UNIVERSITY OF CALIFORNIA, SAN DIEGO

Automated Bee Waggle Dance Detection

A thesis submitted in partial satisfaction of the
requirements for the degree
Master of Science

in

Computer Science

by

Tushar Bansal

Committee in charge:

Professor Yoav Freund, Chair
Professor Ryan Kastner
Professor James Nieh

2018

The thesis of Tushar Bansal is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2018

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I am grateful to my thesis advisor Prof. Yoav Freund for constantly guiding me and steering me in the right direction. Prof. Freund was very supportive from the beginning and always gave me the encouragement and freedom to try new ideas.

I would like to thank Prof. James Nieh for providing his expert advise on the topic whenever required. Prof. Nieh ensured full support towards the project throughout its duration. I would also like to thank Bahram Kheradmand and Dr. Heather Broccard-Bell for their constant guidance and efforts to obtain the data for the analysis.

In addition, I would like to thank Prof. Ryan Kastner for taking the time to review this thesis.

Finally, I would like thank Jane Storck for her constant support and guidance throughout the thesis filling process.

ABSTRACT OF THE THESIS

Automated Bee Waggle Dance Detection

by

Tushar Bansal

Master of Science in Computer Science

University of California, San Diego, 2018

Professor Yoav Freund, Chair

A major limitation on performing detailed behavioral analysis of honey bee colonies is that there is currently no efficient way to carry it out. Due to the time required in manually analyzing the data, the current approach and small sample sizes limit the statistical power of these analyses. An automated system can provide a breakthrough in the way this research is performed. Waggle dances are an important aspect of understanding the behavior of honey bees as it serves as a way to communicate among themselves. In this thesis, we develop an automated system using computer vision and learning techniques to solve two problems i) Single bee tracking and waggle detection and ii) Multiple bee waggle detection. Our approach shows that it is possible to train learning algorithms to detect when and where a waggle happens in the hive.

Chapter 1

Introduction

In 1967, von Frisch decoded how the honeybees communicated with each other and passed on information[vF67, BS05]. Among this was the communication using the waggle dance, which the forager bees perform to pass on the information about the food sources to other bees. There have been some important advances since von Frisch's original experiments[JRR05] and the pattern is still not completely understood. Though the fundamentals of the theory haven't changed significantly, there are many unanswered questions on how and when bees collect this data for waggle dances.

Understanding the impact of variables like multiple food sources, pesticides etc. has become an important problem in understanding the behaviour of bees. The bottleneck in performing this research is the manual process of annotating various aspects of waggle dances from the video. Currently there is no system available for annotating the aspects of a waggle dance automatically. Hundreds of hours could be spent on simply annotating the videos for any valuable research. Also, human errors and subtle difference in definitions of different research groups introduces more problems in generalizing the theories.

We intend to resolve this issue by developing a system that annotates various aspects of bee waggle dance with very minimal human intervention. We use machine learning and computer vision techniques to solve two major tasks i) Single bee tracking and waggle detection ii) Multiple bee waggle detection. Both of these problems are important to understand the behaviour of bees at a micro and macro level.

1.1 Problem Statement

1.1.1 Single bee tracking and waggle detection

The aim of the problem is to study different characteristics of a waggle dance and the way bees gather information and communicate with each other. If we can come up with a mechanism to track a bee in the hive without manual labeling, then we would be able to conduct multiple experiments decoding the communication among bees. Though a lot of progress has been made in decoding the waggle dances, we can infer more about the impact of various factors like predators, multiple food sources etc. on bees by looking at the changing communication patterns.

To understand the communication thoroughly, we need to track the location, speed, orientation, waggle movements/direction etc. of the bee. We conclude that all of these variables can be reduced to computing the location, orientation and waggle detection for the bee. From the data, we are given videos of a bee hive with a marked bee, and we have to track the above stated components for the marked bee.

1.1.2 Multiple bee waggle detection

We perform this study for understanding the impact of xenobiotics on pollination services. It is therefore essential that the true costs and benefits of using specific classes of pesticides in large-scale agriculture may be assessed. For this analysis, instead of looking at particular bees, we would like to analyze the impact at the colony level.

To study the impact at the colony level, we want to detect the location of waggles in a video. Ideally, the application will also be able to count how many waggle circuits are performed in each dance. This will facilitate the collection and analysis of long-term behavioral data, which is currently not possible. However, this analysis isn't covered in this thesis due to time constraints. The current problem poses a challenge in terms of the complexity and speed as earlier we were only dealing with a single bee, but now we have to detect waggle which could be anywhere in the frame. We need an approach that can detect waggles anywhere in the frame while being fast enough to do so at a sufficiently usable speed.

Chapter 2

Data

The data for this analysis is provided by Prof. Nieh's lab in the form of videos of bee hives which included waggle dances specific to the task we were targeting. Although very similar, the two problems required data with different characteristics for us to be able to perform the tasks we intended to. The videos were captured at 30 fps. The properties of the data for both tasks are discussed below:

2.1 Single bee tracking and waggle detection

The data for this problem is collected in the form of videos of bee hives that are particularly focused on a marked bee for which the tracking and detection is to be performed. Since, we are only concerned about tracking a single bee in the frame, the bee is marked with color at thorax at the artificial feeder used to feed the bees. The marked bee is then filmed when it gets back to the bee hive.

The marked bee can communicate with other bees, the location and quantity/quality of food in the feeder using multiple waggle dances. To standardize the process of tracking and waggle detection over these videos, we decided on few properties of an input video:

- **Camera:** The initial data collected had moving camera because the frame was focused on the marked bee and the bee performed waggle dances at different locations in a hive, thus requiring the camera to move along with it. However, with moving camera, it was more challenging to get the exact displacement of bee, so we decided to have a fixed camera with a zoomed-out frame to capture the waggle dances across the hive.

- **Bee Marking:** Due to the video frames now being zoomed-out, it was harder to capture the marked bees out of hundreds of continuously moving bees. To overcome this, we ensured that the thorax of the bees were marked with bright distinguishable colors, and no more than a single bee was marked with the same color. We do this to make sure that the tracker is keeping track of the same bee all the time as bees tend get close to each other, sometimes also overlapping each other.

- **Occlusion:** The initial videos used a color dropper to mark the bees who were performing waggle dance. However, due to marking being done at the time of video recording, many frames had the bee marker occluded by the dropper, which led to eventual loss of data. The new data collected had the bees marked at the feeding station instead, which substantially improved the quality of data.

2.2 Multiple bee waggle detection

For multiple bee waggle detection, the data was collected in form of videos with a full section of beehive in the view. Since, for this problem we are trying to detect all waggles in the frame, no bees were marked specifically. From the analysis in the previous problem, we figured out that it might be helpful to increase the frame rate of the videos to capture the transition of waggle dances better, so videos for this part of the analysis were captured at 120 fps.

Our final working data consisted of around 10,000 video frames with information on 32

Figure 2.1: The figure shows the screenshots from the ground truth video for this problem.

waggle dances. For each waggle dance, the data is gathered at four stages - waggle start (wS), waggle end (wE), return start (rS), return end (rE). The data point is gathered at the thorax of the bee at each stage. The annotations were done using the tracker software [tra].

Figure 2.2: The figure shows the screenshots from the ground truth videos. The screenshots are focused here on the bee performing the waggle.

Chapter 3

Single bee tracking and waggle detection

Since we are interested in tracking a single bee, it is important that we are able to segregate the bee marker in most of the frames. This is generally not a problem but due to the full frame being captured instead of the close up of bee (section 2.1), this can be a problem sometimes as the marker may get hidden by interference from other bees or the movement of the movement of the marked bee (like waggle dancing). Therefore, it is necessary to have a big enough marker with a bright color that can be easily identified by the algorithm in most frames. We discuss in depth later, the methods we employ to make this process more robust.

3.0.1 Approach

The aim to the project is to capture the useful information like trajectory, waggle time etc. for a bee. The problem essentially reduces down to finding three features for every frame of the video:

- Position of the bee in the frame.
- Orientation of the bee.
- Whether the bee is performing the waggle dance.

All the other key parameters can be represented as a function of above features. The problem has been worked on by [WF15], where the authors use similar techniques to track multiple bees in a frame. However, the problem they look into is much more complex on a much more advanced data. Due to lack of sufficient data, and the unavailability of the exact approach used by [WF15], we try to solve the problem using above listed parameters.

Position of the bee

Before performing the analysis, we ask the user for a manual input for two positions:

- Position of the marker on the head of the bee
- Position of the abdomen of the marked bee

These initial inputs give us an accurate estimate of the initial position which we can use to significantly improve our future predictions. It also tells us the initial orientation the bee. We plan to extract the marker color automatically from these initial points but we havent found a robust method yet, since the current marker is very small so initial input by the user is very noisy to capture the color.

With the marker color known (currently manual), we convert each frame of the video to hsv and search for the range of marker color values in the the frame. This could lead to selection of other pixels that have similar color as the marker but to ensure correct selection we choose the pixels which are closest to the marker position in the previous frame. If the new position is substantially far (¿10 pixels) we skip that frame and look for the closest marker position in the next frame.

Orientation of the bee

To retrieve the orientation of the bee, we first use the above stated algorithm to find the position of the thorax of the marked bee in the current frame. Since it is more noisy to work in a RBG frame, we first convert the image to grayscale and then threshold it to remove the background so we are only left with pixels corresponding to bees.

Now, since we have the thorax position of the marked bee, we fit a rectangle over the marked bee that encloses the marked bee completely. The length and width of the rectangle are fixed for a video and are obtained from the user giving the initial thorax and abdomen coordinates.

To get the enclosing rectangle, we create a rectangle on the new estimated marker position of the bee with the same orientation as the previous frame. After, this we rotate the rectangle along the marker position by +/- 15 degrees (stride of 5 degrees). Among these possible 7 orientations, the one that encloses maximum white pixels (we only have binary pixels after threshold on grayscale) gives the current orientation of the bee. For this algorithm to work, it is important that the marker is at the thorax of the bee and the bee doesn't make too many sharp turns ($>$15 degrees) in consecutive frames. It was observed that in cases where a sharp turn is observed, the algorithm generally corrects itself in next few frames as the net degree of the freedom increases (+15 degrees) with every frame.

Waggle Detection

This is the most crucial part the algorithm because there is no clear definition on what exactly classifies as a waggle. A shaking abdomen is what we try to capture in this part. Due to lack of large amounts of labeled data, we are restricted to use any learning based approaches which could have been very helpful given the ambiguity of a clear mathematical definition of the problem. We use two features to detect a waggle: 1) Sharpness of the bee 2) Orientation plot.

When we estimate the orientation by fitting a rectangle on the bee. We compute the sharpness of the image of the bee enclosed by the rectangle. Sharpness is measured by first applying the laplacian filter on the image and then taking the variance of the resulting image. The technique is from [JLPPFV00]. When the bee waggles the edges and the body of the bee will get blurred thus substantially decreasing the sharpness. The results from this technique gives very volatile results so we convolve (take a weighted average along [i-3...i+3]) the series to smoothen the results.

From the orientation plot shown below 3.1, we can clearly observe a pattern where the bee performs a waggle at the peaks and valleys of the plot, before changing its direction. For the first step, we compute all the peak and the valley points of the graph. Now since we know that the waggle happens at roughly the same angle, we traverse from these valley points and select nearby frames where the bee orientation is constant (the mean orientation from the peak to current is within 10 degrees). However, this is not a robust method as can be seen in 3.2. To make sure that a waggle happens at this point, we check the sharpness graph to see if the sharpness for frames was among the bottom 35% of the range of sharpness in the whole video. If sharpness was lower than this threshold, then we predict a waggle.

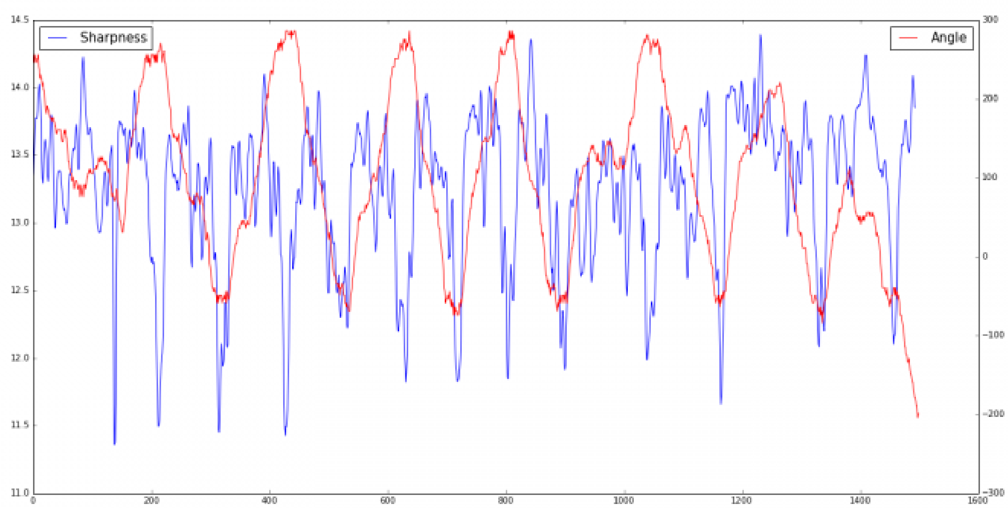

Figure 3.1: The figure shows the sharpness (blue) and the orientation (red) of the marked bee for a video, as learned by our algorithm. The bee performs waggle dances at the peaks and valleys (similar orientation) of the red plot. During the waggle, we observe a drop in the sharpness.

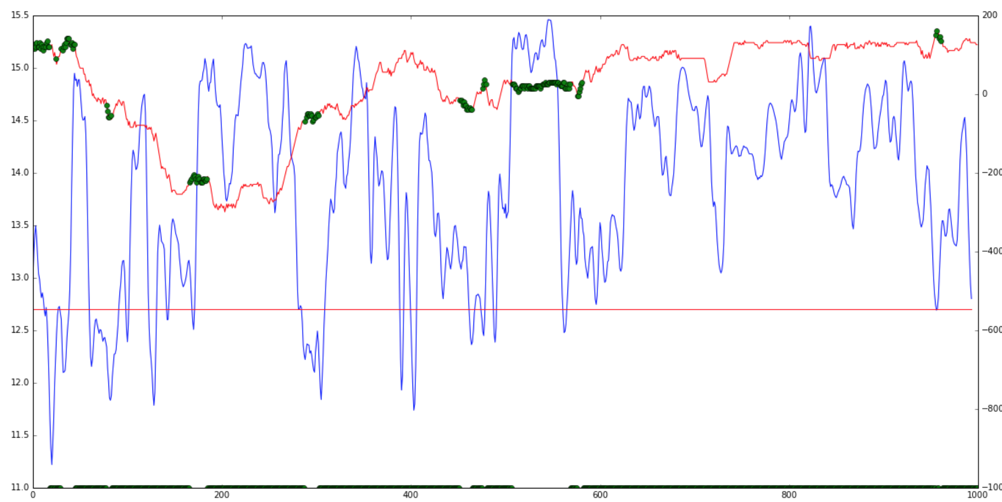

Figure 3.2: The figure shows an example where the current approach fails. As we can see, the sharpness (blue) and orientation (red) pattern is not evident as seen from the figure before. The marked bee here doesn't seem to perform waggle dance in the expected way. The green dots are the algorithms predictions for waggle dances.

Chapter 4

Multiple bee waggle detection

For this problem, we are not interested in tracking particular bees but capture only the number of waggles in a video. We don't have any marked bees in the videos, so the approach from the previous problem cannot be applied straightaway to this problem. However, since for this problem we have some labeled data, we can use machine learning approaches for learning a model. Our approach consists of two steps 1) Feature extraction 2) Learning a model for waggle detection.

4.1 Pre-processing

To look the whole video frame for waggles will be challenging for any learning task, and would constitute as a variant of object/action detection in a frame. However, it is easier to look for patterns in smaller windows in the frame. Instead of learning model over the entire frame, we train our models over windows of size 80x80 pixels. These windows are extracted from the frame with a stride of 20. The window size is selected so that it encloses a full bee with any orientation with it's thorax being at the center of the window.

4.2 Approach

4.2.1 Feature Extraction

An important property for a successful detection algorithm is rotation, translation and scale invariance as the bee performing the waggle could be anywhere in the frame with any orientation. Currently, our algorithm doesn't account for scale invariance, since all the data obtained was from same hive and same distance between camera and hive providing similar scale. However, our approach is invariant to translation and rotation. Since we are working in the small window space with a small stride, the approach will be able to handle translations. For introducing rotational invariance, we have to train our model on features that are invariant to rotation.

Our features consist of two characteristics i) intensity distribution of the window ii) temporal aspects of a waggle. From the single bee tracking problem, we can infer that sharpness is a key factor in detecting a waggle. We tried using the same sharpness filter here too but the results were too noisy. However, instead of computing sharpness in a pre-defined way, we instead let the model to learn it. We capture features using the intensity histogram over the grayscale 80x80 windows. Note that intensity histograms are rotation invariant.

Capturing the temporal aspect in a window is important because the waggle movement in time is very different from other bee movements. To provide the model this information, for all the positives and negative examples, we compute the difference of the current window from the window at same location five frames before. The difference was then scaled to 0-255.

4.2.2 Algorithm

For training, we first obtain the positive and negative windows. We obtain the waggle start and end from the original data. For positive samples, we consider windows with same position as the original data but 5 frames ahead for waggle start and 5 frames before for waggle end. This is to make sure that the positive labels are positions where the waggles are actually happening and also to account for any human errors in labels.

As a learning approach, I use an ensemble of random forests and xgboost. Using each of separately also did well on detecting the waggles but also led to many false positives. To tackle this, we introduce an ensemble method which works well on reducing the false positives.

Chapter 5

Results

5.1 Single bee tracking and waggle detection

We tried our defined approach over a number of videos. Due to the lack of sufficient labelled data, we had to rely on the opinion of field experts to evaluate our method. Our approach generalizes well on the videos with a well defined waggle pattern i.e. when the marked bee performs the dance at a consistent orientation in a periodic way. The approach since not being learning based falls short when the waggle dances have more randomness involved.

Below shown are the prediction of waggles for the marked bee at two different frames:

Figure 5.1: Results corresponding to waggles in 2.1

5.2 Multiple bee waggle detection

The current results only predict whether there is a waggle in the small window or not. This problem is more challenging as we have to consider all possible locations of waggle instead of just a single bee. Even though we have labelled data for learning, the process introduces significant false positives. Using an ensemble of random forests and xgboost helps in reducing the false positives significantly. We tune the parameters using a validation set over the labelled data. Increasing the number of estimators in random forests did improve the results but at the cost of computation time.

The screenshot of results of waggle detection (green circles) over small windows in a video is shown in 5.2.

Figure 5.2: Results corresponding to waggles in 2.2

Chapter 6

Conclusion

For the single bee tracking and waggle detection, the performance was acceptable for performing large scale analysis on the communication patterns of honey bees which is motivation for the research. The cases with more randomness involved in the waggle dances are not very handled well by this method and thus pose a limitation. In future, it would be worthy to look into this direction for improvement. A possible solution to the problem could be to gather more labelled data for this problem which could help in moving to an end-to-end learning based technique.

For multiple bee waggle detection, the algorithm is robust in detecting the waggles in a frame but is falls short over several challenges i) too many false positives ii) computation time for detection. Using an ensemble method helps in reduction of false positives but the current method is not completely resistant to them. Since we are iterating from each small window of the frame, there is a bottleneck in this step. Another extension of the work would be to tie up detected waggles across different frames and classify them as a single waggle.

Appendix A

Link to data and results

A.1 Single bee tracking and waggle detection

- Data video 1
- Data Video 2
- Result Video 1
- Result Video 2

A.2 Multiple bee waggle detection

- Result Video 1

A.3 Code:

- Github link: https://github.com/bansaltushar92/waggle_dance

Bibliography

[BB05] L. Bourdev and J. Brandt. Robust object detection via soft cascade. 2:236–243 vol. 2, June 2005.

[BS05] Jacobus C. Biesmeijer and Thomas D. Seeley. The use of waggle dance information by honey bees throughout their foraging careers. *Behavioral Ecology and Sociobiology*, 59(1):133–142, 2005.

[JLPPFV00] J. Chamorro-Martinez J. L. Pech-Pacheco, G. Cristobal and J. Fernandez-Valdivia. Diatom autofocusing in brightfield microscopy: a comparative study. *Proceedings 15th International Conference on Pattern Recognition, ICPR-2000*, 3:314–317, 2000.

[JRR05] A. D. Smith-D. R. Reynolds R. Menzel J. R. Riley, U. Greggers. The flight paths of honeybees recruited by the waggle dance. *Nature*, 435, 2005.

[JSR04] Guo Jing, Chng Eng Siong, and D. Rajan. Foreground motion detection by difference-based spatial temporal entropy image. A:379–382 Vol. 1, 2004.

[tra] `https://github.com/bansaltushar92/waggle_dance`.

[vF67] K. von Frisch. The dance language and orientation of bees. *Harvard Univ. Press, Cambridge, Massachusetts*, 1967.

[VJ01] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. 1:I–511–I–518 vol.1, 2001.

[WF15] Couvillon Margaret Rojas Ral-Landgraf Tim Wario Fernando, Wild Benjamin. Automatic methods for long-term tracking and the detection and decoding of communication dances in honeybees. *Frontiers in Ecology and Evolution*, 3:103, 2015.

[WF17] Rojas R Landgraf T Wario F, Wild B. Automatic detection and decoding of honey bee waggle dances. *PLoS ONE*, 12, 2017.