flyTracker: real-time analysis of insect courtship

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

Abstract
We have developed a software package designed to automate the analysis of the courtship ritual of male Drosophila flies. Our system comprises three linked packages that (a) track objects in live video feeds or from files; (b) replays video files with and without animation/simulations of the tracked log files and allows expert annotation to be added and; (c) which learns classification rules from expert training data and applies it to new datasets. In this paper we focus on the tracking package.

Keywords
Tracking, bounding box, occlusion, Drosophila.

1 Introduction
The small fly *Drosophila melanogaster*, often referred to as the fruit fly, and still used in high-school biology genetics classes has been studied by biologists for close to 100 years. It is one of the key model organisms for studying biological processes including behavior and neuronal function. The courtship ritual is one of the commonly used behavioral assays for the organisms yet despite intense interest it is still quantitatively measured by visual observation by trained staff. An overview of the biological motivation and summary of the technical advances is presented in the accompanying manuscript by the same authors [1]. In this short paper we focus on the technical challenges and solutions we considered in building the tracking system.

Our requirements analysis identified a number of essential and desirable features. The key ones were that the system should: (1) be modular and extendable. (2) be platform independent. (3) run from live video feed or video files. (4) be able to visualize all processes.

We chose JAVA 1.4 as a development language since it would enable all of our key objectives. We had available, and tested all components on, a wide range of platforms from multi-cpu linux clusters to single processor laptops and workstations.

2 Materials and methods
The tracking module is the most important of the modules. To simplify the task we chose a single camera platform, the AXIS 205 net camera (www.axis.com). These small cameras have built in web servers that control the camera settings and return single snapshot images, live video feeds or movie files in response to a http style URL. We initially recorded several courtship encounters on the system under different pixel resolutions and frame rates to see what the minimum specifications would be for a human observer. All settings were for grayscale video footage taken at 320 by 240 pixels and 10 frames per second.

We then collected a series of test examples that were 5 minutes each and where a varying amount of courtship activity was noted during casual observation.

3 Tracking

3.1 Initial Processing
The task for the tracker component is to convert each frame of video into a log file that describes the identity, the number, position and angle of orientation of each fly sized blobs located in the video. It does require a basic parameter – either the expected number of flies or the expected size range of flies (given one the other can be estimated from the video).

Although all footage we used was in grayscale, we identified at an early stage that color is often used to identify flies (e.g. through genetic determinants of eye color). Therefore all processing was performed in RGB color space.

3.2 Filtering
To make the blob detection simpler, some initial image smoothing was required. We investigated the use of Gaussian Convolution, Sobel Filters and Average Convolution. In terms of both speed and image quality we found the average convolution methods to be the most appropriate.

3.3 Thresholding
Next we considered a number of thresholding methods to identify pixels in each frame that we should consider for segmentation. We evaluated mean thresholding, adaptive thresholding, histogram threshold detection, background removal and pseudo adaptive thresholding (see below). Adaptive thresholding and histogram threshold detection gave the best results but were expensive computationally. Background removal was implemented as an option but was not required for the task of removing ill defined shadows etc. Mean and pseudo adaptive thresholding worked well and were computationally cheap. We chose the later since it was better and handling uneven illumination. We also included a simple erosion operator to split barely touching objects at this stage.

In pseudo adaptive thresholding, the frame is split into a small number of non overlapping sub frames and each of these is mean thresholded. This results in a system that is resilient to small changes in light intensity over the image but is much faster to compute (around 2ms on our systems compared to 20ms for adaptive thresholding). Best results are when the image is split into between 9 and 25 sub frames.

3.4 Blob Detection
We considered a model based approach but decided that it would be too slow and inflexible. Moreover, one of our longer-term applications involves the analysis of behavior in a microgravity environment where new emergent behaviors are possible. Thus a predefined model would be inappropriate. We adopted a labeled feature approach which worked well with the small blob size processing an entire frame in approximately 5 ms. We then fit a bounding box to the blobs (with some simple constraints for maximum and minimum size of a fly sized blob).

3.5 Blob Tracking

To track the flies we implemented the Bounding Box method of [2]. We added an extra rule that is invoked when there are unmatched blobs in consecutive frames (i.e. the fly moves fast enough to move beyond its previous bounding box in a single frame). In this situation we resort to a best-fit algorithm on the positional data of the unmatched blobs. To obtain size, positional, and orientation data the area, centre of mass and the vector between the centre and point furthest from the centre are calculated.

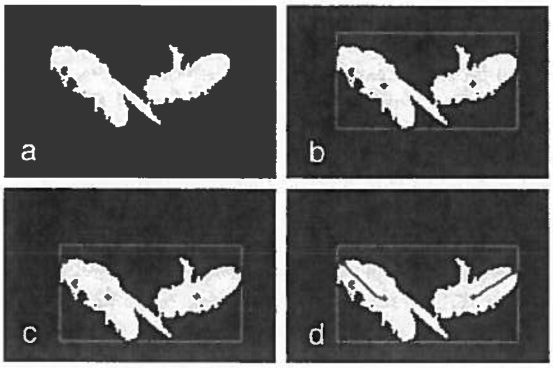

Figure 1 Resolution of occluding flies: (a) Two touching flies within a bounding box. (b) K-means clustering (K=2) finds the centers of the two flies (green blobs). (c) Locate the point furthest away from the centers (magenta). (d) The angle between these points is taken as the orientation.

When flies occlude we attempt to maintain the size, position, orientation and identity information. First, K-means clustering is performed within the bounding box (K=2). Next, the means obtained are matched on a best-fit basis with the positions of the corresponding blobs from the previous frame to maintain the identity and position. The size is retained from the previous frame. Calculating the furthest point from each mean within the bounding box is used to generate the orientation (figure 1). Combined, these methods handle the merging and splitting of more than two objects.

4 Evaluation

The tracking package generates a text log file. The log contains a list of blobs in each frame, their identity, area, X and Y coordinates and their angles of orientation. The change in position between each frame is usually less than 20 pixels which is feasible for the resolution, frame rate and fly size used. Viewing the footage confirms that the flies move about half a body length per frame on average. Occasionally the movement is higher, around 40 to 60 pixels per frame. This is likely due to periods of footage where the male is chasing the female around the container as the movement here is very rapid or jumps in the video footage. Using the methods described here we observed just a single label error in 3000 frames of footage and further testing is on-going.

5 Conclusions

Instead of trying to deal with occlusion by using image based separation techniques we demonstrate that it is possible to deal with occlusion at a higher level. The methods described here reduce the computational time required sufficiently for the tracking of multiple fast moving and often occluding objects to be possible on common PC hardware. With respect to our primary aims, we have used the system to track two interacting flies in real time from a live video feed.

This work was supported by grant no. S18944 from the BBSRC (JDA).

References:

1. Armstrong, J.D., Baker, D.A, Heward, J.A. and Lukins, T.C. (2005) Sex, Flies and no Videotape. *(this volume).*
2. 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).