the tetrahedron formation for all edges $k = 1, \ldots, 6$.

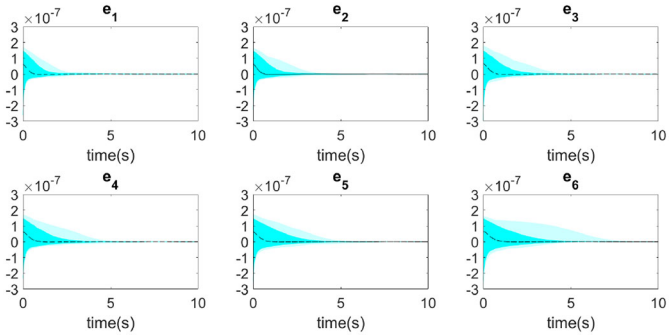

Fig. 6. Monte Carlo simulation results of all visible area error e_k based on 1000 random samples of initial positions. The cyan area shows the error trajectories of all 95% of the samples and the light cyan area shows the error trajectories of all samples.

error does not converge exponentially, as shown in light cyan. This conforms with our theoretical analysis that our proposed controller ensures the exponential convergence of the visible area error e_k within the neighborhood of the desired shape.

VI. Conclusion

We present the design of the IBVR distributed rigid formation controller for single-integrator agents. Particularly, the proposed distributed controller only uses local image measurement without any global information. For future work, we will study the use of physics-based models such as port-Hamiltonian systems [17] to closely represent the agent to the reality and develop an adaptive algorithm to improve the robustness w.r.t visual relative information inconsistency [32].

References

[1] K.-K. Oh, M.-C. Park, and H.-S. Ahn, "A survey of multi-agent formation control," *Automatica*, vol. 53, pp. 424–440, Mar. 2015.

[2] B. D. O. Anderson, B. Fidan, C. Yu, and D. Walle, "UAV formation control: Theory and application," in *Recent Advances in Learning and Control*, V. D. Blondel, S. P. Boyd, and H. Kimura, Eds. London, U.K.: Springer, 2008, pp. 15–33.

[3] M. H. Trinh, S. Zhao, Z. Sun, D. Zelazo, B. D. O. Anderson, and H.-S. Ahn, "Bearing-based formation control of a group of agents with leader-first follower structure," *IEEE Trans. Autom. Control*, vol. 64, no. 2, pp. 598–613, Feb. 2019.

[4] M. Marcantoni, B. Jayawardhana, M. P. Chaher, and K. Bunte, "Secure formation control via edge computing enabled by fully homomorphic encryption and mixed uniform-logarithmic quantization," *IEEE Control Syst. Lett.*, vol. 7, pp. 395–400, 2023.

[5] S. Zhao and D. Zelazo, "Bearing rigidity theory and its applications for control and estimation of network systems: Life beyond distance rigidity," *IEEE Control Syst. Mag.*, vol. 39, no. 2, pp. 66–83, Apr. 2019.

[6] G. Y. Xiaolei Li, M. J. Er, G. Yang, and N. Wang, "Bearing-based formation manoeuvre control of nonholonomic multi-agent systems," *Int. J. Syst. Sci.*, vol. 50, no. 16, pp. 2993–3002, 2019.

[7] L. Chen and Z. Sun, "Gradient-based bearing-only formation control: An elevation angle approach," *Automatica*, vol. 141, Jul. 2022, Art. no. 110310.

[8] N. P. K. Chan, B. Jayawardhana, and H. G. de Marina, "Angle-constrained formation control for circular mobile robots," *IEEE Control Syst. Lett.*, vol. 5, pp. 109–114, 2021.

[9] S. Leorna, T. Brinkman, and T. Fullman, "Estimating animal size or distance in camera trap images: Photogrammetry using the pinhole camera model," *Method. Ecol. Evol.*, vol. 13, no. 8, pp. 1707–1718, 2022.

[10] B. A. Cartwright and T. S. Collett, "Landmark maps for honeybees," *Biolog. Cybern.*, vol. 57, pp. 85–93, Aug. 1987.

[11] K. McGuire, G. de Croon, C. De Wagter, K. Tuyls, and H. Kappen, "Efficient optical flow and stereo vision for velocity estimation and obstacle avoidance on an autonomous pocket drone," *IEEE Robot. Autom. Lett.*, vol. 2, no. 2, pp. 1070–1076, Apr. 2017.

[12] Z. Miao, H. Zhong, J. Lin, Y. Wang, Y. Chen, and R. Fierro, "Vision-based formation control of mobile robots with FOV constraints and unknown feature depth," *IEEE Trans. Control Syst. Technol.*, vol. 29, no. 5, pp. 2231–2238, Sep. 2021.

[13] X. Liang, H. Wang, Y.-H. Liu, W. Chen, and T. Liu, "Formation control of nonholonomic mobile robots without position and velocity measurements," *IEEE Trans. Robot.*, vol. 34, no. 2, pp. 434–446, Apr. 2018.

[14] K. Yao et al., "Image-Based Visual Servoing switchable leader-follower control of heterogeneous multi-agent underwater robot system," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2023, pp. 5200–5206.

[15] J. Lin, Z. Miao, H. Zhong, W. Peng, Y. Wang, and F. Rafael, "Adaptive image-based leader–follower formation control of mobile robots with visibility constraints," *IEEE Trans. Ind. Electron.*, vol. 68, no. 7, pp. 6010–6019, Jul. 2021.

[16] X. Chen and Y. Jia, "Adaptive leader-follower formation control of non-holonomic mobile robots using active vision," *IET Control Theory Appl.*, vol. 9, pp. 1302–1311, May 2015.

[17] M. R. Rosa and B. Jayawardhana, "Distributed adaptive formation control for uncertain point mass agents with mixed dimensional space," *IEEE Control Syst. Lett.*, vol. 7, pp. 2725–2730, 2023.

[18] B. Anderson, C. Yu, B. Fidan, and J. Hendrickx, "Rigid graph control architectures for autonomous formations," *IEEE Control Syst. Mag.*, vol. 28, no. 6, pp. 48–63, Jan. 2009.

[19] L. Krick, M. E. Broucke, and B. A. Francis, "Stabilization of infinitesimally rigid formations of multi-robot networks," in *Proc. 47th IEEE Conf. Decis. Control*, 2008, pp. 477–482.

[20] R. Olfati-Saber and R. M. Murray, "Distributed cooperative control of multiple vehicle formations using structural potential functions," *IFAC Proc. Vol.*, vol. 35, no. 1, pp. 495–500, 2002.

[21] S. Mitani et al., "Int-ball: Crew-supportive autonomous mobile camera robot on ISS/JEM," in *Proc. IEEE Aerosp. Conf.*, 2019, pp. 1–15.

[22] J. Yoo, I.-W. Park, V. To, J. Q. H. Lum, and T. Smith, "Avionics and perching systems of free-flying robots for the international space station," in *Proc. IEEE Int. Symp. Syst. Eng. (ISSE)*, 2015, pp. 198–201.

[23] M. G. Bualat, T. Smith, E. E. Smith, T. Fong, and D. Wheeler, "Astrobee: A new tool for ISS operations," in *Proc. SpaceOps Conf.*, 2018, pp. 1–11. [Online]. Available: https://arc.aiaa.org/doi/abs/10.2514/6.2018-2517

[24] J. Shahmoradi, E. Talebi, P. Roghanchi, and M. Hassanalian, "A comprehensive review of applications of drone technology in the mining industry," *Drones*, vol. 4, no. 3, p. 34, 2020.

[25] K. Edgerton, G. Throneberry, A. Takeshita, C. Hocut, F. Shu, and A. Abdelkefi, "Numerical and experimental comparative performance analysis of emerging spherical-caged drones," *Aerosp. Sci. Technol.*, vol. 95, Dec. 2019, Art. no. 105512.

[26] W. Yamada, K. Yamada, H. Manabe, and D. Ikeda, "iSphere: Self-luminous spherical drone display," in *Proc. 30th Annu. ACM Symp. User Interface Softw. Technol.*, 2017, pp. 635–643.

[27] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[28] R. Cipolla and P. J. Giblin, "The visual motion of curves and surfaces," *Philosoph. Trans. Roy. Soc. London*, vol. 356, pp. 1103–1121, May 1998.

[29] J. Harris and H. Stoecker, *The Handbook of Mathematics and Computational Science*. New York, NY, USA: Springer, 1998.

[30] Y.-T. Cao, J.-M. Wang, Y.-K. Sun, and X.-J. Duan, "Circle marker based distance measurement using a single camera," *Lect. Notes Softw. Eng.*, vol. 1, no. 4, pp. 376–380, 2013.

[31] E. Kelasidi, S. Moe, K. Y. Pettersen, A. M. Kohl, P. Liljebäck, and J. T. Gravdahl, "Path following, obstacle detection and obstacle avoidance for thrusted underwater snake robots," *Front. Robot. AI*, vol. 6, p. 57, Jul. 2019.

[32] H. Garcia de Marina, M. Cao, and B. Jayawardhana, "Controlling rigid formations of mobile agents under inconsistent measurements," *IEEE Trans. Robot.*, vol. 31, no. 1, pp. 31–39, Feb. 2015.