



Published in final edited form as:

*Nat Neurosci.* 2014 February ; 17(2): 296–303. doi:10.1038/nn.3600.

## Flies and humans share a motion estimation strategy that exploits natural scene statistics

Damon A. Clark<sup>1,2,\*</sup>, James E. Fitzgerald<sup>3,4,\*</sup>, Justin M. Ales<sup>5,6,\*</sup>, Daryl M. Gohl<sup>1</sup>, Marion A. Silies<sup>1</sup>, Anthony M. Norcia<sup>5</sup>, and Thomas R. Clandinin<sup>1</sup>

<sup>1</sup>Department of Neurobiology, Stanford University, 299 W. Campus Dr. Stanford, CA 94305

<sup>3</sup>Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305

<sup>5</sup>Department of Psychology, Stanford University, 450 Serra Mall, Stanford, CA 94305

### Abstract

Sighted animals extract motion information from visual scenes by processing spatiotemporal patterns of light falling on the retina. The dominant models for motion estimation exploit intensity correlations only between pairs of points in space and time. Moving natural scenes, however, contain more complex correlations. Here we show that fly and human visual systems encode the combined direction and contrast polarity of moving edges using triple correlations that enhance motion estimation in natural environments. Both species extract triple correlations with neural substrates tuned for light or dark edges, and sensitivity to specific triple correlations is retained even as light and dark edge motion signals are combined. Thus, both species separately process light and dark image contrasts to capture motion signatures that can improve estimation accuracy. This striking convergence argues that statistical structures in natural scenes have profoundly affected visual processing, driving a common computational strategy over 500 million years of evolution.

### Introduction

The statistical distribution of light intensities across space is a core feature of any environment<sup>1–3</sup>. These spatial distributions can be sampled over time to extract information about visual motion, a critical behavioral cue for many animals. The dominant computational models of motion processing estimate motion by correlating light intensity

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:[http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

<sup>2</sup>Present address: Department of Molecular, Cellular, and Developmental Biology, Yale University, 219 Whitney Ave. New Haven, CT 06511

<sup>4</sup>Present address: Center for Brain Science, Harvard University, 52 Oxford St. Cambridge, MA 02138

<sup>6</sup>Present address: School of Psychology and Neuroscience, University of St Andrews, St Mary's Quad, South Street, St Andrews, KY16 9JP, Scotland, UK

\*These authors contributed equally to this work.

#### Author Contributions

DAC, JEF, and JMA designed and carried out the experiments. JEF performed the natural scenes analysis and developed the theoretical modeling framework. DAC performed the fly experiments. DAC and JMA performed the human experiments. DMG and MAS contributed reagents and fly edge selectivity measurements. DAC, JEF, JMA, AMN, and TRC interpreted the data and wrote the paper.

between pairs of points separated in space and time or, equivalently, by measuring local motion energy<sup>4,5</sup>. These pair correlations provide information about the direction and speed of moving edges. However, natural scenes contain additional information about motion that these signals do not capture<sup>6–8</sup>. For example, the contrast polarity, being either dark or light, is also a fundamental feature of moving edges yet is explicitly discarded by pair correlations. Here we show that both the fly and human visual systems take advantage of this additional information, available in specific correlations between three points in space and time, to detect motion.

There are two dominant models of motion perception. The first of these is the Hassenstein-Reichardt Correlator (HRC), which computes spatiotemporal correlations directly by multiplying local contrast signals at two points in space, one at a later time than the other. These products are then summed in anti-symmetric fashion to produce an average signal whose sign and amplitude indicates the direction and magnitude of motion (Fig. S1)<sup>4</sup>. Motion energy, a second correlational model, begins with linear, oriented spatiotemporal receptive fields that are sensitive to particular directions of motion (Fig. S1)<sup>5</sup>. Subsequent circuit operations square and then sum these responses to produce a motion signal. Based on neural and behavioral measurements, motion energy models have been favored in vertebrates<sup>9</sup>, while HRC models have been favored in invertebrates<sup>10</sup>. Nonetheless, the two models are sometimes mathematically equivalent<sup>5,11</sup>, and both ultimately compute correlations only between pairs of points in space and time (see SI).

Experiments using a variety of artificial stimuli have demonstrated that both vertebrates and invertebrates can detect motion even when there are no systematic correlations in intensity between pairs of points<sup>12–16</sup>. An optimal motion estimator would incorporate prior statistical information about the environment and its motion, and would compute many types of stimulus correlations to take advantage of higher-order statistics in moving natural scenes<sup>6</sup>. In particular, analysis of optimal estimators suggests that natural luminance asymmetries<sup>2,17</sup> would allow animals to estimate motion using triple correlations<sup>6</sup>. Here we show that two very different visual systems, those of flies and humans, employ triple correlations to estimate motion in a manner that distinguishes light and dark edges. These results suggest that the separate processing of dark and light in the visual pathways of many organisms can increase the fidelity of motion perception.

## Results

To demonstrate how the motion of natural scenes generates spatiotemporal correlations, we approximated full-field motion by rigid translations of natural images (Fig. 1ai, 1bi). Minimal motion energy and HRC-based models rely exclusively on information extracted from pairwise correlations across the image. One simple example of this correlation structure is the difference between rightward and leftward correlations (Fig. 1aai). In this case, the local correlation, on average, indicated the correct direction of motion (red areas in Fig. 1bii), but because of the variability within the image<sup>18</sup>, this signal also suggested leftward motion in some regions (blue areas in Fig. 1bii). In this example, the standard deviation of the local motion signal, computed across pixels, was 3.6 times the mean. Spatiotemporal averaging can suppress this variability<sup>18</sup>, but at the expense of resolution.

Here we consider two triple correlation structures involving three points in space and time, which we refer to as *diverging* (Fig. 1aiii) and *converging* (Fig. 1aiv). The diverging case incorporates two spatial points at the later time (one point *diverges* into two points), while the converging case incorporates two spatial points at the earlier time (two points *converge* onto one point). These two triple correlations maintain the local spatial and temporal resolutions of the comparable pair correlation. Like the pairwise estimator, these triple correlations were highly variable across the image, but their average signified the direction of motion (Fig. 1biii–biv). Importantly, the motion signals provided by these triple correlations incompletely overlapped with the motion signals derived from pair correlations (Fig. 1bii–biv). Thus, 3-point motion signals provide additional information about motion, beyond what can be obtained from the pairwise signal<sup>6</sup>.

Because the accuracy of motion estimation is scene-dependent, one must determine whether motion estimators that capture specific spatiotemporal correlations perform reliably across an ensemble of scenes. We used a published natural image database<sup>19</sup> to determine how well pair and triple correlations can predict image velocity (Fig. 1c–d). We considered motion estimators with human-like spatial resolution (see Online Methods, Fig. 1ci) and with *Drosophila*-like spatiotemporal sampling (see Online Methods, Figs. 1di, S1). In both cases, we approximated the distribution of image velocities with a Gaussian having zero mean, using standard deviations that were 5°/s and 90°/s respectively, comparable to estimated natural speeds<sup>20,21</sup>. The output of local pairwise motion estimators was weakly correlated with the image velocity (Fig. 1cii–dii). This correlation was improved by averaging two pairwise estimators that survey neighboring spatial points (Fig. 1cii–dii). The diverging and converging triple correlations were more weakly correlated with the velocity (Fig. 1ciii–civ, diii–div), but typically improved motion estimation when summed with the pairwise estimate using optimal weighting coefficients (Fig. 1ciii–civ, diii–div). Strikingly, for *Drosophila*-like sampling, the increased accuracy afforded by the converging 3-point motion estimator (Fig. 1div) exceeded that afforded by a neighboring pairwise estimator (Fig. 1dii). Because these two 3-point estimators sampled the same two spatial and temporal points as the pairwise estimator, this improvement is available without sacrificing either spatial or temporal precision. Importantly, these triple correlations require asymmetric contrast distributions for their functionality<sup>6</sup> (Fig. S2), thereby capitalizing on the strong asymmetries present in natural contrast distributions that are absent from many artificial stimuli (Fig. S2 and S3).

Triple correlations, unlike pair correlations, can encode whether a moving edge is light or dark (Fig. 2). Light and dark edges are defined by whether light intensity at a single point in space increases (light edge, green in Fig. 2) or decreases (dark edge, violet in Fig. 2) as the edge moves across that point in space. The net pairwise correlation motion signal was positive when either edge type moved to the right and was negative when either edge type moved to the left (Fig. 2a). Because moving light and dark edges induce the same pair correlations, the contrast polarity of the moving edge cannot be discerned from pairwise correlations alone (see also S1). Conversely, 3-point estimators do capture edge polarity information. The diverging 3-point estimator produced a positive signal to rightward-moving light edges and leftward-moving dark edges, and a negative signal to leftward-

moving light edges and rightward-moving dark edges (Fig. 2b). Thus, triple correlations jointly encode the direction of motion and the contrast polarity of the moving edge. Critically, this joint encoding implies that the contrast polarity of each moving edge can be deduced in local regions of space from triple correlations once the direction of motion is determined. For example, when the motion is rightward, positive diverging triple correlations signal the presence of a light edge and negative diverging triple correlations signal the presence of a dark edge. We observed similar response patterns in pair and triple correlations when we simulated moving edges by translating natural scenes (Fig. 2c). Triple correlations are able to improve motion estimates in natural scenes (Fig. 1) precisely because in natural scenes, light and dark moving edges generate asymmetric triple correlation signals (see SI).

The early insect visual system contains distinct substrates specialized for detecting moving light and dark edges<sup>22,23</sup>. Because triple correlations can improve motion estimation accuracy and can discriminate between the motion of light and dark edges, we next determined whether flies actually detect these signals. Following previous psychophysical approaches<sup>14</sup>, we constructed “glider” stimuli that enforce positive or negative correlations of the same form as the pairwise, diverging, and converging correlations shown in Figure 1 (Fig. 3a and S4). Importantly, the 3-point “gliders” contained no net 2-point correlations, and *vice versa*. Thus, by construction, these glider stimuli separate the motion information contained in 3-point correlations from that specified by 2-point correlations (see Online Methods and Fig. S4).

We presented spatially homogeneous glider stimuli on panoramic screens arranged around tethered flies. In this apparatus, flies walked on an air-cushioned ball, which was tracked to monitor fly turning<sup>23,24</sup> (Fig. 3b). Flies respond to visual rotations by turning in the direction of motion (the optomotor response), thereby allowing the movement of the ball to provide a behavioral measure of the fly’s motion percept<sup>25</sup>. As predicted by the HRC, flies turned in one direction when presented with the positive 2-point correlations and in the opposite direction with the negative 2-point correlations (top, Fig. 3c–d)<sup>4,23,26</sup>. Remarkably, flies also turned in response to the diverging and converging 3-point gliders, with responses that approached 20% of the 2-point glider response (bottom, Fig. 3c–d). These 3-point glider stimuli are, by design, very different from natural motion, as they achieve their correlation specificity by averaging out all 2-point correlations, whereas natural motion contains both 2-point and 3-point correlations (Fig. 1). As with the 2-point stimuli, positive and negative correlations in 3-point gliders evoked turns in opposite directions. In other words, simply inverting the contrast of 3-point glider patterns inverted their perceived directions. Neither the canonical HRC nor the motion energy model predict that flies would respond to 3-point glider stimuli<sup>14</sup> (Fig. 3d, bottom), and a recent modification to the HRC<sup>27</sup> also does not predict the measured responses (Figure S4d). This behavioral response was not a generic consequence of arbitrary triple correlations in the stimulus, as flies responded only weakly to several other glider stimuli (Fig. S5).

Previous studies demonstrate that the disruption of large monopolar cell (LMC) function causes selective behavioral deficits in response to moving light or dark edges<sup>22–24,28</sup>. We therefore determined how the disruption of LMC function affects behavioral responses to

the diverging and converging 3-point gliders by genetically suppressing synaptic output from the three LMCs (L1–L3) that have been associated with motion detection<sup>24</sup> (Fig. 4a). Control strains in which LMC function was normal all responded similarly to wild-type flies for each glider (Fig 4b). When LMC function was genetically disrupted, responses to various glider stimuli increased, decreased, or inverted compared to the controls (Fig 4b).

Triple correlations are differentially associated with moving light and dark edges (Figs. 2b), as are the three LMCs (L1–L3) that provide inputs to motion detecting circuits<sup>22–24</sup>. We thus tested whether 3-point correlation responses predicted the relative strength of responses to moving light and dark edges both in wild-type flies and in flies with disrupted motion pathway inputs. To do this, we used an edge selectivity index corresponding to the behavioral response to light edges minus the response to dark edges divided by their sum<sup>22–24</sup>. Each glider forces a particular correlation to occur consistently throughout the visual field (see Online Methods). We thus used the observed behavioral response of each genotype of flies to a glider stimulus to infer its sensitivity to the associated correlation. We then predicted the response of each genotype to moving light and dark edges as the appropriately weighted sum of its 2- and 3-point glider responses (Fig. 5), where the weighting scheme was determined by counting how often each glider’s correlational element appeared in each edge type (Fig. 5a and S6). Remarkably, we observed a high correlation between the edge selectivity predicted by the weighted responses to glider stimuli and the independent measurements of edge selectivity (Fig. 5b). This striking result suggests that these 3-point correlations are integral to edge selectivity in flies.

In the primate visual system, light and dark are differentially processed in a variety of ways<sup>29</sup>. We therefore examined whether humans have separate pathways to process the direction of motion for light and dark edges, and whether triple correlations are involved in human edge contrast selectivity. Instead of genetic manipulation, as in flies, we used differential adaptation. In a first experiment, we measured scalp EEG signals, and in a second experiment, we measured behavioral responses. We designed stimuli that independently manipulate both edge contrast polarity and motion direction<sup>23</sup>. In particular, we generated two complementary “opposing edge” stimuli in which every light edge moved in one direction and every dark edge moved in the opposite direction (Fig. S7, supp. movie M1). We denoted the stimulus where light edges moved to the right as “A” and the stimulus where light edges moved to the left as “B.” These stimuli were balanced for leftward and rightward motion and for positive and negative edge contrast polarity, but they were imbalanced in the compound feature that combines motion direction with edge polarity.

Human participants were presented with either stimulus A or stimulus B as an adapting stimulus and then probed with a stimulus that rapidly interleaved short segments of A and B, designated A’ and B’ (Figs. 6a and S7, supp. movie M2), while EEG signals were recorded on the scalp (Fig. 6b). If light and dark edge motions are treated equivalently by neural circuitry, then the probe response should be independent of the adapter. If, however, A and B differentially adapt some neural population, then they would be expected to have complementary effects on the probe response. In particular, as each adapter corresponds to one half of the probe, selective adaptation should create a response that coincides with the rate of alternation between A’ and B’ in the probe. Since the two halves of the probe were

constructed to be 180° out of phase, the differentially adapted responses should also be 180° out of phase. Indeed we found that the two adapting stimuli evoked different departures from baseline in the EEG signal (Fig. 6c, S7) (n=7 participants), and the measured phase difference between them was 179±20° (Fig. 6d, S7).

To compare these adaptation effects with those caused by adaptation to drifting sine-wave grating motion<sup>30</sup>, we computed an adaptation index that quantified the fractional change in response due to adaptation<sup>30</sup> (see Online Methods). Previous experiments using drifting sine-wave gratings produced an adaptation index of  $2.31 \pm 0.52$  (mean ± SEM). Here we observed an adaptation index of  $1.71 \pm 0.18$  (mean ± SEM). The difference between these adaptation indices is small and not statistically significant, demonstrating that these adaptation effects have a similar fractional magnitude. This is remarkable because the adapting stimuli in the current experiment contained no net motion, but generated direction-selective adaptation. Thus, these experiments demonstrate that the human visual system contains neurons that differentially adapt to the motion of light and dark edges, thereby suggesting that neural populations are differentially activated by the motion of these two edge types.

We next investigated whether edge-polarity-selective motion pathways in humans were associated with behavioral responses to triple stimulus correlations, as in flies. By asking participants to classify glider stimuli as moving leftward or rightward, we first reproduced previous psychophysical results demonstrating that humans perceive diverging and converging 3-point gliders as motion<sup>14</sup> (black bars, Figs. 7 and S8c, supp. movie M3). We then adapted edge selective motion pathways using opposing edge stimuli as adapters, and probed the perception of glider motion. Results were aggregated using the relative orientation of the adapter and probe (see Online Methods). These stimuli had no effect on converging glider percepts (gray bars Figs. 7 and S8c). However, these opposing edge adapters inverted the perceived direction of a specific glider (the negative left-diverging glider; Fig. 7).

As opposing edge adapters contain both light and dark edge motion, this adaptation effect could emerge from moving light edges, dark edges or both. To resolve this ambiguity, we generated two more types of adapting stimuli, each containing either light or dark moving edges. These stimuli were not motion balanced, and some participants reported a strong motion aftereffect that caused all responses to be in the opposite direction to the adapter (n=2 of 9, see Fig. S8d). Thus, we restricted subsequent analyses to those participants whose responses were not dominated by motion after-effects (n=7 of 9). Light edge adaptation caused perceptual inversion of the same glider affected by opposing edge adaptation (green bars, Fig. 7). In contrast, dark edge adaptation had no effect on glider perception (purple bars, Fig. 7). Because the light and dark edge adapters have different effects on glider perception, yet contain equivalent net motion signals, this specific pattern of adaptation cannot be explained by a simple motion after-effect. Thus, these experiments demonstrate that at least one triple correlation is differentially involved in edge polarity-selective motion processing in humans.



## Discussion

Our results indicate that both flies and humans extract the motion of light and dark edges via distinct processing pathways in a way that allows both organisms to exploit higher-order statistical correlations that are present in moving natural scenes. We demonstrated this in the detailed pattern of human and fly psychophysical responses, in the profile of human neural responses after adaptation, and in the behavioral responses of flies after genetic manipulation of underlying input channels. Our analysis of natural scenes showed that triple correlations provide information that can improve motion estimates without reducing spatial or temporal resolution. Such odd-ordered correlations are also required for discriminating between edge polarities. Building on previous work showing that visual input circuits in *Drosophila* are separately specialized for detecting moving light and dark edges<sup>22,23</sup>, we showed that flies respond to triple correlations, and that the pattern of these responses predicted edge polarity selectivity across genotypes. Our EEG measurements found a neural correlate of polarity specific motion pathways in humans, consistent with previous psychophysical observations<sup>31,32</sup>. In addition, while humans had previously been shown to perceive motion in 3-point glider stimuli<sup>14</sup>, we showed that behavioral responses to specific 3-point correlation stimuli are differentially adapted by light and dark moving edges. Thus, we have shown that higher-order correlations play central roles in distinguishing moving edge polarity in both flies and humans.

Models of motion processing that are restricted to pairwise correlations (e.g. the motion energy model) have been extremely successful in explaining both perception and neural activity in areas V1 and MT<sup>9,33,34</sup>. Nevertheless, there is also strong evidence that such models cannot fully capture motion processing in either primates<sup>14,15</sup> or invertebrates<sup>12,13,35</sup>. For example, humans perceive robust motion in certain stimuli lacking pairwise correlations (“non-Fourier” motion)<sup>15</sup>. Moreover, experiments in non-human primates have demonstrated that non-Fourier motion can elicit direction-selective behavioral responses in the absence of direction-selective neural responses in area MT<sup>36,37</sup>. Several studies have shown that flies can detect certain non-Fourier motion cues<sup>12,13,35</sup>. This prior work utilized motion stimuli that included quadruple (4-point), or higher even-ordered correlations, but lacked triple correlations. As a result, models that account for these effects specifically detect quadruple correlations<sup>35</sup>. Our current results highlight that triple correlations are important for estimating motion when luminance distributions are asymmetric, as they are in natural scenes<sup>38</sup>, and for encoding the contrast polarity of a moving edge. More generally, our results illustrate how non-Fourier motion cues that seem paradoxical in isolation can signify motion in natural environments<sup>6</sup>.

Relatively minor changes to existing biological models of motion detection can provide access to high-order motion correlations. For example, the ON and OFF channels in the vertebrate retina treat contrast increments and decrements differently, in terms of both amplitude and kinetics<sup>29</sup>. This asymmetry, were it to be appropriately retained, could enhance motion estimation by giving downstream neurons access to higher-order correlations. Similarly, the existence of ON direction-selective retinal ganglion cells<sup>39</sup> shows that contrast polarity specific motion signals are present already in the retina. Interestingly, our demonstration that triple correlations both improve motion estimation in

natural environments and associate with edge polarity specific pathways in fly and human visual systems suggests that brains might utilize separate ON and OFF processing channels to extract complex signatures of natural motion. Nevertheless, it is also possible that the brain explicitly computes higher-order correlations. A variety of machine vision algorithms make direct use of higher-order statistics to improve motion estimation<sup>7,8,40</sup>. Here we multiply three signals to build HRC-like models that compute triple correlations. The vertebrate motion energy model computes and squares local spatiotemporal frequency components to compute the motion signal. By extension, multiplying three frequency components can produce the bispectrum, which encodes triple correlations. Thus, simple generalizations of standard models that compute pairwise correlation enable the computation of triple correlations.

Dark and light have long been considered to be perceptually distinct<sup>29</sup>. Our work extends studies in primates showing that motion processing has a component that retains information about edge contrast polarity<sup>31,32,41</sup>. The distinct processing of light and dark moving edges may reflect fundamental differences in the statistics of light and dark objects in the world. For example, luminance asymmetries in natural scenes imply that the contrast magnitude of light edges can vastly exceed the contrast magnitude of dark edges<sup>2,17</sup>. More complex asymmetries also exist. For instance, light and dark are correlated with distance in natural scenes, with near elements tending to be brighter than far ones<sup>42</sup>. Further, the asymmetry in magnitude between light and dark in natural scenes implies that light objects tend to have a smaller spatial extent than dark objects. Therefore, light edges tend to be quickly followed by dark ones, and processing light and dark edges separately could reflect ethological differences in the detection of light and dark edges as the leading or lagging edges of moving objects<sup>43</sup>. Overall, light and dark edges are not symmetric in natural environments, and treating each separately can enhance motion estimation.

We have demonstrated that triple correlations can distinguish the motion of light and dark edges and can improve the accuracy of motion estimation without sacrificing spatiotemporal resolution. Although we have emphasized the utility of triple correlations for wide-field motion estimation, these properties might make triple correlations even more critical in behavioral contexts where light and dark signals hold special significance or where extensive spatiotemporal averaging is not possible. For example insects track the motion of dark objects<sup>44</sup> and exhibit stereotyped escape responses when presented with looming visual stimuli<sup>45</sup>. The stimuli that elicit these behaviors have strong light-dark edge dependencies: a dark object moving across a light background always consists of a dark leading edge and light lagging edge<sup>43</sup>, whereas the edges of looming objects all have the same contrast polarity and move in different directions. Object tracking is also distinct from wide-field motion estimation in humans<sup>15</sup> and could similarly exploit triple correlations to account for object polarity. In addition, depth estimation using binocular disparity has many similarities with motion estimation and also utilizes luminance polarity information through independent disparity-tuned mechanisms for light and dark features<sup>46</sup>. Interestingly, observers better estimate depth when the stimulus contains both light and dark information<sup>46</sup>, possibly reflecting the aforementioned correlation between luminance and distance in natural scenes<sup>42</sup>. Consistent with these observations, neurons in macaque V1



show conjoint disparity and luminance tuning that reflects the pattern of environmental correlations between brightness and depth<sup>47</sup>.

There is strong selective pressure to compute accurate motion estimates. The statistics of the terrestrial world, combined with biological and physical constraints on neural circuitry<sup>48–50</sup>, shape the computations that underpin these motion estimates, resulting in the possibility of similar motion estimation solutions across diverse taxa. The common existence of light and dark edge selective pathways in flies and humans, along with the association of triple correlation computation with those pathways, points to a deep similarity between fly and human motion estimation strategies. Since chordates and arthropods diverged more than 500 million years ago, and their visual systems have very different architectures, it seems unlikely that these similarities are the result of a conserved algorithm derived from a common ancestor. Rather, these commonalities could reflect the purposeful use of high-order statistics to estimate motion, leading flies and humans to converge on remarkably similar computational strategies to process this critical cue.

## METHODS

### Natural Scene Statistics Analysis Methods

**Correlation images**—We began with natural images from van Hateren’s database (image size = 1024×1536 pixels, pixel size = one arc minute)<sup>19</sup>. We linearly converted each image to a contrast scale,  $C_i = (I_i - I_0)/I_0$ , where  $C_i$  is the  $i^{\text{th}}$  pixel contrast,  $I_i$  is its intensity, and  $I_0$  is the image’s average pixel intensity. The pair correlation image (Fig. 1bii) was

$$R_i = C_i(t)C_{i+\Delta}(t+\delta) - C_i(t+\delta)C_{i+\Delta}(t),$$

where  $R_i$  denotes the  $i^{\text{th}}$  pixel of the pair correlation image,  $C(t)$  denotes the image at time  $t$ , summations inside the pixel index denote horizontal shifts,  $\Delta = 10$  arcmin, and  $\delta = 30$  ms. The velocity of motion was  $v = 5.6^\circ/\text{s}$ , so  $C_i(t) = C_{i+} (t + \delta)$ . The diverging (Fig. 1biii) and converging (Fig. 1biv) triple correlation images were

$$\begin{aligned} D_i &= C_i(t)C_i(t+\delta)C_{i+\Delta}(t+\delta) - C_i(t+\delta)C_{i+\Delta}(t)C_{i+\Delta}(t+\delta), \\ N_i &= C_i(t)C_{i+\Delta}(t)C_{i+\Delta}(t+\delta) - C_i(t)C_i(t+\delta)C_{i+\Delta}(t), \end{aligned}$$

where  $D_i$  and  $N_i$  denote the  $i^{\text{th}}$  pixels of the diverging and converging triple correlation images. Symmetric color axes saturate one standard deviation from 0.

**Quantitative comparison of motion estimators**—To consider the accuracy of motion estimation strategies built from raw image correlations, we converted images to contrast, randomly chose a row from an image, sampled the velocity from a zero-mean Gaussian distribution with  $5^\circ/\text{s}$  standard deviation, and computed local pair and triple correlations as above. We performed  $10^7$  simulations for Fig. 1c and S3 and  $10^5$  simulations for Fig. S2.

We also considered strategies with HRC-like spatiotemporal sampling. Since pixels were small relative to *Drosophila*'s sampling, we down-sampled each image to one-degree pixels by averaging. We converted images to contrast and then emulated photoreceptor blurring by filtering across rows with a Gaussian kernel (FWHM = 5.7°). We considered the central row of the filtered image as a one-dimensional image,  $c(x)$ . Given a randomly chosen image and a velocity drawn from a zero-mean Gaussian with 90°/s standard deviation, we modeled photoreceptor responses as

$$V_i(t) = \int d\tau T(\tau) \int dx h(x - x_i) c(x - v(t - \tau)),$$

where  $i$  indexes the photoreceptor,  $x_i$  is the position of the  $i^{\text{th}}$  photoreceptor (spaced 5.1° apart),  $T$  is a causal exponential kernel (timescale = 10 ms), and  $h$  is a Gaussian kernel (FWHM = 5.7°)<sup>51</sup>. We applied reflective boundary conditions to generate images that covered the visual field. The general HRC is

$$R_i(t) = (f * V_i)(t) (g * V_{i+1})(t) - (g * V_i)(t) (f * V_{i+1})(t).$$

We took the filters to be

$$f(t) = te^{-t/\tau}$$

$$g(t) = \frac{df}{dt}$$

where  $\tau = 30 \text{ ms}^{23}$ , and  $g(t)$  is comparable to LMC responses<sup>52</sup>. The converging and diverging third-order correlators used identical filters:

$$N_i(t) = (f * V_i)(t) (f * V_{i+1})(t) (g * V_{i+1})(t) - (f * V_i)(t) (g * V_i)(t) (f * V_{i+1})(t),$$

$$D_i(t) = (f * V_i)(t) (g * V_i)(t) (g * V_{i+1})(t) - (g * V_i)(t) (f * V_{i+1})(t) (g * V_{i+1})(t),$$

where  $N_i(t)$  and  $D_i(t)$  denote the converging and diverging 3-point correlators (Fig. S1). We performed  $4 \cdot 10^5$  simulations (duration = 800 ms, time step = 10 ms) and considered each correlator's final value as its velocity estimate.

We estimated the correlation between each estimator and the velocity using 2-fold cross-validation. For each data partitioning, we randomly assigned half of the simulations to the training set (the rest comprised the test set). We determined the optimal linear coefficients to combine motion estimators from the training set and evaluated performances on the test set. The percentages in Figs. 1, S2, and S3 are fractional correlation increases, relative to the local 2-point correlator. Figs. 1c, S2, and S3 show means and standard deviations across 100 data partitionings (1000 in Fig. 1d).

## Drosophila Methods

**Strains**—*Drosophila melanogaster* were raised and prepared for testing as previously described<sup>23</sup>, grown on molasses-based food in a 12hr-12hr light-dark cycle, and tested for behavior during the 4 hours after lights-on or 4 hours before lights-off. Females were collected on CO<sub>2</sub> 1–2 days after eclosion, then tested 48–72 hours later, using cold to immobilize them prior to gluing with a UV-cured epoxy. Genotypes are listed in Table S1. All inserts were in the isoD1 background<sup>53</sup> or backcrossed 5 times into that background.

**Fly Behavior**—Fly responses to different visual stimuli were measured with an apparatus similar to one previously described<sup>23</sup>, with modified screens<sup>24</sup>. Three screens, each 3 × 3cm, were arranged as adjacent sides of a cube with the fly at the center. They subtended from +135° to –135° azimuthally, and from +45° to –45° vertically. Images were relayed from a digital light projector to a coherent fiber-optic bundle<sup>23</sup>, and then imaged onto back-projection material that constituted the screens. The angular resolution was identical to the previously described rig<sup>23</sup>; each pixel subtended approximately 1°. The luminance in all glider experiments was 6 cd/m<sup>2</sup>; edge selectivity data was taken at 18 cd/m<sup>2</sup>. All flies were tested at a temperature of 34° C, the restrictive temperature for *shibire*<sup>ts</sup><sup>54</sup>.

**Fly Stimuli**—All stimuli were presented with a screen refresh rate of 240 Hz, using custom code to generate stimuli<sup>23</sup>. Pixel intensities were gamma-corrected, and all stimuli were drawn on a virtual cylinder about the fly, using software to perform azimuthal angular corrections. All stimuli were presented both clockwise and counter-clockwise; the response was measured as the difference in response between the two directions. Each glider presentation had a different random seed pattern. Each stimulus was presented more than 30 times, and responses averaged to obtain each fly's response. Presented means and SEMs were computed for the tested flies from each fly's average response to each stimulus.

Binary glider stimuli were created in real time as described in Constructing Gliders (below) and in their original publication<sup>14</sup>. Glider stimuli lasted for 1 second, and were interleaved with 0.5 s periods of uncorrelated updates. All glider pixel updates took place at 40 Hz. The correlations were 1-dimensional, so that each 'pixel' subtended 5° horizontally and the entire extent vertically. The reported response is the integral of the response from 0.5 s to 1 s after stimulus onset. Non-responding flies were excluded from the analysis by requiring that flies exceed a threshold response to at least one of the glider stimuli. That threshold was set at 30°/s during the second 0.5 seconds of the glider stimulus. WT flies respond to the positive parity, 2-point glider with turning rates of 140 ± 50°/s (mean±std), so that this threshold excluded very few flies. In most genotypes, including WT, this procedure excluded no flies at all; it excluded 1 of 20 flies of genotype L1/*shi*<sup>ts</sup> and 2 of 19 of genotype L2+L3/*shi*<sup>ts</sup>. Among three sicklier genotypes, it excluded 5 of 27 (L3/*shi*<sup>ts</sup>), 9 of 23 (L3/+), and 9 of 25 (L2/+). Inclusion of these non-responding flies tended to bring all responses towards 0 and to add variability to the predicted edge selectivity, due to the small denominator in the normalization procedure (see below).

Individual light and dark edge stimuli were created, presented, and analyzed as previously described<sup>23</sup>.

**Fly Metrics & Statistics**—The glider fractional response was computed for each genotype to account for differences in general health, which created variability in the strength of all turning responses. Within a given genotype, all flies' 2- and 3-point stimulus-induced rotations were computed; then each average was divided by the genotype's mean response to the 2-point, parity +1 stimulus to compute the response as a fraction of the genotype's 2-point response.

Edge selectivity (Figure 5) was defined as  $(X_l - X_d)/(X_l + X_d)$ , where  $X_l$  is the experimental light edge response and  $X_d$  is the experimental dark edge response. When either  $X_l$  or  $X_d$  averages to close to zero, instances of this metric can become greater than 1 or less than -1, a property that accounts for the error bars overlapping  $\pm 1$  in Figure 5.  $X_l$  and  $X_d$  were directly computed for the light and dark edge stimuli by integrating the evoked turning response<sup>23</sup>. We generated predictions for  $X_l$  and  $X_d$ , denoted  $\hat{X}_l$  and  $\hat{X}_d$ , from the experimental responses to each glider,

$$\begin{aligned}\hat{X}_l &= 2X_{R+} - 2X_{R-} + X_{D+} - X_{D-} - X_{N+} + X_{N-}, \\ \hat{X}_d &= 2X_{R+} - 2X_{R-} - X_{D+} + X_{D-} + X_{N+} - X_{N-},\end{aligned}$$

where the linear coefficients correspond to the components of each correlation in the edge type (Fig. 5a and S6),  $X_{R+}/X_{R-}$  are the responses to the positive/negative parity 2-point gliders,  $X_{D+}/X_{D-}$  are the responses to the positive/negative parity diverging 3-point gliders, and  $X_{N+}/X_{N-}$  are the responses to the positive/negative parity converging 3-point gliders. The glider predicted edge selectivity was  $(\hat{X}_l - \hat{X}_d)/(\hat{X}_l + \hat{X}_d)$ .

Throughout, the statistical significance of differences were judged by a 2-tailed Student's t-test between experimental and control groups, or between the experiment and null hypothesis for WT data. Significance is reported as the maximum p-value relative to the 2 controls. We did not assess whether the data was normally distributed. We did not use statistical methods to predetermine our sample sizes, but they are comparable to those in the literature<sup>23,24,26,28</sup>.

**Simulated EMD responses to glider stimuli**—Wild-type HRC and ON/OFF model motion detector responses to various gliders (Figs. 3d and S4d) were simulated in Matlab (Natick, MA). The WT HRC parameters were taken directly from a recent study<sup>23</sup>, while the model parameters for the ON/OFF model were taken from its publication<sup>27</sup>. Simulations were run on a 1-dimensional array of 61 photoreceptors, with spacing of  $5.1^\circ$  and Gaussian acceptance angle of  $5.7^\circ$  (FWHM)<sup>55</sup>. Spatial resolution of the simulation was  $1^\circ$ , and the simulation ran in 1ms time steps. Means and standard deviations in figures were computed from 25 instantiations of the gliders. Responses to gliders were computed as the mean response over 2 seconds, excluding an initial second to discard transients. In the case of the ON/OFF model, the response was the mean in the preferred direction minus the mean in the null direction.

## Human EEG Methods

**Subjects**—Informed consent was obtained from the 7 subjects (4 male, 3 female, ages 24–44 years), who participated in the experiments under protocols approved by the Institutional Review Board of Stanford University. We did not use statistical methods to predetermine our sample sizes, but they are comparable to those in the literature<sup>30</sup>.

**Stimulus generation**—Stimulus generation and signal analysis were performed by in-house software, running on a Macintosh G4 platform. Stimuli were presented in a dark and quiet room on a calibrated CRT monitor at a resolution of  $800 \times 600$  pixels viewed from 125 cm (full-width of  $18.0^\circ$ , full-height of  $13.6^\circ$ ), with a 72-Hz vertical refresh rate.

There were two types of adapting stimuli. Both consisted of a  $2^\circ$  period vertical opposing edge presented at 6 Hz (resulting in a velocity of  $12^\circ/\text{s}$ ) with a mean luminance of  $71.5 \text{ cd/m}^2$  and a 97% Michelson contrast. The difference between adapting stimuli was in the motion direction of the bright and dark edges. The probe stimulus comprised a  $4^\circ$  period vertical opposing edge presented at 3 Hz (resulting in a velocity identical to the adapter of  $12^\circ/\text{s}$ ) and was identical for all three conditions (no adaptation, adapter type A, adapter type B). The adapting stimulus was viewed for 20 seconds. Immediately after the adaptation the probe stimulus was displayed and reversed edge polarity directions at 3 Hz for 12 seconds. All stimuli space-time plots are shown in Fig. S7, and movies of adapter A and the probe stimulus are shown in movies M1 and M2.

**Adaptation Procedure**—A session began with a block of 20 probe trials (unadapted) that was followed by 20 trials of an adapt/probe cycle with a randomly chosen adapter (A or B). The adaptation/probe cycle within a block is illustrated in Figure 6a. After a block using the first adapter, the participant was allowed to rest for several minutes in order to dissipate the adaptation effect, after which they were given a block using the alternate adapter.

**Attention Control**—To control for possible time- or stimulus-dependent effects of attention, participants performed a demanding letter discrimination task at fixation during measurement of the steady-state Visual Evoked Potential (ssVEP)<sup>56</sup>.

**Spectral signature of stimulus specific adaptation**—We used a method that isolates responses of directionally selective neurons in the ssVEP<sup>30</sup>. Briefly, in the unadapted state, the ssVEP after each phase of the probe stimulus is identical. After adaptation the responses of individual neurons that are tuned for the adapting stimulus are reduced. The resulting imbalanced, adapted response creates a signal that has a different temporal sequence, either strong/weak or weak/strong, depending on the adapting stimulus. Therefore, the presence of odd harmonic responses that are  $180^\circ$ -phase shifted after adapting to the different adapters is diagnostic of selective adaptation. Because phase is a circular variable we used the circular statistics toolbox (Matlab) to calculate the 95% confidence intervals for the differences in phase<sup>57</sup>.

**EEG signal acquisition and source imaging procedure**—The electroencephalogram (EEG) was recorded, preprocessed, and analyzed identically to prior experiments<sup>58</sup>. The resulting amplitude spectra of the ssVEP were then evaluated at the first

harmonic of the stimulus frequency (3 Hz). Reported averages in Fig. 6 and S7 are from an average of probe responses between 1 and 6 s after the probe began; a time course of the first harmonic response is shown in Fig. S7c.

**Adaptation Index**—The adaptation index provides a measure of the fractional change in the signal due to adaptation. The amplitude of the first harmonic was divided by the summed amplitude of the first and second harmonic, using the single electrode with the maximum response (shown in Fig. 6b). We then averaged the two adapted conditions and divided by the unadapted case.

## Human Psychophysics Methods

**Subjects**—Informed consent was obtained from the 9 subjects (5 male, 4 female, ages 23–55 years), who participated in the experiments under protocols approved by the Institutional Review Board of Stanford University. We did not use statistical methods to predetermine our sample sizes, but they are comparable to those in the literature<sup>14</sup>.

**Stimulus Generation and Presentation**—Dynamic stimuli were presented on a HP-1320 video monitor, with  $800 \times 600$  pixel resolution, 120 Hz refresh rate, mean luminance of  $48.1 \text{ cd/m}^2$ , and contrast of 93.5% (Michelson). A  $5^\circ$ -radius aperture surrounded the fixation cross; stimuli were presented within the aperture and mean luminance outside.

We used four adapting stimuli (Fig S8): static bars, opposing light and dark edges, light edges, and dark edges. The static adapter was a full contrast square wave with a period of  $2^\circ$ . The opposing edges adapter had the same parameters as the adapter in the EEG experiments:  $2^\circ$  period and  $12^\circ/\text{s}$  edge speeds. The single edges consisted of periodic single edges moving across the screen in one direction, repeated in time (see Fig S8). The single edges had identical parameters to the opposing edge stimuli, moving at a speed of  $12^\circ/\text{s}$  and with a period of  $2^\circ$ .

We created a total of 8 3-point glider stimuli: converging/diverging,  $\pm 1$  parity, and centroid moving left/right. The gliders were full contrast, updated at 30 Hz, and consisted of square pixels of  $0.1^\circ$  on a side, updated using the glider update rules to constrain the contrast correlations appropriately in each case (see Constructing Gliders, below). Each row of the glider stimulus was independent of the others, but updated with the identical update rule (see movie M3).

**Adaptation Protocols**—An initial adaptation phase lasted for the first 20 seconds of each experiment. After the initial adaptation, probe glider stimuli were shown in random order for 0.65 seconds, interleaved with 3-second adaptation ‘top ups’, repeating the initial adapter. All adapters were tested along with their mirror-symmetric counterpart, and the results averaged with appropriate sign changes. Each glider was tested 20 times for each adapter, for a total of 40 trials for each subject in the points shown in Fig. 7. The static case contained 20–60 trials for each subject, as it was occasionally run more than once if the subject participated in multiple sessions.



**Data analysis**—For each adaptation and glider condition, the fraction of trials that were perceived as moving to the right was computed for each subject. Means and standard errors presented in the Figure were computed from the cross-subject data, and significance between conditions was assessed with a 2-sample, unpaired Student’s t-test, Bonferoni-corrected. (The bar plots in Fig. 7 represent a linear transform of fraction-to-the-right. 0 maps to all-left, 0.5 maps to random, and 1 maps to all-right.) We did not assess whether data was normally distributed.

### Constructing gliders

Gliders are binary spatiotemporal stimuli with enforced correlational structures that were previously developed and used to produce motion percepts in humans<sup>14</sup>. These authors describe several properties that make gliders ideal for probing mechanisms of motion detection. First, they are, on average, equiluminant in time and space, averaging to a mean gray. Second, the variance at each time and at each point in space is the same. Third, the enforced correlational structure excludes many other correlational structures. For instance, the four 3-point gliders we produce here exclude all 2-point correlations, and exclude all other 3-point correlations. Gliders can produce more complex, higher-order or more distant correlations<sup>14</sup>.

Each glider stimulus begins with a random seed and uses an update rule to determine whether a pixel should be white or black. It iterates the rule across all pixels to produce a full space-time pattern that obeys the correlational pattern determined by the update rule. Figure S4a shows examples of each update rule, and Figure S4c gives examples of each rule’s space-time patterns. Movie M3 shows 4 gliders similar to those used in the human psychophysics experiments.

The simplest update rule is uncorrelated, when each pixel is randomly chosen to be black or white, independent of all other pixels’ values. The second simplest pattern is a 2-point update rule. We coded white pixels as having value +1 and black as having –1, matching our standard definition of fractional contrast. The update rule is then  $C_i(t)C_{i+}(t + \delta) = P$ , where  $i$  is the pixel spacing,  $\delta$  is the frame duration, and  $P$  is the parity of the pattern, which is equal to +1 for even parity or –1 for odd parity. This is equivalently written as  $C_{i+}(t + \delta) = PC_i(t)$ , since  $C = 1/C$  for our values of  $C$ . Each new pixel value is the previous time’s adjacent pixel multiplied by 1 or –1. For even parity, this rule translate the entire pixel pattern in one direction, while for odd parity, it translates and inverts on each update.

The 3-point glider rules are a generalized version of the 2-point rules. The update rule is  $C_i(t)C_{i+}(t)C_{i++}(t + \delta) = P$  (converging case) and  $C_i(t)C_i(t + \delta)C_{i+}(t + \delta) = P$  (diverging case). Each pixel’s value is updated as a function of its surrounding pixels’ values and is determined by the seed state, so these patterns are similar to Conway’s Game of Life, the origin for the term “glider”<sup>59</sup>. Edge cases can result in undetermined pixel values; in such cases, we seeded with random pixel contrasts. In 3-point gliders, the two different parities are contrast inversions of each other: inverting the contrast on a  $P = 1$  pattern turns it into a  $P = -1$  pattern (Figure S4b). Conversely, inverting the contrast of a 2-point glider does not alter its parity, since all pairwise products remain unchanged.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

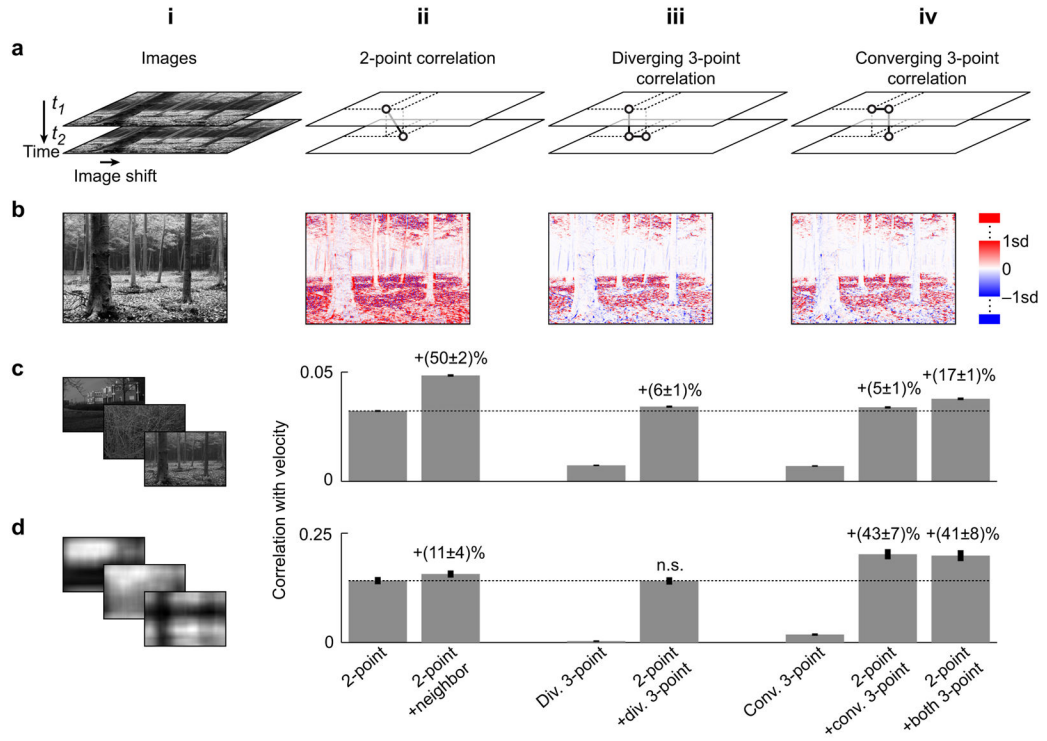
We thank Johannes Burge, Preeti Varghese, Brian Wandell, and the Clandinin lab for helpful comments on this manuscript. DAC was supported by an NIH T32 Vision Research Training Grant and a postdoctoral fellowship from the Jane Coffin Childs Foundation; JEF is supported by an NSF Graduate Research Fellowship and by NSF-0801700; JMA and AMN are supported by NIH-EY015790. DMG was supported by a post-doctoral fellowship from the NIH, and MAS was supported by a postdoctoral fellowship from the Jane Coffin Childs Foundation. In TRC's lab, this work was supported by an NIH Director's Pioneer Award (DP1 OD003530), and by R01 EY022638.

## References

1. Field DJ. Relations between the statistics of natural images and the response properties of cortical cells. *J Opt Soc Am A*. 1987; 4:2379–2394. [PubMed: 3430225]
2. Ruderman DL, Bialek W. Statistics of natural images: Scaling in the woods. *Physical Review Letters*. 1994; 73:814–817. [PubMed: 10057546]
3. Simoncelli EP, Olshausen BA. Natural image statistics and neural representation. *Annual Review of Neuroscience*. 2001; 24:1193–1216.
4. Hassenstein B, Reichardt W. Systemtheoretische Analyse der Zeit-, Reihenfolgen- und Vorzeichenauswertung bei der Bewegungspertzeption des Rüsselkäfers *Chlorophanus*. *Zeitschrift für Naturforschung*. 1956; 11:513–524.
5. Adelson E, Bergen J. Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*. 1985; 2:284–299.
6. Fitzgerald JE, Katsov AY, Clandinin TR, Schnitzer MJ. Symmetries in stimulus statistics shape the form of visual motion estimators. *Proceedings of the National Academy of Sciences*. 2011; 108:12909–12914.
7. Anderson JMM, Giannakis GB. Image motion estimation algorithms using cumulants. *Image Processing, IEEE Transactions on*. 1995; 4:346–357.
8. Sayrol E, Gasull A, Fonollosa JR. Motion estimation using higher order statistics. *Image Processing, IEEE Transactions on*. 1996; 5:1077–1084.
9. Rust NC, Mante V, Simoncelli EP, Movshon JA. How MT cells analyze the motion of visual patterns. *Nature Neuroscience*. 2006; 9:1421–1431. [PubMed: 17041595]
10. Borst A, Haag J, Reiff DF. Fly motion vision. *Annual Review of Neuroscience*. 2010; 33:49–70.
11. Van Santen JPH, Sperling G. Elaborated reichardt detectors. *J Opt Soc Am A*. 1985; 2:300–320. [PubMed: 3973763]
12. Theobald JC, Duistermars BJ, Ringach DL, Frye MA. Flies see second-order motion. *Current Biology*. 2008; 18:R464–R465. [PubMed: 18522814]
13. Zanker JM. Theta motion: a paradoxical stimulus to explore higher order motion extraction. *Vision Research*. 1993; 33:553–569. [PubMed: 8503201]
14. Hu Q, Victor JD. A set of high-order spatiotemporal stimuli that elicit motion and reverse-phi percepts. *Journal of vision*. 2010; 10
15. Lu ZL, Sperling G. Three-systems theory of human visual motion perception: review and update. *JOSA A*. 2001; 18:2331–2370. [PubMed: 11551067]
16. Victor JD, Conte MM. Evoked potential and psychophysical analysis of Fourier and non-Fourier motion mechanisms. *Visual Neuroscience*. 1992; 9:105–123. [PubMed: 1504020]
17. Geisler WS. Visual perception and the statistical properties of natural scenes. *Annu Rev Psychol*. 2008; 59:167–192. [PubMed: 17705683]
18. Dror RO, O'Carroll DC, Laughlin SB. Accuracy of velocity estimation by Reichardt correlators. *JOSA A*. 2001; 18:241–252. [PubMed: 11205969]

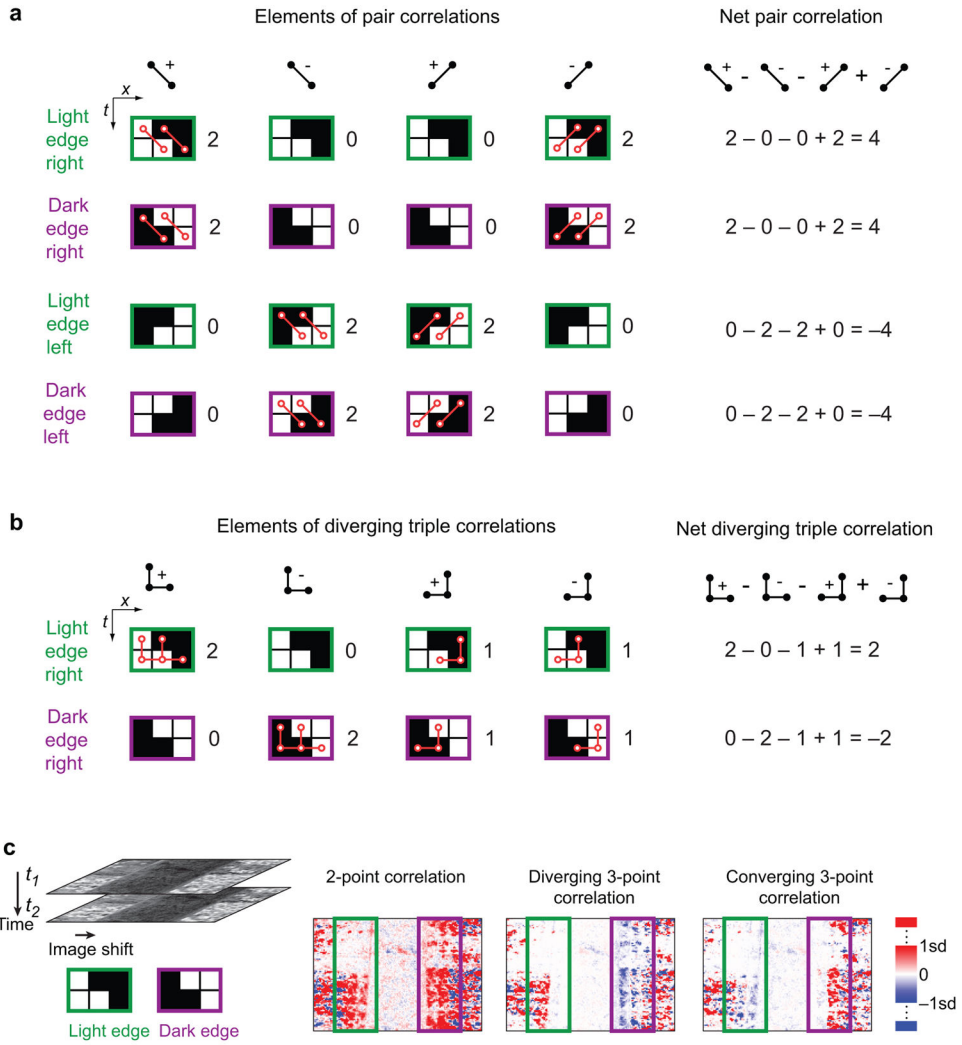
19. van Hateren JH, van der Schaaf A. Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society of London. Series B: Biological Sciences*. 1998; 265:359–366. [PubMed: 9523437]
20. Stocker AA, Simoncelli EP. Noise characteristics and prior expectations in human visual speed perception. *Nature Neuroscience*. 2006; 9:578–585. [PubMed: 16547513]
21. Katsov A, Clandinin T. Motion processing streams in *Drosophila* are behaviorally specialized. *Neuron*. 2008; 59:322–335. [PubMed: 18667159]
22. Joesch M, Schnell B, Raghu S, Reiff D, Borst A. ON and OFF pathways in *Drosophila* motion vision. *Nature*. 2010; 468:300–304. [PubMed: 21068841]
23. Clark DA, Bursztyn L, Horowitz MA, Schnitzer MJ, Clandinin TR. Defining the Computational Structure of the Motion Detector in *Drosophila*. *Neuron*. 2011; 70:1165–1177. [PubMed: 21689602]
24. Silies M, et al. Modular use of peripheral input channels tunes motion-detecting circuitry. *Neuron*. 2013; 79:111–127. [PubMed: 23849199]
25. Buchner E. Elementary movement detectors in an insect visual system. *Biological Cybernetics*. 1976; 24:85–101.
26. Tuthill JC, Chiappe ME, Reiser MB. Neural correlates of illusory motion perception in *Drosophila*. *Proceedings of the National Academy of Sciences*. 2011; 108:9685–9690.
27. Eichner H, Joesch M, Schnell B, Reiff DF, Borst A. Internal structure of the fly elementary motion detector. *Neuron*. 2011; 70:1155–1164. [PubMed: 21689601]
28. Tuthill JC, Nern A, Holtz SL, Rubin GM, Reiser MB. Contributions of the 12 neuron classes in the fly lamina to motion vision. *Neuron*. 2013; 79:128–140. [PubMed: 23849200]
29. Westheimer G. The ON-OFF dichotomy in visual processing: From receptors to perception. *Progress in retinal and eye research*. 2007; 26:636–648. [PubMed: 17851105]
30. Ales JM, Norcia AM. Assessing direction-specific adaptation using the steady-state visual evoked potential: Results from EEG source imaging. *Journal of vision*. 2009; 9
31. Moulden B, Begg H. Some tests of the Marr-Ullman model of movement detection. *Perception*. 1986; 15:139. [PubMed: 3774485]
32. Mather G, Moulden B, O'Halloran A. Polarity specific adaptation to motion in the human visual system. *Vision Research*. 1991; 31:1013–1019. [PubMed: 1858317]
33. Emerson RC, Bergen JR, Adelson EH. Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Research*. 1992; 32:203–218. [PubMed: 1574836]
34. Nishida, Sy. Advancement of motion psychophysics: Review 2001–2010. *Journal of vision*. 2011; 11
35. Quenzer T, Zanker J. Visual detection of paradoxical motion in flies. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*. 1991; 169:331–340.
36. Ilg UJ, Churan J. Motion perception without explicit activity in areas MT and MST. *Journal of Neurophysiology*. 2004; 92:1512–1523. [PubMed: 15084645]
37. Hedges JH, et al. Dissociation of neuronal and psychophysical responses to local and global motion. *Current Biology*. 2011; 21:2023–2028. [PubMed: 22153156]
38. Ratliff CP, Borghuis BG, Kao YH, Sterling P, Balasubramanian V. Retina is structured to process an excess of darkness in natural scenes. *Proceedings of the National Academy of Sciences*. 2010; 107:17368–17373.
39. Barlow H, Levick WR. The mechanism of directionally selective units in rabbit's retina. *The Journal of Physiology*. 1965; 178:477. [PubMed: 5827909]
40. Ibn-Elhaj E, Aboutajdine D, Pateux S, Morin L. HOS-based method of global motion estimation for noisy image sequences. *Electronics Letters*. 1999; 35:1320–1322.
41. Schiller PH, Finlay BL, Volman SF. Quantitative studies of single-cell properties in monkey striate cortex. I. Spatiotemporal organization of receptive fields. *Journal of Neurophysiology*. 1976; 39:1288. [PubMed: 825621]
42. Potetz B, Lee TS. Statistical correlations between two-dimensional images and three-dimensional structures in natural scenes. *JOSA A*. 2003; 20:1292–1303. [PubMed: 12868635]

43. Wiederman SD, Shoemaker PA, O'Carroll DC. A model for the detection of moving targets in visual clutter inspired by insect physiology. *PLoS one*. 2008; 3:e2784. [PubMed: 18665213]
44. Poggio T, Reichardt W. Visual control of orientation behaviour in the fly. *Quarterly Reviews of Biophysics*. 1976; 9:377–438. [PubMed: 790442]
45. Card G, Dickinson MH. Visually Mediated Motor Planning in the Escape Response of *Drosophila*. *Current Biology*. 2008; 18:1300–1307. [PubMed: 18760606]
46. Harris JM, Parker AJ. Independent neural mechanisms for bright and dark information in binocular stereopsis. 1995
47. Samonds JM, Potetz BR, Lee TS. Relative luminance and binocular disparity preferences are correlated in macaque primary visual cortex, matching natural scene statistics. *Proceedings of the National Academy of Sciences*. 2012; 109:6313–6318.
48. Bialek W. Physical limits to sensation and perception. *Annual review of biophysics and biophysical chemistry*. 1987; 16:455–478.
49. Laughlin SB. Energy as a constraint on the coding and processing of sensory information. *Current Opinion in Neurobiology*. 2001; 11:475–480. [PubMed: 11502395]
50. Laughlin SB, Sejnowski TJ. Communication in neuronal networks. *Science*. 2003; 301:1870–1874. [PubMed: 14512617]
51. Potters M, Bialek W. Statistical mechanics and visual signal processing. *Journal de Physique I*. 1994; 4:1755–1775.
52. Juusola M, Uusitalo R, Weckström M. Transfer of graded potentials at the photoreceptor-interneuron synapse. *The Journal of General Physiology*. 1995; 105:117. [PubMed: 7537323]
53. Gohl DM, et al. A versatile in vivo system for directed dissection of gene expression patterns. *Nature Methods*. 2011; 8:231–237. [PubMed: 21473015]
54. Kitamoto T. Conditional modification of behavior in *Drosophila* by targeted expression of a temperature-sensitive shibire allele in defined neurons. *Journal of neurobiology*. 2001; 47:81–92. [PubMed: 11291099]
55. Stavenga D. Angular and spectral sensitivity of fly photoreceptors. II. Dependence on facet lens F-number and rhabdomere type in *Drosophila*. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*. 2003; 189:189–202.
56. Palomares M, Ales JM, Wade AR, Cottreau BR, Norcia AM. Distinct effects of attention on the neural responses to form and motion processing: A SSVEP source-imaging study. *Journal of vision*. 2012; 12
57. Berens P. CircStat: A MATLAB Toolbox for Circular Statistics. *Journal of Statistical Software*. 2009; 31:1–21.
58. Ales JM, Farzin F, Rossion B, Norcia AM. An objective method for measuring face detection thresholds using the sweep steady-state visual evoked response. *Journal of vision*. 2012; 12
59. Gardner M. Mathematical games: The fantastic combinations of John Conway's new solitaire game, 'life'. *Scientific American*. 1970; 223:120–123.
60. Rister J, et al. Dissection of the peripheral motion channel in the visual system of *Drosophila melanogaster*. *Neuron*. 2007; 56:155–170. [PubMed: 17920022]
61. Gao S, et al. The neural substrate of spectral preference in *Drosophila*. *Neuron*. 2008; 60:328–342. [PubMed: 18957224]



**Figure 1.**

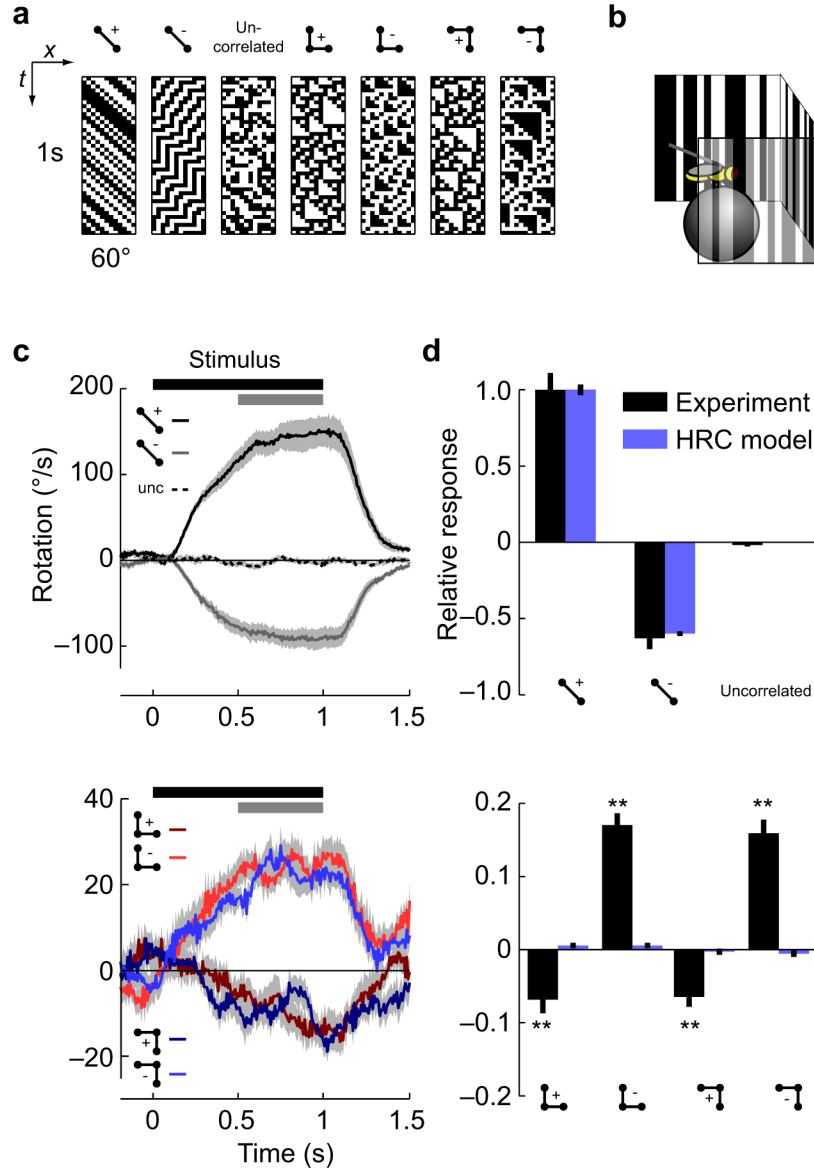
Multiple correlations signify natural image motion. Each row presents a comparison between correlational motion signatures. Columns present: (i) context for each comparison; (ii) properties of pairwise motion estimators; (iii) properties of diverging 3-point estimators; and (iv) properties of converging 3-point estimators. (ai) Motion is approximated by the rigid translation of natural images. (aii–aiv) Cartoon of the correlation structure that each estimator detects. (bi) Example natural image. (bii–biv) Pixelwise contributions to motion estimation are highly variable and differ across estimators. (ci) An ensemble of natural images. (cii–civ) The accuracy with which correlations convey motion is examined across this ensemble. The performance of each estimator is quantified through the Pearson’s correlation between the estimator output and the simulated velocity. We linearly combined estimators to quantify the improvements afforded by multiple correlational signals. The numbers above each bar denote the fractional increase with respect to the 2-point estimate. (d) Same as (c), but with signals spatiotemporally filtered to match motion processing in *Drosophila*. Error bars are standard deviations over cross-validating trials (see Online Methods).



**Figure 2.**

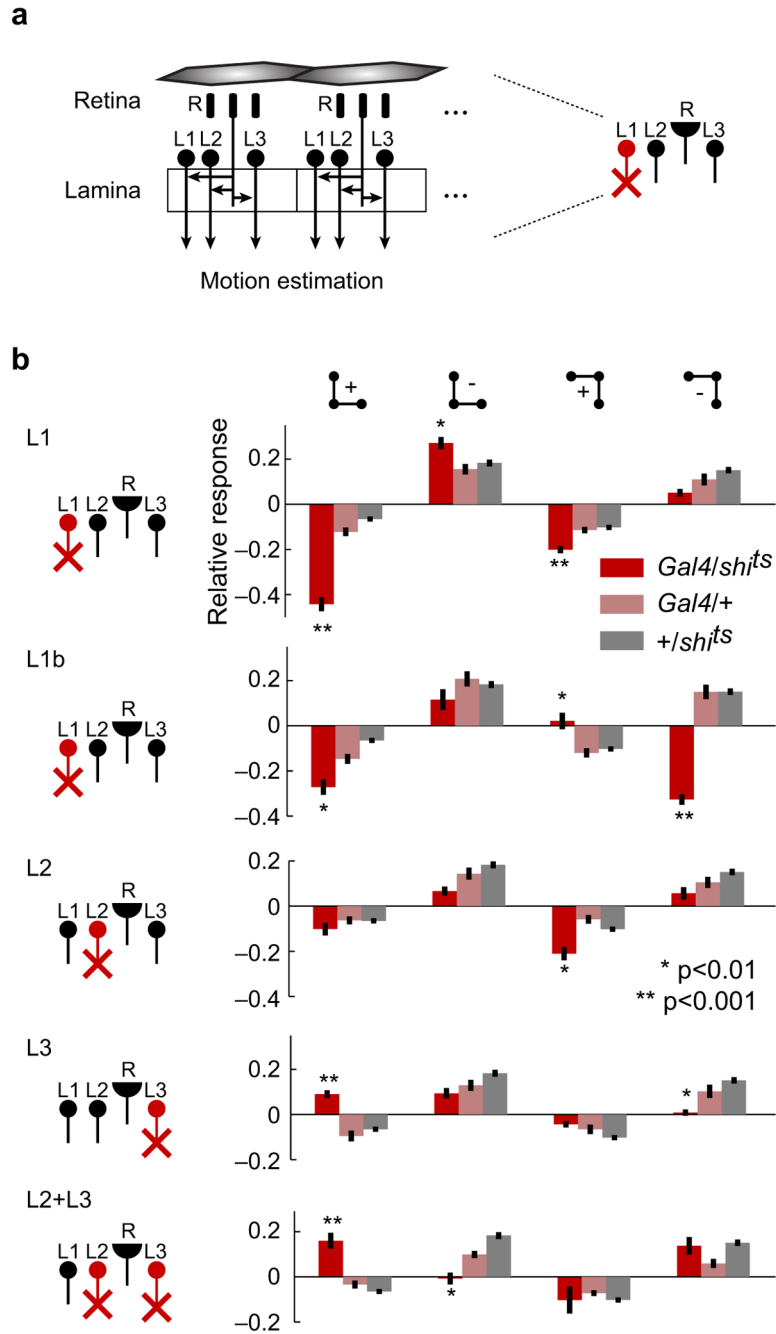
Triple correlations distinguish between light and dark moving edges. (a) A dark point becomes light when a light edge moves across the visual field (rows 1 and 3), and a light point becomes dark when a dark edge moves across the visual field (rows 2 and 4). We decompose the net pair correlation motion signal into four elements whose frequency of occurrence depends upon the motion. This net pair correlation motion signal reflects the direction of motion (compare rows 1 and 2 to rows 3 and 4) and is insensitive to whether the edge was light or dark (compare row 1 to row 2 or row 3 to row 4). (b) We similarly decompose the net diverging and converging triple correlation into four elements (shown for the diverging triple correlation). The sign of the net diverging triple correlation depends both on the contrast polarity of the edge and on the direction of motion (shown for rightward motion). Thus, triple correlations jointly encode the direction and contrast polarity of a moving edge. (c) Natural motion comprises both moving light edges and moving dark edges. Motion signals are associated with each moving edge, but only the 3-point motion signatures distinguish between edge contrast polarities.





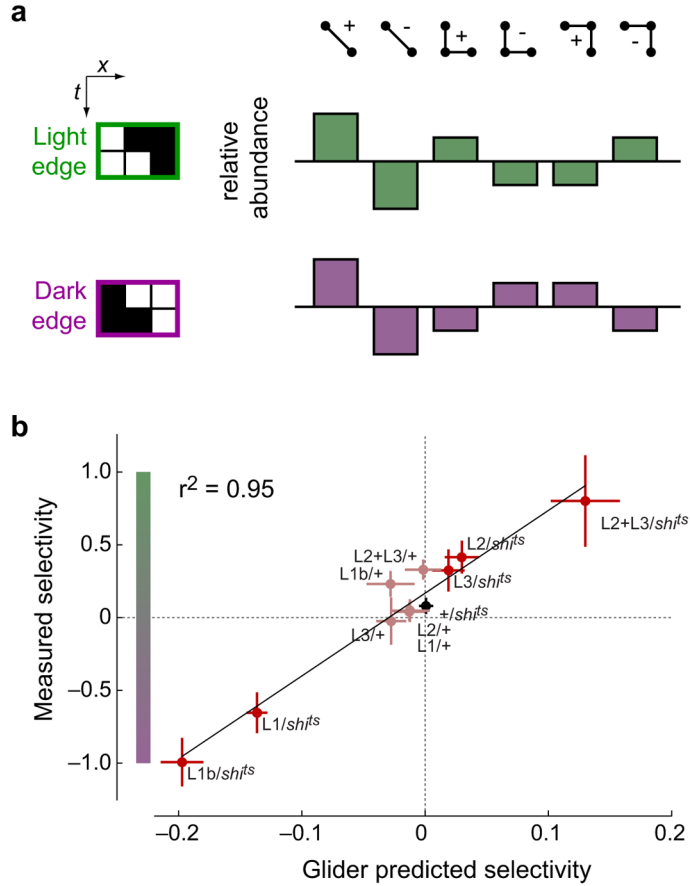
**Figure 3.** *Drosophila* responds to triple correlations. (a) Binary spatiotemporal patterns, glider stimuli with 2- and 3-point contrast correlations, were presented to flies. Space-time plots for each of the 6 gliders, and an uncorrelated stimulus, are shown. (b) During the presentation, we measured flies' turning in response to each glider. Positive rotational velocities represent turning in the direction of the 'centroid' of the pattern (to the right in the space-time plots in (a)). (c) One second periods of glider stimuli were interleaved with uncorrelated stimuli; the timing of the presentation of the gliders is denoted by the thick black bar. Response curves show the mean (solid line) and SEM (shading) over flies. (d) Mean turning velocities were computed for each glider by averaging over 0.5s of the stimulus (gray bar in (c)). Turning responses are presented for wild-type *Drosophila*, alongside the predicted response of a Hassenstein-Reichardt Correlator (HRC) to each glider. N=12 in (c) and (d). '\*\*\*' denotes a

difference from 0 at the  $p < 0.01$  level (two-tailed t-test); from right to left, the marked p-values are  $4.4 \times 10^{-3}$  ( $t_{11}=3.6$ ),  $6.0 \times 10^{-7}$  ( $t_{11}=10.2$ ),  $8.2 \times 10^{-4}$  ( $t_{11}=4.6$ ), and  $3.7 \times 10^{-6}$  ( $t_{11}=8.5$ ). Error bars show SEM.



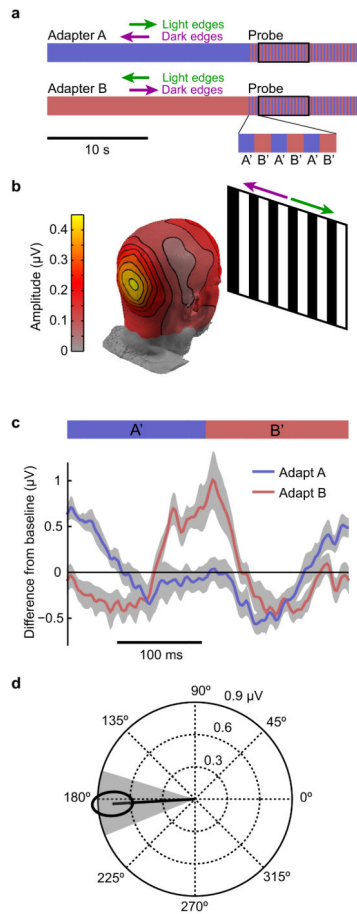
**Figure 4.** Detection of triple correlations associated with specific pathways in *Drosophila*. (a) Left: schematic of the inputs to the fly motion processing pathways. Signals from photoreceptors are relayed through the lamina monopolar cells L1, L2, and L3. Right: a temperature-inducible dominant negative suppressor of synaptic transmission (*shi<sup>ts</sup>*) was used to silence L1, L2 and L3 using cell-specific expression of Gal4 (L1 shown, red). (b) We examined the responses of these disrupted motion detectors to 3-point gliders. Responses are plotted relative to the 2-point positive glider response. The two control genotypes (*Gal4/+* and *+/<sup>ts</sup>*

*shi<sup>ts</sup>*) have all input pathways intact, but contain the genetic constructs for the experimental genotype (*Gal4/shi<sup>ts</sup>*). For the genotypes *Gal4/shi<sup>ts</sup>*, *Gal4/+*, *+/shi<sup>ts</sup>*, from top to bottom, N = (19, 14, 19), (18, 13, 19), (29, 16, 19), (22, 14, 19), and (17, 15, 19). Error bars are  $\pm$  SEM. ‘\*’ and ‘\*\*\*’ represent  $p < 0.01$  and  $p < 0.001$  differences from both control genotypes (two-tailed t-test).



**Figure 5.**

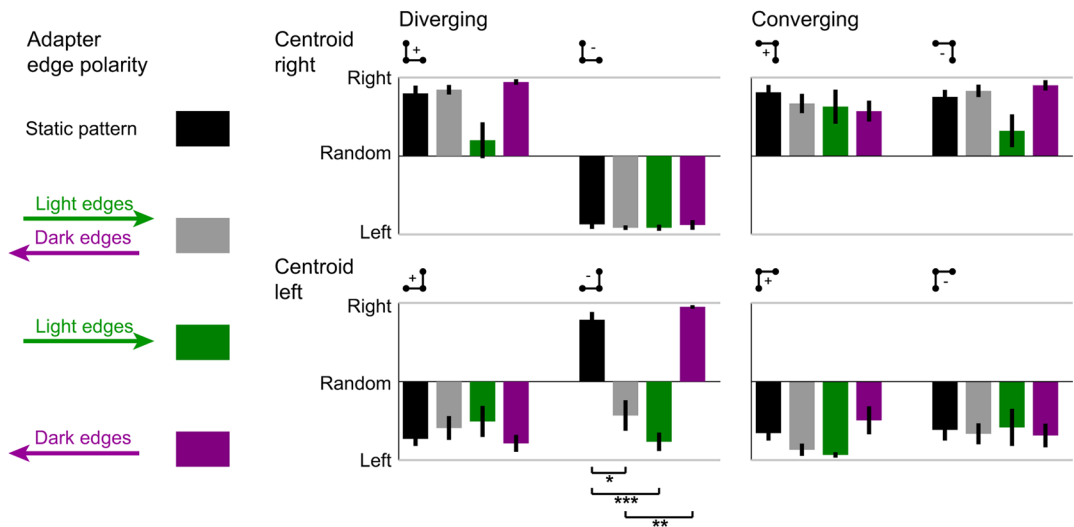
Triple correlations predict the edge selectivity of motion pathways. (a) The frequency of correlational elements in a moving edge depends on its contrast polarity and direction (see Fig. 2), and we compute the relative abundance of each correlation from the difference in frequency of each element in rightward *versus* leftward motion (see Fig. S6). The relative abundances of the four triple correlation elements differ between light and dark edges. (b) We used the relative abundance of each correlational element in each edge type (see also Fig. 2, S6) to weight and sum the response of each genotype to each correlational element (Fig. 4). This generated the glider-predicted responses to each edge type, from which we computed the predicted edge selectivity for each genotype. It correlated highly with the behaviorally measured edge selectivity (see Online Methods). Edge selectivity is computed to be the light minus dark edge responses divided by their sum. Error bars on points are  $\pm$  SEM.



**Figure 6.**

Humans differentially adapt to moving light and dark edges. (a) Schematic of adapter and probe stimulus paradigm (see Figure S4). Black box denotes the time interval used for analysis. (b) Scalp topography of the amplitude of the response at the A'/B' alternation rate (3 Hz). The amplitude peaks near the occipital pole. (c) Time average of the response from the peak electrode to the probe stimulus under the two adaptation regimes. Response to the unadapted state obtained by probe presentation without the adapting stimulus has been subtracted from this signal (see Figure S4). The response to the probe shows complementary modulation by the adapting stimuli at the frequency of probe alternation (3 Hz). Gray area represents  $\pm 1$  SEM. (d) The within subject difference of phase and amplitude at 3 Hz between the two adapting conditions. Ellipse represents 1 SEM while the shaded wedge indicates the 95% confidence interval for the phase.  $N = 7$  subjects in (c) and (d).





**Figure 7.**

Adaptation to moving light and dark edges differentially affects the perception of specific 3-point gliders. Subjects were presented all combinations of four types of adapter stimuli (left column), and 8 gliders (right hand panels), and asked to report the direction of perceived glider motion. Results for each of the 8 glider stimuli are shown, grouped by glider. The color of the bar corresponds to the adapting stimulus: static (black), opposing edges (gray), light edges only (green), or dark edges only (magenta). All stimuli were presented mirror-symmetrically, and responses were aligned to the direction shown in the left hand column. ‘\*’, ‘\*\*’, and ‘\*\*\*’ indicate differences between conditions at  $p=1.6 \times 10^{-3}$  ( $t_{16}=5.6$ ),  $p=8.8 \times 10^{-4}$  ( $t_{14}=6.2$ ), and  $p=2.8 \times 10^{-6}$  ( $t_{14}=10.2$ ) (two-tailed t-test, Bonferroni corrected for 40 comparisons). N=9 subjects for static and opposing edge adaptation; N=7 for light and dark edge adaptation conditions.