Vision-Based Collective Motion: A Locust-Inspired Reductionist Model

David L. Krongauz, Amir Ayali, Gal A. Kaminka

Abstract

Collective motion is an omnipresent, fascinating phenomenon. Swarming individuals aggregate, coordinate and move, utilizing only local social cues projected by conspecifics in their vicinity. Major theoretical studies assumed perfect information availability, where agents rely on complete and exact knowledge of inter-agent distances and velocities. However, the sensory modalities that are responsible for the acquisition of environmental information were often ignored. Vision plays a central role in animal perception, and in many cases of collective motion, it is the sole source of social information. Here we investigate a vision-based collective motion model inspired by locust marching bands. We address two major challenges: the evaluation of distance and velocity and visual occlusions. We compare three strategies an agent can use to interpret partially occluded visual information. These differ in the visual cognition capabilities of the agent and the respective computational requirements. *In silico* experiments conducted in various geometrical conditions show that the three approaches display different rates of convergence to an ordered state: The least computationally-demanding approach, in which no peer recognition takes place, shows slower convergence in geometrically-constrained environments. Our findings provide insights into the visual processing requirements from biological as well as artificial swarming agents in complex settings.

1. Introduction

Collective motion refers to the emergent coordinated and cohesive movement of individuals without centralized (group-wide) control, perception, and communications. In nature, it is ubiquitous in birds [1, 2], fish [3, 4], insects[5, 6], human crowds barnet16,wolf79, and even bacteria [7]. It is a phenomenon that has been of interest to the scientific community for decades, inspiring investigations in mathematics and physics [8, 9], as well as graphics, simulation, and robotics [10, 11, 12, 13].

The leading paradigm is that collective motion results from repeated local myopic interactions between individual swarm members (see, e.g., [9]). Upon perceiving its local physical and social environments, each individual acts using a motion-control procedure [14]. This approach is often modeled by '*self-propelled particles*' (SPP), where simulated agents navigate while being subjected to mutual steering forces, caused by interactions with their neighbors [15]. These interactions feed into each agent's

decision-making scheme, changing its motion [16, 17, 18]. Under the appropriate conditions, this allows the generation of collectively ordered motion, where agents move in approximately a single spatial direction, similar to flocking[1] animals [19, 20, 21].

Collective motion is, therefore, dependent on the swarm members' individual capacity for acquiring and perceiving complete, reliable, and constantly updating information about their social environment. In particular, existing models assume perfect knowledge of neighbors' distances, headings, and velocities (see, for instance, [16, 22]). This is an idealization of natural capabilities for species relying on visual perception as the dominant sensory modality, for two major reasons: First, many species cannot reliably estimate distances (even using stereoscopic vision); and second, partial occlusions of swarm members by others necessarily challenge models that assume perfect knowledge of the social environment.

Estimation of distance based on visual perception is possible but makes certain requirements of the sensory abilities of the observing entity. Some animals can use their binocular, stereoscopic vision capabilities for calculating distance [23]. Many others, lacking the anatomical-physiological required mechanisms, can try making up for this limitation by moving while obtaining the brief visual inputs, thus creating an impression of depth (in computer science, this is called generate *structure from motion*). A key mechanism that is common in animal motion and navigation is the optical flow, which is in mathematical terms the vectorial difference between two frames an individual "sees" [24]. However, as described in [25], it cannot be directly used to estimate distance. Another related mechanism is sensitivity to looming objects in the vicinity [26, 27]. We will elaborate on these issues in Section 2.1.

Occlusions raise a second, independent, challenge. Neighbors that are closer may partially or fully occlude others who are farther, in many perceptual modalities. Using visual modality becomes a central issue in creating a picture of the environment as the density of group members increases.

Recently, several studies have attempted to elucidate the problem of sensory perception in collective behavior modeling [28, 29, 30]. Some of these investigate the role of the visual field in self-organizing systems and vision-based navigation, both in natural [3, 31, 28, 32] and simulated environments, either *in silico* [33, 34, 35, 28, 36, 29, 30] or implemented via robots [37, 38].

Soria et al. [36] explored the effect of a limited visual field on the performance of the Reynolds flocking algorithm [10]. Rosenthal et al. [3] investigated the importance of the visual field (e.g., "angular area") in collective evasion maneuvers manifested in schooling fish. Nonetheless, they also employed globally-acquired information, including exact metric and topological distances between the swarming fish. An extension of the Reynolds model [10] is presented by Kunz et al. [33]. Although their model deals with obstruction of further distanced neighbors, it does not account for possible partial obstruction, i.e., simplifying the perception to only seeing the closest neighbors

[1] We use *flocking* and *collective motion* interchangeably. Flocking describes the emergence of a common movement direction for all the agents out of uncorrelated initial movements. Once the coherent motion is achieved on a group level, it should persist in time. A common heading direction with a narrow spread thus emerges and defines this asymptotic flocking state.

of each agent. Moreover, neighbors' exact distances and headings are employed to calculate the motion of each agent. The work of Collignon et al. [28], though not dealing with flocking, studies zebrafish-inspired agents that move in a 2D environment. To solve the problem of range estimation, they introduce higher dimensional solid angles representing the original 3D characteristics of the simulated fish.

The vision-based model of collective motion presented by Bastien et al. [29] and extended to obstacle avoidance by Qi et al. [30] offers a different approach for perception modeling. They introduce specifically designed complex integro-differential operators that, depending on the visual input, govern an agent's motion, which in turn creates various collective behaviors.

To avoid the problem of ranging, this work and others restricted themselves to a special case of disk-shaped agents [37, 38]. However, the natural morphologies of locusts and fish are elongated and very different from circular disks. Wang et al. [39] demonstrated implicit cooperation between robotic swarm members using visual inputs. Utilizing an elaborate scheme of positioning poles and sets of LED lights, positioned in known distances over the poles, the robots are able to estimate the relative positioning, velocity, state, and other features of their neighbors. Using the acquired information, the robots exhibit, various swarming behaviors, including flocking (collective motion), formations and foraging.

Indeed, the models above either employ exact global information that is unavailable for natural animal perception, assume special disk-shaped agents, which weakens their connection to nature where morphologies are far from the disk shape, or employ more sophisticated visual processing by combing several visual cues for estimations of relative distance and velocity.

In contrast, this paper presents a reductionist approach, which on the one hand stays true to the elongated morphology of natural animals, but on the other reduces the visual inputs to the bare minimum perceivable in two dimensions (i.e., no vertical angles, no visual markers at different heights). We draw our inspiration from marching locust nymphs, a quintessential example of collective motion, by an organism lacking binocular depth perception, and further characterized by a substantially simple and small brain (compared to other flocking organisms as birds and fish); see [15] (and references within). In natural conditions, locust swarms constitute vast numbers of individuals, demonstrating coordinated and synchronized mass-motion [40, 41, 42]. Locust swarms lend themselves to controlled laboratory experiments that were successful in providing valuable insights into the mechanism of collective motion, and specifically, the interactions between the individual and the group (e.g., Ariel et al. [21], Bazazi et al. [43], Knebel et al. [44]). Based on these and other reports, it is largely accepted that vision is a key sensory modality underlying the local interactions among a group of locusts. The individual locust makes its decisions based on visual information which lacks direct knowledge of distance, has reduced information of neighbors' headings, and does not offer precise velocity information. Previous work has established that the Pause-and-Go motion scheme is central to the repeated decision-making underlying the behavior of locusts in a swarm [15, 21, 45], i.e., a representation of the local environment, utilized for deciding if and at what direction to move, is constructed by the locusts when standing. This makes occlusions, as well as distance and velocity assessments, major hindrances.

From a signal processing perspective, vision is a complex, multifaceted process. Without dedicated processing, it can result in information loss during acquisition and significant computational errors introduced in the perception process. We should, therefore, limit and simplify complex properties by adopting a reductionist model. As elaborated in Section 2, we propose a model of vision-based collective motion inspired by the flocking behavior of migrating locusts. We use the geometrical characteristics of visual perception to present a control algorithm solely based on the information accessible via the agent's visual field. Additionally, we examine the effect of occlusions on visual perceptions and propose several approaches that the focal agent might use when faced with visual stimuli from partially obstructed neighbors.

Experiments performed *in silico* demonstrate group-order convergence from disordered to coherent movement. In Section 3 we explore the resultant group dynamics for various group sizes, vision radii, and steering-parameter factor values. Furthermore, we compare the different strategies for addressing occlusions and study their influence on the flocking rate. The swarm's order parameter, measured by the level of alignment between the agents, reaches high asymptotic values for all occlusion approaches when tested in a toroidal arena. However, in arenas with a higher degree of constraints, the rate, and degree of alignment are lower for the least computationally demanding strategies.

2. A Reductionist Model of Visual Perception for Collective Motion

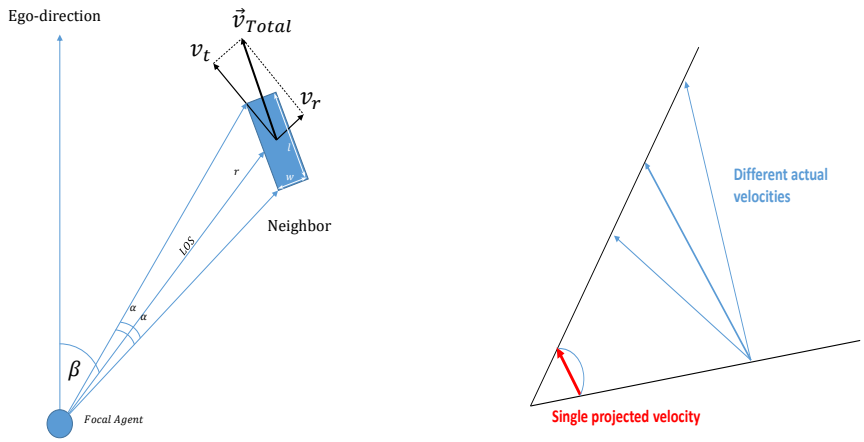
We present a vision-based collective motion model. First, in Section 2.1 we define our simplification to visual perception, and the vision-based algorithm governing the agents' movement. Next, in Section 2.2, we add occlusions of individuals to the model of vision. Three alternative approaches of interpreting and dealing with partially occluded visual information are presented.

2.1. The Principal Vision-Based Reductionist Model

The agent's field of view is assumed to be omnidirectional (360 degrees) and is limited to a constant radius R around the focal agent. This assumption is inspired by locust vision, which is considered to be almost omnidirectional [26, 46].

As depicted in Figure 1a, focal agent P heading "up" with a single neighbor n moving with velocity $\boldsymbol{v_{Total}}$ and located at a distance $r < R$ measured between the COMs of the two agents. The total velocity is combined from the tangential $\boldsymbol{v_t}$ and radial $\boldsymbol{v_r}$ velocity components relating to the line of sight (LS), a line connecting the center-of-mass (COM) of the neighbor and the focal agent. The angular position of the neighbor relative to the heading direction is denoted as (bearing, β) and the angular area is subtended on the vision sensor α n practice. This angle is calculated by finding the vectors from the focal agent towards the vertices on the edge, that subtends the largest angle on the sensor; we shall denote it as the *effective edge*. For convenience, the full subtended angle is regarded as 2α.

Animals with stereoscopic vision, perceive their environment as three-dimensional via processes in the brain, such as *parallax* that rely on both of our eyes focusing on the same object from two slightly different positions, or *motion parallax*, where the observed object remains static while the eye moves. However, for each of our eyes, light



(a) Schematic representation of a focal agent, a neighboring agent, and its visual parameters as observed by the focal agent.

(b) Projection of velocities. Inherent information loss in visual acquisition of velocities.

Figure 1

rays are reflected from a three-dimensional object and hit a two-dimensional retina. In other words, visual stimuli is one dimension less than the object it originated from. Applying the same logic to our world of two-dimensional agents, visual input would have to be one-dimensional. Essentially, a whole dimension of information about the environment is lost when employing vision. The same goes for our simulated two-dimensional world and the way we define vision in the simulation: since the agents are two-dimensional, the visual information an agent perceives, using its non-stereoscopic visual sensor, is one-dimensional. The geometric angles explained above constitute, by definition, this one-dimensional type of vision.

Due to the loss of information, the principal challenge when relying on vision is estimating the velocity heading and magnitude of each neighbor. This requires knowledge of the distance to a neighbor [25]. However, it is generally impossible to infer inter-agent distances from projections without additional information. As shown in Fig. 1b, different actual velocities may be projected to identical observed velocity vectors and distances. Previous works have often assumed a special case, where the agents are circular. This case allows precise calculation of inter-agent distances from projective information since, for disks, distances are in a one-to-one correspondence to the projected visual area. However, as the realistic morphologies of swarming locusts are elongated, we do not make this assumption here.

In our study, we deliberately set our agents to be of elongated shape, in our case rectangular. We consider a group of N rectangular agents with width w and length l, moving in a two-dimensional environment at speed v with heading parallel to their length axis. The maximal speed the agents can accumulate is v_{max}. The position coordinates $\boldsymbol{x_i}$ of the agent i are updated at discrete time-steps according to the motion

equation,

$$x_i\left(t+\Delta t\right)=\ x_i\left(t\right)+\ v_i\left(t\right)\cdot\Delta t, \tag{1}$$

with velocity $v_i(t)$ updated at each time-step causing the agent to steer towards a desired velocity with steering-parameter factor η,

$$v_i\left(t+\Delta t\right)=v_{desired}\left(t\right)\cdot\eta+\ v_i\left(t\right)\cdot\left(1-\eta\right), \tag{2}$$

where $v_{desired}$ is calculated based on a decision algorithm similar to the Vicsek Model[16]. The objective of the algorithm is to estimate the velocities of the neighbors at each time frame, based on visually available geometrical information. Then, the average of these velocities will constitute as $v_{desired}$.

We start our calculation of $\boldsymbol{v}$ by estimating the distance r. Since the orientation of the neighbor is unknown to the observer, we deliberately use a simplistic approximation: the neighbor's effective edge is always perpendicular to the LS, which means the triangle comprised of the focal point P and the two vertices of the effective edge d (see Fig. 1a) is assumed to be an equilateral. Thus, the LS constitutes both median and altitude to effective edge and bisector of the subtended angle 2α (See Fig. 1a). Therefore, the distance r is given by

$$r=\frac{1}{2}d\cot\alpha, \tag{3}$$

differentiating Eq. 3 we get the radial velocity to be

$$v_r=\frac{d}{dt}r=-\frac{1}{2}\frac{d\,\dot{\alpha}}{\sin^2(\alpha)} \tag{4}$$

where $\dot{\alpha}$ denotes the time derivative of the subtended angle. Expressing $d=2r\tan\alpha$ from Eq. 3 and substituting into Eq. 4 results in

$$v_r=-\frac{2\dot{\alpha}r}{\sin 2\alpha}, \tag{5}$$

the tangential velocity is obtained using the time derivative of the bearing angle

$$\begin{aligned}\dot{\beta}=\frac{d}{dt}\beta=\omega=\frac{v_t}{r}\\ \Rightarrow v_t=-\dot{\beta}\cdot r,\end{aligned} \tag{6}$$

where $\dot{\beta}$ is the derivative in time of the bearing angle, which is equivalent to the instantaneous angular velocity ω. Time derivatives are approximated by the difference of parameter values in two consecutive frames. By calculating both components, we get $v=v_r+v_t$. This process is repeated for all the neighbors, resulting in $v_{desired}$ of the focal agent.

We assume every flock member possesses knowledge of the conspecific's size (effective edge d is used in the calculation), i.e., an animal has an inherent representation

of the typical dimensions of individuals from its species [47]. Combining this knowledge with the angular area angle (α) one can estimate the distance (as shown in Eq. 3). The α angel is also present in loom calculations [46, 48, 49, 50].

We emphasize that this is a baseline model. It assumes all the neighbors are fully visible and does not account for possible obstructions of sight. In other words, agents are presumed to be "transparent", in the sense that no obstruction of farther neighbors occurs. Obviously, this assumption is a gross oversimplification of the fundamental capabilities and limitations of visual perception in nature. In the next section, we explore possible general ways to relax this assumption.

2.2. Addressing Occlusions: Three Approaches

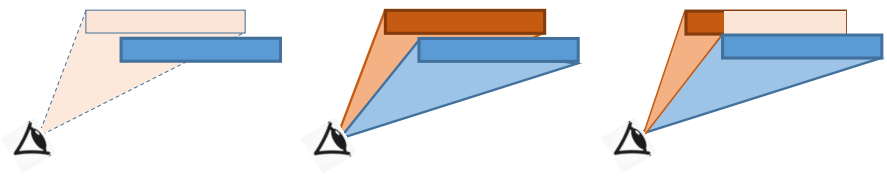
Occlusions are an inherent part of the visual modality in natural systems. Animals often swarm in dense flocks [51], and thus conspecifics located closer to the observing animal are inevitably blocking, partially or entirely, the animals standing behind them.

Figure 2: The visual social environment from the point of view of the individual locust in a swarm; a composite illustration.

Complete and partial occlusion of neighbors not only reduces the information available to the focal agent but can also introduce large estimation errors. Neighbors that are completely occluded are not taken into account in the collective motion model. However, partially-occluded neighbors generate additional errors, as the projected area of the subtended angle is used as a proxy for distance. For example, suppose a neighbor is partially occluded, such that only a small portion of it is observed, and thus it may be perceived to be distant. If the occluding animal moves to uncover it, its full length is now revealed, and it will now be seen as being close within a very short time, implying high velocity towards the observer and a potential collision. The accumulation of such frequent errors may disturb the stability of the swarm.

In accounting for visual occlusions, we posit there are three general strategies that may be taken (illustrated in Fig. 3). Suppose the focal agent may be able to recognize peers, and thus differentiate between entirely-visible individuals and parts (that is, not

complete individuals). This allows it to ignore partially-visible neighbors (Fig. 3-a). It may also be able to cognitively extrapolate parts to a whole, inferring the position and orientation of the partially-occluded peer from its visible parts (Fig. 3-b). Or, as an alternative, it may perceive any visible part of a neighbor as a distinct whole individual. These strategies place very different requirements on the cognitive-computational processes of visual perception in the focal agent, as discussed in detail below.



(a) OMID (b) COMPLID (c) PARTID

Figure 3: Occlusion methods. (a) OMID – partially occluded neighbor (orange) is omitted from the field of view. (b) COMPLID – Orange neighbor is completed from the seen segment. (c) PARTID – partially seen segment is regarded as a neighbor.

Approach 1: Omission of the Occluded (OMID).. The first approach disregards any visual information originating in partially occluded agents (see Fig. 3a). This requires the animal to possess a dedicated *peer recognition* mechanism, i.e., be able to recognize fully-imaged conspecifics (and ignore anything else). Mechanisms of selective attention in visual perception are known to exist in humans and are achieved in the human brain in multiple stages of perception [52, 53]. Neurobiological studies showed the existence of selective attention, as well as target detecting abilities amid cluttered backgrounds, also in insects' visual neurons [54, 55]. Such visual processing mechanisms may, in principle, be utilized during collective motion.

It is not known whether locust visual perception mechanisms are able to recognize peers. Experiments reported in [56] show that an individual locust responds by walking when exposed to visual images composed of randomly-moving dots that are projected via computer screens to both eyes. As the dots are positioned randomly and do not mimic the shape or the colors of locust nymphs, these results seem to indicate that the motion is triggered in the individual sans a dedicated peer recognition mechanism.

Approach 2: Completion of the Occluded (COMPLID).. In the second approach, partially occluded agents are "completed" as if they are fully visible to the observer. In other words, a neighbor that has even the smallest visible segment from the focal agent's perspective would be treated as if no occlusion is present when processing its visually extractable information. COMPLID utilizes *peer recognition* as in OMID, in addition, it requires that the agents will be able to assess the obscured part of their neighbors (if needed) based on the visible part of a neighbor. This completion assumes an agent's visual extrapolation that reconstructs neighbors' outlines using their visible features.

Completing partially visible targets obscured by other objects is a long-studied process in visual perception. The filling-in of details and image regions partially obscured by interceding objects [57, 58] is an established neurophysiological process that gives the organism an ability to identify a complete form based upon observed parts of the contour and is described by the term "visual completion" [59]. This mechanism produces an internal representation called "illusory contour", which extrapolates the physical stimulus to the full geometrical shape of the object [60, 61, 62]. Visual completion of occluded objects has been shown in varied and phylogenetically distant species: birds, fishes, cephalopods, bees, etc., and is accepted as one of the fundamental components of vision in nature [60, 63, 64]. Hence, the second approach is to extrapolate partially occluded neighbors to complete the full image.

Approach 3: Every Part is a Full Agent (PARTID). The third approach treats all visual stimuli related to a neighbor as if they represent a full-body conspecific. Contrary to OMID and COMPLID, in this approach agents lack *peer recognition* abilities. Instead, the visual field is divided into segments, where each segment contains the same optical flow vectors. The agent assumes that each segment represents a different neighbor. In other words, any visual information is considered complete at face value without any additional interpretation. Hence, apart from the ability to accurately extract optical flow vectors, no further advanced visual perception mechanisms are required. Since the optical flow is essentially the vectorial difference between two consecutive frames and does not consist of any form of object recognition by itself, PARTID would be the least computationally demanding approach if implemented in real life.

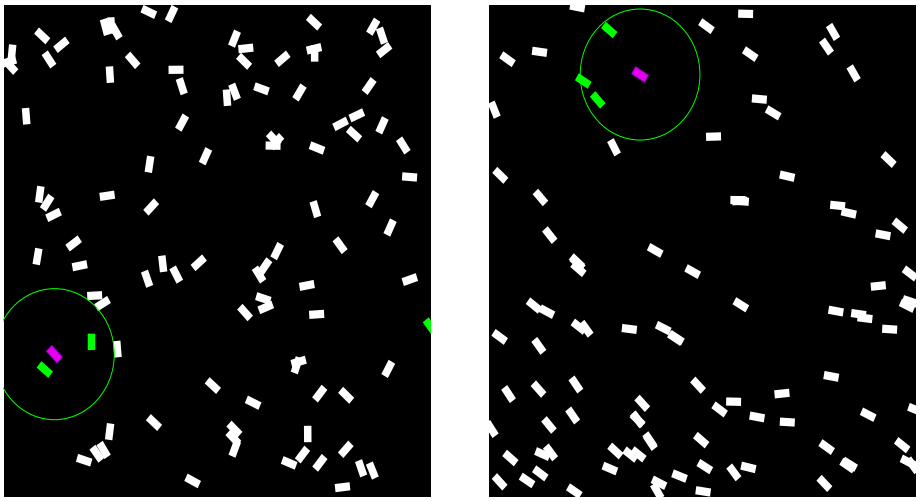
However, in this approach, the potential error in the assessment of the environment is the largest, as compared to OMID and COMPLID, since partially occluded agents are taking less area on the visual field, which translates to a significantly larger distance estimation. The same goes for velocity estimations, which are tightly dependent on the distance. Although in this approach an agent is acquired with object recognition abilities, it is assumed that observable parameters $(\alpha, \dot{\alpha},\ \beta,\ \dot{\beta}\ )$ are still fully available and extractable. As noted, this approach requires relatively limited visual processing and is easier to implement in robotic systems.

PARTID takes inspiration from biological mechanisms, in which an organism performs an action based on visual stimuli originating from an object that is not recognized. For example, locusts have a pair of visually-sensitive neurons that encode looming stimuli and cause the locust to produce escape behaviors [46]. The visual stimuli directly affect the behavior of the individual without passing through object recognition mechanisms [56].

3. Methods

In order to evaluate swarm behavior using different occlusion-handling approaches, we developed a two-dimensional (2D) collective motion simulator based on a basic simulation engine [65] (see Fig. 4). The principal model and the three occlusion-handling approaches were implemented in the simulation, as we describe below. As described in Section 2.1 the agent's steering is based upon the mean of the neighbors' velocities. These velocities, in turn, are derived from the angular measurements of

each perceived neighbor: the subtended angle and the angular velocity of each neighbor. These inputs serve as the agent's subjective perception.



(a) Toroidal arena snapshot at $t = 10[frames]$. Agents are initialized at random positions and random velocities. The purple-colored agent is an arbitrarily marked focal agent with its respective neighbors colored green.

(b) Toroidal arena snapshot at $t = 2000[frames]$. An apparent flocking behavior is displayed, with agents moving roughly in a single direction.

Figure 4: Simulator snapshots.

3.1. Simulating Kinematics

The agents are moving in 2D by updating their coordinates at each iteration in accordance with their current velocities. The location and orientation of each rectangular agent are computed from the coordinates of its COM. It is assumed in our model that velocity heading is always along the long axis of the body. The velocity magnitude can vary between 0 and a fixed maximal speed value, i.e., the agents can accelerate up to a maximal speed. Additionally, a steering parameter is introduced to moderate the influence of $v_{desired}$ at each point in time, thus reducing the sharpness of turns and accelerations performed by the agents. This serves as the motion backbone for all the approaches described below.

3.2. Simulating Perception

We compare the emergent collective motion resulting from the different occlusion-handling approaches. The perception of each agent is calculated from the exact values stored in the simulation. This allows the emulation of different perception approaches by varying the visual parameters of the agents and the effects of occlusions. Each simulated focal agent is given the information it would have perceived in the 2D environment, in each of the following perception approaches.

Simulating the Principal Model. The α angle is calculated using the neighbor's vertices of the edge that subtends the largest angle on the agent. The angle between

the two vectors pointing from the focal agent's COM to the respective vertices equals to α. The β angle is simply the angle between the focal agent's velocity vector and the neighbor's COM. The calculation of β is the same for all the different approaches. However, since occlusions are non-existent in the principal model, the focal agent receives visual parameters of all the neighbors, including those that are completely occluded by other more closely distanced neighbors.

Simulating OMID. We simulate this capacity by calculating the effective α for each neighbor, i.e., the subtended angle on the visual field after canceling occluded individuals, meaning that the occluded agent is ignored in its entirety during the calculation of $v_{desired}$. The effective α calculation starts by sorting all the neighbors by their exact distances, then sequentially for each neighbor, finding the angle "edges" and storing them in a 1D array representing the visual field, spanning from $-\pi$ to π. In case of an angle segment overlaps previously-stored segments, the neighbor is ignored, and the calculation proceeds to the next neighbor.

Simulating COMPLID. We simulate this capacity by calculating the neighbors' "edges" in a similar fashion to the process in the OMID approach. Then, subtended angles that are fully contained in others (i.e., capturing neighbors that are completely occluded by others in the visual field) are eliminated.

Simulating PARTID. The implementation of the last approach in the simulation starts similarly to the OMID and OCLID approaches. We iterate over the neighbors, from the closest to the furthest. Each neighbor's effective edge is calculated and then checked against an array of edges. If a partial overlap (i.e., partial occlusion) exists with the current edge and one or two of already checked edges, the effective α is calculated using only the non-overlapping segment.

3.3. Controlled (independent) simulation variables

In the simulation experiments, we controlled the following independent variables:

Population size (N) number of simulated agents participating in each experiment.

Agents length/width sets the agent's dimensions, that in turn changes the effective edge (d).

Vision radius (R) the radius inside which agents are considered as neighbors (measured in the same units as agent length).

Steering-parameter (η) sets the weight of $v_{desired}$ relative to the current velocity (v) of an agent.

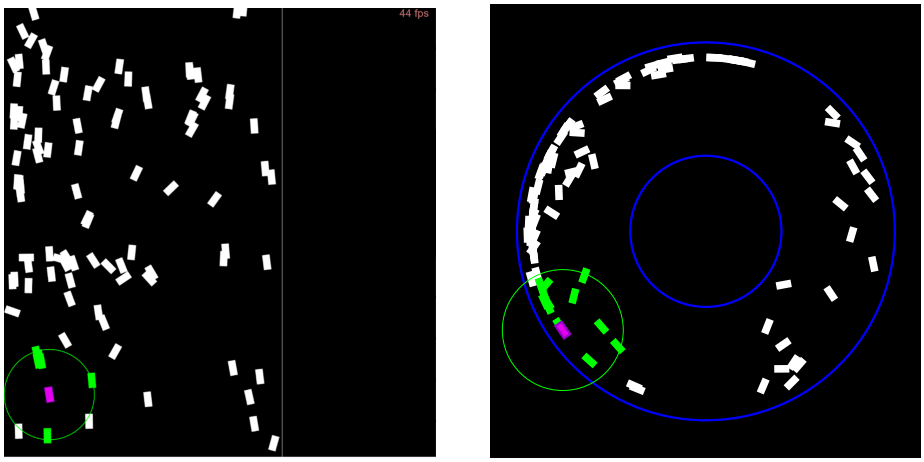
Maximal speed (v_{max}) sets the maximal speed of the agents. In case the magnitude of $v_{desired}$ exceeds v_{max} at the end of the model's calculation, the simulator will limit it to a magnitude of v_{max} while keeping the same vector direction.

We utilized different areas (*arenas*) to test the agents. The key difference between these is the periodicity of the boundaries:

- The first type (shown in Fig. 4) is a square arena with periodic boundary conditions (torus), i.e., once an agent's COM passes the maximal/minimal coordinates or the X/Y axes, it reappears on the other side respectively. The arena dimensions are L, W, which stand for the length and width of the arena.

- The second type is the "infinite corridor" arena (shown in Fig. 5a). In this arena, the vertical boundaries bounce incoming agents by reflecting their velocity's x-component, while the horizontal boundaries are set to be periodic, similarly to the toroidal arena. The distances between the boundaries are varied in the experiments and detailed below.

- The third is the ring arena, modeled after the ring shaped arena widely used in *in vivo* locusts flocking experiments (e.g., [5, 66, 45], shown in Fig. 5b). The agents are repelled from the boundaries with varying repelling force, depending on the size of the radial velocity component (relative to the arena center), i.e., an agent traveling to the external circular boundary will be repelled from with a force proportional to the size of the agent's radial velocity component. This was designed in a way to mimic the behavior of living locusts that align themselves to the ring walls so as to avoid possible collision with them [66]. The dimensions of the ring arena are characterized by the two radii of the inner and outer circles that form the ring's boundaries.



(a) Snapshot of corridor arena. The vertical boundaries are repelling, while the horizontal ones are periodic.

(b) Ring arena snapshot.

Figure 5: Corridor and ring arenas.

For the first set of experiments, we consider a group of N agents, with random initial positions and velocities, positioned in a square arena of size $W \times H$ with periodic boundaries. We studied the dynamics for a range of hyperparameter values. We present the results for representative parameters' values with typical agents' body length (l) of 20 pixels, arena-size of $30 \times 30 [BL^2]$, number of swarm members of $N = 60, 120, 180,$

and maximal velocity normalized to $v_{max} = 1$. Vision radii are measured in units of body-length (BL), ranging from $R = 1$ to $R = 6[BL]$. These ranges for group size and vision radius were chosen to achieve flock density corresponding to available biological literature of common flock densities in nature [51].

3.4. Measured (dependent) simulation outcome variables

The ideal flocking is the situation in which all agents are synchronously moving in the same direction. To compare the approaches, we use the *polarization* measure of order, denoted ϕ [16, 67]. It is defined by

$$\phi = \frac{1}{N} \left| \sum_{i}^{N} \frac{v_i}{|v_i|} \right|, \tag{7}$$

where N is the population size, and $v_i, |v_i|$ correspond to the ith agent's velocity and speed. By taking the absolute value of the average heading, we get a scalar value representing at any given time the degree of global order in the system. This parameter measures the degree of global alignment by averaging the normalized velocities of the agents (i.e., headings). From now onward it will be defined as the order parameter. At maximal chaotic state, ϕ has the value 0, while for a fully flocked system, with all agents moving with an identical heading, it approaches the value of 1.

4. Results

Two sets of experiments were conducted to evaluate the approaches. The first set, presented in Section 4.1, evaluated the performance of the principal approach in different settings, testing its sensitivity to various controlled settings. The second set of experiments, reported in Section 4.2, compares the performance of the three occlusion approaches alongside the principal model in different settings and arena types.

4.1. Flocking in the Principal Model

We begin by testing the principal model in a toroidal arena (in many aspects, closest to open-field conditions), varying independent simulation variables chosen in accordance with the existing biological literature. We conducted experiments with representative hyperparameter values of $N = 120, R = 3$ body lengths [BL], $\eta = 0.01$, and *agent length* set to 20. The population size was set based on observed locust marching-bands densities [51], the sensing radius was set according to empirical observations of locust nymphs not reacting to visual stimuli located farther than 2-3 BLs [21], and the steering-parameter was chosen according to empirical analysis.

We repeated 50 independent runs. At each run, the agents' positions are initialized with random velocities and headings representing a disordered group-level state.

Figure 6 shows the time evolution of a simulated swarm's mean polarization (the order parameter ϕ). The order parameter rises from $\phi \approx 0.1$ to an asymptotic value of $\phi \approx 0.9$, which approximately approaches an ideally parallel ordered state.

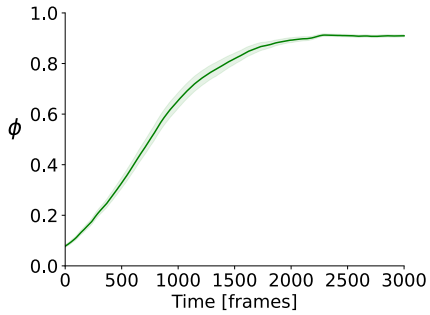


Figure 6: Swarm's alignment level vs time. Data points are the mean order parameter of the swarm at each simulation frame, with standard error margin. Experiments performed in the torus arena.

4.1.1. Influence of Vision Radius and Population Size

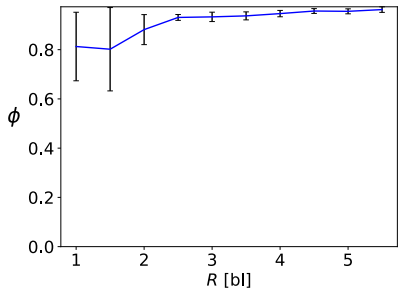
We investigate the principal model's sensitivity to various parameter values, i.e., the dependence of the order parameter ϕ on major controlled quantities: the vision radius R, and the population size N. For each of the parameters, we present two types of graphs. The first is the asymptotic parameter sensitivity, obtained by measuring the order parameter ϕ at a fixed time of $t = 3000$ [frames] for different parameter values. Empirically, it appears that by that point in time, the swarm has reached its asymptotic converged state and no significant improvement of the order parameter follows. The second type of graph is the time-evolution of the order parameter measured for various parameter values, measured at each frame until the end of the experiment at $t = 3000$ [frames]. For each global parameter, all other parameters' values were fixed arbitrarily (typical fixed values were $N = 180, R = 3 - -4$ [BL], $\eta = 0.01$, agent length=$20 - 30$). For each combination of parameters, the experiments were repeated for 50 independent runs and averaged.

Fig. 7 shows the dependence of the order parameter on the vision radius R and the population size N respectively. As seen in Fig. 7a the order parameter increases monotonically as the vision radius grows. For radii $R \geq 50 \sim 2.5\,[BL]$, at $R = 90$ the improvement of the convergence process reaches saturation and the evolution towards a fully flocked state becomes the steepest.

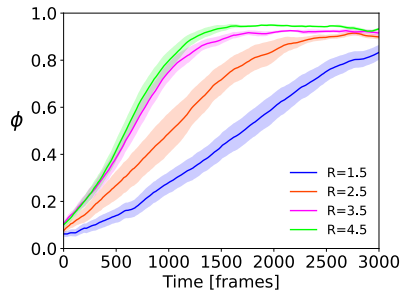
Figures 7c,7d depict the population size sensitivity analysis. As seen in Fig. 7c for relatively small N sizes, increasing the population leads to a higher asymptotic order parameter, though it reaches saturation for $N \geq 100$. Moreover, in Fig. 7d a significant qualitative difference is seen between the convergence of the order parameter for $N = 20$ compared with bigger population sizes.

4.2. Comparison of the three occlusion approaches

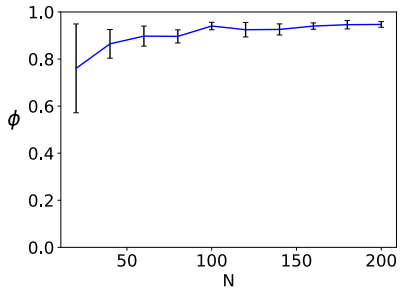
We now turn to investigate the role of occlusions by neighbors. The three approaches are evaluated in comparison with the *principal* model (which does not account for occlusions). By way of a reminder: the first approach (OMID) ignores any neighbor that is partially obstructed by another neighbor; the second approach (COMPLID) extrapolates any partially visible neighbor to its full size as if the occluding neighbors



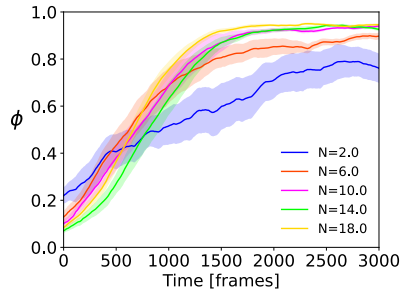
(a) Asymptotic vision radius (R) sensitivity analysis.



(b) Time-evolution of the order parameter of a swarm for different vision radii.



(c) Asymptotic population size (N) sensitivity analysis.



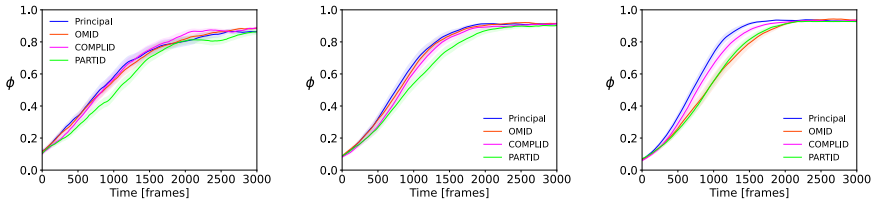
(d) Time-evolution of the order parameter of a swarm for different population sizes.

Figure 7: Vision radius and population size sensitivity analysis. For asymptotic analysis: data points represent the order parameter at t=3000 [frames] for varying R and N respective varying values, with standard error margin. For time-evolution analysis: data points are the mean order parameter of the swarm at each simulation frame, with standard error margin. Torus arena.

do not occlude; and finally, the last approach (PARTID), assumes that any unique set of visible parameters $(\alpha, \dot{\alpha},\ \beta,\ \dot{\beta}\ )$ originates in a completely visible neighbor, so the visible parts of partially occluded neighbors are treated as if they are a complete neighbor. PARTID produces the highest potentially erroneous estimations of distance and velocity of a single neighbor since the subtended angle $(\alpha, \dot{\alpha})$ parameters are greatly affected by occlusions.

Torus Arena. In Fig. 8 time-evolution in the *torus* arena of the order-parameter ϕ is compared between the four approaches. The graphs show the mean order parameter for each point in time. Data is averaged over 30 runs. Three population sizes of $N = 60, 120, 180; R = 3\,[BL], l = 30, \eta = 0.01$. As seen from the figures, all three occlusion approaches, alongside the original model, reach similar asymptotic order-parameter values ($\phi \sim 0.9$), indicative of reaching similar degrees of synchronization after flocking. It is seen from Fig. 8b and Fig. 8c that PARTID has a slower rate of convergence. It can be viewed as additional evidence that this approach results in ex-

cessively noisy perception. At higher densities (Fig. 8c) the rate of convergence of all three perceptive approaches lag behind the Principal model. At higher densities, rates of convergence become steeper. At the same time, the asymptotic order parameter remains very close for all the methods, and even different densities. Finally, the ordering of the rates of convergence at 180 indicates that COMPLID converges faster than OMID. Completing parts of neighbors, instead of omitting them, leads to an effectively larger number of neighbors, which leads to stronger alignment.



(a) (b) (c)

Figure 8: Comparison of the occlusion approaches in the torus arena. Data points are the mean order parameter of the swarm at each simulation frame, with standard error margin. Compared population sizes: (a) 60 agents, (b) 120 agents, (c) 180 agents. Convergences are practically indistinguishable between approaches. A higher population leads to a slightly steeper transition to a flocked state.

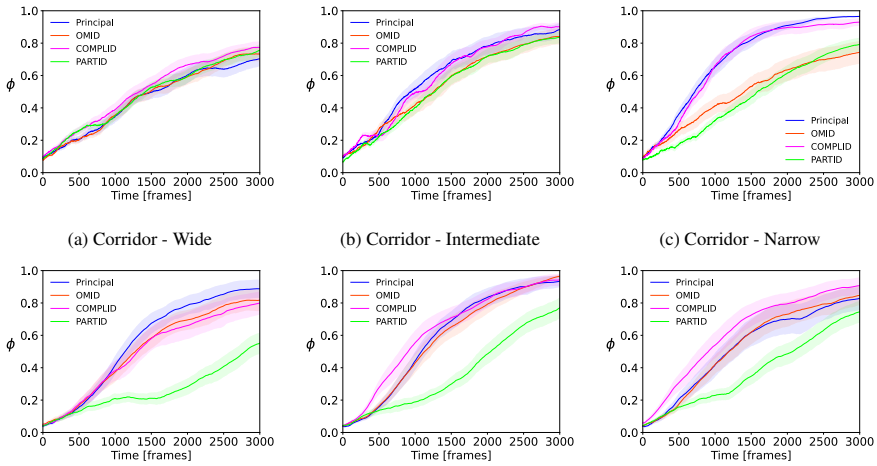
Infinite Corridor and Ring Arenas. In order to further test the model's robustness, we investigate the behavior of the different approaches (including the principal model, which does not handle occlusions), examining flocking in other geometrical environments. Specifically, we examine the behavior in an infinite corridor and in a ring environment (Figure 9) similar to ring arenas present in biological locust experiments as noted above.

Three versions of each arena type are tested: *wide, intermediate, and narrow*. The geometry of the arenas will be characterized by the ratio of arena-width to single agent body-length. In our chosen measurement units of body-length (BL) this parameter is simply the arena width. A swarm of 100 agents was run in each arena configuration. For the infinite corridor, the distance between the periodical boundaries (length) was 20 [BL] for all the experiments. The widths were: 10, 23, 33 [BL] respectively. For the ring arena the radius of the inner circle is 1.66 [BL] and the outer circle radii tested were: 5, 8.33, 11.66 [BL].

Analyzing the infinite corridor arena, we observe an interesting distinction displayed in the *narrow* corridor. All perceptive approaches show a rise in the ordering parameter: The Principal model and the COMPLID approach converge significantly faster and to a higher asymptotic value than OMID and PARTID. The fact that this difference is significant only in the narrow environment indicates that the harsher geometrical constraints cause some approaches to struggle. A potential reason is the fact that in narrow arenas, interactions with the boundaries are much more frequent. Principal and COMPLID converge better even in highly perturbed geometry, this indicates that acquiring accurate information from the maximal amount of seen neighbors is beneficial in constrained environments. As a reminder, COMPLID is able to essentially

approximate the information available to the principal model apart from the agents that are fully occluded, thus, the information it perceives is quite similar to the all-seeing Principal approach.

In the ring arena, we observe a common theme in all the different widths. Although the rest of the approaches seem to converge successfully to an ordered state with similar rates, the PARTID seems to struggle with the constraints imposed by the circular boundaries. The agents in the ring arena are faced with external perturbations by the boundaries, that interfere with the alignment of the model itself in a way that is critically detrimental to the PARTID approach.



(a) Corridor - Wide (b) Corridor - Intermediate (c) Corridor - Narrow

(d) Ring - Wide (e) Ring - Intermediate (f) Ring - Narrow

Figure 9: (a),(b),(c): Infinite corridor arena. Wide / Intermediate / Narrow arena dimensions are 20×33 / 23 / 10 [BL], respectively (agent BL 20-30). Data points are the mean order parameter of the swarm at each simulation frame, with standard error margin. (d),(e),(f): Ring arena. Wide / Intermediate / Narrow ring external border radii are 11.66 / 8.33 / 5, respectively. The internal ring border is constant for all three types and equals 50. (agent BL 20–30). Data points are the mean order parameter of the swarm at each simulation frame, with standard error margin.

5. Discussion

In this work, we have introduced a non-stereoscopic perception framework and developed a novel, purely vision-based flocking algorithm. Taking a reductionist approach, we considered 2D agents, aiming to generate a collective motion with the bare minimum of information. This model employs geometrical aspects of vision, such as subtended visual angle, observable angular velocity, and other derived parameters. The emergence of the collective state and its dynamics were achieved in several arenas. In addition, the flocking was robust enough to emerge under perturbations due to occlusions.

The approach does not assume direct measurement of inter-agent distances or velocities. Instead, we infer distances and velocities from observed angles and their rates of change. Such measurements are fully available in non-stereoscopic vision. This was one of the main motivations of the current study, and as noted, it is essential for modeling natural flocking organisms that lack stereoscopic vision, the prime examples being insects or fish, as well as similar artificial agents. The inability to directly assess range hinders even the acquisition of velocity direction of the neighbors, presenting a significant challenge to efficient flocking. To our knowledge, this is the first time a non-stereoscopic vision-based model, estimating neighbors' velocities only from angular information, has demonstrated flocking.

The collective state is reached by the gradual alignment of the agent's velocity vector with the velocity vectors of its neighbors. The algorithm computes the average of the velocities of an agent's neighbors that are located within a vision radius. It was observed that locusts [21] mostly react to their neighbors up to a limited visual distance. Visual radius thus limits the agent's knowledge about his neighbors. We expect a smaller vision radius to reduce collective coordination between the flock members, while a larger one to increase their coordination. A natural assertion would expect better convergence to occur when using higher vision radii. Experimental results have confirmed this dependence of convergence rates on R, as shown in Figs. 7a,7b. Rates of convergence were higher for larger R values. Also, the final asymptotic degrees of the order-parameter were higher for larger radii. A biologically relevant vision radius significance is evident in Fig. 7a. The asymptotic order parameter reaches saturation at approximately 2-3 BL. In other words, increasing R beyond 3 BL does not improve asymptotic order.

Further, exploring the limitations imposed by natural environments, we investigated the effect of mutual occlusions (of one neighbor by another) on visual-based flocking-related decision-making. To this end, visual cues coming from partially occluded objects constitute lost or distorted information. Thus, the problem of reaching a collective state becomes harder. Previous attempts to account for occlusions and their role in vision-based flocking have been very limited. To the best of our knowledge, the present work constitutes one of the first attempts to investigate this question thoroughly.

We tested and compared three different approaches to dealing with the problem of occlusions in different arenas. The first (COMPLID) completes the outline of a partially hidden neighbor. Such abilities were demonstrated in various species, including insects [60, 64]. However, this approach is the most complex of the three (cognitive-wise) as it requires object detection combined with extrapolating capabilities. The

second (OMID) ignores any partial information altogether. Thus, it calls for differentiating between fully vs. partially observed neighbor. This approach requires object detection; however, it is somewhat simpler than COMPLID since it only filters out erroneous visual stimuli rather than correcting them in the observer's mind. The last one (PARTID) treats each segment of a neighbor as if it represents a full-length body. Hence, the latter is the simplest of the three since it requires minimal cognitive processing from the individual — neither object detection nor object recognition, each set of distinguishable visual parameters is regarded as a neighbor. This results in PARTID being the most erroneous perception of the surrounding agent, e.g., segments change due to closer neighbors revealing or occluding them, which in turn is perceived as neighbors moving away or towards the focal agent.

The difference in the required computational power under the different approaches is very significant, as in nature, organisms demonstrating collective motion are very often limited in this respect (small brains, simple neuronal substrates). Hence, finding the least computationally demanding algorithm that is still capable of reaching flocking can potentially explain the actual mechanisms involved in the flocking of these relatively simple species.

In the torus arena, all three perception approaches of occlusions have successfully demonstrated the flocking transition from a chaotic initial state to an ordered collective state. Analyzing in more detail, we saw slight differences in convergence rates, where PARTID consistently performs slower than the others since it creates a particularly high misperception rate. This deficiency of PARTID is more pronounced in higher densities when a higher number of neighbors are present around the focal agent. This, in turn, leads to an increased prevalence of occlusions and consequently to the increased relative noisiness of PARTID.

We tested our approaches in geometrically constrained environments (arenas): corridors of varying width, and rings of varying thickness, where the boundaries lead to velocity reflections. The convergence process becomes significantly perturbed by rebounds of agents colliding with the boundaries.

Proceeding to examine the corridor arena, in the wide corridor all four behaved statistically similarly as seen in Fig. 9a. In the intermediate width, a slight separation between the approaches emerges 9b. Finally, at the narrow corridor a strongly pronounced difference is displayed between COMPLID (at faster convergence) and OMID and PARTID group with slower convergence. It implies that COMPLID is the more robust method, withstanding frequent disruptions by the boundary, and performing just as well as the baseline Principal model.

Moving on, to examine the ring arena, the results demonstrate a distinct weakness of the PARTID approach, especially pronounced in the wide ring. The other two perceptive approaches behaved similarly, though not identical. One can surmise, that ring geometry exudes high geometric influence on the collective motion, thereby leveling out the consequences of different methods, with the sole exception of PARTID. It could be suggested, that the frequent change of direction present in the ring arena, causes agents to re-appear and disappear from occluding neighbors, therefore strongly perturbing PARTID perception, while COMPLID and OMID cope well with these changes, due to the better treatment of occluded neighbors.

To examine the possible implications of our findings for the real world, one ought

to consider the biological plausibility of the three different approaches; in other words, the likeliness that the approaches describe the actual visual processing of animals. The topography of natural terrain has creeks, valleys, ridges, and other lateral constraints resulting in effectively constrained geometry. As is well established (see [68] and references within), marching locusts bands keep flock formation successfully despite such constraints. Considering our results from constrained arenas, we infer that PARTID is an oversimplification of the perceptive mechanisms in locust vision. Advanced capabilities are required for coping with occluded neighbors, such as the proposed OMID or COMPLID approaches. COMPLID, as referenced in Sec. 2.2, has empirical evidence of similar capabilities in animals, that show the ability to infer complete outlines from only partial visual clues

The present work studies different aspects of vision-based collective motion in swarms. The biological inspiration was to study purely visual and non-stereoscopic inputs, without direct distance measurements, and accounting for occlusions. We showed that the reductionist model is sufficient in principle to achieve collective behavior in a swarm, for different geometries. However, its reliability is dependent on the mechanisms used to address occlusions. This is an important area of continued investigation. The results here also are relevant for navigation and flocking in robotic swarms, where often robots lack stereo vision and ranging sensory capabilities (e.g., LIDAR devices) and are expensive in energy use and computation [69]. As directions for further study, the minimalist model could serve as the basis both for the development of more robust mechanisms or for solving more challenging environments. It may also inform the use of alternative interaction modalities observed in nature or technologically applicable.

5.1. Conclusions

This work establishes that a non-stereoscopic, purely vision-based model can achieve flocking without direct kinematic information — neither of distances nor of velocities. It is an essential step for modeling natural flocking organisms that lack stereoscopic vision, e.g., insects or fish, as well as similar artificial agents. Also, the stability of the model for some simplified approaches to neighbor occlusion was proven. Analyzing the differing emergence rates of collective motion supports the necessity of higher-level visual processes such as outline completion and conspecific recognition.

Acknoledgments

We gratefully acknowledge partial funding support by ISF Grant #2306/18. We would like to thank Michael Krongauz for his invaluable contribution to this research and the development of the model. Thanks to K. Ushi.

References

[1] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, V. Zdravkovic, Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study, Proceedings of the National

Academy of Sciences of the United States of America 105 (4) (2008) 1232–1237. doi:10.1073/PNAS.0711437105.

[2] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, V. Zdravkovic, Empirical investigation of starling flocks: a benchmark study in collective animal behaviour, Animal Behaviour 76 (1) (2008) 201–215. doi:10.1016/J.ANBEHAV.2008.02.004.

[3] S. B. Rosenthal, C. R. Twomey, A. T. Hartnett, H. S. Wu, I. D. Couzin, Revealing the hidden networks of interaction in mobile animal groups allows prediction of complex behavioral contagion, Proceedings of the National Academy of Sciences of the United States of America 112 (15) (2015) 4690–4695. doi:10.1073/PNAS.1420068112.

[4] N. O. Handegard, K. M. Boswell, C. C. Ioannou, S. P. Leblanc, D. B. Tjostheim, I. D. Couzin, The dynamics of coordinated group hunting and collective information transfer among schooling prey, Current Biology 22 (13) (2012) 1213–1217. doi:10.1016/J.CUB.2012.04.050.

[5] J. Buhl, D. J. Sumpter, I. D. Couzin, J. J. Hale, E. Despland, E. R. Miller, S. J. Simpson, From disorder to order in marching locusts, Science 312 (5778) (2006) 1402–1406. doi:10.1126/SCIENCE.1125142/SUPPL_FILE/BUHL.SOM.PDF.

[6] B. Uvarov, et al., Grasshoppers and locusts. A handbook of general acridology Vol. 2. Behaviour, ecology, biogeography, population dynamics., Centre for Overseas Pest Research, 1977.

[7] H. P. Zhang, A. Be'er, E. L. Florin, H. L. Swinney, Collective motion and density fluctuations in bacterial colonies, Proceedings of the National Academy of Sciences of the United States of America 107 (31) (2010) 13626–13630. doi:10.1073/PNAS.1001651107.

[8] A. Deutsch, G. Theraulaz, T. Vicsek, Collective motion in biological systems (2012).

[9] T. Vicsek, A. Zafeiris, Collective motion (8 2012). doi:10.1016/j.physrep.2012.03.004.

[10] C. W. Reynolds, Flocks, herds and schools: A distributed behavioral model, in: Proceedings of the 14th annual conference on Computer graphics and interactive techniques, 1987, pp. 25–34.

[11] M. J. Mataric, Interaction and intelligent behavior, Ph.D. thesis, Massachusetts Institute of Technology (1994).

[12] G. A. Kaminka, N. Fridman, Simulating urban pedestrian crowds of different cultures, ACM Transactions on Intelligent Systems and Technology 9 (3) (2018) 27:1–27:27. doi:10.1145/3102302.

[13] H. Hamann, Swarm Robotics: A Formal Approach, Springer, 2018.

[14] J. E. Herbert-Read, A. Perna, R. P. Mann, T. M. Schaerf, D. J. Sumpter, A. J. Ward, Inferring the rules of interaction of shoaling fish, Proceedings of the National Academy of Sciences of the United States of America 108 (46) (2011) 18726–18731. doi:10.1073/pnas.1109355108.

[15] G. Ariel, A. Ayali, Locust collective motion and its modeling, PLOS Computational Biology 11 (12) (2015) e1004522. doi:10.1371/JOURNAL.PCBI.1004522.

[16] T. Vicsek, A. Czirk, E. Ben-Jacob, I. Cohen, O. Shochet, Novel type of phase transition in a system of self-driven particles, Physical Review Letters 75 (6) (1995) 1226. doi:10.1103/PhysRevLett.75.1226.

[17] A. Kolpas, J. Moehlis, I. G. Kevrekidis, Coarse-grained analysis of stochasticity-induced switching between collective motion states, Proceedings of the National Academy of Sciences of the United States of America 104 (14) (2007) 5931–5935. doi:10.1073/PNAS.0608270104.

[18] C. Escudero, C. A. Yates, J. Buhl, I. D. Couzin, R. Erban, I. G. Kevrekidis, P. K. Maini, Ergodic directional switching in mobile insect groups, Physical Review E - Statistical, Nonlinear, and Soft Matter Physics 82 (1) (2010) 011926. doi:10.1103/PHYSREVE.82.011926/FIGURES/1/MEDIUM.

[19] I. Aoki, internal dynamics of fish schools in relation to inter-fish distance, Nippon Suisan Gakkaishi 50 (5) (1984) 751–758. doi:10.2331/suisan.50.751.

[20] N. W. Bode, J. J. Faria, D. W. Franks, J. Krause, A. J. Wood, How perceived threat increases synchronization in collectively moving animal groups, Proceedings of the Royal Society B: Biological Sciences 277 (1697) (2010) 3065–3070. doi:10.1098/RSPB.2010.0855.

[21] G. Ariel, Y. Ophir, S. Levi, E. Ben-Jacob, A. Ayali, Individual pause-and-go motion is instrumental to the formation and maintenance of swarms of marching locust nymphs, PLOS ONE 9 (7) (2014) e101636. doi:10.1371/JOURNAL.PONE.0101636.

[22] F. Cucker, S. Smale, Emergent behavior in flocks, IEEE Transactions on automatic control 52 (5) (2007) 852–862.

[23] V. Nityananda, J. C. Read, Stereopsis in animals: evolution, function and mechanisms, Journal of Experimental Biology 220 (14) (2017) 2502–2512.

[24] T. Hamada, Vision, action, and navigation in animals, Visual Navigation: From Biological Systems to Unmanned Ground Vehicles 2 (1997) 1.

[25] J. R. Serres, F. Ruffier, Optic flow-based collision-free strategies: From insects to robots, Arthropod structure & development 46 (5) (2017) 703–717.

[26] S. J. Judge, F. C. Rind, The locust dcmd, a movement-detecting neurone tightly tuned to collision trajectories, Journal of Experimental Biology 200 (16) (1997) 2209–2216. doi:10.1242/JEB.200.16.2209.

[27] B. B. Guest, J. R. Gray, Responses of a looming-sensitive neuron to compound and paired object approaches, Journal of neurophysiology 95 (3) (2006) 1428–1441.

[28] B. Collignon, A. Séguret, J. Halloy, A stochastic vision-based model inspired by zebrafish collective behaviour in heterogeneous environments, Royal Society Open Science 3 (1) (1 2015). doi:10.1098/RSOS.150473.

[29] R. Bastien, P. Romanczuk, A model of collective behavior based purely on vision, Science Advances 6 (6) (2 2020). doi:10.1126/SCIADV.AAY0792/SUPPL_FILE/AAY0792_SM.PDF.

[30] J. Qi, L. Bai, Y. Xiao, Y. Wei, W. Wu, The emergence of collective obstacle avoidance based on a visual perception mechanism, Information Sciences 582 (2022) 850–864. doi:10.1016/J.INS.2021.10.039.

[31] A. Strandburg-Peshkin, C. R. Twomey, N. W. Bode, A. B. Kao, Y. Katz, C. C. Ioannou, S. B. Rosenthal, C. J. Torney, H. S. Wu, S. A. Levin, I. D. Couzin, Visual sensory networks and effective information transfer in animal groups, Current Biology 23 (17) (2013) R709–R711. doi:10.1016/J.CUB.2013.07.059/ATTACHMENT/9C177EE8-47CA-4A66-BCD2-FBB3BDE6028B/MMC3.MP4.

[32] D. Pita, B. Collignon, J. Halloy, E. Fernández-Juricic, Collective behaviour in vertebrates: A sensory perspective, Royal Society Open Science 3 (11) (11 2016). doi:10.1098/RSOS.160377.

[33] H. Kunz, C. K. Hemelrijk, Simulations of the social organization of large schools of fish whose perception is obstructed, Applied Animal Behaviour Science 138 (3-4) (2012) 142–151. doi:10.1016/J.APPLANIM.2012.02.002.

[34] B. H. Lemasson, J. J. Anderson, R. A. Goodwin, Collective motion in animal groups from a neurobiological perspective: The adaptive benefits of dynamic sensory loads and selective attention, Journal of Theoretical Biology 261 (4) (2009) 501–510. doi:10.1016/J.JTBI.2009.08.013.

[35] B. H. Lemasson, J. J. Anderson, R. A. Goodwin, Motion-guided attention promotes adaptive communications during social navigation, Proceedings of the Royal Society B: Biological Sciences 280 (1754) (3 2013). doi:10.1098/RSPB.2012.2003.

[36] E. Soria, F. Schiano, D. Floreano, The influence of limited visual sensing on the reynolds flocking algorithm, Proceedings - 3rd IEEE International Conference on Robotic Computing, IRC 2019 (2019) 138–145doi:10.1109/IRC.2019.00028.

[37] N. Moshtagh, N. Michael, A. Jadbabaie, K. Daniilidis, Vision-based, distributed control laws for motion coordination of nonholonomic robots, IEEE Transactions on Robotics 25 (4) (2009) 851–860. doi:10.1109/TRO.2009.2022439.

[38] F. Berlinger, M. Gauci, R. Nagpal, Implicit coordination for 3d underwater collective behaviors in a fish-inspired robot swarm, Science Robotics 6 (50) (1 2021). doi:10.1126/SCIROBOTICS.ABD8668/SUPPL_FILE/ABD8668_SM.PDF.

[39] X. Wang, F. Wang, Z. Nie, Y. Ai, T. Hu, optiswarm: Optical swarm robots using implicit cooperation, IEEE Sensors Journal (2022).

[40] A. Ayali, The puzzle of locust density-dependent phase polyphenism, Current opinion in insect science 35 (2019) 41–47.

[41] D. A. Cullen, A. J. Cease, A. V. Latchininsky, A. Ayali, K. Berry, J. Buhl, R. De Keyser, B. Foquet, J. C. Hadrich, T. Matheson, et al., From molecules to management: mechanisms and consequences of locust phase polyphenism, in: Advances in Insect Physiology, Vol. 53, Elsevier, 2017, pp. 167–285.

[42] L. Zhang, M. Lecoq, A. Latchininsky, D. Hunter, Locust and grasshopper management, Annu. Rev. Entomol 64 (1) (2019) 15–34.

[43] S. Bazazi, J. Buhl, J. J. Hale, M. L. Anstey, G. A. Sword, S. J. Simpson, I. D. Couzin, Collective motion and cannibalism in locust migratory bands, Current biology 18 (10) (2008) 735–739.

[44] D. Knebel, A. Ayali, M. Guershon, G. Ariel, Intra-versus intergroup variance in collective behavior, Science advances 5 (1) (2019) eaav0695.

[45] D. Knebel, C. Sha-ked, N. Agmon, G. Ariel, A. Ayali, Collective motion as a distinct behavioral state of the individual, iScience 24 (4) (2021) 102299. doi:10.1016/J.ISCI.2021.102299.

[46] J. R. Gray, E. Blincow, R. M. Robertson, A pair of motion-sensitive neurons in the locust encode approaches of a looming object, Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology 196 (12) (2010) 927–938. doi:10.1007/S00359-010-0576-7/FIGURES/8.

[47] A. Ben-Nun, A. Ayali, Self body-size perception in an insect, Naturwissenschaften 100 (5) (2013) 479–484.

[48] M. A. Goodale, C. G. Ellard, L. Booth, The role of image size and retinal motion in the computation of absolute distance by the mongolian gerbil (meriones unguiculatus), Vision research 30 (3) (1990) 399–413. doi:10.1016/0042-6989(90)90082-V.

[49] R. D. Santer, F. C. Rind, R. Stafford, P. J. Simmons, Role of an identified looming-sensitive neuron in triggering a flying locust's escape, Journal of Neurophysiology 95 (6) (2006) 3391–3400. doi:10.1152/JN.00024.2006.

[50] S. E. De Vries, T. R. Clandinin, Loom-sensitive neurons link computation to action in the drosophila visual system, Current Biology 22 (5) (2012) 353–362. doi:10.1016/J.CUB.2012.01.007.

[51] L. Bennett, P. Symmons, A review of estimates of numbers in some types of desert locust (schistocerca gregaria (forsk.)) populations, Bulletin of Entomological Research 61 (4) (1972) 637–649.

[52] S. J. Luck, M. A. Ford, On the role of selective attention in visual perception, Proceedings of the National Academy of Sciences 95 (3) (1998) 825–830.

[53] R. L. Canosa, Real-world vision: Selective perception and task, ACM Transactions on Applied Perception (TAP) 6 (2) (2009) 1–34.

[54] J. R. Dunbier, S. D. Wiederman, P. A. Shoemaker, D. C. O'Carroll, Facilitation of dragonfly target-detecting neurons by slow moving features on continuous paths, Frontiers in Neural Circuits 0 (OCTOBER 2012) (2012) 1–29. doi:10.3389/FNCIR.2012.00079/BIBTEX.

[55] H. Wang, J. Peng, S. Yue, A directionally selective small target motion detecting visual neural network in cluttered backgrounds, IEEE Transactions on Cybernetics 50 (4) (2020) 1541–1555. doi:10.1109/TCYB.2018.2869384.

[56] I. Bleichman, P. Yadav, A. Amir, Visual processing and collective motion-related decision making in desert locusts, Proc. R. Soc. Lond. BSee also bioRxiv ; https://doi.org/10.1101/2022.09.19.508462 (2022).

[57] G. Kanizsa, P. Renzi, S. Conte, C. Compostela, L. Guerani, Amodal completion in mouse vision., Perception 22 (6) (1993) 713–721. doi:10.1068/p220713.

[58] M. Singh, Modal and amodal completion generate different shapes, Psychological Science 15 (7) (2004) 454–459. doi:10.1111/j.0956-7976.2004.00701.x.

[59] V. Bruce, P. R. Green, M. A. Georgeson, Visual perception: Physiology, psychology, & ecology, 4th Edition, Hove & London: Psychology Press, 2003.

[60] A. Nieder, Seeing more than meets the eye: processing of illusory contours in animals, Journal of Comparative Physiology A 2002 188:4 188 (4) (2002) 249–260. doi:10.1007/S00359-002-0306-X.

[61] I. R. Lin, C. C. Chiao, Visual equivalence and amodal completion in cuttlefish, Frontiers in Physiology 8 (FEB) (2017) 40. doi:10.3389/FPHYS.2017.00040/BIBTEX.

[62] E. Mascalzoni, L. Regolin, Animal visual perception, Wiley Interdisciplinary Reviews: Cognitive Science 2 (1) (2011) 106–116. doi:10.1002/WCS.97.

[63] M. A. Cox, M. C. Schmid, A. J. Peters, R. C. Saunders, D. A. Leopold, A. Maier, Receptive field focus of visual area v4 neurons determines responses to illusory surfaces, Proceedings of the National Academy of Sciences of the United States of America 110 (42) (2013) 17095–17100. doi:10.1073/PNAS.1310806110.

[64] G. A. Horridge, S. Zhang, D. O'Carroll, Insect perception of illusory contours, Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 337 (1279) (1992) 59–64.

[65] Ernst schmidt - coding.
URL www.ernst-schmidt.com

[66] G. Amichay, G. Ariel, A. Ayali, The effect of changing topography on the coordinated marching of locust nymphs, PeerJ 4 (2016) e2742.

[67] A. Czirók, A. L. Barabási, T. Vicsek, Collective motion of self-propelled particles: Kinetic phase transition in one dimension, Physical Review Letters 82 (1) (1999) 209. doi:10.1103/PhysRevLett.82.209.

[68] J. Dkhili, U. Berger, L. M. I. Hassani, S. Ghaout, R. Peters, C. Piou, Self-organized spatial structures of locust groups emerging from local interaction, Ecological Modelling 361 (2017) 26–40.

[69] M. Keidar, G. A. Kaminka, Efficient frontier detection for robot exploration, IJRR 33 (2) (2014) 215–236.