Supplementary Information for:

Leg-tracking and automated behavioral classification in *Drosophila*

Jamey Kain[1], Chris Stokes[1], Quentin Gaudry[2], Xiangzhi Song[3], James Foley[1], Rachel Wilson[2], Benjamin de Bivort[1,4,5]

[1]The Rowland Institute at Harvard, Cambridge, Massachusetts, USA. [2]Department of Neurobiology, Harvard Medical School, Boston, Massachusetts, USA. [3]College of Chemistry & Chemical Engineering, Central South University, Changsha, P. R. China. [4]Center for Brain Science, Harvard University, Cambridge. Massachusetts, USA.

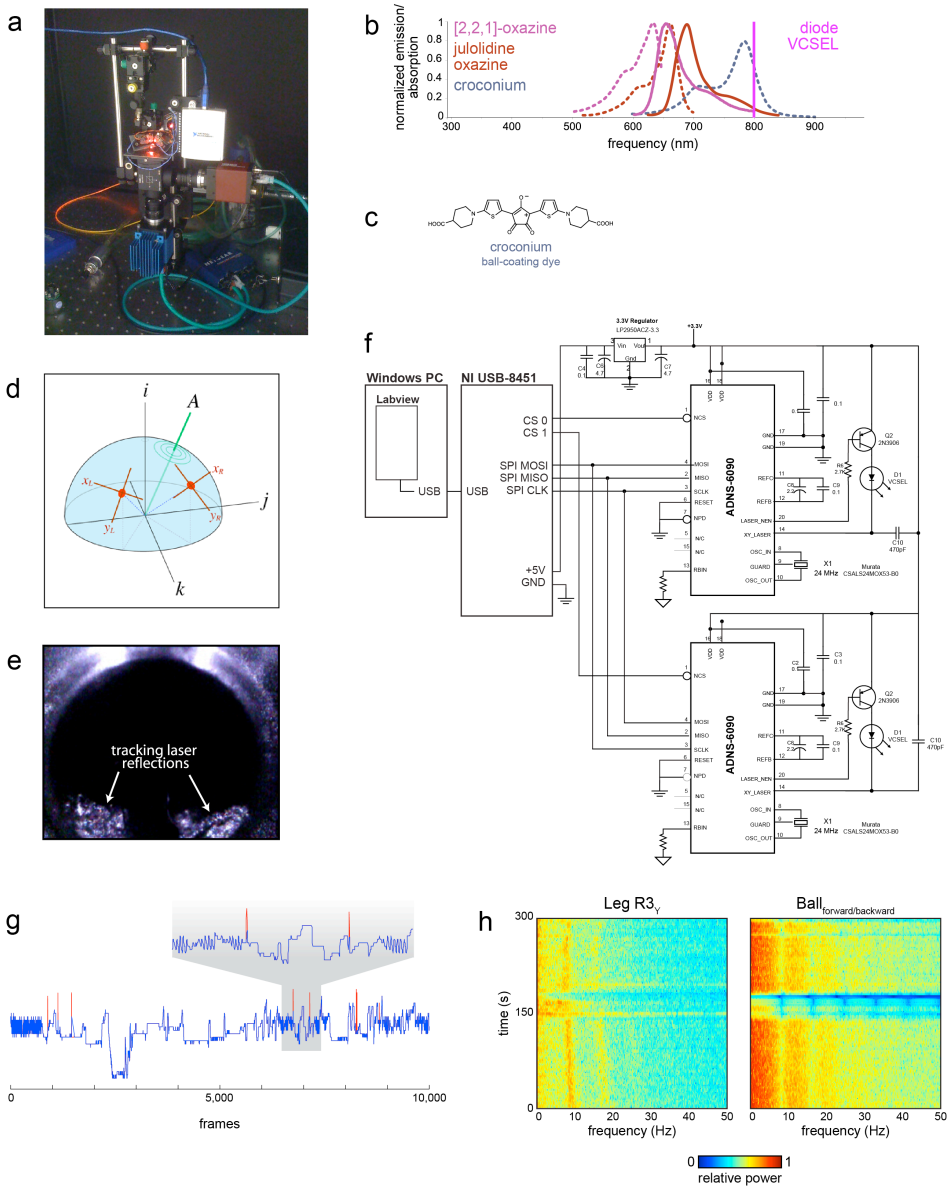
[5]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge Massachusetts, USA.

Correspondence should be addressed to B.D. (debivort@rowland.harvard.edu).

CONTENTS

SUPPLEMENTARY FIGURE

Figure S1 (de Bivort et al.)



Supplementary Figure S1| Additional apparatus details.

(a) Photograph of the leg-tracker setup.

(b) Absorption and emission wavelengths as in Figure 1e, with the optional croconium ball dye (see Supplementary Methods) included.

(c) Chemical structure of the croconium dye.

(d) Motion of the ball around its axis of rotation (A) is described by a basis set of three component rotational axes (i, j, k), and inferred from the ball's motion at the tracking spots (red).

(e) The appearance of the floating ball in infrared as illuminated by the tracking sensor lasers.

(f) Wiring diagram allowing the Avago ADNS6090 tracking sensors to operate once extracted from their original application in computer mice, as well as communicate with the SPI hub.

(g) A representative vector data trace before (red) and after (blue) clean up using a median filter and removal and interpolated replacement of rare erroneous readings. See Supplementary Methods.

(h) Sphere motion is sufficiently sensitive to respond to single strides, as shown by time-frequency analysis of an individual leg vector (left) and the pitch component of ball motion (right). Color indicates relative power. The ~9 and ~18Hz leg-motion modes drive corresponding modes in the motion of the ball.

SUPPLEMENTARY METHODS

Apparatus design details and use

Previous reports of *Drosophila* spherical treadmills used styrofoam, high-density polyethylene, or polyurethane foam spheres [9, 10, 12, 28]. This was not an option for our setup because we needed an optically clear sphere for imaging the dye spots of the leg from below. At first we thought we would need a spherical treadmill that matched the inertial mass of the tiny fruit fly. This led us to try aerogel, a rigid, ultralight and translucent material produced through supercritical drying [29]. However, it was time-consuming and difficult to make uniform spheres of aerogel, the material absorbed infrared and was very hygroscopic.

Next, we tried clear, acrylic spheres (1/4'', McMaster-Carr), which were much heavier than a fruit fly, but considerably more spherical than our in-house manufactured aerogel spheres. We were able to drastically reduce the mass of the sphere (down to <7mg) by spliting them in half, hollowing them out using Dremel rotary tools, and then glueing them back together. However, we found that the spheres were too light, and that the fly had trouble running because each stride would displace the sphere instead of spinning it. Therefore, matching the inertial mass of the sphere to the fly was not the correct approach and ultimately, we found the clear, acrylic spheres without modification yielded behaviors that, to our eyes, were the most naturalistic.

We introduced small, scattering surface imperfections on the spheres by lightly rolling them over 400 grit sandpaper to allow the infrared sensors to efficiently track motion. The more scuffed the sphere, the better it was tracked, but at the cost of poorer transmission of the fluores-

cence from the leg dyes to the cameras below. Thus, we incrementally scuff a sphere and strike a balance between infrared tracking and fluorescence transmission.

Alternatively, the spheres can be coated in the near-infrared absorbing/reflecting croconium dye (Piperdinium, 4-carboxy-1-[5-[3-[5-(4-carboxy-1-piperdinyl0-2-thienyl]-2-hydroxy-4,5-dioxo-2-cyclopenten-1-ylidene]-2(5H)-thienylidene]-, inner salt) to make it visible to the sensors. The croconium dye was prepared using the procedure reported [30]. However, all of the data presented in this work was collected using the sandpaper method.

The housing of the treadmill sphere bearing consisted of the rim of a 150μL Eppendorf tube. During trials, the ball was floated on air from a cylinder of compressed air (Airgas). The air was first bubbled through water to prevent static charge from accumulating, then passed through an adjustable flow regulator, and finally into a plenum to evenly distribute the flow into the air bearing. We used an airflow rate of approximately 440mL per minute.

Infrared sensors (Avago chip #ADNS-6090) for tracking the sphere are available in high-performance gaming mice. The wiring diagram of the circuit that allows these sensors to operate once removed from their original circuit boards, as well as communicate with a Serial Peripheral Interface (SPI) hub (National Instruments) is shown in Supplementary Figure S1f. (See the section below for details on transforming the four vectors from the two sensors into the three motion vectors of the sphere.) Measurement of angular displacement in the three components of rotational motion of the ball was calibrated by manually turning a calibration sphere (rigidly attached to a long, stiff, wire handle) through π radians on the roll and yaw axes and $\pi/2$ radians on the pitch axis, numerous times and averaging the integrated sensor motion.

The dyes were excited by a HeNe laser (632.8 nm, Thorlabs) which passes through a beam expander to ensure the space surrounding the legs was fully illuminated. In preliminary

experiments, we determined that groups of agitated flies were attracted to red LED light sources centered at 625nm, but not the HeNe source. Since a sphere will act as a lens, a second sphere was positioned within the optics housing to re-collimate the light before reaching the cameras. In conjunction with a beam splitter, two Prosilica GigE cameras (Allied Vision Technologies) were used for imaging the dyed legs from through the spheres, and acquired images at 80Hz typically. Band-pass filters (Edmund Optics) were used to optimize each camera's detection for either 221ox (HQ650/20m band-pass filter, Chroma) or julox (HQ700/20m band-pass filter, Chroma). Additionally, a short-pass filter (NT49-829, Edmund Optics) was used to block stray light from both the infrared tracking sensors and potential two-photon sources.

For the optomotor feedback experiments we used a standard 15 inch LCD monitor to display alternating blue (RGB=0000FF, width=4.8cm) and black (RGB=000000, width=2.4cm) colored vertical bars. Optic flow consisted of these vertical bars translating out (10cm/s) from a line of expansion. The monitor was positioned six inches from the tethered fly. The optomotor feedback trial comprised the conditions: 20min in the dark, 20min closed-loop, 20min dark, 20min open-loop, repeat from the beginning once. Closed-loop allowed the turns of the fly to control the position of the line of expansion for the translating vertical bars based on the detected turning-in-place (i) vector. Open-loop disabled the feedback control from the fly and allowed the point of origin for the diverging vertical bars to drift (0.5cm/s) to the point that the entire screen was filled with bars moving either left-to-right or right-to-left. During the 20min open-loop trial, the drift would switch direction every 5min. The 5 animals recorded in the optomotor experiment were different than the 8 animals recorded for day-to-day idiosyncrasy experiments, and the 5 animals used to train the classifier, as well as several animals recorded for day-to-day analysis that yielded only one run of satisfactory data.

Transforming the tracking sensor data into motion vectors

The X, Y coordinates from the sensors, integrated over 10 frames in a sliding window fashion as a noise reducing measure, were transformed to generate the 3 motion vectors of the sphere as follows:

- The axis of rotation of the ball $\boldsymbol{A}$ has components whose relative proportions are the proportional extents to which the fly is running forward/backward, sidestepping, or turning in place (respectively, pitch, roll and yaw). Specifically, rotation around i indicates turning in place, j forward/backward walking, and k sidestepping. We need to determine a, b, c such that $\boldsymbol{A}=\{ai,bj,ck\}$ - these are the relative proportions of each type of motion.
- Data from the sensors comes as two vectors $\boldsymbol{L}$ and $\boldsymbol{R}$ each in its own coordinate system (basis): $\boldsymbol{L}=\{x_L, y_L,0\}$, $\boldsymbol{R}=\{x_R, y_R,0\}$. First, convert the basis of these vectors to the global ball basis ($B=\{\{i,0,0\},\{0,j,0\},\{0,0,k\}\}$) using a change of basis matrix $\boldsymbol{C}$ derived from the positions of the sensors. Call the converted vectors $\boldsymbol{L}_B$ and $\boldsymbol{R}_B$. These will always be latitude lines with respect to $\boldsymbol{A}$. Therefore $\boldsymbol{A}$ is parallel to a vector which is perpendicular to both $\boldsymbol{L}_B$ and $\boldsymbol{R}_B$. Therefore $\boldsymbol{A}$ is parallel to $\boldsymbol{A}'=\boldsymbol{L}_B \times \boldsymbol{R}_B$.
    - $\boldsymbol{A}'_i$ is the % yaw
    - $\boldsymbol{A}'_j$ is the % pitch
    - $\boldsymbol{A}'_k$ is the % roll.

- From $\boldsymbol{A}$, we don't know the speed at which the ball is turning. This is derived from the length of either $\boldsymbol{L}$ or $\boldsymbol{R}$ as a function of their distance from $\boldsymbol{A}$. We calculate it based on whichever sensor is farther from $\boldsymbol{A}$, as this will be a less noisy measurement. Assume, without loss of generality, that $|\boldsymbol{L}|>|\boldsymbol{R}|$, i.e. that $\boldsymbol{L}$ is farther from $\boldsymbol{A}$. The angle between $\boldsymbol{A}$ and $\boldsymbol{L}$ is given by $\cos^{-1}((\boldsymbol{A}\bullet\boldsymbol{L})/(|\boldsymbol{A}||\boldsymbol{L}|))$. The surface speed of a sphere moving with speed 1 at an angular distance θ from the axis of rotation is $\sin(\theta)$, thus the speed of rotation of the ball (and the magnitude of the vector whose components are the fly's behavioral components) is

$$M = |\boldsymbol{L}|/\sin(\cos^{-1}((\boldsymbol{A}\bullet\boldsymbol{L})/(|\boldsymbol{A}||\boldsymbol{L}|)))$$

- While $\boldsymbol{L}_B \times \boldsymbol{R}_B$ is mathematically parallel to $\boldsymbol{A}$, when $\boldsymbol{R}_B$ is nearly parallel to $\boldsymbol{L}_B$, $\boldsymbol{A}'$ will approach 0, and its orientation will become highly sensitive to noise in the measurement of $\boldsymbol{L}$ and $\boldsymbol{R}$. This occurs when $\boldsymbol{A}$ lies in the plane of the bearing housing, i.e. perpendicular to i. To the extent that $\boldsymbol{L}_B$ and $\boldsymbol{R}_B$ are parallel, we want to use a different formula than their cross product. Since $\boldsymbol{L}_B$ and $\boldsymbol{R}_B$ are parallel, we can replace their information with their mean, and we can let "the extent to which they are parallel" be the cosine of the angle between $\boldsymbol{L}_B$ and $\boldsymbol{R}_B$:

$$\beta = \cos(\cos^{-1}((\boldsymbol{L}_B\bullet\boldsymbol{R}_B)/(|\boldsymbol{L}_B||\boldsymbol{R}_B|))) = (\boldsymbol{L}_B\bullet\boldsymbol{R}_B)/(|\boldsymbol{L}_B||\boldsymbol{R}_B|)$$

- This is the weighting factor given to the term of our final equation when $\boldsymbol{L}_B$ and $\boldsymbol{R}_B$ are parallel. The term itself is simply the cross product of i with the mean of $\boldsymbol{L}_B$ and $\boldsymbol{R}_B$, which by assumption are both perpendicular to $\boldsymbol{A}$. Thus, the noise-resistant ball-basis axis of rotation is given by:

$$\boldsymbol{A}' = \beta*(\boldsymbol{L}_B+\boldsymbol{R}_B)/2 \times \{(|\boldsymbol{L}_B|+|\boldsymbol{R}_B|)/2, 0, 0\} + (1-\beta)*\boldsymbol{L}_B \times \boldsymbol{R}_B$$

Developing the *k*-nearest-neighbors classifier

We implemented feed-forward neural networks and support vector machines (SVM) as potential non-linear classifiers for our instrument data before settling on *k*-nearest neighbors analysis (KNN). Those alternative methods showed promise, but neural networks are hard to interpret and have path-dependent performance variability. Compared to KNN, SVM required more computationally intensive optimization of two parameters (in the case of the most generally useful Gaussian radial bias kernel). Moreover, the performance of SVM was not obviously better than that of the simpler KNN. Importantly, we intend to present our technique to building the classifier as a general approach, and the intuitive underpinnings, computational simplicity, and successful performance of KNN were compelling. The classifier was built as follows.

A training set was built by manually scoring 4000 or 8000 video frames of a fly behaving on the ball while simultaneously recording the 15 raw data vectors. This was done independently by two investigators to assess the variability in manual scoring. Five different flies were scored.

Data were recorded as fast as possible during experiments, resulting in varying inter-frame intervals, with an average recording rate of ~90Hz. Prior to any additional processing, raw data was linearly interpolated to a uniform rate of 100Hz. The videos acquired for training purposed were recorded at 30Hz, and therefore the frames used in conjunction with the manual scores were selected for their correspondence with the time stamps of the video frames. Illumination of the HeNe laser was used as a trigger to synchronize the video and instrument data.

Raw data contained rare frames with erroneous leg-track positions. These typically arose when the dye spots moved out of the HeNe laser excitation, e.g. if they were occluded by the ab-

domen. Any trial with more than ~1% error frames was discarded. Raw data for the KNN training, and all other experiments, was filtered to remove error positions as follows. A 3-frame median filter was applied to all raw data vectors eliminating single-frame errors. The first frames of errors persisting longer than a single frame were identified by their values deviating from adjacent frames by more than 5 standard deviations of the inter-frame differences. The final frames of these extended errors were identified as the next frame within 5 standard deviations of the value prior to the error, or within 1 standard deviation of the median position across all frames. Values thus flagged as erroneous were replaced with values interpolated linearly from flanking frames. This approach was conservative, but removed essentially all conspicuous errors (Fig. S1g).

To explore which higher-order features were required for sufficient KNN performance, we used cross-validation at the level of flies – i.e. setting aside all of one fly's data, training the KNN classifier with the raw values and higher-order features from the remaining four flies from the training set. Correlation coefficient was used as the KNN metric. Performance was assessed using the withheld fly's raw data and higher-order features. Data from different flies were made comparable by *z*-score normalization within the data from each fly (to eliminate any positioning differences between mountings), then concatenated and *z*-score normalized again (to assure that all training variables would contribute equally to the classifications). The data from the withheld fly was z-score normalized, and then re-normalized using the mean and standard deviation of concatenated training set data.

Performance of the KNN with respect to one set of manual scores was assessed using the unbalanced accuracy, i.e. the overall percent of classifications matching the manual score. Also evaluated for reference was a so-called "plausible accuracy" which tolerates errors in assignment

between behaviors that are only arbitrarily distinguishable (e.g. the precise distinction between standing still and making small postural adjustments is necessarily an arbitrary cutoff). Misclassifications tolerated in this analysis include the above, postural adjustment with the locomotory behaviors, the locomotory behaviors with complex motion, leg1- and head-grooming, and leg3- and abdomen-grooming. We have observed that flies will groom both their hind-legs and abdomen in the same stroke, so this reflects biological overlap in addition to definitional continuity. The plausible accuracy between the manual scoring data sets was 96%, and between the KNN classifier scores and manual (JK) scores 91%. Manual (BD) scores were slightly more difficult for KNN to classify, yielding an 88% plausible accuracy.

Higher-order features characterizing local temporal dynamics in the data were calculated with a sliding window of +/-5frames. Other window sizes were evaluated and did not improve performance. Vector derivatives were calculated as the difference of values from frame t-1 to t+1. Local cross-correlations were calculated in a pair-wise fashion between all 15 raw values in a sliding window of +/-5frames. Vectors of higher-order features were padded with 0s as necessary. The KNN k parameter was optimized to 16 using just the raw data, and re-optimized to 24 once the best training set (raw data + derivatives + local standard deviations) was identified. KNN performance was quite robust to the choice of $4<k<100$. Lastly we noticed that the model was prone to switch between behavioral types more frequently than the manual scorers. We explored two approaches to address this. First we implemented a sequential frame Bayesian hidden Markov Model with priors extracted from the training data set. This provided little improvement and thus was not kept. However, a simple low-pass filter that replaced each frame's classification with the most abundant classification in a +/-5frame sliding window did improve performance, albeit slightly (5%).

Raw data collected post-KNN-training was interpolated, filtered for errors, normalized, augmented with higher-order features, and low-pass filtered as above, with the exception that the training samples included data from all 5 training flies, i.e. none were withheld.

SUPPLEMENTARY REFERENCES

28. Gotz, KG. & Wenking, H. J. Visual control of locomotion in the walking fruitfly Drosophila. Comp. *Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **85**, 235-266 (1973).

29. Kistler, SS. Coherent expanded aerogels and jellies. *Nature* **127**, 741 (1931).

30. Song, X. & Foley, JW. A new water-soluble near-infrared croconium dye. *Dyes and Pigments* **78**, 60-64 (2008).