Symmetries in stimulus statistics shape the form of visual motion estimators

James E. Fitzgerald[a,b,1], Alexander Y. Katsov[c], Thomas R. Clandinin[c], and Mark J. Schnitzer[a,d,e,f,g,1]

[a]Bio-X, James H. Clark Center, Departments of [b]Physics, [c]Neurobiology, [d]Applied Physics, and [e]Biology, and [f]CNC Program, Stanford University, Stanford, CA 94305; and [g]Howard Hughes Medical Institute, Stanford, CA 94305

Edited* by William T. Newsome, Stanford University, Stanford, CA, and approved June 7, 2011 (received for review October 20, 2010)

The estimation of visual motion has long been studied as a paradigmatic neural computation, and multiple models have been advanced to explain behavioral and neural responses to motion signals. A broad class of models, originating with the Reichardt correlator model, proposes that animals estimate motion by computing a temporal cross-correlation of light intensities from two neighboring points in visual space. These models provide a good description of experimental data in specific contexts but cannot explain motion percepts in stimuli lacking pairwise correlations. Here, we develop a theoretical formalism that can accommodate diverse stimuli and behavioral goals. To achieve this, we treat motion estimation as a problem of Bayesian inference. Pairwise models emerge as one component of the generalized strategy for motion estimation. However, correlation functions beyond second order enable more accurate motion estimation. Prior expectations that are asymmetric with respect to bright and dark contrast use correlations of both even and odd orders, and we show that psychophysical experiments using visual stimuli with symmetric probability distributions for contrast cannot reveal whether the subject uses odd-order correlators for motion estimation. This result highlights a gap in previous experiments, which have largely relied on symmetric contrast distributions. Our theoretical treatment provides a natural interpretation of many visual motion percepts, indicates that motion estimation should be revisited using a broader class of stimuli, demonstrates how correlation-based motion estimation is related to stimulus statistics, and provides multiple experimentally testable predictions.

neuroscience | vision | sensory perception | motion energy | *Drosophila*

The statistics of sensory inputs influence perception and shape neural encoding of auditory (1), somatosensory (2), and visual stimuli (3). The adaptation of neural codes to stimulus statistics in vertebrates and invertebrates is thought to increase neural coding efficiency and improve perceptual sensitivity in a changing sensory world (4). These ideas have prompted scrutiny of natural statistics (5) and perceptual effects under ethological sensory stimulation (6). Here we consider the problem of visual motion estimation and explore how the statistics of visual contrast influence estimation strategies.

To investigate how strategies for motion estimation are influenced by the statistics of visual stimulation and the observer's prior expectations, we examine the problem of motion estimation using Bayesian inference. Bayesian inference is one means of relating neural computation to stimulus statistics (7), and multiple researchers have used this approach to study motion estimation in biology (8–12) and machine vision (13, 14). We consider how the broad class of motion estimation models represented by a Volterra series (15–17) is influenced by the statistical structure of stimuli. We show how the Bayes optimal motion estimator can be written as a Volterra series and discuss how prior expectations and sensory inputs interact to shape motion estimates. These analyses reveal the prominent role that symmetries in stimulus statistics play in guiding motion perception in the presence of noise.

Results

A Volterra Series Generalizes Pairwise Models of Motion Estimation. The visual system interfaces with the external world only through photoreceptors, so the ultimate goal of any visual estimator of global motion is to use noisy photoreceptor voltages, $\{V_n\}$, induced by unknown dynamic stimuli, $C(\theta, t) = (I(\theta, t) - I_0)/I_0$, to infer motion in the visual field, $\psi(t)$ (Fig. 1*A*). Here $n = 1, \ldots, N$, where N is the number of photoreceptors, $I(\theta, t)$ is the intensity profile of the stimulus, I_0 is the average intensity, $C(\theta, t)$ is the luminance contast, and $\psi(t)$ is the rotational trajectory of the organism. In this description, $\psi(t)$ accounts for all coherent motion, thereby allowing $C(\theta, t)$ to have arbitrary structure and dynamics. This formulation reflects our focus on global motion, but the strategies identified will be able to compute local motion as well. Assuming linear photodetection, we use photovoltages and contrasts interchangeably. Our treatment does not consider optic flow induced by global translational motion.

Seminal work by Hassenstein and Reichardt developed a model of motion processing that predicted a relationship between the motion response and specific parameters of the visual stimulus (18–22). In this model's simplest form, the visual contrast at one location is multiplied with the time-delayed contrast from a neighboring location. Subtracting the mirror-symmetric operation yields a quantity related to the stimulus velocity (Fig. 1*B*). The Reichardt model has subsequently been generalized to a class of pairwise models with a common quadratic nonlinearity (Fig. 1*C*) (23–25). This class includes Adelson and Bergen's motion energy model (23). It is natural to ask: Is the second-order nonlinearity necessary or sufficient to estimate motion, and what benefits are offered by other nonlinearities?

To explore this question we use a Volterra series to model higher-order nonlinearities (15–17). The Volterra series represents a broad class of motion estimators as a sum of terms correlating visual signals from multiple spatiotemporal points. Formally, the Volterra series' velocity estimate is

$$\dot{\psi}_S(t) = \kappa^{(0)} + \sum_{l=1}^{\infty} \sum_{i_1=1}^{N} \cdots \sum_{i_l=i_{l-1}}^{N} \int dt_1 \ldots \int dt_l \times \frac{1}{j_1! \ldots j_N!} \kappa^{(l)}_{i_1,\ldots,i_l}(t_1, \ldots, t_l) V_{i_1}(t - t_1) \ldots V_{i_l}(t - t_l), \quad [1]$$

where the $\{\kappa\}$ are known as Volterra kernels. The sum over the order l separates terms according to how many photovoltages are being correlated. The $l = 0$ term describes a bias that is independent of input, the $l = 1$ terms describe linear estimators, and when some terms $l \geq 2$ are nonzero, the estimator is nonlinear. The sums over photoreceptor indexes, $\{i_k\}$, and time integrals allow arbitrary spatiotemporal correlations to be used

Author contributions: J.E.F. designed research; J.E.F., A.Y.K., T.R.C., and M.J.S. performed research; and J.E.F., A.Y.K., T.R.C., and M.J.S. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

[1]To whom correspondence may be addressed. E-mail: jamesef@stanford.edu or mschnitz@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1015680108/-/DCSupplemental.

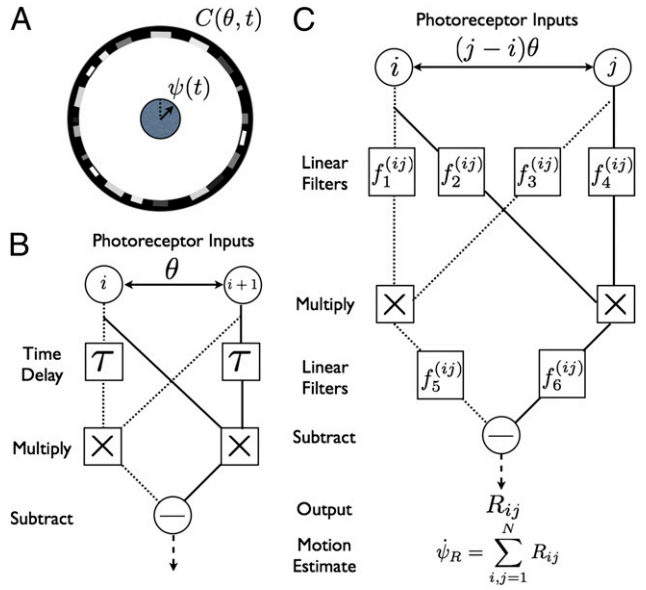

Fig. 1. Fundamentals of motion estimation. (*A*) The task of the organism is to estimate its rotational velocity $\dot{\psi}(t)$ from the unknown contrast pattern $C(\theta, t)$. (*B*) The simplest Reichardt correlator uses time delays τ to estimate motion. The mirror symmetry makes the model sensitive to motion in either direction. (*C*) The generalized Reichardt hypothesis allows arbitrary filters $\{f_1^{(ij)}, f_2^{(ij)}, f_3^{(ij)}, f_4^{(ij)}\}$ before the multiplicative nonlinearity as well as filters $\{f_5^{(ij)}, f_6^{(ij)}\}$ postnonlinearity.

for motion estimation, and j_n is the number of times n appears in $\{i_k\}$ (e.g., $\kappa^{(4)}_{1,2,3,3} \Rightarrow j_1 = j_2 = 1, j_3 = 2, j_4 = 0$, etc.). The general Reichardt hypothesis can be succinctly restated as $\kappa^{(l)}_{i_1,\ldots,i_l}(t_1, \ldots, t_l) = 0, l \neq 2$. This definition of the Reichardt hypothesis allows arbitrary second-order kernels, thereby encapsulating both the Reichardt model (18) and the most commonly used motion energy model (23).

To study motion perception with the Volterra series, we consider motion estimation as a Bayesian inference, show how the Bayesian computation can be expressed as a Volterra series, and examine how this series depends on prior expectations regarding stimulus statistics. We also examine how statistical properties of the experimental stimulus influence the importance of various terms in the series.

Motion Estimation Is Statistical Inference. Suppose the organism rotates with trajectory $\psi(t)$ so that its angular velocity is $\dot{\psi}(t)$. We model the stimulus representation by photoreceptor voltages in the linear regime of photodetection (8),

$$V_n(t) = \int dt' T(t') \int d\theta M(\theta - \theta_n) C(\theta - \psi(t-t'), t-t') + \delta V_n(t). \quad [2]$$

The temporal filter T captures how the photoreceptors integrate signal through time, whereas the spatial filter M describes their angular acceptance profiles. Together the filters imply that if the animal is stationary, then the average voltage response of a photoreceptor to a stimulus is a spatiotemporal smoothing of the stimulus. The form of each filter is irrelevant for most results. The voltages are distributed around their mean by zero-mean Gaussian fluctuations, $\delta V_n(t)$. Here we model these fluctuations as independent in the frequency domain so that $V_n(\omega) \sim P_{\text{Gauss}} \equiv \mathcal{N}(\bar{V}_n(\omega|C, \psi), N(\omega))$, where the mean voltage in the Fourier domain is $\bar{V}_n(\omega|C, \psi) = T(\omega) \int dt e^{i\omega t} \int (dk/2\pi) e^{-ik(\theta_n - \psi(t))} M(-k) C(k, t)$, and $N(\omega)$ is the power spectrum of the noise. If the animal is rotating at constant speed in a static world, then adjacent photoreceptors will on average receive the same signal separated by a time delay that encodes the animal's speed. This is the basic motion signal that the visual system must extract. The dependence of Eq. **2** on $\psi(t)$ and $C(\theta, t)$ allows the general situation of an animal moving with variable speed in a dynamic world.

As already mentioned, although the neural implementation may be complex, ultimately the organism must use its photovoltages to compute an estimate of the velocity. If this estimate is chosen to minimize the expected value of the squared error, then the optimal estimator is given by the conditional mean, $\dot{\psi}_e[\{V\}](t) = \mathbb{E}_{\psi(t)|\{V\}}[\dot{\psi}(t)|\{V\}]$, where the expectation operator $\mathbb{E}_X = \mathbb{E}_{P(X)} = \int dP(X)$ averages a random variable over $P(X)$. This estimator depends on the probability distribution over possible trajectories given the recorded voltages. Because the voltages depend both on the trajectory, $\psi(t)$, and on the stimulus, $C(\theta, t)$, the dependence upon the unknown stimulus must be removed by averaging over possible stimuli. This averaging should give high weight to stimuli that are consistent with the sensory evidence and fit prior expectation regarding the visual world. Mathematical methods to accomplish this are provided by Bayesian statistics, so to evaluate this expectation, we apply the product rule of probability, Bayes' theorem, and the functional form of P_{Gauss} to obtain

$$\dot{\psi}_e[\{V\}](t) = \frac{1}{Z[\{V\}]} \mathbb{E}_{C(\theta,t),\psi(t)} \dot{\psi}(t) \times \exp\left[-\sum_{n=1}^{N} \int \frac{d\omega}{2\pi} \left(\frac{|\bar{V}_n(\omega|C, \psi)|^2}{2N(\omega)} - \frac{V_n(\omega)\bar{V}_n^*(\omega|C, \psi)}{N(\omega)} \right) \right] \quad [3]$$

(*SI Text*), where $\bar{V}_n^*$ is the complex conjugate of $\bar{V}_n$, the expectation is over the prior $P_{\text{prior}}[C, \psi]$, and

$$Z[\{V\}] = \mathbb{E}_{C(\theta,t),\psi(t)} \times \exp\left[-\sum_{n=1}^{N} \int \frac{d\omega}{2\pi} \left(\frac{|\bar{V}_n(\omega|C, \psi)|^2}{2N(\omega)} - \frac{V_n(\omega)\bar{V}_n^*(\omega|C, \psi)}{N(\omega)} \right) \right] \quad [4]$$

normalizes the probability. Unless otherwise stated, by prior we mean the joint distribution over contrast and trajectory. The dependence on sensory evidence is captured by the exponential factors in Eqs. **3** and **4** whereas prior expectations are summarized by $P_{\text{prior}}[C, \psi]$. Eqs. **3** and **4** were used by Potters and Bialek (8) as a starting point to study the effect of the signal-to-noise ratio on motion estimation (8, 26). We focus on how the prior contributes to motion estimation.

Multipoint Correlators Implement Bayesian Estimation. Using a Volterra series to express the optimal Bayesian motion estimator as a series of multipoint correlators yields a mathematical form in which the biological plausibility is clearer than in Eq. **3**. Writing this series amounts to expanding Eq. **3** in powers of V (*SI Text*),

$$\dot{\psi}_e[\{V\}] = \dot{\psi}_e[\{V\}]_{|\{V\}=0} + \sum_{l=1}^{\infty} \sum_{i_1=1}^{N} \cdots \sum_{i_l=i_{l-1}}^{N} \int d\omega_1 \ldots \int d\omega_l \times \frac{1}{j_1! \ldots j_N!} \left(\frac{\delta^{(l)} \dot{\psi}_e[\{V\}]}{\delta V_{i_1}(\omega_1) \ldots \delta V_{i_l}(\omega_l)} \right)_{|\{V\}=0} V_{i_1}(\omega_1) \ldots V_{i_l}(\omega_l) \quad [5]$$

to identify the kernels as functional derivatives. This relationship between derivatives and Volterra kernels reiterates the analogy between the Volterra series expansion of an operator and the power series expansion of a function. It is convenient to introduce $\Psi[\{V\}](t)$ to represent the numerator of Eq. **3** and the shorthand notation $f_{,i_1,\ldots,i_l} \equiv \delta^{(l)} f / (\delta V_{i_1}(\omega_1) \ldots \delta V_{i_l}(\omega_l))$. Then the linear kernels are

$$k_i^{(1)} = \frac{\Psi_{,i} Z - \Psi Z_{,i}}{Z^2}_{|\{V\}=0} \quad [6]$$

and the kernels for the pairwise motion estimators are

$$k_{i,j}^{(2)} = \frac{1}{Z^4}(\Psi_{,i,j} Z^3 - \Psi_{,i} Z_{,j} Z^2 - \Psi_{,j} Z_{,i} Z^2 - \Psi Z_{,i,j} Z^2 + 2\Psi Z_{,i} Z_{,j} Z)_{|\{V\}=0}. \quad [7]$$

Because the voltage appears linearly in the exponential, the functional derivatives are particularly easy to evaluate:

$$\Psi_{,i_1,\ldots,i_l} = \mathbb{E}_{C(\theta,t),\psi(t)} \frac{\dot{\psi}(t)}{(2\pi)^l} \frac{\bar{V}_{i_1}^*(\omega_1|C,\psi)\ldots\bar{V}_{i_l}^*(\omega_l|C,\psi)}{N(\omega_1)\ldots N(\omega_l)} \times \exp\left[-\sum_{n=1}^{N}\int\frac{d\omega}{2\pi}\left(\frac{|\bar{V}_n(\omega|C,\psi)|^2}{2N(\omega)} - \frac{V_n(\omega)\bar{V}_n^*(\omega|C,\psi)}{N(\omega)}\right)\right] \quad [8]$$

$$Z_{,i_1,\ldots,i_l} = \mathbb{E}_{C(\theta,t),\psi(t)} \frac{1}{(2\pi)^l} \frac{\bar{V}_{i_1}^*(\omega_1|C,\psi)\ldots\bar{V}_{i_l}^*(\omega_l|C,\psi)}{N(\omega_1)\ldots N(\omega_l)} \times \exp\left[-\sum_{n=1}^{N}\int\frac{d\omega}{2\pi}\left(\frac{|\bar{V}_n(\omega|C,\psi)|^2}{2N(\omega)} - \frac{V_n(\omega)\bar{V}_n^*(\omega|C,\psi)}{N(\omega)}\right)\right]. \quad [9]$$

Eqs. **5**–**9** provide a theoretical framework connecting stimulus statistics to motion estimation with multipoint correlators, and from hereon we focus on distilling the implications of these equations. Note that in Eqs. **6** and **7** the numerator is a sum of terms that distribute the l derivatives among Ψ and Z. Further, each term in the numerator is order 2^l containing Ψ, Z, or their derivatives. This general structure is present in higher-order functional derivatives as well (*SI Text*) and is a core component of our argument.

Eq. **5** represents a series of multipoint correlators. All prior information is encoded in the correlator kernels, whose structure has likely been obtained over evolutionary timescales. Reichardt and motion energy models compute correlations differently; either instantiation can be embedded in these kernels. In principle the Bayesian estimator uses correlators of all orders, but a correlator vanishes unless there are correlations over the timescales required for the stimulus to traverse the associated photoreceptors. In a biological system other constraints may force the series to be truncated. Still, the model's symmetry properties provide interesting conclusions that do not require the complete expansion. If the prior is time-translation invariant, then Eq. **5** is a Volterra series (*SI Text*). Without this condition, an explicit time dependence of the estimator makes the analogy incomplete. For simplicity, we refer to Eq. **5** as a Volterra series.

Experiments Couple Stimuli to the Motion Estimator. A typical psychophysical experiment investigates how a specific motion $\chi(t)$ is perceived. A common and simple choice for χ is uniform rotation, but ethological χ may contain complex dynamics. We examine what happens when this motion is to be inferred by observing a stimulus drawn probabilistically from $P_{\text{stim}}[C(\theta, t)]$. We assume that the experimenter has complete control over this stimulus, so the experimental stimulus does not necessarily reflect the organism's prior expectations. We also do not assume that the Volterra kernels have the form suggested by Eq. **5**. The estimate of $\dot{\chi}(t)$ varies according to stimulus and photoreceptor noise instantiations, so an aggregate response is usually obtained by averaging either many individual responses or all instances of $\chi(t)$ in a long trial. In both cases, the aggregate response is

$$\dot{\chi}_e(t) = \mathbb{E}_{C(\theta,t)}\mathbb{E}_{\{V(t)\}|C(\theta,t),\chi(t)}\left(\dot{\psi}_S^{\text{even}}(t) + \dot{\psi}_S^{\text{odd}}(t)\right), \quad [10]$$

where we have decomposed the motion estimator as the sum of the even-ordered terms ($\dot{\psi}_S^{\text{even}}(t)$) and odd-ordered terms ($\dot{\psi}_S^{\text{odd}}(t)$). $\dot{\chi}_e(t)$ is the expected velocity perceived by the organism, the contrast expectation is over P_{stim}, and the voltage expectation is over P_{Gauss}. The expectation over voltages is analytically tractable using functional integral methods. However, the resulting structure can be more simply understood by noting that the Gaussianity of P_{Gauss} implies that each term is a product of average voltages and power spectra. Because power spectra arise from expectations over a pair of fluctuations, even-ordered terms contain an even number of $\bar{V}$s and odd-ordered terms contain an odd number (*SI Text*).

We must also average over experimental instantiations of P_{stim}. Conclusions regarding this expectation are limited by the fact that the experimenter may choose P_{stim} arbitrarily. However, because many stimuli contain common features, general stimulus dependencies can still be extracted. Here we focus on symmetry. One interesting symmetry comes from considering the effects of inverting contrast, $C(\theta, t) \to -C(\theta, t)$. From Eq. **2** it is obvious that $\bar{V}_n \to -\bar{V}_n$ under contrast inversion. Thus, if P_{stim} is invariant (symmetric) to this transformation, then upon averaging over P_{stim} the contributions from $\pm C$ add for terms containing an even number of $\bar{V}$ and cancel for terms with an odd number. Such terms correspond to those in the Volterra series with even or odd l, respectively. Thus,

$$\dot{\chi}_e(t) = \dot{\chi}_e^{\text{even}}(t), \text{ CIS}. \quad [11]$$

Without contrast inversion symmetry (CIS), the cancellation of contributions from $\dot{\psi}_S^{\text{odd}}$ need not occur. Thus, for odd-ordered terms in the Volterra series to be detected, there must be statistical asymmetries in the experimental stimulus.

Many experimental and theoretical stimuli are drawn from probability distributions that have contrast inversion symmetry. Key examples include sinusoidal gratings (22, 27), square waves (19, 21), Gaussian contrast distributions (8, 12), and binary stimuli with dark–light equivalence (spatially dense patterns) (24, 28). Consequently, these experimental stimuli cannot determine whether odd-ordered estimators contribute to biological motion estimation. More generally, this argument demonstrates a fundamental connection between the ability of an algorithm to empirically describe motion estimation and the experimental stimuli used to detect it. Even when an organism uses identical motion estimators, different experimental stimuli will be perceived differently, and the average response will be describable by different models. It is essential that this result be considered when interpreting experimental results to prevent models from being erroneously discounted or excluded.

Motion Estimators Reflect Symmetries in the Prior. We now return to Bayesian motion estimation and examine the effects of expected stimulus symmetries on Volterra kernels. If there is time-reversal symmetry (TRS), the prior is invariant to the transformation, $\{C(\theta, t), \psi(t)\} \to \{C(\theta, -t), \psi(-t)\}$. Time reversal inverts the sign of the velocity, so for psychophysical experiments time-reversal symmetry is satisfied when clockwise and counterclockwise motions are equally likely and when the probability of a stimulus is the same when it is played forward or in reverse. Suppose time-reversal symmetry and consider $\Psi_{|V=0}$. Because the prior is even and $\dot{\psi}(t)$ is odd under time reversal, the overall behavior of the expression inside the expectation is determined by the exponential term. Eq. **2** implies that $\bar{V}_n(\omega|C,\psi) \to (T(\omega)/T^*(\omega))\bar{V}_n^*(\omega|C,\psi)$, so that the exponential factor in Eq. **8** is even under time reversal. Thus, the expression inside the expectation is odd and on averaging over $P_{\text{prior}}[C, \psi]$ the contributions from $\{C(\theta, t), \psi(t)\}$ and $\{C(\theta, -t), \psi(-t)\}$ cancel. Time-reversal symmetry implies that in the absence of visual input the organism estimates there to be no motion. As shown by Eqs. **8** and **9**, higher-order derivatives contain additional copies of $\bar{V}_n^*(\omega|C,\psi)$, and time reversal's implications for higher-order estimators depend on T. Note that if $T(\omega) = T^*(\omega)$, then the temporal integration filter is either acausal or instantaneous.

Next suppose that prior expectations do not discriminate between dark and light stimuli so that P_{prior} is invariant to contrast inversion. When evaluated at $\{V\} = 0$, the exponentials in Ψ, Z, and their derivatives are even under this transformation. In addition to the exponential term, each lth-order derivative contains l copies of $\bar{V}$ and thus the expression inside the expectation is even under contrast inversion if l is even and is odd if l is odd. Upon averaging over $P_{\text{prior}}[C, \psi]$ the contributions of $\{C(\theta, t), \psi(t)\}$ and $\{-C(\theta, t), \psi(t)\}$ add for even-ordered derivatives and cancel for odd-ordered derivatives. Because each term in an odd-ordered kernel contains at least one odd-ordered derivative of Ψ or Z, if the prior is symmetric under contrast inversion, then all odd-ordered estimators vanish. If $P_{\text{prior}}[C, \psi]$ is asymmetric, then performing the average over $P_{\text{prior}}[C, \psi]$ does not necessarily result in the odd-order estimators vanishing:

$$k^{(l)}_{i_1,\ldots,i_l} \begin{cases} = 0, l \text{ odd and prior CIS} \\ \neq 0, \text{ otherwise.} \end{cases} \quad [12]$$

The major difference between Eqs. **11** and **12** is that Eq. **12** follows from symmetry properties of the prior whereas Eq. **11** results from symmetry properties of the experimental stimulus. An asymmetry in the prior is required for odd-ordered estimators to be useful whereas an asymmetry in the experiment is required to detect them.

To exploit asymmetries, pairwise estimators should be supplemented with odd-order estimators. Given the asymmetry of natural statistics (5) and mounting evidence that visual processing reflects stimulus statistics (4, 6, 9, 11), we propose departures from the Reichardt correlator and motion energy models in the form of low-order odd multipoint correlators. More generally, we propose that organisms estimate motion with a series of multipoint correlators whose form is related to the animal's prior expectations and whose efficacy can be improved by adaptation to the statistics of stimulation. Without adaptation the series is fixed, but the stimulus influences the contribution of each term. Low-order estimators dominate the response at small contrasts or low signal-to-noise ratio (8).

How the Contrast Prior Influences Motion Estimation. To more fully understand how multipoint correlators reflect the properties of the prior, we first consider the correlators' default structure without including the prior and then discuss what happens when a prior is introduced. When evaluated at $\{V\} = 0$, the argument of the exponential in Eqs. **8** and **9** is quadratic in the average voltages. To separate the dependence of the average voltage on the contrast from the trajectory, it is convenient to Fourier transform the function $C(\theta, t)$ to $C(k, t)$. This transformation shows that the exponential factor is proportional to a Gaussian probability distribution on the variables $C(k, t)$ that has zero mean and a covariance matrix that depends upon the random variable, $\psi(t)$, representing the trajectory. We thus introduce the function $P_0[C \mid \psi]$ to denote the exponential factor that shows up in Eqs. **8** and **9** at zero voltage and note that $P_0[C|\psi] \propto \mathcal{N}(0, \Sigma_\psi)$, where Σ_ψ is the ψ-dependent covariance matrix. We want to evaluate expectations of the form appearing in Eqs. **8** and **9**. In either case, if we define

$$\mathcal{P}[C|\psi] \equiv \frac{1}{\mathcal{Z}_\psi} P_{\text{prior}}[C|\psi] P_0[C|\psi], \quad [13]$$

then we can rewrite all contrast expectations in the form $\mathbb{E}_{\mathcal{P}[C|\psi]} C(k_1, t_1) \ldots C(k_l, t_l)$, where $\mathcal{Z}_\psi$ normalizes the probability. These expectations correspond to moments of $\mathcal{P}[C|\psi]$.

An lth order estimator is nonzero only if $\mathcal{P}[C|\psi]$ has a nonzero lth moment. Intuitively, it is possible to estimate motion with a multipoint correlator only if the animal can leverage associated correlations. For example, it would be useless to correlate spatially separated signals at different times if the stimulus changes too quickly to provide a reliable motion signal. Because we have not yet evaluated the expectation over $\psi(t)$, a nonzero lth moment does not necessarily provide an lth-order estimator. Even if a stimulus correlation exists, if the animal's dynamics are too fast to use this correlation, then the associated multipoint correlator does not provide a relevant signal. Furthermore, the lth-order estimator receives contributions from lower-order moments that may cancel the lth-order moment (see Eqs. **6** and **7** for two examples).

With a uniform prior, $\mathcal{P}$ is a zero-mean Gaussian and all nonzero moments can be written as the product of matrix elements from the covariance matrix. This result shows the utility of the Reichardt correlator and the motion energy model. Priors modify even-ordered correlators and lead to odd-ordered estimators by providing odd-ordered moments to $\mathcal{P}$. Whereas linear motion estimators may occur in principle, their applicability is severely limited by their reliance on strongly stereotyped and predictable visual input (*SI Text*).

Structure and Utility of Multipoint Correlators. Generically, the Bayesian estimation strategy contains a bias, linear filters, pairwise correlators, and multipoint correlators. We now explore the interpretation of these terms.

The zeroth order term in Eq. **5** describes an input-independent bias in the estimator. If the stimulus is more likely to rotate clockwise than counterclockwise, then the estimate should likewise be biased toward clockwise motion. Although Watson and Ahumada proposed a linear motion detector (29), subsequent work argued for an implicit nonlinearity to detect moving sine gratings (24). Our theory contains linear estimators in circumstances that are not likely to be ethological (*SI Text*).

The simplest two- and three-point correlators to detect rightward motion are schematized in Fig. 2*A*. If a square wave moves across the detectors at a speed equal to the photoreceptor spacing divided by the time delay τ, then each correlator arm receives the same input. For the two-point correlator any signal is equivalent because multiplying two dark or two bright points together gives a positive response indicating motion to the right. For the three-point correlator, multiplying three dark spots gives a negative response and multiplying three bright spots gives a positive response. This makes the response sensitive to the phase of the wave. To understand how asymmetries avoid this problem, con-

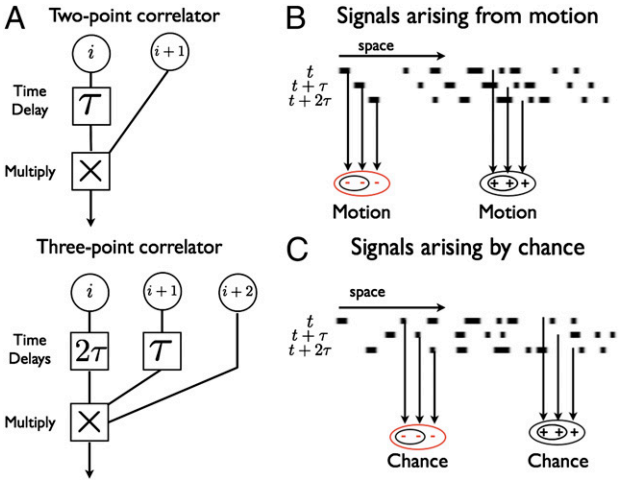

Fig. 2. The utility of odd-ordered estimators. (*A*) Schematic showing two- and three-point correlators to detect rightward motion. (*B*) When motion is rightward, the two-point correlator returns a positive response to light or dark spots (small black ovals). The three-point correlator gives a positive response to light spots (large black oval) and a negative response to dark spots (large red oval). This difference yields an ambiguity between rightward moving light spots and leftward moving dark spots. (*C*) Motion signals also arise randomly. This coincidence is less likely for dark spots, so the estimator should interpret negative three-point output as motion to the right.

sider negative contrast sparse random dot stimuli. Here contrast refers to the Michelson contrast, $C = (I_{fg} - I_{bg})/(I_{fg} + I_{bg})$, and negative contrast implies that the bright signals are more common than dark signals ($P_{dark} < 0.5$). The ambiguity is resolved with the observation that signals could arise from either motion (Fig. 2*B*) or chance (Fig. 2*C*). Because the animal is less likely to detect dark spots by chance, the animal should associate negative output with rightward motion.

These considerations are captured by multipoint correlators that we numerically evaluate using the theory (Fig. 3). Fig. 3*A* demonstrates the symmetry breaking induced by $P_{dark} \neq 0.5$ and shows example stimuli. The timescales of motion are chosen to be short so the full structure of the correlator can be seen at small temporal offsets. The shape of the Reichardt correlator is independent of whether the random dots are dense, sparse and of positive contrast, or sparse and of negative contrast (Fig. 3*B*). However, the third-order correlator is quite different in these three situations. For dense random dots, the third-order estimator vanishes, whereas the signs of kernel elements are reversed for sparse random dots with positive and negative contrast (Fig. 3*C*). This inversion reflects the confidence switch discussed above. We examine the structure of the third-order correlator using 2D cross-sections obtained by fixing one of the three kernel variables (Fig. 3*D*). In the correlators, the extrema are determined by the speed of motion and the contour shape is governed by the timescales associated with random dot coherence and motion persistence. As the speed decreases, the extrema shift toward large offsets. As the timescales lengthen, the contours straighten, reflecting distant signals becoming correlated with the current velocity.

Discussion

By representing the Bayesian estimator with a Volterra series, our theory connects two important ways of looking at neural processing. The theory is analytically tractable in simple situations and allows complex statistical models to be treated numerically. These properties make the theory a flexible tool for studying motion estimation in diverse contexts.

We predict that animals' visual systems may use higher-order multipoint correlators to convert regularities of ethological stimuli into improved motion estimates. By choosing the prior to reflect natural statistics, our theory makes explicit predictions for the form of the pairwise and higher-order estimators, thereby allowing a direct comparison of theory to phenomenological results. Future theoretical and experimental work should explore this relationship in detail. In the natural world, both the frequency of occurrence and the consequences of misestimation are critical determinants of ethological relevance. For example, to improve estimation of potentially threatening stimuli, it may be advantageous for the animal to sacrifice accuracy when estimating motion of common stimuli. This trade-off can be modeled theoretically by scaling the squared error by a stimulus-dependent cost function. Minimizing this error is equivalent to replacing the prior with the product of

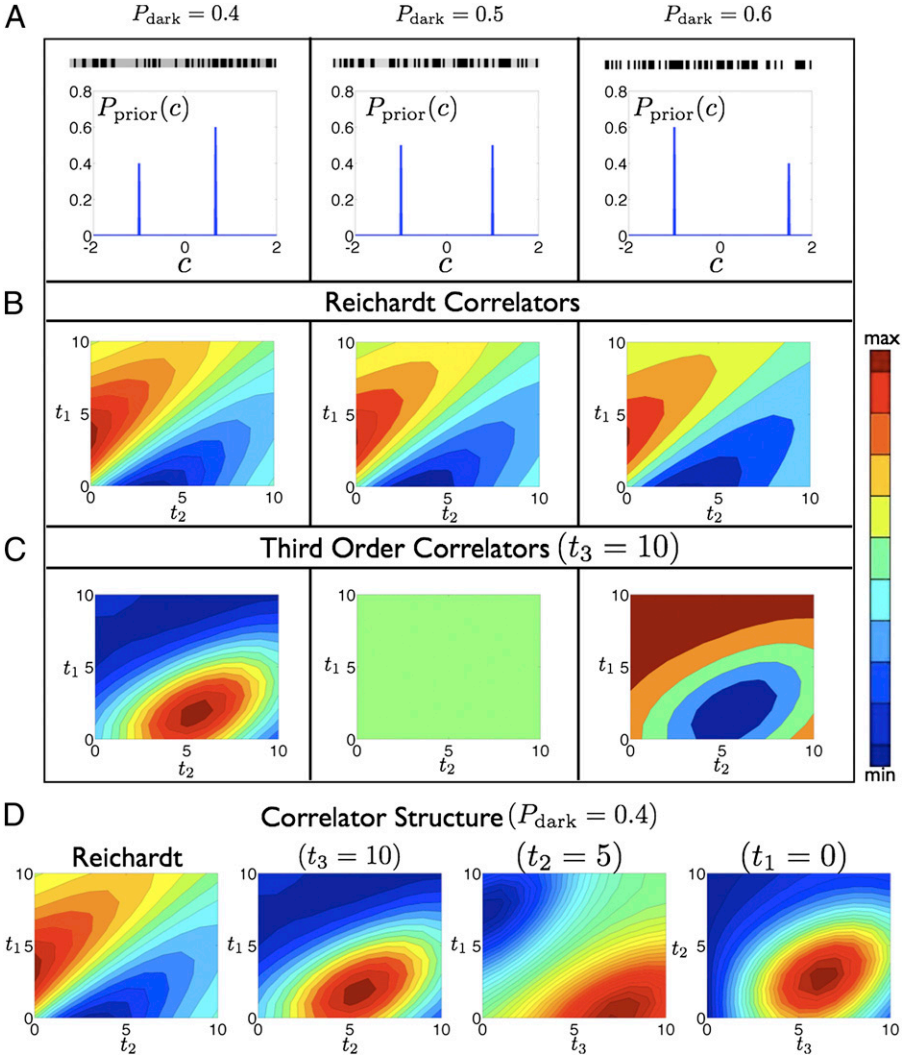

Fig. 3. The structure of multipoint correlators. (*A–C*) *Left*, probability of a dark signal to be 0.4; *Center*, 0.5; and *Right*, 0.6. (*A*) To demonstrate the asymmetry, we show an example stimulus and cartoon of the contrast prior. (*B*) The Reichardt correlator is insensitive to P_{dark}. (*C*) The third-order correlator varies strongly with P_{dark}. The sign of the kernel inverts as P_{dark} crosses 0.5, indicating the switch of confidence (Fig. 2). (*D*) Structure of the Reichardt and third-order correlator, fixing the third, second, or first coordinate.

the prior and cost function. Thus, our theory also describes this scenario through an appropriate reinterpretation of the prior.

Our theory extends to higher-order correlations the prediction that pairwise spatiotemporal correlations in the stimulus can induce motion percepts (21). Intriguingly, certain stimuli, termed non-Fourier motion stimuli, lack pairwise correlations but induce motion percepts in many animals, including humans (30–32). Because these stimuli have higher-order correlations, our theory explains the perception of non-Fourier motion in terms of naturally occurring correlations in the visual field that are represented in higher-order terms of our analysis and conceptually unifies Fourier and non-Fourier motion. Nonlinearities in the visual system coupled with a pairwise motion estimator can describe some of these percepts (30). Our theory demonstrates when such nonlinearities enhance motion estimation. Recent experiments suggest that biological systems use an even broader range of higher-order statistics to estimate motion (28, 33). Using stimuli void of lower-order correlations, Hu and Victor showed that humans experience motion percepts built from third- and fourth-order correlations (33). Consistent with multipoint correlators, most percepts invert in response to contrast inversion or its generalization to even-ordered estimators (33). Percepts that do not invert may reflect competition between correlators. An asymmetric contrast-response curve in *Drosophila* further suggests contributions from odd-ordered correlators (28).

The theoretical dependence of motion estimation on the statistics of stimuli underscores the importance of ethological experiments. Multipoint correlators use the statistics of the environment to improve motion estimation (*SI Text*). Natural stimuli have regularities that the nervous system may exploit in its estimation strategy (6). Probing the system with artificial stimuli misconstrues the relative importance of each motion estimator. The stimuli introduced by Hu and Victor (33) provide an excellent technique to specifically probe higher-order multipoint correlators. However, because these stimuli induce multiple higher-order correlations, they do not allow individual multipoint correlators to be measured unambiguously. Whereas we have emphasized the perceptual role of symmetry and symmetry breaking, the stimulus structure determines the contribution of each estimator to the percept under general circumstances. Motion estimation thus should be reinvestigated using a broad range of stimuli to determine the extent to which the organism adapts its strategy to moments of the stimulus distribution or stimulus complexity. In particular, it will be interesting to test whether exposure to a visual environment that is strongly dark–light asymmetric will cause an increase in the contribution of odd-ordered correlators to motion estimation.

Materials and Methods

We approximated expectations using 100 samples of 100 s in duration with 10-ms increments, drawn from the prior distribution. This method assumed the prior was stationary and noise was white (*SI Text*). We calculated kernels up to 100 ms in duration and averaged over all 200-ms windows. $\{\bar{V}_n(t)\}$ were calculated with exponential T (30-ms time constant), Gaussian M (2° SD), and a 4° photoreceptor spacing. Our contrast prior consisted of random dots from a zero-mean binomial distribution set by the probability of dark dots, $P_{\text{dark}} \in \{0.01, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.99\}$. Pixels were 1°, and the intensity of bright pixels was chosen to maintain constant I_0. The stimulus rotated leftward or rightward at 100°/s. At every step, dot contrasts were updated with probability P_t, and the direction of motion was updated with probability P_v (Fig. 3, $P_t = 0.1$, $P_v = 0.2$). When contrast–inversion symmetry was enforced, we included each sample with original and inverted contrasts.

ACKNOWLEDGMENTS. The authors thank Damon Clark, Eran Mukamel, Daniel Fisher, and Michael Peskin for discussion and comments on the manuscript. J.E.F. acknowledges support from a National Science Foundation Graduate Research Fellowship. A.Y.K. acknowledges support from a Stanford Graduate Fellowship. T.R.C. acknowledges support from a National Institutes of Health Pioneer Award (DP1OD003530). M.J.S. acknowledges support from the W. M. Keck Foundation and a National Institutes of Health Pioneer Award (DP1OD003560).

1. Dean I, Harper NS, McAlpine D (2005) Neural population coding of sound level adapts to stimulus statistics. *Nat Neurosci* 8:1684–1689.
2. Maravall M, Petersen RS, Fairhall AL, Arabzadeh E, Diamond ME (2007) Shifts in coding properties and maintenance of information transmission during adaptation in barrel cortex. *PLoS Biol* 5:e19.
3. Simoncelli EP (2003) Vision and the statistics of the visual environment. *Curr Opin Neurobiol* 13:144–149.
4. Fairhall AL, Lewen GD, Bialek W, de Ruyter Van Steveninck RR (2001) Efficiency and ambiguity in an adaptive neural code. *Nature* 412:787–792.
5. Geisler WS (2008) Visual perception and the statistical properties of natural scenes. *Annu Rev Psychol* 59:167–192.
6. Dror RO, O'Carroll DC, Laughlin SB (2001) Accuracy of velocity estimation by Reichardt correlators. *J Opt Soc Am A Opt Image Sci Vis* 18:241–252.
7. Knill DC, Pouget A (2004) The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends Neurosci* 27:712–719.
8. Potters M, Bialek W (1994) Statistical mechanics and visual signal processing. *J Phys I France* 4:1755–1775.
9. Ascher D, Grzywacz NM (2000) A Bayesian model for the measurement of visual velocity. *Vision Res* 40:3427–3434.
10. Stocker AA, Simoncelli EP (2006) Noise characteristics and prior expectations in human visual speed perception. *Nat Neurosci* 9:578–585.
11. Thiel A, Greschner M, Eurich CW, Ammermüller J, Kretzberg J (2007) Contribution of individual retinal ganglion cell responses to velocity and acceleration encoding. *J Neurophysiol* 98:2285–2296.
12. Lalor EC, Ahmadian Y, Paninski L (2009) Optimal decoding of stimulus velocity using a probabilistic model of ganglion cell populations in primate retina. *J Opt Soc Am A Opt Image Sci Vis* 26:B25–B42.
13. Simoncelli EP, Adelson EH, Heeger DJ (1991) Probability distributions of optical flow. *Proceedings of the Conference on Computer Vision Pattern Recognition* (IEEE Computer Society, Washington, DC), pp 310–315, 10.1109/CVPR.1991.139707.
14. Black MJ, Anadan P (1993) A framework for the robust estimation of optical flow. *Proceedings of the Fourth International Conference on Computer Vision* (IEEE Computer Society, Washington, DC), pp 231–236, 10.1109/ICCV.1993.378214.
15. Poggio T, Reichardt W (1973) Considerations on models of movement detection. *Kybernetik* 13:223–227.
16. Buchner E (1976) Elementary movement detectors in an insect visual system. *Biol Cybern* 24:85–101.
17. Poggio T, Reichardt W (1980) On the representation of multi-input systems: Computational properties of polynomial algorithms. *Biol Cybern* 37:167–186.
18. Hassenstein B (1951) Ommatidienraster und afferente bewegungsintegration - versuche an dem russelkafer chlorophanus viridis. *Z Vgl Physiol* 33:301–326.
19. Hassenstein B, Reichardt W (1956) Systemtheoretische analyse der zeit-, reihenfolgen- und vorzeichenauswertung bei der bewegungsperzeption des russelkafers chlorophanus. *Z Naturf* 11b:513–524.
20. Hassenstein B (1959) Optokinetische wirksamkeit bewegter periodischer muster (nach messungen am russelkafer chlorophanus-viridis). *Z Naturf Part B* 14:659–674.
21. Reichardt W (1961) Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. *Principles of Sensory Communication*, ed WA Rosenblith (MIT Press, Cambridge, MA), pp 303–317.
22. Egelhaaf M, Reichardt W (1987) Dynamic response properties of movement detectors: Theoretical analysis and electrophysiological investigation in the visual system of the fly. *Biol Cybern* 56:69–87.
23. Adelson EH, Bergen JR (1985) Spatiotemporal energy models for the perception of motion. *J Opt Soc Am A* 2:284–299.
24. van Santen JPH, Sperling G (1985) Elaborated Reichardt detectors. *J Opt Soc Am A* 2: 300–321.
25. Rajesh S, O'Carroll D, Abbott D (2002) Elaborated Reichardt correlator for velocity estimation tasks. *Proc SPIE* 4937:241–253.
26. Haag J, Denk W, Borst A (2004) Fly motion vision is based on Reichardt detectors regardless of the signal-to-noise ratio. *Proc Natl Acad Sci USA* 101:16333–16338.
27. Egelhaaf M, Borst A, Reichardt W (1989) Computational structure of a biological motion-detection system as revealed by local detector analysis in the fly's nervous system. *J Opt Soc Am A* 6:1070–1087.
28. Katsov AY, Clandinin TR (2008) Motion processing streams in Drosophila are behaviorally specialized. *Neuron* 59:322–335.
29. Watson AB, Ahumada AJ, Jr (1983) A look at motion in the frequency domain. NASA Technical Memorandum TM-84352. Available at http://vision.arc.nasa.gov/personnel/al/bib.htm.
30. Taub E, Victor JD, Conte MM (1997) Nonlinear preprocessing in short-range motion. *Vision Res* 37:1459–1477.
31. Orger MB, Smear MC, Anstis SM, Baier H (2000) Perception of Fourier and non-Fourier motion by larval zebrafish. *Nat Neurosci* 3:1128–1133.
32. Theobald JC, Duistermars BJ, Ringach DL, Frye MA (2008) Flies see second-order motion. *Curr Biol* 18:R464–R465.
33. Hu Q, Victor J (2010) A set of high-order spatiotemporal stimuli that elicit motion and reverse-phi percepts. *J Vision* 10(3):1–16.