Sex, Flies and no Videotape.

J.D. Armstrong, D.A. Baker, J.A. Heward and T.C. Lukins
School of Informatics, University of Edinburgh, Edinburgh

Abstract

The characterization and quantification of animal behavior is one of the most resource intensive activities in modern biological science. One of the best known examples is the courtship ritual of the small fly *Drosophila melanogaster* which has been discussed in the academic literature for almost 100 years. It is an exemplary behavior for the functional dissection of the nervous system as it is a natural behavior that requires the animal to integrate and learn from multiple sensory modalities. Object tracking software is extremely good at tracking single moving objects and describing their behavior particularly where previous examples have been well characterized. However, tracking multiple, interacting objects is more challenging. Here we describe the application of a bounding box segmentation algorithm that works well with Drosophila courtship behavior. We show that once segmented, just a few simple parameters are sufficient to classify the behavior in real-time, requiring no off-line storage or post processing.

Keywords
Drosophila courtship mating tracking

1 Introduction

Drosophila melanogaster males (and although less obviously, also the females) display a complex courtship ritual that uses a range of cues from multiple sensory modalities, is naively instinctive but then subsequently modified by experience. The behavior itself was first reviewed in 1915 [7] who cited what were presumably the earliest studies in the field by Barrows in 1907, Lutz in 1911 and Payne in 1911. We are now close to 100 years since these first publications in the field and interest has increased each year. There are now established laboratory methods to disentangle many of the interlinked sensory cues resulting in a very naturally embedded assay for perception of sensory stimuli and behavioral adaptation (for review see [5]). When combined with the genetic toolbox used by Drosophila researchers to mutate genes or modify cellular physiology in identified neurons (for review see [1]), we have one of the most sophisticated assays for mapping molecules to naturally occurring behavior.

However, there is a drawback - it is expensive, almost prohibitively so. In the 100 years or so of research into courtship behavior, the methods employed in laboratories to quantify the behavior itself have remained essentially the same. Courtship activity is affected by time of day therefore video recording has been employed to capture datasets in parallel which are later analyzed in serial by expert observers sometimes with the use of ethnographic annotation software packages. Otherwise the methods are more or less identical (e.g. [3]).

There are a number of behavioral events that can be recorded but most studies focus on three measures. The first, and most general measure is the Courtship Index (CI) and is the combined percentage of time the male fly under investigation spends exhibiting any recognizable courtship behavior towards a target fly or object during the observation period. These activities include orientation of the body, following, tapping with the legs, licking, wing extension or attempted copulation. The second measure sometimes reported is the Sex Appeal Parameter (SAP). This is the percentage of the observation time that wig extension is observed. Finally, latency is the time taken from introduction to the first interaction between the two animals.

Thus, for medium throughput genetic screens, assays that segregate a group of animals based on their behavior are most commonly employed. As insects move through the apparatus their behavior influences choices that alters their exit or final location. Perhaps the best known example of these is the choice point of the olfactory test [8] where flies have a choice of moving towards one odor or another. After they have chosen, flies are simply collected and counted. More complex examples are the mazes used to measure phototaxis or gravitaxis (for review see [2]) where the flies move sequentially through a series of T shaped choice points that are orientated with respect to light or gravitational stimuli and finally exit into a series of collection tubes and are simply counted. Such techniques allow for large numbers of flies to be tested but lack any analysis of individual behavior.

For mammals, almost all behavioral experiments require detailed quantitative and qualitative analysis of the behavior of individuals. The emphasis for automation has been on rodent behavior with many methods employed for tracking and describing behavior from video footage. Here, simple methods for the movement of objects in two dimensions can deal with many of the problems of occlusion and fitting active contours to the animal can be used to describe in high detail and with high precision the body movements over time. When combined these techniques can describe accurately the movement of a rat through a maze and identify specific behaviors.

The analysis and annotation of human behavior has also attracted much recent research within the Computer Vision community. Particularly in determining potential dangerous and anti-social activity as recorded in real-time on CCTV footage. This is primarily a two-stage process. Firstly reliably tracking and differentiating multiple people must be robustly handled [4]. Secondly, on the basis of good segmentation, analysis and statistical modeling of the relative interactions between people can occur, leading to a classification of the type of interaction [6,9].

Clearly an automated, real-time, behavioral tracking system for Drosophila would be extremely useful. Beyond the obvious benefit to anyone who has spent a considerable proportion of their graduate studies scoring video footage of insect courtship there are a number of other very important benefits. Observers will differ in their opinions of exactly when an insect is courting. Over long periods, observer fatigue is reported. Collecting very detailed information about courtship activity such as a fine breakdown of all the component behaviors, the velocities of the animals involved is possible but extremely time consuming. We have no empirical evidence for many of these issues thus cannot evaluate their importance but they are widely acknowledged by experts in the field.

Here we describe the development of a solution directed at all of the problems described for courtship behavior tests in Drosophila. Our aim is to produce a measure for courtship activity that is not observer dependant, will scale to medium throughput activities (i.e. 1000's of tests per day with a small team) and will also collect some of the technical information such as velocity that is difficult to do with manual annotation of video footage. We designed our test to demonstrate whether we could measure CI in a variety of samples and how well this correlated to that of human observers.

1 Methods

1.1 Video Capture

We chose to use a networked video camera (Axis 205 www.axis.com) as it supports a range of resolutions and frame rates all of which can be controlled by sending a simple URL call from a web browser or application. The camera then sends back a video file in MJPEG format. No specialist lighting or optics were used. We used incident lighting in the observation room and moved the apparatus around to limit any shadows or high contrast reflections. Unless stated otherwise, the resolution used was 320x240 pixels, captured at a rate of 10 frames per second.

1.1 Drosophila courtship

We focused our attention on normal behavior. Thus the flies described here are all Canton-S wild-type. To obtain a range of CIs we used mated and virgin target female flies in combination with naïve and experienced males (e.g. experienced males will not court mated females thus giving a low CI).

Flies were raised on standard cornmeal medium at 25'C on a 12 hour light:dark cycle. All courtship recordings were done in the morning and within 4 hours of lights on.

We used cylindrical plexiglass observation chambers, 10mm in diameter and 8mm deep. The chamber allowed for enforced separation of the insects prior to recording. For details refer to [5]. Flies were aspirated into the chambers without anesthesia and given 15 minutes to relax after transfer. The divider between the male and the target fly was then removed and activity recorded for 5 minutes.

1.1 Computation

Clearly any algorithm can be performed in real-time given sufficient computational power. We used a range of operating systems and hardware all of which was purchased between 2003 and 2004. Given that some of our development computers are multiple CPU architectures we reconfirmed all real-time tests using single CPU computers. 1) Apple 12" Powerbook 867 MHz G4 running Mac OS 10.3 with 640MB Ram. 2) Dell Optiplex GX270 2.6GHz P4 running Windows XP Pro with 512GB Ram.

1.1 Expert annotation

We designed a small JAVA annotation program that replayed the video and recorded keyboard activity to a log file. We recruited several experts to observe video footage and indicate via pressing keys when courtship activity started and stopped during the 5 minute period and thus a CI. For a small number of samples we also recorded wing extension thus obtaining an SAP score.

1 Results

1.1 Image Processing and Thresholding

The grayscale video footage we obtained did not require background subtraction since there was significant intensity difference between the fly and the background to identify it. Adaptive thresholding would have suited this task but was found to be too computationally intensive for live video work. Instead, a pseudo-adaptive thresholding procedure was implemented that split frames into smaller non-overlapping regions before mean thresholding each. We found a significant speed improvement with no loss of sensitivity for the flies (data not shown).

1.1 Segmentation

We searched the image space for blobs of approximate fly shapes and return their centers and longest angles to give a rough position and longitudinal body axis (flies can be approximated to ovals). Identifying flies within the binary image is achieved by filtering the list of blobs using one or more of the following criteria; 1) Position within frame, 2) Area of the blob.

1.1 Tracking

Flies move in 3D and are equally likely to be walking on the ceiling of the observation chamber as they are on the floor. Hence occlusion of flies directly on top of each other is common. As a result, methods for dealing with occluded blobs often applied to rodents such as erosion, water-fill etc did not perform well. We tested the bounding box algorithm of [5] and found it performed exceptionally well with the footage we obtained (see figure 1 and supplementary video file). For details of this, please see the accompanying short paper by Heward et al in these proccedings.

Using out default settings for video capture, tracking mistakes are extremely rare with approximately one error per 3000 frames.

1.1 Descriptive analysis

The most distinctive features of the fly courtship ritual (see video for an example) are following behavior, the relative angles of body orientation relative to the other fly and the extension of a single wing during the courtship song. We decided to focus our attention on relative body orientation and following behavior.

We found the relative body orientation and distance between the flies to be very predictive of the expert annotation. Just on the distribution of relative angle and relative distance on a frame by frame basis the majority of courting and non courting frames can be distinguished (figure 2).

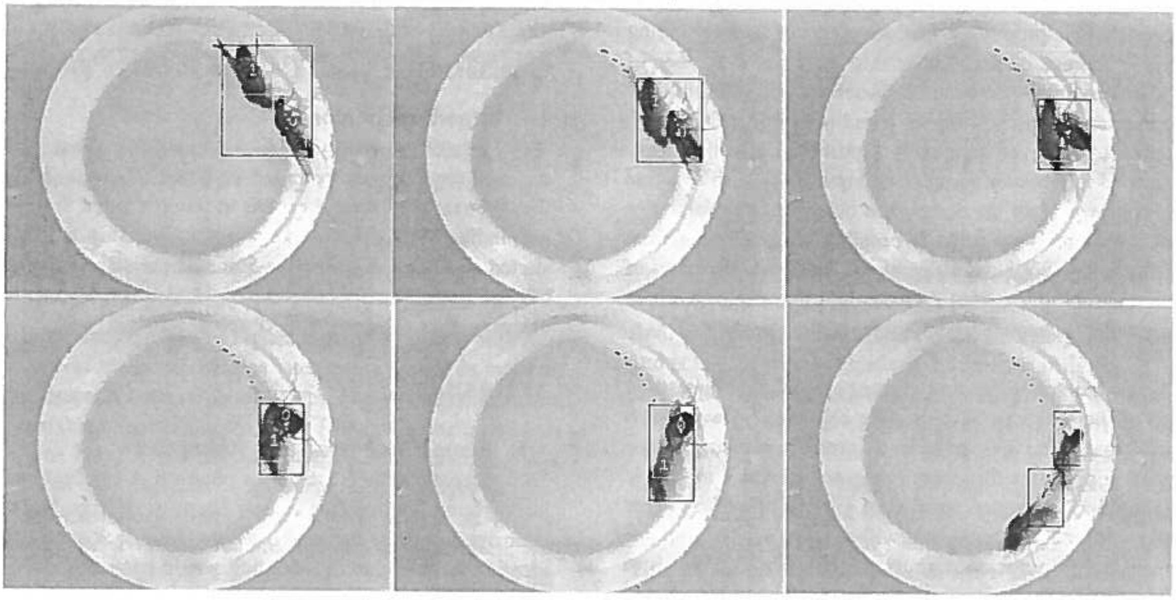

Figure 1 Six successive frames with bounding boxes following the flies. In each frame, the previous boundary is in cyan and the current boundary in dark blue. When the boxes merge (frames 1-5) the position of each fly is estimated using K-means (K=2).

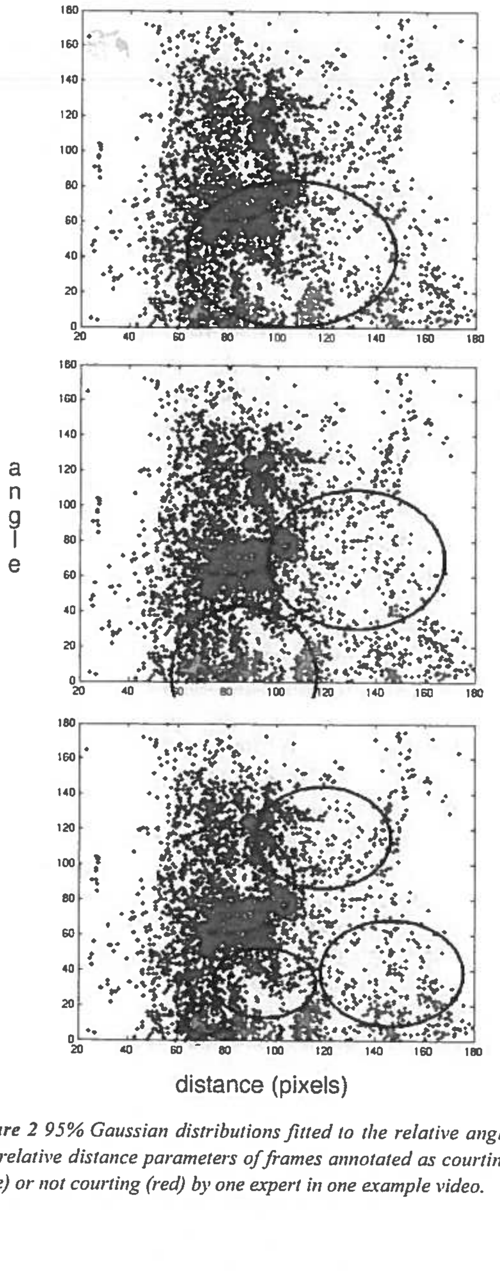

Figure 2 95% Gaussian distributions fitted to the relative angle and relative distance parameters of frames annotated as courting (blue) or not courting (red) by one expert in one example video.

1.1 Automatic Classification

We implemented and tested a Genetic Algorithm and K-means Clustering approach to extract predictive rules using the expertly annotated footage. The learning algorithms were given the X, Y distances, the relative distance and the relative angle between the longest axes for each blob, trained on two sets of expertly annotated datasets leaving residual datasets aside for validation. The expert annotation was offset by up to one second to compensate for slight time lags when the experts react to changes in behavior. For the GA, parameters from the previous nine frames were also available for rule design.

The CI values returned using the learned rules had correlations with the two expert datasets used of 96% and 98%. This is in line with the observed variability between our experts. Further validation is on-going.

1.6 Performance

The computational complexity increases with the number of pixels, the frame rate and with the number of objects to be segmented in the arena (all in a linear manner). Preliminary experiments suggest that our existing packages will handle 20-30 frames per second and three fly – ménage a trios - type experiments (figure 3) on standard desktop systems.

The tracker normally operates in one of three modes. The first is a live capture mode where the algorithm aims to match the speed of the live video feed. If insufficient compute or network resources are available then frames will be dropped and this is logged. In the second mode, the tracker works on previously recorded MJPEG files and will process them as fast as resources permit. In the third mode we can replay the video file at any specific rate and observe the tracker in operation.

The supplementary video file is in the third of those modes, replay frame rate was set to 10 frames per second (same as the original recording). This was recorded on a 867Mhz G4 based Apple laptop using SnapzPro2 (www.ambrosia.com) demonstrating that this system can easily handle both the tracking and video capture tasks in real time. The specific example was chosen as the flies are very actively courting resulting in a large number of partial occlusion events, highlighting the value of the bounding box algorithm.

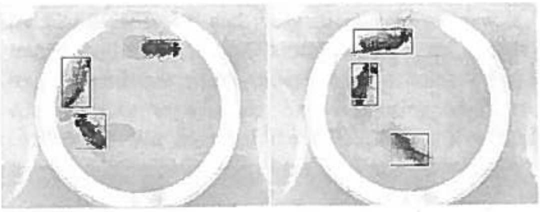

Figure 3 *The algorithms scale to multiple flies, e.g. ménage a trios type experiments.*

2 Conclusions

We have demonstrated that we can segment and process the video data at a rate suitable for live analysis of *Drosophila* courtship behavior. The application of the bounding box tracking algorithm has proven extremely useful in dealing with the occlusion problem. Moreover both the Genetic Algorithm and K-means classifier systems result in simple rule sets that perform within the accuracy bounds of experts. When all are combined the system can capture and process in real time to produce a reliable CI thus freeing research staff to focus on much finer details.

Acknowledgements

We thank Jean-Christophe Billeter, Tony Dornan, Stephen Goodwin, Kevin O'Dell and Jo Young for their help as expert annotators and system testers. This work was supported by grant no. S18944 from the BBSRC (JDA).

References

1. Armstrong J.D. and Goodwin S.F. (2004) Studying neuronal function using the Drosophila genetic system In Davies RW and Morris B (Eds) *Molecular Biology of the Neuron* Oxford University Press
2. Beckingham K.M., Texada M.J., Baker D.A., Munjal R. and Armstrong J.D. (2006) Genetics of Graviperception in Animals. *Advances in Genetics.* in press.
3. Ejima, A., Smith, B.P., Lucas, C., Levine, J.D. and Griffith, L.C. (2005) Sequential learning of pheromonal cues modulates memory consolidation in trainer-specific associative courtship conditioning. *Curr Biol* **15**, 194-206.
4. Fuentes, L.M. and Velastin, S.A. (2001) People tracking in surveillance applications in 2nd IEEE International Workshop on Performance Evaluation on Tracking and Surveillance, PETS 2001 (Kauai, Hawaii-USA, 2001).
5. O'Dell, K.M. (2003) The voyeurs' guide to Drosophila melanogaster courtship. *Behav Processes* **64**, 211-223.
6. Oliver, N.M., Rosario, B. and Pentland, A.P. (2000). A Bayesian Computer Vision System for Modelling Human interactions. IEE Transactions on Pattern Analysis and Machine Intelligence, vol.22
7. Sturtevant, A.H., (1915). Experiments on sex recognition and the problems of sexual selection in Drosophila. *Anim. Behav.* **5**, 351–366.
8. Tully T. and Quinn W.G. (1985). Classical conditioning and retention in normal and mutant Drosophila melanogaster. *J Comp Physiol [A]* **157**, 263-277.
9. Van Vu T., Brémond F. and Thonnat, M. (2003). Automatic Video Interpretation: A Recognition Algorithm for Temporal Scenarios Based on Precompiled Scenario Models. Proceedings of ICVS