***Evaluating Vision-Based Drone Swarms Performance under Visual Occlusions***

Oleksii Smovzhenko, Andrii Pysarenko

National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”

Kyiv, Ukraine

*Abstract.* We evaluate the performance of uncrewed aerial vehicle (UAV) swarms under visual occlusions caused by neighboring agents. We propose a simplified visibility model and performance metrics for evaluation. Our findings show that neighbor selection based on Delaunay triangulation outperforms other strategies for larger swarms, maintaining stable formation and alignment. Results are validated in simplified point mass dynamics and more realistic quadrocopter dynamics environments.

Keywords: UAV swarm, vision-based localization, formation control, decentralized coordination, artificial potential field.

# Introduction

An uncrewed aerial vehicle (UAV) swarm is a collective of UAVs that operate collaboratively to achieve common objectives. Unlike individual UAVs, swarms leverage collective behavior to perform tasks efficiently and address complex problems. These swarms are designed as fully distributed systems, where each UAV perceives its local environment to fulfill designated tasks, contributing to global swarm objectives. Recent reviews highlight their increasing deployment in applications such as agriculture, construction, logistics, search and rescue, humanitarian aid, telecommunications, environmental monitoring, and entertainment. However, UAV swarms face problems, including communication limitations and dependency on the Global Navigation Satellite System (GNSS), which may introduce a single point of failure.

Vision-based localization enables each UAV to estimate the position and orientation of nearby agents using onboard visual sensors, without relying on external infrastructure like GNSS. This approach supports decentralized coordination, allowing swarms to operate in GNSS-denied, indoor, or cluttered environments [1-2]. Our study introduces a novel simplified visibility model to assess how visual occlusions affect swarm performance in dense formations, a considerable issue for scalable swarm operations.

This research evaluates the impact of visual occlusions on UAV swarms relying solely on vision-based localization. We propose a simplified visibility model and analyze the effectiveness of neighbor selection strategies (metric, topological, and Delaunay) in mitigating occlusion-related problems. The paper is organized as follows: section 2 describes the methodology, including the visibility model and neighbor selection strategies; section 3 presents simulation results and comparisons; section 4 discusses findings in the context of prior work; and section 5 concludes with implications and future directions.

# Method

We simulate a vision-based UAV swarm navigating collision-free toward a goal, modeling agents as homogeneous entities in a 3D space. We consider a set of  agents labeled by . The set excluding agent  is defined as . Each agent's state is described by its position and velocity . The relative position of agent  with respect to agent  is:

 (1)

where , and  denotes the Euclidean norm. This equation defines the spatial relationship between agents, essential for localization. An adjacency between agents  and  is denoted , represented in an adjacency matrix  with entries of one if  and zero otherwise. This matrix captures connections within the swarm. The motion of each agent at step  is:

 (2)

where  is the velocity at step , and  is the time step.

Various algorithms exist for UAV swarm control, including reinforcement learning [3]. However, we use an artificial potential field inspired by Reynolds’ flocking algorithm [4], due to its lower hardware requirements compared to machine learning approaches. The potential field balances attraction for cohesion and repulsion for separation. The velocity command for an agent is:

 (3)

This combines social (cohesion and separation) and migration (goal-directed) components. The social component is:

 (4)

where  and  are gains regulating cohesion and separation, and  is the set of neighbors. Cohesion pulls agents together, while separation prevents collisions. The migration component is:

 (5)

where  is the relative position of the goal, and  is the migration gain. This directs the swarm toward the target. The final velocity is limited by:

 (6)

This ensures speeds do not exceed a maximum threshold.

The neighbor set  is defined based on selection strategies. Metric neighbor selection includes agents within a maximum radius :

 (7)

Vision-based neighbor selection further restricts this to agents within  and not occluded by others. Treating agents as spheres, occlusion is simplified to a planar check among three agents' centers:

 (8)

where  and  are angular half-sizes of agents  and  from agent 's center, and  is their angular separation. This model captures visual perception constraints.

Delaunay selection uses Delaunay triangulation, where agent  is a neighbor of  if they share a ridge:

 (9)

where  is the set of agent positions, and  is the triangulation. This typically yields up to 12 neighbors in 3D space. Topological selection includes the  nearest neighbors:

 (10)

We assess swarm performance using three metrics computed at each time step. The minimum nearest neighbor distance, indicating collision avoidance, is:

. (11)

A collision occurs if , where  is the agent radius. Alignment, measuring velocity vector correlation, is:

 (12)

A value of one indicates perfect alignment, zero indicates disorder. The union metric, assessing swarm connectivity, is:

 (13)

where  is the number of connected components in the adjacency matrix. A value of zero indicates complete fragmentation.

We conducted 10 migration experiments to evaluate swarm performance across neighbor selection strategies. For topological selection, , matching the average Delaunay neighbor count. Agents are spawned randomly in a cube with predefined density, assigned a constant migration velocity  along the horizontal axis. Performance metrics are averaged over the last 25% of steps. Parameters are listed in Table 1.

We combine metric, topological, and Delaunay strategies with vision-based selection to evaluate vision-based localization. Two simulation environments are used: a point mass environment for statistical analysis (up to 150 agents) and Gazebo with PX4 for realistic quadrocopter dynamics (up to 70 agents, due to hardware limits). Gazebo's lockstep feature synchronizes velocity commands. The Gazebo environment validates point mass results in a realistic setting.

Table 1

Experimental parameters

| Description | Notation | Value |
| --- | --- | --- |
| Agent radius |  | 0.25 m |
| Perception radius |  | 10 m |
| Maximum neighbors |  | 12 |
| Maximum speed |  | 1 m/s |
| Separation gain |  | 1 m/s |
| Cohesion gain |  | 2 m/s |
| Migration gain |  | 0.5 m/s |
| Time delta |  | 0.1 s |
| Simulation duration |  | 120 s |

##### Results

We evaluated swarm performance across sizes  and neighbor selection strategies in a point mass simulation to identify the most effective strategy. Results are shown in Fig. 1.

Purely vision-based selection degrades as swarm size increases, with reduced average minimum distance and alignment. This can lead to collisions and route deviations, increasing battery usage. For swarms larger than 30 agents, visual occlusions and frequent neighbor changes cause trajectory instability. Vision+metric and vision+topological strategies maintain stable alignment and neighbor counts but suffer from reduced union, indicating fragmentation. Vision+metric shows fragmentation from 10 agents, while vision+topological shows it from 30 agents. Vision+Delaunay provides consistent performance for swarms of 50+ agents, maintaining stable minimum distance, alignment, and neighbor count with perfect union. However, for swarms with fewer than 30 agents, it exhibits lower alignment due to insufficient neighbors. Overall, vision+Delaunay excels for larger swarms, despite slightly lower alignment.

We validated the vision+Delaunay strategy in Gazebo with quadrocopter dynamics for . Results are shown in Fig. 2. The close agreement between environments confirms that vision+Delaunay enables effective formation control in realistic settings.

A group of graphs on a white background

AI-generated content may be incorrect.

1. Performance metrics (minimum distance, alignment, union, neighbor count) for different neighbor selection strategies across swarm sizes in point mass simulations

A comparison of graphs and charts

Description automatically generated with medium confidence

1. Comparison of performance metrics (minimum distance, alignment, union) between point mass and quadrocopter dynamics for vision+Delaunay across swarm sizes

##### Discussion

The vision+Delaunay strategy outperforms others for large swarms, as it mitigates occlusion effects by prioritizing neighbors forming triangulation ridges, ensuring robust connectivity [2]. Vision+metric and vision+topological strategies, while simpler, lead to fragmentation, as seen in reduced union metrics, consistent with findings in [1]. The poor performance of vision+Delaunay for small swarms (less than 30 agents) likely stems from insufficient neighbor counts, limiting alignment, a problem also noted in [2] for sparse formations. Compared to prior work, our simplified visibility model reduces computational complexity by approximating occlusions in a planar framework, unlike the more complex models in [1]. However, the model assumes spherical agents, which may not capture real-world UAV shapes, and the artificial potential field requires careful parameter tuning, as noted in [4].

##### Conclusion

We evaluated UAV swarm performance under visual occlusions using a novel simplified visibility model, focusing on vision-based localization. Simulations across swarm sizes reveal that vision+metric and vision+topological strategies risk fragmentation, while vision+Delaunay enhances performance for swarms with 50+ agents, maintaining connectivity and stability, though it underperforms for smaller swarms. The model's planar occlusion check and spherical agent assumption may limit its applicability to complex UAV shapes, and the artificial potential field's sensitivity to tuning poses problems for real-world deployment. These findings suggest vision+Delaunay is well-suited for large-scale applications like search and rescue in cluttered environments. Future work could explore reinforcement learning [3] to improve scalability and address occlusion problems, or refine the visibility model to account for non-spherical agents.

##### References

1. F. Schilling, F. Schiano, and D. Floreano, “Vision-Based drone flocking in outdoor environments,” *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 2954–2961, Feb. 2021
2. F. Schilling, E. Soria, and D. Floreano, “On the Scalability of Vision-Based Drone Swarms in the Presence of Occlusions,” *IEEE Access*, vol. 10, pp. 28133–28146, Jan. 2022
3. Y. Albrekht and A. Pysarenko, “Exploring the power of heterogeneous UAV swarms through reinforcement learning,” *Technology Audit and Production Reserves*, vol. 6, no. 2(74), pp. 6–10, Dec. 2023
4. C. W. Reynolds, “Flocks, herds and schools: A distributed behavioral model,” *ACM SIGGRAPH Computer Graphics*, vol. 21, no. 4, pp. 25–34, Aug. 1987