

# Video Analysis Algorithms for Automated Categorization of Fly Behaviors

Md. Alimoor Reza<sup>1</sup>, Jeffrey Marker, Siddhita Mhatre<sup>2</sup>, Aleister Saunders<sup>2</sup>,  
Daniel Marendra<sup>2</sup>, and David Breen<sup>1,\*</sup>

<sup>1</sup> Department of Computer Science

<sup>2</sup> Department of Biology

Drexel University, Philadelphia, PA 19104, USA

david@cs.drexel.edu

**Abstract.** The fruit fly, *Drosophila melanogaster*, is a well established model organism used to study the mechanisms of both learning and memory *in vivo*. This paper presents video analysis algorithms that generate data that may be used to categorize fly behaviors. The algorithms aim to replace and improve a labor-intensive, subjective evaluation process with one that is automated, consistent and reproducible; thus allowing for robust, high-throughput analysis of large quantities of video data. The method includes tracking the flies, computing geometric measures, constructing feature vectors, and grouping the specimens using clustering techniques. We also generated a Computed Courtship Index (CCI), a computational equivalent of the existing Courtship Index (CI). The results demonstrate that our automated analysis provides a numerical scoring of fly behavior that is similar to the scoring produced by human observers. They also show that we are able to automatically differentiate between normal and defective flies via analysis of their videotaped movements.

## 1 Introduction

The fruit fly, *Drosophila melanogaster*, is a well-established model organism used to study the mechanisms of both learning and memory. The ability to study learning and memory behavior in *Drosophila* has significantly increased our understanding of the underlying molecular mechanisms of this behavior, and allows for the rational design of therapeutics in diseases that affect cognition. The existing human-observer-based techniques used to assess this behavior in flies are powerful. However, they are time-consuming, tedious, and can be subjective. An automated method based on video analysis of fly movements can replace and improve a labor-intensive, possibly inconsistent evaluation process. Automating the process will provide a reliable and reproducible analysis of fruit fly courtship behavior. Additionally, it also enables quantification of many aspects of the behaviors that are not easily perceived by a human observer. Moreover, the

---

\* Corresponding author.

automation promises the possibility of robust, high-throughput analysis of substantial quantities of video data, which would be useful for large-scale genetic and drug screens, and disease modeling.

In addition to the extensive genetic tools available for the analysis of fly neural circuitry associated with learning and memory [1,2], both adult and larval flies exhibit a number of behaviors that can be altered with training [3]. One of these behaviors that is well established in the field is courtship conditioning. Courting behavior by male flies in *Drosophila* follows a linear, stereotyped, and well documented set of behaviors [4], and this behavior is modified by previous sexual experience [5]. Courtship conditioning is a form of associative learning in *Drosophila*, where male courtship behavior is modified by exposure to a previously mated female that is unreceptive to courting [5,6]. Thus, after 1 hour of courting a mated female, males suppress their courtship behavior even towards subsequent receptive virgin females for 1-3 hours [5,7]. This courtship suppression is measured by the Courtship Index (CI), which is calculated by dividing the total amount of time each male fly spends courting by the total duration of a testing period [5,6]. CI is the standard metric used to assess learning and memory in courtship suppression analysis.

A computational approach to fly behavior quantification and characterization has been developed based on the analysis of videos of courting fruit flies [8]. The approach includes identifying individual flies in the video, quantifying their size (which is correlated with their gender), and tracking their motion, which also involves computing the flies' head directions. Geometric measures are then computed, for example distance between flies, relative orientation, velocities, contact time between the flies, and the time when one fly looks at another. This data is computed for numerous experimental videos and produces high-dimensional feature vectors that represent the behavior of the flies. We formulated a computational equivalent (Computational Courtship Index) of the existing CI, based on the feature vector values, and compared it with CI. Clustering techniques, e.g., k-means clustering, are then applied to the feature vectors in order to computationally group the specimens based on their courtship behavior. Our results show that we are able to reproduce CI values and automatically differentiate between normal and memory/learning defective flies using only the data derived from our video analysis.

## 2 Previous Work

Significant research has been conducted on tracking objects (cars, humans, animals, and insects) in video frames [9]. Isard et al. developed a real-time tracking system, BraMBLe (A Bayesian framework for multiple blob tracker) that combines blob tracking and particle filtering approaches to develop a Bayesian multiple-blob tracker [10]. Khan et al. identify objects entering and leaving a tracking arena and keep track of the trail of the targets using particle filter based approaches [11]. Branson et al. address tracking multiple targets (i.e., three mice) even with occlusion [12]. Tao et al. present a CONDENSATION-like

sampling algorithm [13] for tracking multiple objects which supports addition, deletion, and occlusion of objects [14]. Comaniciu et al. focus on tracking based on target representation and localization [15].

Several studies have been done on fly behavior modeling and analysis. Dankert et al. describe an automated social behavior (aggression and courtship) measurement of *Drosophila* [16] using a vision-based tracking system. Though the system has a courtship analysis component, it has not been used to assess courtship suppression in learning and memory. Branson et al. present a method for high-throughput computer-vision-based tracking that quantifies individual and social behaviors of fruit flies [17] and showed that an ethogram-based behavior quantifications can differentiate between gender and genotype. Valente et al. developed a system based on spatial and temporal probabilistic measures for analyzing the trajectory of *Drosophila melanogaster* in a circular open field arena [18].

Though many aspects of fly behaviors have been previously studied [16,17,18], to the best of our knowledge, no work has been done to computationally quantify and categorize the learning and memory capabilities of individual fly specimens by analyzing videos of their courtship behavior. Additionally, the technical contributions of our work include: 1) automated fly tracking, 2) determination of fly head direction, 3) formulation of a computational equivalent of the existing Courtship Index, and 4) automated grouping of fly specimens.

### 3 Computational Pipeline

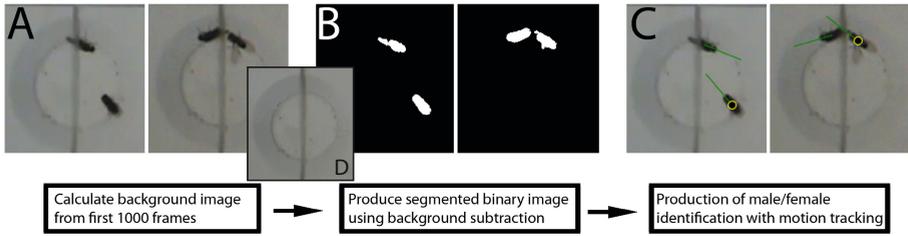
We developed a computational pipeline to analyze videos of courting fruit flies in order to characterize the memory and learning behaviors of the flies. A sample video of 10-minute duration is processed in the pipeline. Images are extracted from the video and processed by the pipeline components: segmentation, filtering, fly tracking, and geometric measures computation. A feature vector is generated for each video segment using the computed geometric measures. The feature vectors are used to generate a Computed Courtship Index (CCI). A clustering method (k-means clustering) is used to then group the samples according to their behaviors as defined by the feature vectors.

#### 3.1 Extract Chamber Images from the Video

A fly pair is video-taped inside a circular chamber. An input video may contain 3 to 6 of these chambers. Individual frames are extracted from the video at a rate of 30 frames-per-second. The extracted images are cropped to isolate each chamber. The rectangular coordinates surrounding a chamber are manually specified in the first frame of a segment. Each chamber is isolated using the same coordinates from the subsequent frames. Two example frames are presented in Figure 1 A.

#### 3.2 Segmentation

A binary image is produced from each cropped image, where a group of white pixels represents a fly object and black pixels represent the background. This stage utilizes background detection, image subtraction and thresholding.



**Fig. 1.** Steps in the fly identification and tracking process

**Background Detection.** This component processes a set of  $N$  images to find the most frequent color that appears in each pixel, producing a single background image. In general, using the first 1000 cropped images was found to be sufficient for accurately detecting the background. One such background image is presented in Figure 1 D.

**Image Subtraction and Thresholding.** Each of the approximately 18,000 frames in a video segment are subtracted from the background image. The resulting images are thresholded to compute a corresponding segmented binary image. If the intensity difference between the background and foreground pixel exceeds a threshold, the binary pixel is set to white. Empirically, a threshold value of 0.15 (in a 0 to 1 range) was found to be sufficient for identifying fly pixels in the sample videos.

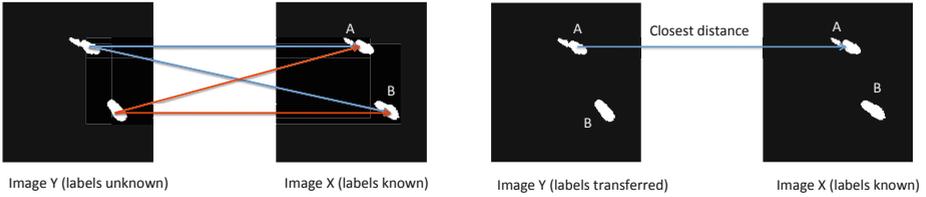
### 3.3 Filtering

The filtering stage removes noise and fills voids from the binary segmented images, and identifies the white-region representation(s) of the fly pair, as seen in Figure 1 B.

**Filling Voids.** The first step in the filtering stage fills voids inside any collections of white pixels. The voids inside the white regions are removed using a flood-filling process, starting from the top left corner pixel. The 4-connected black region is found. The white regions are defined as the complement of the flood-filled black region; thus removing interior voids in the white regions.

**Fly Object Identification.** We assume that the two flies will produce two sets of white pixel regions that have approximately the same area. If the areas of the two largest white regions differ greatly we assume that the flies are touching each other and form one large white pixel region. Collections of white pixels are identified in the void-filled binary images. The area of each white region is calculated. One or two white regions are selected in each image based on the following conditions:

1. If the ratio between the second largest white region and largest white region is less than or equal to a given value, e.g.,  $\frac{1}{15}$ , the largest white region



**Fig. 2.** Propagating fly object labels from Image X to Image Y

is identified as a single object. This single object represents the two flies in contact with each other. In our experience using this ratio consistently allowed us to identify when flies are in contact.

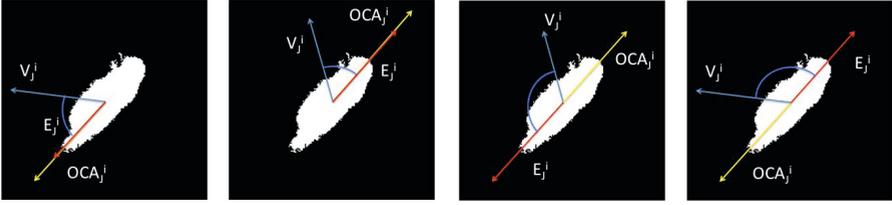
2. If the current image only has one white region, it is selected as the fly object, and indicates that once again the flies are together.
3. Otherwise, the largest two white pixel regions are selected to represent the two flies.

### 3.4 Fly Object Tracking

The two fly objects are tracked for all of the frames of a video segment. The first step of the process is to divide the 18,000 filtered, binary images into smaller subsequences. A *subsequence* is a contiguous series of images where each image contains two fly objects (white regions). A *subsequence* with fewer than 15 (i.e., half of the extraction frame-rate) images is not processed. Let us define a subsequence starting from image number  $M$  to image number  $N$  (in increasing order of extraction). The images at the start and end of this subsequence are denoted by  $I_M$  and  $I_N$  respectively. The Euclidean distance between the centroids of the two white regions in each image is computed. The frame with the maximum inter-centroid distance is identified as a starting image for this subsequence. Let us assume that the starting image is frame  $S$  and the image is denoted by  $I_S$ . Each white region in image  $S$  is labeled and these labels are propagated from image  $S$  both forwards and backwards in the subsequence. A minimum distance criterion is used for establishing the label correspondences between images.

A four-way distance calculation is performed between the two fly objects in the source frame and the two fly objects in an adjacent frame. The minimum of the four distances determines one of the label correspondences, i.e. the label of the object in a source image that has the smallest Euclidean distance is transferred to the closest fly object in the adjacent image. The second fly object in the adjacent image is given the label of the other fly object in the source image, as seen in Figure 2. The labels for the first and second white regions of image  $I_S$  are initialized to be  $A$  and  $B$  respectively. The fly objects in the immediately adjacent images,  $I_{S-1}$  and  $I_{S+1}$ , are assigned labels from the fly objects in  $I_S$ . This labeling information is then propagated out to the remaining images in the subsequence.

Once the white regions are tracked and labeled ( $A$  or  $B$ ) in all of the images in a subsequence, we attempted to also identify the gender of the fly objects by



**Fig. 3.** Definition of a fly object’s Oriented Central Axis  $OCA_J^i$  from its velocity vector  $V_J^i$  and eigen vector  $E_J^i$

assuming that the larger object is female. Our experiments showed that this assumption was only valid approximately 80% of the time. Since this identification was not deemed sufficiently reliable, we did not use this size/gender information during subsequent calculations.

### 3.5 Fly Object Head Direction Determination

Several geometric quantities are needed for the head direction calculation. The **Velocity Vector** of  $fly^i$  in frame J is denoted by  $V_J^i$ . The centroid of  $fly^i$  in frame J is denoted by  $C_J^i$ , a 2-dimensional point. The velocity vector  $V_J^i$  of frame J for  $fly^i$  is calculated using the following equation,

$$V_J^i = \frac{C_{J+1}^i - C_J^i}{J + 1 - J} \text{ pixels/frame.}$$

The denominator in the previous equation equals 1, since we compute the velocity vector from two consecutive frames. The velocity vectors  $\{V_M^i, V_{M+1}^i, V_N^i\}$  are calculated for all frames of a subsequence for  $fly^i$ .

The **Eigen Vector** of a frame J for  $fly^i$  is denoted by  $E_J^i$ .  $E_J^i$  is the eigen vector that corresponds to the largest eigen value of the covariance matrix generated from the locations of the white pixels representing  $fly^i$ .

The **Oriented Central Axis**  $OCA_J^i$  is the Eigen Vector  $E_J^i$  of  $fly^i$  pointing in the direction of the fly’s velocity vector  $V_J^i$ . It is determined from the sign of the dot product between  $V_J^i$  and  $E_J^i$  with the following equation,

$$OCA_J^i = \begin{cases} E_J^i & \text{if } (V_J^i \cdot E_J^i) > 0 \\ -E_J^i & \text{if } (V_J^i \cdot E_J^i) < 0 \\ 0 & \text{if } (V_J^i \cdot E_J^i) = 0 \end{cases} \quad (1)$$

$OCA_J^i$  is calculated for all images in the subsequence. The three quantities,  $V_J^i$ ,  $E_J^i$  and  $OCA_J^i$ , are presented in Figure 3 for a single fly object.

The **Head Direction Vector** of  $fly^i$  in frame J is denoted by  $HD_J^i$ .  $HD_J^i$  can be defined either as  $E_J^i$  or  $-E_J^i$ . The head direction vector of the starting frame,  $HD_S^i$ , in the subsequence is first calculated. Head direction vectors for the remaining frames in the subsequence are derived from  $HD_S^i$  utilizing the following calculations.

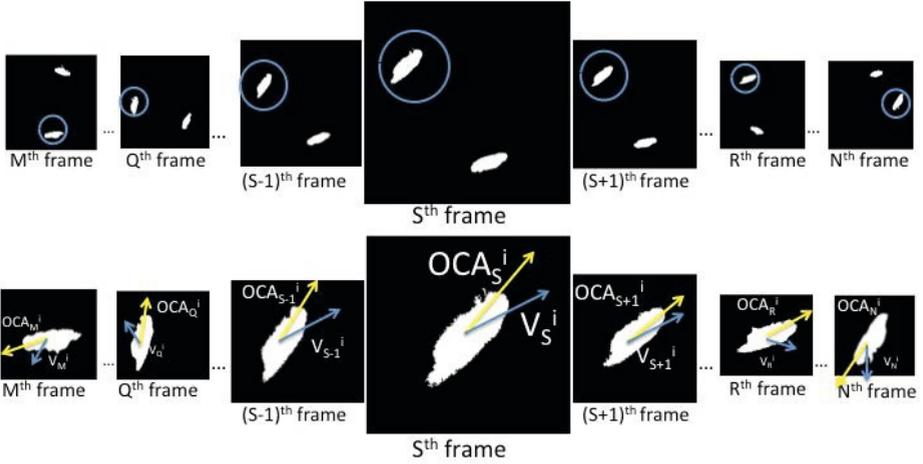


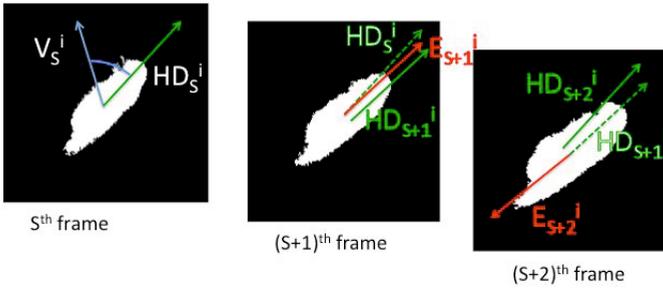
Fig. 4. Longest positive dot product sequence identification

**Longest Positive Dot Product Series.** In a courtship assay, a fly moves in the forward (i.e. head direction) for most of the time in the assay chamber. Given this observation, we search for the longest time period when a fly moves in a consistent direction. The computation of the longest series of consecutive positive dot products between the adjacent  $OCA_j^i$ s will give us the greatest number of frames where the angle between each of the adjacent pairs of  $OCA_j^i$ s is less than 90 degrees. In other words, the adjacent pairs of  $OCA_j^i$ s in the series are oriented in the same direction. The blue circles in Figure 4 identify the series of fly objects for which we are computing head directions. The bottom row of Figure 4 shows the vectors  $OCA_j^i$  (yellow arrows) and  $V_j^i$  (blue arrows) for the subsequence from frame M to frame N. The *longest positive dot product series* extends from frame Q to frame R. We compute the pairwise dot product between each of the adjacent  $OCA_j^i$ s and select the longest series of dot products in which all the products are positive. In this series (from frame Q to frame R) the adjacent pairs of  $OCA_j^i$ s are oriented in the same direction.

**Start Frame Head Direction.** Given the series of longest positive dot products of the  $OCA_j^i$ s in the subsequence, we choose the frame with the maximum speed in the series as a start frame. Let us denote this start frame to be frame number S. The head direction vector  $HD_S^i$  of  $fly^i$  in frame number S is defined as  $OCA_S^i$ , since it is the central axis vector pointing in the same direction as  $V_S^i$ .

**Propagate the Head Direction Vectors.** The head direction vector  $HD_X^i$  for  $fly^i$  in frame number X immediately adjacent to frame number Y, whose head direction vector  $HD_Y^i$  for  $fly^i$  is known, is computed by

$$HD_X^i = \begin{cases} E_X^i & \text{if } (HD_Y^i \cdot E_X^i) \geq 0 \\ -E_X^i & \text{Otherwise.} \end{cases}$$



**Fig. 5.** Head direction propagation from start frame S in the forward direction

The head direction vector  $HD_X^i$  is the Eigen vector of  $fly^i (E_X^i)$  that is in the same direction as  $HD_Y^i$ . Given the head direction vector  $HD_S^i$  for  $fly^i$  of start frame S, the head directions are propagated to the remaining frames in the subsequence. Figure 5 shows the head direction propagation in the forward direction from start frame S. Figure 1 C presents the centroids and head directions for two flies overlaid on the original processed image.

### 3.6 Compute Geometric Measures

Five geometric measures are computed for each video segment. They are 1) *percentage of frames when one fly is looking at the other*, 2) *percentage of frames when flies are together*, 3) *distribution of the distances between the fly centroids*, 4) *distribution of the head direction angles between the flies*, and 5) *distribution of the flies' speeds*. For the first measure the head direction vector of fly object A is intersected with the pixels of fly object B, and the head direction of fly object B is intersected with fly object A. The number of frames where these intersections are non-null are counted and divided by the number of frames of the video segment. For the second measure the number of frames that contain one fly object are divided by the number of frames of the video segment. For the third measure the Euclidean distance between fly object centroids is calculated in each frame and accumulated in a histogram. For the fourth measure the angle between fly object head directions is calculated in each frame and accumulated in a histogram. For the fifth measure  $V_S^i$  is calculated in each frame and its length is accumulated in a histogram.

### 3.7 Create Feature Vector

Four types of video clips of 10-minute duration are processed. They are: 1) *first 10 minutes*: first 10 minutes in the normal training period of a male fly, 2) *last 10 minutes*: last 10 minutes in the normal training period of a male fly,

3) *immediate recall 10 minutes*: immediate recall 10 minutes<sup>1</sup> of a normal-trained male fly, and 4) *sham 10 minutes*: immediate recall 10 minutes of a sham-trained<sup>2</sup> male fly. From the geometric measures an 8-dimensional feature vector that represents the behavior of the flies is computed for each video segment. The elements of the feature vector are: 1) *percentage of frames when one fly is looking at the other*, 2) *percentage of frames when flies are together*, 3,4) *mean and standard deviation of inter-centroid distances*, 5,6) *mean and standard deviation of head direction angles*, and 7,8) *mean and standard deviation of speeds multiplied by 100*. The last two numbers in the feature vector are multiplied by 100 so that all eight numbers have the same order of magnitude.

### 3.8 Generate Computed Courtship Index

A computational equivalent of the Courtship Index (CI), the Computed Courtship Index (CCI) is computed based on the values of the 8-dimensional feature vectors extracted from the  $N$  sample videos. The matrix  $\mathbf{A}$  is an  $N \times 8$  dimensional matrix, where each row is a data sample. The vector  $\mathbf{b}$  is a column vector whose elements are the manually-derived CI values for each of the samples. The vector  $\mathbf{x}$  contains the coefficients that define the linear relationship between the feature vectors and the CI values, and can be found by minimizing the standard Euclidean norm of the error,

$$\| \mathbf{Ax} - \mathbf{b} \|_2 .$$

Since our linear system is overdetermined, we use Singular Value Decomposition (SVD) to find the least squares solution of the minimization problem. The CCI for sample  $n$  may then be computed as:

$$CCI_n = PLA * x_1 + PFT * x_2 + MCD * x_3 + SD CD * x_4 + MHA * x_5 + SDHA * x_6 + MS * x_7 + SDS * x_8. \quad (2)$$

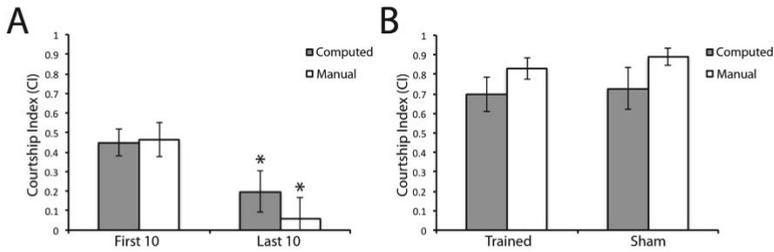
where the  $x_{ks}$  are the elements of the least squares solution vector, and the remaining variables are the components of the feature vector described in Section 3.7, e.g. PLA is *Percentage of frames when one fly is Looking At the other*.

### 3.9 Clustering

We also group the samples in terms of their learning and memory response based solely on extracted features. This is accomplished by applying k-means clustering [19] to the *learning* and *memory* feature vectors derived from the

<sup>1</sup> To obtain the immediate recall memory response of a male fly, it is transferred into a separate courtship chamber within 2 minutes after its training phase. In the new courtship chamber the male is assayed for 10 minutes with a virgin female unlike the training phase, in which the male courts a fertilized female.

<sup>2</sup> *Sham training* is a training phase 60 minutes in length where the male fly is not paired with a female fly.



**Fig. 6.** Comparison of average CI and CCI values for flies with good learning and memory ability. Error bars represent  $\pm$  SEM. \* indicates  $p < 0.05$ .

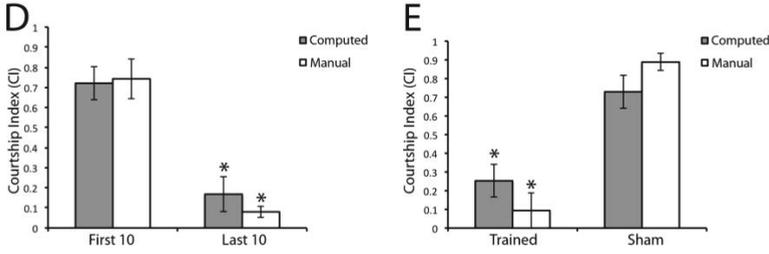
raw 8-dimensional feature vectors. The Euclidean norm is used during cluster analysis to measure the distance between two feature vectors. When clustering samples by learning ability the feature vector calculated from a sample's last 10 minutes of a one-hour training session is subtracted from the feature vector computed from the first 10 minutes. This is done because the difference between a fly's first and last 10-minute behaviors characterizes its ability to learn. This produces a *learning* feature vector for each fly. Similarly, when clustering samples by memory ability the average vector calculated from all of the feature vectors for the sham-trained flies is subtracted from the immediate recall feature vector for a trained fly. This produces a *memory* feature vector for each fly.

## 4 Results

Video sequences for 22 fly specimens, which were recorded with a Sony DCR-SR47 Handycam fitted with a Carl Zeiss Vario-Tessarare lens, were processed and analyzed.

### 4.1 Evaluation of the Computational Courtship Index

We evaluated the effectiveness of our video analysis algorithms to produce a Computational Courtship Index (CCI) by comparing the CCI values to traditional CI values. CCI values were generated with a 10-fold cross-validation process. In the first study, we selected 7 different flies that individually showed normal learning and memory by manual CI analysis (Figures 6, A, B white columns). These flies showed a significant decrease in their overall courtship behavior in the last ten minutes of a one-hour training session compared to the first 10 minutes, both in the CI (Figure 6 A, white columns) and CCI (Figure 6 A, gray columns) values. Trained flies showed a significant reduction in courtship behavior as compared to sham-trained flies, indicative of normal immediate-recall memory of training (Figure 6 B, white columns). The CCI values also show a significant decrease for the trained flies as compared to the sham-trained flies (Figure 6 B, gray columns). There is no significant difference between the manually calculated CIs and the CCIs produced via video analysis.



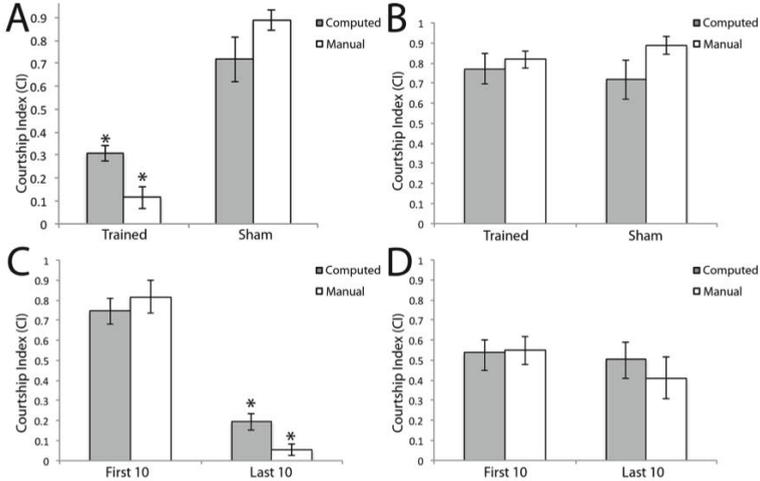
**Fig. 7.** Comparison of average CI and CCI values for flies with poor memory ability

We next determined whether the CCI could also detect flies with memory deficits. We selected 7 different flies that individually showed a successful learning response to training, but no memory of that training. In both cases, the manual and computed CIs showed a significant difference between the first and last 10 minutes of training (Figure 7 A), indicative of a successful learning response. However, for both the manual and computed CIs, there was no significant change in CIs from trained flies compared to sham controls (Figure 7 B), indicating that these flies have no memory of training. Again, there is no significant difference between the CIs calculated manually and the CCIs calculated via video analysis. Taken together, both sets of data suggest that the CCI is capable of determining successful learning and memory from those flies capable of performing this behavior robustly. Additionally the CCI is able to determine unsuccessful memory in the context of successful learning.

#### 4.2 Automatic Grouping via k-Means Clustering

We also categorized fly specimens by learning and memory capabilities based solely on the feature vectors derived from video analysis with no prior human intervention. We applied k-means clustering [19] to the *learning* and *memory* feature vectors described in Section 3.9. In this study, we increased our total fly specimen size to 22 with varying CIs for both learning and memory. Using the *memory* feature vectors k-means cluster analysis (for  $k = 2$ ) separated the fly specimens into two groups based on their memory capabilities. We evaluated the effectiveness of the analysis by computing average CI and CCI values for the two groups. The group with good memory is indicated by the higher difference in CI/CCI values (Figure 8 A) between trained flies and sham-trained flies. The poor memory group is indicated by the small difference in CI/CCI values (Figure 8 B) between trained flies and sham-trained flies.

We next performed this same k-means cluster analysis based on the learning capabilities of the flies, followed by a calculation of both the average CI and the CCI for each of these clustered groups. Again, it was clear that these two groups represent flies with good learning (Figure 8 C), and poor learning (Figure 8 D). Significant difference between the CI/CCI values between the first 10 minutes and last 10 minutes of the courtship suppression assay training phase indicates good learning. A small difference indicates poor or no learning.



**Fig. 8.** Automated grouping of fly specimens by memory (A (good) and B (poor)) and learning (C (good) and D (poor)) ability. Error bars represent  $\pm$  SEM. \* indicates  $p < 0.05$ .

This analysis does not take into account the genotypes of the flies, only learning and memory capabilities as captured by our feature vectors derived via video analysis. Our data suggest that the automated analysis is capable of successfully creating two groups of flies with differing learning and memory capabilities without human intervention.

## 5 Conclusion

We have presented a set of algorithms that may be applied to videos of courting fruit flies. The algorithms compute a feature vector for a video segment that captures the behavior of the flies and may be used to categorize the learning and memory capabilities of the specimens. The algorithms include identifying individual flies in the video, tracking their motion, and computing geometric measures. Our results show that we are able to reproduce conventional CI values and automatically differentiate between normal and memory/learning-defective flies using the data derived from our video analysis.

Future work includes improving the performance of the computations via parallelization, automatic image cropping, more robust segmentation, and multivariate non-linear regression analysis for the CCI computation. We anticipate that our analysis will be capable of distinguishing between more subtle differences in fly behavior. Analyzing larger samples of flies and clustering them into several groups may allow for a more detailed analysis of learning and memory in both wild type and mutant flies. Finally, our automated system can also allow for increased throughput of pharmacological screenings in *Drosophila* models of human disease, such as Alzheimer's disease.

**Acknowledgements.** We would like to thank the members of the Marena and Saunders labs for their helpful contributions to this paper. This work was supported by grants from the NIH, R01NS057295 (AJS), R21RR026074 (DRM).

## References

1. Wu, J., Luo, L.: A protocol for mosaic analysis with a repressible cell marker (marcm) in *Drosophila*. *Nat. Protoc.* 1, 2583–2589 (2006)
2. Yu, H., Chen, C., Shi, L., Huang, Y., Lee, T.: Twin-spot marcm to reveal the developmental origin and identity of neurons. *Nature Neuroscience* 12, 947–953 (2009)
3. Pitman, J., DasGupta, S., Krashes, M., Leung, B., Perrat, P., Waddell, S.: There are many ways to train a fly. *Fly* 3, 3 (2009)
4. Bastock, M., Manning, A.: The courtship of *Drosophila melanogaster*. *Behaviour* 8, 85–111 (1955)
5. Siegel, R., Hall, J.: Conditioned responses in courtship behavior of normal and mutant *Drosophila*. *Proceedings of the National Academy of Sciences* 76, 3430 (1979)
6. Siwicki, K., Riccio, P., Ladewski, L., Marcillac, F., Darteville, L., Cross, S., Ferveur, J.: The role of cuticular pheromones in courtship conditioning of *Drosophila* males. *Learning & Memory* 12, 636 (2005)
7. Kamyshev, N., Iliadi, K., Bragina, J.: *Drosophila* conditioned courtship: Two ways of testing memory. *Learning & Memory* 6, 1 (1999)
8. Reza, M.: Automated categorization of *Drosophila* learning and memory behaviors using video analysis. Master’s thesis. Drexel University, Philadelphia, PA (2011)
9. Maggio, E., Cavallaro, A.: *Video Tracking: Theory and Practice*. Wiley (2011)
10. Isard, M., MacCormick, J.: BraMBLe: A Bayesian multiple-blob tracker. In: *Proc. IEEE International Conference on Computer Vision*, vol. 2, pp. 34–41 (2001)
11. Khan, Z., Balch, T., Dellaert, F.: Mcmc-based particle filtering for tracking a variable number of interacting targets. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27, 1805–1918 (2005)
12. Branson, K., Rabaud, V., Belongie, S.: Three brown mice: See how they run. In: *IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, PETS* (2003)
13. Isard, M., Blake, A.: Condensation – conditional density propagation for visual tracking. *International Journal of Computer Vision* 29, 5–28 (1998)
14. Tao, H., Sawhney, H.S., Kumar, R.: A Sampling Algorithm for Tracking Multiple Objects. In: Triggs, B., Zisserman, A., Szeliski, R. (eds.) *ICCV-WS 1999. LNCS*, vol. 1883, pp. 53–68. Springer, Heidelberg (2000)
15. Comaniciu, D., Ramesh, V., Meer, P.: Kernel-based object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25, 564–577 (2003)
16. Dankert, H., Wang, L., Hoopfer, E., Anderson, D., Perona, P.: Automated monitoring and analysis of social behavior in *Drosophila*. *Nature Methods* 6, 297 (2009)
17. Branson, K., Robie, A., Bender, J., Perona, P., Dickinson, M.: High-throughput ethomics in large groups of *Drosophila*. *Nature Methods* 6, 451 (2009)
18. Valente, D., Golani, I., Mitra, P.: Analysis of the trajectory of *Drosophila melanogaster* in a circular open field arena. *PLoS One* 2, e1083 (2007)
19. Everitt, B., Landau, S., Leese, M., Stahl, D.: *Cluster Analysis*, 5th edn. John Wiley & Sons, Chichester (2011)