

Article

Wolf-Pack-Inspired Distributed Encirclement for UAV–USV Systems via Visual Connectivity Preservation

Jinheng Xiao ¹, Shihan Kong ^{2,*} , Jinan Sun ³, Yingnan Li ⁴, Fang Wu ⁴ and Junzhi Yu ^{2,*} 

¹ School of Aerospace Engineering, Beijing Institute of Technology, Beijing 100081, China; 1120220683@bit.edu.cn

² The State Key Laboratory for Turbulence and Complex Systems, School of Advanced Manufacturing and Robotics, Peking University, Beijing 100871, China

³ National Engineering Research Center for Software Engineering, Peking University, Beijing 100871, China; sjn@pku.edu.cn

⁴ SPIC Nuclear Energy Co., Ltd., Yantai 100088, China; liyingnan@spic.com.cn (Y.L.); wufang@spic.com.cn (F.W.)

* Correspondence: kongshihan@pku.edu.cn (S.K.); yujunzhi@pku.edu.cn (J.Y.)

Abstract

This paper presents a wolf-pack-inspired distributed encirclement framework for heterogeneous unmanned aerial vehicle–unmanned surface vehicle (UAV–USV) teams operating with limited communication and intermittent visual sensing. The proposed growth-based visual connectivity encirclement (GB-VCE) method separates high-level role evolution from low-level motion execution. Each agent accumulates local evidence from target observation, visual connectivity contribution, motion stability, and platform characteristics, enabling smooth transitions among leader, relay, searcher, and encircler roles. Visual links and relay relations are treated as coordination resources rather than auxiliary sensing constraints, allowing target-related information to propagate through the team without centralized fusion or persistent all-to-all communication. Simulations with paired random seeds show that GB-VCE improves encirclement accuracy, target-information coverage, convergence consistency, and heterogeneous role complementarity compared with fixed-role, static-leader, UAV-only, and USV-only baselines. The results indicate that biomimetic role growth and visual connectivity preservation provide an interpretable and robust coordination principle for air–sea cooperative encirclement.

Keywords: heterogeneous UAV–USV systems; distributed encirclement; communication-constrained environments; biological coordination; visual connectivity preservation; wolf-pack predation



Academic Editors: Stanislav N. Gorb, Zhendong Dai and Ille C. Gebeshuber

Received: 17 June 2026

Revised: 13 July 2026

Accepted: 16 July 2026

Published: 21 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Encirclement is a demanding cooperative behavior in which multiple agents must maintain target observability, safe spacing, angular distribution, and robustness to target motion. The problem is particularly challenging for heterogeneous UAV–USV teams: UAVs provide fast repositioning and wide-area visual relay, whereas USVs offer persistent surface-level monitoring but slower dynamics. In maritime missions, continuous communication and global state fusion are often unavailable because of occlusion, bandwidth limits, platform motion, and environmental disturbance. Thus, the central question is how to generate distributed encirclement from local perception while preserving the flow of target-related information.

Table 1 lists the comparison of representative methods for cooperative encirclement and heterogeneous coordination. Existing approaches include geometric enclosing control, optimization and learning-based coordination, and connectivity-aware multi-robot control. Chen et al. [1] developed a monocular vision-based target-enclosing method for USVs under field-of-view constraints and uncertain dynamics. Chen et al. [2] further investigated bearing-only cooperative enclosing for multiple USVs with unknown dynamics and sideslip. Mnih et al. [3] established a deep reinforcement learning framework, while Van Hasselt et al. [4] improved value-estimation stability through double Q-learning. Orr and Dutta [5] reviewed the application of multi-agent deep reinforcement learning to multi-robot systems. Busoniu et al. [6] summarized the principles and challenges of multi-agent reinforcement learning. Lowe et al. [7] proposed a multi-agent actor–critic framework for mixed cooperative–competitive environments. Rashid et al. [8] introduced monotonic value-function factorization for cooperative multi-agent learning, whereas Foerster et al. [9] developed counterfactual policy gradients for multi-agent credit assignment. Although these learning-based approaches improve adaptability, their functional roles are generally implicit. Müller and Hutter [10] studied cooperative vision-based navigation for heterogeneous robot teams. Pesce and Montana [11] investigated connectivity-driven communication for multi-agent coordination, while Shetty et al. [12] proposed decentralized connectivity maintenance under motion and sensing uncertainties. However, these methods rarely combine visual connectivity with an explicit and biologically interpretable role-evolution mechanism for heterogeneous encirclement.

Table 1. Comparison of representative methods for cooperative encirclement and heterogeneous coordination.

Method	Main Idea	Advantages	Limitations
Geometric target enclosing [1,2]	Maintain desired bearing, distance, or angular distribution around the target	Interpretable control structure and direct encirclement objective	Often assumes continuous observation or sufficient state information
Learning-based coordination [3–5]	Learn cooperative policies through reinforcement learning or multi-agent value/policy decomposition	Adaptive to complex dynamics and environmental uncertainty	Role assignment is often implicit and difficult to interpret
Vision/connectivity-aware coordination [10–12]	Preserve visual links, relay relations, or network connectivity among agents	Improves robustness under constrained communication	Usually focuses on connectivity maintenance rather than encirclement role evolution
Temporal role allocation [13,14]	Introduce continuity or stability into multi-agent role switching	Reduces frequent role oscillation	Not specifically designed for heterogeneous visual encirclement
Biological collective hunting [15]	Use natural role specialization and adaptive group hunting as inspiration	Provides interpretable principles for distributed encirclement	Requires formalization for robotic control
Proposed method [16–18]	Growth-based role evolution with visual connectivity preservation	Adaptive, distributed, interpretable, and suitable for UAV–USV encirclement	Requires evaluation of growth state and local visual topology

The references listed in each row denote representative studies or conceptual foundations related to the corresponding methodological category. For the proposed-method row, these references indicate biological inspiration and visual-connectivity-related foundations, rather than previous presentations of the proposed GB-VCE framework. The original contribution of this work lies in integrating growth-based role evolution and visual connectivity preservation into a unified distributed encirclement framework for heterogeneous UAV–USV systems.

Biological collective hunting offers a useful design principle. Wolf-pack and mixed-species hunting studies show that leadership and contribution are not fixed attributes; rather, they emerge from local observation, spatial position, prey response, and the changing contribution of individuals [15]. Goeckner et al. [16] demonstrated that graph-based

learning can improve resilient distributed coordination in multi-robot systems. Ning and Xie [13] reviewed adaptive decision-making mechanisms in multi-agent reinforcement learning, while Wang et al. [14] investigated deep-reinforcement-learning-based multi-robot coordination in complex environments. These studies suggest that UAV-USV roles should evolve from accumulated evidence rather than be determined only by instantaneous conditions. In this paper, visual connectivity is therefore treated as a coordination primitive: agents that maintain target-information paths can become relay agents, agents with strong observation and stability can become leaders, and agents with poor information access can search or recover connectivity.

Unlike conventional role-allocation methods based on fixed roles or instantaneous conditions, the proposed mechanism uses accumulated observation, connectivity, stability, and platform information to achieve continuous role evolution. Moreover, visual connectivity is directly incorporated into relay selection and target-information propagation rather than being treated only as a connectivity constraint.

Figure 1 summarizes the biological abstraction used in this work. Cai et al. [17] showed that a biologically inspired gene-regulation network can generate emergent UAV-swarm behaviors from local interactions. Shi et al. [18] investigated communication-connectivity restoration in multi-robot systems. Inspired by these studies, the proposed two-layer framework separates role growth and task-state selection from executable motion control, allowing target search, relay maintenance, tracking, and encirclement behaviors to emerge within the same distributed framework.

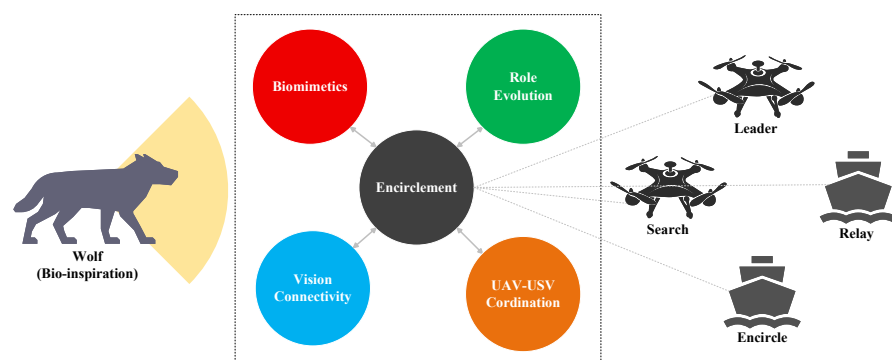


Figure 1. Biologically inspired encirclement mechanism.

Compared with representative cooperative encirclement approaches, the principal advantage of GB-VCE lies in extending encirclement from geometric formation regulation to adaptive functional organization. Conventional geometric methods primarily regulate inter-agent distance, bearing, or angular distribution, whereas GB-VCE additionally determines which agents should undertake target guidance, visual relay, search recovery, and geometric enclosure according to their locally evaluated contributions. Unlike fixed-role and static-leader strategies, these functions are not permanently assigned but are reorganized as target observability, visual topology, motion stability, and platform suitability vary. Moreover, the growth-based accumulation of role evidence provides temporal continuity, reducing the sensitivity of instantaneous role-selection schemes to transient sensing and topology fluctuations. In contrast to learning-based coordination, the resulting functional differentiation remains explicit and physically interpretable. Finally, visual connectivity is not treated merely as a feasibility constraint; its contribution is incorporated into role evaluation, so that agents preserving critical target-information paths are actively assigned relay functions. This enables heterogeneous UAVs and USVs to establish complementary roles while maintaining distributed encirclement under constrained communication.

The innovations of this paper are as follows:

1. A wolf-pack-inspired distributed encirclement framework is proposed for heterogeneous UAV–USV teams under communication-constrained visual sensing, enabling cooperative target acquisition, relay maintenance, and encirclement without centralized fusion.
2. A growth-based role-evolution mechanism is developed, in which leader, relay, searcher, and encircler roles emerge from accumulated observation quality, visual connectivity contribution, motion stability, and platform bias.
3. A visual-connectivity-preserving motion-control strategy is designed and validated through multi-seed simulations, showing improved encirclement accuracy, target-information coverage, role stability, and heterogeneous complementarity over baseline methods.

2. Related Work

2.1. Biomimetic Encirclement and Collective Hunting

Biomimetic coordination seeks transferable principles from biological systems rather than literal imitation. For encirclement, the relevant principles are local perception, progressive spatial closure, role differentiation, and resilient information sharing. Xie et al. [19] reviewed the development and future directions of flocking-control models for multi-agent systems. Yan and Su [20] analyzed the role of alignment interactions in the collective behavior of self-propelled particles. Gable et al. [15] showed that the contribution of individual wolves changes with prey size, prey abundance, and provisioning conditions. Sampaio et al. [21] further demonstrated that leadership and collective success in mixed-species hunting groups depend on social influence and group composition. Oh et al. [22] systematically reviewed multi-agent formation-control methods. Cao et al. [23] summarized the theoretical foundations of distributed multi-agent coordination. Ge et al. [24] reviewed distributed coordination under intermittent sampling and communication. Feng et al. [25] discussed resilient cooperative control and optimization, while Wang et al. [26] investigated resilient consensus control under disturbances and failures. These studies motivate an encirclement mechanism that integrates geometric enclosure, visual connectivity preservation, and temporally continuous role growth. These results motivate an encirclement mechanism that combines geometric enclosure, visual connectivity preservation, and temporally continuous role growth.

2.2. Target Enclosing Control and Learning-Based Coordination

Target enclosing has been addressed through bearing control, circular formation control, collision avoidance, and learning-based planning. Chen et al. [1] considered field-of-view constraints and uncertain dynamics in monocular vision-based USV target enclosing. Chen et al. [2] addressed cooperative enclosing using bearing-only measurements under unknown USV dynamics and sideslip. Ma et al. [27] applied deep reinforcement learning to multi-robot target encirclement with collision avoidance. Zhang et al. [28] extended this approach to multi-target encirclement. Lowe et al. [7] developed a multi-agent actor–critic framework for decentralized policy learning. Rashid et al. [8] proposed value-function factorization for cooperative decision making, whereas Foerster et al. [9] introduced counterfactual policy gradients for multi-agent credit assignment. However, the functional division learned by these approaches is generally implicit. Chung et al. [29] reviewed aerial swarm robotics and emphasized the importance of distributed sensing and coordination. Qu et al. [30] applied multi-agent reinforcement learning to cooperative encirclement by autonomous surface vehicles. These studies indicate that sensing and interaction topology are as important as encirclement geometry in air–sea cooperative tasks.

The proposed GB-VCE differs by making roles explicit, interpretable, and driven by local visual connectivity evidence.

3. Architecture and Mathematical Model

3.1. Overall Architecture of the Biological Distributed Encirclement Framework

The proposed GB-VCE framework shown in Figure 2 realizes cooperative encirclement of a maneuvering surface target by a heterogeneous UAV-USV team. The architecture contains three coupled modules: visual perception and connectivity evaluation, growth-based role evolution, and role-dependent motion execution. UAVs contribute fast relocation and wide-area visual relay, whereas USVs provide persistent surface-level encirclement. Instead of assigning a permanent leader, the system lets leadership and relay functions emerge from accumulated observation quality, connectivity contribution, and motion stability.

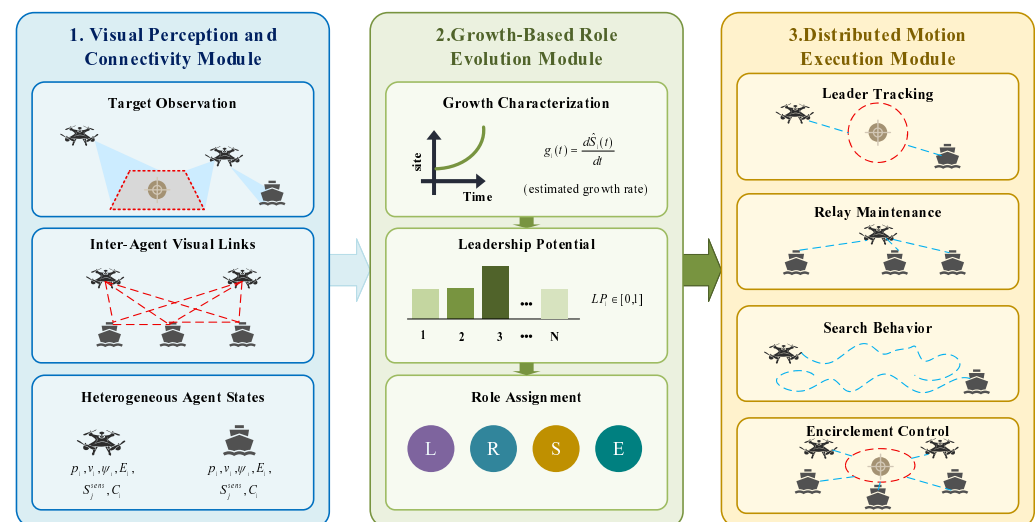


Figure 2. Overall architecture of the proposed UAV-USV biological distributed encirclement framework.

The framework follows a two-layer decision structure shown in Figure 3. The upper layer evaluates biological growth state and role potential, while the lower layer transforms the selected role into executable acceleration and heading commands. This separation keeps the method interpretable: the leader guides target-related encirclement, relay agents preserve visual information propagation, searchers recover target or teammate visibility, and encirclers regulate their angular distribution around the target.

At each sampling instant, local measurements update the target observation quality $O_i(t)$, the visual graph $\mathcal{G}_v(t)$, and the connectivity contribution $C_i(t)$. These variables drive the growth state $z_i(t)$ and leadership potential $P_i(t)$, after which the role variable is selected as

$$r_i(t) \in \{L, R, S, E\}, \quad (1)$$

where L , R , S , and E denote leader, relay, searcher, and encircler. The resulting closed loop is summarized by

$$\{O_i(t), \mathcal{G}_v(t), C_i(t)\} \rightarrow \{g_i(t), z_i(t), P_i(t)\} \rightarrow r_i(t) \rightarrow (a_i^c(t), \alpha_i^c(t)) \rightarrow \mathbf{x}_i(t + \Delta t). \quad (2)$$

where $g_i(t)$, $z_i(t)$, and $P_i(t)$ denote the growth drive, accumulated growth state, and leadership potential, respectively. This loop links perception, role evolution, and motion execution without centralized state fusion.

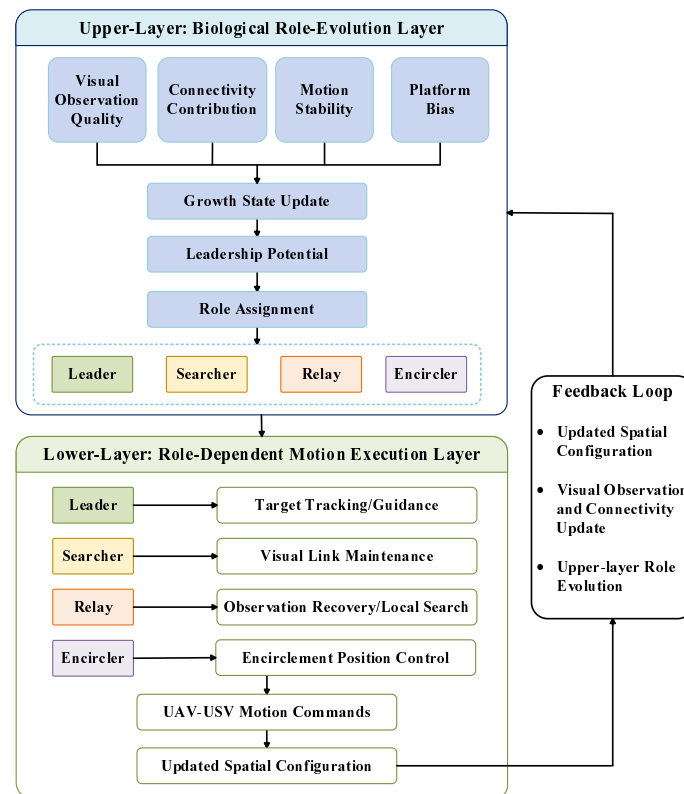


Figure 3. Two-layer distributed decision structure.

3.2. Heterogeneous Agent, Target, and Visual Connectivity Model

Let the heterogeneous team be

$$\mathcal{A} = \mathcal{A}_{\text{UAV}} \cup \mathcal{A}_{\text{USV}}, \quad N = |\mathcal{A}|. \quad (3)$$

For agent i , the planar state is denoted by $\mathbf{x}_i = [\mathbf{p}_i^\top, \psi_i, v_i, \omega_i]^\top$, where $\mathbf{p}_i \in \mathbb{R}^2$ is position, ψ_i is heading, and v_i and ω_i are linear and angular velocities. The target state is $\mathbf{x}_T = [\mathbf{p}_T^\top, \mathbf{v}_T^\top]^\top$. Heterogeneity is represented by platform-dependent limits on sensing range, field of view, speed, acceleration, and turning rate.

A visual link from agent i to object j is valid only when range and bearing constraints are both satisfied. With relative vector $\mathbf{d}_{ij} = \mathbf{p}_j - \mathbf{p}_i$, the distance condition is $\|\mathbf{d}_{ij}\| \leq R_i$, and the field-of-view condition is

$$|\text{wrap}(\text{atan2}(d_{ij,y}, d_{ij,x}) - \psi_i)| \leq \frac{\Phi_i}{2}. \quad (4)$$

Here, $\text{wrap}(\cdot)$ maps an angle to the interval $(-\pi, \pi]$. The visual accessibility indicator is therefore

$$\delta_{ij}(t) = \mathbb{I}(\|\mathbf{d}_{ij}\| \leq R_i, |\beta_{ij}(t)| \leq \Phi_i/2), \quad (5)$$

and the visual graph is $\mathcal{G}_v(t) = (\mathcal{A}, \mathcal{E}_v(t))$, where $(i, j) \in \mathcal{E}_v(t)$ if a bidirectional or task-valid visual relation exists. Target observation quality is normalized as

$$O_i(t) = \delta_{iT}(t) \exp\left(-\frac{\|\mathbf{p}_i - \mathbf{p}_T\|^2}{\sigma_R^2}\right) \cos^+ \beta_{iT}(t), \quad (6)$$

where $\cos^+(x) = \max(0, \cos x)$. Since all factors in Equation (6) are bounded, the resulting observation-quality index satisfies $0 \leq O_i(t) \leq 1$. This normalization provides a dimensionless and comparable measure across heterogeneous UAV and USV sensing conditions,

facilitates its integration with the other role-evaluation terms, and supports consistent threshold selection in the subsequent role-evolution process. Connectivity contribution $C_i(t)$ is computed from the visual graph by combining local degree, bridge contribution, and accessibility to target-observing agents. This term rewards agents that preserve the visual information path rather than merely approaching the target.

3.3. Growth-Based Role Evolution Model

The biomimetic role mechanism treats leadership as a growth process. The instantaneous growth drive of agent i is

$$g_i(t) = w_O O_i(t) + w_C C_i(t) + w_S S_i(t) + w_H H_i, \quad (7)$$

where $O_i(t)$, $C_i(t)$, $S_i(t)$, and H_i denote target observation quality, connectivity contribution, motion stability, and heterogeneous platform bias, respectively; w_O , w_C , w_S , and w_H are the corresponding nonnegative weights.

$$z_i(t + \Delta t) = (1 - \lambda_g) z_i(t) + \Delta t g_i(t), \quad (8)$$

where λ_g is the decay coefficient. This accumulation prevents role changes caused by short visual fluctuations.

Motivated by the observation that leadership and individual contribution in collective hunting vary with local perception, spatial position, and task-dependent contribution [15], and by adaptive role-decision principles in multi-agent systems [13], we define a composite leadership-potential function in this study as

$$P_i(t) = \eta_z z_i(t) + \eta_O O_i(t) + \eta_C C_i(t) - \eta_D D_i(t), \quad (9)$$

where η_z , η_O , η_C , and η_D are weighting coefficients for growth state, observation quality, connectivity contribution, and penalty term, respectively. $D_i(t)$ penalizes excessive distance or unstable motion. The leader is selected by hysteretic maximization

$$L(t) = \arg \max_{i \in \mathcal{A}} P_i(t), \quad P_{L(t)}(t) > P_{L(t-)}(t) + h_L, \quad (10)$$

where h_L is a switching margin. Relay agents are selected from nodes with high $C_i(t)$, searchers are assigned when both observation and connectivity are weak, and the remaining agents are encirclers. This creates explicit and temporally smooth roles, consistent with adaptive biological hunting rather than instantaneous thresholding [13].

3.4. Distributed Encirclement Control Model

For an encircler, the desired position is defined as a point on a circle centered at the estimated target position, so that different encirclers occupy different angular slots around the target:

$$\mathbf{p}_i^d(t) = \hat{\mathbf{p}}_T(t) + R_E \begin{bmatrix} \cos \theta_i^d(t) \\ \sin \theta_i^d(t) \end{bmatrix}, \quad (11)$$

where R_E is the desired encirclement radius and θ_i^d is generated from the angular distribution of active encirclers. Leader and relay agents use the same target reference but add stronger observation and connectivity terms. Search agents move toward the predicted target region or toward positions that increase visual coverage.

The commanded planar acceleration is written as

$$\mathbf{u}_i(t) = \mathbf{u}_i^{\text{role}}(t) + \mathbf{u}_i^{\text{conn}}(t) + \mathbf{u}_i^{\text{align}}(t) + \mathbf{u}_i^{\text{safe}}(t), \quad (12)$$

where the four terms correspond to role objective tracking, visual connectivity maintenance, direction alignment, and collision avoidance. The command is then mapped to the heterogeneous execution model subject to platform limits

$$a_i^c = \text{sat}_{a_i}(\mathbf{u}_i^\top \mathbf{e}_{\psi_i}), \quad \alpha_i^c = \text{sat}_{\alpha_i}(k_\psi \text{wrap}(\psi_i^d - \psi_i) - k_\omega \omega_i). \quad (13)$$

The feasibility of encirclement is evaluated by success rate, radius error, stable encirclement time, target-information coverage, role switching frequency, and control smoothness. These metrics jointly assess spatial enclosure, visual information preservation, and biological role stability.

4. Methods

GB-VCE comprises three complementary modules for distributed heterogeneous encirclement. The connectivity-aware task-state switching module regulates the transition among target acquisition, connectivity recovery, and cooperative encirclement according to target-information availability, thereby avoiding premature coordination based on incomplete observations. The growth-based role-evolution module adaptively organizes agents into leader, relay, searcher, and encircler roles using accumulated local evidence, providing greater adaptability than fixed role assignment and greater temporal stability than instantaneous selection. The role-dependent control module maps these functional roles to platform-specific motion objectives, enabling UAVs and USVs to exploit their complementary sensing and mobility capabilities more effectively than a unified control strategy. Together, these modules improve information consistency, role reconfiguration, and heterogeneous coordination compared with simpler methods that rely on direct encirclement, static roles, or identical agent behaviors.

4.1. Target Acquisition and Connectivity-Aware Task-State Switching

In heterogeneous pursuit-encirclement missions, target visibility and information accessibility may be intermittent. Therefore, GB-VCE first uses a connectivity-aware task-state switch to distinguish target acquisition, connectivity recovery, and cooperative encirclement. Full encirclement is activated only when at least one valid target observation exists and target-related information is accessible to a sufficient portion of the swarm.

The three task states are defined as follows: target acquisition when no agent observes the target, connectivity recovery when the target is observed but visual information has not propagated sufficiently, and cooperative encirclement when observation and accessibility conditions are both satisfied.

Let $\mathcal{A} = \{1, 2, \dots, N\}$ denote the heterogeneous team. An observation by agent i is effective when $O_i(t) > \epsilon_O$, yielding

$$\mathcal{S}_T(t) = \{i \in \mathcal{A} \mid O_i(t) > \epsilon_O\} \quad (14)$$

A binary target-availability variable is then introduced as

$$\zeta_T(t) = \begin{cases} 1, & \mathcal{S}_T(t) \neq \emptyset, \\ 0, & \mathcal{S}_T(t) = \emptyset. \end{cases} \quad (15)$$

Here, $\zeta_T(t)$ indicates whether valid target information is currently available.

A single isolated observation is insufficient for collective encirclement; therefore, target-information accessibility is evaluated on the visual graph. Let $\mathbf{A}_{\text{vis}}(t)$ denote the visual connectivity adjacency matrix. For agent i , the accessibility indicator $\xi_i(t)$ is defined as

$$\xi_i(t) = \begin{cases} 1, & \text{agent } i \text{ is connected to at least one agent in } \mathcal{S}_T(t) \text{ through } \mathbf{A}_{\text{vis}}(t), \\ 0, & \text{otherwise.} \end{cases} \quad (16)$$

Based on $\xi_i(t)$, the target-information coverage ratio of the swarm is defined as

$$\Gamma_T(t) = \frac{1}{N} \sum_{i=1}^N \xi_i(t) \quad (17)$$

$\Gamma_T(t)$ measures the fraction of agents with direct or relayed access to target information.

The control layer uses an estimated target reference. When the target is visible, the estimate is updated as

$$\hat{\mathbf{p}}_T(t) = \mathbf{p}_T(t), \quad \hat{\mathbf{v}}_T(t) = \mathbf{v}_T(t), \quad \text{if } \zeta_T(t) = 1 \quad (18)$$

and the latest observation time is recorded as $t_{\text{last}} = t$. When the target is temporarily lost but has been observed previously, a constant-velocity prediction model is used

$$\hat{\mathbf{p}}_T(t) = \hat{\mathbf{p}}_T(t - \Delta t) + \Delta t \hat{\mathbf{v}}_T(t - \Delta t), \quad \text{if } \zeta_T(t) = 0, \quad t - t_{\text{last}} \leq T_{\text{lost}} \quad (19)$$

where Δt is the sampling interval and T_{lost} is the maximum allowable duration for short-term target-state prediction. This treatment allows the agents to continue searching around the most probable target region after a temporary loss of visual contact.

If the target has never been observed, or if the target has been lost for longer than T_{lost} , the extrapolated target estimate is no longer regarded as reliable. In this case, the target reference is reset to the global search center $\mathbf{p}_c$, and the estimated velocity is set to zero

$$\hat{\mathbf{p}}_T(t) = \mathbf{p}_c, \quad \hat{\mathbf{v}}_T(t) = \mathbf{0}, \quad \text{if } \zeta_T(t) = 0, \quad t - t_{\text{last}} > T_{\text{lost}} \quad (20)$$

This reset prevents pursuit of unreliable extrapolations and reorganizes the team around a distributed search pattern.

Each agent uses an information-consistent target reference:

$$\mathbf{p}_{T,i}^c(t) = \begin{cases} \mathbf{p}_T(t), & \xi_i(t) = 1, \\ \hat{\mathbf{p}}_T(t), & \xi_i(t) = 0, \end{cases} \quad (21)$$

and

$$\mathbf{v}_{T,i}^c(t) = \begin{cases} \mathbf{v}_T(t), & \xi_i(t) = 1, \\ \hat{\mathbf{v}}_T(t), & \xi_i(t) = 0. \end{cases} \quad (22)$$

Thus, disconnected agents use predicted references instead of unavailable true target states. The task state is jointly determined by $\zeta_T(t)$ and $\Gamma_T(t)$. If $\zeta_T(t) = 0$, leader election is suspended and all agents search.

The threshold Γ_{th} determines whether dissemination is sufficient. If $\zeta_T(t) = 1$ but $\Gamma_T(t) < \Gamma_{\text{th}}$, the team enters connectivity recovery and selects the strongest observer

$$i_L(t) = \arg \max_{i \in \mathcal{S}_T(t)} O_i(t) \quad (23)$$

The observing leader maintains target contact, while other agents execute relay or search motions until visual information becomes sufficiently accessible.

If $\zeta_T(t) = 1$ and $\Gamma_T(t) \geq \Gamma_{th}$, the cooperative-encirclement state activates the growth-based role assignment.

The complete target acquisition and connectivity-aware task-state switching procedure is summarized in Algorithm 1.

Algorithm 1 Connectivity-aware target acquisition and task-state switching.

Require: $\mathcal{A}, \{O_i(t)\}_{i=1}^N, \epsilon_O, \mathbf{A}_{vis}(t), \Gamma_{th}, (\hat{\mathbf{p}}_T, \hat{\mathbf{v}}_T), t_{last}, T_{lost}, \mathbf{p}_c$
Ensure: $\mathcal{M}(t), \{(\mathbf{p}_{T,i}^c, \mathbf{v}_{T,i}^c)\}_{i=1}^N, \{r_i(t)\}_{i=1}^N$, leader candidate $i_L(t)$

- 1: $\mathcal{S}_T(t) \leftarrow \{i \in \mathcal{A} \mid O_i(t) > \epsilon_O\}$
- 2: Compute $\zeta_i(t)$ from $\mathbf{A}_{vis}(t)$ and $\mathcal{S}_T(t)$; set $\Gamma_T(t) \leftarrow N^{-1} \sum_{i=1}^N \zeta_i(t)$
- 3: **if** $\mathcal{S}_T(t) = \emptyset$ **then**
- 4: $\zeta_T(t) \leftarrow 0, \mathcal{M}(t) \leftarrow$ target acquisition state
- 5: Suspend leader election; assign $r_i(t) \leftarrow \mathcal{R}_S, \forall i \in \mathcal{A}$
- 6: **if** target has been observed before and $t - t_{last} \leq T_{lost}$ **then**
- 7: $\hat{\mathbf{p}}_T(t) \leftarrow \hat{\mathbf{p}}_T(t - \Delta t) + \Delta t \hat{\mathbf{v}}_T(t - \Delta t)$
- 8: **else**
- 9: $\hat{\mathbf{p}}_T(t) \leftarrow \mathbf{p}_c, \hat{\mathbf{v}}_T(t) \leftarrow \mathbf{0}$
- 10: **end if**
- 11: Set $(\mathbf{p}_{T,i}^c(t), \mathbf{v}_{T,i}^c(t)) \leftarrow (\hat{\mathbf{p}}_T(t), \hat{\mathbf{v}}_T(t)), \forall i \in \mathcal{A}$
- 12: **else**
- 13: $\zeta_T(t) \leftarrow 1$
- 14: $\hat{\mathbf{p}}_T(t) \leftarrow \mathbf{p}_T(t), \hat{\mathbf{v}}_T(t) \leftarrow \mathbf{v}_T(t), t_{last} \leftarrow t$
- 15: Set $(\mathbf{p}_{T,i}^c(t), \mathbf{v}_{T,i}^c(t)) \leftarrow (\mathbf{p}_T(t), \mathbf{v}_T(t))$ for agents with $\zeta_i(t) = 1$, and $(\hat{\mathbf{p}}_T(t), \hat{\mathbf{v}}_T(t))$ for agents with $\zeta_i(t) = 0$
- 16: $i_L(t) \leftarrow \arg \max_{i \in \mathcal{S}_T(t)} O_i(t)$
- 17: **if** $\Gamma_T(t) < \Gamma_{th}$ **then**
- 18: $\mathcal{M}(t) \leftarrow$ connectivity-recovery state
- 19: Assign $r_{i_L}(t) \leftarrow \mathcal{R}_L$; assign non-leader agents to relay-oriented or search-oriented recovery
- 20: **else**
- 21: $\mathcal{M}(t) \leftarrow$ cooperative-encirclement state
- 22: Activate growth-based adaptive role assignment
- 23: **end if**
- 24: **end if**
- 25: **return** $\mathcal{M}(t), \{(\mathbf{p}_{T,i}^c(t), \mathbf{v}_{T,i}^c(t))\}_{i=1}^N, \{r_i(t)\}_{i=1}^N, i_L(t)$

4.2. Growth-Based Adaptive Role Assignment

Once cooperative encirclement is active, agents are assigned to functional roles according to accumulated growth evidence rather than fixed labels.

The growth state integrates observation quality, connectivity contribution, stability, and platform bias, thereby filtering transient sensing errors and topology fluctuations.

Leader selection uses accumulated leadership potential with switching margin and holding-time constraints. Relay agents preserve visual information paths, searchers recover weak observation/connectivity, and the remaining agents form the encirclement ring.

In Algorithm 2, the auxiliary variable $\tilde{g}_i(t)$ denotes the instantaneous growth drive before temporal smoothing. The relay score $\Psi_i^{\text{Re}}(t)$ evaluates whether agent i is suitable for maintaining visual information paths. Here, $d_{iL}(t) = \|\mathbf{p}_i(t) - \mathbf{p}_{i_L}(t)\|$ and $d_{iT}(t) = \|\mathbf{p}_i(t) - \mathbf{p}_{T,i}^c(t)\|$ denote the distances from agent i to the current leader and to its target reference, respectively. N_{Re} is the prescribed number of relay agents selected from the remaining agents according to the relay score. The variables $\tau_i^S(t)$ and $\sigma_i^S(t)$ denote the duration of poor observation/connectivity and the binary searcher-state flag,

respectively. The thresholds O_{low} , O_{rec} , C_{low}^T , and C_{rec}^T define the entry and recovery conditions for the searcher role, where $C_i^T(t)$ measures the local accessibility of target-related information. The sets $\mathcal{A}_{\text{rem}}$, $\mathcal{A}_{\text{Re}}$, $\mathcal{A}_{\text{res}}$, $\mathcal{A}_S$, and $\mathcal{A}_E$ denote the remaining agents after leader selection, selected relay agents, residual agents after relay selection, searchers, and encirclers, respectively.

Algorithm 2 Growth-based adaptive role assignment.

Require: $\mathcal{M}(t)$, $\mathcal{A}$, $i_L(t - \Delta t)$, $\{r_i(t - \Delta t)\}_{i=1}^N$, $\{g_i(t - \Delta t)\}_{i=1}^N$, $\{P_i(t - \Delta t)\}_{i=1}^N$

Ensure: $\{r_i(t)\}_{i=1}^N$, $i_L(t)$

```

1: if  $\mathcal{M}(t) \neq \mathcal{M}_{\text{CE}}$  then
2:    $\{r_i(t)\}_{i=1}^N \leftarrow \{r_i(t - \Delta t)\}_{i=1}^N$ ,  $i_L(t) \leftarrow i_L(t - \Delta t)$ 
3:   return  $\{r_i(t)\}_{i=1}^N$ ,  $i_L(t)$ 
4: end if
5: for  $i = 1$  to  $N$  do
6:    $\tilde{g}_i(t) \leftarrow f_g(O_i(t), C_i(t), S_i(t), H_i)$ 
7:    $g_i(t) \leftarrow \lambda_g g_i(t - \Delta t) + (1 - \lambda_g) \tilde{g}_i(t)$ 
8:    $P_i(t) \leftarrow f_L(g_i(t), O_i(t), C_i(t), S_i(t), H_i)$ 
9: end for
10:  $i_L^{\text{cand}} \leftarrow \arg \max_{i \in \mathcal{A}} P_i(t)$ ,  $\Delta P_L \leftarrow P_{i_L^{\text{cand}}}(t) - P_{i_L(t - \Delta t)}(t)$ 
11:  $i_L(t) \leftarrow \begin{cases} i_L^{\text{cand}}, & \Delta P_L \geq \Delta P_{\text{th}} \text{ and } T_L(t) \geq T_L^{\text{min}}, \\ i_L(t - \Delta t), & \text{otherwise.} \end{cases}$ 
12:  $r_{i_L(t)}(t) \leftarrow \mathcal{R}_L$ ,  $\mathcal{A}_{\text{rem}} \leftarrow \mathcal{A} \setminus \{i_L(t)\}$ 
13: for  $i \in \mathcal{A}_{\text{rem}}$  do
14:    $\Psi_i^{\text{Re}}(t) \leftarrow f_{\text{Re}}(C_i(t), d_{iL}(t), d_{iT}(t), O_i(t))$ 
15: end for
16:  $\mathcal{A}_{\text{Re}} \leftarrow \text{TopK}_{i \in \mathcal{A}_{\text{rem}}}(\Psi_i^{\text{Re}}(t), N_{\text{Re}})$ 
17:  $\mathcal{A}_{\text{res}} \leftarrow \mathcal{A}_{\text{rem}} \setminus \mathcal{A}_{\text{Re}}$ 
18: for  $i \in \mathcal{A}_{\text{res}}$  do
19:    $\tau_i^S(t) \leftarrow \begin{cases} \tau_i^S(t - \Delta t) + \Delta t, & O_i(t) < O_{\text{low}} \text{ and } C_i^T(t) < C_{\text{low}}^T, \\ 0, & \text{otherwise.} \end{cases}$ 
20:    $\sigma_i^S(t) \leftarrow \begin{cases} 1, & \tau_i^S(t) \geq T_S^{\text{in}}, \\ 0, & O_i(t) \geq O_{\text{rec}} \text{ or } C_i^T(t) \geq C_{\text{rec}}^T, \\ \sigma_i^S(t - \Delta t), & \text{otherwise.} \end{cases}$ 
21: end for
22:  $\mathcal{A}_S \leftarrow \{i \in \mathcal{A}_{\text{res}} \mid \sigma_i^S(t) = 1\}$ ,  $\mathcal{A}_E \leftarrow \mathcal{A}_{\text{res}} \setminus \mathcal{A}_S$ 
23:  $r_i(t) \leftarrow \mathcal{R}_{\text{Re}}, \forall i \in \mathcal{A}_{\text{Re}}; r_i(t) \leftarrow \mathcal{R}_S, \forall i \in \mathcal{A}_S; r_i(t) \leftarrow \mathcal{R}_E, \forall i \in \mathcal{A}_E$ 
24: return  $\{r_i(t)\}_{i=1}^N$ ,  $i_L(t)$ 

```

The output of this module is not a direct motion command, but a role label for each agent. These role labels are passed to the role-dependent motion-control module in Section 4.3, where they are converted into executable motion objectives. The complete procedure is summarized in Algorithm 2.

4.3. Role-Dependent Motion Control

The role-dependent motion-control module converts the task state and adaptive role assignment into physically feasible motion commands for the heterogeneous UAV-USV swarm. In contrast to a fixed formation controller, the proposed controller does not impose an identical motion objective on all agents. Instead, each agent generates its motion command according to its current biological role, locally accessible target reference, visual-neighbor topology, and platform-dependent dynamic constraints. This design allows target guidance, information relay, search recovery, and encirclement formation to be executed within a unified distributed control process.

At each control step, the current role selects a role-specific virtual velocity: target guidance for the leader, information-path maintenance for relays, exploration for searchers, and ring tracking for encirclers. Agent-wise target references ensure distributed information consistency.

For agent i , the heading-alignment term is obtained by averaging the unit heading vectors of its currently visible neighbors:

$$\bar{\mathbf{v}}_i(t) = \frac{1}{|\mathcal{N}_i^v(t)|} \sum_{j \in \mathcal{N}_i^v(t)} \begin{bmatrix} \cos \psi_j(t) \\ \sin \psi_j(t) \end{bmatrix}, \quad (24)$$

where $\mathcal{N}_i^v(t)$ denotes the visual-neighbor set of agent i . If no visual neighbor is available, $\bar{\mathbf{v}}_i(t)$ is set to zero. The term $k_{\text{align}} \bar{\mathbf{v}}_i(t)$ is added to the role-dependent desired velocity to improve local directional consistency without overriding the primary role objective.

Connectivity-maintenance and heading-alignment corrections are then added before bounded acceleration commands are generated under UAV/USV-specific limits.

Algorithm 3 summarizes the role-dependent controller.

Algorithm 3 Role-dependent motion control.

Require: Agent states $\{\mathbf{p}_i, \psi_i, v_i, \omega_i\}_{i=1}^N$, roles $\{r_i\}_{i=1}^N$, target references $\{(\mathbf{p}_{T,i}^c, \mathbf{v}_{T,i}^c)\}_{i=1}^N$, visual graph $\mathbf{A}_{\text{vis}}$, leader i_L , platform limits

Ensure: Control inputs $\{a_i^c, \alpha_i^c\}_{i=1}^N$ and updated states

- 1: **for** $i = 1$ to N **do**
- 2: Compute $\mathbf{r}_{iT} \leftarrow \mathbf{p}_i - \mathbf{p}_{T,i}^c$, $d_{iT} \leftarrow \|\mathbf{r}_{iT}\|$, $\mathbf{e}_{r,i} \leftarrow \mathbf{r}_{iT} / (d_{iT} + \varepsilon)$, $\mathbf{e}_{t,i} \leftarrow [-e_{r,i}^y, e_{r,i}^x]^\top$
- 3: **if** $r_i = \mathcal{R}_L$ **then**
- 4: $\mathbf{p}_L^d \leftarrow \mathbf{p}_{T,i}^c + \gamma_L R_E [-v_{T,i}^{c,y}, v_{T,i}^{c,x}]^\top / (\|\mathbf{v}_{T,i}^c\| + \varepsilon)$
- 5: $\mathbf{v}_{i,\text{role}}^{\text{ref}} \leftarrow k_L (\mathbf{p}_L^d - \mathbf{p}_i) + k_{LT} \mathbf{v}_{T,i}^c + k_{Lo} \mathbf{e}_{t,i}$
- 6: **else if** $r_i = \mathcal{R}_{\text{Re}}$ **then**
- 7: $\mathbf{p}_{\text{mid}} \leftarrow \lambda_L \mathbf{p}_{i_L} + (1 - \lambda_L) \mathbf{p}_{T,i}^c$
- 8: $\mathbf{v}_{i,\text{role}}^{\text{ref}} \leftarrow k_{Re} (\mathbf{p}_{\text{mid}} - \mathbf{p}_i) + k_{\text{conn}} \mathbf{v}_i^{\text{conn}} + k_{RT} \mathbf{v}_{T,i}^c$
- 9: **else if** $r_i = \mathcal{R}_S$ **then**
- 10: $\theta_i^S \leftarrow \omega_S t + \phi_i$, $\mathbf{p}_i^S \leftarrow \mathbf{p}_{T,i}^c + r_S [\cos \theta_i^S, \sin \theta_i^S]^\top$
- 11: $\mathbf{v}_{i,\text{role}}^{\text{ref}} \leftarrow k_{S1} (\mathbf{p}_i^S - \mathbf{p}_i) + k_{S2} (\mathbf{p}_{T,i}^c - \mathbf{p}_i)$
- 12: **else**
- 13: $\varphi_i^d \leftarrow 2\pi k_i / N_E$, $\mathbf{p}_i^E \leftarrow \mathbf{p}_{T,i}^c + R_E [\cos \varphi_i^d, \sin \varphi_i^d]^\top$
- 14: $\mathbf{v}_{i,\text{role}}^{\text{ref}} \leftarrow k_E (\mathbf{p}_i^E - \mathbf{p}_i) + k_o \mathbf{e}_{t,i} + k_{ET} \mathbf{v}_{T,i}^c$
- 15: **end if**
- 16: $\mathcal{N}_i^v \leftarrow \{j \in \mathcal{A} \setminus \{i\} \mid A_{\text{vis},ij} = 1\}$
- 17: Compute $\mathbf{v}_i^{\text{cm}}$ from pairwise visual-neighbor distance regulation, and compute $\bar{\mathbf{v}}_i$ from average neighbor heading alignment
- 18: $\mathbf{v}_i^{\text{des}} \leftarrow \mathbf{v}_{i,\text{role}}^{\text{ref}} + k_{\text{cm}} \mathbf{v}_i^{\text{cm}} + k_{\text{align}} \bar{\mathbf{v}}_i$
- 19: $s_i \leftarrow \|\mathbf{v}_i^{\text{des}}\|$, $\psi_i^{\text{ref}} \leftarrow \text{atan2}(v_{i,y}^{\text{des}}, v_{i,x}^{\text{des}})$, $e_{\psi,i} \leftarrow \text{wrap}(\psi_i^{\text{ref}} - \psi_i)$
- 20: $v_i^{\text{ref}} \leftarrow \min(v_{\chi_i}^{\text{max}}, s_i) \max(\eta_{\min}, \cos e_{\psi,i})$
- 21: $\omega_i^{\text{ref}} \leftarrow \text{sat}(k_{\psi}^{\text{ref}} e_{\psi,i}, -\omega_{\chi_i}^{\text{max}}, \omega_{\chi_i}^{\text{max}})$
- 22: $a_i^c \leftarrow \text{sat}(k_v (v_i^{\text{ref}} - v_i), a_{\chi_i}^{\min}, a_{\chi_i}^{\text{max}})$
- 23: $\alpha_i^c \leftarrow \text{sat}(k_{\omega} (\omega_i^{\text{ref}} - \omega_i), -\alpha_{\chi_i}^{\text{max}}, \alpha_{\chi_i}^{\text{max}})$
- 24: Apply velocity-boundary protection to a_i^c and α_i^c
- 25: Update x_i, y_i, ψ_i, v_i , and ω_i using the discrete-time second-order motion model
- 26: **end for**
- 27: **return** $\{a_i^c, \alpha_i^c\}_{i=1}^N$ and updated states

5. Experimental Setup

5.1. Simulation Environment

All simulations were implemented in MATLAB 2023a on a local workstation equipped with an Intel R12 12700H CPU, 16 GB of RAM, and an NVIDIA RTX 3070 GPU. Table 2 lists the environment and platform parameters.

Table 2. Simulation environment parameters.

Parameter	Value	Description
Environment size	$240 \times 240 \text{ m}^2$	Length $\times$ width
N_{UAV}	3	Number of UAVs
N_{USV}	5	Number of USVs
N	8	Total number of agents
$\mathbf{x}_T(0)$	$[40, -30, 1.2, 0.4]^\top$	Initial target state $[x_T, y_T, v_T^x, v_T^y]^\top$
$v_{\text{UAV}}^{\max}$	3.6 m/s	Maximum linear velocity of UAVs
$v_{\text{USV}}^{\max}$	2.0 m/s	Maximum linear velocity of USVs
$\omega^{\max}$	1.2 rad/s	Maximum angular velocity
$a_{\text{UAV}}^{c,\max}$	2.2 m/s^2	Maximum forward acceleration of UAVs
$a_{\text{UAV}}^{c,\min}$	-2.8 m/s^2	Maximum braking acceleration of UAVs
$a_{\text{USV}}^{c,\max}$	0.85 m/s^2	Maximum forward acceleration of USVs
$a_{\text{USV}}^{c,\min}$	-1.10 m/s^2	Maximum braking acceleration of USVs
$\alpha_{\text{UAV}}^{\max}$	3.8 rad/s^2	Maximum angular acceleration of UAVs
$\alpha_{\text{USV}}^{\max}$	1.10 rad/s^2	Maximum angular acceleration of USVs
R_{UAV}	38 m	Visual sensing range of UAVs
R_{USV}	30 m	Visual sensing range of USVs
ϕ_i	85°	Half field-of-view angle
R_{enc}	16 m	Desired encirclement radius
R_{close}	3.5 m	Close-distance threshold for encirclement evaluation
Γ_{th}	0.30	Target-information coverage threshold
S_{th}	0.70	Stable encirclement success threshold
$e_{r,\text{th}}$	7.0 m	Stable encirclement radius-error threshold
K_{hold}	12 steps	Required holding steps for stable encirclement

5.2. Algorithm Parameters

The key motion-control and role-evolution parameters are listed in Tables 3 and 4. To improve reproducibility, the evaluated ranges and selection rationale are provided for the parameters that most directly affect convergence, visual connectivity preservation, and role stability.

The experimental tuning was performed in a staged manner. The motion-control gains were tuned while keeping the role-evolution parameters fixed. The objective of this stage was to ensure that agents could follow their role-dependent reference motions without obvious oscillation or excessive control effort. The role-evolution parameters were tuned while using the selected motion-control gains. This stage focused on reducing transient role switching and maintaining sufficient target-information coverage. The combined parameter set was checked through preliminary paired-seed simulations under the same randomized initial perturbations used in the main experiments.

The final parameters were selected according to four criteria: stable encirclement success, radius error, target-information coverage, and role switching stability. Parameter settings that improved one metric but caused severe degradation in another metric were not used. For example, excessively large tracking gains accelerated convergence but increased

oscillatory motion, while overly small leader-switching margins improved responsiveness but caused frequent leader changes.

Table 3. Key motion-control hyperparameters.

Parameter	Value	Description	Evaluated Range	Rationale
k_v	1.80	Linear-velocity tracking gain	1.0–2.5	Balance tracking speed and smoothness
k_ω	2.40	Angular-velocity tracking gain	1.5–3.0	Reduce heading lag without angular oscillation
k_L	1.10	Leader target-tracking gain	0.6–1.4	Maintain target guidance and formation flexibility
k_{Re}	0.80	Relay guidance gain	0.4–1.2	Preserve visual relay paths without overconstraint
k_{align}	0.12	Local heading-alignment gain	0.05–0.25	Improve local consistency while avoiding swarm rigidity
T_{lost}	8.0 s	Maximum duration for short-term target prediction	4–12 s	Avoid stale target estimates after long target loss

The selected values were chosen through preliminary paired-seed simulations to balance convergence speed, control smoothness, and visual connectivity preservation.

In general, larger tracking gains accelerate convergence but may increase control effort and oscillatory motion. Smaller tracking gains improve smoothness but delay stable encirclement. A larger growth decay coefficient, switching margin, or holding time reduces frequent role changes, but may slow adaptation to visual-topology changes. The observation and connectivity weights determine the balance between target tracking and information preservation: excessive observation weighting tends to favor target-facing agents, whereas excessive connectivity weighting may delay spatial encirclement by overemphasizing relay behavior. Therefore, the selected values represent a trade-off among convergence speed, radius accuracy, visual connectivity preservation, and role stability.

Table 4. Key role-evolution hyperparameters.

Parameter	Value	Description	Evaluated Range	Rationale
η_g	1.10	Growth driving gain	0.7–1.5	Control response speed of role evidence
λ_g	0.30	Natural decay coefficient of growth state	0.1–0.5	Smooth role evolution and suppress transient changes
w_O	0.42	Weight of target observation quality	0.30–0.55	Emphasize target observation in role growth
w_C	0.28	Weight of visual connectivity contribution	0.15–0.40	Encourage information propagation and relay contribution
Δ_L	0.10	Leadership switching margin	0.05–0.20	Prevent frequent leader switching while retaining adaptivity
T_{hold}	6 steps	Minimum holding steps for leader switching	3–10 steps	Require persistent superiority before leader replacement
N_{Re}	2	Number of relay agents	1–3	Provide relay support without reducing encirclers excessively

The selected values were chosen to balance target observation, visual connectivity preservation, role stability, and heterogeneous role complementarity.

5.3. Evaluation Metrics

We evaluate the performance of each method using the following metrics:

- **Final success metric:** The composite encirclement score at the final simulation step.
- **Mean success score in the last 20 steps:** The average composite score during the final stage of each run.
- **Post-stabilization mean score:** The average composite score after stable encirclement is first reached.
- **Radius error:** The deviation between the achieved encirclement radius and the desired radius.
- **Stable encirclement time:** The time required for the team to first reach stable encirclement.
- **Stable reached rate:** The proportion of runs in which stable encirclement is successfully achieved.
- **Coverage ratio:** The fraction of agents with direct or relayed access to target information.
- **Visual connectivity:** The average strength of valid visual links in the team.
- **Geometric score:** A normalized score measuring radius accuracy and angular distribution.
- **Information score:** A normalized score measuring target-information availability within the team.
- **Role switches:** The total number of role changes during one run.
- **Leader switches:** The number of leader changes during one run.
- **Control effort:** The accumulated motion-control cost during the simulation.
- **Composite value index:** An overall normalized index combining convergence, radius accuracy, information availability, visual connectivity, and control effort.
- **Mean rank and best count:** Ranking-based indicators used to compare UAV-only, USV-only, and UAV-USV configurations across multiple metrics.

5.4. Statistical Significance Analysis

To verify whether the observed performance gains were statistically reliable, a paired Wilcoxon signed-rank test was conducted. Since all compared methods were evaluated under identical random seeds and perturbed initial conditions, the results form paired samples. The significance level was set to $p < 0.05$.

The test focuses on the comparison between GB-VCE and the static-leader baseline, because static-leader is the stronger baseline and reaches stable encirclement in most runs, whereas fixed-role fails to reach stable encirclement in all 100 runs. The tested metrics include final success metric, stable encirclement time, final radius error, post-stabilization target-information coverage, and mean information score.

6. Results and Performance Analysis

6.1. Computational-Time Analysis

The computational cost of the proposed GB-VCE framework was evaluated in terms of wall-clock simulation time. All simulations were executed on the workstation described above, and the reported values were obtained from 100 independent runs under identical experimental settings. In the strategy-comparison study, the average computational times of the fixed-role, static-leader, and GB-VCE models were 1.1371 ± 0.1591 s, 2.0811 ± 0.3374 s, and 1.5198 ± 0.2040 s per 200 s simulated mission, respectively. The corresponding average computation time of GB-VCE was approximately 1.518 ms per control step. In the heterogeneous-value study, the UAV-only, USV-only, and heterogeneous UAV-USV configurations required 1.9405 ± 0.3426 s, 1.4383 ± 0.2609 s, and 1.3390 ± 0.1754 s per run, respectively. In all evaluated cases, the average computation time per control step remained substantially below the adopted sampling interval of 0.2 s. These results indicate that the computational burden of GB-VCE is compatible with online execution under the tested simulation settings.

Table 5 summarizes the computational time of the evaluated simulation studies over 100 independent runs.

Table 5. Computational time of the evaluated simulation studies over 100 independent runs.

Simulation Study	Method/Configuration	Time Per Run (s)	Average Time Per Step (ms)
Strategy comparison	Fixed-Role	1.1371 ± 0.1591	1.136
	Static-Leader	2.0811 ± 0.3374	2.079
	Proposed GB-VCE	1.5198 ± 0.2040	1.518
Heterogeneous-value study	UAV-only	1.9405 ± 0.3426	1.939
	USV-only	1.4383 ± 0.2609	1.437
	UAV-USV	1.3390 ± 0.1754	1.338

Values are reported as mean $\pm$ standard deviation over 100 independent runs. Each run represents a 200 s simulated mission with a sampling interval of 0.2 s, corresponding to 1001 control steps. Wall-clock timing includes simulation and control computations but excludes figure generation and file-output operations.

6.2. Encirclement Performance Comparison

The proposed GB-VCE method was evaluated against fixed-role and static-leader baselines using 100 independent paired-seed runs. The use of 100 runs reduces the influence of random initial positions and target-motion perturbations and provides a sufficiently large sample for reliable statistical comparison. For each seed, all compared methods were evaluated under the same perturbed initial conditions, so the observed differences mainly reflect the coordination strategy rather than random variation.

Table 6 shows that GB-VCE achieves the strongest spatial encirclement performance. It obtains the highest final success metric (0.644 ± 0.060), the highest post-stabilization score (0.640 ± 0.009), and the lowest final radius error (1.040 ± 1.034 m). Compared with the static-leader baseline, GB-VCE reduces stable encirclement time from 68.29 ± 41.15 s to 40.65 ± 11.25 s and reaches stability in all 100 runs, indicating faster and more consistent convergence.

Table 6. Statistical comparison of encirclement performance over 100 independent paired-seed runs.

Metric	Fixed-Role	Static-Leader	Proposed GB-VCE
Final success metric $\uparrow$	0.271 ± 0.017 [0.267, 0.274]	0.598 ± 0.139 [0.570, 0.625]	0.644 ± 0.060 [0.632, 0.655]
Mean success score in last 20 steps $\uparrow$	0.336 ± 0.002 [0.335, 0.336]	0.544 ± 0.028 [0.539, 0.550]	0.634 ± 0.022 [0.629, 0.638]
Post-stabilization mean score $\uparrow$	0.336 ± 0.002 [0.335, 0.336]	0.543 ± 0.033 [0.536, 0.549]	0.640 ± 0.009 [0.638, 0.641]
Final radius error (m) $\downarrow$	4.831 ± 0.088 [4.814, 4.848]	2.851 ± 1.064 [2.642, 3.059]	1.040 ± 1.034 [0.837, 1.242]
Post-stabilization radius error (m) $\downarrow$	4.078 ± 0.067 [4.065, 4.091]	3.044 ± 0.429 [2.960, 3.128]	2.343 ± 0.297 [2.285, 2.402]
Mean radius error in last 20 steps (m) $\downarrow$	4.078 ± 0.067 [4.065, 4.091]	2.826 ± 0.741 [2.680, 2.971]	2.142 ± 0.371 [2.069, 2.214]
Stable encirclement time (s) $\downarrow$	N.R. (0/100)	68.29 ± 41.15 [60.10, 76.48] (97/100)	40.65 ± 11.25 [38.44, 42.85] (100/100)
Stable reached rate $\uparrow$	0.00	0.97	1.00

Values are mean $\pm$ standard deviation with 95% confidence interval in brackets. N.R. denotes that stable encirclement was not reached. For stable encirclement time, statistics are computed only over successful runs; parentheses indicate success count out of 100. $\uparrow$ indicates that a higher value is better, whereas $\downarrow$ indicates that a lower value is better. Bold values identify the best result for each metric according to this direction.

Table 7 further explains the source of this improvement. GB-VCE achieves the highest final coverage ratio (0.726 ± 0.066) and post-stabilization coverage (0.692 ± 0.018), and substantially improves the mean information score in the last 20 steps from 0.490 ± 0.044 for static-leader to 0.716 ± 0.016 . Although its role switches are higher than in fixed-role, fixed-role never reaches stable encirclement; compared with static-leader, GB-VCE reduces role switches from 273.96 ± 54.53 to 130.27 ± 27.71 .

Table 7. Statistical comparison of information coverage, visual connectivity, role adaptation, and control smoothness over 100 independent paired-seed runs.

Metric	Fixed-Role	Static-Leader	Proposed GB-VCE
Final coverage ratio $\uparrow$	0.625 ± 0.000 [0.625, 0.625]	0.705 ± 0.219 [0.662, 0.748]	0.726 ± 0.066 [0.713, 0.739]
Post-stabilization coverage $\uparrow$	0.605 ± 0.005 [0.604, 0.606]	0.616 ± 0.050 [0.607, 0.626]	0.692 ± 0.018 [0.688, 0.695]
Mean visual connectivity $\uparrow$	0.020 ± 0.001 [0.020, 0.020]	0.085 ± 0.032 [0.079, 0.092]	0.085 ± 0.018 [0.082, 0.089]
Mean geometric score in last 20 steps $\uparrow$	0.511 ± 0.005 [0.510, 0.512]	0.645 ± 0.073 [0.631, 0.660]	0.728 ± 0.038 [0.720, 0.735]
Mean information score in last 20 steps $\uparrow$	0.254 ± 0.003 [0.253, 0.254]	0.490 ± 0.044 [0.482, 0.499]	0.716 ± 0.016 [0.713, 0.719]
Mean saturated visual score in last 20 steps $\uparrow$	0.168 ± 0.006 [0.167, 0.170]	0.458 ± 0.139 [0.430, 0.485]	0.368 ± 0.056 [0.357, 0.379]
Role switches $\downarrow$	52.16 ± 45.53 [43.24, 61.08]	273.96 ± 54.53 [263.27, 284.65]	130.27 ± 27.71 [124.84, 135.70]
Leader switches $\downarrow$	0.00 ± 0.00 [0.00, 0.00]	0.00 ± 0.00 [0.00, 0.00]	8.14 ± 2.95 [7.56, 8.72]
Control smoothness $\downarrow$	0.691 ± 0.086 [0.674, 0.707]	1.342 ± 0.105 [1.322, 1.363]	1.083 ± 0.085 [1.067, 1.100]

Values are mean $\pm$ standard deviation with 95% confidence interval in brackets. $\uparrow$ indicates that a higher value is better, whereas $\downarrow$ indicates that a lower value is better. The best value for each metric is bold according to its optimization direction. Although fixed-role has fewer role switches and lower smoothness, it never reaches stable encirclement (see Table 6).

Table 8 reports the paired Wilcoxon signed-rank test results between GB-VCE and the static-leader baseline. The results show that GB-VCE achieves statistically significant improvements in the main performance dimensions. Specifically, GB-VCE improves the final success metric from 0.5978 to 0.6435 ($p = 6.81 \times 10^{-3}$), reduces the stable encirclement time from 68.29 s to 40.85 s ($p = 1.22 \times 10^{-13}$), and decreases the final radius error from 2.8505 m to 1.0397 m ($p = 1.45 \times 10^{-12}$). In addition, GB-VCE significantly improves post-stabilization target-information coverage and mean information score. These results confirm that the observed performance gains are statistically supported rather than caused by random variations.

Table 8. Statistical significance test between GB-VCE and the static-leader baseline.

Metric	GB-VCE	Static-Leader	Improvement	<i>p</i> -Value
Final success metric $\uparrow$	0.6435	0.5978	+7.65%	6.81×10^{-3}
Stable encirclement time (s) $\downarrow$	40.85	68.29	+40.18%	1.22×10^{-13}
Final radius error (m) $\downarrow$	1.0397	2.8505	+63.52%	1.45×10^{-12}
Post-stabilization coverage $\uparrow$	0.6919	0.6164	+12.24%	2.84×10^{-15}
Mean information score in last 20 steps $\uparrow$	0.7158	0.4903	+45.99%	3.90×10^{-18}

The Wilcoxon signed-rank test was conducted using paired-seed simulation results. For benefit-type metrics, a positive improvement means that GB-VCE achieves a higher value. For cost-type metrics, including stable encirclement time and radius error, a positive improvement means that GB-VCE achieves a lower value. $\uparrow$ denotes a benefit-type metric for which higher values are better, and $\downarrow$ denotes a cost-type metric for which lower values are better.

Representative Dynamics and Statistical Dispersion

Figures 4–8 visualize the same trend from time-series, distributional, and snapshot perspectives.

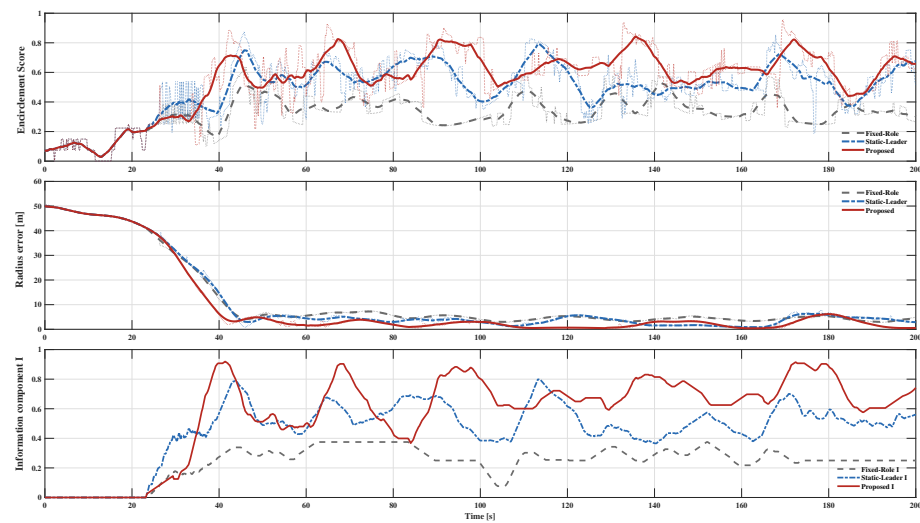


Figure 4. Temporal evolution of representative encirclement-performance metrics for the fixed-role, static-leader, and proposed GB-VCE methods over paired-seed simulations.

The performance curves in Figure 4 show that GB-VCE rises faster and stabilizes at a higher score than the baselines. The fixed-role method remains limited by its inability to redistribute sensing and relay functions, whereas the static-leader method improves convergence but is more sensitive to initial visibility and leader placement.

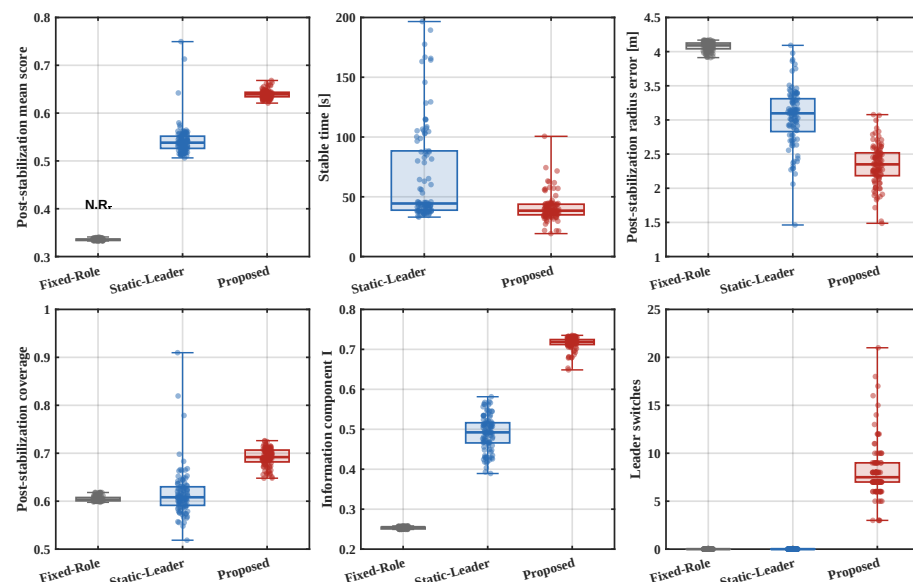


Figure 5. Box-scatter distributions of final success score, stable encirclement time, radius error, information coverage, visual connectivity, and control-related metrics for the compared methods.

The box-scatter results in Figure 5 indicate that GB-VCE has both higher median performance and smaller dispersion, especially for final success and radius-error metrics. This supports the conclusion that growth-based role evolution improves not only average performance but also run-to-run reliability.

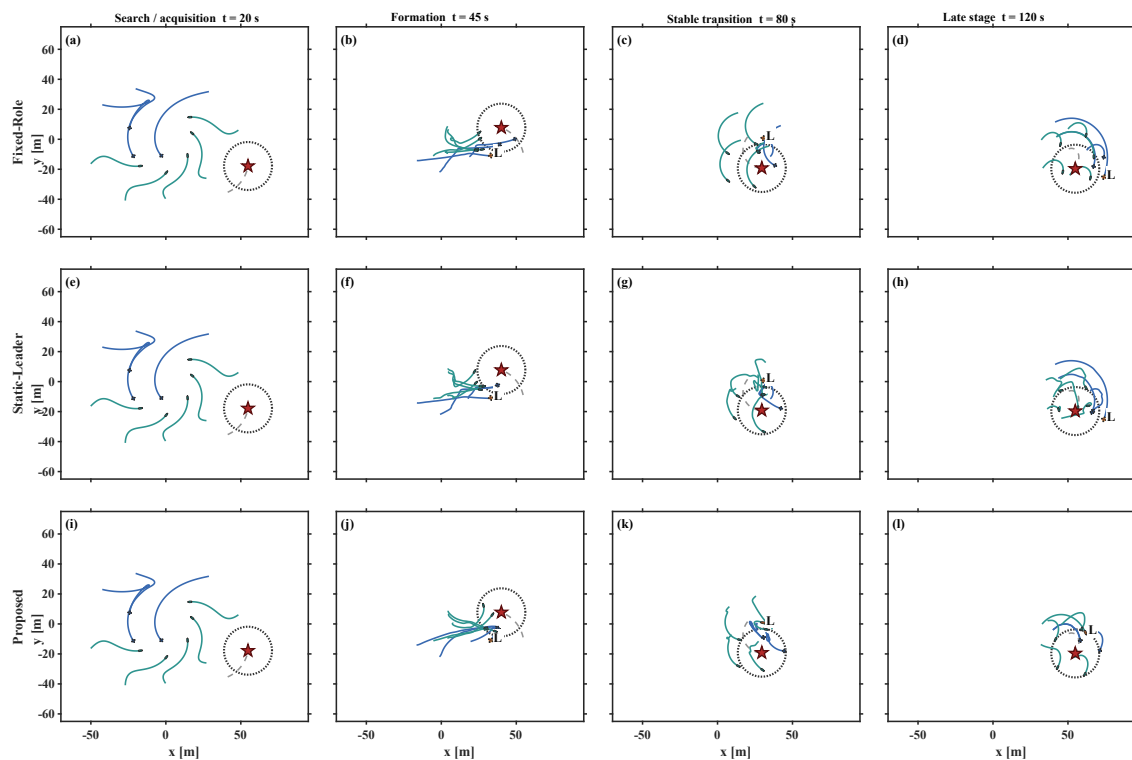


Figure 6. Representative encirclement snapshots of fixed-role (a–d), static-leader (e–h), and GB-VCE (i–l) methods at different mission stages.

Figure 6 qualitatively shows the formation process, while the quantitative results in Tables 6 and 7 confirm the smaller final radius error, higher geometric score, and better information coverage of GB-VCE. GB-VCE forms a tighter and more symmetric encirclement ring around the target than the fixed-role and static-leader baselines.

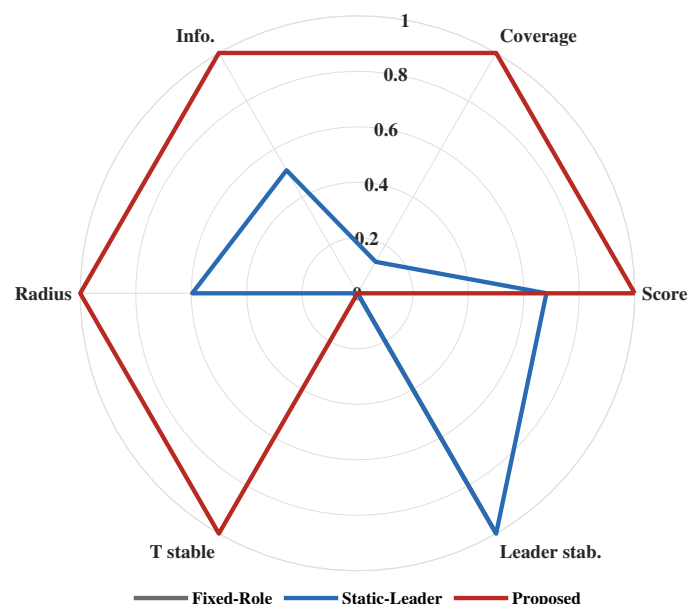


Figure 7. Normalized radar comparison of the fixed-role, static-leader, and proposed GB-VCE methods across success score, radius accuracy, information coverage, role stability, visual connectivity, and control smoothness.

The normalized radar comparison in Figure 7 illustrates the trade-off among success score, radius accuracy, information coverage, role stability, and control smoothness. GB-

VCE provides the most balanced profile, which is important for heterogeneous UAV-USV deployment where a single metric cannot capture mission quality.

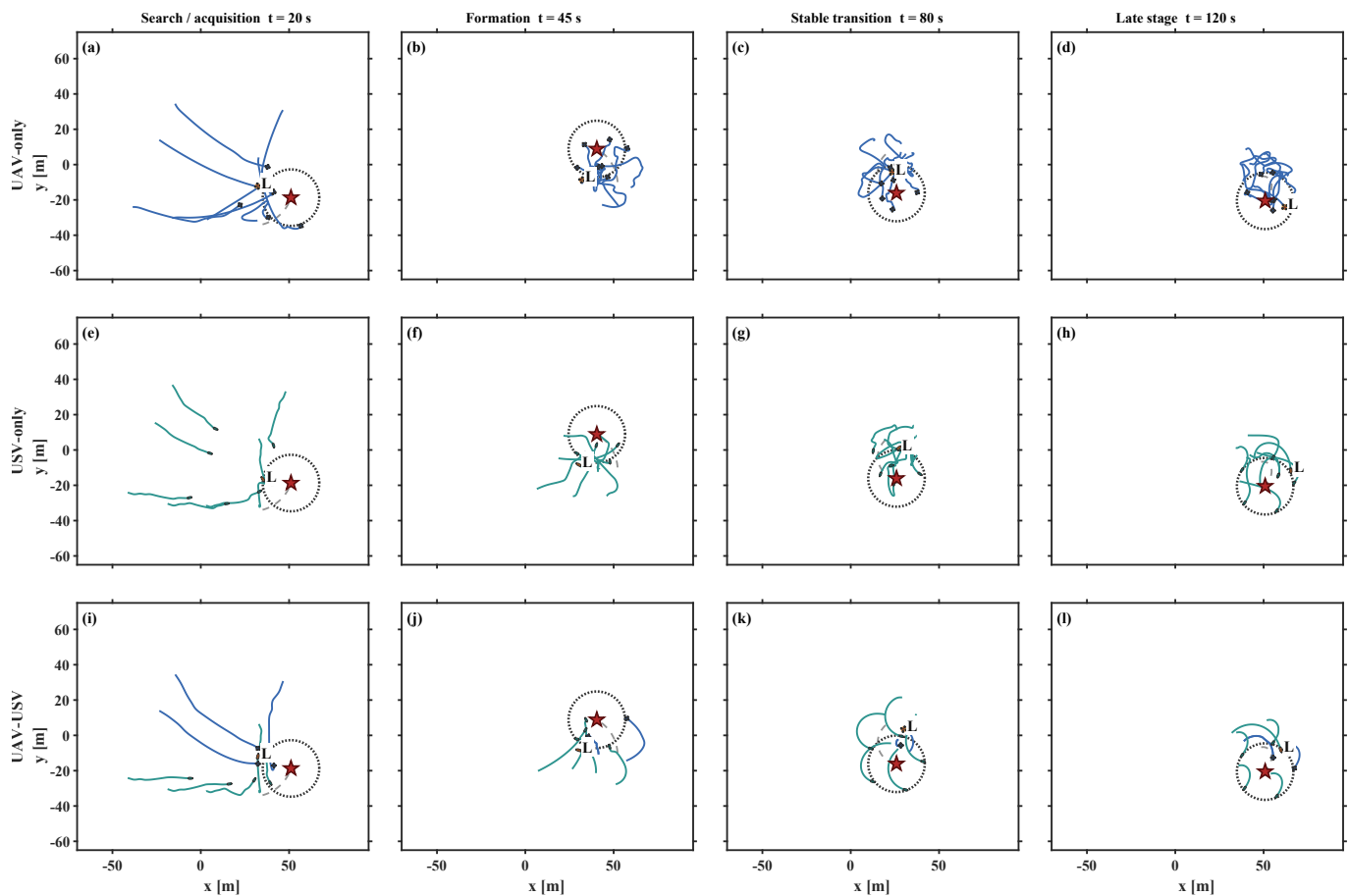


Figure 8. Recovery-to-encirclement snapshots showing visual relay formation and target-information propagation. UAV-only (a–d), USV-only (e–h) and UAV-USV (i–l).

6.3. Heterogeneous Value Verification

To isolate the value of platform heterogeneity, the proposed role-evolution mechanism was also tested with UAV-only, USV-only, and UAV-USV teams under the same number of agents and paired seeds.

Figure 8 compares the recovery-to-encirclement process for UAV-only, USV-only, and heterogeneous UAV-USV teams. UAV-only teams are fast and agile, bridging information gaps quickly but with less stable coverage. USV-only teams maintain stable encirclement but converge slowly. The heterogeneous UAV-USV team combines both strengths: UAVs act as fast information relays, while USVs provide a stable encirclement backbone, achieving both rapid convergence and robust, complete coverage.

Figure 9 compares the average performance of UAV-only, USV-only, and UAV-USV teams in terms of convergence time, post-stabilization score, radius error, information coverage, visual connectivity, control effort, and composite value. The dynamic curves in Figure 9 show that the UAV-only team converges quickly but with higher control activity, while the USV-only team is smoother but slower. The heterogeneous UAV-USV team achieves the most favorable balance between convergence, radius accuracy, and stable behavior.

The bar comparison in Figure 10 confirms that heterogeneity is most beneficial for task-critical encirclement accuracy and composite value rather than for every individual

metric. This is consistent with the biological interpretation: mixed functional capabilities improve group-level performance through complementarity.

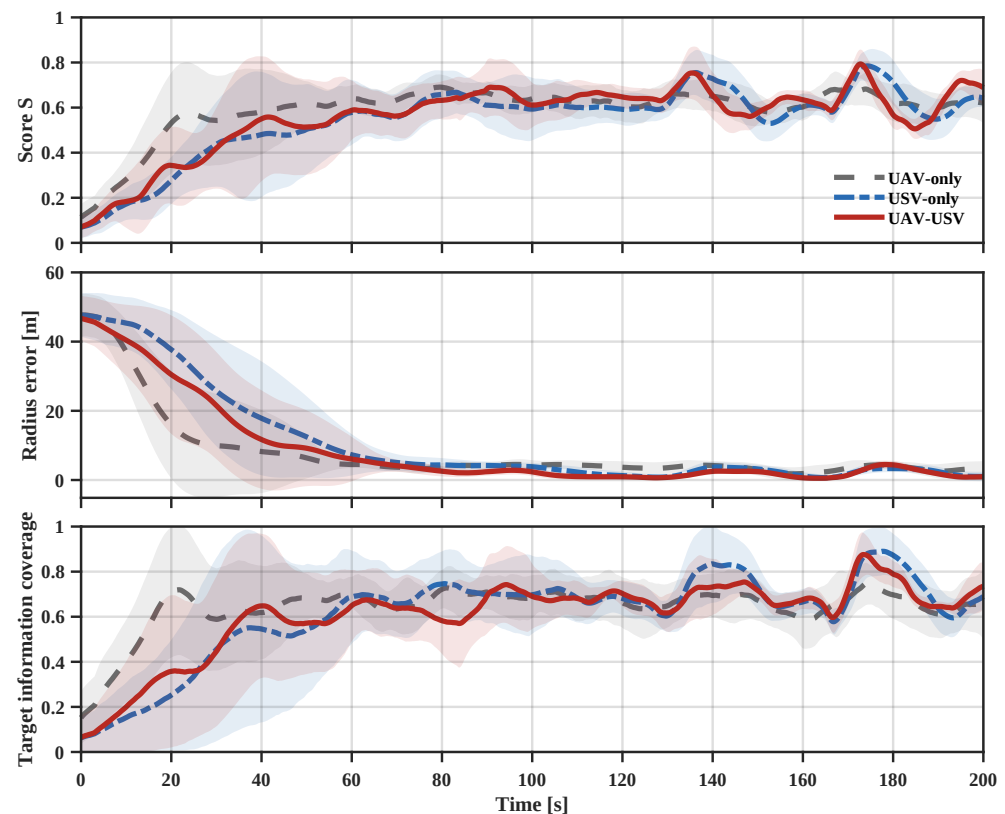


Figure 9. Temporal performance comparison of UAV-only, USV-only, and heterogeneous UAV-USV teams.

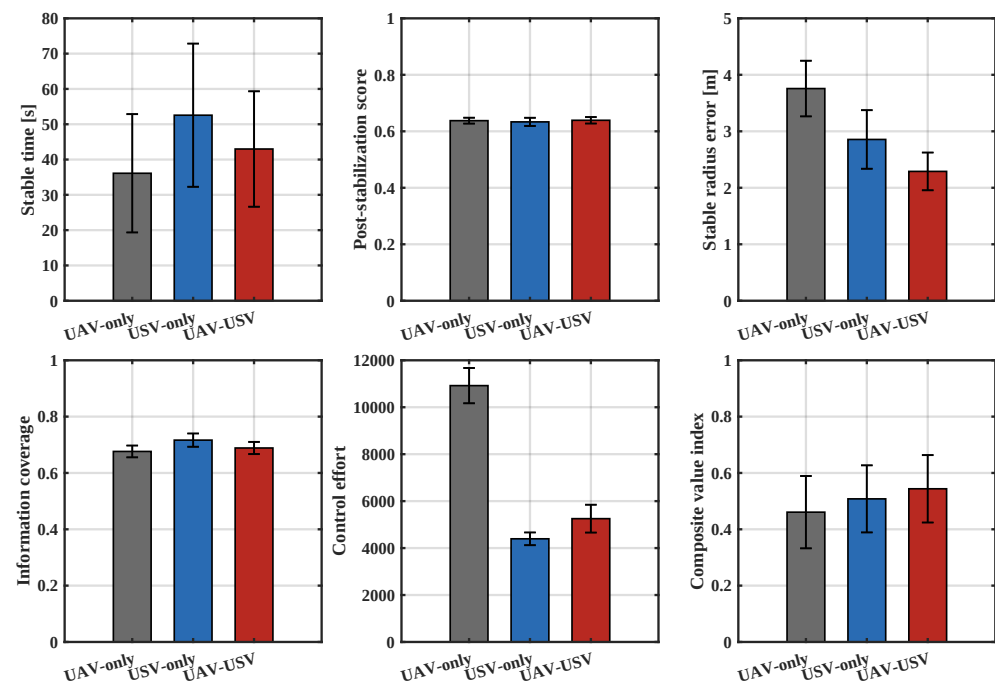


Figure 10. Statistical comparison of homogeneous and heterogeneous platform configurations over multi-seed simulations.

Table 9 quantifies this effect. UAV-USV obtains the highest post-stabilization score and the lowest stable radius error among the three platform configurations. Although UAV-only

reaches stable encirclement faster and USV-only yields lower control effort, the heterogeneous team provides the best overall trade-off.

Table 9. Multi-seed statistical comparison of different platform configurations.

Metric	UAV-Only	USV-Only	UAV-USV	Best	Direction
Stable encirclement time [s]	36.11 ± 16.76	52.56 ± 20.28	42.98 ± 16.35	UAV-only	Lower
Post-stabilization score	0.6377 ± 0.0104	0.6333 ± 0.0145	0.6390 ± 0.0114	UAV-USV	Higher
Stable radius error [m]	3.757 ± 0.492	2.856 ± 0.519	2.290 ± 0.334	UAV-USV	Lower
Information coverage	0.6764 ± 0.0209	0.7166 ± 0.0236	0.6886 ± 0.0215	USV-only	Higher
Mean visual connectivity	0.1423 ± 0.0422	0.0982 ± 0.0295	0.0847 ± 0.0192	UAV-only	Higher
Control effort	10,920.4 ± 751.5	4394.6 ± 270.7	5253.3 ± 592.8	USV-only	Lower
Composite value index	0.4608 ± 0.1283	0.5080 ± 0.1191	0.5438 ± 0.1197	UAV-USV	Higher
Mean rank	2.17	2.05	1.78	UAV-USV	Lower
Best count	28	28	44	UAV-USV	Higher
Success rate [%]	100	100	100	All	Higher

The values are reported as mean ± standard deviation over 100 independent random seeds. Bold values indicate the best result for each metric. The composite value index is calculated from normalized convergence, encirclement accuracy, information availability, visual connectivity, and control-effort-related metrics.

The gain is calculated relative to the best homogeneous baseline under the same random seed, providing a conservative comparison of heterogeneous complementarity. Figure 11 shows that the UAV-USV team consistently improves the most task-critical metric, radius error, relative to homogeneous baselines. The conservative per-seed best-baseline comparison should therefore be interpreted as a strict test: even under this reference, the heterogeneous system retains a clear advantage in encirclement accuracy.

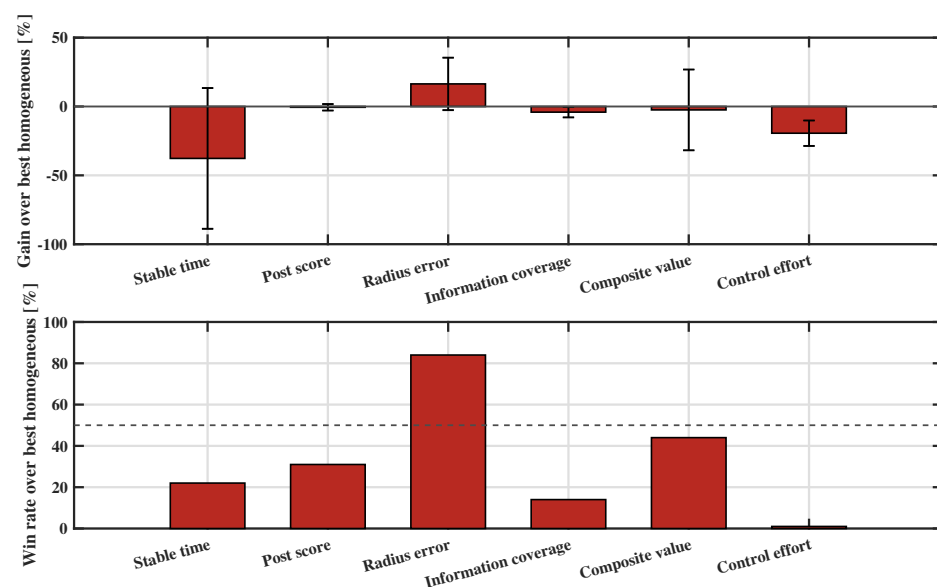


Figure 11. Performance gain of the heterogeneous UAV-USV team over homogeneous baselines.

Role Complementarity in the Heterogeneous Team

Figure 12 reveals a clear emergent division of labor. UAVs spend most of the stable phase as relay agents, using their mobility and observation flexibility to maintain visual information paths. USVs spend most of the stable phase as encircling agents, exploiting their smoother motion to maintain the surrounding ring. This role complementarity explains why the heterogeneous team obtains the lowest radius error and highest composite value: UAVs support information flow and fast adjustment, while USVs provide persistent geometric enclosure.

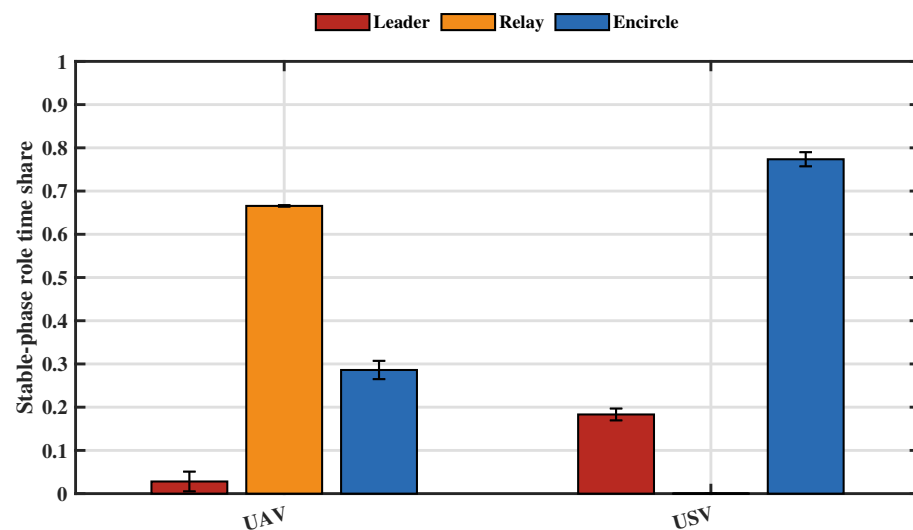


Figure 12. Stable-phase role allocation shares of UAVs and USVs in the heterogeneous team.

Overall, the experiments verify that GB-VCE improves encirclement by coupling biomimetic role growth with visual connectivity preservation. The method achieves better spatial accuracy, more reliable target-information propagation, smoother role differentiation, and stronger heterogeneous complementarity than the evaluated baselines. The organization and definitions of the simulation data supporting the reported statistical results are provided in Appendix A.

7. Conclusions and Future Work

This paper proposed GB-VCE, a wolf-pack-inspired distributed encirclement framework for heterogeneous UAV–USV teams operating under visual connectivity constraints. The method translates biological role differentiation into a two-layer robotic coordination architecture: the upper layer accumulates local growth evidence to select leader, relay, searcher, and encircler roles, while the lower layer converts the selected role into bounded motion commands for heterogeneous platforms. By using visual links and relay relations as coordination resources, GB-VCE avoids dependence on centralized fusion or persistent all-to-all communication.

Simulation results over 100 paired-seed runs demonstrate that GB-VCE improves both convergence and stable encirclement quality. Compared with fixed-role and static-leader baselines, the proposed method achieves the highest final success metric, the lowest final radius error, faster stable encirclement, and higher target-information coverage. The heterogeneous verification further shows that UAV–USV cooperation provides a better balance than UAV-only or USV-only teams: UAVs mainly emerge as relay agents that maintain information flow, while USVs mainly serve as encircling agents that stabilize the surrounding geometry. These results confirm that biomimetic role growth can transform platform heterogeneity into measurable cooperative performance gains.

The current study still has several limitations. The validation is simulation-based, and the target dynamics, sensing uncertainty, communication delay, localization error, environmental disturbances, and aerodynamic and hydrodynamic effects are simplified. In addition, the current framework considers planar motion and single-target encirclement. Future work will introduce a high-fidelity air–sea simulation environment, hardware-in-the-loop validation, and physical UAV–USV experiments. Three-dimensional UAV motion, multi-target encirclement, and more realistic sensing, communication, and environmental models will also be investigated.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/biomimetics11070513/s1>. Table S1: Strategy Comparison All Seed Metrics. Table S2: Strategy Comparison Significance Tests. Table S3: Strategy Comparison Metric Definitions. Table S4: Heterogeneous Study All Trials. Table S5: Heterogeneous Gain All Seeds. Table S6: Heterogeneous Role Complementarity.

Author Contributions: Conceptualization, J.X. and S.K.; methodology, J.X.; software, J.X.; validation, S.K., J.X. and J.Y.; formal analysis, J.X. and S.K.; investigation, F.W., Y.L. and J.S.; resources, J.Y.; data curation, F.W. and J.S.; writing—original draft preparation, J.X. and S.K.; writing—review and editing, S.K.; visualization, S.K.; supervision, J.S.; project administration, J.Y.; funding acquisition, J.Y. and S.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant 62573007, Grant 62433021, Grant 62533001, and 62533014, in part by the Beijing Natural Science Foundation under Grant 4242038.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The simulation data supporting the findings of this study are described in Appendix A and provided in the accompanying Supplementary Materials. The supplementary files include per-seed results for the strategy-comparison and heterogeneous-value studies, statistical-test outputs, metric definitions, heterogeneous-gain and role-complementarity data, and computational-time records. Additional information is available from the corresponding authors upon reasonable request.

Conflicts of Interest: Yingnan Li and Fang Wu were employed by the company SPIC Nuclear Energy Co. Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Dual-Use Research of Concern Statement: Current research is limited to cooperative robots, which is beneficial for swarm intelligence researches and does not pose a threat to public health or national security. Authors acknowledge the dual use potential of the research involving autonomous robot control technologies and confirm that all necessary precautions have been taken to prevent potential misuse. As an ethical responsibility, authors strictly adhere to relevant national and international laws regarding DURC. Authors advocate for responsible deployment, ethical considerations, regulatory compliance, and transparent reporting to mitigate misuse risks and foster beneficial outcomes.

Appendix A. Simulation Data Supporting the Statistical Analysis

The simulation data were generated from 100 paired random seeds for each evaluated method or platform configuration. Each run represented a 200 s simulated mission with a sampling interval of 0.2 s, corresponding to 1001 simulation steps. Under the same random seed, the compared methods or platform configurations were evaluated using matched initial-condition and target-motion perturbations.

Two groups of simulation data were generated. The strategy-comparison dataset contains the results of the fixed-role, static-leader, and proposed GB-VCE methods.

The heterogeneous-value dataset contains the results of the UAV-only, USV-only, and heterogeneous UAV–USV configurations. The recorded variables include encirclement performance, convergence time, radius error, target-information coverage, visual connectivity, role transitions, control effort, heterogeneous gain, and computational time.

The Supplementary Data files and their correspondence to the results reported in the main text are summarized in Table A1.

The complete per-seed data are provided in machine-readable CSV format. The strategy-comparison dataset contains 300 records, corresponding to 100 seeds and three methods. The heterogeneous-value dataset also contains 300 records, corresponding to 100 seeds and three platform configurations. Summary statistics, confidence intervals, significance-test results, and heterogeneous-gain measures reported in the main text were calculated from these per-seed records.

Table A1. Supplementary Data files supporting the reported simulation results.

File	Content
S1_StrategyComparison_AllSeedMetrics.csv	Per-seed performance results for the fixed-role, static-leader, and proposed GB-VCE methods over 100 paired random seeds.
S2_StrategyComparison_SignificanceTests.csv	Paired statistical-test results for the principal performance metrics, including sample sizes and <i>p</i> -values.
S3_StrategyComparison_MetricDefinitions.csv	Definitions, units, and evaluation directions of the principal metrics used in the strategy-comparison study.
S4_HeterogeneousStudy_AllTrials.csv	Per-seed performance results for the UAV-only, USV-only, and heterogeneous UAV–USV configurations.
S5_HeterogeneousGain_AllSeeds.csv	Per-seed performance gains of the heterogeneous UAV–USV configuration relative to the best homogeneous configuration.
S6_HeterogeneousRoleComplementarity.csv	Per-seed UAV and USV role-time shares for the leader, relay, and encircler roles.

Table A2. Principal data fields used in the statistical analyses.

Field	Unit	Description
Seed	–	Random-seed index used to generate the matched simulation condition.
Method/Configuration	–	Evaluated control method or platform configuration.
FinalSuccessMetric	–	Composite encirclement score at the final simulation step.
StableEncirclementTime	s	First time at which the prescribed stable-encirclement conditions were continuously satisfied.
RadiusError	m	Deviation between the achieved and desired encirclement radii.
TargetInformationCoverage	–	Fraction of agents with direct or relayed access to target information.
MeanVisualConnectivity	–	Mean visual connectivity measure of the agent team.
RoleSwitches	–	Total number of role transitions during one simulation run.
LeaderSwitches	–	Total number of leader replacements during one simulation run.
ControlEffort	–	Accumulated motion-control cost over the simulated mission.
CompositeValueIndex	–	Normalized aggregate index combining convergence, accuracy, information availability, visual connectivity, and control effort.

References

- Chen, Z.; Chen, X.; Liu, Z.; Huang, F. Monocular vision-based target enclosing control for USVs with FOV constraint and uncertain dynamics. *IEEE Trans. Ind. Inform.* **2026**, *22*, 4681–4692.
- Chen, X.; Huang, F.; Pan, Y.-J.; Chen, Z. Bearing-only-based cooperative target enclosing control for multiple uncrewed surface vehicles with unknown dynamics and sideslip. *IEEE J. Ocean. Eng.* **2025**, *50*, 1015–1029.

3. Mnih, V.; Kavukcuoglu, K.; Silver, D.; Rusu, A.A.; Veness, J.; Bellemare, M.G.; Graves, A.; Riedmiller, M.; Fidjeland, A.K.; Ostrovski, G. Human-level control through deep reinforcement learning. *Nature* **2015**, *518*, 529–533.
4. Van Hasselt, H.; Guez, A.; Silver, D. Deep reinforcement learning with double Q-learning. In Proceedings of the 30th AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–17 February 2016; Volume 30, pp. 2094–2100.
5. Orr, J.; Dutta, A. Multi-agent deep reinforcement learning for multi-robot applications: A Survey. *Sensors* **2023**, *23*, 3625.
6. Busoniu, L.; Babuska, R.; De Schutter, B. Multi-agent reinforcement learning: An overview. *Innov. Multi-Agent Syst. Appl.* **2010**, *310*, 183–221.
7. Lowe, R.; Wu, Y.; Tamar, A.; Harb, J.; Abbeel, P.; Mordatch, I. Multi-agent actor-critic for mixed cooperative-competitive environments. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 6379–6390.
8. Rashid, T.; Samvelyan, M.; De Witt, C.S.; Farquhar, G.; Foerster, J.; Whiteson, S. QMIX: Monotonic value function factorization for deep multi-agent reinforcement learning. In Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018; Volume 80, pp. 4295–4304.
9. Foerster, J.; Farquhar, G.; Afouras, T.; Nardelli, N.; Whiteson, S. Counterfactual multi-agent policy gradients. In Proceedings of the 32nd AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018; Volume 32, pp. 2974–2982.
10. Müller, M.; Hutter, M. Cooperative vision-based navigation for heterogeneous robot teams. *Rob. Auton. Syst.* **2019**, *117*, 233–245.
11. Pesce, E.; Montana, G. Learning multi-agent coordination through connectivity-driven communication. *Mach. Learn.* **2023**, *112*, 483–514.
12. Shetty, A.; Hussain, T.; Gao, G.X. Decentralized connectivity maintenance for multi-robot systems under motion and sensing uncertainties. *Navigation* **2023**, *70*, 552.
13. Ning, Z.; Xie, L. A survey on multi-agent reinforcement learning and its application. *J. Autom. Intell.* **2024**, *3*, 73–91.
14. Wang, D.; Deng, H.; Pan, Z. Multirobot coordination with deep reinforcement learning in complex environments. *Expert Syst. Appl.* **2021**, *180*, 115128.
15. Gable, T.D.; Johnson-Bice, S.M.; Homkes, A.T.; Windels, S.K.; Bump, J.K. Differential provisioning roles, prey size, and prey abundance shape the dynamic feeding behavior of gray wolves. *Commun. Biol.* **2023**, *6*, 1045.
16. Goeckner, A.; Sui, Y.; Martinet, N.; Li, X.; Zhu, Q. Graph neural network-based multi-agent reinforcement learning for resilient distributed coordination of multi-robot systems. In Proceedings of the 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Abu Dhabi, United Arab Emirates, 14–18 October 2024; pp. 5732–5739.
17. Cai, Y.; Li, H.; Hong, J.; Xu, P.; Cheng, H.; Zhu, X.; Hu, B.; Hao, Z.; Fan, Z. VG-Swarm: A vision-based gene regulation network for UAVs swarm behavior emergence. *IEEE Robot. Autom. Lett.* **2023**, *8*, 1175–1182.
18. Shi, G.; Ishat-E-Rabban, M.; Bonner, G.; Tokekar, P. Fast k-connectivity restoration in multi-robot systems for robust communication maintenance: Algorithmic and learning-based solutions. *Auton. Robot.* **2025**, *49*, 34.
19. Xie, J.; Zhong, Y.; Cui, Y. Advances and future directions in flocking control models for multi-agent systems. In Proceedings of the 2025 5th International Conference on Automation Control, Nanjing, China, 6–8 June 2025; pp. 477–482.
20. Yan, R.; Su, J. Mechanism of the nonequilibrium phase transition in self-propelled particles with alignment. *Phys. Rev. Res.* **2025**, *7*, 33082.
21. Sampaio, E.; Sridhar, V.H.; Francisco, F.A.; Weitekamp, C.A.; Romanczuk, P.; Gingins, S.; Thünken, T.; Bshary, R. Multidimensional social influence drives leadership and composition-dependent success in octopus–fish hunting groups. *Nat. Ecol. Evol.* **2024**, *8*, 2072–2084.
22. Oh, K.-K.; Park, M.-C.; Ahn, H.-S. A survey of multi-agent formation control. *Automatica* **2015**, *53*, 424–440.
23. Cao, Y.; Yu, W.; Ren, W.; Chen, G. An overview of recent progress in the study of distributed multi-agent coordination. *IEEE Trans. Ind. Inform.* **2013**, *9*, 427–438.
24. Ge, X.; Han, Q.-L.; Zhang, X.-M.; Ding, D. Distributed coordination control of multi-agent systems under intermittent sampling and communication: A comprehensive survey. *Sci. China Inf. Sci.* **2025**, *68*, 151201.
25. Feng, Z.; Dong, X.; Hu, G.; Lyu, J. *Resilient Cooperative Control and Optimization of Multi-Agent Systems*; Academic Press: Cambridge, MA, USA, 2025.
26. Wang, J.; Deng, X.; Guo, J.; Zeng, Z. Resilient consensus control for multi-agent systems. *Sensors* **2023**, *23*, 2904.
27. Ma, J.; Lu, H.; Xiao, J.; Zeng, Z.; Zheng, Z. Multi-robot target encirclement control with collision avoidance via deep reinforcement learning. *J. Intell. Robot. Syst.* **2020**, *99*, 371–386.
28. Zhang, T.; Li, Q.; Zhang, C.-S.; Liang, H.-W.; Li, P.; Wang, T.M.; Li, S.; Zhu, Y.-L.; Wu, C. Multi-target encirclement with collision avoidance via deep reinforcement learning. In Proceedings of the 2022 International Conference on Robotics and Automation (ICRA); IEEE: Piscataway, NJ, USA, 2022; pp. 6900–6906.

29. Chung, S.-J.; Paranjape, A.A.; Dames, P.; Shen, S.; Kumar, V. A survey on aerial swarm robotics. *IEEE Trans. Robot.* **2018**, *34*, 837–855.
30. Qu, X.; Li, C.; Jiang, S.; Liu, G.; Zhang, R. Multi-agent reinforcement learning-based cooperative encirclement control of autonomous surface vehicles against multiple targets. *J. Mar. Sci. Eng.* **2025**, *13*, 1558.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.