

## Research Paper

# Comparative analysis of tracking and behavioral patterns between wild-type and genetically modified fruit flies using computer vision and statistical methods

Fei Ying Lu <sup>a</sup>, Xiang Liu <sup>a,\*</sup>, Hai Feng Su <sup>b,\*</sup>, Shuo Hong Wang <sup>c</sup><sup>a</sup> Shanghai University of Engineering Science China<sup>b</sup> Shanghai University China<sup>c</sup> Harvard Medical School, Cambridge USA

## ARTICLE INFO

## Keywords:

Multi-object tracking  
Behavioral analysis  
Filter tracking  
3D reconstruction

## ABSTRACT

Collective animal behavior occurs in groups and swarms at almost every biological scale, from single-celled organisms to the largest animals on Earth. The intriguing mysteries behind these group behaviors have attracted many scholars, and while it is known that models can reproduce qualitative features of such complex behaviors, this requires data from real animals to demonstrate, and obtaining data on the exact features of these groups is tricky. In this paper, we propose the Hidden Markov Unscented Tracker (HMUT), which combines the state prediction capability of HMM and the high-precision nonlinear processing capability of UKF. This prediction-driven tracking mechanism enables HMUT to quickly adjust tracking strategies when facing sudden changes in target motion direction or rapid changes in speed, reducing the risk of tracking loss. Videos of fruit fly swarm movement in an enclosed environment are captured using stereo cameras. For the captured fruit fly images, the thresholded AKAZE algorithm is first used to detect the positions of individual fruit flies in the images, and the motion of the fruit flies is modeled using a multidimensional hidden Markov model (HMM). Tracking is then performed using the Unscented Kalman Filter algorithm to obtain the flight trajectories of the fruit flies in two camera views. Finally, 3D reconstruction of the trajectories in both views is achieved through polar coordinate constraints, resulting in 3D motion data of the fruit flies. Additionally, the efficiency and accuracy of the proposed algorithm are evaluated by simulating fruit fly swarm movement using the Boids algorithm. Finally, based on the tracked fruit fly flight data, behavioral characteristics of the fruit flies are analyzed from two perspectives. The first is a statistical analysis of the differences between the two behaviors. The second dimension involves clustering trajectory similarity using the DTW method based on fruit fly flight trajectories, further analyzing the similarity within clusters and differences between clusters.

## 1. Introduction

*Drosophila* is a common laboratory model organism that has been widely used in research in biology, genetics, neuroscience, and other fields as it has a short lifespan, strong reproductive ability, and a genome that has been completely sequenced.

Faraz et al. (2022) proposed a nanoparticle segmentation method based on U-Net, utilizing deep learning to identify nanoparticles across all frames of video sequences. They then iteratively track and reconstruct their motion trajectories. Through this method, quantitative data on the motion are inferred. Leveraging traditional convolutional neural networks such as R-CNN, YOLOv3, and YOLOv3 Tiny, Zhang et al.

(2020) have developed a system for real-time detection and tracking of golf balls in video sequences. For the tracking phase, they employ a discrete Kalman filter to predict the golf ball's location based on observational data, enhancing the precision and reliability of tracking the golf ball's trajectory. Aktaş and Ateş (2021) address the challenge of tracking targets in aerial images, where the targets appear relatively small and the visual information in the image sequences is insufficient for traditional object detection models to yield ideal results. Against this backdrop, they propose leveraging both spatial and temporal information from the image sequences. They refine the Region Proposal Network (RPN) stage by adjusting the anchors and optimizing the Intersection over Union (IOU) for small targets, enhancing detection and

\* Corresponding authors.

E-mail addresses: [xliu@sues.edu.cn](mailto:xliu@sues.edu.cn) (X. Liu), [hfsu@shu.edu.cn](mailto:hfsu@shu.edu.cn) (H.F. Su).

tracking accuracy under these constrained conditions. Zhu et al. (2023) have developed a Multilayer Knowledge Distillation Network (MKDNet), designed to perform knowledge extraction across three levels within a unified framework. This approach significantly improves the feature representation, recognition, and localization of small targets in tracking scenarios, refining the process of extracting crucial information for efficient tracking of small objects.

Ahmadi and Salari (2016) proposed a low signal-to-noise ratio infrared image sequence weak small target detection and tracking algorithm based on frequency domain information and spatial domain information. The Dual-tree Complex-Wavelet Transform (DT-CWT) is applied, and a Constant False Alarm Rate (CFAR) detector is used in the frequency domain to identify the location of the target in each frame. Finally, based on the spatial domain information within the frame, a Support Vector Machine (SVM) is used for classification, combining frequency domain and spatial domain information to improve tracking accuracy. Han et al. (2019) proposed the Spatial-Temporal Context-Aware tracker algorithm (STCAT). The authors pointed out that most Discriminative Correlation Filter (DCF) models only consider the spatial features of the target, neglecting the use of inter-frame and historical information, which affects tracking performance during occlusions and deformations. To address this issue, they incorporated temporal constraints on top of the traditional DCF, using spatial-temporal context information to make the tracker aware, achieving the goal of adapting to target appearance changes. Since changes between adjacent frames are minimal, this approach can maintain relatively stable tracking. Liu et al. (2019) proposed that small targets in the tracking process are prone to blurring and appearance changes, and based on this, they proposed an algorithm suitable for small target tracking.

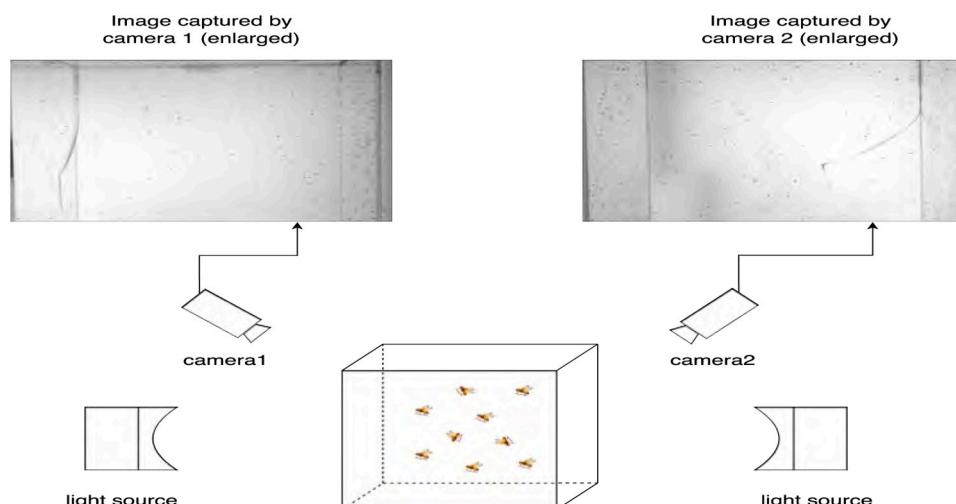
Despite significant progress in small target tracking technology in recent years, existing methods still face numerous challenges in dealing with dense group tracking, adapting to target state modeling, and complex behavior analysis. The tracking accuracy of high-density small target groups is severely affected by individual occlusions and similar movement patterns, with missed detections and false alarms being particularly prominent. Secondly, the ability to accurately predict the target's state in dynamic environments is crucial for maintaining the continuity and accuracy of tracking (Yick et al., 2005; Rossi et al., 2008). The performance of tracking algorithms largely depends on their predictive capabilities, that is, the accuracy of predicting the future position and state of the target in a changing environment. A lack of precise predictions can lead to tracking interruptions or mis-tracking, negatively impacting the overall effectiveness of tracking.

Typical examples of group movements include flocks of birds, schools of fish, and swarms of insects. These groups demonstrate the

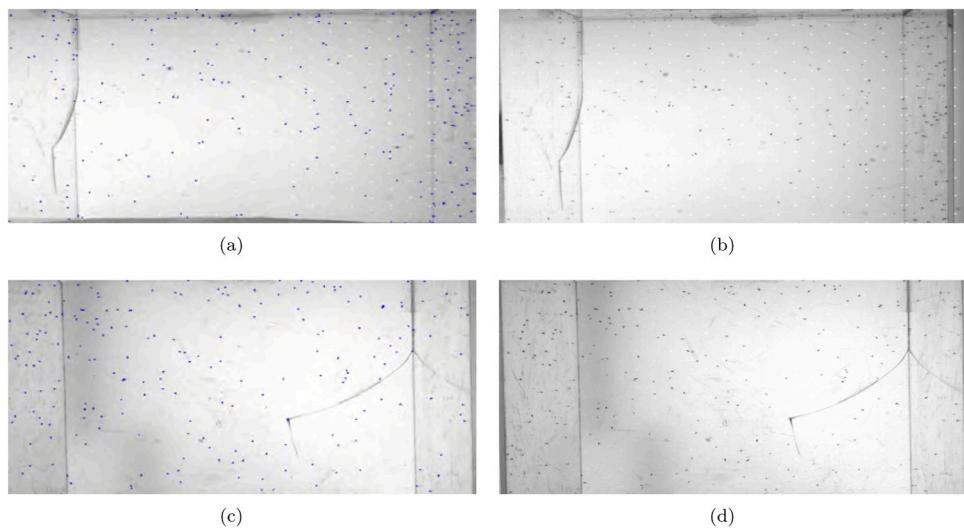
complexity of coordinated actions, and in-depth research into them can help reveal the secrets of group behavior. Buhl et al. (2006) tracked the movement behavior of locusts by defining the angle and quantifying the motion of locusts, identifying the critical density at which locust swarms begin to march in coordination. This work furthered our understanding of the transmission of directional information during their operation, advancing research into collective behavior. Makris et al. (2009) utilized Ocean Acoustic Waveguide Remote Sensing (OAWRS) to detect schools of fish in the ocean, observing the formation process of fish schools during spawning periods, and quantifying the corresponding migration speed and density. They discovered leadership phenomena within fish schools and a rapid transition from disordered to highly synchronized movement when population density reached a critical value. Numerical models suggest that the collective behavior of animals may be caused by simple local interaction rules among individuals. Based on this, researchers conducted extensive related studies. Ballerini et al. (2008) performed three-dimensional reconstructions of bird flocks to obtain their positions in space, finding that birds' interactive behaviors occur among a fixed number of close neighbors, proposing that interactions among individuals depend on topological rather than metric distance, as assumed by most models and theoretical hypotheses. They also confirmed that compared to standard metric methods, topological interactions significantly improve the cohesion of aggregation.

Analyzing the behavior of animals and capturing the biological characteristics therein cannot be separated from the quantitative analysis of their movement data, which requires accurate and quantitative comparison and analysis with the actual data (Branson et al., 2009; Grover et al., 2008; Wu et al., 2011; Jezovit et al., 2021; Gernat et al., 2018). In order to obtain the three-dimensional movement trajectories of Drosophila, two sub-processes are required: the first one is two-dimensional tracking on different views, i.e., temporal tracking within each view in order to recognize the correspondence of the target over time, and the second one is stereo matching on different views in order to reconstruct the target's three-dimensional coordinates. There are some challenges for tracking targets such as fruit flies, which are similar in appearance and small in size, so it is not feasible to use local texture information as features for tracking. In addition, when there are a large number of them, occlusion may occur (Straw et al., 2011; Kohlhoff et al., 2011; Wu et al., 2009).

In this paper, we use a binocular camera to realize the tracking of the flight of Drosophila, and obtain its long-distance flight trajectory. At the same time, we reconstructed the trajectories in 3D based on the trajectories in the two viewpoints, and analyzed the motion characteristics of Drosophila. Our contributions mainly include the following three points:



**Fig. 1.** Experiment settings.



**Fig. 2.** Detection results of images taken by both cameras.

- We proposed the Hidden Markov Unscented Tracker (HMUT), an effective method for integrating multi-source data, enhancing the algorithm's comprehensive understanding of the target state, and increasing its adaptability to changes in the environment.
- Perform multi-dimensional dynamic clustering on fruit fly trajectories, divide motion trajectories based on time windows, and further calculate the similarity of trajectory segments.
- We quantified the effect of mutant alleles in the Drosophila ellipsoid-body open (ebo) gene on its behavior and compared it to normal Drosophila.

Through the above improvements and operations, we have realized the long-distance tracking of Drosophila populations, and analyzed and compared the two types of Drosophila statistically by selecting different features, and quantified the differences and similarities between the two, the details of which will be discussed in the following section.

## 2. Material and methods

The Ellipsoid body (EB) of Drosophila is part of its central nervous system and has an important role in its behavioral patterns (Renn et al., 1999). The EB receives and sends outputs from many areas of the brain as direct inputs. This high level of connectivity makes the EB a center for the integration of multiple information streams, including visual, motor, and mechanosensory sensations, allowing it to functionally adapt complex behaviors (Yan et al., 2023). By observing how the ellipsoid-body open allele mutation changes behavior patterns in fruit flies, it is possible to gain a deeper and more sophisticated understanding of how the brain controls complex behaviors (Solanki et al., 2015; Martín-Peña et al., 2014). Flies carrying mutations in the gene ebo, *ebo*<sup>678</sup> (Heisenberg et al., 1985; Hanesch, 1987; Strauss and Heisenberg, 1993; Ilíus et al., 2007), as gift from Professor M Heisenberg, maintain in our lab. Drosophila in incubated vials were poured from the top of a square acrylic box with dimensions of 40cm\*40cm\*40cm and backlighting. Two geometrically calibrated and time-synchronized high-speed cameras were placed on the side of the box to capture the flies after they became familiar with their environment, with the cameras placed about 900mm from the arena. The camera parameters were: IO Industries Canada, Flare 4 M 180-CL, resolution: 2048w\*2040h pixels, frame rate: 100fps. We also tapped the walls of the box before shooting to maximize the number of fruit flies in flight. Fig. 1 show the setup of the experimental environment. The details of tracking and reconstruction will be presented in the next section.

### 2.1. AKAZE

The AKAZE algorithm uses the M-LDB algorithm to describe the feature points (Alcantarilla and Solutions, 2011), and utilizes the gradient and intensity information extracted from the nonlinear space to improve the matching speed of the feature points while ensuring the matching accuracy. Fig. 2 shows the detection results of fruit fly images using AKAZE.

### 2.2. Hidden Markov Model

HMM (Hidden Markov model) has efficient learning and inference ability (Rabiner and Juang, 1986), which represents the motion pattern of the target as a Markov process, and deduces the inference ability of HMM to the prediction ability under various scenarios by constructing two types of variables (hidden variables and observed variables) (Neal, 1993). In HMM, the invisibility of the state introduces model uncertainty, and HMM outputs observation and prediction errors during the inference process, which can accumulate rapidly during the continuous inference process. Kalman filtering is one of the most widely used estimation methods (Woods and Radewan, 1977), but it is challenging to apply it in nonlinear systems. Therefore, UKF is introduced here to combine with HMM. For non-Gaussian distributions, UKF can approximate the nonlinearity to the second order without explicitly computing the Jacobi matrix (Julier and Uhlmann, 2004). In summary, the HMM inference is used to obtain the most probable state of the next time frame T, and the confidence of the state is passed to the UKF to estimate a more accurate position of the target based on the nonlinear system dynamics, and then to determine the trajectory of the target.

With the AKAZE algorithm, we obtain the location coordinates of individual Drosophila in the image and the corresponding feature descriptors, and the next step is to model the movement of Drosophila using the HMM. The HMM consists of two state sets and three probability matrices, namely, the hidden state  $S$ , the observed state  $O$ , the initial state probability matrix  $\pi$ , the state transfer probability matrix  $A$ , and the observed probability matrix  $B$ .

The Viterbi algorithm (Forney, 1973) is a dynamic programming algorithm for finding the most probable state sequence in HMM. The basic idea is to calculate the maximum probability of each possible state at each moment using the known observation sequence and the model parameters  $M = (A, B, \pi)$ , by stepwise recursion to obtain the most probable state sequence  $Q$ . Define the variable  $\delta_i(i)$  to denote the probability maximum among all sequences with hidden state  $i$  at moment  $t$ . The definition formula is:

$$\delta_i = \max_{\{k_1, k_2, \dots, k_{t-1}\}} P(k_1, k_2, \dots, k_{t-1}, q_t = s_i, o_1, o_2, \dots, o_t | \mathbf{k}); \quad i = 1, 2, \dots, N; t = 1, 2, \dots, T \quad (1)$$

Define  $\psi_t(i)$  as the previous state  $j$  of the probabilistic maximum path recorded for each hidden state  $i$ . The optimal path (the sequence of hidden states corresponding to the observed states) solving process is divided into initialization, recursion, termination and optimal path backtracking.

### 1. Initialization:

$$\delta_k = \pi_k \cdot b_k(o_1), \psi_1(k) = 0; k = 1, 2, \dots, N \quad (2)$$

### 2. Recursion:

$$\delta_j = \left[ \max_{\{1 \leq k \leq N\}} \delta_{t-1}(k)a_{kj} \right] b_j(o_t), \quad j = 1, 2, \dots, N; \quad t = 2, \dots, T \quad (3)$$

$$\psi_t(j) = \arg \max_{\{1 \leq k \leq N\}} [\delta_{t-1}(k)a_{kj}]; \quad j = 1, 2, \dots, N; \quad t = 2, \dots, T \quad (4)$$

### 3. Final output:

$$P_* = \max_{\{1 \leq k \leq N\}} [\delta_T(k)]; \quad q_{T*} = \arg \max_{\{1 \leq k \leq N\}} [\delta_T(k)] \quad (5)$$

### 4. Optimal state sequence determination:

$$q_{*t} = \psi_{t+1}(q_{t+1*}); \quad t = (T-1), (T-2), \dots, 1 \quad (6)$$

Finally, the optimal hidden state sequence can be obtained.

### 2.3. Unscented Kalman filter

For a general nonlinear discrete-time dynamic system, where the process and measurement model can be expressed as follows.

$$\begin{cases} X_i = f(X_{i-1}) + \Gamma_{i-1} W_{i-1} \\ Z_i = h(X_i) + V_i \end{cases} \quad (7)$$

Step1: Initialization

$$\begin{cases} \hat{X}_0 = E\{X_0\} \\ P_0 = E\{(X_0 - \hat{X}_0)(X_0 - \hat{X}_0)^T\} \end{cases} \quad (8)$$

where  $\hat{X}_0$  is initial state and  $P_0$  is the initial estimation error covariance.

Step2: Sigma points calculation

$$\chi_{k-1}^{(0)} = \hat{X}_{k-1} \quad (9)$$

$$\chi_{k-1}^{(i)} = \hat{X}_{k-1} + \sqrt{(n+\kappa)p_{k-1}}, \quad i=1, 2, \dots, n \quad (10)$$

$$\chi_{k-1}^{(i)} = \hat{X}_{k-1} - \sqrt{(n+\kappa)p_{k-1}}, \quad i=n+1, n+2, \dots, 2n \quad (11)$$

$n$  is the state dimension,  $\kappa = \frac{n^2}{n+1} - n$  is the composite scaling factor.  $n$  and  $\kappa$  are tuning parameters. The parameter  $n$  is set to  $0 \leq n \leq 1$  and a good default setting on  $\kappa$  is  $\kappa = 0$ .

Step3: State prediction

$$\chi_{k/k-1}^{(i)} = f(\chi_{k-1}^{(i)}, k-1), \quad i = 0, 1, 2, \dots, 2n \quad (12)$$

$$\hat{X}_{k/k-1} = \sum_{i=0}^{2n} \omega_i^{(m)} \chi_{k/k-1}^{(i)} \quad (13)$$

$$P_{XX} = \sum_{i=0}^{2n} \omega_i^{(c)} (\chi_{k/k-1}^{(i)} - \hat{X}_{k/k-1}) (\chi_{k/k-1}^{(i)} - \hat{X}_{k/k-1})^T \quad (14)$$

$$P_{k/k-1} = P_{XX} + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^T \quad (15)$$

where  $\omega_i^m$  and  $\omega_i^c$  are weights, which are defined as:

$$\omega_0^{(m)} = \frac{\kappa}{n + \kappa} \quad (16)$$

$$\omega_0^{(c)} = \frac{\kappa}{n + \kappa} + (1 - \alpha^2 + \beta) \quad (17)$$

$$\omega_i^{(m)} = \omega_i^{(c)} = \frac{1}{2(n + \kappa)}, \quad i = 1, 2, \dots, 2n \quad (18)$$

We introduce the parameter  $\beta \geq 0$  to incorporate higher order information about the distribution. Here we set  $\beta = 2$ .

Step4: Measurement prediction

$$\zeta_{k/k-1}^{(i)} = h(\chi_{k/k-1}^{(i)}, k), \quad i = 0, 1, 2, \dots, 2n \quad (19)$$

$$\hat{Z}_{k/k-1} = \sum_{i=0}^{2n} \omega_i^{(m)} \zeta_{k/k-1}^{(i)} \quad (20)$$

Step5: Kalman Gain calculation

$$\begin{aligned} P_{XZ} &= \sum_{i=0}^{2n} \omega_i^{(c)} (\chi_{k/k-1}^{(i)} - \hat{X}_{k/k-1}) \\ &\quad (\zeta_{k/k-1}^{(i)} - \hat{Z}_{k/k-1})^T \end{aligned} \quad (21)$$

$$\begin{aligned} P_{ZZ} &= \sum_{i=0}^{2n} \omega_i^{(c)} (\zeta_{k/k-1}^{(i)} - \hat{Z}_{k/k-1}) \\ &\quad (\zeta_{k/k-1}^{(i)} - \hat{Z}_{k/k-1})^T + R_k \end{aligned} \quad (22)$$

$$K_k = P_{XZ} P_{ZZ}^{-1} \quad (23)$$

Step6: Filtering update

$$\hat{X}_k = \hat{X}_{k/k-1} + K_k (Z_k - \hat{Z}_{k/k-1}) \quad (24)$$

$$P_k = P_{k/k-1} + K_k P_{ZZ} K_k^T \quad (25)$$

Step7: For the next sample implements steps 2–6.

### 2.4. 3D reconstruction

Based on the two-dimensional trajectories obtained from the two camera views, the three-dimensional motion information of Drosophila was further obtained. Before 3D reconstruction, we need to solve the problem of pairing and aligning the trajectories in different viewpoints. We have obtained the 2D trajectories of Drosophila in two different viewpoints, and the next step is to compute the trajectories of the same target in two viewpoints, that is to implement the pairing of trajectories in two viewpoints. On the other hand, considering the number of Drosophila and other factors, there may be occlusion and missed detection. To address the above problems, this paper adopts DTW to solve the problem. The DTW method calculates the minimum path based on the distance between the points of two sequences to derive the similarity. It can solve the problem of length inconsistency between time series by automatic regularization.

Triangulation is a method used to reconstruct a three-dimensional scene from a two-dimensional image with two or more viewpoints (Yuen and MacDonald, 2005; Avidan and Shashua, 2000). The basic

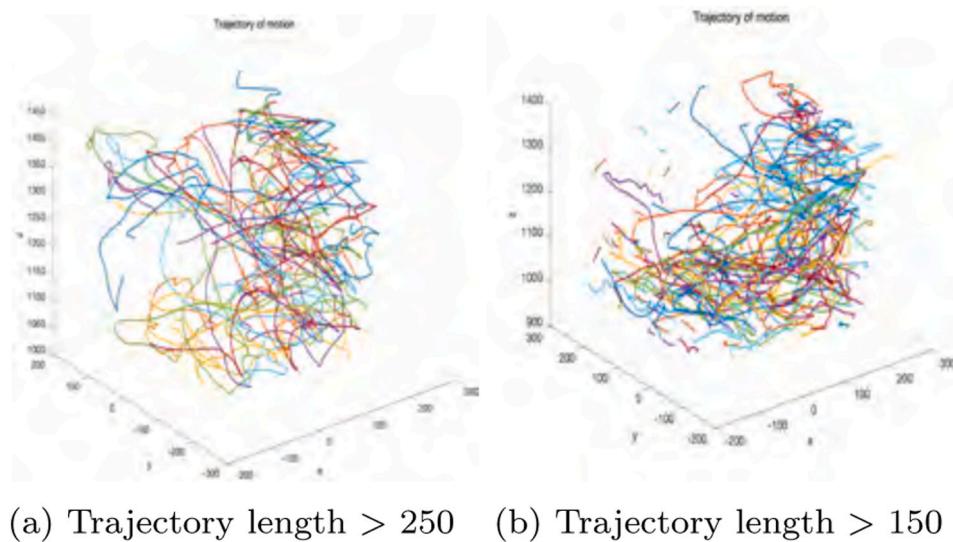


Fig. 3. Visualization of Drosophila trajectories.

**Table 1**  
HMM result.

u	3	4	5	7
accuracy	60.31	73.02	81.44	90.24

**Table 2**  
Comparison of object detection performance.

Method	Accuracy	Recall	F1 score
SIFT	0.93	0.91	0.92
SURF	0.90	0.88	0.89
ORB	0.87	0.85	0.86
AKAZE	0.95	0.91	0.93
T-AKAZE	0.96	0.95	0.95

principle of this method is to measure the position of a target in different viewpoints and use geometric relations to calculate the coordinates of the target in three-dimensional space. We obtained a sequence of trajectories aligned with the target in two viewpoints by the DTW method. In addition, the internal and external parameters of the camera are obtained using the checkerboard grid calibration method. Finally, the coordinates of the target in 3D space were calculated using the 2D coordinates of the target in different viewpoints and the camera parameters.

### 3. Results

Through the above steps, the three-dimensional coordinates of Drosophila movements were obtained, and the flight trajectories of normal flies and *ebo*<sup>678</sup> flies were plotted in Fig. 3. By calculating the statistical characteristics of the kinematics and spatial distribution of normal flies and *ebo*<sup>678</sup> flies, we found that there were notable differences in their flight behaviors. In addition, there were inter-individual interactions between Drosophila colonies in flight, and the tendency of their interactions decreased exponentially with the increase of the distance between individuals.

For the result  $o$  obtained by HMM inference, we set a tolerance value  $u$ . If the angle change occurring at the next moment is at  $[o - u, o + u]$ , then we consider the target as the possible next motion position. That is, the range is limited to a sector by the HMM. Table 1 shows the effect of different tolerance values on the inference results of HMM. Finally, considering the computational effort and accuracy, we set the value of  $u$

**Table 3**  
Performance measure with  $E_{ca}$ .

N	40	60	80	100
Kalman Filter	0.082	0.227	0.403	0.621
Unscented Kalman Filter	0.060	0.147	0.314	0.500
Particle Filter	0.056	0.122	0.301	0.496
HMET	0.058	0.098	0.292	0.483

to 5.

At the same time, this paper also compares the performance of feature matching algorithms, including SIFT (Lowe, 2004), SURF (Bay et al., 2006), ORB (Rublee et al., 2011), and AKAZE (Alcantarilla and Solutions, 2011) algorithms. The evaluation is based on accuracy, recall rate, and F1 score, and from Table 2, we can conclude that the T-AKAZE algorithm achieves good results.

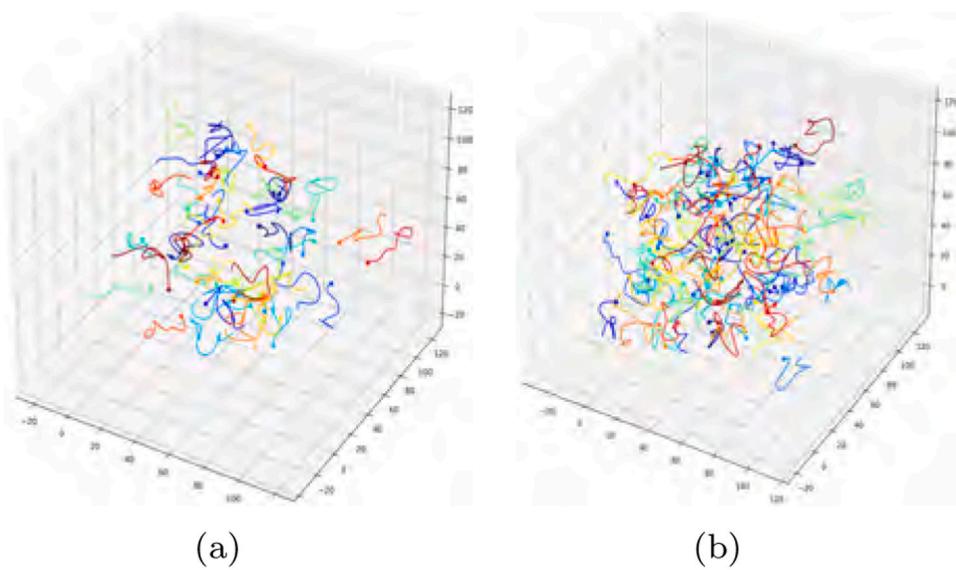
We use Correspondence and association errors ( $E_{ca}$ ) to measure tracking performance, which is defined as follows:

$$E_{ca} = \frac{N_c + N_a}{T} \quad (26)$$

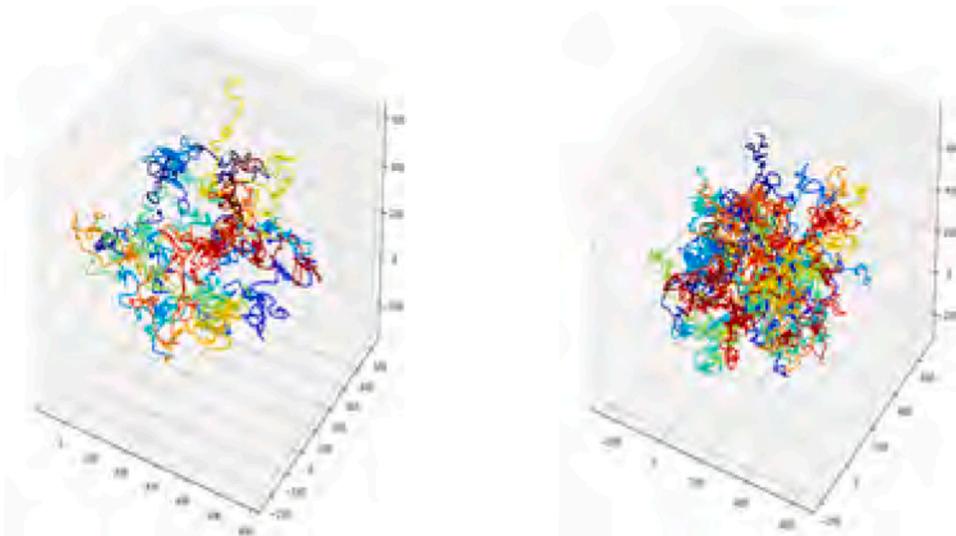
where  $N_c$  is the number of incorrect stereo matches,  $N_a$  is the number of incorrect temporal associations, and  $t$  is the number of frames. The evaluation results of the different methods are given in Table 3, and we can conclude that the proposed method is effective in tracking the simulated fruit flies.

In this paper, the positional trajectories of the target objects in the simulation data are generated by the Boids model (Reynolds, 1987), which is composed of four main components: the individual model, perception system, behavior system, and complex behavior. This algorithm simulates simple interaction rules among individuals within a group, leading to the emergence of complex group behaviors.

We set different group sizes  $N = 50, 100, 150, 200, 250$ , to investigate the dynamics of the group under varying densities. Figs. 4 and 5 displays the trajectories and tracking results generated by simulating with the Boids algorithm for  $N = 50$  and  $N = 100$ . Each group size is simulated in a three-dimensional space for 200 frames to generate a dataset under different density conditions. These datasets will be used to evaluate and validate the performance of the algorithm we propose, particularly its accuracy and robustness in handling the dynamics of groups of varying sizes. During the simulation process, the initial positions and velocities of each fruit fly individual are randomly generated, ensuring diversity and generalization capability in our experiments. By



**Fig. 4.** The flight conditions of two types of fruit flies under different time lengths.



**Fig. 5.** Simulation results for Boids.

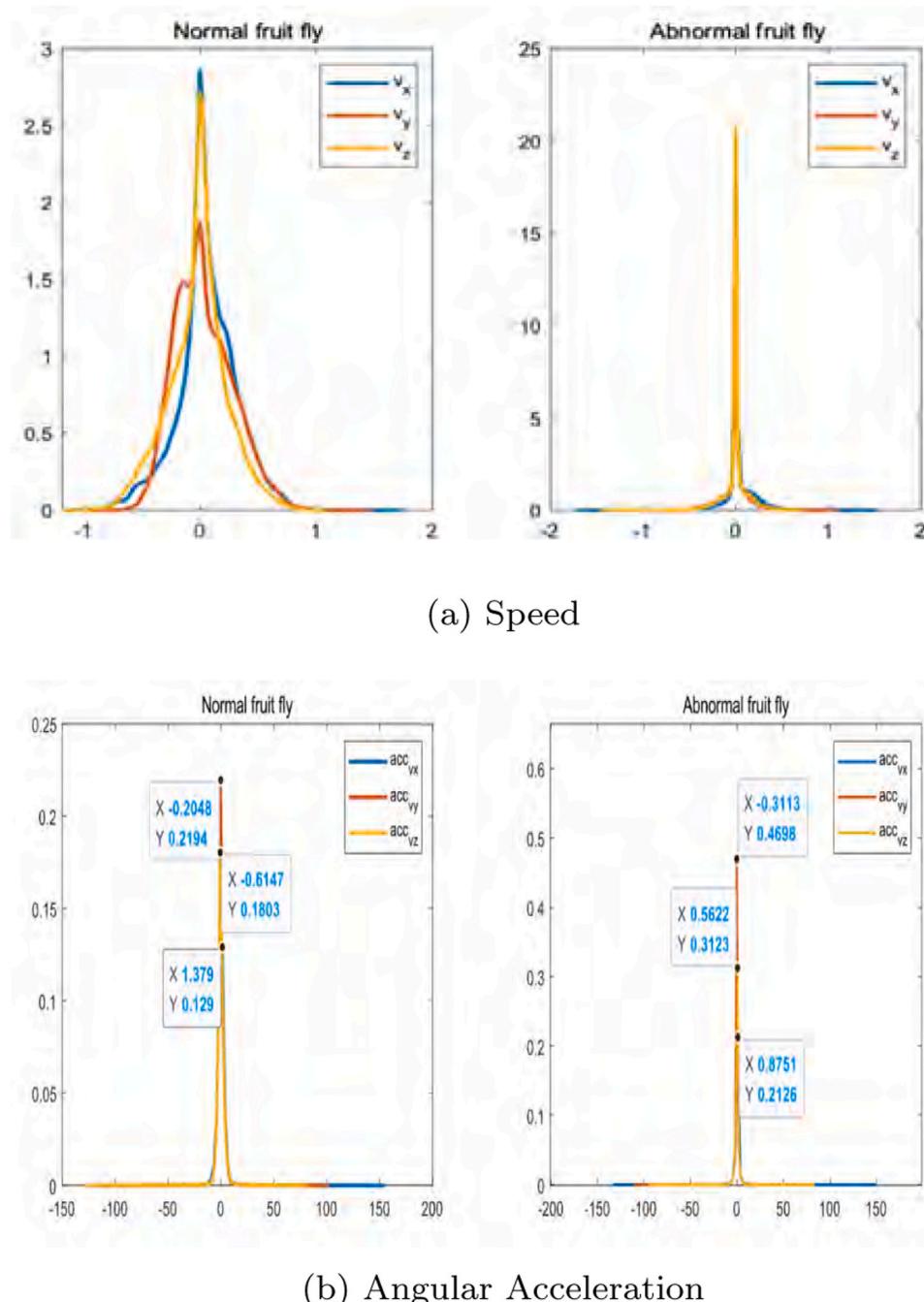
adjusting the parameters of the Boids model, such as the minimum distance between individuals, velocity matching range, and the degree of group cohesion, we can simulate fruit fly group movement characteristics that closely resemble those in the real world. The generated datasets not only capture the movement trajectories of the fruit fly groups in three-dimensional space but also reflect the complex interactions and behavior patterns within the group.

The design of the simulation experiment and the data generation process provide us with a rich testing scenario to comprehensively evaluate the effectiveness of the proposed algorithm in handling complex dynamic group behaviors. By comparing the movement data of groups at different scales, we can gain a deeper understanding of the algorithm's performance under various conditions, offering significant experimental evidence for further optimization and application.

#### 4. Discussion

When studying the walking state of *Drosophila melanogaster*, scientists usually focus on aspects such as *Drosophila*'s movement patterns

and the control of movement by gait characteristics. These studies have contributed to a deeper understanding of the animal's locomotor mechanisms. In contrast, fewer studies have been conducted on the flying state of *Drosophila melanogaster*. However, some studies have shown that the flying behavior of *Drosophila melanogaster* also has considerable research value. For example, it has been found that *Drosophila melanogaster* often adopts a straight line flight interspersed with a 90-degree turn at fly speed, and this flight trajectory is thought to be a characteristic of Levi's flight (Viswanathan et al., 1996). Levi's flight is an efficient search strategy that helps animals find food or habitat in unknown environments. From the perspective of neuroscience, the flight movements of *Drosophila* involve complex neural circuits and signaling (Lehmann and Bartussek, 2017; Chan et al., 1998), and the analysis of flight movements can provide insights into the effects of the allele *ebo*<sup>678</sup> on neuromodulatory mechanisms, which in turn can lead to a better understanding of the role of the nervous system in complex motor behaviors. In addition, aerodynamic properties and sensory feedback mechanisms in flight are crucial for navigation and obstacle avoidance in *Drosophila* (Srinivasan, 2010; Mronz and



**Fig. 6.** Presentation of statistical results of Drosophila flight data.

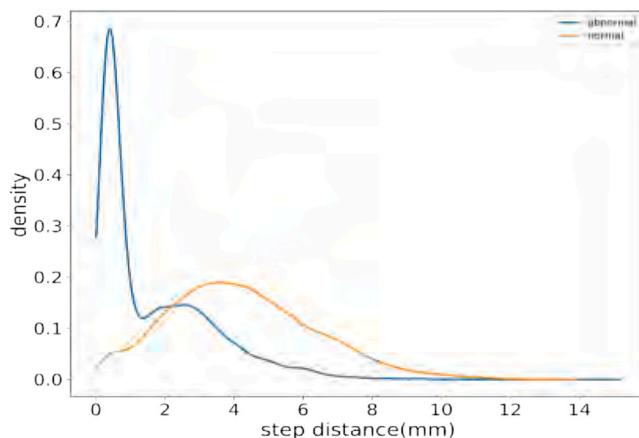
Lehmann, 2008), and mutations in the allele *ebo*<sup>678</sup> may alter these abilities in Drosophila during flight. Therefore, studying the flying state of *Drosophila melanogaster* could provide new perspectives for developing an understanding of its behavior and adaptations.

Firstly, the flight trajectories of *Drosophila* were clustered by the position coordinates under the 3D trajectory, in order to find out the interaction behavior generated during its flight on the one hand, and to analyze the state transitions occurring from the initial state to the final stable phase on the other. There are two main categories of possible interactions between *Drosophila*: the first is where the flies fly in opposite directions, with their neural mechanisms issuing collision avoidance commands and consequent angular shifts; the second is where groups of flies fly together, i.e. at a relatively constant distance from each other. In the former scenario, the distances, angular velocities,

steering angles and velocities are further analyzed; in the latter scenario, the entire *Drosophila* population is first clustered, then the changes in distance and angular acceleration between classes are analyzed, and finally the differences between normal and *ebo*<sup>678</sup> flies are compared.

Here, we analyzed the characteristics of normal flies and *ebo*<sup>678</sup> flies on the property of speed. Fig. 6 (a) and (b) shows the probability density functions of the two types of *Drosophila* on the three speed components of xyz. We can obviously see that the variance in the three components is larger in normal *Drosophila* than in the allele *ebo*<sup>678</sup> flies, which means that the speed of *ebo*<sup>678</sup> flies mostly fluctuates in a small range and the variation is not very significant, compared to normal *Drosophila*, which has a more complete neural mechanism in the brain and its speed fluctuates a bit more.

On the other hand, we present the probability density distribution of



**Fig. 7.** PDF of Drosophila step size.

the two types of fruit flies regarding the property of acceleration, and we can conclude that the probability density distribution of acceleration of the *ebo*<sup>678</sup> flies is relatively concentrated, which means that the speed of these flies fluctuates slightly, and their flight behavior does not change much during the flight and is relatively stable.

Filtering the trajectory lengths, further analysis was conducted to examine the motion characteristics during continuous movement intervals. Fig. 8 illustrate the changes in fly angular acceleration within this time range for normal and *ebo*<sup>678</sup> flies. By comparing the statistical results of angular acceleration between the two types of fruit flies, we observe that during certain periods of the entire flight process of normal fruit flies, there are significant fluctuations in angular acceleration. In contrast, *ebo*<sup>678</sup> flies exhibit sustained large fluctuations in angular acceleration throughout the entire motion process, rather than smooth, periodic changes. This observation suggests the impact of allele *ebo*<sup>678</sup> mutation on the flight dynamics of fruit flies, revealing similarities and fluctuations in angular acceleration. These flies exhibit weaker abilities for long-term flight planning, remaining in a state of continuous, intense changes to avoid collisions with obstacles in their environment.

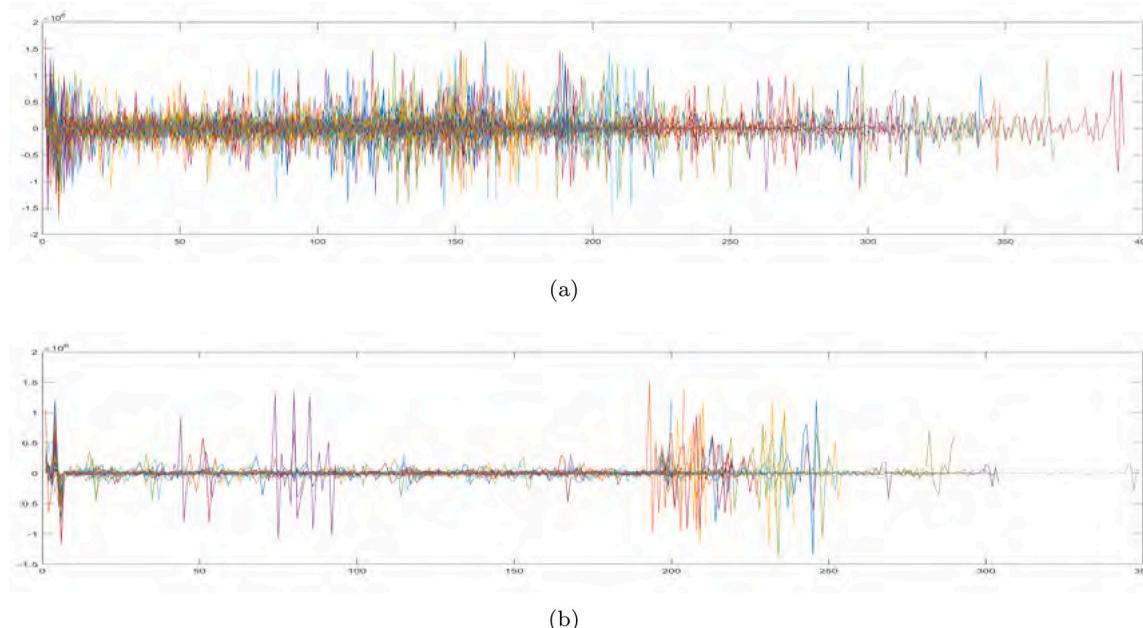
Fig. 7 depict the probability density functions (PDFs) of step distances for both normal and *ebo*<sup>678</sup> flies. Here, step distance is defined as

the distance of fly movement between consecutive frames. From the graphs, it can be observed that the step distance distribution for normal fruit flies (represented by the orange line) is relatively broad, indicating greater variability in step distances, with the peak density corresponding to larger step distances compared to *ebo*<sup>678</sup> flies. This suggests that normal fruit flies exhibit more variable step lengths during movement, with a tendency towards longer step distances. In contrast, the step distance distribution for *ebo*<sup>678</sup> flies (represented by the blue line) is more concentrated, with higher peak density values and shorter corresponding step distances. This indicates that *ebo*<sup>678</sup> flies display more consistent and generally shorter step distances Fig. 8.

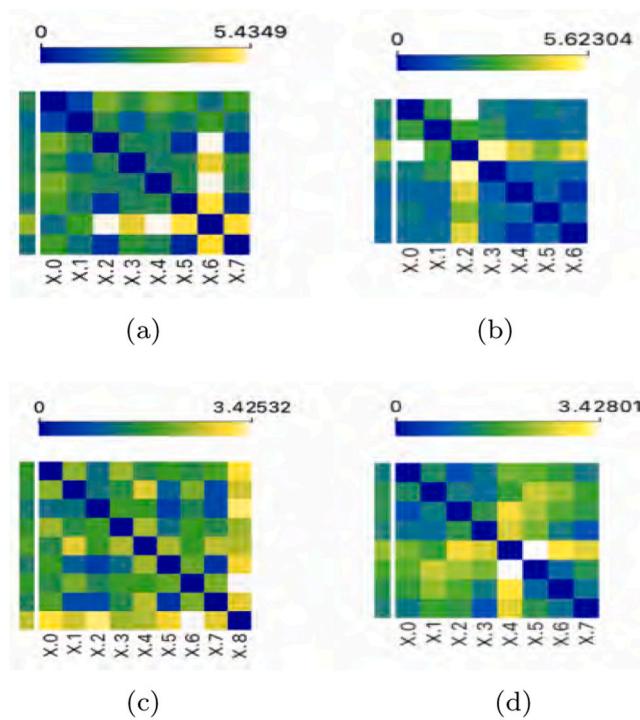
The ellipsoid body of the fruit fly is part of the central complex, a structure in the fly brain closely associated with navigation and motor control. Therefore, individuals with mutations in the *ebo*<sup>678</sup> allele may affect the fly's motor patterns or its ability to perform complex movements. In this context, we observe significant differences in step distance distribution between *ebo*<sup>678</sup> flies and normal fruit flies, with the former exhibiting more consistent and generally shorter step distances. This could suggest that genetic mutation in the ellipsoid body region influences the spatial orientation and motor coordination of fruit flies, resulting in more constrained movement patterns. This consistency and reduction in step distances may reflect alterations in the fly's behavior while exploring the environment, possibly due to changes in perception, motor planning, or balance.

Fig. 9 displays the differences in angular acceleration correlation between normal fruit flies and *ebo*<sup>678</sup> flies, with each subplot representing a time segment of length  $\tau$ . The color intensity within the blocks indicates the magnitude of the Fréchet distance. The top row of two subplots shows greater variability in angular acceleration among *ebo*<sup>678</sup> flies individuals within the same time window. This variability could reflect behavioral variations caused by genetic mutation, affecting the flies' flight patterns and making their flight behavior more uncoordinated, such as exhibiting more individual differences in behaviors like collision avoidance or foraging. The bottom row of two subplots presents relatively lower Fréchet distance values, indicating that the flight behaviors among normal fruit flies are more similar within these time windows. This similarity might be due to their adoption of comparable strategies for obstacle avoidance, foraging behaviors, or social interactions during flight.

In addition, we calculated the nearest neighbor distances for both



**Fig. 8.** Visualization of differences in angular acceleration among normal and *ebo*<sup>678</sup> flies.



**Fig. 9.** Correlation after clustering by dtw.

types of fruit flies during flight. Firstly, trajectories with a duration less than 150 frames were filtered out. Fig. 10 illustrates the nearest neighbor distances for each frame for the two types of fruit flies. By comparing the nearest neighbor distances of *ebo*<sup>678</sup> flies and normal fruit flies over long-term trajectories, differences in their flight behavior and social space are revealed. Despite the significant fluctuations in nearest

neighbor distances during flight for *ebo*<sup>678</sup> flies (ranging from 0 to 120 mm), their maximum spatial separation from other individuals is notably smaller than that of normal fruit flies (ranging from 0 to 175 mm). This comparison indicates that although *ebo*<sup>678</sup> flies may undergo rapid spatial positional changes at local time scales during flight, overall, they maintain a more compact group structure.

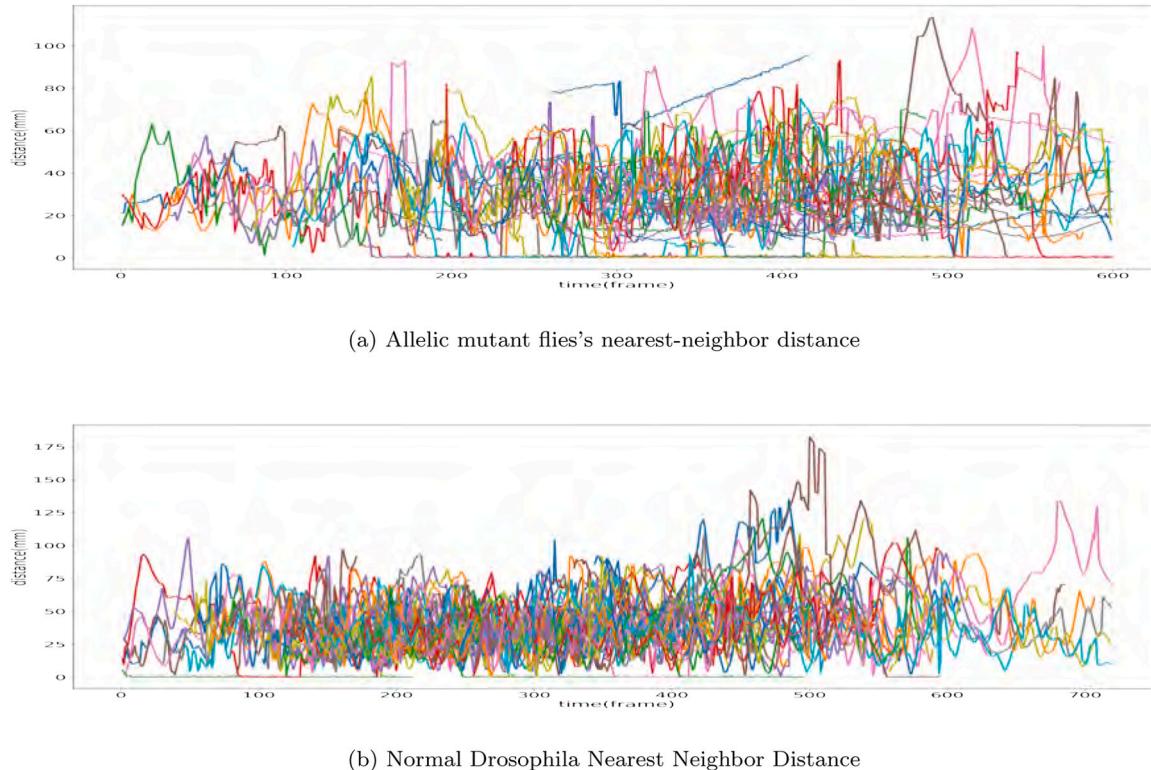
Overall, the trajectory segments of normal fruit flies show significantly higher angular acceleration correlation within the same time window compared to those of *ebo*<sup>678</sup> flies. This difference may suggest that mutant alleles in the Drosophila's *ebo* affects the behavior patterns of fruit flies in the EB region, subsequently impacting their social interaction patterns.

However, we acknowledge the challenges such as occlusion and missed detections inherent in this setup. To address these limitations and further enhance our tracking system, we propose utilizing a triocular camera setup for capturing the motion imagery of targets. Triocular cameras offer superior capability in capturing target positions and motion trajectories, thereby mitigating instances of occlusion and missed detections. Additionally, leveraging the triocular setup is expected to augment the inference capabilities of HMM. Therefore, this enhancement not only reduces occlusion and missed detection but also enhances the inference capabilities of HMM, ultimately improving the overall tracking accuracy and reliability of our system.

## 5. Conclusion

In this study, we present a new method that combines machine learning techniques and computer vision algorithms to predict fruit fly movement trajectories. Our main goal is to further quantify the behavioral characteristics of flying fruit flies by tracking their flight trajectories as well as acquiring 3D data of their movements.

The process of this paper is mainly divided into 5 steps, the first step is to process the images acquired by the camera to detect the positional coordinates of the fruit fly in each frame as accurately as possible, calculate the angular change of the fruit fly's movement, and use the



**Fig. 10.** The nearest-neighbor distance among two types of Drosophila.

HMM to infer the angle of transition of the fruit fly's next movement. Utilizing with this angle a set of candidate positions are generated in a specified area around the current position of the fruit fly. Next, we use the UKF to fuse the position information of the fruit flies together to predict the flight position of the fruit flies. After obtaining the fruit fly motion data from both cameras, 3D reconstruction is performed to further analyze the behavioral characteristics of the fruit flies. Finally, we quantified the differences between normal flies and *ebo*<sup>678</sup> flies in terms of step distance, nearest neighbor distance, angular acceleration, and trajectory clustering.

## 6. Future directions

Within the scope of our current research on the HMUT, future investigative pathways emerge prominently in the quest for optimization and enhancement of tracking algorithms. The optimization endeavors could include exploring more efficient feature extraction methods that boost processing velocity and precision, coupled with the study of adaptive algorithm parameters that automatically adjust to the dynamism of environmental shifts. Moreover, integrating a diverse spectrum of sensor data, like sonar or radar, may significantly uplift the robustness of tracking in varied settings, paving the way for algorithmic applications in real-time performance exigencies, including potential hardware acceleration advancements.

The latent utility of the HMUT algorithm extends well beyond the tracking of fruit flies within a sealed container, portending transformative applications across domains such as autonomous driving and robotic navigation, where environmental perception and obstacle evasion are critical. In public safety monitoring systems, persistent tracking of specific targets can be instrumental. Moreover, the algorithm's applicability to the study of behavioral patterns in biological and ecological research offers profound insights. In medical imaging and monitoring, particularly in endoscopy or the tracking of microscopic entities, the algorithm's precision can be crucial. With the expansion of such applications, the algorithm's generalizability to diverse scales and speeds of targets, as well as its adaptability to various environmental conditions, emerges as an unresolved inquiry. Computational resource demands pose significant considerations, especially for deployment within low-power or embedded systems. The robustness of the algorithm under extreme or noisy conditions necessitates further exploration. Tackling these unresolved issues will not only refine the HMUT algorithm but also broaden its applicability, potentially revolutionizing the field of tracking technology across numerous domains.

## CRediT authorship contribution statement

**Fei Ying Lu:** Writing – original draft, Visualization, Methodology. **Xiang Liu:** Writing – review & editing, Resources. **Hai Feng Su:** Resources, Data curation. **Shuo Hong Wang:** Resources, Data curation.

## Data availability

The data that has been used is confidential.

## References

- Ahmadi, K., Salari, E., 2016. Small dim object tracking using frequency and spatial domain information. *Pattern Recognit.* 58, 227–234.
- M. Aktaş and H.F. Ateş. Small object detection and tracking from aerial imagery. In: Proceedings of the Sixth International Conference on Computer Science and Engineering (UBMK), 688-693.IEEE, 2021.
- Alcantarilla, P.F., Solutions, T., 2011. Fast explicit diffusion for accelerated features in nonlinear scale spaces. *IEEE Trans. Pattern Anal. Mach. Intell.* 34 (7), 1281–1298.
- Avidan, S., Shashua, A., 2000. Trajectory triangulation: 3d reconstruction of moving points from a monocular image sequence. *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (4), 348–357.
- Ballerini, M., Cabibbo, N., Candelier, R., Cavagna, A., Cisbani, E., Giardina, I., Lecomte, V., Orlandi, A., Parisi, G., Procaccini, A., et al., 2008. Interaction ruling animal collective behavior depends on topological rather than metric distance: evidence from a field study. *Proc. Natl. Acad. Sci.* 105 (4), 1232–1237.
- Bay, H., Tuytelaars, T., Van Gool, L., 2006. Surf: Speeded up robust features. In: *Proceedings of the Computer Vision-ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I* 9. Springer, pp. 404–417.
- Branson, K., Robie, A.A., Bender, J., Perona, P., Dickinson, M.H., 2009. High-throughput ethomics in large groups of drosophila. *Nat. Methods* 6 (6), 451–457.
- Buhl, J., Sumpter, D.J.T., Couzin, I.D., Hale, J.J., Despland, E., Miller, E.R., Simpson, S.J., 2006. From disorder to order in marching locusts. *Science* 312 (5778), 1402–1406.
- Chan, W.P., Prete, F., Dickinson, M.H., 1998. Visual input to the efferent control system of a fly's "gyroscope". *Science* 280 (5361), 289–292.
- Faraz, K., Grenier, T., Ducotter, C., Epicier, T., 2022. Deep learning detection of nanoparticles and multiple object tracking of their dynamic evolution during *in situ* stem studies. *Sci. Rep.* 12 (1), 2484.
- Forney, G.D., 1973. The viterbi algorithm. *Proc. IEEE* 61 (3), 268–278.
- Gernat, T., Rao, V.D., Middendorf, M., Dankowicz, H., Goldenfeld, N., Robinson, G.E., 2018. Automated monitoring of behavior reveals bursty interaction patterns and rapid spreading dynamics in honeybee social networks. *Proc. Natl. Acad. Sci. USA* 115 (7), 1433–1438.
- Grover, D., Tower, J., Tavaré, S., 2008. O fly, where art thou? *J. R. Soc. Interface* 5 (27), 1181–1191.
- Han, Y., Deng, C., Zhao, B., Zhao, B., 2019. Spatial-temporal context-aware tracking. *IEEE Signal Process. Lett.* 26 (3), 500–504.
- U. Hanesch. Der zentralkomplex von *Drosophila melanogaster* (PhD thesis), 1987.
- Heisenberg, M., Borst, A., Wagner, S., Byers, D., 1985. *Drosophila* mushroom body mutants are deficient in olfactory learning. *J. Neurogenet.* 2 (1), 1–30.
- Ilius, M., Wolf, R., Heisenberg, M., 2007. The central complex of *Drosophila melanogaster* is involved in flight control: studies on mutants and mosaics of the gene ellipsoid body open. *J. Neurogenet.* 21 (4), 321–338.
- Jezovit, J.A., Alwash, N., Levine, J.D., 2021. Using flies to understand social networks. *Front. Neural Circuits* 15, 755093.
- Julier, S.J., Uhlmann, J.K., 2004. Unscented filtering and nonlinear estimation. *Proc. IEEE* 92 (3), 401–422.
- Kohlhoff, K.J., Jahn, T.R., Lomas, D.A., Dobson, C.M., Crowther, D.C., Vendruscolo, M., 2011. The ifly tracking system for an automated locomotor and behavioural analysis of *Drosophila melanogaster*. *Integr. Biol.* 3 (7), 755–760.
- Lehmann, F.-O., Bartussek, J., 2017. Neural control and precision of flight muscle activation in drosophila. *J. Comp. Physiol. A* 203, 1–14.
- Liu, C., Ding, W., Yang, J., Murino, V., Zhang, B., Han, J., Guo, G., 2019. Aggregation signature for small object tracking. *IEEE Trans. Image Process.* 29, 1738–1747.
- Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* 60, 91–110.
- Makris, N.C., Ratilal, P., Jagannathan, S., Gong, Z., Andrews, M., Bertsatos, I., Godo, O. R., Nero, R.W., MichaelJech, J., 2009. Critical population density triggers rapid formation of vast oceanic fish shoals. *Science* 323 (5922), 1734–1737.
- Martín-Peña, A., Acebes, A., Rodríguez, J.-R., Chevalier, V., Casas-Tinto, S., Triphan, T., Strauss, R., Ferrús, A., 2014. Cell types and coincident synapses in the ellipsoid body of drosophila. *Eur. J. Neurosci.* 39 (10), 1586–1601.
- Mronz, M., Lehmann, F.-O., 2008. The free-flight response of *drosophila* to motion of the visual environment. *J. Exp. Biol.* 211 (13), 2026–2045.
- R.M. Neal. Probabilistic Inference Using Markov Chain Monte Carlo Methods.1993.
- Rabiner, L., Juang, B., 1986. An introduction to hidden markov models. *IEEE ASSP Mag.* 3 (1), 4–16.
- Renn, S.C.P., Armstrong, J.D., Yang, M., Wang, Z., An, X., Kaiser, K., Taghert, P.H., 1999. Genetic analysis of the drosophila ellipsoid body neuropil: organization and development of the central complex. *J. Neurobiol.* 41 (2), 189–207.
- C.W. Reynolds. Flocks, herds and schools: A distributed behavioral model. In: *Proceedings of the Fourteenth Annual Conference on Computer Graphics and Interactive Techniques*, 25-34, 1987.
- Rossi, C., Abderrahim, M., Diaz, J.C., 2008. Tracking moving optima using kalman-based predictions. *Evolut. Comput.* 16 (1), 1–30.
- E. Rublee, V. Rabaud, K. Konolige, and G. Bradski.Orb: An efficient alternative to sift or surf. In: *Proceedings of the International conference on computer vision*, 2564-2571. Ieee, 2011.
- Solanki, N., Wolf, R., Heisenberg, M., 2015. Central complex and mushroom bodies mediate novelty choice behavior in drosophila. *J. Neurogenet.* 29 (1), 30–37.
- Srinivasan, M.V., 2010. Honey bees as a model for vision, perception, and cognition. *Annu. Rev. Entomol.* 55 (1), 267–284.
- Strauss, R., Heisenberg, M., 1993. A higher control center of locomotor behavior in the drosophila brain. *J. Neurosci.* 13 (5), 1852–1861.
- Straw, A.D., Branson, K., Neumann, T.R., Dickinson, M.H., 2011. Multi-camera real-time three-dimensional tracking of multiple flying animals. *J. R. Soc. Interface* 8 (56), 395–409.
- Viswanathan, G.M., Afanasyev, V., Buldyrev, S.V., Murphy, E.J., Prince, P.A., Stanley, H. E., 1996. Lévy flight search patterns of wandering albatrosses. *Nature* 381 (6581), 413–415.
- Woods, J., Radewan, C., 1977. Kalman filtering in two dimensions. *IEEE Trans. Inf. Theory* 23 (4), 473–482.
- H.S. Wu, Q. Zhao, D. Zou, and Y.Q. Chen.Acquiring 3d motion trajectories of large numbers of swarming animals. In: *Proceedings of the IEEE Twelfth International Conference on Computer Vision Workshops, ICCV Workshops*, 593-600.IEEE, 2009.
- Wu, H.S., Zhao, Q., Zou, D., Chen, Y.Q., 2011. Automated 3d trajectory measuring of large numbers of moving particles. *Opt. Express* 19 (8), 7646–7663.

- Yan, W., Lin, H., Yu, J., Wiggin, T.D., Wu, L., Meng, Z., Liu, C., Griffith, L.C., 2023. Subtype-specific roles of ellipsoid body ring neurons in sleep regulation in *Drosophila*. *J. Neurosci.* 43 (5), 764–786.
- J. Yick, B. Mukherjee, and D. Ghosal. Analysis of a prediction-based mobility adaptive tracking algorithm. In: Proceedings of the Second International Conference on Broadband Networks, 2005, 753-760.IEEE, 2005.
- Yuen, D.C.K., MacDonald, B.A., 2005. Vision-based localization algorithm based on landmark matching, triangulation, reconstruction, and comparison. *IEEE Trans. Robot.* 21 (2), 217–226.
- X. Zhang, T. Zhang, Y. Yang, Z. Wang, and G. Wang. Real-time golf ball detection and tracking based on convolutional neural networks. In: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2808–2813.IEEE, 2020.
- Zhu, Y., Li, C., Liu, Y., Wang, X., Tang, J., Luo, B., Huang, Z., 2023. Tiny object tracking: a large-scale dataset and a baseline. *IEEE Trans. Neural Netw. Learn. Syst.*