Automatically Extracting Drosophila Courtship Behaviour Statistics from Video

J.A. Heward, T.C. Lukins, M.A. Dewar and J.D. Armstrong
School of Informatics, University of Edinburgh
[jheward|tlukins|mdewar1|jda]@inf.ed.ac.uk

Abstract

Many laboratories work with the fruit fly (Drosophila melanogaster) as the ideal model organism for performing genetic studies. The short life cycle, low maintenance, and controllable population size make it ideal for efficient analysis. However, dealing with the shear volume of assays in a consistent, robust and reliable manner is a major challenge. In this work we present an approach to improve and automate this process using Computer Vision based techniques. We show how this can be performed to extract relevant behavioural statistics which can then be compared favourably against expert human annotations.

1 Introduction

Laboratories engaged in genetic studies perform many hundreds of assays in order to link targeted mutations with phenotypical development and behaviour. These are often carried out by human observers who must watch and annotate hours of video footage. Such an approach is consequently tedious and prone to error - with no real consistency between observers, nor across labs. This in turn makes it harder to reconcile results and compare studies. A far better solution would be to automatically analyse incoming video data to determine and tag relevant information, and so improve both the throughput and standardise the output of the assays. This motivation has recently led to commercial systems [1], but still requires expert knowledge to tailor and set-up for specific behaviours.

The fruit fly (*Drosophila melanogaster*) is widely used as a model organism in part due to the advantages it has of a short life-cycle and ease of maintenance. Many studies of it focus of sex-specific behaviour [2] which forms a useful distinct subset of recognisable interactions. Quantifying the extent to which these occur forms the basis of a number of derived statistics. Of particular interest is the courtship index (CI) which is a measurement of the fraction of the observed trial during which time the male actively courts the female (or any other target). To a greater extent this is defined mostly by the orientation of the flies and the degree of following that occurs. It can include recognising other behaviours such as wing vibration, which may indicate the intensity of the event - but these subtleties can be very hard to discern. This is why CI based on simple relative positions is generally accepted and used.

Tracking such small and mobile objects as *Drosophila* presents many challenges. The scale of the problem can broadly be constrained by the nature and extent of the experimental space. For example [3] look to track the in-flight motion using stereo infra-red cameras in a highly controlled environment to successfully track the actual trajectory of flies in response to their visual landscape. At the other end of the spectrum, other studies investigate fly motion dynamics by relying on tethered flies, such as [4] who have looked at the high-speed possibilities (at 6000Hz) of analysing flight by using advanced Computer Vision techniques to track the individual wings beats.

Between these extremes lie the smaller type of assay in which the flies are generally not captured in-flight, but are instead captured closer up at effectively higher-resolution to enable unique features to be determined. Work by [5] looked at a complete system for resolving location and behaviours in the Mexican fruit fly (*Anastrepha ludens*). They used a smaller scale stereo system to first localise the flies (distances relative to other regions containing food, water, etc.) and to also extract a number of image based features based on eigen-analysis. The combined feature vector was then used with a hybrid data fusion and K-Nearest-Neighbour classifier to achieve reasonable (94%) accuracy at determining behaviours from the set oft: [resting, walking, eating, drinking, flying, ovipositing].

Related research in directly determining the CI of courting flies was presented in our earlier work [6]. The significance of that research was that multiple flies (and their interaction) were considered, which made the task considerably more difficult - in that the flies can crossover and occlude one another. Despite these issues, successful tracking using a bounding box based segmentation algorithm performed very well, as the flies could be clearly differentiated in most cases, with their general orientation calculated simply as the line from the centre of mass to the furthermost point on the outline. One of the major problems still remained in resolving which end - head or tail - was which (an issue directly tacked by our work). Further work that can can resolve a greater number of targets and interactions must rely on more advanced techniques for multiple tracking, such as the work performed on desert ants (*Aphaenogaster cockerelli*) by [7] using a sophisticated particle filtering approach.

2 Approach

The primary contribution of our work is to improve the robustness of automatic fly courtship analysis, by considering the additional properties associated with the flies' motion. The main problems to overcome with this are: *(i)* resolving the head-tail "flip" that can occur when considering an ambiguous extracted fly contour, and *(ii)* retaining the track on exactly the same fly throughout the sequence. Given successful tracking, we then classify the CI based on the relative features between the flies.

Building on earlier work (based on [8] and [6]) we first extract the outline contours of the two flies. This follows a fairly standard image processing pipeline of first masking a rectangular region of interest containing the flies, followed by a Gaussian 3×3 smoothing operation and mean thresholding to reveal candidate blobs. These are cleaned to remove speckling by a set of morphological erode and dilate operations, after which the remaining blobs are described by ellipse fitting in terms of major/minor axis length and area. The best blobs are are then filtered by considering their size and compactness (Figure 1).

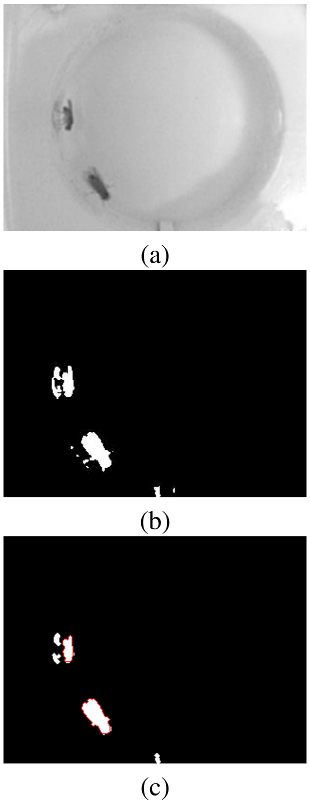


Figure 1. Segmentation by smoothing *(a)*, thresholding *(b)*, cleaning and filtering by size/compactness *(c)*.

Next, to resolve the orientation of these fly blobs we calculate the angle of the ellipse major axis to the x-axis. This is, however, ambiguous - in that it can have two values depending on which way the fly is pointing. It was decided that rather than seeking to discover which end was the head via image processing, to instead exploit the fact that the flies generally never move backwards and turn consistently (i.e. do not suddenly invert their direction). The algorithm we use to determine this orientation thus considers the motion and change in direction from the previous time-step. If this effectively results in a instantaneous reversal of the fly direction, or backward motion, then this registers as a head-tail "flip" to be corrected.

Additionally, in order to retain the correct tracking of individual flies we consider the circumstances that can occur between candidate blobs over time, especially in the case of occlusion where only a single candidate blob is found. Our algorithm identifies 3 possibilities:

- *One-to-one*: this is where a fly blob matches the previous frame with an exact correspondence. Note that the blob may contain more than two flies, in which case we should already know the number of targets contained in the blob from tracking previous merge operations. If there a single target in the blob then we simple extract the features directly. If there is more than one target contained in the blob we must resolve the features. This is performed by means of the algorithm presented in [6].
- *Merge*: this is where N blobs in the previous frame merge to become a single blob in the current frame. Here we record all the flies as contained within the blob and attempt to uniquely resolve their features as mentioned above.
- *Split*: this is where one blob in the previous frame splits in to N blobs in the current frame. The assignment of blobs to targets is carried out using the following algorithm: for each target extrapolate predicted current position using a linear extrapolation of the previous position velocity, and acceleration. Then assign each target to the closest current blob using the predicted position, with the constraint that each blob must contain at least one target. In addition to this measure we look at the length of the period of occlusion that has occurred for an target (i.e. the number of frames between a merge and split involving a particular target). If this is within some threshold then we assume that the flies have run over the top of each other (as is common) and we allocate the targets using the position and velocity when the occlusion began.

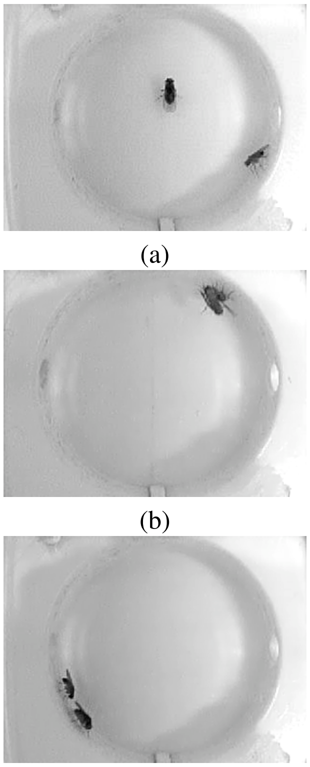
Finally, the features most representative of courting can be chosen based on courtship ethograms given in the literature (see i.e. [9]). The most representative features were found to be the distance between the two flies and the orientation of the male with respect to the female. By building these relative features from the tracking of individual flies, we are able to automatically classify the CI using decision-tree based supervised learning techniques.

3 Experimentation

Our data was captured from a grayscale camera capturing at 320×240 $pixels$ at $10 f.p.s.$ The male and female flies were loaded into a $2.5cm$ diameter clear plastic arena ($0.5cm$ in height) with a movable wall to separate them. At the start of the experiment the wall was removed and the flies allowed to interact for 5 minute from which we then extract a 1 minute example (600 frames in length). In total we performed 9 trials (with an additional trial created by rotating one of the sequences by 180 degrees). These same trial videos were also compiled together to form a single 10 minute video that was manually annotated by 10 expert observers for the CI. Example frames are shown in Figure 2.

The basic tracking solution was able to successfully segment and determine the location of candidate fly blobs. However, as expected it confused orientation in 54% of the extracted flies. By employing the orientation algorithm we were then able to correct this down to 10%. Additional manual correction was then necessary to resolve these final "flips" and to correct for the times where the labelling between male and female were switched due to occlusion (this happened for 29% of the frames in the corrected data).

From the 10 expert observers we derive a single set of CI annotations by discarding annotations where less than 3 observers agree. Using these aggregate annotations, we then train the decision tree using data from 7 minutes of the video and test against the remaining 2 minutes (excluding the rotated video), achieving 83% accuracy against the aggregated annotations. This is well within the inter-expert variability of the annotations. A benefit of using the decision tree as a classifier is that the inferred rules are readable, so can also be directly validated by the domain experts.



(a)

(b)

(c)

Figure 2. Example frames showing separate flies *(a)* **occluded flies** *(b)* **and courting flies** *(c)*.

4 Conclusion

In this work we have presented Computer Vision approach that attempts to automate and determine the courtship index between pairings of *Drosophila*. We show how it is possible to generate reasonably reliable statistics for the CI. However, while this improves the current state-of-the-art, but does not yet completely resolve the deceptively simplistic tasks of orientation and occlusion.

For future work we are to consider more robust statistical means of tracking over occlusion and in retaining a fix on the head of the fly. Furthermore, assessment of the performance of the fly tracking in relation to the annotations of the experts raises a number interesting of open issues regarding the consistency and variability of correctly marked-up data.

References

[1] L.P.P.J. Noldus, A.J. Spink, and R.A.J. Tegelenbosch. Ethovision: A versatile video tracking system for automation of behavioural experiments. *Behavior Research Methods, Instruments, & Computers*, 33(3):398–414, 2001.

[2] K.M.C. O'Dell. The voyeurs' guide to drosophila melanogaster courtship. *Behavioural Processes*, 64(2), 2003.

[3] L.F. Tammero and M.H. Dickinson. The influence of visual landscape on the free flight behaviour of fruit fly drosophila melanogaster. *The Journal of Experimental Biology*, 205(327–343), 2002.

[4] C.F. Graetzel, S.N. Fry, and B.J. Nelson. A 6000 hz computer vision system for real-time wing beat analysis. In *The First IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics.*, pages 278–283, 2006.

[5] S. Hernandez-Rodriguez, L. Altamirano-Robles, J.R. Carey, and P. Liedo. Automatization of the mexican fruit fly activity recognition process using 3d and gray-level features. In *Proc. 2nd IASTED International Conference on Advances in Computer Science and Technology*, pages 118–123, 2006.

[6] J.D. Armstrong, D.A. Baker, J.A. Heward, and T.C. Lukins. Sex, flies and no videotape. In *Proceedings of Measuring Behavior: 5th International Conference on Methods and Techniques in Behavioral Research*, 2005.

[7] Z. Khan, T. Balch, and F. Dellaert. MCMC-based particle filtering for tracking a variable number of interacting targets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(11):1805–1918, 2005.

[8] L.M. Fuentes and S.A. Velastin. People tracking in surveillance applications. In *2nd IEEE International Workshop on Performance Evaluation on Tracking and Surveillance*, 2001.

[9] M. Reif, K. E. Linsemair, and M. Heisenberg. Evolutionary significance of courtship conditioning in drosophila melanogaster. *Animal Behaviour*, 63:143–155, 2002.