Automated Worm Tracking and Classification

Wei Geng, Pamela Cosman, Clare Huang
{wgeng,pcosman,s4huang}@code.ucsd.edu
Department of Electrical and Computer Engineering
University of California, San Diego
La Jolla, CA 92093-0407

William R. Schafer
wschafer@ucsd.edu
Department of Biology
University of California, San Diego
La Jolla, CA 92093-0449

Abstract— **The locomotion of *C. elegans* (a microscopic worm) provides valuable information about mutant genes and their effect on behavior. In order to investigate detailed movement and body posture characteristics of these living animals, advanced automated tracking algorithms are required. Here we describe a novel procedure of tracking an individual worms body part movement accurately by combining head and tail recognition with tracking. In addition, we describe a classification system to distinguish mutant types from each other. We demonstrate that its performance can be improved by incorporating new image features introduced by this tracking method.**

I. Introduction

The nematode *Caenorhabditis elegans* is widely used for studies of nervous system function and development. *C. elegans* is a free-living worm, approximately 1 μm long, that lives in the soil and eats bacteria. It has a simple nervous system with 302 neurons; the precise position, cell lineage, and synaptic connectivity of these neurons is known. Despite its anatomical simplicity, the *C. elegans* nervous system mediates diverse and intricate patterns of behavior.

C. elegans sense organs perceive and respond to a wide range of environmental conditions, including heavy/light touch, temperature, volatile odorants, osmotic and ionic strength, food, and other nematodes. Each sensory modality regulates many aspects of behavior, including mating, the rate and direction of movement, and the rates of feeding, egg-laying, and defecation. Because a particular neuron can be positively identified based on its position, it is possible to eliminate the function of an individual neuron or group of neurons through single cell laser ablation. Moreover, because of their short generation time, completely sequenced genome, and accessibility to germline transformation, these animals are highly amenable to molecular and classical genetics. Thus, in *C. elegans* it is relatively straightforward to evaluate the functions of particular neurons or gene products by characterizing the effects of mutations or neuronal ablations on the animal's behavior.

This approach to understanding nervous system function requires precise quantitative measurements of behavioral abnormalities seen in mutant and cell-ablated animals. Although gross behavioral defects can often be discriminated qualitatively by simple observation, precisely defining these differences can be challenging without quantitative measurements of parameters such as the curvature and amplitude of body bends. Furthermore, many important and significant behavioral abnormalities are reliably detected only using quantitative methods. For example, oviposition occurs on a long time scale that precludes evaluation by human observation [13], [15].

A variety of computer-driven systems for automated recording and/or analysis of *C. elegans* behavior have been described. Some systems have been designed to observe multiple animals at low magnification and track position over time [8], [7] (it is necessary to record at low magnification to keep all tracked animals in the field of view of the microscope). Such systems can measure large-scale behavioral features such as the rate and direction of movement and the frequency of reversals in direction. However, with low magnification, it is not possible to obtain more detailed information about body posture and morphology. To gain such information, we developed a system to follow an individual animal at high magnification [1]. A tracking program directs the movement of a motorized stage to maintain the worm in the center of the field of view. In this way, it is possible to follow the position of the animal over long time periods and comprehensively measure multiple features that define behavioral and morphological abnormalities of nematode mutants. With 94 such features, we classified representative mutant types using a binary decision tree algorithm (CART). Although this system performed well at distinguishing visibly different mutant phenotypes, it was less effective at distinguishing types with more subtle differences.

In this paper, we present a new automatic method for tracking worms and classifying the phenotypes based on morphological operations and on head and tail recognition. We extend our previous work [1], [9], [10] in several ways. We identify and track the head and tail separately, and extract features that characterize them. This enables us to successfully classify wild type and 15 mutant types (previous work only used wild type and 5 mutant types) using an improved classifier. Our approach can be divided into several stages. After video images acquisition, the images are segmented to isolate the worm body from the background and remove noise and undesired components. Next, the head and tail are recognized for entire video sequences. Feature extraction is applied, to extract useful information from the segmented objects and the head/tail locations. Finally, a Random Forests

0-7803-8104-1/03/$17.00 ©2003 IEEE

classifier operates on the extracted characteristics. Each of these steps will be presented in the next sections.

II. Data Acquisition and Segmentation

Routine culturing of C. elegans was performed as described [5]. All worms analyzed in these experiments were young adults; fourth-stage larvae were picked the evening before the experiment and tracked the following morning. Experimental animals were allowed to acclimate for 5 minutes before their behavior was analyzed. We used wild type worms and fifteen mutants (*goa-1*, *nic-1*, *egl-19*, *unc-2*, *unc-36*, *unc-29*, *unc-38*, *tph-1*, *unc-63*, *unc-43*, *dgk-1*, *dar-1*, *flp-1*, *cat-2*, *eat-4*).

C. elegans locomotion was tracked with a stereomicroscope mounted with a CCD video camera [1]. A computer-controlled tracker was used to maintain the worms in the center of the optical field of the stereomicroscope during observation. To record the locomotion of an animal, a frame was snapped every 0.5 second for at least five minutes. Among those image pixels with values less than or equal to the average value minus three times the standard deviation, the largest connected component was found. The image was then trimmed to the smallest axis-aligned rectangle that contained this component, and saved as eight-bit grayscale data. The dimensions of each image, and the coordinates of the upper left corner of the bounding box surrounding the image were also saved simultaneously as the references for the location of an animal in the tracker field at the corresponding time point when the images are snapped. The stereomicroscope was fixed to its largest magnification (50 X) during operation. Depending on the type and the posture of a worm, the number of pixels per trimmed image frame varied. The number of pixels per millimeter was fixed at 312.5 pixel/mm for all worms. A total of 1,596 video sequences of at least five minutes each were analyzed containing 100 videos of *goa-1*, *nic-1*, *egl-19*, *unc-36*, *unc-38*, *tph-1*, *unc-63*, *unc-43*, *dar-1*, *flp-1*, *cat-2*, *eat-4* each, 99 of *unc-2*, 98 of *unc-29*, and 99 of *dgk-1*.

To obtain the clean binary image, the background intensity level (b) of the grayscale image was found by taking the maximum of the values of the four corner points of the trimmed image (at least one corner points is always outside the worm body). Then a 5x5 moving window was scanned over the trimmed image, and the mean (m) and standard deviation (s) of the pixels inside the window were computed at every pixel position. If $m < 0.7b$ or $s > 0.3m$, the pixel was considered to be in the worm body (assigned value 1). To clean up the spots inside the worm body, a morphological closing operator (binary dilation followed by erosion) was applied. To remove unwanted isolated objects, the connected components were labeled, and the largest component was selected to guarantee that there will be only one object, the worm, in the binary image.

Following binarization, a morphological skeleton was obtained by applying a skeletonizing algorithm [14]. Redundant pixels on the skeleton were eliminated by thinning. To avoid branches on the ends of skeletons, the skeleton was first shrunk from all its end points simultaneously till only two end points were left. These represent the longest end-to-end path on the skeleton. A clean skeleton can be obtained by growing out these two remaining end points along the unpruned skeleton by repeating dilation.

III. Tracking and Head/Tail Recognition

Even though a simple tracking system was able to follow worm centroid movement, head and tail information were not extracted in our earlier work. Because of the highly deformable nature of the worm's body, many conventional image matching and tracking algorithms do not apply to this problem. Human observers use three spatial and temporal clues to recognize the head and tail sections. Even though the entire worm body could travel a large distance (in camera coordinates) between two consecutive frames taken 0.5 seconds apart, the head and tail locations relative to the body centroid (worm body coordinates) tend to change little [Fig. 1]. The other two clues are: the worm's tail area is darker than the head (having to do with fat distribution), and the head moves more frequently than the tail (having to do with foraging behaviors). The detailed procedure, illustrated in Fig. 2, is as follows:

1) The above segmentation procedure was applied to recorded grayscale images. For each video frame, the grayscale image and its corresponding binary image and skeleton were stored.

2) The two skeleton end points are potential head and tail locations. We assign the end points to two groups for each uninterrupted video segment according to the following rules: Let $endpt1x(t)$ denote the x coordinate of the end point in frame t that was assigned to group 1. Similar definitions hold for $endpt1y(t)$, $endpt2x(t)$ and $endpt2y(t)$. Now we use $endpt1(t) = [endpt1x(t), endpt1y(t)]$ to denote the vector of spatial coordinates for end point 1 in frame t. Let $endptA(t+1)$ and $endptB(t+1)$ denote the vectors of spatial coordinates for the two end points in frame $t+1$ that have not yet been assigned to group 1 or group 2. Let $(M, N) = \arg\min_{(m,n)} dist(endptm(t+1), endptn(t))$, $(m \subset \{A, B\}, n \subset \{1, 2\})$ Then $endptM(t+1)$ will be assigned to group N provided that $(\overline{M}, \overline{N}) \neq \arg\max_{(m,n)} dist(endptm(t+1), endptn(t))$, $(m \subset \{A, B\}, n \subset \{1, 2\})$. The condition statement avoids accidental flipping of head and tail locations. If the condition is not met, the current frame is marked as "undecided" and the grouping process restarts from the next frame to avoid potentially spreading errors.

3) To isolate the head and tail sections from the rest of the body, we identify two points on the skeleton that are at 1/3 skeleton-length away from each end point. We compute the best fit line to 9-pixel-long segments from the skeleton list surrounding the two identified pixels. The lines are then

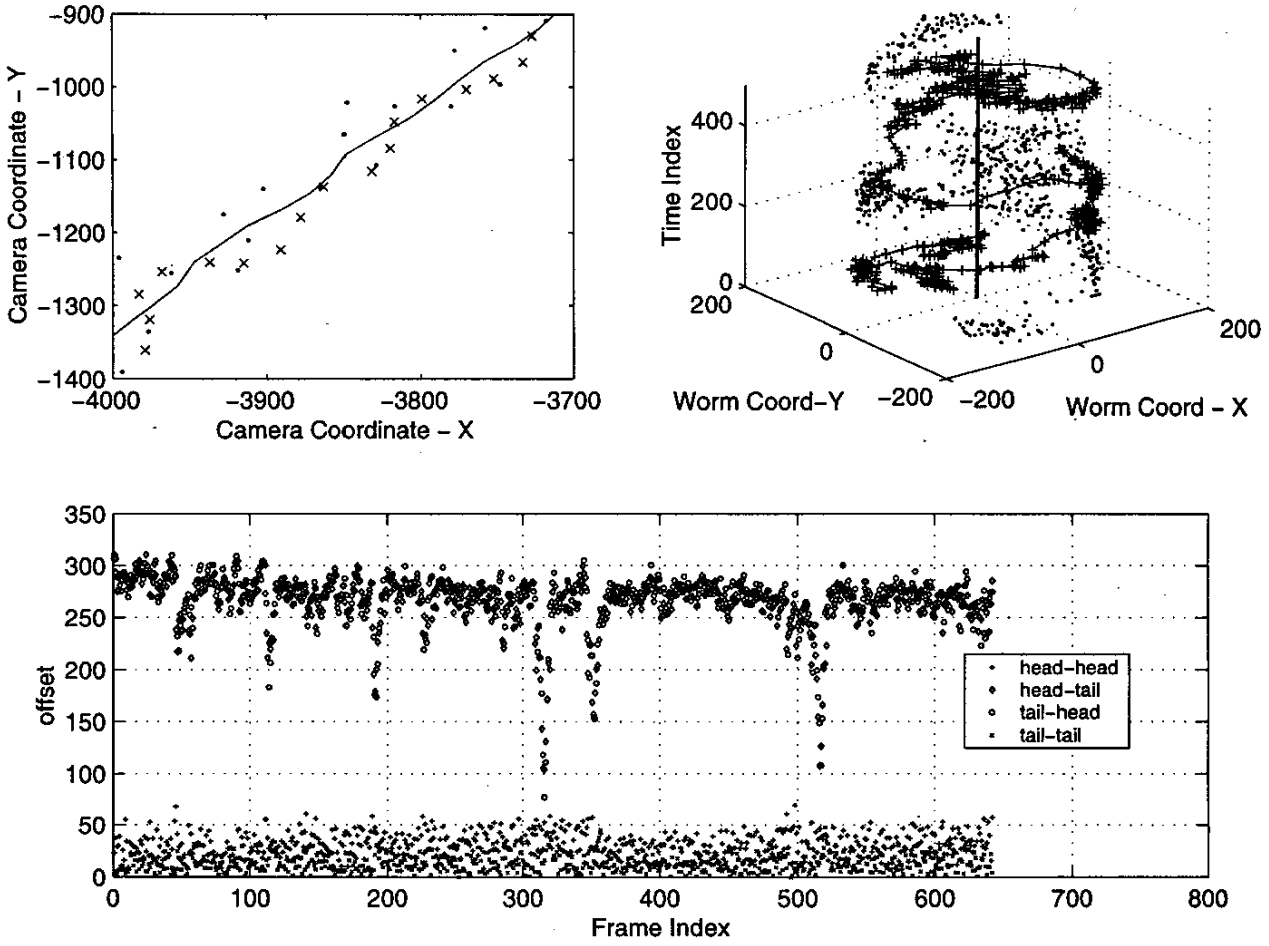

Fig. 1. Worm movement characteristics and their usage for tracking. (A) A portion of track in camera view. The solid line represents the worm body centroid movement. x and · represent the worm's tail and head location respectively, as they wiggle around the travel direction. (B) 3-D plot of head and tail movement in worm coordinates. The centroid movement is represented as the vertical line in the $(0, 0, t)$ location. The tail locations (+) are connected, showing the circular movement around the centroid. The head locations, marked by dots, tend to locate opposite the corresponding tail locations. (C) Bottom plot shows the location offset of heads and tails in worm coordinates for two consecutive frames. The head-head and tail-tail correspondences have smallest offsets of the four.

rotated by 90 degrees to get perpendicular lines. Lines that are +5 and 5 degrees off from the perpendicular are also generated. The line with the shortest distance traversing the binary image is chosen as the separation line between the head/tail and the rest of the body. The end sections are separated from the rest by deleting binary pixels along the separation lines.

4) Referring back to the grayscale image, we calculate the median brightness of the two end sections for each frame. The means of these values for group 1 and 2 are calculated for the segment. The group with higher average brightness is labeled as the head (if the brightness difference is $> 20\%$).

5) Mutant types with digestive abnormalities have smaller brightness differences between head and tail. For these (brightness difference $\leq 20\%$), a secondary decision rule compares the local movement distance for the two end points. The group with higher total movement distance is labeled as the head. The above procedure was applied independently for each video segment. Segments are separated by missing frames, failed segmentation, or undecided frames.

IV. Feature Extraction

The software for binarization, skeletonization, and feature extraction was coded either in C or MATLAB and implemented on UNIX machines. Some features (e.g., the area of the worm, that is, the number of pixels which make up the single binary object in the frame) could be computed on a single frame; these were computed for all 600 frames in the sequence. The average, maximum, and minimum value were

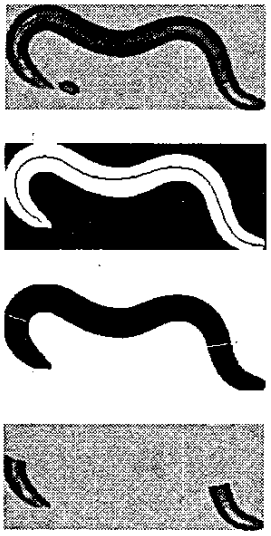

Fig. 2. Image processing and head/tail extraction procedure. (A) Original grayscale image. Notice there is an egg object nearby that needs to be removed by cleaning. (B) Binary image after segmentation and cleaning. The worm skeleton generated from the thinning and pruning process is superimposed on the binary image. The two skeleton end points are candidates for head and tail locations. (C) Two perpendicular lines (to skeleton line fitting) at 1/3 of skeleton location. Deleting the pixels along these separation lines divides the worm body into head, tail and middle sections. (D) Head and tail sections of the grayscale image.

computed for these 600 measurements. Some of the max/min values are outliers introduced by noise or errors during image capture and processing. To avoid using outliers, we used the 90th and 10th percentile values out of the population of 600 numbers. For the remainder of this paper, the terms max and min denote the 90th and 10th percentile values. Other features could not be extracted from a single frame, for example, the movement within 10 seconds (20 frames). Since there are approximately 600 frames total in a sequence, the movement over 20 frames could be computed 30 times if we take groups of 20 frames in a non-overlapping fashion, or it could be calculated 581 times using overlapping 20-frame intervals. We used the overlapping approach. As before, the average, max, and min were computed from this set of numbers.

The measured features included the minimum, maximum, and average values of the following: distance moved in 0.5, 5, 10, 15, 20, 25, 30 seconds and 5 min, number of reversals in 10, 20, 40, 60, 80, 100, 120 sec and 5 min, worm area, worm length, thickness at center and head/tail, ratio of thickness to length, fatness, eccentricity and lengths of major/minor axes of best-fit ellipse, height and width of minimum enclosing rectangle (MER), ratio of MER width and height, ratio of worm area to MER area, angle change rate, head/tail/center brightness, local head/tail/center movement relative to centroid, and head to tail angle. The area, angle change rate, and movement features were calculated for the head, tail, center, and entire worm body respectively. We now describe in detail how several of these features were extracted from the image data.

1) *Body length and area*: The worm length is the number of pixels on the skeleton. The head/center/tail area are measured by counting the binary pixels in these sections.

2) *Body thickness/width/fatness*: The worm thickness was measured at the center, head, and tail positions of the worm skeleton (the center position was the middle value of the skeleton pixel list; head/tail positions were 7 pixels away from head/tail end points identified by the tracking algorithm). The thickness measurement methods were as described in [1]. The center (head/tail) width was defined as the average width of the center (head/tail) section. Fatness is the ratio of worm area to length.

3) *Local movement*: Many features characterize global movement using the absolute distance traveled by the worm centroid over various fixed time intervals. We also measured the offset of the head relatvie to the centroid across the frames as an indication of the worm head's movement. This offset was defined as the movement of the head when the centroids were aligned from one frame to the next. The tail's movement was also measured.

4) *Angle change rate*: The angle change, an important feature for distinguishing different worm types, is defined as $R = \frac{\frac{1}{n-1}\sum_{i=1}^{n-1}\theta_i}{L}$, $\theta_i = arctan\frac{y_{i+2}-y_{i+1}}{x_{i+2}-x_{i+1}} - \frac{y_{i+1}-y_i}{x_{i+1}-x_i}$, $(x_i, y_i), (x_{i+1}, y_{i+1})...$ are consecutive points 5 pixels apart along the worm skeleton. A larger angle change rate means that a worm has sharper body bends. Angle change rates were also calculated separately for head, tail, and center regions.

5) *Brightness*: Variations in fat distribution and absorption of nutrients cause some mutant types to be more transparent than others. Transparency can be measured by the median pixel value of the head, center, tail, and whole body regions.

6) *Symmetry*: To measure unbalanced muscle behavior of uncoordinated mutants, we characterized the way a worm body deviates from a perfect sine wave. These features include the amplitude (defined as the absolute distance between points on the skeleton to the line connecting the head and tail), the sum of signed distances to the head-tail line, the angle between the head-centroid line and the tail-centroid line, and the head-centroid and tail-centroid distances.

7) *Reversals*: Reversals are characterized by the distance and frequency of the worm moving back into its recent previous path. We kept a moving window to record the previous 20 centroid locations. A reversal was detected when the new centroid was closer to any of the 19 previous centroid locations than to the most recent past.

V. Classification and Feature Selection

Based on qualitative descriptions of uncoordinated mutant phenotypes, we expected the features measured by our system would provide useful quantitative definitions for specific mutant types [11]. To assess the ability of these image

features to provide effective characterization of *C. elegans* mutant phenotypes, we used the Random Forests algorithm [12], [2], [3]. Random Forests uses an ensemble learning scheme. Instead of generating a single classification tree, many trees (a *forest*) are generated independently by bootstrapping from the original data. A simple majority vote is taken for prediction. In addition to constructing each tree with a different bootstrap sample of the data, Random Forests adds additional randomness by splitting each node using a random subset of predictors instead of using the best split among all features as is done in CART [4]. These two layers of randomness turn out to perform very well compared to many other classifiers including discriminant analysis, support vector machines and neural networks, and the method is robust against overfitting [2]. An estimate of the error rate can be obtained by predicting using "Out-of-bag" (OOB) data, which are the data (around 36% of the data) that were not used in each bootstrap sample. The classification error rate is thus defined as the aggregated OOB prediction error rate. With enough trees, the OOB error rate is quite accurate [6]. There are two free parameters (number of trees in the forest, number of random features considered at each split).

Random Forests also provides four measures of feature importance that can be used for model reduction. Among these four, we chose to measure the importance of a feature using the average lowering of the margin across all samples when this feature is randomly permuted, because this was more robust against noise (Liaw, personal communication). For each sample, the margin is defined as the proportion of votes for its true class minus the maximum of the proportion of votes for each of the other classes.

VI. Results and Discussion

Tracking and Head/Tail Recognition: The tracking and head/tail recognition were tested on 161 5-minute video sequences (sampled at 2Hz) from 16 mutant types including more than 111,000 image frames. The videos were played back with the worm's tail marked by the algorithm for a human observer to verify. Experimental results are shown in Table I. The columns show mutant type, total number of frames in the video, number of frames that had head recognized as tail due to the tail section being lighter, and number of frames that had head recognized as tail due to grouping errors. The average error percentage is only around 2% for 111,233 frames tested.

Classification: For the classification, the forest was made up by 5,000 trees ($n_{trees} = 5000$). At each split, 15 features were randomly selected to be considered for splitting ($m_{try} = 15$), which is approximately the square root of the total 253 features used. The confusion matrix, represented by OOB errors, is shown in Table II. The classification success rates are listed along the main diagonal while the off-diagonal entries represent the misclassification error rates. The average success percentage is 90.9%, showing a high degree of success at identifying the correct mutant type even if presented with a single example recording.

TABLE I
HEAD AND TAIL IDENTIFICATION RESULTS.

Worm type	Total Frame	Recognition Wrong	Grouping Wrong	Error Percent
goa-1I	6193	0	1	0
unc-29	4679	0	0	0
w.t.	6057	5	0	0
egl-19	5503	1	0	0
cat-2	4908	0	5	0
dar-1	4954	0	0	0
dgk-1	4892	0	0	0
eat-4	5014	0	0	0
fb-1	4942	0	30	0.8%
nic-1	11817	1	0	0
unc-38	4853	3	0	0
unc-63	4926	0	0	0
unc-43	9908	251	0	2.5%
tph-1	10552	51	0	0.7%
unc-2	10361	69	0	0.7%
unc-36	10559	2613	3	25%
Total	**111233**	**3030**	**33**	**2%**

Fig. 3(a) shows the effect of m_{try} on the error rate with 5000 trees constructed. The errors are stable between 10 and 100 features ($error = [0.088, 0.095]$) and trend upward slightly afterwards. Fig. 3(b) shows the effect of n_{trees} when 15 features are selected at each split. The error converges quickly after 800 trees ($error = [0.090, 0.096]$) are constructed. Both plots indicate that the results are not sensitive to the selection of these two parameters.

Applications: Hundreds of genes have been identified in *C. elegans* that affect behavior in specific ways. Our long-term aim is to collect data on much larger numbers of mutant types and effectively classify them according to their phenotypic similarity. With an increasing data set, it becomes more challenging to identify features that effectively classify and distinguish the large variety of worms. The image processing methods developed here, including new features that require accurately identifying the head and tail, allowed us to achieve high classification accuracy even for a data set involving 16 different mutant types with subtle distinctions that are hard to classify by eye. These methods will be applied to other mutant types.

These methods also hold promise for investigating the clustering of behavioral patterns seen in different mutant worm types. In a recent study [10], we investigated the natural clustering of *C. elegans* behavioral phenotypes using data collected by our automated tracking system. From a complex data set consisting of 253 features (including the head and tail features described here) measured from behavioral recordings of 797 individuals representing 8 distinct genotypes, we used principal component analysis to represent each mutant type as a cloud of data points in low-dimensional feature space. We also used k-means clustering and Euclidean distance measurements to explore

TABLE II

CLASSIFICATION RESULTS FROM RANDOM FOREST.

Type	w.t.	*goa-1*	*nic-1*	*unc-36*	*unc-38*	*unc-29*	*egl-19*	*unc-2*	*tph-1*	*unc-63*	*dgk-1*	*unc-43*	*dar-1*	*ftp-1*	*eat-4*	*cat-2*
w.t.	93								1		2		1			3
goa-1	1	93								2	1	1			2	
nic-1			98	1						1						
unc-36				97		1		2								
unc-38			3	2	83	11				1						
unc-29		1			21	74		1		1						
egl-19				2			97									1
unc-2			1	1				97								
tph-1			1					3	89	1			2	1	3	
unc-63			6	1	1	3				87			1		1	
dgk-1	1	1								3	91				3	
unc-43							1					98		1		
dar-1					1		1		2				88	1	7	
ftp-1							3		1	1	5	1	1	85	2	1
eat-4			4	1					4	1					90	
cat-2	4									1	1		1	2		91

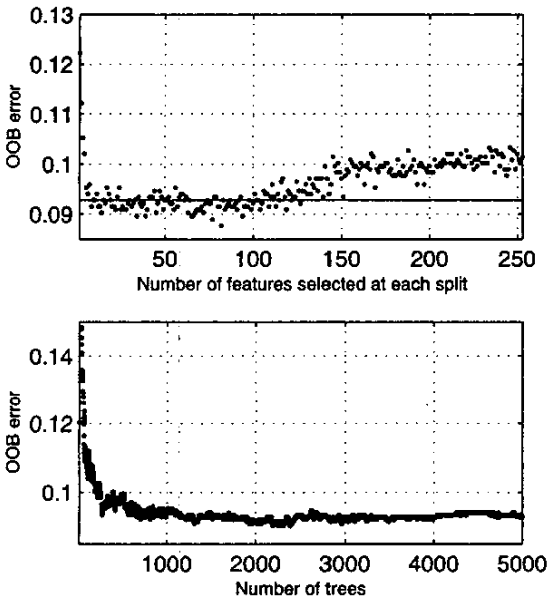

Fig. 3. Effect of parameters on OOB error rate.

Fig. 4. (A) Effect of number of features (m_{try}) selected at each split on OOB error rate. 5000 trees are used. The horizontal line represents the error rate using 15 features. (B) Effect of number of trees (n_{trees}) in the forest on OOB error rate. 15 features are used at each split.

the natural structure of the behavioral data and to compare the similarities of mutant phenotypic patterns. Encouragingly, the phenotypic classes inferred from the cluster analysis matched the known molecular similarities of the mutants that were grouped together; for example, two nicotinic receptor mutants formed a single cluster, as did two calcium channel mutants. Together, these results provided a precise and comprehensive definition of several important *C. elegans* behavioral phenotypes, and demonstrate that mutant phenotypes can be clustered using a complex behavioral signature based on quantitative image features.

Ackknowledgment: This work was supported in part by NIH grant 1 R21 DA15823-01.

REFERENCES

[1] J. Baek, P. Cosman, Z. Feng, J. Silver, and W. R. Schafer, "Using machine vision to analyze and classify *C. elegans* behavioral phenotypes quantitatively," *J. Neurosci Meth*, 118, pp. 9 21, 2002.

[2] L. Breiman, "Random forests", *Machine Learning*, 45(1), pp. 5-32, 2001.

[3] L. Breiman, "Manual on setting up, using and understanding random forests v3.1", *online material*, 2002.

[4] L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone, "Classifi cation and Regression Trees", *Wadsworth*, Belmont, CA, 1984.

[5] S. Brenner, "The genetics of Caenorhabditis elegans", *Genetics*, 77, pp. 77-94, 1974.

[6] T. Bylander, "Estimating generalization error on two-class datasets using out-of-bag estimates", *Machine Learning*, 48, pp. 287-297, 2002.

[7] R.Dhawan, D.B. Dusenbery, and P.L.Williams, "Comparison of lethality, reproduction, and behavior as toxicological endpoints in the nematode Caenorhabditis elegans", *J. Toxicology and Environmental Health*, Parrt A, 58, pp. 451-462, 1999.

[8] M. de Bono, C. I. Bargmann, "Natural Variation in a Neuropeptide Y Receptor Homolog Modifi es Social Behavior and Food Response in *C. elegans*", *Cell*, 94, pp. 679-689, 1998. *Wiley*, New York, Wiley, 2002.

[9] W. Geng, P. Cosman, J-H Baek, C. Berry, and W.R. Schafer, "Image Features and Natural Clustering of Worm Body Shapes and Motion", *Proc.IASTED International Conference on Signal and Image Processing(SIP2003)*, pp. 342-347, 2003.

[10] W. Geng, P. Cosman, J-H Baek, C. Berry, and W.R. Schafer, "Quantitative Classifi cation and Natural Clustering of *C. elegans* Behavioral Phenotypes", *Genetics*, to appear, 2003.

[11] T. Hastie, R. Tibshirani, and J. Friedman, "The Elements of Statistical Learning", *Springer*, New York, 2002. *J. Genetics*, 103, pp. 43-64, 1983.

[12] A. Liaw, M. Wiener, "Classifi cation and Regression by randomForest", *R newsletter*, Vol. 2/3, 2001. *Dev. Biol*, 100, 64-119, 1983.

[13] L. Waggoner, G.T. Zhou, R.W. Schafer, W.R. Schafer, "Control of alternative behavioral states by serotonin in Caenorhabditis elegans", *Neuron*, 21, pp. 203-214, 1998.

[14] T.Y. Zhang, C.Y. Suen, "A Fast Parallel Algorithm for Tinning Digital Patterns", *Comm. ACM*, vol.27, no.3, pp. 236-239, 1984.

[15] G.T. Zhou, W.R. Schafer, R.W. Schafer, "A three-state biological point process model and its parameter estimation", *IEEE Trans. Signal Process*, 46, pp. 2698-2707, 1998.