

Vision-based Formation Control with Control Barrier Function

Seung-Beom Won¹ and Hyo-Sung Ahn^{1*}

¹School of Mechanical and Robotic Engineering, Gwangju Institute of Science and Technology (GIST), Gwangju, 61005, Republic of Korea (sbwon1916@gm.gist.ac.kr; hyosung@gist.ac.kr) * Corresponding author

Abstract: This study explores a vision-based formation control algorithm for nonholonomic mobile robots operating under visibility constraints without relying on inter-robot communication. The camera used to monitor the leading agent has field of view (FOV) constraints, making it crucial to maintain real-time visibility for effective leading agent tracking. To address these challenges, we designed a barrier function based on the camera's pixel coordinates and showed that the control barrier function based quadratic programming satisfies both the visibility maintenance and nonholonomic property. The effectiveness of the proposed method is validated through simulations that demonstrate its ability to maintain visibility and achieve the desired formation.

Keywords: Image-based visual servoing, Visibility maintenance, Control barrier function.

1. INTRODUCTION

Formation control is a type of multi-robot control and has been extensively studied over the past decade. This approach has various applications, such as exploration [1], mapping [2], and surveillance [3]. Most existing work on formation control assumes that each robot shares global position information or a common reference frame and that inter-robot communication is reliable. However, this makes it difficult to apply traditional formation control methods in environments where GPS is not available or communication is degraded. To overcome these problems, control techniques that utilize various sensors beyond onboard sensing, such as sonar and LiDAR, have been explored. Among these, visual cameras have become a popular option for formation control that uses only available relative onboard sensing due to their low cost, high versatility, and advances in computer vision technique [4].

There are two main approaches to vision-based formation control: Position-Based Visual Servoing (PBVS) and Image-Based Visual Servoing (IBVS). The PBVS method uses image features to calculate relative pose and then controls the system based on the reconstructed 3D Cartesian space information. In contrast, the IBVS method directly controls the error between the current and desired image coordinates. IBVS is computationally efficient and inherently more robust to camera calibration and target modeling errors because IBVS does not require pose reconstruction [5].

Since vision-based formation control does not rely on robot communication, the leading agent must remain within the camera's field of view (FOV) throughout the control process. Maintaining visibility for real-time visual feedback of the leading agent is crucial and necessary for generating control inputs. If the follower has non-holonomic characteristics, the leading agent may move out of the camera's FOV. It makes tracking impossible. Existing IBVS-based studies overcome these challenges

by designing specific feedback control laws based on the target's image coordinates. In [6], the visibility constraint was characterized by a convex and polyhedral set that encodes both position and orientation information, and formation control is performed based on the concept of control invariance. [7] proposed a method based on a barrier Lyapunov function and a potential function, and it generates control inputs that satisfy the constraints to perform formation control. However, these approaches have difficulties in ensuring system stability when one agent is required to follow multiple agents simultaneously. Furthermore, the proposed formations are limited to linear or V-shaped with a leader that moves at a constant speed rather than forming a specific shape.

To address these limitations, this paper presents a control barrier function that represents the FOV constraints and a vision-based formation control algorithm that utilizes this function. The proposed algorithm uses the control barrier function to generate control inputs that satisfy both the camera's FOV and the robot's nonholonomic constraints. It allows a single agent to follow multiple agents simultaneously by simply adding constraint conditions and enables diverse shapes of formation, unlike existing methods.

The remainder of this paper is organized as follows: Section 2 provides the necessary preliminaries, that include notations and background information on formation control and control barrier function. In Section 3, we present a vision-based formation control algorithm with a designed control barrier function and simulation results. Last, Section 4 is a conclusion.

2. PRELIMINARIES

In this paper, we use following notations. The set of real number is denoted by \mathbb{R} , and the n -dimensional Euclidean space is denoted by \mathbb{R}^n . The set of positive real numbers is denoted by \mathbb{R}_+ . For a given finite set A , $|A|$ denotes the cardinality of A . For a given vector v , $\|v\|$ denotes the Euclidean norm of v .

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Republic of Korea government (MSIT) (2022R1A2B5B03001459)

2.1 Graph Representation

A directed graph $G = (V, E)$ consists of the vertex set V and the edge set $E \subseteq V \times V$. An edge e_{ij} represents a unique connection between two vertices i and j . The notation e_{ij} is used when the edge originates from vertex i and terminates at vertex j . We assume that the graph contains no self-loops, i.e., $e_{ii} \notin E$ for all $i \in V$, no multiple edges are joining any two of the vertices, and is acyclic¹. The set of neighbors of vertex i is defined as $N_i = \{j \in V : e_{ij} \in E\}$.

2.2 Acyclic Minimally Persistent Graph

In formation control, each vertex represents the position of an agent in Euclidean space; thus, the position vector of an agent i can be expressed as $p_i \in \mathbb{R}^n$. Then, we call $\mathbf{p} = [p_1^T \dots p_{|V|}^T]^T \in \mathbb{R}^{n|V|}$ a *realization* of G in \mathbb{R}^n . The combination of the graph and its realization, (G, \mathbf{p}) , is called a *framework*.

The frameworks (G, \mathbf{p}) and (G, \mathbf{q}) with two different realizations \mathbf{p} and \mathbf{q} are considered *equivalent* if

$$\|p_i - p_j\| = \|q_i - q_j\|, \forall e_{ij} \in E.$$

The realizations \mathbf{p} and \mathbf{q} are *congruent* if

$$\|p_i - p_j\| = \|q_i - q_j\|, \forall i, j \in V.$$

If no other point exists for p_j , called p'_j , then p_j is said to be a *fitting* for L , where $L = \{d_{ij} \in \mathbb{R}_+ : e_{ij} \in E\}$ is a given distance set. The following is a mathematical representation of fitting.

$$\begin{aligned} & \{(i, j) \in E : \|p_i - p_j\| = d_{ij}\} \\ & \subseteq \{(i, j) \in E : \|p_i - p'_j\| = d_{ij}\}. \end{aligned}$$

If $\forall j \in V$, p_j is fitting for L , then \mathbf{p} is referred to as a *fitting realization* for L .

When a desired formation framework is given, verifying whether its realization is unique is essential. If the realization is not unique, the designed controller might form a different formation from the intended one despite satisfying the given framework. The property that guarantees a framework with a directed graph has a unique realization is referred to as *persistence*.

Definition 1 ([8]) For a given framework (G, \mathbf{p}) in \mathbb{R}^n , let L be a set of distances given by $L = \{d_{ij} : d_{ij} = \|p_i - p_j\|, e_{ij} \in E\}$. The framework (G, \mathbf{p}) is *persistent* in \mathbb{R}^n if there exists ϵ such that every realization $\mathbf{p}' \in \mathbb{R}^{2|V|}$ fitting for L , and \mathbf{p}' satisfying $d(\mathbf{p}, \mathbf{p}') < \epsilon$ is congruent to \mathbf{p} , where $d(\mathbf{p}, \mathbf{p}') = \max_{i \in V} \|p_i - p'_i\|$. If (G, \mathbf{p}) is persistent for almost all realizations of G , then G is *generically persistent*, and a generically persistent graph G is *minimally persistent* if none of the edges can be removed without losing persistence.

This paper considers the formations with acyclic minimally persistent graphs to ensure their uniqueness of realization. The following is the property of an acyclic minimally persistent graph.

Proposition 1 ([9]) An acyclic graph having more than one vertex is minimally persistent in \mathbb{R}^n if and only if

¹A graph is acyclic if it contains no cycles, it means there is no path that starts and ends at the same vertex.

- One vertex (the leader) has an out-degree of 0.
- One vertex (the first follower) has an out-degree of 1, and the corresponding edge is incident to the leader.
- All other vertices have an out-degree of 2.

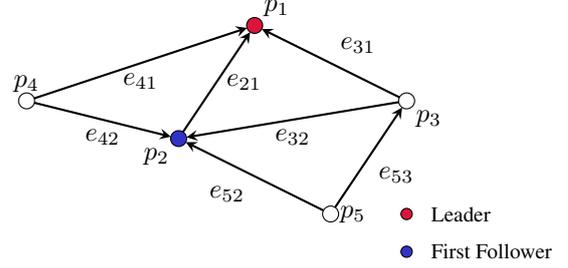


Fig. 1.: Acyclic minimally persistent graph.

2.3 Robot Model

When an edge $(i, j) \in E$ exists, the robot R_i , $i \in V$, tracks the moving target denoted by the leading robot R_j , $j \in N_i$. We assume that all robots are nonholonomic agents. However, for simplicity, robot R_i deals with robot $R_j \in N_i$ as holonomic robots, even though robot R_j is nonholonomic.

The kinematic model of robot R_i is described by

$$\dot{X}_i = \begin{bmatrix} \dot{x}_i \\ \dot{y}_i \\ \dot{\theta}_i \end{bmatrix} = \begin{bmatrix} \cos \theta_i & 0 \\ \sin \theta_i & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_i \\ \omega_i \end{bmatrix} = G(X_i)u_i, \quad (1)$$

and the kinematic model of leading robot R_j , estimated by robot R_i , is represented by

$$\begin{bmatrix} \dot{x}_j \\ \dot{y}_j \end{bmatrix} = \begin{bmatrix} v_{j_x} \\ v_{j_y} \end{bmatrix}, \quad (2)$$

where v_{j_x} and v_{j_y} represent the x-axis and y-axis linear velocity of the leading robot R_j . Each states are described in a R_i 's local reference frame.

2.4 Monocular Camera Model

A monocular camera is mounted on the R_i to track the leading agent R_j . The camera's optical axis is aligned with the R_i 's heading direction. Suppose that the R_i 's body-fixed frame is denoted as F_i , F_c denotes the camera frame, and the coordinate of the optical center in F_i is $(0, 0, 0)$. The feature point coordinate (relative position) of R_j , which is used to calculate tracking errors, can be described in R_i 's body-fixed frame F_i as follows:

$$\begin{bmatrix} x_{F_i} \\ y_{F_i} \\ z_{F_i} \end{bmatrix} = \begin{bmatrix} \cos \theta_i & \sin \theta_i & 0 \\ -\sin \theta_i & \cos \theta_i & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_j - x_i \\ y_j - y_i \\ h \end{bmatrix}, \quad (3)$$

where (x_i, y_i) is a state of R_i 's local reference frame, $h > 0$ is the fixed height of the leading agent's feature point between the optical center, and $[x_{F_i}, y_{F_i}, z_{F_i}]^T$ represents the coordinate of the feature point in F_i .

A feature point can be converted into a camera frame point by using the following principle:

$$\begin{bmatrix} m \\ n \\ 1 \end{bmatrix} = \frac{1}{z_{F_c}} AP \quad (4)$$

with

$$A = \begin{bmatrix} a_m & 0 & m_0 \\ 0 & a_n & n_0 \\ 0 & 0 & 1 \end{bmatrix}; P = \begin{bmatrix} x_{F_c} \\ y_{F_c} \\ z_{F_c} \end{bmatrix}, \quad (5)$$

where A is the intrinsic matrix of the camera, and P represents the coordinate of the feature point in the camera frame F_c .

Using equation (6), we can find the relationship between the frames F_c and F_f .

$$\begin{bmatrix} m \\ n \end{bmatrix} = \frac{1}{x_{F_i}} \begin{bmatrix} a_m & 0 & m_0 \\ 0 & a_n & n_0 \end{bmatrix} \begin{bmatrix} -y_{F_i} \\ -h \\ x_{F_i} \end{bmatrix} \quad (6)$$

2.5 Time Varying Control Barrier Function

Consider an affine control system

$$\dot{X} = f(X) + G(X)u \quad (7)$$

where f and G are locally Lipschitz, $X \in \mathbb{R}^n$ and $u \in U \subset \mathbb{R}^m$ are the state and control input, respectively.

In safety-critical situations, it is essential to maintain the stability of a dynamic system and keep it within a secure region. This can be achieved by controlling the system's state within a safe set, which is represented mathematically as:

$$S(t) = \{X \in \mathbb{R}^n \mid h(t, x) \geq 0\} \quad (8)$$

where $h : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}$ is a continuously differentiable function.

Definition 2 ([10]) A function $h : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}$ is a time-varying control barrier function (CBF) defined on a set D , if $S(t) \subseteq D \subset \mathbb{R}^n$ and there exists an extended class \mathcal{K} function² $\alpha : \mathbb{R} \rightarrow \mathbb{R}$ such that

$$\sup_{u \in U} \{L_f h(t, X) + L_G h(t, X)u + \frac{\partial h}{\partial t}(t, X)\} \geq -\alpha(h(t, X)), \quad (9)$$

for all $X \in D$ and for all $t \geq 0$.

Lemma 1 ([10]) Let $h : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}$ be the time-varying CBF defining a time-varying safe set (8). Then, for the system (7), a Lipschitz continuous control input u satisfying

$$L_f h(t, X) + L_G h(t, X)u + \frac{\partial h}{\partial t}(t, X) \geq -\alpha(h(t, X)), \quad (10)$$

for all $X \in S(t)$ and for all $t \geq 0$, renders $S(t)$ forward invariant for all $t \geq 0$.

²An extended class \mathcal{K} function is a continuous function $\alpha : (-b, a) \rightarrow (-\infty, \infty)$ for some $a, b > 0$, strictly increasing and $\alpha(0) = 0$.

3. TRACKING ALGORITHM DESIGN

3.1 Visibility Constraint

Due to the limited FOV of a camera, the pixel coordinates (m, n) must adhere to the following constraints:

$$m_{\min} \leq m \leq m_{\max}, \quad n_{\min} \leq n \leq n_{\max} \quad (11)$$

where m_{\min} , m_{\max} , n_{\min} , and n_{\max} are fixed parameters that are determined by the pixel resolution of the camera.

We design a barrier function, $h(t, X_i)$, that reflects the above FOV constraints, where X_i is the state of the tracking robot R_i .

$$h(t, X_i) = 1 - \left(\frac{m(t, X_i)}{a}\right)^k - \left(\frac{n(t, X_i)}{b}\right)^k \geq 0 \quad (12)$$

where $a = -m_{\min} = m_{\max}$, $b = -n_{\min} = n_{\max}$ and k is arbitrary constant. The larger k , the better the barrier function approximates the camera frame.

3.2 Vision-based Formation Control Algorithm

The vision-based formation control algorithm utilizes a quadratic optimization approach incorporating a nominal controller and control barrier functions. The nominal controller generates control inputs for formation control without considering visibility constraints. Subsequently, a quadratic optimization process determines the control inputs that satisfy the camera FOV and nonholonomic constraints while minimizing the variation from the nominal controller's inputs. The quadratic optimization formulation for each agent is as follows:

$$\begin{aligned} u_i^* &= \arg \min_{u_i} \|u_i - \hat{u}_i\|^2 \\ \text{s.t. } & L_f h(t, X_i) + L_G h(t, X_i)u_i + \frac{\partial h(t, X_i)}{\partial t} \\ & \geq -\alpha(h(t, X_i)) \end{aligned} \quad (13)$$

where \hat{u}_i represents the control input from the nominal controller, and the optimized result u_i^* is the actual control input.

While any formation control law can be used as a nominal controller, these approaches usually do not consider the agents' nonholonomic characteristics. Since we assume that all agents have nonholonomic properties, traditional formation control techniques are not appropriate. To achieve effective control, using a nominal controller that includes nonholonomic dynamics is more efficient because it minimizes the deviation from the actual control input. The following equation presents a control law that reflects a nonholonomic nature, as proposed in [11].

$$\begin{aligned} \hat{u}_i &= \begin{bmatrix} v_i \\ \omega_i \end{bmatrix} = \begin{bmatrix} h_i^T \\ h_i^{\perp T} \end{bmatrix} f_i \\ &= \begin{bmatrix} \cos \theta_i & \sin \theta_i \\ -\sin \theta_i & \cos \theta_i \end{bmatrix} f_i, \end{aligned} \quad (14)$$

where $h_i \in \mathbb{R}^n$ is the unit-length heading vector of agent i , and f_i is the original distance-based formation control input proposed by [9].

The relative leading agent position is required to compute \hat{u} . In this paper, the follower tracks a fixed feature on the leading agent using its camera. It means that the height of the leading agent's feature point is constant. Consequently, as shown in Equation (6), the relative position of the leading agent can be determined based on the camera data. Additionally, to calculate $\frac{\partial h(t, X)}{\partial t}$ in Equation (13), the velocity of the leading agent is necessary. By using the relative leading agent position, visual odometry (VO), and inertial measurement unit (IMU), we can obtain the relative leading agent position in a local reference frame of R_i . This allows us to estimate the leading agent's velocity by calculating the change in its position over time.

3.3 Simulation Results

The first simulation represents a scenario where a follower tracks a leader. The leading agent R_1 begins at position (3, 0), while the tracking agent R_2 starts at position (1, 0) with an orientation of -36° . The camera mounted on the tracking robot has a FOV constraint of $\pm 55^\circ$, and the desired distance between the tracking robot and the leading robot is 2.8 units. In this scenario, the leading agent moves in a uniform circular motion with a radius of 3 units. Fig. 2 illustrates the tracking performance using only the nominal controller, whereas Fig. 3 shows the trajectory of the tracking robot when the proposed algorithm is applied. As shown in the "after 5s" node of Fig. 2, R_1 falls outside the FOV of R_2 . In contrast, Fig. 3 demonstrates that the proposed algorithm consistently maintains the FOV constraint, which is further supported by the analysis in Fig. 4.

The barrier function presented in this paper can reach a maximum value of 1. According to the definition of a CBF, it must always remain bigger than 0 to satisfy the constraints. Therefore, the valid range for the barrier function, where the constraints are satisfied, is between 0 and 1. Fig. 4a represents the barrier function values corresponding to Fig. 2, while Fig. 4b corresponds to the barrier function values shown in Fig. 3. As observed in Fig. 4a, the barrier function drops below 0 at the beginning and around 5 seconds, which is consistent with the behavior of the "after 5s" node in Figure 2. On the other hand, Fig. 4b shows that the barrier function consistently remains above 0, indicating that the proposed algorithm successfully satisfies the FOV constraint.

In the second scenario, the first follower R_2 follows the leader R_1 , and the other agent R_3 simultaneously follows both agents connected by an edge, forming an acyclic minimally persistent graph. Unlike R_2 , R_3 must follow both agents simultaneously, so the optimization problem for R_3 has two constraints. Each constraint restricts the position of R_1 and R_2 on a camera frame; thus, they do not move out of the FOV. The desired lengths of the edges are as follows:

$$\|e_{21}\| = 2.8, \quad \|e_{31}\| = 3.6, \quad \|e_{32}\| = 2.1.$$

The trajectory shown in Fig. 5 demonstrates the application of the proposed algorithm for formation control

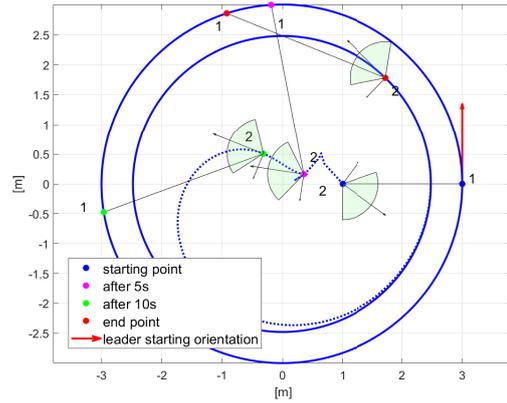


Fig. 2.: The trajectory using the nominal controller alone.

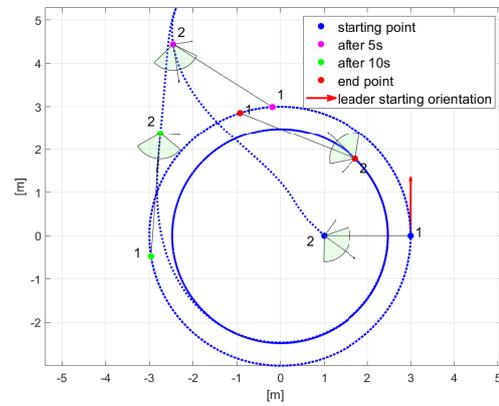
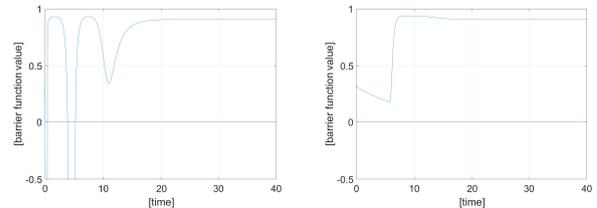


Fig. 3.: The trajectory using the proposed algorithm.



(a) Fig. 2's barrier function value (b) Fig. 3's barrier function value

Fig. 4.: Barrier function values of the first scenario.

over time. Fig. 6a displays the value of the barrier function over time, illustrating that the constraint is consistently satisfied, even when R_3 follows two agents simultaneously. Fig. 6b also indicates that the formation error converges to zero over time, where formation error is the difference between the actual edge length and the desired edge length.

4. CONCLUSION

This paper investigates a vision-based formation control algorithm for nonholonomic mobile robots under visibility constraints where communication is absent. We design a barrier function that reflects the camera visibility

constraint and propose quadratic programming based on it to form a formation. The proposed barrier function can be applied to various applications such as surveillance, tracking, etc.

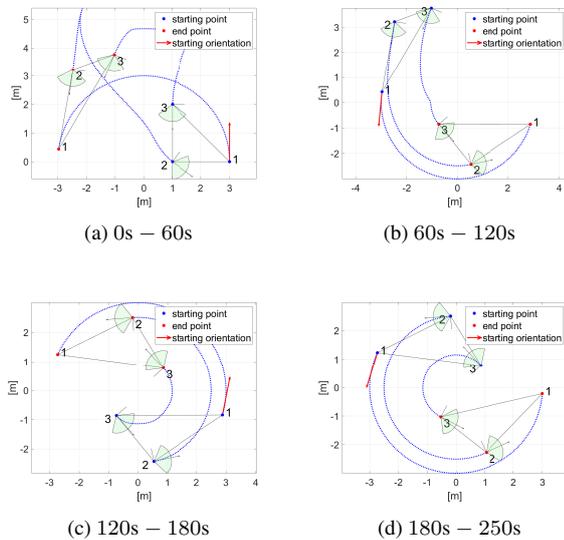


Fig. 5.: The formation trajectory using the proposed algorithm.

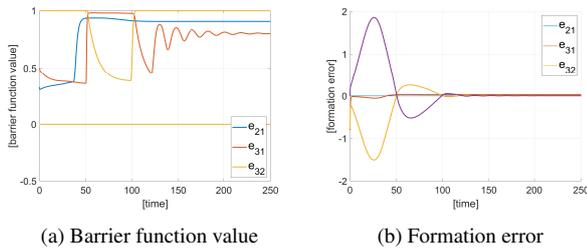


Fig. 6.: Barrier function value and formation error.

REFERENCES

- [1] M. Corah, C. O'Meadhra, K. Goel, and N. Michael, "Communication efficient planning and mapping for multi-robot exploration in large environments," *IEEE Robot. Autom. Lett.*, vol. 4, no. 2, pp. 1715–1721, Apr. 2019.
- [2] L. Paull, M. Seto, J. J. Leonard, and H. Li, "Probabilistic cooperative mobile robot area coverage and its application to autonomous seabed mapping," *Int. J. Robot. Res.*, vol. 37, no. 1, pp. 21–45, 2018.
- [3] Y. Yang, Y. Xiao, and T. S. Li, "A survey of autonomous underwater vehicle formation: Performance, formation control, and communication capability," *IEEE Commun. Surveys Tuts*, vol. 23, no. 2, pp. 815–841, 2021.
- [4] R. Tron, J. Thomas, G. Loianno, K. Daniilidis, and V. Kumar, "A distributed optimization framework for localization and formation control: Applications

to vision-based measurements," *IEEE Control Syst. Mag.*, vol. 36, no. 4, pp. 22–44, Aug. 2016.

- [5] 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.
- [6] F. Morbidi, F. Bullo, and D. Prattichizzo, "Visibility maintenance via controlled invariance for leader-follower vehicle formations," *Automatica*, vol. 47, no. 5, pp. 1060–1067, 2011.
- [7] J. Lin, Z. Q. Miao, H. Zhong, W. X. Peng, Y. N. Wang, and R. Fierro, "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.
- [8] J. M. Hendrickx, B. D. O. Anderson, J.-C. Delvenne, and V. D. Blondel, "Directed graphs for the analysis of rigidity and persistence in autonomous agent systems," *International Journal of Robust and Nonlinear Control*, vol. 17, no. 10-11, pp. 960–981, 2007.
- [9] M. C. Park and H. S. Ahn, "Distance-based acyclic minimally persistent formations with non-steepest descent control," *International Journal of Control Automation & Systems*, vol. 14, no. 1, pp. 163–173, 2016.
- [10] K. Garg and D. Panagou, "Robust Control Barrier and Control Lyapunov Functions with Fixed-Time Convergence Guarantees," in *2021 American Control Conference (ACC)*, pp. 2292–2297, May 2021.
- [11] S. Zhao, D. V. Dimarogonas, Z. Sun, and D. Bauso, "A general approach to coordination control of mobile agents with motion constraints," *IEEE Trans. Autom. Control*, vol. 63, no. 5, pp. 1509–1516, May 2018.