Honeybee Swarm Dynamics: Investigating the Relationship Between Individual Decision Making and Collective Foraging

Javon Hickmon [1] **Golnar Gharooni Fard** [2] **Orit Peleg** [3 4]

Abstract

Honeybees are social insects that utilize pheromone signals to direct each other, resulting in emergent swarm behavior. Research has shown that bees can solve the shortest path problem and locate their queen by directing these pheromone signals, or "scenting," to create a communication network collectively (2); however, little work has been done to study how the behavior of the bees changes when in the presence of the swarm. This study aims to investigate how the decisions of individual honeybees affect the foraging and aggregation process of the collective . We utilize state-of-the-art video object tracking and segmentation tools to gather time- series data on the scenting behaviors of individual bees when separated from their queen. Combining this with additional computer vision approaches described in the literature, We analyze the behavior patterns of the individuals as the swarm forages for and aggregates around the queen. Through this analysis, We anticipate the time-series data from individual honeybees will display a predictable relationship between scenting and exploring that allows for the emergent pathfinding capability of the swarm . This improved tracking of individual honeybees will allow us to increase the accuracy of the multi-agent reinforcement learning model to predict the behaviors of the swarm. These improvements directly impact swarm robotics and tasks such as patrolling, disaster recovery, and search and rescue, where it is important to create a low- resource distributed system that can rapidly adapt to new operating conditions.

[1]Department of Computer Science, University of Washington, Seattle, WA [2]Department of Computer Science, University of Colorado, Boulder, CO [3]BioFrontiersInstitute, University of Colorado Boulder, Boulder, CO [4]Santa Fe Institute, Santa Fe, NM.

Code available at: https://github.com/peleg-lab/IndividualTracker

1. Introduction

A vast body of research has surrounded the topic of honeybee swarms. To become a coherent swarm, bees locate their queen by tracking her pheromones. Previous work has shown that bees collectively create a scenting-mediated communication network by arranging in a specific spatial distribution where there is a characteristic distance between individuals and directional signaling away from the queen. Rather than depositing static information in the environment, individual bees locally sense and globally manipulate the physical fields of chemical concentration and airflow. Individual bees act as receivers and senders of signals by using the Nasonov scenting behavior, releasing pheromones from the glands and fanning their wings to direct the signals backward. In this network, scenting bees stand at a characteristic distance from their neighbors while dispersing signals, which suggests a concentration threshold in the activation mechanism of individual bees' scenting behavior. The scenting events are highly correlated with the collective aggregation around the queen (3).

Despite previous research showing the correlation between scenting events and collective aggregation around the queen, little work has been done to study how the behavior of individual bees changes when in the presence of the swarm. This was because we were only able to detect the states of the bees; however, previous methods were not able to able to track them over longer periods of time.

2. Methods

To track individual bees and understand their behavior in the presence of the swarm, we established a environment in which worker bees search for a stationary caged queen in a semi–two-dimensional (2D) arena. We recorded the search and aggregation behavior of the bees from an aerial view with tests containing one, two, three, and four worker bees. To extract data from the recordings, we then developed a markerless, semi-automatic, highly accurate tracking and segmentation algorithm using a variety of computer vision methods (3). This pipeline improves upon previous work by allowing us to detect and track individual bees throughout the video, while also identifying the positions and orienta-

Algorithm 1 Honeybee Tracking and Segmentation

```
background = get_background(video)
for frame ∈ video do
    processed = preprocess(frame, background)
    contours = form_contours(processed)
    tracks = create_tracks(contours, prev_tracks)
    groups = [group in tracks]
    if len(groups) > 0 then
        split_tracks = split_groups(tracks)
        tracks.remove(groups)
        tracks.extend(split_tracks)
    end if
    frame = draw_tracks(frame, tracks)
    show_frame(frame)
    prev_tracks = split_tracks
end for
```

tions of scenting bees each frame.

2.1. The Algorithm

get_background(video) - Since the background of our video is stationary we utilize a standard method of video background subtraction by defining the background to be the median of each frame. By using this temporal median approach, we can approximate the background pixels and utilize that for background subtraction.

Figure 1. get_background()

preprocess(frame, background) - To preprocess our frame, we utilize traditional image processing techniques to isolate the honeybees. We start by subtracting the background from the current frame. From there, we blur the resulting image in order to reduce noise for our next step, thresholding. We apply a binary threshold to the grayscale image to isolate the bees. After this step, we apply a variety of morphological image operations in order to reduce the noise and improve the consistency of detecting the honeybees within the thresholded image. As a final step, depending on the quality of the resulting video we remove pixels of irrelevant artifacts such as the queen's cage or the border of the arena.

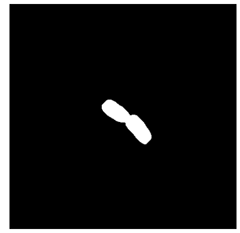


Figure 2. preprocess()

form_contours(processed) - We utilize the method described by Suzuki et al. (4) to form contours from the topological features of the preprocessed frame. We then leverage the topological structural analysis of the method to lift the bee contours from the background. Finally, an area filter is applied to remove noise from the resulting contours.

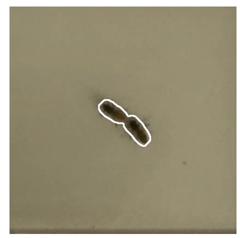
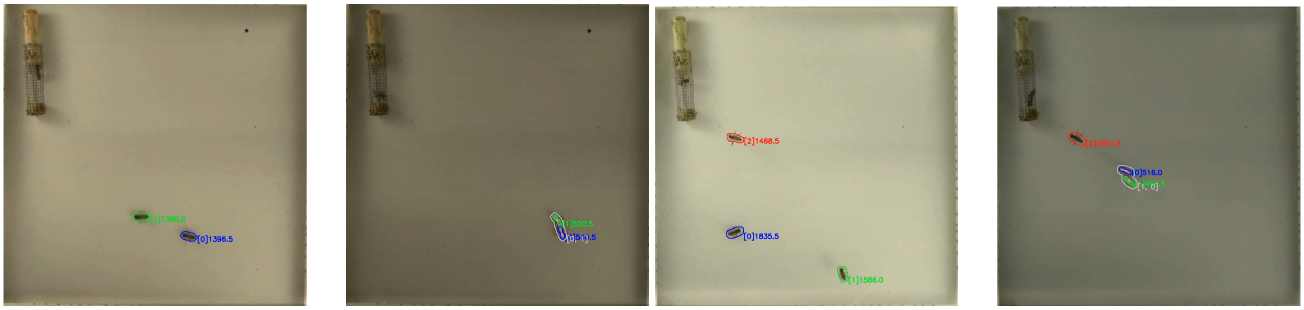


Figure 3. form_contours()

create_tracks(contours, prev_tracks) - A "track" is an object that contains labels, a center point, and a contour. We form these track objects and assign labels based on a greedy search algorithm that identifies the nearest previous track and takes that label.

Figure 4. create_tracks()

split_groups(tracks) - This function is the bulk of the bee tracking pipeline. In order to segment the individual bees from the group, we leverage the watershed algorithm. The watershed transformation treats the image it operates upon like a topographic map, with the brightness of each point representing its height, and finds the lines that run along the tops of ridges. Specifically we implement one of the variants of watershed, non-parametric marker-based segmentation algorithm, described by Meyer (1). Since the watershed algorithm requires landmark points to start, we utilize the previous center points for each of the track objects.



(a) Two honeybees detected by the tracker

(b) Two clustered honeybees detected by the tracker

(c) Three honeybees detected by the tracker

(d) Three clustered honeybees detected by the tracker

Figure 5. Honeybee tracking and segmentation results

If the current frame is the first frame and group separation is needed, then we enter a manual landmark selection mode where users can select the centers of the honeybees then run the watershed algorithm. After separating the honeybees, we then use the same greedy algorithm to relabel the contours and form the track objects.

Despite this process functioning a majority of the time, there are edge cases where it can fail such as when bees are on top of one another or if they rotate while being grouped. In order to mitigate these issues, we implemented a progressive color thresholding technique where bee contours are recalculated based on thresholding based on the color at the watershed landmark. We start with a large range of color values and progressively restrict the range until the watershed algorithm can separate the contours.

This additional step fixes a myriad of edge cases but occasionally the algorithm still cannot separate the contours. In this scenario, we enter a manual mode where users can select the landmarks for the watershed algorithm. Including this manual mode ensures that the resulting tracking objects are highly accurate and temporally consistent.

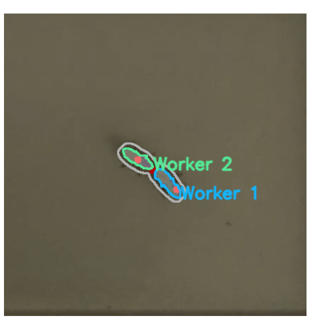


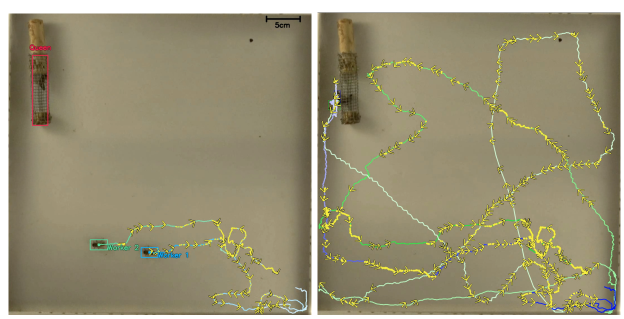
Figure 6. split_groups()

2.2. Figures

This algorithm is able to detect, track, and segment multiple honeybees throughout the entirety of the video analysis. We can see in Figure 5b and Figure 5d that the labels of the honeybees are preserved despite their bodies overlapping. The results of this highly-accurate segmentation are shown in Figure 5a and 5c. The bees are segmented properly, but their labels are also temporally consistent, so the labels for the bees at the start of the video are the exact same labels at the end of the video. This temporal consistency is exactly why the algorithm exceeds the performance of other state-of-the-art methods.

3. Results

By combining our novel algorithm with the honeybee scenting detection system developed in Nguyen et. al. (3) we are able to generate long-term trajectories for individual honeybees, and overlay these trajectories with individual scenting events as shown in Figure 7.



(a) Short-term trajectories

(b) Long-term trajectories

Figure 7. Honeybee trajectories

From this trajectory data, we then generate a number of plots to visualize the distribution of honeybee scenting events. These plots allow us to better understand the correlation between the time-series data and the density of scenting events.

3.1. Time Series Data

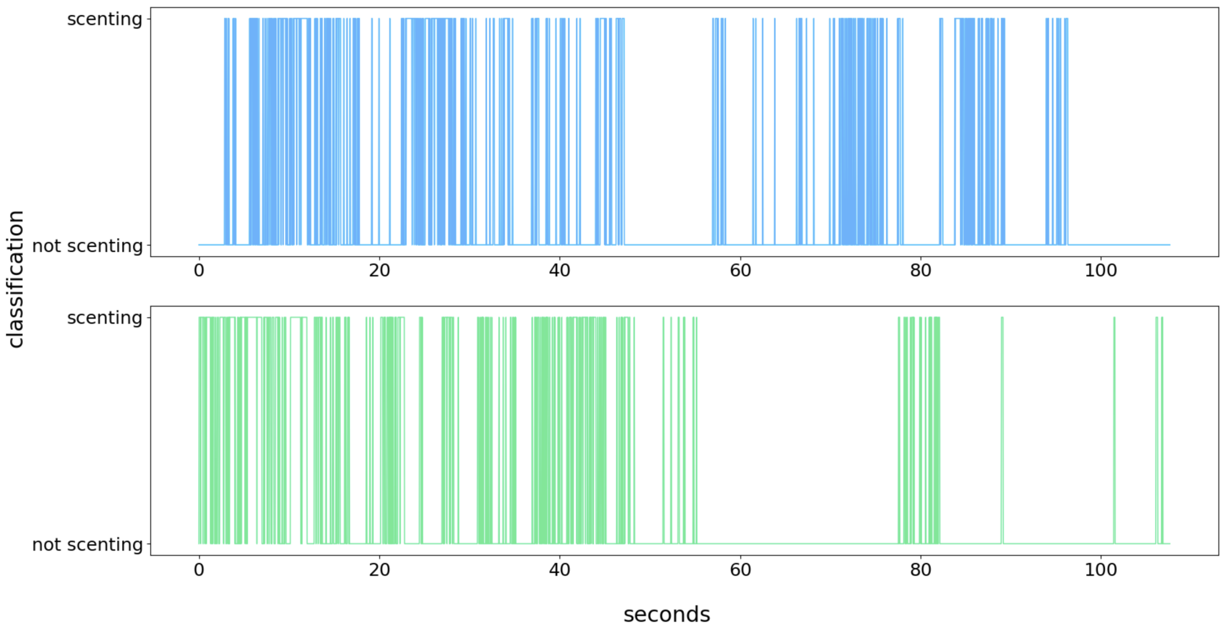


Figure 8. Time Series Data

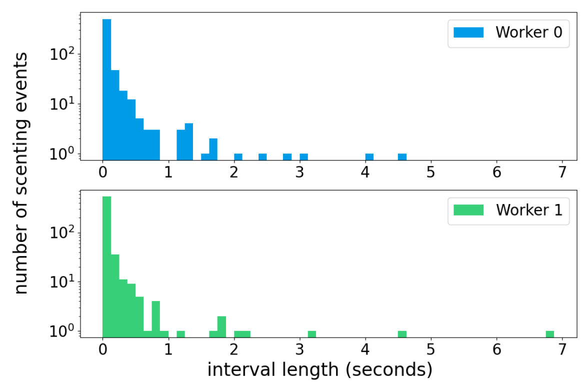
We then leverage the data generated from our trajectory plots to create a time series of the scenting events for each individual honeybee throughout the duration of the video. In the time series shown in Figure 8, each bee is classified as scenting or not scenting. Our classifier detects scenting events within 1/30th of a second, so the densely packed areas are characteristic of large bursts of consistent scenting.

3.2. Histogram of Scenting Intervals

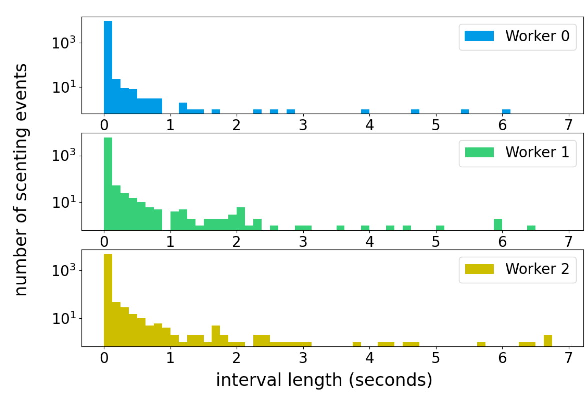
Our histograms plot the distribution of scenting events with the specified interval length, where the x-axis is the length (in seconds) of the interval between scenting events and the y-axis is the number of scenting events that had that interval length.

An interval is defined as the length of time between the current scenting event and the previous one. We decided to group intervals by the nearest 1/8th of a second so the plot would be easily interpretable (note that the y-axis is on a logarithmic scale).

We can clearly see most of the scenting events are extremely quick; however, there are still a significant amount of longer scenting intervals. To understand the results, we perform a correlation analysis.



(a) Two honeybees intervals_histogram



(b) Three honeybees intervals_histogram

Figure 9. Honeybee tracking data

3.3. Correlation Analysis

To better understand the correlation within the time series data created, we perform a correlation analysis. To simplify the analysis process we present the correlation results for our sample test with 2 worker honeybees. For this analysis, we utilize two metrics of statistical correlation:

1. **The ϕ Coefficient (Mean Square Contingency Coefficient):** This coefficient measures the strength and direction of a linear relationship between two binary variables. The binary nature of this correlation metric means that it aligns perfectly with our binary time series data.

   This metric is defined by the following calculation:

$$\phi = \frac{n_{11}n_{00} - n_{10}n_{01}}{\sqrt{n_{1.}n_{0.}n_{.0}n_{.1}}}$$

   Where each variable is generated from the following table of binary random variables x and y:

	y = 1	y = 0	total
x = 1	n_{11}	n_{10}	$n_{1.}$
x = 0	n_{01}	n_{00}	$n_{0.}$
total	$n_{.1}$	$n_{.0}$	n

2. **The Jaccard Similarity Coefficient:** This coefficient measures the similarity and diversity between two sets by calculating the intersection of the sets, and divinding by the union. This metric is defined by the following equation given two sets A and B:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

The ϕ and Jaccard coefficients are defined within the range of 1 and -1. In this definition, 1 indicates a perfect positive association such that when one variable is 1, the other variable is also 1 for every variable in the set. Similarly, -1 indicates a perfect negative association such that when one variable is 1, the other variable is 0 for every variable in the set. Finally, 0 indicates no correlation between the two sets of variables. Calculating the ϕ and Jaccard coefficients where w_0 is the binary time series data generated from worker bee 0, and w_1 is the data form worker 1, we yield the following results:

$$\phi_{w_0,w_1} = 0.09756434$$

$$J(w_0, w_1) = 0.15450$$

These numbers show a positive correlation between the time series scenting data of worker 0 and 1, meaning that the raw scenting data shows a linear relationship between the two sets. Despite this result, there is usually a delay in the time that a bee starts scenting, and the time that another bee starts following that scent.

As a result of this behavior, we also perform a cross-correlation analysis, where we analyze the relationship between the time series at different time delays. We start at a delay of 1 frame (i.e. 1/30th of a second), and increase the delay by 1 frame until we get to a 300 frame delay (i.e. 10 seconds). We ensure to perform a positive time delay, where the time series of worker 1 is delayed, and a negative time delay, where the time series of worker 0 is delayed. Along with this, we remove the leftover portions on the ends of each time series so each set is of the same size. Performing this cross-correlation analysis yields the following results:

$$\text{Best positive offset} = 83 \text{ frames}$$

$$+\hat{\phi}_{w_0,w_1} = 0.2250229957206299$$

$$+\hat{J}(w_0, w_1) = 0.2299084435401831$$

$$\text{Best negative offset} = 31 \text{ frames}$$

$$-\hat{\phi}_{w_0,w_1} = 0.15488867376573087$$

$$-\hat{J}(w_0, w_1) = 0.15488867376573087$$

4. Conclusion

The best overall result was from a positive offset of 83 frames or 2.77 seconds. Since the correlation is so high for both metrics, we can infer that worker 1 was overall following the scent of worker 0 throughout most of the video (since a positive offset represents a delay in worker 1's scenting data, i.e. pushing the entire time series backwards in time). Upon a brief visual analysis of the footage, one can confirm that this is the case. A correlation this high also indicates that the scenting events of worker 0 often directly result in the scenting of worker 1. This important distinction means that the individual decision making process of the honeybees results in a somewhat predicable relationship between scenting and exploration (or not scenting). This confirms our proposed hypothesis.

5. Next Steps

We aim to run similar tests with a larger number of bees so we can compare our scenting data from a variety of trials in order to better understand the behavior of individuals in the presence of the swarm.

Along with this, we hope to improve the tracking algorithm to be more robust to variation, and require less human interaction yet still result in the same high tracking accuracies.

References

[1] Fernand Meyer, *Color image segmentation*, 1992 international conference on image processing and its applications, IET, 1992, pp. 303–306.

[2] Dieu My T Nguyen, Golnar Gharooni Fard, Ashley Atkins, Paul Bontempo, Michael L Iuzzolino, and Orit Peleg, *Honey bees find the shortest path: a collective flow-mediated approach*, Artificial Life and Robotics **28** (2023), no. 1, 1–7.

[3] Dieu My T Nguyen, Michael L Iuzzolino, Aaron Mankel, Katarzyna Bozek, Greg J Stephens, and Orit Peleg, *Flow-mediated olfactory communication in honeybee swarms*, Proceedings of the National Academy of Sciences **118** (2021), no. 13, e2011916118.

[4] Satoshi Suzuki et al., *Topological structural analysis of digitized binary images by border following*, Computer vision, graphics, and image processing **30** (1985), no. 1, 32–46.