

Article

Visual-Attention-Based Neighbor Selection for Artificial-Potential-Field UAV Formation Control

Miao Yu 

School of Mathematics and Statistics, Xidian University, Xi'an 710071, China; myu@xidian.edu.cn

Abstract

The development of Unmanned Aerial Vehicle (UAV) formation technology is rapid, and formation flying under complex conditions has also received more attention. However, neighbor selection without reliable radio communication remains challenging because fixed-radius or fixed-topology methods may process redundant neighbor states. Based on this, this paper designs two visual-attention-based neighbor-selection algorithms, namely Visual Attention Potential Field (VAPF) and Cluster Visual Attention Potential Field (CVAPF). Both algorithms use a zoom-lens visual attention rule to retain informative neighbors before the artificial-potential-field control input is evaluated. The algorithm VAPF selects informative UAV-level neighbors for the APF controller and supports aggregation behavior in simulation. Algorithm CVAPF extends the selection rule to clusters through a dual layer communication architecture, which improves synchronization while gathering formations. In the tested redundant sensing scene, BOIDS and Optimized-flocking process 43.38 and 15.20 neighbors on average, whereas VAPF and CVAPF reduce the values to 6.31 and 3.60 while preserving the formation behavior observed in the simulations.

Keywords: formation control; distributed artificial potential field; visual attention mechanism; neighbor selection

1. Introduction

Unmanned Aerial Vehicle (UAV) formation control is important for cooperative sensing and operation in complex environments [1–4]. Representative applications include photogrammetry [5], humanitarian response [6], and reconnaissance or surveillance [7]. At the algorithmic level, multi-agent consensus provides a basic framework for coordinating many UAVs through local state exchange [8,9]. In this paper, the target is to support APF-driven formation behavior while reducing the number of neighbor states processed by each UAV when reliable radio communication is unavailable.

Formation control in complex environments has been studied from distributed control, multi-UAV coordination, and real-flight perspectives [10–13]. However, these settings are difficult when communication infrastructure is weak or unavailable. Vision-assisted sensing offers a practical way to estimate neighboring UAV states and support autonomous formation flight; for example, 3D UAV pose estimation can provide target position and attitude information [14], and dense-environment flight demonstrates the value of distributed local control [15]. For this reason, the neighbor-selection rule becomes a central design issue: a UAV should process enough visual neighbors for APF-driven aggregation and synchronization, but avoid processing redundant observations that do not improve the controller.



Academic Editor: Sangman Moh

Received: 29 May 2026

Revised: 8 July 2026

Accepted: 13 July 2026

Published: 14 July 2026

Copyright: © 2026 by the author.

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.

The research objective of this paper is to design a visual-attention-based neighbor-selection topology for artificial-potential-field UAV formation control. The central research question is whether a UAV can suppress angularly redundant visual neighbors before controller evaluation while still retaining the visually informative neighbors needed for aggregation and synchronization. The main challenges are therefore summarized as follows: (1) selecting informative neighbors without processing all visible UAVs, (2) maintaining formation behavior when radio-based fixed topologies are unavailable, and (3) improving motion order after local aggregation without requiring every UAV to process all visible neighbors.

To address these challenges, this paper designs a visual attention neighbor-selection topology for APF-based formation control. The significance of this design is that each UAV can retain visually informative neighbors for aggregation and synchronization while reducing the number of neighbor states that must be processed. Under the Visual Attention Potential Field (VAPF) algorithm, UAVs select neighbors through visual attention, and the APF controller then uses the selected neighbor states to drive formation motion. Because pure aggregation can still lead to locally disordered trajectories after UAVs approach one another, this paper further introduces the Cluster Visual Attention Potential Field (CVAPF) algorithm. CVAPF uses cluster-level visual attention, a two-layer communication architecture, and cluster merging to improve synchronization without requiring every UAV to process all visible neighbors.

The contribution of this paper is as follows.

- (1) A leader flight planning model is designed so that UAV.1 tracks preset velocities along a reference trajectory and provides a moving reference for the formation.
- (2) Two attention-based neighbor-selection algorithms are developed for APF-based formation control. VAPF uses zoom-lens visual attention to select informative UAV-level neighbors before APF control evaluation. CVAPF extends this idea to cluster-level attention and uses a two-layer communication architecture with cluster merging to reduce redundant neighbor processing at the cluster level.
- (3) Geometric boundary conditions are discussed for collision avoidance and communication maintenance under the stated assumptions. These conditions are used as practical design constraints for the attention-generated topology rather than as a full stability proof.

Modeling level: The proposed algorithms are developed at the formation control layer, where each UAV updates its motion from relative neighbor states selected by the visual attention rule. The UAV motion is therefore represented by a kinematic point-mass model, and visible neighbor positions and velocities are treated as available state inputs to the controller. This abstraction allows the analysis to concentrate on the central question of this paper: how the attention-generated topology changes the number and relevance of neighbor states processed before the artificial-potential-field input is evaluated. Deployment-level sensing errors, communication imperfections, and low-level multirotor dynamics are left for future implementation-oriented studies.

The rest of this paper is organized as follows. Section 2 reviews related work. Section 3 introduces the preliminaries. Section 4 presents the proposed method, including leader flight planning, VAPF, CVAPF, and the boundary conditions. Section 5 reports simulation results and discussion, including the comparative neighbor processing and attention ratio analyses. Section 6 concludes the paper.

2. Related Work

This section summarizes the literature most directly related to the proposed neighbor-selection topology. It first reviews flocking and artificial-potential-field methods that

motivate the control component, then reviews metric, topology-based, and visual attention neighbor-selection models that motivate the perception component.

Flocking and artificial-potential-field models provide the most relevant methodological basis for this work. Classical BOIDS (bird-oid) flocking [16] and visual movement models [17,18] show that collective motion can emerge from local rules. Optimized-flocking of autonomous drones further shows that collision avoidance, velocity alignment, sensing delay, and motion constraints must be considered together in real aerial robot swarms [19]. In parallel with these flocking models, artificial-potential-field (APF) methods provide a controller-level way to convert relative position or target information into attractive and repulsive virtual forces. This idea has been widely used for autonomous trajectory generation, local navigation, and obstacle avoidance [19–21]. APF-based methods may suffer from local minima when target attraction and obstacle repulsion cancel in obstacle-rich planning environments. We therefore treat this issue as an APF limitation outside the present focus on inter-UAV neighbor selection before potential-field evaluation. These studies justify the potential-field controller used in this paper, but they do not directly solve how a vision-based UAV should suppress angularly redundant neighbors before the controller is evaluated.

The remaining gap is how to choose interaction neighbors when the topology is generated from visual perception rather than from radio links or a fixed metric neighborhood. Metric models select neighbors within a distance threshold [22,23], and UAV formations often use predefined or directly constructed communication topologies [24]. However, these rules can become inefficient or unavailable when radio communication is not reliable. Visual models provide another route: visual collective-behavior frameworks [25], studies of individual and collective vision [26], and zoom-lens/spotlight attention models [27] show that agents can process only selected visual sectors. This paper therefore relates most closely to BOIDS and optimized drone flocking [16,19] on the control side, and to visual attention models [27] on the perception side. The specific gap addressed here is the lack of a neighbor-selection topology that combines potential-field formation control with attention-based suppression of redundant visual neighbors.

Recent large-scale aerial-swarm planning and robot-swarm assembly studies further show that scalability, resilience, and low per-agent processing load are central requirements for practical swarm systems [21,28,29]. Existing orderly-flight strategies include preset formation assignment [20] and trajectory planning for each UAV [30,31]. These studies are complementary to the present work: they address formation planning, assignment, or swarm assembly, whereas this paper focuses on the visual neighbor-selection topology used before the potential-field controller is evaluated.

3. Preliminaries

This section introduces the mathematical notation and basic mechanisms used by the proposed algorithms. The goal is to make the later leader-planning, VAPF, and CVAPF equations readable without repeatedly redefining common graph, potential-field, and visual attention concepts.

The section is organized into three parts. First, graph notation defines the neighbor sets used by the distributed topology; second, the artificial-potential-field components define attraction, velocity alignment, and repulsion; third, the visual attention rule explains how redundant visual neighbors are removed before the controller is evaluated.

3.1. Knowledge of Graph Theory

Graph notation is used to describe which UAV states are processed by each distributed controller. In this paper, the graph is not assumed to be a fixed radio topology; it can be

generated by the visual attention selection rule. Distributed communication in formation requires the establishment of communication topology graph $\mathcal{G}$. Consider a directed communication graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$, $\mathcal{V} = \{1, 2, 3, \dots, n\}$ as a set of nodes, $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is set of edges (directed edges), and adjacency matrix $\mathcal{A} = (a_{ij})_{n \times n}$ with a non-negative element (this paper does not consider the weight of each edge). $N_i = \{j \in \mathcal{V} | (j, i) \in \mathcal{E}, i \neq j\}$ is the set of parent nodes of i . $\|\cdot\|$ represents the Euclidean norm.

3.2. Artificial Potential Field

This subsection defines the potential-field components used after VAPF and CVAPF have selected their neighbors and explains why these components are arranged in distance-dependent regions. Existing APF-based navigation and obstacle-avoidance studies commonly construct control inputs from attractive or repulsive virtual potentials [32–34]. In this paper, the same gradient-based idea is used for inter-UAV formation interaction: selected neighboring UAVs act as moving potential-field sources, and the follower UAV evaluates only the sources retained by the visual attention topology.

The four potential-field regions in Figure 1 are designed to reproduce the main local interaction roles used in flocking and formation control. At long range, the gravitational field provides cohesion by attracting separated UAVs toward selected neighbors. At intermediate range, the black-hole field acts as a shaped attraction term that strengthens gathering before the UAV reaches the velocity alignment zone; it should be understood as a designed transition potential rather than as a separate physical phenomenon. Near the desired interaction range, the homogeneous field penalizes velocity differences and therefore plays the role of flocking alignment. At very short range, the anti-gravity field dominates the other terms so that separation and collision avoidance have priority over aggregation.

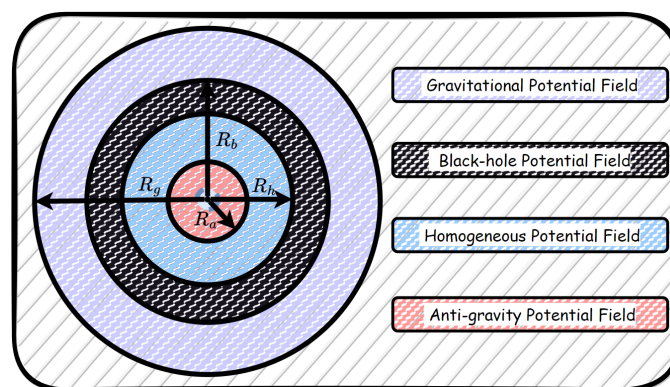


Figure 1. Artificial potential field generated by an individual UAV. The arrows indicate radial distance directions from the source.

The following equations define these regions in a continuous order from long-range attraction to short-range repulsion. The first region is the long-range cohesion region: its quadratic potential produces a force proportional to the relative displacement, so a selected neighbor pulls the UAV toward the formation when the distance is large. When $\|x - x_{goal}\| \in [R_b, R_g]$, the gravitational potential field $Y_g(x)$ is represented as

$$Y_g(x) = \frac{1}{2}k_g\|x - x_{goal}\|^2, \quad (1)$$

where k_g is the coefficient related to the strength of the gravitational potential field; x_{goal} represents the position of the potential field source; R_b represents the radius of the black-

hole potential field range; and R_g represents the radius of the gravitational potential field range. The corresponding force $f_g(x)$ is expressed as

$$f_g(x) = -\nabla Y_g(x) = -k_g(x - x_{goal}). \quad (2)$$

Thus, the role of this region is only cohesion; velocity matching and collision avoidance are activated by the inner regions below.

The second region is a transition attraction region. Its purpose is to keep the attractive force active while preventing an abrupt switch between long-range cohesion and the inner alignment region. When $\|x - x_{goal}\| \in [R_h, R_b]$, the black-hole potential field $Y_b(x)$ is represented as

$$Y_b(x) = -\frac{1}{2}k_b(R_b - \|x - x_{goal}\|)^2 + \frac{1}{2}k_g\|x - x_{goal}\|^2, \quad (3)$$

where k_b is the coefficient related to the strength of the black-hole potential field; R_h represents the radius of the homogeneous potential field range. The corresponding force $f_b(x)$ is expressed as

$$\begin{aligned} f_b(x) &= -\nabla Y_b(x) \\ &= -k_b[R_b \frac{x - x_{goal}}{\|x - x_{goal}\|} - (x - x_{goal})] - k_g(x - x_{goal}). \end{aligned} \quad (4)$$

This shaped term is named the black-hole field in this manuscript because it increases the tendency of separated UAVs to gather before they enter the alignment range; it is a designed APF component, not a physical black-hole model.

After a neighbor enters the desired interaction range, position attraction is no longer the only objective. The homogeneous field is defined on velocity difference so that nearby UAVs reduce relative speed and move coherently. When $\|x - x_{goal}\| \in [R_a, R_h]$, the homogeneous potential field $Y_h(v)$ is represented as

$$Y_h(v) = \frac{1}{2}k_h\|v - v_{goal}\|^2, \quad (5)$$

where k_h is the coefficient related to the strength of the homogeneous potential field; v_{goal} represents the velocity of the potential field source; and R_a represents the radius of the anti-gravity potential field range. The corresponding force $f_h(v)$ is expressed as

$$f_h(v) = -\nabla Y_h(v) = -k_h(v - v_{goal}). \quad (6)$$

This term corresponds to the alignment rule in flocking models: the UAV does not simply move closer to the neighbor, but adjusts its velocity toward the neighbor velocity.

The innermost region is reserved for safety. When the distance is below R_a , the repulsive potential grows rapidly as the distance decreases, so short-range separation dominates the attraction and alignment terms. When $\|x - x_{goal}\| \in (0, R_a]$, the anti-gravity potential field $Y_a(x)$ is represented as

$$Y_a(x) = \frac{1}{2}k_a\left(\frac{1}{\|x - x_{goal}\|} - \frac{1}{R_a}\right)^2, \quad (7)$$

where k_a is the coefficient related to the strength of the anti-gravity potential field. The corresponding force $f_a(x)$ is expressed as

$$f_a(x) = -\nabla Y_a(x) \quad (8)$$

$$= k_a \left\{ \frac{1}{(x - x_{goal})^3} - \frac{1}{R_a \|x - x_{goal}\| (x - x_{goal})} \right\}.$$

The above APF design does not claim to eliminate the general local minimum problem of APF path planning. Local minima usually arise in environments with static or dynamic obstacles, where target attraction and obstacle repulsion can cancel and trap the robot before it reaches the goal. The present manuscript does not model external obstacles or global path planning; it studies how visual attention selects inter-UAV neighbors before the formation control input is evaluated. Therefore, the local minimum issue is treated as an APF limitation outside the current obstacle-free neighbor-selection setting.

3.3. Visual Attention Selects Neighbors

This subsection defines the visual attention rule used to construct the neighbor set before the potential-field controller is evaluated. The rule keeps visually informative targets and masks angularly redundant targets that are covered by already selected perceived cones.

The zoom-lens model is an attention mechanism model that originates from the principle of camera imaging. When the camera focuses on an object, it also defocus around the object. Similar to the human eye tracking moving targets, this attention model is consistent with biological intuition. Focusing can help the eye capture the state of objects, and defocusing can filter out other unimportant information. The Sensed angle and Perceived angle in this paper is a conical field of view in 3D space. Sensed angle is the range we focus on, which precisely wraps around the target; Perceived angle is the range within which we defocus, which means that other targets within that range will not be focused (Perceived angle = $AR \cdot$ Sensed angle). Therefore, in the UAV formation, each UAV selects neighbors according to the following steps (refer to Figure 2): Step 1: Select the UAV with the largest Sensed angle as the neighbor node, while defocusing within the Perceived angle. Step 2: Select UAVs outside of the Perceived angle. Step 3: Repeat the above process until no UAV can be selected. Step 4: Output the selected UAVs as the neighbor set N_i used by the potential-field controller. In Figure 2, this step represents the final retained visual targets after redundant targets inside perceived cones have been discarded. This attention mechanism has some properties in cluster systems.

- (1) The average distance of individual nearest neighbors in a cluster is called cluster sparsity (CP); the total number of individuals in a cluster is called cluster size (CS). When AR is a fixed value, as CP and CS increase, the number of neighbors chosen by individuals will also increase. We call this phenomenon scale-dependence. When $AR_i = distance_i$ (distance from individual to node i), as CP and CS increase, the number of neighbors selected by individuals will saturate within a certain range. We call this phenomenon scale-free, which is a method that conforms to biological visual processing information: biological processing information has boundaries.
- (2) Individuals tend to focus on isolated nodes or parts with low CP . This phenomenon is consistent with biological behavior, with less attention given to individuals in clusters and more attention given to isolated individuals, just as wolves tend to prey on sheep that leave the flock.

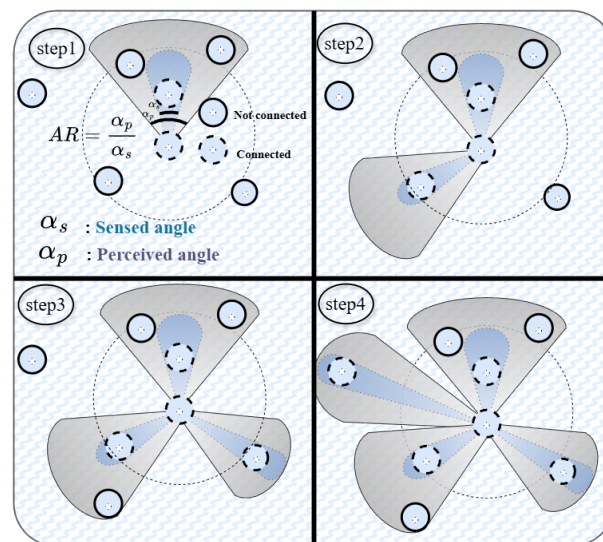


Figure 2. Visual attention mechanism for selecting neighboring nodes. Light-blue sectors denote sensed angles, gray sectors denote perceived defocusing angles, black dashed outlines mark selected neighboring UAVs, and the four panels show the sequential selection process.

4. Proposed Method

This section presents the proposed leader-planning, VAPF, and CVAPF methods. The objective is to show how visual attention neighbor selection is coupled with artificial-potential-field control at both the individual-UAV and cluster levels.

The section proceeds as follows. First, a leader trajectory-tracking model is introduced to provide the formation reference. Second, VAPF is defined as an individual-level attention-potential-field method. Third, CVAPF extends the same idea to clusters through cluster attention, two-layer communication, and merging. Finally, geometric boundary conditions are discussed as practical constraints for collision avoidance and communication maintenance.

- (1) Develop a virtual UAV algorithm for the formation leader, which can manage the speed and trajectory of the formation. (*Subsection A*)
- (2) Propose two visual-attention-based neighbor-selection algorithms (VAPF and CVAPF) for generating the APF communication topology. (*Subsections B and C*)
- (3) Calculate the corresponding boundary conditions used as practical constraints for the attention-generated neighbor topology. (*Subsection D*)

4.1. Leader Flight Planning

The leader-planning model defines the moving reference followed by the rest of the formation. It is not intended to solve global path planning; instead, it converts a given reference trajectory and desired velocity sequence into the leader's discrete-time motion. Starting from the formation leader (UAV.1) and ending at the target position, we generate a continuous trajectory, denoted as L (we will not discuss how to plan the trajectory). Divide the trajectory into n_l segments to obtain $n_l + 1$ points, denoted as $\{l^{(k)}\}_{k=1}^{n_l+1}$. In the k_l -th segment of the trajectory, we preset the expected velocity of UAV.1 as $v_L^{(k+1)}$. We represent UAV.1 motion modeling as

$$x_1[n+1] = x_1[n] + \min\{\|v_1[n]\|, \|v_{max}\|\} \frac{v_1[n]}{\|v_1[n]\|} \Delta t, \quad (9)$$

$$v_1[n+1] = v_1[n] + \min\{\|u_1[n]\|, \|u_{max}\|\} \frac{u_1[n]}{\|u_1[n]\|} \Delta t, \quad (10)$$

$$u_1[n+1] = \sum_{k=1}^{n_l} \left\{ k_L^{(k+1)} (v_1[n] - v_L^{(k+1)}) - \tau(\|(l^{(k+1)} - x_1[n])(x_1[n] - l^{(k)})\|) \right\}, \quad (11)$$

where $x_1[n]$ represents the position of UAV.1; $v_1[n]$ represents the speed of UAV.1; $\|v_{max}\|$ represents the maximum speed of the UAVs; $\|u_{max}\|$ represents the maximum output of the controller; $k_L^{(k)}$ represents the potential field intensity generated by the k -th point on the trajectory L . $\tau(\cdot)$ satisfies the following: $\tau(x) = 1, x > 0$; otherwise, $\tau(x) = 0$. The function $\tau(\cdot)$ is therefore an activation indicator: only the trajectory segment that contains the current leader position contributes to the leader controller. The gains $k_L^{(k)}$ determine how strongly UAV.1 tracks the desired velocity on each segment, and they can be increased when faster velocity correction is required or decreased when smoother motion is preferred.

4.2. Visual Attention Artificial Potential Field (Algorithm VAPF)

VAPF applies the visual attention rule directly to individual UAVs. After the selected neighbors are obtained, the controller separates them by distance so that attraction, velocity alignment, and repulsion are activated in different distance regions. Select attention ratio as AR , neighboring node N_i according to the attention mechanism above, and divide N_i into the following sets based on $\|x_j - x_i\|$.

$$\begin{aligned} N_i^g &= \{j | j \in N_i, \|x_j - x_i\| \in [R_h^{ji}, R_g^{ji}]\}, \\ N_i^h &= \{j | j \in N_i, \|x_j - x_i\| \in [R_a^{ji}, R_h^{ji}]\}, \\ N_i^a &= \{j | j \in N_i, \|x_j - x_i\| \in (0, R_a^{ji}]\}, \end{aligned} \quad (12)$$

where N_i^g represents the set of neighboring nodes of UAV. i in the gravitational black-hole potential field; N_i^h represents the set of neighboring nodes of UAV. i in the homogeneous potential field; and N_i^a represents the set of neighboring nodes of UAV. i in the anti-gravity potential field. For UAV. $i, i \neq 1$, their motion models can be represented as

$$x_i[n+1] = x_i[n] + \min\{\|v_i[n]\|, \|v_{max}\|\} \frac{v_i[n]}{\|v_i[n]\|} \Delta t, \quad (13)$$

$$v_i[n+1] = \gamma \tilde{v}_i[n+1] + \frac{(1-\gamma)}{\min\{n, n_h\} + 1} \sum_{k=0}^{\min\{n, n_h\}} \tilde{v}_i[n-k], \quad (14)$$

$$\tilde{v}_i[n+1] = \tilde{v}_i[n] + \min\{\|u_i[n]\|, \|u_{max}\|\} \frac{u_i[n]}{\|u_i[n]\|}, \quad (15)$$

$$u_i[n] = \{k_1 f_g^{[n]}(i) + k_2 f_h^{[n]}(i) + k_3 f_a^{[n]}(i)\}, \quad (16)$$

where $x_i[n]$ represents the position of UAV. i ; $v_i[n]$ represents the speed of UAV. i ; $\gamma \in [0, 1]$ represents the filtering coefficient; n_h represents the number of historical speed records; the denominator $\min\{n, n_h\} + 1$ is the number of velocity records in the finite-history average; k_1, k_2, k_3 are non negative constants, and satisfy $k_1 + k_2 + k_3 = 1$; $f_g^{[n]}(i)$ represents the force acting on UAV. i in the gravitational black-hole potential field; $f_h^{[n]}(i)$ represents the force acting on UAV. i in the homogeneous potential field; $f_a^{[n]}(i)$ represents the force acting on UAV. i in the anti-gravity potential field; and $f_g^{[n]}(i), f_h^{[n]}(i), f_a^{[n]}(i)$ represent the following. The three weights k_1, k_2, k_3 allocate control authority among attraction, velocity alignment, and collision

avoidance. In practice, k_3 should be large enough to make repulsion dominant at short range, while k_1 and k_2 are tuned to balance aggregation speed and velocity synchronization.

$$f_g^{[n]}(i) = \sum_{j \in N_i^g} \left\{ -k_b^{ji} \left[-R_b^{ji} \frac{\Delta x_{ji}[n-1]}{\|\Delta x_{ji}[n-1]\|} + \Delta x_{ji}[n-1] \right] \right. \\ \left. * \tau(R_b^{ji} - \|\Delta x_{ji}[n-1]\|) + k_g^{ji} \Delta x_{ji}[n-1] \right\}, \quad (17)$$

where $\Delta x_{ji}[n] = x_j[n] - x_i[n]$; R_g^{ji} represents the radius of the gravitational black-hole potential field generated by UAV.j on UAV.i; R_h^{ji} represents the radius of the homogeneous potential field generated by UAV.j on UAV.i; R_b^{ji} represents the radius of the black-hole potential field generated by UAV.j on UAV.i; k_b^{ji} represents the intensity coefficient of the black-hole potential field generated by UAV.j on UAV.i; and k_g^{ji} represents the intensity coefficient of the gravitational black-hole potential field generated by UAV.j on UAV.i.

$$f_h^{[n]}(i) = \sum_{j \in N_i^h} k_h^{ji} \Delta v_{ji}[n-1], \quad (18)$$

where R_a^{ji} represents the radius of the anti-gravity potential field generated by UAV.j on UAV.i; k_h^{ji} represents the intensity coefficient of the homogeneous potential field generated by UAV.j on UAV.i; and $\Delta v_{ji}[n] = v_j[n] - v_i[n]$.

$$f_a^{[n]}(i) = \sum_{j \in N_i^a} \frac{\xi_{ji}[n-1]}{\Xi_i[n-1]} \left\{ -\frac{1}{R_a^{ji} \|\Delta x_{ji}[n-1]\| \Delta x_{ji}[n-1]} + \frac{1}{(\Delta x_{ji}[n-1])^3} \right\}, \quad (19)$$

where $\xi_{ji}[n]$ represents UAV.j for UAV.i attention coefficient; $\frac{\xi_{ji}[n]}{\Xi_i[n]}$ represents UAV.j for UAV.i attention weight. The calculation formula for $\Xi_i[n]$ is as follows

$$\Xi_i[n] = \sum_{j \in N_i^a} \xi_{ji}[n] \\ = \sum_{j \in N_i^a} k_a^{ji} \frac{\exp\{\|\Delta v_{ji}[n]\| \cos \angle(\Delta v_{ji}[n], \Delta x_{ji}[n])\}}{\|\Delta x_{ji}[n]\| + \varepsilon}. \quad (20)$$

The attention weight of UAV.i regarding UAV.j is positively correlated with the relative velocity projected onto their line of sight and negatively correlated with their distance; k_a^{ji} represents the intensity coefficient of the anti-gravity potential field generated by UAV.j on UAV.i; and $\varepsilon > 0$ is a constant. The notation $\angle(a, b)$ denotes the angle between two vectors, so $\cos \angle(\Delta v_{ji}, \Delta x_{ji})$ extracts the component of relative velocity along the line of sight. The angular part of visual relevance has already been handled by the zoom-lens selection rule, which removes targets hidden inside previously selected perceived cones. After this angular filtering, the remaining question is whether a selected neighbor needs a short-range response: a closer UAV is easier to observe and more directly related to collision avoidance, while a UAV with larger approaching velocity along the line of sight requires stronger short-range response. This is why the attention coefficient uses distance and line-of-sight relative velocity as the minimal first-order state variables. Other measurable neighbor cues, such as bearing angle, bearing-rate change, distance-change rate, acceleration, or state-estimation confidence, could also be considered in future sensing-level models. They are not included here because they would require additional measurement assumptions and tuning beyond the formation control layer considered in this paper. Here ε prevents

division by zero when two UAVs are very close, and it also limits the numerical gain of the attention coefficient.

Remark 1. *The attention coefficient should be interpreted as a lightweight relevance score rather than a complete visual perception model. It is consistent with the biological intuition that nearer and faster-approaching objects receive more attention, while keeping the controller implementable with relative position and velocity estimates.*

4.3. Cluster Visual Attention Potential Field (Algorithm CVAPF)

CVAPF is introduced because individual-level attention alone can still leave locally disordered trajectories after UAVs aggregate. The cluster version treats a group of nearby and kinematically compatible UAVs as one higher-level visual object, so neighbor selection is performed between clusters while local synchronization is maintained inside each cluster. The algorithm CVAPF extends the cooperation between UAVs to the cooperation between clusters. Algorithm CVAPF has the following advantages over the algorithm VAPF.

- (1) Because the algorithm CVAPF has a cluster merging mechanism, the time for the formation to reach the steady state is shorter than that of the algorithm VAPF.
- (2) After the formation reaches a steady state, the synchronization performance of formation is better than that of the algorithm VAPF, due to the two-layer communication architecture of the algorithm CVAPF.
- (3) With cluster merging, the number of external neighbor objects processed by CVAPF can be reduced because several UAVs are treated as one cluster-level visual object.

The following will explain the methods for generating cluster behavior.

4.3.1. Cluster Attention Neighbor Node Selection

Figure 3 illustrates how the individual zoom-lens rule in Figure 2 is lifted to the cluster level. Unlike BOIDS and Optimized-flocking, which select individual neighbors using local interaction rules, CVAPF first groups nearby UAVs and then applies visual attention to cluster-level objects. Each cluster is treated as a visual object whose angular span is determined by its boundary UAVs. After the nearest cluster is selected, its perceived cone masks angularly redundant clusters behind the same visual sector. This operation keeps the inter-cluster topology sparse while preserving the clusters that are geometrically most informative for aggregation (reference Figure 3). The distance from cluster $\mathcal{J}$ to cluster $\mathcal{I}$ is represented as

$$\begin{aligned}\|\Delta X_{\mathcal{JI}}\| &= \min \|x_i - x_j\|, i \in \mathcal{I}, j \in \mathcal{J}, \\ &= \|x_{i^*}^* - x_{j^*}^*\|, i^* \in \mathcal{I}, j^* \in \mathcal{J}.\end{aligned}\quad (21)$$

$$\Delta V_{\mathcal{JI}} = v_{j^*}^* - v_{i^*}^*, \quad (22)$$

where UAV. i^* and UAV. j^* are the two closest UAVs to Cluster $\mathcal{I}$ and Cluster $\mathcal{J}$. For UAV- $\mathcal{I}$ ($1 \notin \mathcal{I}$), it represents the set of UAV clusters. A single UAV is recorded as a cluster with a quantity of 1.

Step 1, select the cluster $\mathcal{J}^{[1]}$ closest to cluster $\mathcal{I}$.

$$\mathcal{J}^{[1]} = \arg \min_{\mathcal{J} \in \mathcal{U}} \|\Delta X_{\mathcal{JI}}\|. \quad (23)$$

The set of all clusters is denoted as U . $X_{\mathcal{I}}; X_{\mathcal{J}}$ are the position of clusters $\mathcal{I}$ and $\mathcal{J}$; Cluster $\mathcal{I}$ can receive information from cluster $\mathcal{J}^{[1]}$ at this point, denoted as

$$\mathcal{N}_{\mathcal{I}}^{[1]} = \mathcal{J}^{[1]}. \quad (24)$$

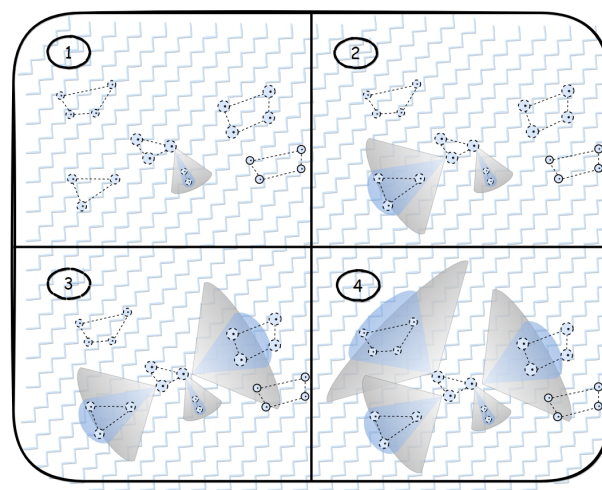


Figure 3. Cluster attention mechanism for selecting neighboring nodes. The numbers 1–4 indicate the cluster-neighbor selection order, light-blue sectors denote sensed angles, gray sectors denote perceived defocusing angles, and UAVs connected by dashed black lines belong to the same cluster.

The Perceived angle generated by $\mathcal{I}$ on $\mathcal{J}^{[1]}$ is denoted as $\Theta_{\mathcal{I}\mathcal{J}^{[1]}}$, and the Sensed angle is denoted as $\theta_{\mathcal{I}\mathcal{J}^{[1]}}$. The cone corresponding to a Sensed angle is denoted by $C_{\theta_{\mathcal{I}\mathcal{J}^{[1]}}}$, and the cone corresponding to a Perceived angle is denoted by $C_{\Theta_{\mathcal{I}\mathcal{J}^{[1]}}}$. We have

$$\Theta_{\mathcal{I}\mathcal{J}^{[1]}} = AR_C \cdot \theta_{\mathcal{I}\mathcal{J}^{[1]}}, \quad (25)$$

where $AR_C > 0$ is the cluster-level attention ratio. A larger AR_C masks a wider angular region after one cluster has been selected, so fewer external clusters are processed; a smaller AR_C keeps more clusters and increases the amount of state information used by the controller.

$$\theta_{\mathcal{I}\mathcal{J}^{[1]}} = \max \left\{ \arccos \left(\frac{(\Delta x_{ji^*}, \Delta x_{j'i^*})}{\|\Delta x_{ji^*}\| \|\Delta x_{j'i^*}\|} \right) + \arcsin \frac{r}{\|\Delta x_{ji^*}\|} + \arcsin \frac{r}{\|\Delta x_{j'i^*}\|} \right\}, \quad (26)$$

where $\Delta x_{ji^*} = x_j - x_{i^*}$, $i^* \in \mathcal{I}$; $j, j' \in \mathcal{J}^{[1]}$; r is the radius of the UAV,

$$x_{i^*} = \arg \min_{x_i} \left\{ \|x_j - x_i\| \mid j \in \mathcal{J}^{[1]} \right\}. \quad (27)$$

Step k , select the cluster $\mathcal{J}^{[k]}$ closest to cluster $\mathcal{I}$,

$$\mathcal{J}^{[k]} = \arg \min_{\mathcal{J} \in U - \bigcup_{m=1}^{k-1} \mathcal{J}^{[m]}} \|\Delta X_{\mathcal{J}\mathcal{I}}\|, \quad (28)$$

by (26) and (27), we have $\theta_{\mathcal{I}\mathcal{J}^{[k]}}$. Here $C_{\Theta_{\mathcal{I}}^{[k-1]}}$ denotes the union of perceived cones already selected before step k .

$$\mathcal{N}_{\mathcal{I}}^{[k]} = \begin{cases} \mathcal{J}^{[k]}, C_{\theta_{\mathcal{I}\mathcal{J}^{[k]}}} \cap C_{\Theta_{\mathcal{I}}^{[k-1]}} = \emptyset. \\ \emptyset; \text{otherwise.} \end{cases} \quad (29)$$

The perceived cone union of cluster $\mathcal{I}$ is updated as

$$C_{\Theta_{\mathcal{I}}}^{[k]} = \bigcup_{\mathcal{J}^{[k]} \in \mathcal{N}_{\mathcal{I}}^{[k]}} C_{\Theta_{\mathcal{I}\mathcal{J}^{[k]}}}, \quad (30)$$

The set of neighbors selected by cluster $\mathcal{I}$ can be represented as

$$\mathcal{N}_{\mathcal{I}} = \bigcup_k \mathcal{N}_{\mathcal{I}}^{[k]}. \quad (31)$$

Internal communication of cluster $\mathcal{I}$ or cluster $\mathcal{J}$, selecting AR_C^{IN} as the attention ratio

$$N_i^{IN} = \{j | j \in N_i \cap \mathcal{I}\}.$$

The intra-cluster ratio AR_C^{IN} controls how many members of the same cluster are used for velocity alignment. In our design, external cluster links drive aggregation, whereas N_i^{IN} provides local synchronization after members have merged into the same cluster.

Remark 2. The size of the inter-cluster topology changes dynamically with the number of clusters. When clusters merge, fewer external cluster-level objects need to be processed by the visual attention rule, while UAVs within the same cluster mainly use local links for velocity synchronization. This is the main difference between CVAPF and a fixed-radius neighbor rule: CVAPF does not process all visible UAVs independently after clustering, but separates sparse inter-cluster interaction from low-cost intra-cluster synchronization.

4.3.2. Cluster Motion Update

Cluster motion update converts selected cluster neighbors into control inputs applied through representative closest UAV pairs. The purpose is to make external cluster interactions sparse while still using intra-cluster velocity alignment to keep merged members moving coherently. Based on $\|\Delta X_{\mathcal{I}\mathcal{J}}\|$, split the neighboring node $\mathcal{N}_{\mathcal{I}}$ into the following sets

$$\begin{aligned} \mathcal{N}_{\mathcal{I}}^g &= \left\{ \mathcal{J} | \mathcal{J} \in \mathcal{N}_{\mathcal{I}}, \|\Delta X_{\mathcal{I}\mathcal{J}}\| \in [\mathcal{R}_h^{\mathcal{J}\mathcal{I}}, \mathcal{R}_g^{\mathcal{J}\mathcal{I}}] \right\}, \\ \mathcal{N}_{\mathcal{I}}^h &= \left\{ \mathcal{J} | \mathcal{J} \in \mathcal{N}_{\mathcal{I}}, \|\Delta X_{\mathcal{I}\mathcal{J}}\| \in [\mathcal{R}_a^{\mathcal{J}\mathcal{I}}, \mathcal{R}_h^{\mathcal{J}\mathcal{I}}] \right\}, \\ \mathcal{N}_{\mathcal{I}}^a &= \left\{ \mathcal{J} | \mathcal{J} \in \mathcal{N}_{\mathcal{I}}, \|\Delta X_{\mathcal{I}\mathcal{J}}\| \in (0, \mathcal{R}_a^{\mathcal{J}\mathcal{I}}] \right\}. \end{aligned} \quad (32)$$

Here $\mathcal{R}_g^{\mathcal{J}\mathcal{I}}$, $\mathcal{R}_h^{\mathcal{J}\mathcal{I}}$, and $\mathcal{R}_a^{\mathcal{J}\mathcal{I}}$ are the cluster-level gravitational, homogeneous, and anti-gravity radii generated by cluster $\mathcal{J}$ for cluster $\mathcal{I}$, respectively. They play the same role as R_g^{ji} , R_h^{ji} , and R_a^{ji} in the individual-level VAPF model. The anti-gravity set is restricted to $(0, \mathcal{R}_a^{\mathcal{J}\mathcal{I}}]$ so that cluster-level repulsion is activated only in the short-range safety region.

For cluster UAV i , $i \neq 1$, their motion models can be represented as

$$\tilde{v}_i[n+1] = \tilde{v}_i[n] + \alpha \sum_{j \in N_i^{IN}} (v_j[n] - v_i[n]) + \min\{\|U_i[n]\|, \|U_{max}\|\} \frac{U_i[n]}{\|U_i[n]\|}; \quad (33)$$

$$U_i[n+1] = \left\{ \mathcal{K}_1 \mathcal{F}_g^{[n]}(i) + \mathcal{K}_2 \mathcal{F}_h^{[n]}(i) + \mathcal{K}_3 \mathcal{F}_a^{[n]}(i) \right\}, \quad (34)$$

where α is the intra-cluster velocity alignment gain, and $\|U_{max}\|$ is the maximum allowed cluster-level control input; $\mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3$ are non-negative constants, and satisfy $\mathcal{K}_1 + \mathcal{K}_2 + \mathcal{K}_3 = 1$; $\mathcal{F}_g(i)$ represents the force exerted by external clusters on UAV- i in the gravitational black-hole potential field; $\mathcal{F}_h(i)$ represents the force exerted by exter-

nal clusters on UAV- i in the homogeneous potential field; and $\mathcal{F}_a(i)$ represents the force exerted by external clusters on UAV- i in the anti-gravity potential field. The constants $\mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3$ have the same physical roles as k_1, k_2, k_3 , but they are applied to interactions between clusters rather than interactions between individual UAVs. For clusters $\mathcal{I}$ and $\mathcal{J}$, $\mathcal{F}_g^{[n]}(i), \mathcal{F}_h^{[n]}(i), \mathcal{F}_a^{[n]}(i)$ represent the following.

$$\mathcal{F}_g^{[n]}(i) = \mathbf{0}, i \neq i^*,$$

$$\mathcal{F}_g^{[n]}(i) = \sum_{\mathcal{J} \in \mathcal{N}_{\mathcal{I}}^g} \left\{ -k_b^{\mathcal{J}\mathcal{I}} \left[-R_b^{\mathcal{J}\mathcal{I}} \frac{\Delta X_{\mathcal{J}\mathcal{I}}^{[n-1]}}{\|\Delta X_{\mathcal{J}\mathcal{I}}^{[n-1]}\|} + \Delta X_{\mathcal{J}\mathcal{I}}^{[n-1]} \right] \cdot \tau(R_b^{\mathcal{J}\mathcal{I}} - \|\Delta X_{\mathcal{J}\mathcal{I}}^{[n-1]}\|) + k_g^{\mathcal{J}\mathcal{I}} \Delta X_{\mathcal{J}\mathcal{I}}^{[n-1]} \right\}, \text{otherwise}, \quad (35)$$

where $R_g^{\mathcal{J}\mathcal{I}}$ represents the radius of the gravitational black-hole potential field generated by UAV- $\mathcal{J}$ on UAV- $\mathcal{I}$; $R_h^{\mathcal{J}\mathcal{I}}$ represents the radius of the homogeneous potential field generated by UAV- $\mathcal{J}$ on UAV- $\mathcal{I}$; $R_b^{\mathcal{J}\mathcal{I}}$ represents the radius of the black-hole potential field generated by UAV- $\mathcal{J}$ on UAV- $\mathcal{I}$; $k_b^{\mathcal{J}\mathcal{I}}$ represents the intensity coefficient of the black-hole potential field generated by UAV- $\mathcal{J}$ on UAV- $\mathcal{I}$; and $k_g^{\mathcal{J}\mathcal{I}}$ represents the intensity coefficient of the gravitational black-hole potential field generated by UAV- $\mathcal{J}$ on UAV- $\mathcal{I}$.

$$\mathcal{F}_h^{[n]}(i) = \mathbf{0}, i \neq i^*,$$

$$\mathcal{F}_h^{[n]}(i) = - \sum_{\mathcal{J} \in \mathcal{N}_{\mathcal{I}}^h} k_h^{\mathcal{J}\mathcal{I}} \Delta V_{\mathcal{I}\mathcal{J}}, \text{otherwise}, \quad (36)$$

where $R_a^{\mathcal{J}\mathcal{I}}$ represents the radius of the anti-gravity potential field generated by UAV- $\mathcal{J}$ on UAV- $\mathcal{I}$; $k_h^{\mathcal{J}\mathcal{I}}$ represents the intensity coefficient of the homogeneous potential field generated by UAV- $\mathcal{J}$ on UAV- $\mathcal{I}$.

$$\mathcal{F}_a^{[n]}(i) = \mathbf{0}, i \neq i^*,$$

$$\mathcal{F}_a^{[n]}(i) = \sum_{\mathcal{J} \in \mathcal{N}_{\mathcal{I}}^a} \frac{\xi_{\mathcal{J}\mathcal{I}}[n-1]}{\Xi_{\mathcal{I}}[n-1]} \left\{ -\frac{1}{R_a^{\mathcal{J}\mathcal{I}} \|\Delta X_{\mathcal{J}\mathcal{I}}^{[n-1]}\| \Delta X_{\mathcal{J}\mathcal{I}}^{[n-1]}} + \frac{1}{(\Delta X_{\mathcal{J}\mathcal{I}}^{[n-1]})^3} \right\}, \text{otherwise}, \quad (37)$$

where $\xi_{\mathcal{J}\mathcal{I}}$ represents UAV- $\mathcal{J}$ for UAV- $\mathcal{I}$ attention coefficient; $\frac{\xi_{\mathcal{J}\mathcal{I}}}{\Xi_{\mathcal{I}}}$ represents UAV- $\mathcal{J}$ for UAV- $\mathcal{I}$ attention weight. The calculation formula for $\Xi_{\mathcal{I}}$ is as follows.

$$\begin{aligned} \Xi_{\mathcal{I}}[n] &= \sum_{\mathcal{J} \in \mathcal{N}_{\mathcal{I}}^a} \xi_{\mathcal{J}\mathcal{I}}[n] \\ &= \sum_{\mathcal{J} \in \mathcal{N}_{\mathcal{I}}^a} k_a^{\mathcal{J}\mathcal{I}} \frac{\exp\left\{ \|\Delta V_{\mathcal{J}\mathcal{I}}^{[n-1]}\| \cos \angle(\Delta V_{\mathcal{J}\mathcal{I}}^{[n-1]}, \Delta X_{\mathcal{J}\mathcal{I}}^{[n-1]}) \right\}}{\|\Delta X_{\mathcal{J}\mathcal{I}}^{[n-1]}\| + \varepsilon}, \end{aligned} \quad (38)$$

The cluster-level attention weight follows the same relevance logic as the individual-level coefficient: distance measures short-range interaction urgency, and $\cos \angle(\Delta V_{\mathcal{J}\mathcal{I}}, \Delta X_{\mathcal{J}\mathcal{I}})$ extracts whether the two closest representative UAVs are approaching along their line of sight. $k_a^{\mathcal{J}\mathcal{I}}$ represents the intensity coefficient of the anti-gravity potential field generated by UAV- $\mathcal{J}$ on UAV- $\mathcal{I}$. Thus, a cluster that is closer or approaching faster receives a larger

repulsive weight, which is consistent with the purpose of avoiding close-range conflicts between clusters.

Figure 4 summarizes the data flow of CVAPF. The upper layer uses the cluster visual attention rule to choose only a small number of external clusters and applies potential-field forces through representative closest UAV pairs. The lower layer uses intra-cluster links to align velocities among UAVs that have already merged into the same cluster. Therefore, aggregation is controlled at the cluster level, while smooth local motion is controlled at the member level.

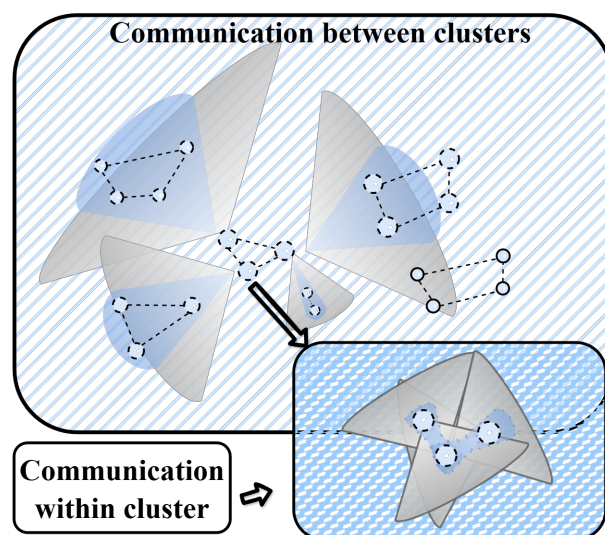


Figure 4. Dual layer communication architecture. Light-blue sectors denote sensed angles, gray sectors denote perceived defocusing angles, and UAVs connected by dashed black lines belong to the same cluster.

Remark 3. The dual layer communication architecture designed by the CVAPF algorithm is designed for distributed communication between and within clusters. Distributed communication between clusters is used to support aggregation between clusters, while distributed communication within clusters is used to support velocity consistency inside the cluster. In this way, CVAPF evaluates fewer external visual objects after merging while preserving local synchronization links among members of the same cluster.

4.3.3. Cluster Merging

Cluster merging changes the interaction topology, so the merge operation is made conservative. The following three conditions require spatial proximity, pairwise velocity compatibility, and internal velocity consistency before two clusters are fused. When groups approach each other, they will merge if they satisfy the following three conditions.

The merging rules are introduced to prevent two clusters from being fused before they are physically close and kinematically compatible.

Condition 1. The distance between clusters satisfies the following:

$$\|\Delta X_{\mathcal{JI}}\| < \mathcal{R}_{\mathcal{JI}}. \quad (39)$$

Condition 2. The speed of the two closest UAVs between clusters should be sufficiently close.

$$1 - \frac{v_i^* v_j^*}{\max\{\|v_i^*\|^2, \|v_j^*\|^2\}} \leq \epsilon_a, i^* \in \mathcal{I}, j^* \in \mathcal{J}. \quad (40)$$

Condition 3. The UAVs within each cluster have similar speeds.

$$1 - \min_{i_1, i_2 \in \mathcal{I}; j_1, j_2 \in \mathcal{J}} \left\{ \frac{v_{i_1} v_{i_2}}{\max\{\|v_{i_1}\|^2, \|v_{i_2}\|^2\}}, \frac{v_{j_1} v_{j_2}}{\max\{\|v_{j_1}\|^2, \|v_{j_2}\|^2\}} \right\} \leq \epsilon_b, \quad (41)$$

$\mathcal{R}_{\mathcal{I}\mathcal{J}}$ represents the distance that UAV- $\mathcal{I}$ and UAV- $\mathcal{J}$ can merge, ϵ_a, ϵ_b are any small positive constants; and smaller ϵ_a and ϵ_b make merging stricter because they require stronger velocity similarity, while larger values allow earlier merging at the cost of possible transient disorder. When (39)–(41) are satisfied, we have

$$\mathcal{I} := \mathcal{I} \cup \mathcal{J}. \quad (42)$$

Remark 4. The merging between clusters requires the two clusters to be close enough, and the speed between the two closest UAVs in the two clusters must be sufficiently similar. And each cluster needs to achieve autonomy before it can be merged, which means that the speed of all UAVs within the cluster must be highly similar before it can be satisfied. These conditions are all aimed at ensuring that cluster merging should be a gradual process. This process is similar to two trains that need to be docked; they need to be close enough in distance and speed. And the speed of each carriage in each train needs to be similar.

4.4. Calculate the Boundary Conditions

This subsection gives geometric boundary conditions used for parameter selection. These conditions are practical constraints for avoiding close-range collision and maintaining local visual links under the stated assumptions, not a complete stability proof for arbitrary switching topologies.

The discussion is divided into two parts. First, a worst-case head-on encounter is used to define a safety margin for the anti-gravity field. Second, this margin is used to derive a sufficient angular condition for preserving local communication links inside a small three-UAV unit.

4.4.1. Avoid Collision (Reference Figure 5)

Consider the most extreme situation: when two clusters are moving towards each other at maximum speed $\|v_{max}\|$ and their initial distance is $R_a^{\mathcal{I}\mathcal{J}}$. We hope that they do not collide and the minimum distance can still maintain communication properties when constructing communication topologies. When we analyze the cluster as a whole, we do not consider the forces within the cluster. At this point, consider the cluster model as rigid movement (in reality, the position movement of each UAV within the cluster is not rigid movement. However, we only consider the two closest UAVs in the two clusters, so it does not affect it). Under this worst-case setting, the anti-gravity term is isolated because it is the component responsible for short-range separation.

$$U_{\mathcal{I}}[n+1] = \mathcal{K}_3 \left\{ \frac{1}{(\Delta X_{\mathcal{I}\mathcal{J}}^{[n]})^3} + \frac{1}{R_a^{\mathcal{I}\mathcal{J}} \|\Delta X_{\mathcal{I}\mathcal{J}}^{[n]}\| \Delta X_{\mathcal{I}\mathcal{J}}^{[n]}} \right\}, \quad (43)$$

$$U_{\mathcal{J}}[n+1] = -U_{\mathcal{I}}[n+1]. \quad (44)$$

We need to constrain clusters to avoid collisions

$$R_{safe} = \min_{\mathcal{I}, \mathcal{J}} R_a^{\mathcal{I}\mathcal{J}} - \min\{\|X_{\mathcal{J}}[k] - X_{\mathcal{I}}[k]\|\} > 0. \quad (45)$$

In this constraint, R_{safe} is the remaining distance margin after the strongest head-on repulsive interaction. Positive R_{safe} means that the minimum simulated inter-cluster distance is still smaller than the activation radius $R_a^{\mathcal{I}\mathcal{I}}$ by a nonzero safety margin, so the anti-gravity field can stop the two closest UAVs before collision.

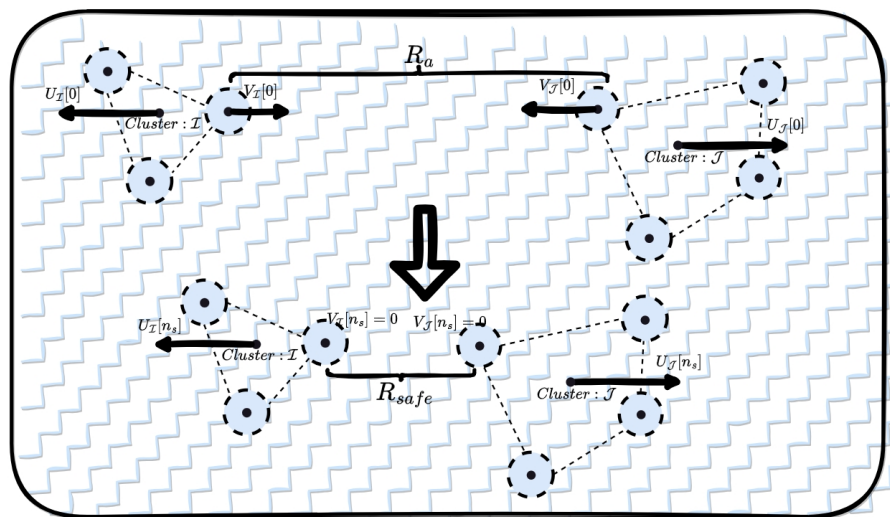


Figure 5. Cluster collision avoidance. Blue disks denote UAVs, and UAVs connected by dashed black lines belong to the same cluster. Solid black arrows denote velocity or control-input directions, and bracketed distances show the anti-gravity activation radius R_a and the remaining safe distance R_{safe} .

4.4.2. Maintain Communication

The communication maintenance condition links the safety margin to the visual attention parameter inside a cluster. The purpose is to prevent the perceived cone of one selected UAV from masking all other local candidates in the smallest connected unit.

Using the safety margin above, we derive a sufficient geometric condition for maintaining local visual links in the communication topology generated within the cluster based on attention mechanism (refer to Figure 6). We consider three UAVs as a minimum unit, and they meet the following conditions. Here we consider one of the UAVs as the leader (not affect the generation of communication topology). Using the same cone notation at the UAV level, $C_{\theta_{ij}}$ is the sensed cone of UAV. j observed from UAV. i , and $C_{\Theta_{ij}}$ is the corresponding perceived cone. The condition below prevents the perceived cone of one selected UAV from covering the sensed cone of another candidate UAV, so at least two local visual links can be preserved in the three-UAV unit.

$$C_{\Theta_{ij}} \cap C_{\theta_{ij}} = \emptyset, \quad (46)$$

Because the smallest strongly connected unit in Figure 6 is arranged with an angular separation of $\pi/3$, this non-overlap condition leads to the following sufficient bound:

$$AR_C^{IN} \cdot \arcsin\left(\frac{r}{R_{safe}}\right) + \arcsin\left(\frac{r}{R_{safe}}\right) \leq \frac{\pi}{3}. \quad (47)$$

Within the smallest unit, at least one UAV can receive information from the other two UAVs. Moreover, at least one UAV in the formation can receive information from the leader, which provides a sufficient local condition for preserving a visual information path to the leader under the assumed geometry.

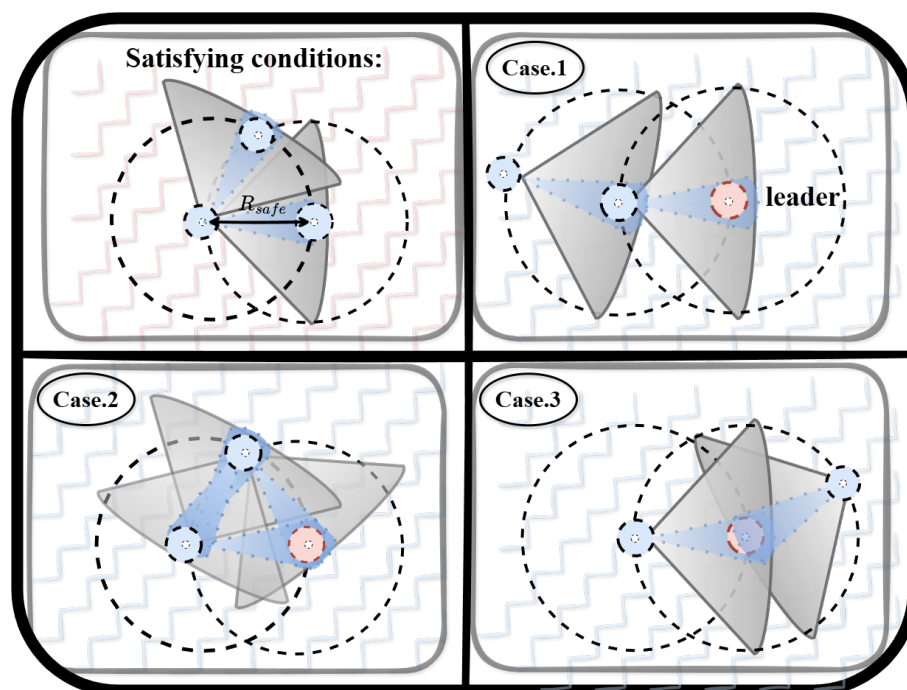


Figure 6. Communication maintenance condition in a three-UAV unit. Light-blue sectors denote sensed angles, gray sectors denote perceived defocusing angles. Dashed circular boundaries around UAVs denote the safety radius, UAVs with dashed borders are the focused UAVs retained by visual attention.

Remark 5. Equation (47) is a sufficient geometric condition for preserving local visual links in the three-UAV unit shown in Figure 6. Extending this unit by adding nearby UAVs gives a practical construction rule for local connectivity under the stated geometric assumptions, but it should not be interpreted as a complete proof for arbitrary switching topologies. This argument should be interpreted as a geometric sufficient-condition analysis for maintaining local visual links under the stated assumptions, rather than as a full graph-theoretic stability proof for arbitrary switching topologies. Equation (47) also gives a practical rule for parameter selection: increasing AR_C^{IN} reduces redundant intra-cluster observations, but it must remain below the bound implied by r and R_{safe} to avoid disconnecting the local topology.

5. Simulation Results and Discussion

This section evaluates the proposed methods in simulation with emphasis on the scope of this paper: visual neighbor selection and redundancy suppression. The simulations are not intended to be a full robustness benchmark for sensor noise, communication delay, obstacle-rich environments, or multirotor dynamics. To demonstrate the effectiveness of the proposed method, we designed three simulation examples. Example 1 verifies the effectiveness of Algorithm 1 in formation flight. Example 2 verifies the effectiveness of Algorithm 2 in formation flight. Example 3 is designed specifically for the main focus of this paper: comparing the neighbor processing load and visual redundancy of different neighbor-selection rules. It does not aim to provide a full robustness, convergence, or flight-dynamics benchmark. We compare the advantages and disadvantages of the two algorithms. The effectiveness of the algorithm has been demonstrated through simulation using Python 3.8.5.

Algorithm 1. Leader Flight Planning + Algorithm VAPF**Require:** Position x_i and speed v_i of each UAV.**Ensure:** Formation behavior evaluated under the proposed VAPF update.

- 1: Initialization $r, AR, n_l, \left\{k_L^k\right\}_{k=1}^{n_l+1}, \left\{v_L^k\right\}_{k=1}^{n_l+1}, \left\{l^{(k)}\right\}_{k=1}^{n_l+1}, \|v_{max}\|, \|u_{max}\|, \gamma, n_h, k_1, k_2, k_3, R_g^{ji}, R_h^{ji}, R_b^{ji}, R_a^{ji}, k_g^{ji}, k_h^{ji}, k_b^{ji}, k_a^{ji}, \varepsilon$.
- 2: At the $n + 1$ -th moment
- 3: Initialization $x_i[n], v_i[n], \sum_{k=0}^{\min\{n, n_h\}} \tilde{v}_i[n - k]$.
- 4: Flight planning for leader (UAV.1).
- 5: Calculate $u_1[n + 1] \Rightarrow v_1[n + 1] \Rightarrow x_1[n + 1]$
- 6: **for** UAV. i ; $i \neq 1$ **do**
- 7: Using visual attention mechanism to select neighbor N_i .
- 8: Divide N_i into N_i^g, N_i^h , and N_i^a based on distance.
- 9: Calculate $f_g^{[n]}(i), f_h^{[n]}(i), f_a^{[n]}(i) \Rightarrow u_i[n] \Rightarrow \tilde{v}_i[n + 1] \Rightarrow v_i[n + 1] \Rightarrow x_i[n + 1]$.
- 10: **end for**

Algorithm 2. Leader Flight Planning + Algorithm CVAPF**Require:** Position x_i and speed v_i of each UAV.**Ensure:** Formation behavior evaluated under the proposed CVAPF update.

- 1: Initialization $r, AR_C, AR_C^{IN}, n_l, \left\{k_L^k\right\}_{k=1}^{n_l+1}, \left\{v_L^k\right\}_{k=1}^{n_l+1}, \left\{l^{(k)}\right\}_{k=1}^{n_l+1}, \|v_{max}\|, \|u_{max}\|, \gamma, n_h, \alpha, \|U_{max}\|, \mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3, R_g^{\mathcal{J}\mathcal{I}}, R_h^{\mathcal{J}\mathcal{I}}, R_b^{\mathcal{J}\mathcal{I}}, R_a^{\mathcal{J}\mathcal{I}}, k_g^{\mathcal{J}\mathcal{I}}, k_h^{\mathcal{J}\mathcal{I}}, k_b^{\mathcal{J}\mathcal{I}}, k_a^{\mathcal{J}\mathcal{I}}, \varepsilon, \mathcal{R}_{\mathcal{J}\mathcal{I}}, \epsilon_a, \epsilon_b$.
- 2: At the $n + 1$ -th moment
- 3: Initialization $x_i[n], v_i[n], \sum_{k=0}^{\min\{n, n_h\}} \tilde{v}_i[n - k]$.
- 4: Flight planning for leader (UAV.1).
- 5: Calculate $u_1[n + 1] \Rightarrow v_1[n + 1] \Rightarrow x_1[n + 1]$
- 6: **for** UAV. i ; $i \neq 1$ **do**
- 7: **if** Cluster UAV- $\mathcal{I}$ meets merging conditions (41) **then**
- 8: **for** $\mathcal{J}$ in $\mathcal{U}$ **do**
- 9: **if** Cluster UAV- $\mathcal{I}$ and UAV- $\mathcal{J}$ meet condition (39) and (40) **then**
- 10: Cluster UAV- $\mathcal{I}$ and UAV- $\mathcal{J}$ merge to form a new cluster UAV- $\mathcal{I}_{new}$
- 11: **end if**
- 12: **end for**
- 13: **end if**
- 14: Using cluster visual attention mechanism to select neighbor $\mathcal{N}_{\mathcal{I}}$, and using visual attention mechanism to select neighbor $\mathcal{N}_{\mathcal{I}}^{IN}$.
- 15: Divide $\mathcal{N}_{\mathcal{I}}$ into $\mathcal{N}_{\mathcal{I}}^g, \mathcal{N}_{\mathcal{I}}^h$, and $\mathcal{N}_{\mathcal{I}}^a$ based on $\|\Delta X_{\mathcal{J}\mathcal{I}}\|$.
- 16: Calculate $\mathcal{F}_g^{[n]}(i), \mathcal{F}_h^{[n]}(i), \mathcal{F}_a^{[n]}(i) \Rightarrow U_i[n] \Rightarrow \tilde{v}_i[n + 1] \Rightarrow v_i[n + 1] \Rightarrow x_i[n + 1]$.
- 17: **end for**

The simulation setup is reported before the examples so that the parameter choices are explicit. Tables 1 and 2 give the leader settings and the common initial follower states used in Examples 1 and 2. In leader flight planning, the initial parameters we set refer to Table 1. These parameters will be used in the following two examples.

Table 1. Leader flight planning simulation parameters.

Items	Simulation Parameters
Initial Position	$x_1[0] = (0, 0, 20)^T$
Initial Speed	$v_1[0] = (0, 0, 20)^T$
Maximum Speed	$\ v_{max}\ = 20$
Maximum Input	$\ u_{max}\ = 40$
Divide Quantity	$n_l = 580$
Potential Field Intensity Coefficient	$\left\{ k_L^{(k+1)} \right\}_{k=1}^{n_l} = 1$
Preset Expected Speed	$\left\{ v_L^{(k+1)} \right\}_{k=1}^{n_l} = (0, 0, 20)^T$
Division Point Position	$\left\{ l^{(k+1)} \right\}_{k=0}^{n_l} = (0, 0, 20 + k)^T$

The parameters used in the following examples are selected according to their physical roles rather than by exhaustive optimization. The interaction radii follow the order $0 < r < R_a < R_h < R_b < R_g$: R_a defines the short-range safety region, R_h marks the velocity alignment region, R_b separates shaped intermediate attraction from long-range attraction, and R_g is the largest distance at which a selected neighbor affects the APF controller. The same ordering principle is used for cluster-level interaction radii. The weights k_1, k_2, k_3 and $\mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3$ are normalized to keep the input scale bounded, with the repulsive weight chosen largest so that collision avoidance dominates at short range. The filtering parameters γ and n_h trade response speed for smoothness, while the merging thresholds ϵ_a and ϵ_b determine how strictly clusters must match velocities before fusion.

The initial speed of the remaining UAVs is 0, and their initial positions refer to Table 2.

Table 2. Initial position of the follower UAVs.

Items	Simulation Parameters
Initial Position of UAV.2	$x_2[0] = (37.45, 95.07, 73.19)^T$
Initial position of UAV.3	$x_3[0] = (5.81, 86.62, 60.11)^T$
Initial position of UAV.4	$x_4[0] = (83.24, 21.23, 18.18)^T$
Initial position of UAV.5	$x_5[0] = (43.19, 29.12, 61.18)^T$
Initial position of UAV.6	$x_6[0] = (45.60, 78.51, 19.96)^T$
Initial position of UAV.7	$x_7[0] = (60.75, 17.05, 6.51)^T$

Example 1. Algorithm 1 (leader flight planning + algorithm VAPF). In this example, we consider seven UAVs that rely on visual attention mechanisms to establish communication topologies and transmit state information. The unique parameter $AR = 4, r = 1$ used to construct communication topology. The various parameters in the motion of the UAVs are $\gamma = 0.9; n_h = 5; \epsilon = 1; k_1 = 0.2; k_2 = 0.2; k_3 = 0.6; R_a^{ji} = 10, R_h^{ji} = 40, R_b^{ji} = 55, R_g^{ji} = 200, k_g^{ji} = k_b^{ji} = k_h^{ji} = k_a^{ji} = 1$.

Figure 7 shows the experimental results of Algorithm 1 after simulation. Among them, Figure 7a represents the flight trajectory of the formation starting from its initial position within 30 seconds, and the color of the trajectory is mapped to time. The trajectory indicates that Algorithm 1 produces aggregation behavior in the tested scenario. However, as the formations approached each other, each UAV seemed to become hasty, causing the UAVs to repeatedly move from dispersion to aggregation, as shown in Figure 7b. In this simulation,

Algorithm 1 shows strong aggregation behavior and no inter-UAV collision is observed. From a local perspective, the trajectory of the follower UAV is disorderly, as shown in Figure 7c.

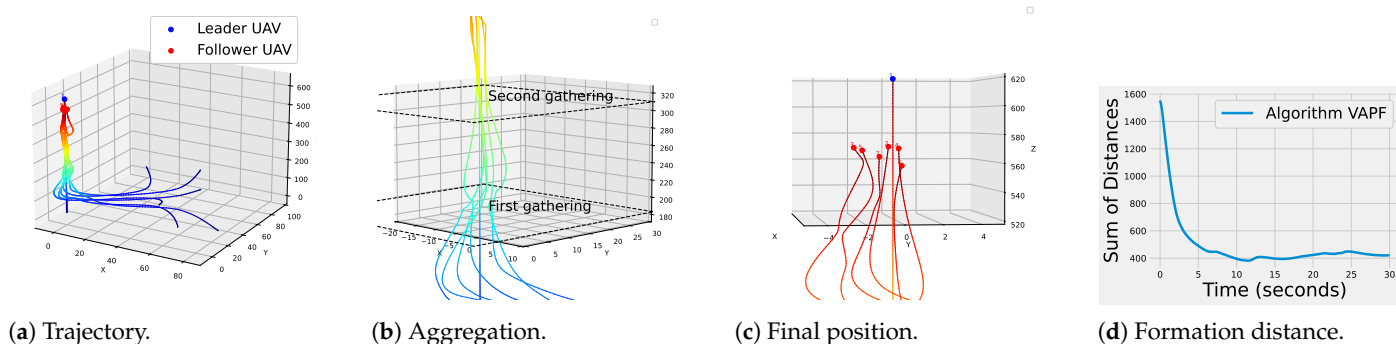


Figure 7. Algorithm VAPF simulation results. The color of the trajectory is mapped to time.

Example 2. Algorithm 2 (leader flight planning + algorithm CVAPF). In this example, the attention ratio of establishing a communication topology between clusters is $AR_C = 3$, $r = 1$, and the attention ratio of establishing a communication topology between UAVs within a cluster is $AR_C^{IN} = 2$. The parameters for cluster motion and merging are $\gamma = 0.9$; $n_h = 5$; $\epsilon = 1$; $\alpha = 0.5$, $K_1 = 0.2$; $K_2 = 0.2$; $K_3 = 0.6$; $R_a^{JI} = 10$, $R_h^{JI} = 40$, $R_b^{JI} = 55$, $R_g^{JI} = 200$, $k_g^{JI} = k_b^{JI} = k_h^{JI} = k_a^{JI} = 1$; $R_{JI} = 40$; $\epsilon_a = 0.9$; $\epsilon_b = 0.95$.

Figure 8 shows the experimental results of Algorithm 2 after simulation. Among them, Figure 8a represents the flight trajectory of the formation starting from its initial position within 30 seconds, and the color of the trajectory is mapped to time. The trajectory shows that Algorithm 2 improves aggregation and synchronization in the tested scenario. As formation mergers occur, the trajectory of the formation becomes increasingly synchronized, as shown in Figure 8b. From a local perspective, the formation consistency driven by Algorithm 2 is more orderly than the VAPF trajectory in Figure 7c, as shown in Figure 8c.

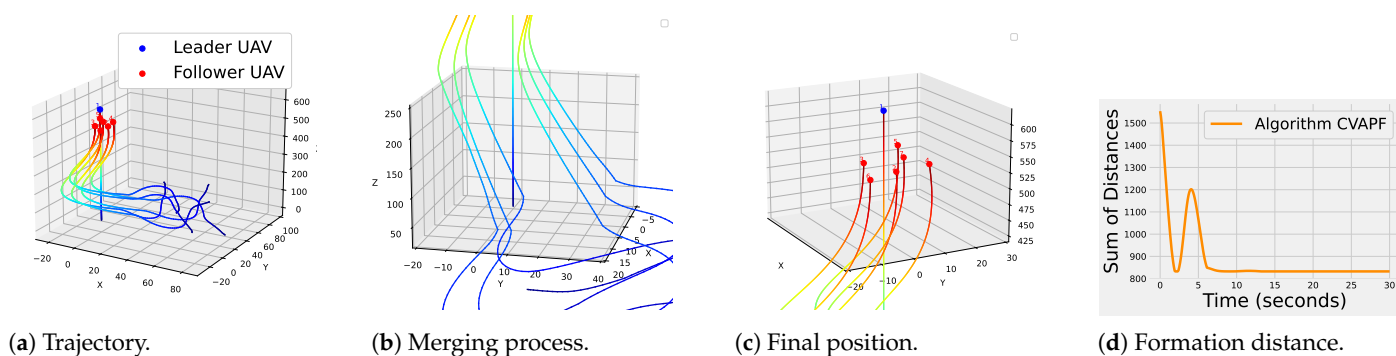


Figure 8. Algorithm CVAPF. The color of the trajectory is mapped to time.

Remark 6. Algorithm 1 has better aggregation performance than Algorithm 2 (refer to Figures 7d and 8d). The sum of distances $\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \|x_i - x_j\|$ for the formation in Algorithm 1 is approximately 400, which is only half of Algorithm 2. In the tested scenario, Algorithm 1 produces a more compact final aggregation, while Algorithm 2 produces more synchronized trajectories because cluster merging reduces repeated external neighbor processing and strengthens intra-cluster velocity alignment. Due to the setting of the condition (41) for cluster merging, the distance between UAVs is far greater than that of Algorithm 1.

Example 3. Comparative neighbor processing experiment.

To address whether visual attention provides a measurable advantage over standard neighbor-selection rules, we further test a redundant sensing scene with 45 UAVs in the same spatial configuration. The compared methods are baseline BOIDS [16], which processes all UAVs inside a fixed sensing radius; baseline Optimized-flocking [19], which uses a shorter effective interaction radius motivated by the optimized autonomous-drone flocking model; VAPF; and CVAPF. This experiment evaluates only the neighbor-selection topology, so all methods are compared under the same positions, velocities, and sensing scene.

The qualitative comparison in Figure 9 shows that the two baseline methods retain many angularly overlapping neighbors, whereas VAPF and CVAPF suppress redundant neighbors in the same visual sector. The quantitative results in Table 3 support the same conclusion. The average processed neighbors per UAV decrease from 43.38 in BOIDS and 15.20 in the Optimized-flocking baseline to 6.31 in VAPF and 3.60 in CVAPF. In this experiment, a redundant neighbor means a farther UAV whose visual cone is occluded by a nearer UAV in the same viewing direction. Such an occluded UAV does not provide additional effective visual information to the reference UAV, but including it as a neighbor still increases the controller's state-processing load. Therefore, the redundancy ratio is computed by scanning the candidate neighbors from near to far and counting the farther UAVs that fall inside cones already occupied by nearer UAVs. This test gives redundancy ratios of 0.318 for BOIDS and 0.273 for Optimized-flocking, while VAPF and CVAPF have no additional occlusion redundancy after attention-based selection. Therefore, under the same sensing scene, the proposed visual attention topology preserves the APF-driven formation behavior observed in Examples 1 and 2 while substantially reducing redundant neighbor processing.

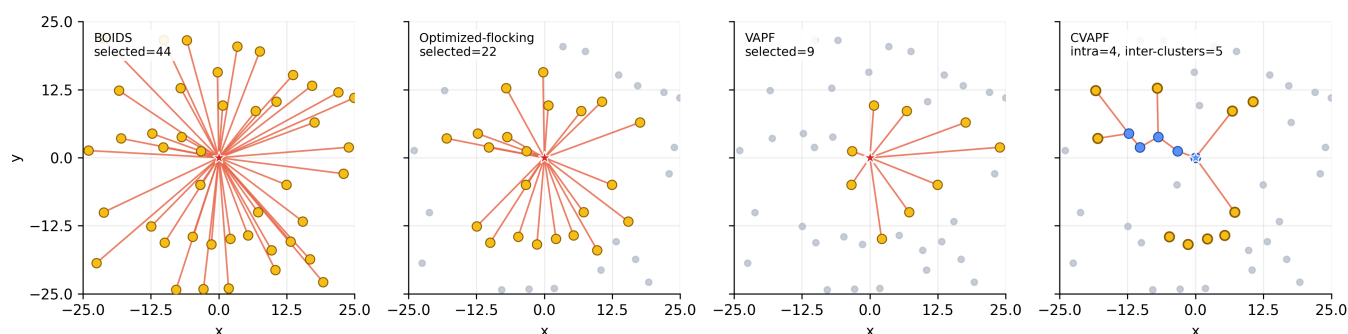


Figure 9. Neighbor selection in a redundant UAV sensing scene. The star denotes the inspected UAV or reference cluster, yellow points denote selected neighbors or selected neighbor clusters, gray points denote unselected UAVs, red lines denote processed neighbor links, and blue points denote intra-cluster selected UAVs.

Table 3. Quantitative comparison of neighbor processing load and redundancy ratio in the redundant sensing scene.

Method	Avg. Neighbors	Redundant Ratio
Baseline BOIDS	43.38	0.318
Baseline Optimized-flocking	15.20	0.273
VAPF	6.31	0.000
CVAPF	3.60	0.000

To clarify the influence of the attention ratio, we further vary AR in the same redundant sensing scene and keep all other settings unchanged. This analysis is not intended as a full controller sensitivity study; it isolates the neighbor-selection parameter that directly controls angular masking. A larger AR creates a wider perceived cone after a target is selected, so more angularly redundant candidates are suppressed and the average number of processed neighbors decreases.

Table 4 shows the expected monotonic trend. For VAPF, the average processed neighbors decrease from 22.73 to 4.73 as AR increases from 1 to 5. For CVAPF, the value decreases from 7.96 to 3.29 because cluster merging already suppresses many external objects before the attention ratio is varied. These results explain the practical role of AR : it trades retained visual information for lower neighbor processing load, and it should be chosen together with the communication maintenance bound in (47).

Table 4. Average processed neighbors under different attention ratios in the redundant sensing scene.

Method	$AR = 1$	$AR = 2$	$AR = 3$	$AR = 4$	$AR = 5$
VAPF	22.73	13.80	8.89	6.31	4.73
CVAPF	7.96	5.22	4.27	3.60	3.29

6. Conclusions

Considering formation control in complex environments, this paper proposes two visual-attention-based neighbor-selection algorithms based on a visual attention mechanism to generate communication topology, namely VAPF and CVAPF. The VAPF algorithm uses an individual-level zoom-lens rule to select informative UAV-level neighbors before APF control evaluation, which supports aggregation in the tested scenario. The CVAPF algorithm extends this neighbor-selection idea to the cluster level through a dual layer communication architecture and cluster merging. It reduces repeated external-neighbor processing and improves the synchronization behavior observed in the simulations. The simulations show that the proposed neighbor-selection algorithms preserve APF-driven formation behavior while reducing average processed neighbors to 6.31 and 3.60, compared with 43.38 for BOIDS and 15.20 for Optimized-flocking. Future work will study robustness under more realistic sensing and flight conditions.

Funding: This research received no external funding.

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Yan, P.; Guo, J.; Su, X.; Bai, C. Long-Term Tracking of Evasive Urban Target Based on Intention Inference and Deep Reinforcement Learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, *35*, 16886–16900. [\[CrossRef\]](#)
2. Gupta, A.; Afrin, T.; Scully, E.; Yodo, N. Advances of UAVs toward future transportation: The state-of-the-art, challenges, and opportunities. *Future Transp.* **2021**, *1*, 326–350. [\[CrossRef\]](#)
3. Erdos, D.; Erdos, A.; Watkins, S.E. An experimental UAV system for search and rescue challenge. *IEEE Aerosp. Electron. Syst. Mag.* **2013**, *28*, 32–37. [\[CrossRef\]](#)
4. Zhao, J.; Wang, H.; Bellotto, N.; Hu, C.; Peng, J.; Yue, S. Enhancing LGMD's looming selectivity for UAV with spatial-temporal distributed presynaptic connections. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *34*, 2539–2553.
5. Colomina, I.; Molina, P. Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2014**, *92*, 79–97. [\[CrossRef\]](#)
6. Tatsidou, E.; Tsiamis, C.; Karamagioli, E.; Boudouris, G.; Pikoulis, A.; Kakalou, E.; Pikoulis, E. Reflecting upon the humanitarian use of unmanned aerial vehicles (drones). *Swiss Med. Wkly.* **2019**, *149*, w20065. [\[CrossRef\]](#) [\[PubMed\]](#)

7. Wang, H.; Zhao, H.; Zhang, J.; Ma, D.; Li, J.; Wei, J. Survey on unmanned aerial vehicle networks: A cyber physical system perspective. *IEEE Commun. Surv. Tutor.* **2019**, *22*, 1027–1070. [\[CrossRef\]](#)
8. Chen, J.; Li, J.; Yuan, X. Global fuzzy adaptive consensus control of unknown nonlinear multiagent systems. *IEEE Trans. Fuzzy Syst.* **2019**, *28*, 510–522. [\[CrossRef\]](#)
9. Chen, J.; Xie, J.; Li, J.; Chen, W. Human-in-the-loop fuzzy iterative learning control of consensus for unknown mixed-order nonlinear multi-agent systems. *IEEE Trans. Fuzzy Syst.* **2023**, *32*, 255–265.
10. Zhang, J.; Cui, Y.; Ren, J. Dynamic mission planning algorithm for UAV formation in battlefield environment. *IEEE Trans. Aerosp. Electron. Syst.* **2022**, *59*, 3750–3765. [\[CrossRef\]](#)
11. Causa, F.; Vetrella, A.R.; Fasano, G.; Accardo, D. Multi-UAV formation geometries for cooperative navigation in GNSS-challenging environments. In *Proceedings of the 2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*; IEEE: New York, NY, USA, 2018; pp. 775–785.
12. Lin, Z.; Wang, L.; Han, Z.; Fu, M. Distributed formation control of multi-agent systems using complex Laplacian. *IEEE Trans. Autom. Control* **2014**, *59*, 1765–1777. [\[CrossRef\]](#)
13. Cetin, O.; Yilmaz, G. Real-time autonomous UAV formation flight with collision and obstacle avoidance in unknown environment. *J. Intell. Robot. Syst.* **2016**, *84*, 415–433. [\[CrossRef\]](#)
14. Zheng, Y.; Zheng, C.; Shen, J.; Liu, P.; Zhao, S. Keypoint-Guided Efficient Pose Estimation and Domain Adaptation for Micro Aerial Vehicles. *IEEE Trans. Robot.* **2024**, *40*, 2967–2983. [\[CrossRef\]](#)
15. Quan, L.; Yin, L.; Zhang, T.; Wang, M.; Wang, R.; Zhong, S.; Zhou, X.; Cao, Y.; Xu, C.; Gao, F. Robust and efficient trajectory planning for formation flight in dense environments. *IEEE Trans. Robot.* **2023**, *39*, 4785–4804. [\[CrossRef\]](#)
16. Reynolds, C.W. Flocks, herds and schools: A distributed behavioral model. In *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques*, Anaheim, CA, USA, 27–31 July 1987; pp. 25–34.
17. Das, A.K.; Fierro, R.; Kumar, V.; Ostrowski, J.P.; Spletzer, J.; Taylor, C.J. A vision-based formation control framework. *IEEE Trans. Robot. Autom.* **2002**, *18*, 813–825. [\[CrossRef\]](#)
18. Reynolds, C.W. An evolved, vision-based behavioral model of coordinated group motion. In *From Animals to Animats 2: Proceedings of the Second International Conference on Simulation of Adaptive Behavior*; MIT Press: Cambridge, MA, USA, 1993.
19. Vársárhelyi, G.; Virágh, C.; Somorjai, G.; Nepusz, T.; Eiben, A.E.; Vicsek, T. Optimized flocking of autonomous drones in confined environments. *Sci. Robot.* **2018**, *3*, eaat3536. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Wu, J.; Luo, C.; Luo, Y.; Li, K. Distributed UAV swarm formation and collision avoidance strategies over fixed and switching topologies. *IEEE Trans. Cybern.* **2021**, *52*, 10969–10979. [\[CrossRef\]](#)
21. Sun, G.; Zhou, R.; Ma, Z.; Li, Y.; Groß, R.; Chen, Z.; Zhao, S. Mean-shift exploration in shape assembly of robot swarms. *Nat. Commun.* **2023**, *14*, 3476. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Niizato, T.; Gunji, Y.P. Metric–topological interaction model of collective behavior. *Ecol. Model.* **2011**, *222*, 3041–3049. [\[CrossRef\]](#)
23. Bialek, W.; Cavagna, A.; Giardina, I.; Mora, T.; Silvestri, E.; Viale, M.; Walczak, A.M. Statistical mechanics for natural flocks of birds. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 4786–4791. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Gupta, L.; Jain, R.; Vaszkun, G. Survey of important issues in UAV communication networks. *IEEE Commun. Surv. Tutor.* **2015**, *18*, 1123–1152. [\[CrossRef\]](#)
25. Bastien, R.; Romanczuk, P. A model of collective behavior based purely on vision. *Sci. Adv.* **2020**, *6*, eaay0792. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Davidson, J.D.; Sosna, M.M.; Twomey, C.R.; Sridhar, V.H.; Leblanc, S.P.; Couzin, I.D. Collective detection based on visual information in animal groups. *J. R. Soc. Interface* **2021**, *18*, 20210142. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Li, K.; Li, L.; Groß, R.; Zhao, S. A collective perception model for neighbor selection in groups based on visual attention mechanisms. *New J. Phys.* **2024**, *26*, 012001. [\[CrossRef\]](#)
28. Zhou, Y.; Quan, L.; Xu, C.; Xu, G.; Gao, F. RE-Formation: Resilient and Efficient Formation Planning in Large-Scale Distributed Aerial Swarms. *IEEE Trans. Autom. Sci. Eng.* **2025**, *22*, 21212–21227. [\[CrossRef\]](#)
29. Hou, J.; Zhou, X.; Pan, N.; Li, A.; Guan, Y.; Xu, C.; Gan, Z.; Gao, F. Primitive-swarm: An ultra-lightweight and scalable planner for large-scale aerial swarms. *IEEE Trans. Robot.* **2025**, *41*, 3629–3648. [\[CrossRef\]](#)
30. Yao, D.; Wang, X.; Lin, H.; Wang, Z. Leader trajectory planning method considering constraints of formation controller. *J. Syst. Eng. Electron.* **2023**, *34*, 1294–1308. [\[CrossRef\]](#)
31. Zhao, C.; Liu, J.; Sheng, M.; Teng, W.; Zheng, Y.; Li, J. Multi-UAV trajectory planning for energy-efficient content coverage: A decentralized learning-based approach. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 3193–3207. [\[CrossRef\]](#)
32. Shan, S.; Shao, J.; Zhang, H.; Xie, S.; Sun, F. Research and validation of self-driving path planning algorithm based on optimized A*-artificial potential field method. *IEEE Sens. J.* **2024**, *24*, 24708–24722. [\[CrossRef\]](#)

33. Bounini, F.; Gingras, D.; Pollart, H.; Gruyer, D. Modified artificial potential field method for online path planning applications. In Proceedings of the 2017 IEEE Intelligent Vehicles Symposium (IV), Los Angeles, CA, USA, 11–14 June 2017; pp. 180–185. [\[CrossRef\]](#)
34. Zhou, L.; Li, W. Adaptive Artificial Potential Field Approach for Obstacle Avoidance Path Planning. In Proceedings of the 2014 Seventh International Symposium on Computational Intelligence and Design, Hangzhou, China, 13–14 December 2014; Volume 2, pp. 429–432. [\[CrossRef\]](#)

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.