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

Damon A Clark[1,4,5], James E Fitzgerald[2,4,5], Justin M Ales[3–5], Daryl M Gohl[1], Marion A Silies[1], Anthony M Norcia[3] & Thomas R Clandinin[1]

**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. We found 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 extracted triple correlations with neural substrates tuned for light or dark edges, and sensitivity to specific triple correlations was retained even as light and dark edge motion signals were combined. Thus, both species separately process light and dark image contrasts to capture motion signatures that can improve estimation accuracy. This convergence argues that statistical structures in natural scenes have greatly affected visual processing, driving a common computational strategy over 500 million years of evolution.**



The statistical distribution of light intensities across space is a core feature of any environment[1–3]. These spatial distributions can be sampled over time to extract information about visual motion, a critical behavioral cue for many animals. Prominent computational models of motion processing estimate motion by correlating light intensity between pairs of points separated in space and time or, equivalently, by measuring local motion energy[4,5]. 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[6–8]. For example, contrast polarity, being either dark or light, is a fundamental feature of moving edges, yet it is explicitly discarded by pair correlations. We found 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 point than the other. These products are then summed in an antisymmetric fashion to produce an average signal whose sign and amplitude indicates the direction and magnitude of motion (**Supplementary Fig. 1**)[4]. Motion energy, a second correlational model, begins with linear, oriented spatiotemporal receptive fields that are sensitive to particular directions of motion (**Supplementary Fig. 1**)[5]. Subsequent circuit operations square and then sum these responses to produce a motion signal. On the basis of neural and behavioral measurements, motion energy models have been favored in vertebrates[9], whereas HRC models have been favored in invertebrates[10]. Nonetheless, the two models are sometimes mathematically equivalent[5,11], and both ultimately compute correlations only between pairs of points in space and time (see **Supplementary Modeling**).

Experiments using a variety of artificial stimuli have demonstrated that both vertebrates and invertebrates sometimes detect motion even when there are no systematic correlations in intensity between pairs of points[12–16]. 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[6]. In particular, analysis of optimal estimators suggests that natural luminance asymmetries[2,17] would allow animals to estimate motion using triple correlations[6]. We found 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. Our 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 determine how the motion of natural scenes generates spatiotemporal correlations, we approximated full-field motion by rigid translations of natural images (**Fig. 1a**,**b**, column 1)[19]. Minimal motion energy and HRC-based models rely exclusively on information extracted from pairwise correlations across the images. One simple example of this correlation structure takes the difference between rightward

[1]Department of Neurobiology, Stanford University, Stanford, California, USA. [2]Department of Physics, Stanford University, Stanford, California, USA. [3]Department of Psychology, Stanford University, Stanford, California, USA. [4]Present addresses: Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut, USA (D.A.C.), Center for Brain Science, Harvard University, Cambridge, Massachusetts, USA (J.E.F.), School of Psychology and Neuroscience, University of St Andrews, St Andrews, UK (J.M.A.). [5]These authors contributed equally to this work. Correspondence should be addressed to D.A.C. (damon.clark@yale.edu), J.E.F. (jamesfitzgerald@fas.harvard.edu) or J.M.A. (jma23@st-andrews.ac.uk).

Received 12 August 2013; accepted 14 November 2013; published online 5 January 2014; doi:10.1038/nn.3600

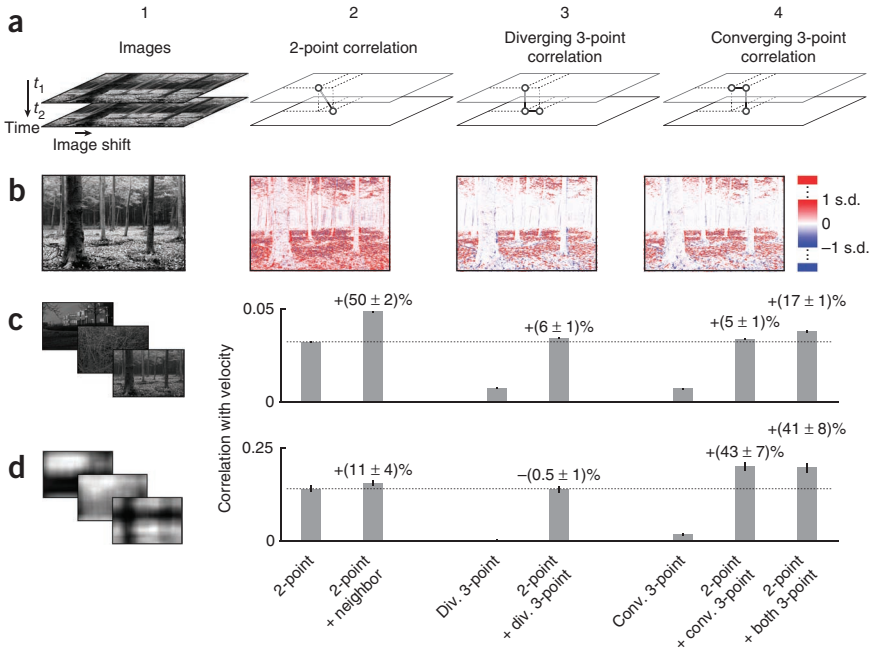



**Figure 1** Multiple correlations signify natural image motion. Each row presents a comparison between correlational motion signatures. Columns present context for each comparison (1), properties of pairwise motion estimators (2), properties of diverging three-point estimators (3) and properties of converging three-point estimators (4). (**a**) Motion is approximated by the rigid translation of natural images (1). Cartoons of the correlation structure that each estimator detects are shown (2–4). (**b**) Example natural image (1)[19]. Pixel-wise contributions to motion estimation were highly variable and differed across estimators (2–4). Red and blue pixels indicate opposite directions of motion. (**c**) Examples from an ensemble of natural images (1). The accuracy with which correlations convey motion was examined across this ensemble (2–4). The performance of each estimator was quantified through the Pearson's correlation coefficient 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 increases with respect to the two-point estimate. (**d**) Data presented as in **c**, but with signals spatiotemporally filtered to match motion processing in *Drosophila*. Error bars represent s.d. over cross-validating trials (Online Methods).



and leftward correlation signals (**Fig. 1a**, column 2). In this case, the net local correlation signal, on average, correctly indicated rightward motion, but, as a result of the variability in the image[18], this signal also suggested leftward motion in some regions (**Fig. 1b**, column 2). In this example, the standard deviation (s.d.) of the local motion signal, computed across pixels, was 3.6-fold greater than the mean. Spatiotemporal averaging can suppress this variability[18], but at the expense of resolution.

We considered two triple correlation structures involving three points in space and time, which we refer to as diverging (**Fig. 1a**, column 3) and converging (**Fig. 1a**, column 4). The diverging case incorporates two spatial points at the later time (one point diverges into two points), whereas 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. Similar to the pairwise estimator, these triple correlation signals were highly variable across the image, but their average signified the direction of motion (**Fig. 1b**, columns 3 and 4). Notably, the motion signals provided by these triple correlations incompletely overlapped with the motion signals derived from pair correlations (**Fig. 1b**, columns 2–4). Thus, three-point motion signals provide additional information about motion, beyond what can be obtained from the pairwise signal[6].

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[19] to determine how well pair and triple correlations can predict image velocity (**Fig. 1c**,**d**). We considered motion estimators with human-like spatial resolution (**Fig. 1c**, column 1, and Online Methods) and with *Drosophila*-like spatiotemporal sampling (**Fig. 1d**, column 1, **Supplementary Fig. 1** and Online Methods). In both cases, we approximated the distribution of image velocities with a Gaussian having zero mean, using s.d. that were 5° $s^{-1}$ and 90° $s^{-1}$, respectively, comparable to estimated natural speeds[20,21]. The output of local pairwise motion estimators was weakly correlated with the image velocity (**Fig. 1c**,**d**, column 2). This correlation was improved by averaging two pairwise estimators that survey neighboring spatial points (**Fig. 1c**,**d**, column 2). The diverging and converging triple correlations were more weakly correlated with the velocity, but typically improved motion estimation when summed with the pairwise estimate using optimal weighting coefficients (**Fig. 1c**,**d**, columns 3 and 4). For *Drosophila*-like sampling, the increased accuracy afforded by the converging three-point motion estimator (**Fig. 1d**, column 4) exceeded that afforded by a neighboring pairwise estimator (**Fig. 1d**, column 2). Because these two three-point estimators sampled the same two spatial and temporal points as the pairwise estimator, this improvement was available without sacrificing either spatial or temporal precision. Notably, these triple correlations required asymmetric contrast distributions for their functionality[6] (**Supplementary Fig. 2**), thereby capitalizing on the strong asymmetries present in natural contrast distributions that are absent from many laboratory stimuli (**Supplementary Figs. 2** and **3**).

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) or decreases (dark edge) 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 **Supplementary Modeling**). Conversely, three-point estimators do capture edge polarity information. The diverging three-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. 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 deter-

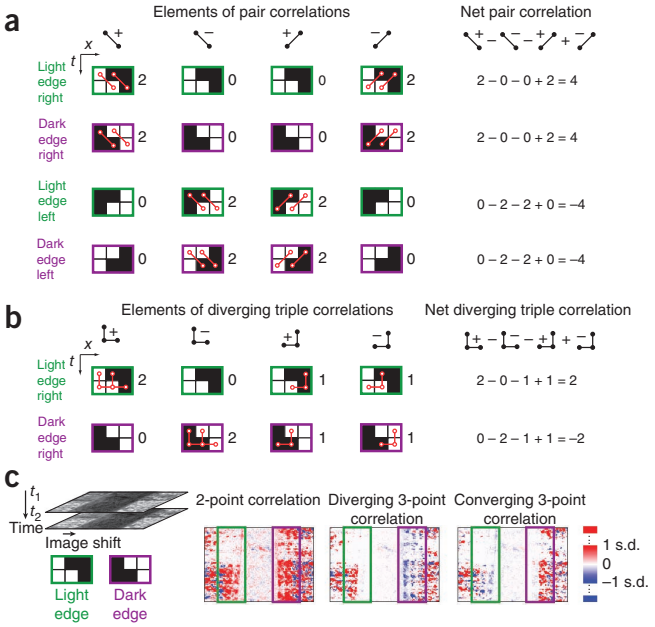
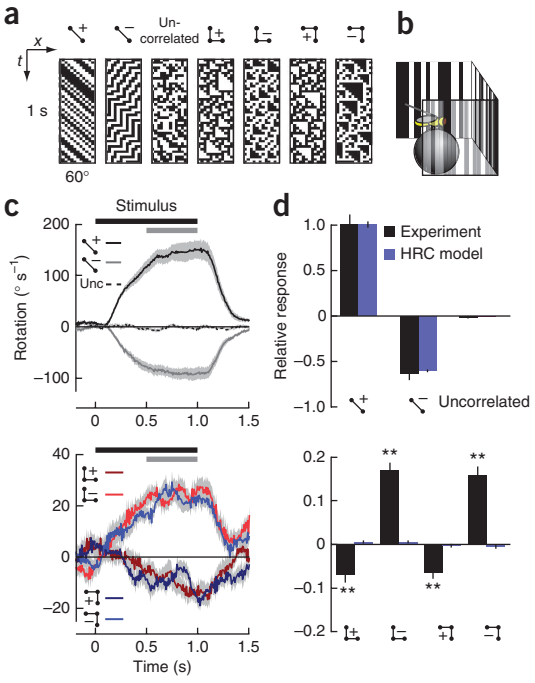



**Figure 2** Triple correlations distinguish between light and dark moving edges. (**a**) A dark point becomes light when a light edge moves across the visual field (rows 1 and 3), and a light point becomes dark when a dark edge moves across the visual field (rows 2 and 4). We decomposed the net pair correlation motion signal into four elements whose frequency of occurrence depends on the motion. This net pair correlation motion signal reflects the direction of motion (compare rows 1 and 2 with rows 3 and 4) and is insensitive to whether the edge was light or dark (compare row 1 with row 2 or row 3 with row 4). (**b**) We similarly decomposed the net diverging and converging triple correlations 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 comprised both moving light edges and moving dark edges. Motion signals were associated with each moving edge, but only the three-point motion signatures distinguished between edge contrast polarities.

mined. 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 (**Fig. 1**) precisely because, in natural scenes, light and dark moving edges generate asymmetric triple correlation signals (see **Supplementary Modeling**).

The early insect visual system contains distinct substrates specialized for detecting moving light and dark edges[22,23]. Because triple correlations can improve motion estimation accuracy and discriminate between the motion of light and dark edges, we next determined whether flies actually detect these signals. Following previous psychophysical approaches[14], 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 **Supplementary Fig. 4**). Notably, the three-point gliders contained no net two-point correlations, and vice versa. Thus, by construction,




these glider stimuli separate the motion information contained in three-point correlations from that specified by two-point correlations (Online Methods and **Supplementary Fig. 4**).

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[23,24] (**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[25]. As predicted by the HRC, flies turned in one direction when presented with the positive two-point correlations and in the opposite direction with the negative two-point correlations (**Fig. 3c,d**)[4,23,26]. The flies also turned in response to the diverging and converging three-point gliders, with responses that approached 20% of the two-point glider response (**Fig. 3c,d**). These three-point glider stimuli are, by design, very different from natural motion, as they achieve their correlation specificity by averaging out all two-point correlations, whereas natural motion contains both two-point and three-point correlations (**Fig. 1**). As with the two-point stimuli, positive and negative correlations in three-point gliders evoked turns in opposite directions. In other words, simply inverting the contrast of three-point glider patterns inverted their perceived directions.

**Figure 3** *Drosophila* responds to triple correlations. (**a**) Binary spatiotemporal patterns, glider stimuli with two- and three-point contrast correlations, were presented to flies. Space-time plots for each of the six 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**) We interleaved 1-s periods of glider stimuli with uncorrelated (Unc) stimuli; the timing of the presentation of the gliders is denoted by the thick black bar. Response curves show the mean (solid line) and s.e.m. (shading) over flies. (**d**) Mean turning velocities were computed for each glider by averaging over 0.5 s of the stimulus (gray bar in **c**). Turning responses are presented for wild-type *Drosophila*, alongside the predicted response of an 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 $3.7 \times 10^{-6}$ ($t_{11} = 8.5$). Significance was not tested for the two-point gliders. Error bars represent s.e.m.

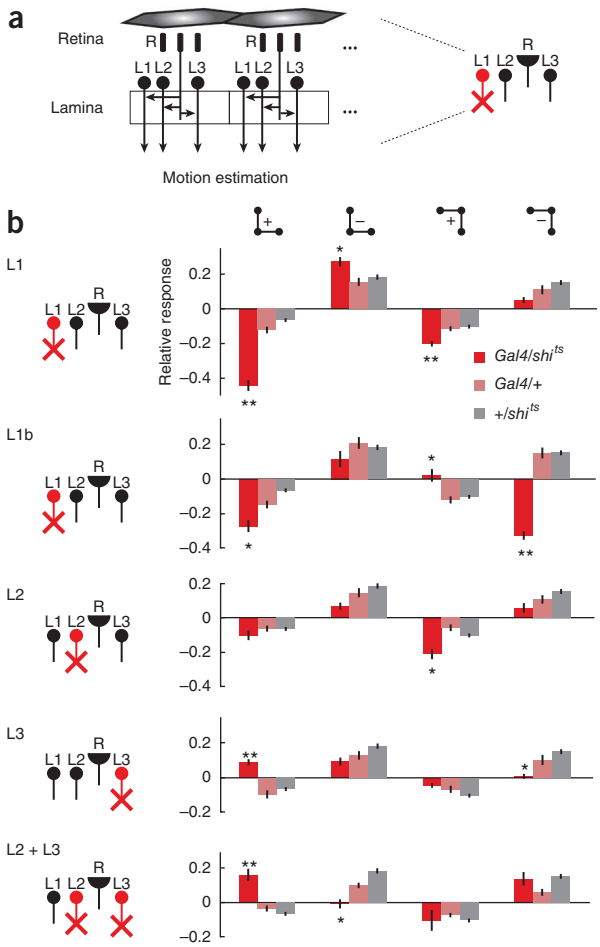



**Figure 4** Detection of triple correlations associated with specific pathways in *Drosophila*. (**a**) Left, schematic of the inputs to the fly motion processing pathways. Signals from photoreceptors are relayed through the lamina monopolar cells L1, L2 and L3. Right, a temperature-inducible dominant-negative suppressor of synaptic transmission ($shi^{ts}$) 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 three-point gliders. Responses are plotted relative to the two-point positive glider response. The two control genotypes (*Gal4/+* and *+/shi*$^{ts}$) have all input pathways intact, but contain the genetic constructs for the experimental genotype (*Gal4/shi*$^{ts}$). For the genotypes *Gal4/shi*$^{ts}$, *Gal4/+* and *+/shi*$^{ts}$, from top to bottom, $N = (19, 14, 19), (18, 13, 19), (29, 16, 19), (22, 14, 19)$ and $(17, 15, 19)$. Error bars represent ±s.e.m. $^{*}P < 0.01$, $^{**}P < 0.001$, different from both control genotypes (two-tailed *t* test).

Neither the canonical HRC nor the motion energy model predict that flies would respond to three-point glider stimuli[14] (**Fig. 3d**), and a recent modification to the HRC[27] also did not predict the measured responses (**Supplementary Fig. 4d**). 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 (**Supplementary Fig. 5**).

Previous studies have shown that the disruption of large monopolar cell (LMC) function causes selective behavioral deficits in response to moving light or dark edges[22–24,28]. We therefore determined how the disruption of LMC function affects behavioral responses to the diverging and converging three-point gliders by genetically suppressing synaptic output from the three LMCs (L1-L3) that have been associated with motion detection[24] (**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 with the controls (**Fig. 4b**).

Triple correlations were differentially associated with moving light and dark edges (**Fig. 2b**), as were the three LMCs (L1-L3) that provide inputs to motion detecting circuits[22–24]. We tested whether three-point correlation responses predict 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[22–24]. Each glider forces a particular correlation to occur consistently throughout the visual field (see Online Methods). We 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 two- and three-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 **Supplementary Fig. 6**). Notably, 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 suggests that these three-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[29]. 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 separate experiments, we measured scalp electroencephalogram (EEG) signals and behavioral responses. We designed stimuli that independently manipulate both edge contrast polarity and motion direction[23]. 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 (**Supplementary Fig. 7** and **Supplementary Movie 1**). We denoted the stimulus where light edges moved to the right as stimulus A and the stimulus where light edges moved to the left as stimulus 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′ (**Fig. 6a**, **Supplementary Fig. 7** and **Supplementary Movie 2**), 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, stimuli 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 adaptor corresponds to one half of the probe, selective adaptation should create a response that coincides with the rate of alternation between stimuli A′ and B′ in the probe. Given that 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 ($n = 7$ participants; **Fig. 6c** and **Supplementary Fig. 7**), and the measured phase difference between them was 179 ± 20° (**Fig. 6d** and **Supplementary Fig. 7**).

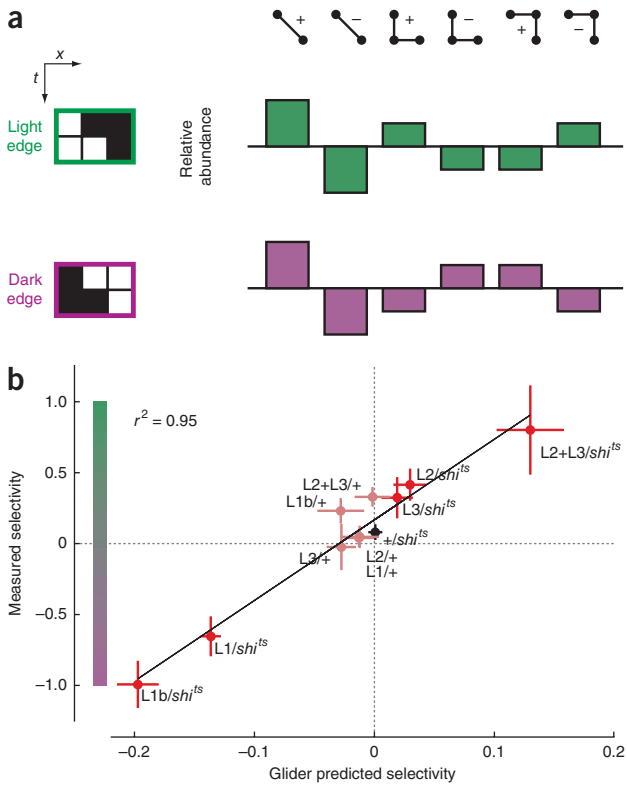



**Figure 5** Triple correlation responses predict the edge selectivity of motion pathways. (**a**) The frequency of correlational elements in a moving edge depends on its contrast polarity and direction (**Fig. 2**), and we computed the relative abundance of each correlational element from the difference in frequency of each element in rightward versus leftward motion (**Supplementary Fig. 6**). The relative abundances of the four triple correlation elements differed between light and dark edges. (**b**) We used the relative abundance of each correlational element in each edge type (**Fig. 2** and **Supplementary Fig. 6**) 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 (Online Methods). Edge selectivity was computed to be the light minus dark edge responses divided by their sum. Error bars represent ±s.e.m.

To compare these adaptation effects with those caused by adaptation to drifting sine-wave grating motion[30], we computed an adaptation index that quantified the fractional change in response as a result of adapation[30] (Online Methods). Previous experiments using drifting sine-wave gratings produced an adaptation index of 2.31 ± 0.52 (mean ± s.e.m.). Here we observed an adaptation index of 1.71 ± 0.18 (mean ± s.e.m.). The difference between these adaptation indices was small and not statistically significant (unpaired $t$ test, $P = 0.27$, $t_{13} = 1.15$), indicating that these adaption effects have a similar fractional magnitude. This is notable because the adapting stimuli in our experiment contained no net motion, yet still generated direction-selective adaptation. Thus, these results indicate that the human visual system contains neurons that differentially adapt to the motion of light and dark edges, suggesting that some 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 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 three-point gliders as motion[14] (**Fig. 7**, **Supplementary Fig. 8c** and **Supplementary Movie 3**). We then adapted edge-selective motion pathways using opposing edge stimuli as adapters and probed the perception of glider motion. The results were aggregated using the relative orientation of the adaptor and probe (Online Methods). These stimuli had no effect on converging glider percepts (**Fig. 7** and **Supplementary Fig. 8c**). 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 adaptor ($n = 2$ of 9; **Supplementary Fig. 8d**). 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 (**Fig. 7**). In contrast, dark edge adaptation had no effect on glider perception (**Fig. 7**). Because the light and dark edge adaptors 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 results suggest

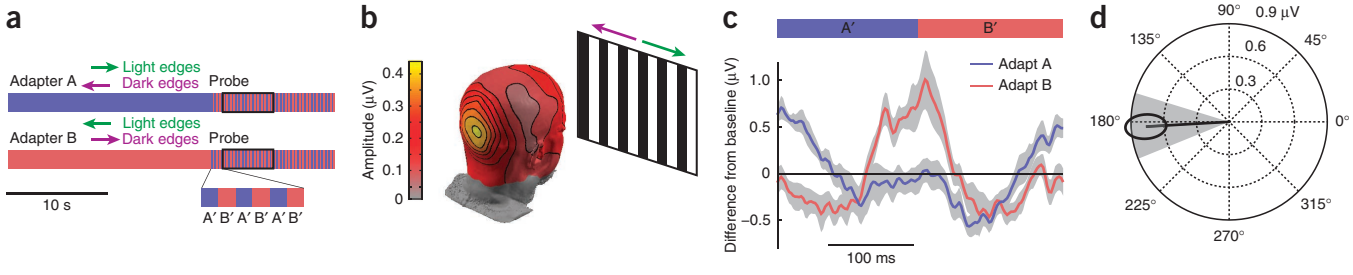



**Figure 6** Humans differentially adapt to moving light and dark edges. (**a**) Schematic of adaptor and probe stimulus procedure (**Supplementary Fig. 7**). Black box denotes the time interval used for analysis. (**b**) Scalp topography of the amplitude of the response at the A′/B′ alternation rate (3 Hz). The amplitude peaked 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 (**Supplementary Fig. 7**). The response to the probe revealed complementary modulation by the adapting stimuli at the frequency of probe alternation (3 Hz). Gray area represents ±1 s.e.m. (**d**) The within-subject difference of phase and amplitude at 3 Hz between the two adapting conditions. Ellipse represents 1 s.e.m. and the shaded wedge indicates the 95% confidence interval for the phase ($n = 7$ subjects in **c** and **d**).

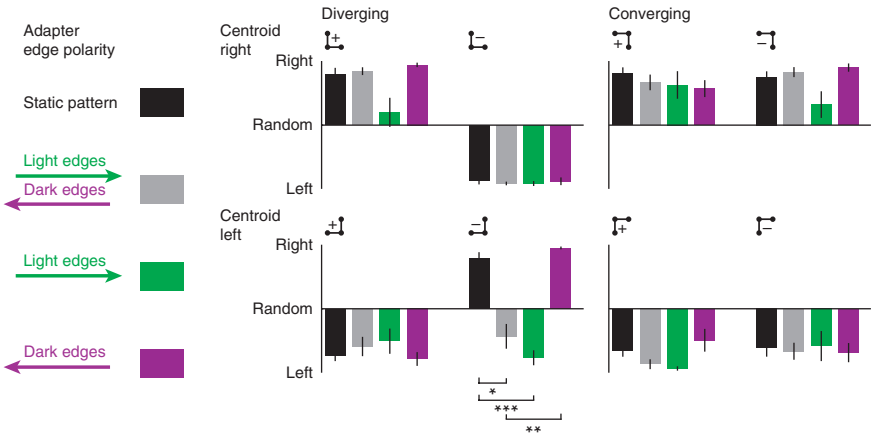



**Figure 7** Adaptation to moving light and dark edges differentially affects the perception of specific three-point gliders. Subjects were presented all combinations of four types of adaptor stimuli (left column) and eight gliders (right), and asked to report the direction of perceived glider motion. Results for each of the eight 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. $^{*}P = 1.6 \times 10^{-3}$ ($t_{16} = 5.6$), $^{**}P = 8.8 \times 10^{-4}$ ($t_{14} = 6.2$) and $^{***}P = 2.8 \times 10^{-6}$ ($t_{14} = 10.2$) (differences between conditions, 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. Error bars represent ± s.e.m.

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. This was apparent 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 revealed 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[22,23], we found that flies respond to triple correlations, and that the pattern of these responses predicted edge polarity selectivity across genotypes. Our EEG measurements revealed a neural correlate of polarity-specific motion pathways in humans, consistent with previous psychophysical observations[31,32]. In addition, although humans have been shown to perceive motion in three-point glider stimuli[14], we found that behavioral responses to specific three-point correlation stimuli were differentially adapted by light and dark moving edges. Thus, our results indicate that higher order correlations are central for distinguishing moving edge polarity in both flies and humans.

Models of motion processing that are restricted to pairwise correlations (for example, the motion energy model) have been extremely successful in explaining both perception and neural activity in areas V1 and MT[9,33,34]. Nevertheless, there is also strong evidence that such models cannot fully capture motion processing in either primates[14,15] or invertebrates[12,13,35]. For example, humans perceive robust motion in certain stimuli lacking pairwise correlations (non-Fourier motion)[15]. 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[36,37]. Several studies have shown that flies can detect certain non-Fourier motion cues[12,13,35]. The studies used motion stimuli that included quadruple (four-point) or higher even-ordered correlations, but lacked triple correlations. As a result, models that account for these effects specifically detect quadruple correlations[35]. Our results highlight that triple correlations are important for estimating motion when luminance distributions are asymmetric, as they are in natural scenes[38], 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[6].

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[29]. 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[39] indicates that contrast polarity-specific motion signals are already present in the retina. Our finding that triple correlations both improved motion estimation in natural environments and associated 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[7,8,40]. We multiplied three signals to build HRC-like models that compute triple correlations. The 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 correlations enable the computation of triple correlations.

Dark and light have long been considered to be perceptually distinct[29]. Our work extends studies in primates showing that motion processing has a component that retains information about edge contrast polarity[31,32,41]. 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[2,17]. 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[42]. Furthermore, the magnitude asymmetry between light and dark signals in natural scenes implies that light

objects tend to have a smaller spatial extent than dark objects. Thus, 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[43]. Overall, light and dark edges are not symmetric in natural environments, and treating each separately can enhance motion estimation.

We found that triple correlations could distinguish the motion of light and dark edges and improve the accuracy of motion estimation without sacrificing spatiotemporal resolution. Although we emphasize the utility of triple correlations for wide-field motion estimation, these properties might make triple correlations even more critical in behavioral contexts in which light and dark signals hold special importance or in which extensive spatiotemporal averaging is not possible. For example, insects track the motion of dark objects[44] and exhibit stereotyped escape responses when presented with looming visual stimuli[45]. 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[43], 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[15] 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[46]. Notably, observers better estimate depth when the stimulus contains both light and dark information[46], possibly reflecting the aforementioned correlation between luminance and distance in natural scenes[42]. 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[47].

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[48–50], 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. Given that chordates and arthropods diverged more than 500 million years ago and that 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. Instead, these commonalities could reflect the purposeful use of high-order statistics to estimate motion, leading flies and humans to converge on similar computational strategies to process this critical cue.

## METHODS

Methods and any associated references are available in the online version of the paper.

*Note: Any Supplementary Information and Source Data files are available in the online version of the paper.*


### ACKNOWLEDGMENTS

We thank J. Burge, P. Varghese, B. Wandell and members of the Clandinin laboratory for helpful comments on this manuscript. D.A.C. was supported by an US National Institutes of Health T32 Vision Research Training grant and a postdoctoral fellowship from the Jane Coffin Childs Foundation. J.E.F. was supported by a National Science Foundation Graduate Research Fellowship and by NSF-0801700. J.M.A. and A.M.N. were supported by a grant from the US National Institutes of Health (EY015790). D.M.G. was supported by a post-doctoral fellowship from the US National Institutes of Health, and M.A.S. was supported by a postdoctoral fellowship from the Jane Coffin Childs Foundation. In T.R.C.'s laboratory, this work was supported by US National Institutes of Health Director's Pioneer Award (DP1 OD003530) and by R01 EY022638.


### AUTHOR CONTRIBUTIONS

D.A.C., J.E.F. and J.M.A. designed and carried out the experiments. J.E.F. performed the natural scenes analysis and developed the theoretical modeling framework. D.A.C. performed the fly experiments. D.A.C. and J.M.A. performed the human experiments. D.M.G. and M.A.S. contributed reagents and fly edge selectivity measurements. D.A.C., J.E.F., J.M.A., A.M.N. and T.R.C. interpreted the data and wrote the paper.

### COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.




1. Field, D.J. Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A* **4**, 2379–2394 (1987).
2. Ruderman, D.L. & Bialek, W. Statistics of natural images: scaling in the woods. *Phys. Rev. Lett.* **73**, 814–817 (1994).
3. Simoncelli, E.P. & Olshausen, B.A. Natural image statistics and neural representation. *Annu. Rev. Neurosci.* **24**, 1193–1216 (2001).
4. Hassenstein, B. & Reichardt, W. Systemtheoretische Analyse der Zeit-, Reihenfolgen- und Vorzeichenauswertung bei der Bewegungsperzeption des Rüsselkäfers Chlorophanus. *Z. Naturforsch.* **11**, 513–524 (1956).
5. Adelson, E.H. & Bergen, J. Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A* **2**, 284–299 (1985).
6. Fitzgerald, J.E., Katsov, A.Y., Clandinin, T.R. & Schnitzer, M.J. Symmetries in stimulus statistics shape the form of visual motion estimators. *Proc. Natl. Acad. Sci. USA* **108**, 12909–12914 (2011).
7. Anderson, J.M.M. & Giannakis, G.B. Image motion estimation algorithms using cumulants. *IEEE Trans. Image Process.* **4**, 346–357 (1995).
8. Sayrol, E., Gasull, A. & Fonollosa, J.R. Motion estimation using higher order statistics. *IEEE Trans. Image Process.* **5**, 1077–1084 (1996).
9. Rust, N.C., Mante, V., Simoncelli, E.P. & Movshon, J.A. How MT cells analyze the motion of visual patterns. *Nat. Neurosci.* **9**, 1421–1431 (2006).
10. Borst, A., Haag, J. & Reiff, D.F. Fly motion vision. *Annu. Rev. Neurosci.* **33**, 49–70 (2010).
11. van Santen, J.P.H. & Sperling, G. Elaborated reichardt detectors. *J. Opt. Soc. Am. A* **2**, 300–321 (1985).
12. Theobald, J.C., Duistermars, B.J., Ringach, D.L. & Frye, M.A. Flies see second-order motion. *Curr. Biol.* **18**, R464–R465 (2008).
13. Zanker, J.M. Theta motion: a paradoxical stimulus to explore higher order motion extraction. *Vision Res.* **33**, 553–569 (1993).
14. Hu, Q. & Victor, J.D. A set of high-order spatiotemporal stimuli that elicit motion and reverse-phi percepts. *J. Vis.* **10**, 9.1–9.16 (2010).
15. Lu, Z.L. & Sperling, G. Three-systems theory of human visual motion perception: review and update. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **18**, 2331–2370 (2001).
16. Victor, J.D. & Conte, M.M. Evoked potential and psychophysical analysis of Fourier and non-Fourier motion mechanisms. *Vis. Neurosci.* **9**, 105–123 (1992).
17. Geisler, W.S. Visual perception and the statistical properties of natural scenes. *Annu. Rev. Psychol.* **59**, 167–192 (2008).
18. Dror, R.O., O'Carroll, D.C. & Laughlin, S.B. Accuracy of velocity estimation by Reichardt correlators. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **18**, 241–252 (2001).
19. van Hateren, J.H. & van der Schaaf, A. Independent component filters of natural images compared with simple cells in primary visual cortex. *Proc. Biol. Sci.* **265**, 359–366 (1998).
20. Stocker, A.A. & Simoncelli, E.P. Noise characteristics and prior expectations in human visual speed perception. *Nat. Neurosci.* **9**, 578–585 (2006).
21. Katsov, A.Y. & Clandinin, T. Motion processing streams in *Drosophila* are behaviorally specialized. *Neuron* **59**, 322–335 (2008).
22. Joesch, M., Schnell, B., Raghu, S., Reiff, D. & Borst, A. ON and OFF pathways in *Drosophila* motion vision. *Nature* **468**, 300–304 (2010).
23. Clark, D.A., Bursztyn, L., Horowitz, M.A., Schnitzer, M.J. & Clandinin, T.R. Defining the computational structure of the motion detector in *Drosophila*. *Neuron* **70**, 1165–1177 (2011).
24. Silies, M. *et al.* Modular use of peripheral input channels tunes motion-detecting circuitry. *Neuron* **79**, 111–127 (2013).
25. Buchner, E. Elementary movement detectors in an insect visual system. *Biol. Cybern.* **24**, 85–101 (1976).
26. Tuthill, J.C., Chiappe, M.E. & Reiser, M.B. Neural correlates of illusory motion perception in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **108**, 9685–9690 (2011).

27. Eichner, H., Joesch, M., Schnell, B., Reiff, D.F. & Borst, A. Internal structure of the fly elementary motion detector. *Neuron* **70**, 1155–1164 (2011).
28. Tuthill, J.C., Nern, A., Holtz, S.L., Rubin, G.M. & Reiser, M.B. Contributions of the 12 neuron classes in the fly lamina to motion vision. *Neuron* **79**, 128–140 (2013).
29. Westheimer, G. The ON-OFF dichotomy in visual processing: From receptors to perception. *Prog. Retin. Eye Res.* **26**, 636–648 (2007).
30. Ales, J.M. & Norcia, A.M. Assessing direction-specific adaptation using the steady-state visual evoked potential: results from EEG source imaging. *J. Vis.* **9**, 8 (2009).
31. Moulden, B. & Begg, H. Some tests of the Marr-Ullman model of movement detection. *Perception* **15**, 139 (1986).
32. Mather, G., Moulden, B. & O'Halloran, A. Polarity-specific adaptation to motion in the human visual system. *Vision Res.* **31**, 1013–1019 (1991).
33. Emerson, R.C., Bergen, J.R. & Adelson, E.H. Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Res.* **32**, 203–218 (1992).
34. Nishida, S.y. Advancement of motion psychophysics: review 2001–2010. *J. Vis.* **11**, 11 (2011).
35. Quenzer, T. & Zanker, J. Visual detection of paradoxical motion in flies. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **169**, 331–340 (1991).
36. Ilg, U.J. & Churan, J. Motion perception without explicit activity in areas MT and MST. *J. Neurophysiol.* **92**, 1512–1523 (2004).
37. Hedges, J.H. *et al.* Dissociation of neuronal and psychophysical responses to local and global motion. *Curr. Biol.* **21**, 2023–2028 (2011).
38. Ratliff, C.P., Borghuis, B.G., Kao, Y.-H., Sterling, P. & Balasubramanian, V. Retina is structured to process an excess of darkness in natural scenes. *Proc. Natl. Acad. Sci. USA* **107**, 17368–17373 (2010).
39. Barlow, H.B. & Levick, W.R. The mechanism of directionally selective units in rabbit's retina. *J. Physiol. (Lond.)* **178**, 477 (1965).
40. Ibn-elhaj, E., Aboutajdine, D., Pateux, S. & Morin, L. HOS-based method of global motion estimation for noisy image sequences. *Electron. Lett.* **35**, 1320–1322 (1999).
41. Schiller, P.H., Finlay, B.L. & Volman, S.F. Quantitative studies of single-cell properties in monkey striate cortex. I. Spatiotemporal organization of receptive fields. *J. Neurophysiol.* **39**, 1288 (1976).
42. Potetz, B. & Lee, T.S. Statistical correlations between two-dimensional images and three-dimensional structures in natural scenes. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **20**, 1292–1303 (2003).
43. Wiederman, S.D., Shoemaker, P.A. & O'Carroll, D.C. A model for the detection of moving targets in visual clutter inspired by insect physiology. *PLoS ONE* **3**, e2784 (2008).
44. Poggio, T. & Reichardt, W. Visual control of orientation behaviour in the fly. *Q. Rev. Biophys.* **9**, 377–438 (1976).
45. Card, G. & Dickinson, M.H. Visually mediated motor planning in the escape response of *Drosophila*. *Curr. Biol.* **18**, 1300–1307 (2008).
46. Harris, J.M. & Parker, A.J. Independent neural mechanisms for bright and dark information in binocular stereopsis. *Nature* **374**, 808–811 (1995).
47. Samonds, J.M., Potetz, B.R. & Lee, T.S. Relative luminance and binocular disparity preferences are correlated in macaque primary visual cortex, matching natural scene statistics. *Proc. Natl. Acad. Sci. USA* **109**, 6313–6318 (2012).
48. Bialek, W. Physical limits to sensation and perception. *Annu. Rev. Biophys. Biophys. Chem.* **16**, 455–478 (1987).
49. Laughlin, S.B. Energy as a constraint on the coding and processing of sensory information. *Curr. Opin. Neurobiol.* **11**, 475–480 (2001).
50. Laughlin, S.B. & Sejnowski, T.J. Communication in neuronal networks. *Science* **301**, 1870–1874 (2003).

## ONLINE METHODS

**Correlation images.** We began with natural images from van Hateren's database (image size = 1,024 × 1,536 pixels, pixel size = 1 arcmin)[19]. We linearly converted each image to a contrast scale, $C_i = (I_i - I_0)/I_0$, where $C_i$ is the $i$th pixel contrast, $I_i$ is its intensity and $I_0$ is the image's average pixel intensity. The pair correlation image (**Fig. 1b**, column 2) was

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

where $R_i$ denotes the $i$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 $v = 5.6°\ s^{-1}$, so $C_i(t) = C_{i+\Delta}(t+\delta)$. The diverging and converging triple correlation images (**Fig. 1b**, columns 3 and 4) were

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

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

where $D_i$ and $N_i$ denote the $i$th pixels of the diverging and converging triple correlation images. Symmetric color axes saturate 1 s.d. 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 $5°\ s^{-1}$ s.d. and computed local pair and triple correlations as above. We performed $10^7$ simulations for **Figure 1c** and **Supplementary Figure 3** and $10^5$ simulations for **Supplementary Figure 2**.

We also considered strategies with HRC-like spatiotemporal sampling. Given that pixels were small relative to *Drosophila*'s sampling, we downsampled each image to 1° 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 $90°\ s^{-1}$ s.d., we modeled photoreceptor responses as

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

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

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

where * denotes the convolution operator. We took the filters to be

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

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

where $\tau = 30$ ms[23] and $g(t)$ is comparable to LMC responses[52]. 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 three-point correlators (**Supplementary Fig. 1**). We performed $4 \times 10^5$ 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 twofold 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 **Figure 1** and **Supplementary Figures 2** and **3** are fractional correlation increases, relative to the local two-point correlator. **Figure 1c** and **Supplementary Figures 2** and **3** show means and s.d. across 100 data partitionings (1,000 in **Fig. 1d**).

**Fly strains.** *Drosophila melanogaster* were raised and prepared for testing as previously described[23], grown on molasses-based food in a 12-h/12-h light-dark cycle and tested for behavior during the 4 h after lights-on or 4 h before lights-off. Females were collected on $CO_2$ 1–2 d after eclosion, then tested 48–72 h later, using cold to immobilize them before gluing with an ultraviolet-cured epoxy. Genotypes are listed in **Supplementary Table 1**. All inserts were in the isoD1 background[53] or backcrossed five times into that background.

**Fly behavior.** Fly responses to different visual stimuli were measured with an apparatus similar to one previously described[23], with modified screens[24]. Three screens, each 3 × 3 cm, 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[23] and then imaged onto back-projection material that constituted the screens. The angular resolution was identical to the previously described rig[23]; each pixel subtended approximately 1°. The luminance in all glider experiments was 6 cd $m^{-2}$; edge selectivity data was taken at 18 cd $m^{-2}$. All flies were tested at a temperature of 34 °C, the restrictive temperature for *shibire$^{ts}$* (ref. 54).

**Fly stimuli.** All stimuli were presented with a screen refresh rate of 240 Hz, using custom code to generate stimuli[23]. 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 s.e.m. 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 below and in their original publication[14]. Glider stimuli lasted for 1 s and were interleaved with 0.5-s periods of uncorrelated updates. All glider pixel updates took place at 40 Hz. The correlations were one-dimensional, so that each 'pixel' subtended 5° horizontally and the entire extent vertically. The reported response is the integral of the response from 0.5–1 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 $30°\ s^{-1}$ during the second 0.5 s of the glider stimulus. Wild-type flies respond to the positive parity, two-point glider with turning rates of 140 ± 50° $s^{-1}$ (mean ± s.d.), so that this threshold excluded very few flies. In most genotypes, including wild type, this procedure excluded no flies at all; it excluded 1 of 20 flies of genotype L1/*shi$^{ts}$* and 2 of 19 of genotype L2+L3/*shi$^{ts}$*. Among three sicklier genotypes, it excluded 5 of 27 (L3/*shi$^{ts}$*), 9 of 23 (L3/+) and 9 of 25 (L2/+). Inclusion of these nonresponding flies tended to bring all responses toward 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[23].

**Fly metrics and 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' two- and three-point stimulus-induced rotations were computed; then each average was divided by the genotype's mean response to the two-point, parity +1 stimulus to compute the response as a fraction of the genotype's two-point response.

Edge selectivity (**Fig. 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 ±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[23].

We generated predictions for $X_l$ and $X_d$, denoted $\hat{X}_l$ and $\hat{X}_d$, from the experimental responses to each glider

$$\hat{X}_l = 2X_{R+} - 2X_{R-} + X_{D+} - X_{D-} - X_{N+} + X_{N-}$$

$$\hat{X}_d = 2X_{R+} - 2X_{R-} - X_{D+} + X_{D-} + X_{N+} - X_{N-}$$

where the linear coefficients correspond to the components of each correlation in the edge type (**Fig. 5a** and **Supplementary Fig. 6**), $X_{R+}/X_{R-}$ are the responses to the positive/negative parity two-point gliders, $X_{D+}/X_{D-}$ are the responses to the positive/negative parity diverging three-point gliders and $X_{N+}/X_{N-}$ are the responses to the positive/negative parity converging three-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 two-tailed Student's *t* test between experimental and control groups, or between the experiment and null hypothesis for wild-type data. Significance is reported as the maximum *P* value relative to the two controls. We did not assess whether the data were normally distributed. We did not use statistical methods to predetermine our sample sizes, but they are comparable to those in the literature[23,24,26,28].

**Simulated EMD responses to glider stimuli.** Wild-type HRC and ON/OFF model motion detector responses to various gliders (**Fig. 3d** and **Supplementary Fig. 4d**) were simulated in Matlab (MathWorks). The wild-type HRC parameters were taken directly from a recent study[23], while the model parameters for the ON/OFF model were taken from its publication[27]. Simulations were run on a one-dimensional array of 61 photoreceptors, with spacing of 5.1° and Gaussian acceptance angle of 5.7° (FWHM)[55]. Spatial resolution of the simulation was 1°, and the simulation ran in 1-ms time steps. Means and s.d. in figures were computed from 25 instantiations of the gliders. Responses to gliders were computed as the mean response over 2 s, 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.

**EEG subjects.** Informed consent was obtained from the seven 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[30].

**EEG 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 × 600 pixels viewed from 125 cm (full-width of 18.0°, full-height of 13.6°), with a 72-Hz vertical refresh rate.

There were two types of adapting stimuli. Both consisted of a 2° period vertical opposing edge presented at 6 Hz (resulting in a velocity of 12° $s^{-1}$) with a mean luminance of 71.5 cd $m^{-2}$ and a 97% Michelson contrast. The difference between the two adapting stimuli was the motion direction of the bright and dark edges. The probe stimulus comprised a 4° period vertical opposing edge presented at 3 Hz (resulting in a velocity identical to the adaptor of 12° $s^{-1}$) and was identical for all three conditions (no adaptation, adaptor type A, adaptor type B). The adapting stimulus was viewed for 20 s. Immediately after the adaptation the probe stimulus was displayed and reversed edge polarity directions at 3 Hz for 12 s. All stimuli space-time plots are shown in **Supplementary Figure 7**, and movies of adaptor A and the probe stimulus are shown in **Supplementary Movies 1** and **2**.

**EEG 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 adaptor (A or B). The adaptation/probe cycle within a block is illustrated in **Figure 6a**. After a block using the first adaptor, the participant was allowed to rest for several minutes to dissipate the adaptation effect, after which they were given a block using the alternate adaptor.

**EEG 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)[56].

**EEG spectral signature of stimulus specific adaptation.** We used a method that isolates responses of directionally selective neurons in the ssVEP[30]. 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. Thus, the presence of odd harmonic responses that are 180° 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[57].

**EEG signal acquisition and source imaging procedure.** The EEG was recorded, preprocessed and analyzed identically to prior experiments[58]. The resulting amplitude spectra of the ssVEP were then evaluated at the first harmonic of the stimulus frequency (3 Hz). Reported averages in **Figure 6** and **Supplementary Figure 7** 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 **Supplementary Figure 7c**.

**EEG 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.

**Psychophysics subjects.** Informed consent was obtained from the nine 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[14].

**Psychophysics stimulus generation and presentation.** Dynamic stimuli were presented on an HP-1320 video monitor, with 800 × 600 pixel resolution, 120-Hz refresh rate, mean luminance of 48.1 cd $m^{-2}$ and a contrast of 93.5% (Michelson). A 5° radius aperture surrounded the fixation cross; stimuli were presented within the aperture and mean luminance outside.

We used four adapting stimuli (**Supplementary Fig. 8**): static bars, opposing light and dark edges, light edges, and dark edges. The static adaptor was a full contrast square wave with a period of 2°. The opposing edges adaptor had the same parameters as the adaptor in the EEG experiments: 2° period and 12° $s^{-1}$ edge speeds. The single edges consisted of periodic single edges moving across the screen in one direction, repeated in time (**Supplementary Fig. 8**). The single edges had identical parameters to the opposing edge stimuli, moving at a speed of 12° $s^{-1}$ and with a period of 2°.

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

**Psychophysics adaptation protocols.** An initial adaptation phase lasted for the first 20 s of each experiment. After the initial adaptation, probe glider stimuli were shown in random order for 0.65 s, interleaved with three-s adaptation 'top ups', repeating the initial adaptor. 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 adaptor, for a total of 40 trials for each subject in the points shown in **Figure 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.

**Psychophysics 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 were computed from the cross-subject data, and significance between conditions was assessed with a two-sample, unpaired Student's *t* test, Bonferroni-corrected. The bar plots in **Figure 7** present a linear




doi:10.1038/nn.3600

transform of fraction-to-the-right, where 'Left' corresponds to 0 to the right, 'Random' corresponds to 0.5 to the right, and 'Right' corresponds to 1 to the right. We did not assess whether data were 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[14]. 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 three-point gliders we produce here exclude all two-point correlations and exclude all other three-point correlations. Gliders can produce more complex, higher order or more distant correlations[14].

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. **Supplementary Figure 4a** shows examples of each update rule, and **Supplementary Figure 4c** gives examples of each rule's space-time patterns. **Supplementary Movie 3** shows four 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 two-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 $C_i(t)C_{i+\Delta}(t+\delta) = P$, where $\Delta$ 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 $C_{i+\Delta}(t+\delta) = PC_i(t)$, given that $C = 1/C$ 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 translates the entire pixel pattern in one direction, whereas for odd parity, it translates and inverts on each update.

The three-point glider rules are a generalized version of the two-point rules. The update rule is $C_i(t)C_{i+\Delta}(t)C_{i+\Delta}(t+\delta) = P$ (converging case) and $C_i(t)C_i(t+\delta)C_{i+\Delta}(t+\delta) = P$ (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[59]. Edge cases can result in undetermined pixel values; in such cases, we seeded with random pixel contrasts. In three-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 (**Supplementary Fig. 4b**). Conversely, inverting the contrast of a two-point glider does not alter its parity, given that all pairwise products remain unchanged.


51. Potters, M. & Bialek, W. Statistical mechanics and visual signal processing. *J. Phys. I* **4**, 1755–1775 (1994).
52. Juusola, M., Uusitalo, R. & Weckström, M. Transfer of graded potentials at the photoreceptor-interneuron synapse. *J. Gen. Physiol.* **105**, 117 (1995).
53. Gohl, D.M. *et al*. A versatile *in vivo* system for directed dissection of gene expression patterns. *Nat. Methods* **8**, 231–237 (2011).
54. Kitamoto, T. Conditional modification of behavior in *Drosophila* by targeted expression of a temperature-sensitive shibire allele in defined neurons. *J. Neurobiol.* **47**, 81–92 (2001).
55. Stavenga, D.G. Angular and spectral sensitivity of fly photoreceptors. II. Dependence on facet lens F-number and rhabdomere type in *Drosophila*. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **189**, 189–202 (2003).
56. Palomares, M., Ales, J.M., Wade, A.R., Cottereau, B.R. & Norcia, A.M. Distinct effects of attention on the neural responses to form and motion processing: A SSVEP source-imaging study. *J. Vis.* **12**, 15 (2012).
57. Berens, P. CircStat: A MATLAB toolbox for circular statistics. *J. Stat. Softw.* **31**, 1–21 (2009).
58. Ales, J.M., Farzin, F., Rossion, B. & Norcia, A.M. An objective method for measuring face detection thresholds using the sweep steady-state visual evoked response. *J. Vis.* **12**, 18 (2012).
59. Gardner, M. Mathematical games: the fantastic combinations of John Conway's new solitaire game, 'life'. *Sci. Am.* **223**, 120–123 (1970).