Data-driven Noise Model for Simulating Swarms of Flying Insects

Jiaping Ren[1], Xinjie Wang[2], Xiaogang Jin[3†], and Dinesh Manocha[4]

[1,2,3]State Key Lab of CAD&CG, Zhejiang University, Hangzhou 310058, China

([1]E-mail: renjiaping1@gmail.com, [2]E-mail: xinjiewang.cg@gmail.com, [3]E-mail: jin@cad.zju.edu.cn)

[4]Department of Computer Science, University of North Carolina at Chapel Hill, NC, USA

(E-mail: dm@cs.unc.edu)

http://gamma.cs.unc.edu/BSwarm (video and appendix)

Abstract: We present a model to generate noise-induced insect movements in a large swarm that are similar to those observed in real-world trajectories. Our approach is based on pre-recorded insect trajectories. After presenting a novel evaluation metric and a statistical validation approach that takes into account various characteristics insect motions, we evaluate well-known noise functions. Finally, we combine Curl noise function with a dynamics model to generate realistic swarm simulations and emergent behaviors of flying insects. We demonstrate their performance on simulating large swarms.

Keywords: data-driven, noise modeling, validation

1. INTRODUCTION

Collective biological behaviors are frequently observed in the real world, such as in the coordinated behavior of large groups of similar animals. Local interactions among the individuals in a group give rise to emergent behaviors or patterns. Such behaviors and interactions have previously been extensively studied in biology, control theory, swarm intelligence, and related areas.

In this paper, we address the problem of accurately simulating the collective behaviors and trajectories of swarms of flying insects. It is generally thought that these behaviors or patterns can be explained using simple interaction rules [1] and inherent noise; the latter concept refers to the random movements of the insects in a swarm [2-4], which can help the insects maintain swarm alignment. Moreover, certain continuum approaches (e.g. the Vicsek model) assume that each insect in a group follows the trajectory of neighboring individuals and that the deviations in their trajectories can be modeled as noise [5].

There are several sources for this noise. At a broad level, they can be classified into intrinsic and extrinsic noises. The intrinsic noise refers to the decision mechanism through which the insects update their positions [2]. On the other hand, the extrinsic noise refers to the effects of the environment [6]. One of the major challenges is to derive a computational model of the inherent noise in terms of simulating the motion of a large swarm of flying insects. Recent developments in computer vision and capturing technologies have made it possible to track the dynamic motions of insects [7, 8]. Such datasets can help us analyze specific swarm phenomena as well as calibrate and evaluate different simulation models [9]. Given insects' small size, it is rather difficult to acquire accurate ground-truth motion trajectories [8]. As a result, we need to develop appropriate probabilistic techniques to model noise – in addition to evaluation metrics to compare the results with captured insect data.

† Xiaogang Jin is the presenter of this paper.

Main Result: We present a data-driven noise model that can be used to generate noise-induced movements that are similar to those observed in real-world trajectory datasets. The main idea is to utilize pre-recorded insect trajectories to derive a general noise model and compute the appropriate parameters for different species. Our approach employs a statistical formulation that inherently accounts for noise in the dynamics model to compute the trajectory of each insect in a swarm. Furthermore, we use seven time-varying metrics to evaluate the collective behaviors of insects and compute the optimal parameters using a genetic algorithm. We observe that the Curl noise [10] can be combined with a dynamics model and used to simulate many emergent behaviors, including locust migrations, aggregations at different scales, competition for mates, phase transition in terms of density passing a critical point, positive phototaxis, and escape responses to predator-like objects. The combined dynamics and noise model can be used to simulate very large swarms with thousands of insects and to handle high swarm densities.

The rest of the paper is organized as follows. We briefly survey the prior research in Section 2. We present our data-driven noise and evaluation model in Section 3. We analyze different noise models in Section 4 and use them to generate different collective behaviors.

2. PRIOR WORK

In this section, we give a brief overview of aggregation behavior models, noise functions used in dynamical models and validation method for multi-agent models.

2.1. Aggregation Behavior Models

There is substantial scientific literature on aggregation behaviors of animal groups, and many force-based multi-agent models have been proposed [11, 12]. These models can be divided into two categories: continuum approach models and discrete approach models. Most continuum approaches are based on partial differential equation [13-15]. Discrete approach is more common than the continuum approach [16] and includes rule-based model and mathematical model. In rule-based models, individual agents apply particular rules to achieve global behaviors [17]. Mathematical models explore collective behavior in a more general manner [18]. In these types of models, individuals interact with one another based on perception forces [19], and these forces include short-range repulsion and long-range attraction [20]. In addition, some models also consider medium-range orientation [1].

2.2. Noise Functions

Noise is a constructive force at the collective level in an insect swarm [4]. In multi-agent models, noise mainly includes white noise and Gaussian white noise, and its variants. White noise obeys a uniform distribution, and self-propelled particle (SPP) [2] adopts this kind of noise. Gaussian white noise is the noise that obeys a Gaussian (or normal) distribution, this noise function is used in most of the Brownian dynamics models [21, 22], Esciderp et al. [23] also use this noise. Meanwhile, the manner in which the noise is introduced into the system will affect the simulation results [24]. Aldana et al. [25] consider intrinsic noise and extrinsic noise based on Vicsek's model [2], but both types of noise are white noise. Gönci et al. [26] use a scalar noise model that is chosen because it is uniformly distributed as a rotation tensor. In computer graphics, there are two kind of noise that can be applied to the simulation of collection motion: Perlin noise and Curl noise. Perlin noise is a type of gradient noise that consists of a collection of lattices of random gradients in which the values between lattices are obtained by interpolation [27]. Curl noise is incompressible velocity fields which is based on Perlin noise and its amplitude can be modulated in space as desired [10]. Chaté et al. [20] propose the notion of angular noise (a scalar) and vectorial noise (a vector), both of them are uniformly distributed.

2.3. Evaluation Methods

Dynamic multi-agent models can produce behaviors qualitatively similar to real biological systems. Lukeman et al. [9] validate their results by overlaying them on original images. However, it cannot be used to accurately evaluate the dynamics of a swarm. Other techniques have been proposed in evaluating the accuracy of human crowds, including parameter optimization approaches [28, 29] that use real-world crowd trajectories. Guy et al. [30] propose an entropy-based evaluation approach to quantify the similarity between real-world and simulated trajectories. However, these approaches are unable to model the inherent noise.

3. DATA-DRIVEN NOISE MODEL

It is well known that insects exhibit noise-induced movements and sudden changes in direction as a protection mechanism [2, 4]. Thus, it is important to develop a parametric noise model that can simulate different insect behaviors [24-26].

We present a data-driven approach that uses real-world insect trajectory datasets to model the noise. In particular, we first introduce an evaluation method to evaluate and compare the simulation results of a dynamics model that computes insect trajectories based on real-world datasets. Next, we use this evaluation metric to compare the accuracy of different noise models, based on the insect trajectory datasets.

3.1. Model Evaluation

The simplest technique for evaluating a model is to render the trajectories and observe the insect movements. However, basing an evaluation of whether a given dynamics model can capture all aspects of insects' emergent behaviors on visual rendering alone is not sufficient [16]. We present a novel quantitative approach to evaluate insect dynamics models by using real-world trajectory datasets. Our approach accounts for some key aspects of insect behaviors and trajectories based on seven time-varying metrics.

It is possible that two different swarms with noisy trajectories may exhibit similar swarm behaviors even when their trajectory positions are quite different. Our approach uses discrete probability density distribution functions (PDF) that are generated from the time-varying metrics and reflects the global characteristics of insect swarms. The influence of a small amount of data abnormality or noise can be ignored.

Our evaluation model is represented by the following equation, which contains seven energy terms:

$$E = 1 - \sum_{\phi \in \Phi} w_\phi E_\phi, \tag{1}$$

where $\Phi = \{v, a, \omega, \alpha, \mu, d, \eta\}$, which consists of seven time-varying metrics (see Section 3.2): v the velocity, a the acceleration, ω the angular velocity, α the angular acceleration, μ the Cartesian jerk, d the shortest distance, and η the velocity difference. These seven metrics are inspired from the biological literature (more details in Section 3.2). E_ϕ denotes the energy term about the metric ϕ, and w_ϕ denotes the weight of E_ϕ.

For a metric ϕ in Φ, E_ϕ is the energy term that represents the difference in discrete PDF between the real-world data and the simulation data. We formulate E_ϕ as

$$E_\phi = \left\| \mathcal{Q}_\phi^{real} - \mathcal{Q}_\phi^{sim} \right\|_1, \tag{2}$$

where $\mathcal{Q}_\phi^{real}$ denotes the discrete PDF of an insect swarm's metrics from real-world captured data and $\mathcal{Q}_\phi^{sim}$ represents the discrete PDF of an insect swarm's metrics from our swarm simulation model. We compute E_ϕ in four steps as follows:

Step 1: Sample the real data and the simulation data for the metric ϕ. For one set of real data or simulation data, compute the metric ϕ for all insects in all frames;

Step 2: Normalize the samples with the z-score method which refers to a mean shift followed by a standard deviation scaling. Because the real-world data and the simulator's output have different quantity scales, we must normalize the samples before comparing. We simply apply the z-score normalization method to the time-varying metric ϕ;

Step 3: Compute the discrete PDFs of the real-world data and the simulator's data with normalized samples from Step 2. For example, we can consider the real-world data: let S be the number of samples, and let $[u_1, u_2]$ be the interval of a given metric ϕ. We divide the interval $[u_1, u_2]$ into M equal sub-intervals. When we consider the ith subinterval $[u_1 + \frac{u_2-u_1}{M}(i-1), u_1 + \frac{u_2-u_1}{M}i]$ with S_i samples, the probability density in the ith interval is given as $\mathcal{Q}^{real}_{\phi,i} = \frac{S_i M}{S(u_2-u_1)}$. The probability density of the simulation data in the ith interval $\mathcal{Q}^{sim}_{\phi,i}$ can be calculated similarly;

Step 4: Compute the energy term E_ϕ: the difference of the discrete PDFs between the real data and the simulation data, and $E_\phi = \sum_{i=1}^{M} \left| \mathcal{Q}^{real}_{\phi,i} - \mathcal{Q}^{sim}_{\phi,i} \right|$.

We normalize the energy terms in Equation 2:

$$E_\phi = \frac{\left\| \mathcal{Q}^{real}_{\phi} - \mathcal{Q}^{sim}_{\phi} \right\|_1 - p_{1\phi}}{p_{2\phi}}, \tag{3}$$

where $p_{1\phi}$ and $p_{2\phi}$ are normalization parameters. The computation of $\left\| \mathcal{Q}^{real}_{\phi} - \mathcal{Q}^{sim}_{\phi} \right\|_1$ is the same as in Equation (2).

3.2. Time-varying metrics

We present seven time-varying metrics that are used to evaluate the trajectories and behaviors of insect swarms. These metrics are designed based on prior work and known characteristics of insect behaviors. A higher evaluation with respect to our metric indicates that the resulting simulation algorithm results in more realistic simulation results.

Velocity: Velocity is a basic metric used to evaluate the motion of an agent. We measure the magnitude of velocity v.

Acceleration: We can consider the acceleration as an effective net force on an insect [7]. We use the magnitude of acceleration a.

Angular velocity & acceleration: Angular rotations of an insect's body result in Coriolis forces, and the trajectory of an insect is affected by that force [31]. Therefore, we account for angular velocity and angular acceleration. The angular velocity is defined as the rate of change of angular displacement $\omega = \frac{\arccos \frac{\mathbf{v_1 v_2}}{|\mathbf{v_1}||\mathbf{v_2}|}}{\Delta t}$, where $\mathbf{v_1}$ and $\mathbf{v_2}$ represent the velocity of one insect in neighboring time points. The angular acceleration is defined as $\alpha = \frac{\Delta \omega}{\Delta t}$.

Cartesian jerk: Insect behavior tends to include some inherent noise [4], whereas humans and large animals typically move in a trajectory with gradual changes. The Cartesian jerk is used to represent the noise of insects' motion. Cartesian jerk is mathematically defined as the rate of change of acceleration [32] and reflects the smoothness of velocity $\mu = \left\| \frac{\Delta \mathbf{v}_1 - \Delta \mathbf{v}_2}{(\Delta t)^2} \right\|_2$, where μ is the magnitude of the second order differential of velocity, $\Delta \mathbf{v}_1$ and $\Delta \mathbf{v}_2$ are the velocity changes of one insect in neighboring time points.

Shortest distance: The density of an insect swarm reflects the group's degree of order [3] and the number of insects per unit volume. But the number of samples for density is limited, which affects discrete PDF computation. We note that the distance to the nearest neighbor for each insect is a reflection of the density of an insect swarm. Therefore, we choose the distance to nearest neighbor [33] as our metric, and term it the shortest distance d. We formulate the shortest distance as follows:

$$d = \min_{k \in \{1,2,\ldots,N\} \backslash \{m\}} \|\vec{p}_k - \vec{p}_m\|_2 ,$$

where m denotes the ID of current insects, k denotes the ID of other insects, N is the number of insects in the swarm, and $\vec{p}$ denotes the position of the insects.

Velocity difference: Unlike bird flocks and fish schools, a single insect in a swarm has little tendency to align with its neighbors [7]. Therefore, it is important to study the difference in velocity between neighboring insects to distinguish insect swarms from groups. If the shortest distance has a large magnitude, the influence of the difference in velocity to the corresponding metric should be relatively weak. As a result, we formulate the velocity difference as $\eta = \frac{|v_{nei} - v|}{d}$, where v_{nei} denotes the magnitude of velocity of the nearest neighbor.

3.3. Model Evaluation with Entropy Weight

In this section, we describe our evaluation algorithm. The overall evaluation has two components: optimizing the dynamics model parameters and optimizing the weights of seven energy terms.

We evaluate dynamics models for insect swarms with estimated optimal parameters (see Figure 1). The performance of a dynamics model for insect swarms is sensitive to the choice of underlying parameters. Therefore, we use a genetic algorithm to compute the optimal parameters by maximizing the evaluation function in Equation 1.

However, when we use the evaluation model to assess the different simulation techniques for insect swarms, it may require assigning different weights to each energy term. Instead, we compute the weights of all the energy terms automatically and then compute the final weighted score to evaluate different insect simulation models for fair comparisons. We use the entropy-based evaluation method described in [34] to compute the weights of the evaluation model in Equation 1 to provide reliable results.

Let m be the number of insect swarm simulation models evaluated, and n the number of different energy terms defined in Equation 1. The resulting energy terms matrix (before normalization) is $X = (x_{ij})_{m \times n}$:

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix}.$$

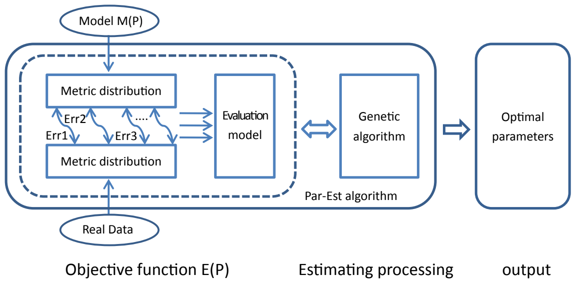


Fig. 1 Parameter-estimation algorithm (Par-Est algorithm). For the given real-world trajectory data of an insect swarm, we compute the discrete PDFs of the seven time-varying metrics using that data. Meanwhile, we use the parameterized dynamics model to simulate the insects and compute the discrete PDFs of the seven time-varying metrics with the simulation data. Next, we evaluate the function given in Equation (1) and use that as an objective function for the genetic algorithm to compute the optimal parameters.

We normalize matrix X as follows:

$$x_{ij} = \frac{x_{ij} - p_{1j}}{p_{2j}},$$

where $p_{1j} = \min_i x_{ij}, p_{2j} = \max_i x_{ij} - \min_i x_{ij}$ are the normalization parameters described in Equation 3. Let $R = (r_{ij})_{m \times n}$, $r_{ij} = 1 - x_{ij}$, and the entropy of an energy term is defined as

$$e_j = -\frac{1}{\ln m} \sum_{i=1}^{m} g_{ij} \ln g_{ij}, j = 1, 2, ..., n,$$

where $g_{ij} = \frac{r_{ij}}{\sum_{i=1}^{m} r_{ij}}$, and $g_{ij} \ln g_{ij} = 0$ when $g_{ij} = 0$. The weight of the ith energy term is calculated by

$$w_j = \frac{1 - e_j}{\sum_{j=1}^{n} (1 - e_j)}. \tag{4}$$

We use this evaluation scheme to compare the performance of prior multi-agent and insect swarm simulation models. The resulting evaluation algorithm that can compare the performance of different models insect swarms is summarized as follows:

Step 1: Initialize the weights in Equation 1 and normalization parameters in Equation 3; then, set the value ranges of the parameters in the dynamics models;

Step 2: Compute the energy terms with optimal parameters of each model to be evaluated using the Par-Est method shown in Figure 1;

Step 3: Compute the weights and normalization parameters with the energy terms matrix generated from Step 2;

Step 4: If the weights of our evaluation model are close to the weights computed in prior iterations, or the current number of iterations reaches the maximum number of iterations, go to Step 5; otherwise go to Step 2;

Step 5: Return the results.

4. RESULTS AND ANALYSIS

In this section, we highlight the performance of our evaluation method for noise modeling and multi-agent simulation algorithm comparison for insect swarms.

We have implemented our evaluation approach in MATLAB and insect swarm simulation in C++, both on a PC with Intel Xeon CPU E3-1230 and 8GB memory.

4.1. Real-World Datasets

We use four insect trajectory datasets to compute the appropriate noise model and estimate the parameters of insect swarm simulations. Both of these datasets were captured in an indoor setting with state-of-the-art motion capture systems. The dataset-1 from [7] was captured in a transparent $91cm$ cubical enclosure, and corresponds to time-resolved measurements of the positions, velocities, and accelerations of individual insects in laboratory swarms of the midge *Chironomus riparius*. The total number of midges vary from 12 to 111 per frame. The three other datasets, dataset-2, dataset-3, and dataset-4 were captured in a cube of $2m$ edge length with hundreds of *Drosophila* (fruit flies) [8]. We choose 500 frames from each of these four datasets.

4.2. Data-driven Noise Model

Our evaluation method can help dynamics models find the suitable noise to make the simulation results more likely to more closely resemble the real biological system, and the form of noise function is not limited. In this section, we use four noise functions as examples to choose the most suitable noise for a certain dynamics model; additionally, the type of dynamics model is not limited.

4.2.1. Noise Function

The description of the four noise functions for the stochastic term is as follows:

White noise: A scalar white noise W, which has a probability distribution with zero mean and finite variance. A 3D white noise $\mathbf{W}$ consists of three Ws that are statistically independent. This noise is used in the SPP model [2].

Gaussian white noise: An approximation of Gaussian white noise $\mathbf{G}$ is generated from two white noises, $\mathbf{W_1}$ and $\mathbf{W_2}$, and expressed as:

$$\mathbf{G} = \lambda \cdot \sqrt{-\mathbf{2} \cdot \log(\mathbf{W_1})} \cdot \sin(\mathbf{2}\pi\mathbf{W_2}),$$

where λ is a strength coefficient. This noise function is used in most of the Brownian dynamics models [21, 22].

Perlin noise: Perlin noise correlates to position $\mathbf{r_i}$. Assume that P is a scalar Perlin noise; thus, a 3D Perlin noise field $\mathbf{P}$ is generated by:

$$\mathbf{P}(\mathbf{r_i}) = \left(\mathbf{P_1}(\frac{\mathbf{r_i}}{\mathbf{scale}}), \mathbf{P_2}(\frac{\mathbf{r_i}}{\mathbf{scale}}), \mathbf{P_3}(\frac{\mathbf{r_i}}{\mathbf{scale}})\right) \cdot \mathbf{gain},$$

where *scale* and *gain* are two noise parameters: *scale* is used to control the smoothness of noise indirectly, and *gain* is used to adjust the magnitude of the noise.

Curl noise: Introduced by Bridson et al. [10], Curl noise is used to simulate continuous noise trajectories.

Inspired by Wang et al. [35], Curl noise $\mathbf{C_i}$ can be described as a force field related to the positions:

$$\mathbf{C}(\mathbf{r_i}) = \nabla \times \mathbf{P}(\mathbf{r_i}).$$

4.2.2. Noise Modeling

We compare four simulation models with four different noise functions (mentioned in Section 4.2.1) but with the same dynamics model (neglecting noise) based on the algorithm described in Section 3.3. We selected the parameters of the dynamics model (neglecting noise) as the common parameters for the four models, along with intensity of the stochastic force for the white and Gaussian noise, and scale and gain for Perlin noise and Curl noise. Figure 2 shows the comparison results of the models based on the four different noise types with four different datasets. Curl noise provides the most accurate results for our four datasets. All the detailed parameters used for these results are given in the Appendix. The weight in the evaluation model shown in Equation 1 with each dataset, the normalization parameter $p_{1\phi}$, $p_{2\phi}$ in Equation 3, and the results with each dataset are also given in the Appendix. Snapshots of the comparison results are shown in Figure 3. Please refer to the supplemental demo video for the comparison results on different insects of the animation results.

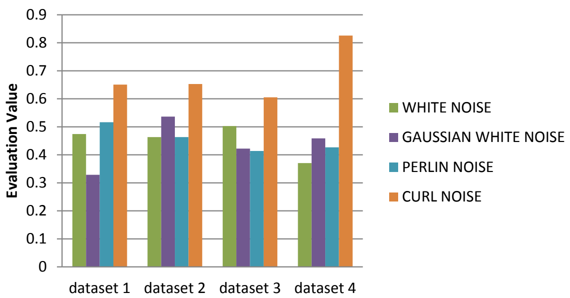
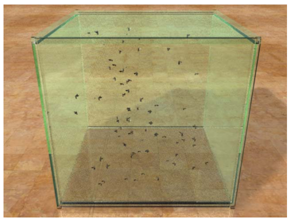


Fig. 2 The comparison results of the force-based model with the four different noise functions. The force-based model with Curl noise function is more accurate with respect to the real trajectory-datasets than the other noise functions.

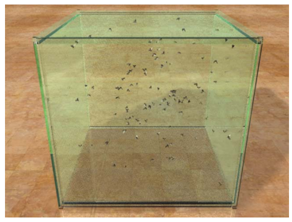
4.3. Comparing Different Multi-Agent Simulation Models

We compare 3 parameterized multi-agent simulation models based on the algorithm described in Section 3.3: dynamics with the noise model selected in Section 4.2.1, dynamics simulation only, and using a noise model only. For each model, we used our parameter estimation algorithm to compute the optimal parameters. Figure 4 shows the results comparing the three models with four different ground truth datasets. Please refer to the appendix for more details about the results.

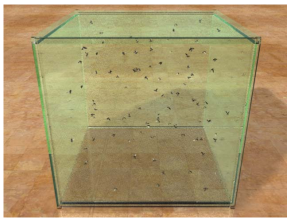
In addition, we have rendered a side-by-side visual comparison for these models by estimating the optimized parameters according to the ground truth dataset 4. Snapshots of the comparison results are shown in Figure 5.



(a) The ground truth data

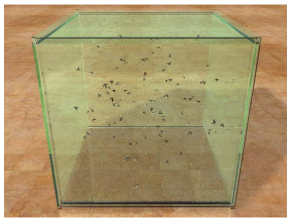


(b) White noise



(c) Gaussian white noise

(d) Perlin noise



(e) Curl noise

Fig. 3 Visual comparisons among different noise functions.

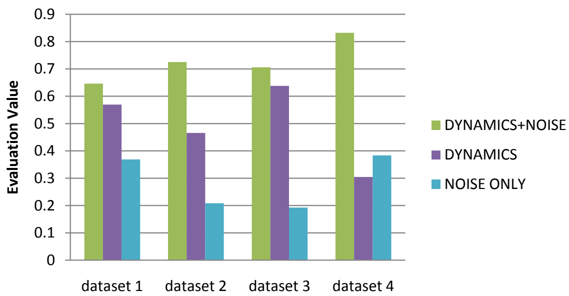


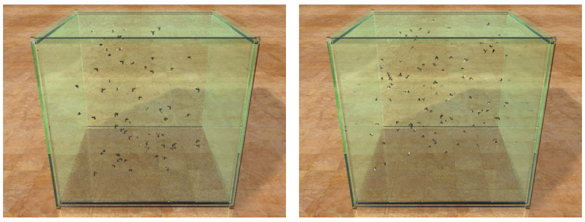
Fig. 4 We compared the simulation results of dynamics with noise, dynamics only and noise only with the real-world datasets, and we determined that our method can improve the accuracy of dynamics.

4.4. Collective Behaviors with Noise Model

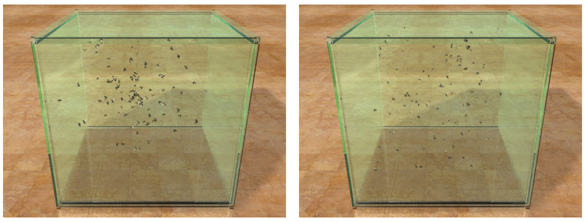
In this section, we highlight different collective behaviors that can be generated using our noise model combined with the dynamics model, which gives a plenty of behaviors driven by dynamics with noise.

Aggregation: We generated aggregation behaviors of midges with different individuals. Figure 6(a) contains 500 midges and Figure 6(b) contains $3,000$ in the same size of space. The swarm flies are with different densities: $9,469.7/m^3$ and $56,818.2/m^3$ (with their body length set to 0.01m) respectively while the swarm center hardly changes.

Locust Migration: Figure 7 shows locusts passing migration. These two locust swarms simulated has: (a) $2,000$ individuals with density $34.2/m^3$; (b) $200,000$ in-



(a) The ground truth data (b) Dynamics with noise



(c) Dynamics Only (d) Noise only

Fig. 5 Visual comparisons among different models.

(a) (b)

Fig. 6 Aggregation: (a) and (b) simulated midges flying in the same size of space with 500 and $3,000$ individuals, respectively.

(a) (b) (c)

Fig. 7 Locust Migration: (a) and (b) show migratory locusts pass through a grassland. The number of locusts in (a) is $2,000$ and (b) is $200,000$; (c) a snapshot of a real scene.

dividuals with density $342/m^3$, both of them with $0.04m$ body length.

Competition for mates: Figure 8 shows male flies competing for a female (in green color). The collective behavior is generated based on known behaviors of males and females [36].

Phase transition: Phase transition happens when the density of the swarm reaches the critical point. in Figure 9(a) and (b), a swarm increases its density until it reaches the critical point; (c) after the critical point, the swarm merely changes its density but changes its direction suddenly.

Positive phototaxis: In this scenario, moths gather around a street lamp (see Figure 10).

Startle/escape response: Figure 11 illustrates the response of a fly swarm to a predator-like object. When the predator-like object is close, each insect will run away.

Swarm of bats in a cave: Some behaviors of bats are

(a) (b)

Fig. 10 Positive phototaxis: (a) a snapshot of moths swarm simulated by noise model. The number of moths is 80; (b) a picture of a real scenario.

(a) (b)

Fig. 12 (a) A swarm of bats rapidly responds due to echolocation in a cave simulated by the noise model; (b) a real snapshot of bats.

similar to insect swarm. Bats can respond rapidly to the wall of the cave due to echolocation. We give a real snapshot of bat flies in a dark cave in Figure 12(b). In Figure 12(a), we simulate echolocation behavior by the noise model.

5. CONCLUSIONS AND FUTURE WORK

We have presented a novel PDF-based evaluation method to analyze the similarity between a simulation movement and a real dataset using entropy theory. We use our evaluation method to select the best suitable data-driven noise function that can be combined with a force-based simulation model. In addition, we compared different multi-simulation techniques using our evaluation method.

Limitations: We use genetic algorithms to estimate the parameters in our current implementation. Because genetic algorithms are probabilistic, they may not give optimal answers. Additionally, our implementation is not optimized and the running times can be considerably improved. Since our method is data dependent, over-fitting may occur if the trajectory dataset is too sparse or insufficient. Ultimately, we would like to evaluate its accuracy or performance on a large number of trajectory datasets.

Future work: We would like to collect more real trajectories of complex insect swarm behaviors, such as escape responses and migration. This can further improve the accuracy of our data-driven models and the overall simulation. We would like to use it for simulation other collective behaviors or use on different insect species.

Acknowledgments Xiaogang Jin was supported by the National Natural Science Foundation of China (Grant nos. 61272298, 61328204). Dinesh Manocha is sup-

(a) (b) (c)

Fig. 8 Competition for mates: (a) the male flies close to the female for chasing her; (b) more males join in; and (c) some time later, most of the male flies chase the female.

(a) (b) (c)

Fig. 9 Phase transition: (a) and (b) the swarm moves with the center position hardly changes before the density reaches the critical point; (c) when the density is $50/m^3$, the phase transition happens.

(a) (b) (c) (d)

Fig. 11 Startle/escape response: When a predator-like object (the sphere) attacks the swarm, individuals escape quickly to avoid it. The individuals aggregate slowly when the predator-like object disappears.

ported in part by ARO Contract W911NF-10-1-0506 and the National Thousand Talents Program of China.

REFERENCES

[1] Iain D. Couzin, Jens Krause, Richard James, Graeme D. Ruxton, and Nigel R. Franks. Collective memory and spatial sorting in animal groups. *Journal of Theoretical Biology*, 218(1):1 – 11, 2002.

[2] Tamás Vicsek, András Czirók, Eshel Ben-Jacob, Inon Cohen, and Ofer Shochet. Novel type of phase transition in a system of self-driven particles. *Physical Review Letters*, 75(6):1226–1229, 1995.

[3] J. Buhl, D. J. T. Sumpter, I. D. Couzin, J. J. Hale, E. Despland, E. R. Miller, and S. J. Simpson. From disorder to order in marching locusts. *Science*, 312(5778):1402–1406, 2006.

[4] Christian A. Yates, Radek Erban, Carlos Escudero, Iain D. Couzin, Jerome Buhl, Ioannis G. Kevrekidis, Philip K. Maini, and David J. T. Sumpter. Inherent noise can facilitate coherence in collective swarm motion. *Proceedings of the National Academy of Sciences*, 106(14):5464–5469, 2009.

[5] Tamas Vicsek. Universal patterns of collective motion from minimal models of flocking. In *Self-Adaptive and Self-Organizing Systems, 2008. SASO'08. Second IEEE International Conference on*, pages 3–11. IEEE, 2008.

[6] Guillaume Grégoire and Hugues Chaté. Onset of collective and cohesive motion. *Physical Review Letters*, 92(2):025702, 2004.

[7] Douglas H. Kelley and Nicholas T. Ouellette. Emergent dynamics of laboratory insect swarms. *Scientific Reports*, 3(1073), 2013.

[8] Hai Shan Wu, Qi Zhao, Danping Zou, and Yan Qiu Chen. Automated 3d trajectory measuring of large numbers of moving particles. *Optics Express*, 19(8):7646–7663, 2011.

[9] Ryan Lukeman, Yue-Xian Li, and Leah Edelstein-Keshet. Inferring individual rules from collective behavior. *Proceedings of the National Academy of Sciences*, 107(28):12576–12580, 2010.

[10] Robert Bridson, Jim Houriham, and Marcus Nordenstam. Curl-noise for procedural fluid flow. *ACM Transactions on Graphics*, 26(3):1–4, 2007.

[11] Iain D Couzin and Jens Krause. Self-organization and collective behavior in vertebrates. *Advances in the Study of Behavior*, 32(1), 2003.

[12] Xinjie Wang, Jiaping Ren, Xiaogang Jin, and Dinesh Manocha. Bswarm: Biologically-plausible dynamics model of insect swarms. In *Proceedings of the 14th ACM SIGGRAPH / Eurographics Symposium on Computer Animation*, SCA '15, pages 111–118, New York, NY, USA, 2015. ACM.

[13] John Toner and Yuhai Tu. Flocks, herds, and schools: A quantitative theory of flocking. *Physical Review E*, 58(4):4828–4858, 1998.

[14] Chad M Topaz, Andrea L Bertozzi, and Mark A Lewis. A nonlocal continuum model for biological aggregation. *Bulletin of Mathematical Biology*, 68(7):1601–1623, 2006.

[15] José A Carrillo, Marco DiFrancesco, Alessio Figalli, Thomas Laurent, Dejan Slepčev, et al. Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations. *Duke Mathematical Journal*, 156(2):229–271, 2011.

[16] James G Puckett, Douglas H Kelley, and Nicholas T Ouellette. Searching for effective forces in laboratory insect swarms. *Scientific Reports*, 4(4766), 2014.

[17] Craig W. Reynolds. Flocks, herds and schools: A distributed behavioral model. *SIGGRAPH Comput. Graph.*, 21(4):25–34, 1987.

[18] Jen Schellinck and Tony White. A review of attraction and repulsion models of aggregation: Methods, findings and a discussion of model validation. *Ecological Modelling*, 222(11):1897–1911, 2011.

[19] Alessandro Attanasi, Andrea Cavagna, Lorenzo Del Castello, Irene Giardina, Stefania Melillo, Leonardo Parisi, Oliver Pohl, Bruno Rossaro, Edward Shen, Edmondo Silvestri, et al. Collective behaviour without collective order in wild swarms of midges. *PLoS Computational Biology*, 10(7):e1003697, 2014.

[20] Hugues Chaté, Francesco Ginelli, Guillaume Grégoire, and Franck Raynaud. Collective motion of self-propelled particles interacting without cohesion. *Physical Review E*, 77(4):046113, 2008.

[21] Jessica Strefler, Udo Erdmann, and Lutz Schimansky-Geier. Swarming in three dimensions. *Physical Review E*, 78(3):031927, 2008.

[22] Pawel Romanczuk, Iain D. Couzin, and Lutz Schimansky-Geier. Collective motion due to individual escape and pursuit response. *Physical Review Letters*, 102(1):010602, 2009.

[23] Carlos Escudero, Christian A Yates, Jerome Buhl, Iain D Couzin, Radek Erban, Ioannis G Kevrekidis, and Philip K Maini. Ergodic directional switching in mobile insect groups. *Physical Review E*, 82(1):011926, 2010.

[24] Maximino Aldana, Victor Dossetti, Christian Huepe, VM Kenkre, and Hernán Larralde. Phase transitions in systems of self-propelled agents and related network models. *Physical Review Letters*, 98(9):095702, 2007.

[25] M Aldana, H Larralde, and B Vázquez. On the emergence of collective order in swarming systems: a recent debate. *International Journal of Modern Physics B*, 23(18):3661–3685, 2009.

[26] Balázs Gönci, Máté Nagy, and Tamás Vicsek. Phase transition in the scalar noise model of collective motion in three dimensions. *The European Physical Journal Special Topics*, 157(1):53–59, 2008.

[27] Ken Perlin. Improving noise. *ACM Transactions on Graphics*, 21(3):681–682, 2002.

[28] David Wolinski, S J Guy, A-H Olivier, Ming Lin, Dinesh Manocha, and Julien Pettré. Parameter estimation and comparative evaluation of crowd simulations. In *Computer Graphics Forum*, volume 33, pages 303–312. Wiley Online Library, 2014.

[29] Glen Berseth, Mubbasir Kapadia, Brandon Haworth, and Petros Faloutsos. SteerFit: Automated parameter fitting for steering algorithms. In *ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '14, pages 113–122, New York, NY, USA, 2014. ACM, ACM.

[30] Stephen J. Guy, Jur van den Berg, Wenxi Liu, Rynson Lau, Ming C. Lin, and Dinesh Manocha. A statistical similarity measure for aggregate crowd dynamics. *ACM Transactions on Graphics*, 31(6):190:1–190:11, 2012.

[31] Wai Pang Chan, Frederick Prete, and Michael H Dickinson. Visual input to the efferent control system of a fly's "gyroscope". *Science*, 280(5361):289–292, 1998.

[32] Tamar Flash and Neville Hogan. The coordination of arm movements: an experimentally confirmed mathematical model. *The Journal of Neuroscience*, 5(7):1688–1703, 1985.

[33] Mubbasir Kapadia, Shawn Singh, Brian Allen, Glenn Reinman, and Petros Faloutsos. Steerbug: an interactive framework for specifying and detecting steering behaviors. In *Proceedings of the 2009 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pages 209–216. ACM, 2009.

[34] Zhi-Hong Zou, Yi Yun, and Jing-Nan Sun. Entropy method for determination of weight of evaluating indicators in fuzzy synthetic evaluation for water quality assessment. *Journal of Environmental Sciences*, 18(5):1020–1023, 2006.

[35] Xinjie Wang, Xiaogang Jin, Zhigang Deng, and Linling Zhou. Inherent noise-aware insect swarm simulation. *Computer Graphics Forum*, 33(6):51–62, 2014.

[36] Ian A. N. Dublon and David J. T. Sumpter. Flying insect swarms. *Current Biology*, 24(18):R828–R830, 2014.