Noise characteristics and prior expectations in human visual speed perception

Alan A Stocker & Eero P Simoncelli

Human visual speed perception is qualitatively consistent with a Bayesian observer that optimally combines noisy measurements with a prior preference for lower speeds. Quantitative validation of this model, however, is difficult because the precise noise characteristics and prior expectations are unknown. Here, we present an augmented observer model that accounts for the variability of subjective responses in a speed discrimination task. This allowed us to infer the shape of the prior probability as well as the internal noise characteristics directly from psychophysical data. For all subjects, we found that the fitted model provides an accurate description of the data across a wide range of stimulus parameters. The inferred prior distribution shows significantly heavier tails than a Gaussian, and the amplitude of the internal noise is approximately proportional to stimulus speed and depends inversely on stimulus contrast. The framework is general and should prove applicable to other experiments and perceptual modalities.

Human perception of visual motion is biased. In many situations, the perceived speed and direction of a moving visual stimulus depends significantly on attributes other than its physical motion. For example, a variety of psychophysical experiments have shown that perceived retinal speed is affected by contrast, with low-contrast stimuli generally appearing to move slower than those of high contrast[1,2]. Although this behavior seems at first glance to be a shortcoming, it can be seen as optimal for an observer who lives in a world in which slower motions are more likely to occur than faster ones and whose judgments are based on noisy measurements[3,4]. This optimal observer model is a probabilistic instantiation of Helmholtz's description of perception as a 'best guess' as to what is in the world, given the observer's current sensory input and prior experience[5].

In the modern framework of statistical estimation, the optimal observer may be precisely formulated in terms of two probability distributions. First, the variability of a set of measurements, $\vec{m}$, is specified as a conditional probability distribution, $p(\vec{m}|v)$, where v is the stimulus speed. The variability is due to a combination of external sources (for example, photon noise) as well as internal sources (for example, neural response variability). When considered as a function of v for a particular measurement, this conditional density is known as a likelihood function. The second component is a prior probability distribution, $p(v)$, which specifies the probability of encountering stimuli moving at any particular speed. According to Bayes' rule, the product of these two components (when appropriately normalized) gives the posterior distribution, $p(v|\vec{m})$, and an optimal observer should select a value of v that is best according to this distribution. Common choices are the mean or the mode. Contrast-induced biases in the perceived speed of moving patterns

arise intrinsically in this model, assuming a prior that favors low speeds: lower contrast stimuli lead to noisier measurements, producing broader likelihood functions, which lead to lower speed estimates (**Fig. 1**).

Despite the intuitive appeal of Bayesian models for perception, they are difficult to validate experimentally because one does not usually know the prior distribution or the likelihood function. In some cases, a prior can be deduced from theoretical considerations or measured from the natural environment in which an observer lives[6,7]. Some authors have developed models for the spatiotemporal structure of natural image sequences[7,8]. If one assumes a retinal coordinate system, it is difficult to deduce a distribution for human retinal image velocities because of the relative effects of body, head and eye movements. Even if such measurements were possible, conditions in the environment change over many timescales, and the observer may thus use a prior that is adapted or even switched abruptly according to sensory context[9]. Finally, a Bayesian perceptual system operates under constraints that may prevent it from representing the true prior. Consequently, the prior distribution used in most Bayesian models to date was chosen for simplicity and/or computational convenience[3,4].

An analogous set of issues arise in determining the likelihood function, which defines the stochastic relationship between the measurements and the quantity that is to be estimated. For speed perception, this relationship can be derived by assuming that image brightness is conserved[10,11] and that measurements are corrupted by additive noise[3,12–14]. Bayesian models built on this foundation have been shown to be roughly consistent with human perception[3,4,15,16]. However, the noise characteristics in these models are again chosen for computational convenience, and are unlikely to provide an accurate description

Howard Hughes Medical Institute, Center for Neural Science and Courant Institute of Mathematical Sciences, New York University, 4 Washington Place Rm 809, New York, New York 10003, USA, Correspondence should be addressed to A.A.S. (alan.stocker@nyu.edu).
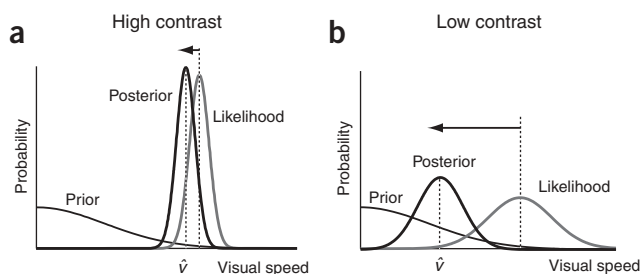
Figure 1 Illustration of a Bayesian estimator accounting for contrast-induced biases in speed perception. (**a**) A stimulus with high contrast leads to relatively precise measurements and thus a narrow likelihood. Multiplication by a prior probability for low speeds induces only a small shift of the posterior relative to the likelihood. (**b**) A low-contrast stimulus is assumed to produce noisier measurements and thus a broader likelihood. Multiplication by the same prior induces a larger shift and thus the low-contrast stimulus is typically perceived as moving slower.

of perception or physiology. Some authors have proposed likelihood models based on the response and noise properties of neurons in primary visual cortex (area V1)[15,17], and these have been shown to provide an improved description of biases observed in human speed perception[15].

In this article, we resolve these issues with an alternative approach. Rather than making assumptions based on theoretical considerations or indirect measurements, we reverse-engineered the shape of the prior distribution and the contrast and speed dependence of the likelihood function directly from perceptual behavior. Specifically, we embedded a Bayesian estimator in a general observer model that includes an optimal decision stage, and we fitted this model to trial-by-trial responses in a two-alternative forced choice (2AFC) speed discrimination experiment. We were able to validate the ability of a Bayesian observer model to account for the data and also to determine the prior distribution and internal noise level associated with the best-fitting Bayesian estimator. A preliminary version of some of this work has been presented earlier[18].

RESULTS

As outlined briefly above, a Bayesian estimator can predict contrast-induced biases in speed perception. The estimation bias is determined both by the likelihood function and the shape of the prior (**Fig. 1**). Because of this ambiguity, experimental measurements of perceptual speed biases in a subject are not sufficient to uniquely constrain both the likelihood and the prior. We show in this section that the two components may be disambiguated by embedding the Bayesian estimator in an observer model that provides a description of both the bias and the variability of subjective responses.

Bayesian observer model for speed discrimination

When human observers are presented with the same moving stimulus on repeated trials, their perception of speed fluctuates. Although it is derived from a probabilistic formulation of the problem, a Bayesian estimator is a deterministic function that maps each measurement to an estimated value $\hat{v}(\vec{m})$ and thus cannot, by itself, account for these fluctuations. Variability in perceived speed arises entirely because of the variability in the measurement, $\vec{m}$. These variations in the measurement lead to variations in the likelihood function, which in turn lead to variations in the posterior distribution, and finally to variations in the estimate. We summarize this entire process with a conditional probability distribution of the estimated speed given the true stimulus

speed, $p(\hat{v}(\vec{m})|v)$ (**Fig. 2**). For the remainder of this article, we simplify notation by leaving out the dependence on $\vec{m}$, referring to the estimate as $\hat{v}$.

The width and position of the conditional distribution of the estimates, $p(\hat{v}|v)$, can be related directly to perceptual quantities of the observer model. Specifically, the mean of the distribution represents the average perceived speed for a given stimulus speed. The width provides a measure of perceptual discriminability: that is, the ability of the observer to distinguish between stimuli moving at similar speeds. Thus, this conditional distribution provides a link between the components of the Bayesian model (prior and likelihood) and two fundamental perceptual quantities (bias and discrimination). Both perceptual quantities may be measured using standard experimental methods. Here, we use a 2AFC experimental protocol, in which the subject was asked to select which of two presented stimuli is perceived to move faster[2]. We assume that on each trial, subjects perform an independent estimate of the speeds of both stimuli and then select the one with the higher estimate. This strategy defines the relationship between the probability of the subject's responses (psychometric function) and the two conditional probability distributions, $p(\hat{v}_1|v_1)$ and $p(\hat{v}_2|v_2)$ (**Fig. 3a**). Finally, this relationship may then be used to directly constrain the prior distribution and the likelihood function using the experimentally gathered speed discrimination data (**Fig. 3b**) (Methods).

Note that our formulation of a Bayesian observer differs from most previous approaches, in which the model is used to describe the average performance of the observer by applying Bayes' rule to the average measurement at a given stimulus speed[4,15]. These models do not account for trial-to-trial variability, which is always present in the data and which provides exactly the additional information that is needed to unambiguously distinguish the contributions of the prior and the likelihood.

Estimating prior and likelihood from experimental data

Five human subjects performed a 2AFC speed discrimination task, in which they chose on each trial which of two simultaneously presented stimuli was moving faster. Stimuli consisted of drifting gratings with a broadband power spectrum of f^{-2} (see examples in **Fig. 3a**) and with variable contrast and speed covering a wide range of values. Applying
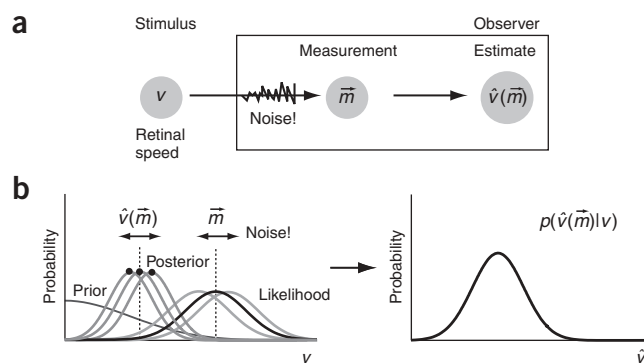
Figure 2 Bayesian estimation and measurement noise. (**a**) For a given retinal stimulus speed v, the measurement $\vec{m}$ contains all the information from which the observer will compute the estimate $\hat{v}(\vec{m})$. Because the measurement $\vec{m}$ is internal to the system, it is corrupted by internal noise and thus will vary from trial to trial over multiple presentations of the exact same stimulus. (**b**) The likelihood will also vary on each trial, as will the posterior distribution and, ultimately, the Bayesian estimate $\hat{v}(\vec{m})$. We denote the distribution of estimates for a given stimulus speed as $p(\hat{v}(\vec{m})|v)$.
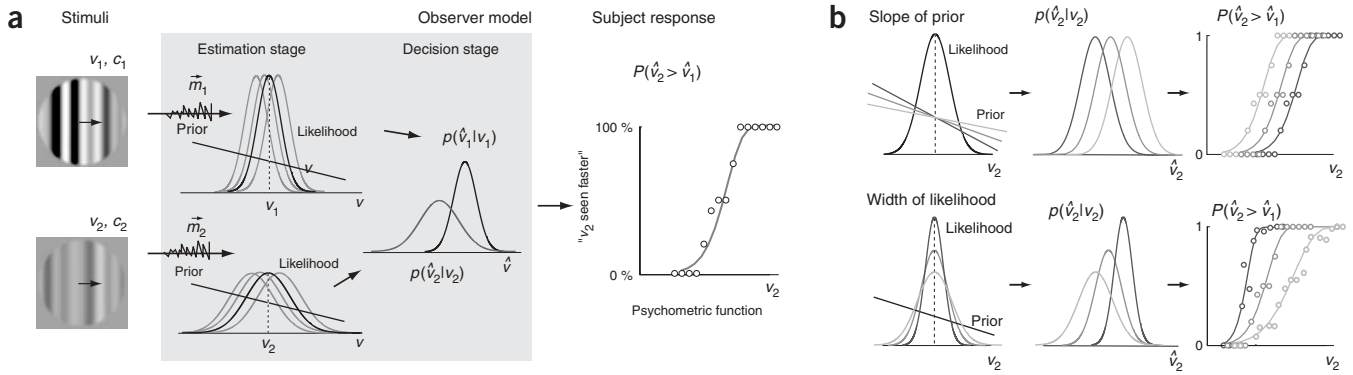
Figure 3 Bayesian observer model for 2AFC speed discrimination experiment. (**a**) On each trial, the observer independently performs an optimal estimate of the speed of each of the two stimuli based on measurements ($\vec{m}_1$, $\vec{m}_2$). These estimates are passed to a decision stage, which selects the grating with the higher estimate. Over many trials, the estimates for each stimulus pair will vary due to noise fluctuations in the measurements, and the average response of the decision stage can be computed using standard methods from signal detection theory (Methods). Plotting this average response as a function of, say, v_1, yields a psychometric function. (**b**) Illustration depicting the relationship between the model parameters and the psychometric function. The slope of the prior affects the position of the distribution of estimates and thus influences only the position of the psychometric function. However, the width of the likelihood affects both the width and the position of the distribution of estimates and thus influences both the position and the slope of the psychometric function.

the observer model (**Fig. 3a**), we solved for a nonparametric description of the prior distribution and the likelihood width (as a separable function of speed and contrast) that maximized the probability of the observed data for each subject (Methods).

The prior distribution recovered for all subjects is maximal at the lowest stimulus speed tested and decreases monotonically with stimulus speed (**Fig. 4**). But the shape differs significantly from that of the Gaussian distribution assumed in previous Bayesian models[3,4,15]. The

central portion of best fitting prior distributions can be approximated by a power law function of speed. But all subjects tested showed a flattening at low speeds, and three of the five subjects showed a flattening at high speeds (for example, subject 1, **Fig. 4**). The remaining two did not show this tendency, at least not over the range of speeds tested (for example, subject 2, **Fig. 4**).

For all subjects, the width of the likelihood is roughly constant with respect to speed (**Fig. 4**, middle column) when considered in a logarithmic speed domain, suggesting that a fixed-width Gaussian in this domain (that is, a log-Normal distribution) might provide an adequate functional description (Methods). The recovered dependence of the likelihood width on contrast is monotonically decreasing (**Fig. 4**, right column). We found that this relationship may be fit by a simple parametric function derived from assumptions about noise and contrast response models of cortical neurons[19] (Methods). This is consistent with previous findings that the introduction of contrast saturation improves the ability of a Bayesian model to fit subjective data[15]. Note that the sensitivity of speed perception on contrast varies from subject to subject.
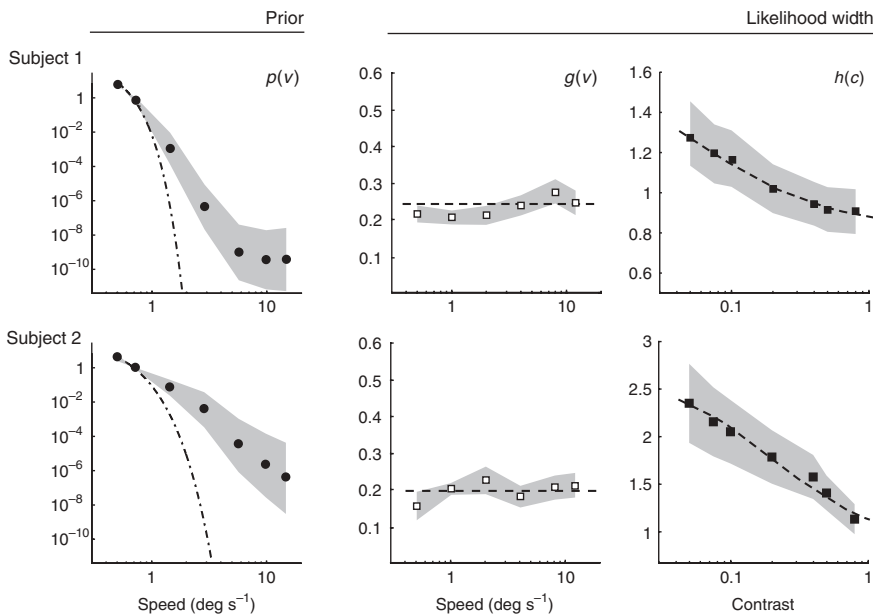
Figure 4 Parameters of the Bayesian observer model fitted to perceptual data of two representative subjects. The extracted prior, $p(v)$, exhibits a much heavier tail than the best-fitting Gaussian distribution (dash-dotted lines), for both subjects. The speed and contrast dependence of the likelihood width ($g(v)$ and $h(c)$) indicate that likelihood is approximately constant in a logarithmic speed domain and decreases monotonically with contrast in a manner consistent with a simple model for neural response characteristics (dashed line; Methods). Shaded areas represent the two standard deviation intervals computed from 30 bootstrapped data sets. Subject 1 was aware of the purpose of the experiment but subject 2 was not. Among all subjects, subject 2 shows the strongest contrast dependence as well as the broadest likelihoods.

Comparison of perceptual data and model

To examine how well the fitted Bayesian observer model accounts for human visual speed perception, we used the model to generate predictions of both average perceived speed and thresholds for speed discrimination. We compared these to values extracted directly by fitting a Weibull function to the psychometric function associated with each stimulus combination (for each subject, there are a total of 72 such functions; provided in **Supplementary Fig. 1** online together with model and Weibull fits). Data for all subjects show that lower-contrast stimuli appeared to
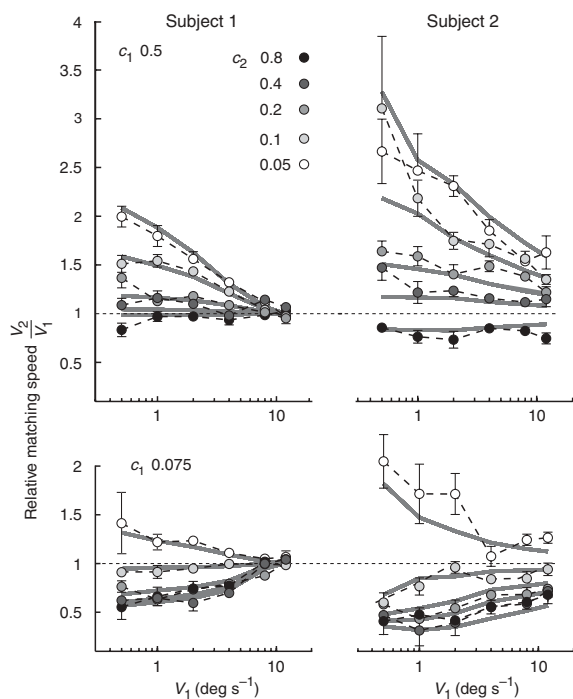
Figure 5 Perceived matching speeds as a function of contrast. Comparison of matching speeds predicted by the fitted Bayesian observer model with those obtained from Weibull fits to the raw data in each experimental condition, for the two representative subjects (**Fig. 4**). Top: relative speed of a test stimulus with different contrast levels $c_2 = [0.05, 0.1, 0.2, 0.4, 0.8]$ perceived to be moving as fast as a high-contrast reference stimulus ($c_1 = 0.5$), as a function of reference stimulus speed v_1. Points indicate the speed of subjective equality estimated from the Weibull fit (that is, the value of v_1 for which the response probability $P(\hat{v}_2 > \hat{v}_1) = 0.5$). Error bars indicate s.d. across 30 bootstrapped sets of the trial data. Data points of constant contrast c_2 are connected with dashed lines and are filled with the same shade. Solid gray lines show the predicted relative matching speed of the fitted Bayesian observer model (**Fig. 4**), averaged over all bootstrap samples. Bottom: same comparison for a low-contrast reference stimulus ($c_1 = 0.075$).

move slower, and the model provides a good account of this behavior. The strength of the contrast effect, however, varies substantially across subjects (**Fig. 5**) and is reduced for higher speeds, effectively vanishing for some subjects (for example, subject 1).

Subjective discrimination thresholds, which are primarily determined by the likelihood width (**Fig. 3b**), are seen to increase monotonically with speed but fail at low speeds to show the proportionality to speed that would be expected from the Weber-Fechner law (**Fig. 6**). This is most easily seen by replotting relative thresholds (**Fig. 6**, bottom panel) for which the Weber-Fechner law predicts a value that is constant with respect to speed. The behavior is consistent with results from previous experiments although all thresholds are higher than those reported for sinewave[20,21] or squarewave[22] gratings.

Comparison to other models

To further validate the extracted prior distributions and likelihood functions, we compared the performance of our fitted Bayesian observer model with previously published Bayesian models that assume a speed-independent Gaussian likelihood function and a Gaussian prior distribution[4,15]. We also considered a semiparametric version of

our model, in which the likelihood width is assumed to be constant in the chosen logarithmic speed domain and to fall with contrast according to a simple parametric model for neural response variability (**Fig. 4**, dashed lines). We fit each of the four models to the data of each of the five subjects and summarized the quality of the fit as the average log-probability of the data over all stimulus conditions. To present these probabilities in a more useful coordinate system and to normalize for the quality of data across the different subjects, we expressed the values for each subject on a relative scale whose minimum and maximum values were specified by two extremal models: the lower bound was computed as the average log-probability of the data for a coin-flipping observer model (that is, one that chooses randomly on each trial) and the upper bound was computed as the average log-probability of the data according to a Weibull function fit to each experimental condition.

For all subjects, the Bayesian observer model, using the reverse-engineered prior distribution and likelihood widths, performs nearly as well as the individual Weibull fits (**Fig. 7**). This is remarkable given the difference in degrees of freedom between the models: two free parameters of the Weibull function are independently fit to each of 72 experimental conditions, yielding a total of 144 free parameters, whereas the nonparametric Bayesian model has only 18. The performance of the
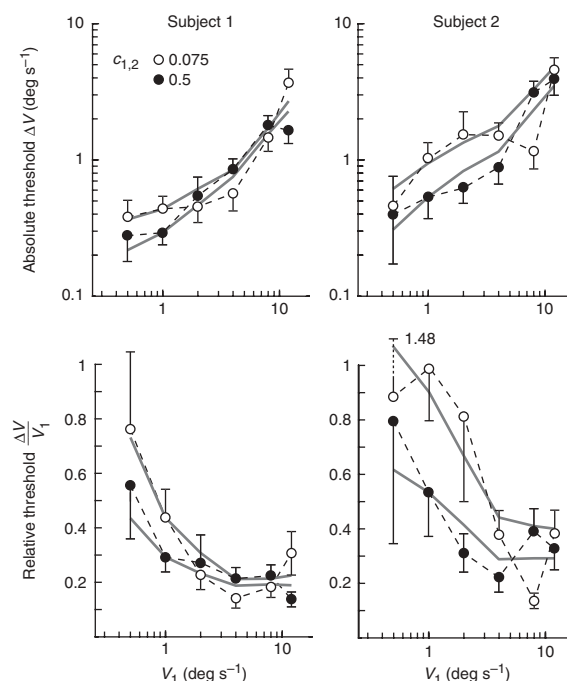
Figure 6 Speed discrimination thresholds. Comparison of speed discrimination thresholds predicted by the fitted Bayesian observer model with those obtained from Weibull fits to the raw data in each experimental condition, for the two representative subjects (**Fig. 4**). Points indicate thresholds ($\Delta v = |v_2 - v_1|$ such that response probability $P(\hat{v}_2 > \hat{v}_1) = 0.75$) as a function of reference stimulus speed v_1 for pairs of stimuli of the same contrast (solid points: $c_1 = c_2 = 0.5$; hollow points: $c_1 = c_2 = 0.075$). Error bars indicate s.d. across 30 bootstrapped sets of the trial data. Solid lines represent discrimination threshold predicted by the fitted Bayesian observer model (**Fig. 4**). Top: absolute thresholds increase monotonically with speed. Bottom: relative discrimination thresholds (absolute threshold divided by v_1) at low speeds deviate from the constant value predicted by the Weber-Fechner law.
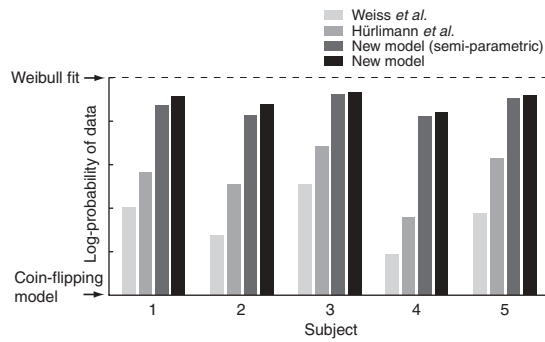
Figure 7 Model comparison: average log-probability of the experimental data. Average log-probability of the experimental data, computed according to four different estimator models fitted to data of each of five subjects. Probabilities for each subject are expressed on a scale that ranges from the value obtained for a random (coin-flipping) model to that obtained from a Weibull function fit to each experimental condition.

semiparametric version of the Bayesian observer model, which has only ten parameters, is only marginally worse. The Gaussian models have fewer free parameters (three for the model with contrast saturation[15] and only two for the other[4]) but show a performance that is significantly worse, in some cases (for example, subject 4) not much better than the coin-flipping model. This is partly because the adaptive staircase procedure leads to an accumulation of data mass around the point of subjective equality where subject responses are essentially random.

To further elucidate the behavior of the different models, we compared their prediction for matching speeds and discrimination thresholds with the values obtained from Weibull fits to the data of subject 1. The Gaussian models[4,15] predict that matching speeds and discrimination thresholds are speed independent (**Fig. 8**, left and center panels). This could provide a reasonable approximation for data

gathered over a small speed and contrast range but does not account for the full range shown here, especially in the case of the discrimination thresholds. The semiparametric model provides a substantially better account of the data (**Fig. 8**, right panel) and performs nearly as well as the full nonparametric model (compare with **Fig. 5** and **Fig. 6**).

DISCUSSION

We have shown that a Bayesian estimator can provide an accurate description of human visual speed perception. Unlike previous Bayesian models[3,4,14,15] (or related estimators based on a regularization framework[23,24,25]), we include an explicit noisy internal measurement stage so as to explain variability in perceived speed and an optimal decision stage in order to mimic trial-by-trial responses in a 2AFC speed discrimination experiment. We collected human speed discrimination data, indirectly manipulating internal noise levels by varying stimulus contrast, and used these measurements to derive the shape of the prior distribution and the width of the likelihood function. In addition to providing a good fit to the data of all five subjects, the model reveals that (i) the likelihood width is proportional to a logarithmic function of speed; (ii) the likelihood width falls monotonically with contrast and is consistent with known contrast response functions and noise characteristics of cortical neurons; and (iii) the prior falls with speed as a power law, except that the slope becomes shallower at the lowest and (for some subjects) the highest speeds tested. Thus, our fitted model confirms the assumption of a low speed prior made in previous Bayesian models[3,4,15] but clearly demonstrates that the prior distributions and likelihood functions assumed by these models do not provide an accurate account of human speed perception.

Bayesian models have also been developed to explain other aspects of human perception[26]. Some studies have extracted subjects' likelihoods or priors from perceptual data. In the cue combination literature, likelihood widths have been estimated from discrimination threshold experiments with single cues[26–31]. A recent study constrains a prior for a sensorimotor estimation task by introducing variability into the visual stimuli and assuming that subject likelihoods are consistent with this variability[32]. Another study constrains a prior by examining detection performance for stimuli drawn from different distributions and hypothesizing that the subject's performance will be best when the stimulus distribution matches their internal prior model[33]. In our experiments, external noise is negligible and our derived likelihood functions and prior directly reflect the internal noise characteristics and the prior expectations of the subjects.

Although the Bayesian observer model provides an excellent fit to the data of all subjects, it is important to recognize its limitations. The conclusions we state are well supported over the tested ranges of speed and contrast but may not hold beyond these. For example, some authors report that the perceived speed of high-speed gratings increases as their contrast is reduced[1,34]. This was not seen in the data of any of our subjects, but we did observe that the contrast-induced bias was substantially reduced at the high end of the speed range (12 deg s^{-1}), disappearing altogether for some subjects (**Fig. 5**). For our stimulus configuration, we found that subjects were
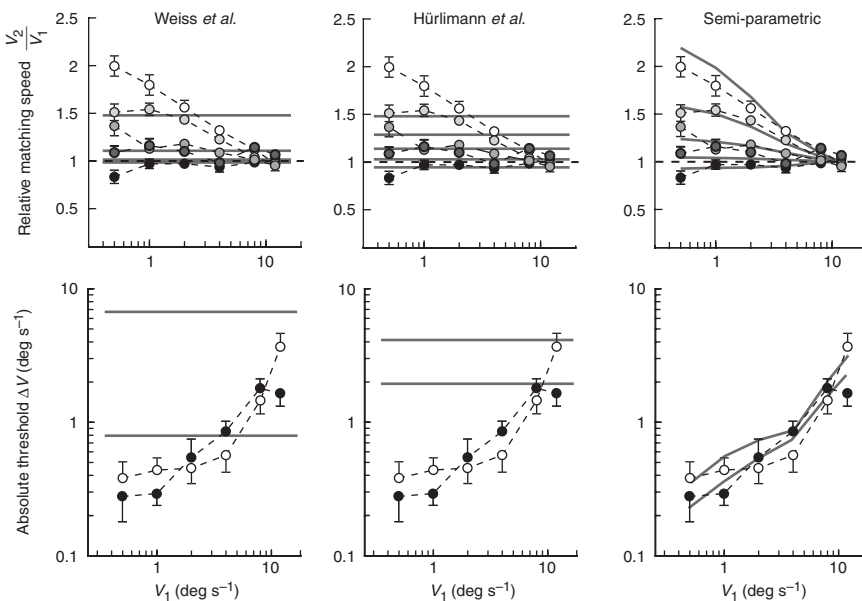
Figure 8 Model comparison: perceptual bias and discrimination predictions. Relative matching speeds and absolute discrimination thresholds predicted by three estimator models fitted to the data of subject 1 (compare with **Fig. 5** and **Fig. 6**). Left and center: Gaussian models of refs. 4 and 16, respectively. Right: fitted Bayesian model incorporating parametric functions $g(v)$ and $h(c)$ for the likelihood width (see text).

unable to make reliable judgments for speeds beyond this range. It is also worth noting that if our data were to show increases in perceived speed for low-contrast high-speed stimuli, the Bayesian model described here would be able to fit these behaviors with a prior that increases at high speeds.

Further validation of the model is needed to substantiate the broader conclusion that humans use Bayesian inference to compute visual speed. Specifically, if our subjects behave as Bayesian observers, we should be able to use their extracted prior and noise characteristics to predict their behavior on different psychophysical motion tasks[4]. This kind of validation may not be straightforward, because it is likely that the likelihood and prior depend on the details of stimulus configuration and viewing conditions. For example, speed discrimination is known to depend on retinal eccentricity[20]. Thus, the reconstructed likelihood and prior for our subjects may be specialized for the particular retinal location used in our experiment. This is not necessarily inconsistent with a Bayesian view as, under natural viewing conditions, it is likely that the visual speed distribution on the retina depends on eccentricity.

An important topic for future investigations is the underlying neurobiological implementation of our observer model. The presentation in this article has been intentionally noncommittal regarding the definition of the measurement vector $\vec{m}$, and it is of interest to associate $\vec{m}$ and the estimate $\hat{v}$ with the responses of particular neurons or populations of neurons underlying visual motion perception. The form of the contrast-dependent measurement noise in our model suggests that the locus of representation for measurements $\vec{m}$ is likely to be cortical. Neurons in area MT are a natural choice: they are highly motion selective[35,36] and their responses have been directly linked to perception[37]. If we associate the measurement $\vec{m}$ with responses of MT neurons, the estimate must be computed in subsequent neural stages[38] and should be consistent with the prior as well as the likelihood associated with the MT population response[39]. In a similar fashion, perceptual judgments have been explained with an optimal decision stage[40] or an optimal discrimination stage[41] operating on a population of noisy MT responses. Alternatively, we can assume that the population response of MT neurons directly reflects the speed estimate[16,42], and the measurement vector $\vec{m}$ is associated with responses of neurons earlier in the system (for example, area V1). This implies that the MT population responses should reflect the influence of the prior, varying with contrast in a way that is consistent with the perceptual biases exhibited by the Bayesian observer model. This behavior could be implemented in a variety of ways. For example, the contrast response functions of individual cells could differ depending on their preferred speed[16]; alternatively, the speed tuning of individual cells could change with contrast. Recent physiological experiments have begun to explore the interaction of speed and contrast in the responses of these cells[38,43].

The current model assumes a set of noisy measurements $\vec{m}$, followed by a deterministic estimator and a decision stages. If these latter stages are to correspond to neural computations, each should presumably introduce additional noise, and this should be included in optimizing the computation of the next stage. Finally, it is well known that sensory neurons adapt their response properties to the ensemble of recently presented stimuli. We have begun to examine ways by which adaptation processes can be incorporated into a more complete Bayesian theory for perception[44].

Bayesian models have attained substantial popularity in recent years and have the potential to form a unifying optimality framework for the understanding of both perception and physiology. But the Bayesian framework is quite general, and in order to realize its potential for explaining biology, it needs to be constrained to the point where it can

make quantitative experimentally testable predictions. The methodology and results introduced in this article provide a step toward this goal, and we believe that they will prove applicable to other areas of perception.

METHODS

Psychophysical experiments. Three male and two female human subjects with normal or corrected-to-normal vision participated in the psychophysical experiments. Experimental procedures were approved by the human subjects committee of New York University and all subjects signed an approved consent form. Two of the subjects (2 and 4) were not aware of the purpose of the study. Subjects were presented simultaneously with two circular patches containing horizontally drifting gratings. Patches were $3°$ in diameter, and were centered $6°$ on either side of a fixation cross. Gratings were broadband with a frequency spectrum spanning six octaves (from 1/3 cycles deg^{-1} to 2 cycles deg^{-1}) with randomized phases and a power spectrum falling as f^{-2} (see examples in **Fig. 3a**). The mean luminance of both gratings and the background was held constant at 38 cd m^{-2}. Subjects were asked to fixate a central fixation mark (cross) while each stimulus pair was presented for 1 s. After presentation, subjects selected the stimulus that appeared to be moving faster by pressing an appropriate button. If they did not respond within a 1-s interval, the trial was repeated. The total blank period between stimulus presentation was approximately 1.5 s, varying slightly with the computational time needed to generate the next stimulus pair.

Each pair of stimuli consisted of a reference and a test grating that were assigned to the left and right patches at random. On each trial, the two gratings moved in the same direction (left or right, randomly chosen on each trial). The reference grating had one of two contrast values ($c_1 = [0.075, 0.5]$) and one of five different speeds ($v_1 = [0.5, 1, 2, 4, 8, 12]$ deg s^{-1}), and the test grating had one of seven different contrast values ($c_2 = [0.05, 0.075, 0.1, 0.2, 0.4, 0.5, 0.8]$) and a variable speed v_2 that was adjusted according to two interleaved adaptive staircase procedures, each starting from one end of the adaptive speed range of each condition. Staircases procedures were of the type 'one-up one-down'. Contrast was defined as the ratio between the maximal intensity amplitude in each grating and the maximum intensity difference that could be displayed on the monitor. Each stimulus parameter triplet [v_1, c_1, c_2] was presented a total of 80 times, and these 80 trials determined a psychometric function for that condition. Individual trials for different conditions were randomly interleaved.

Extracting the prior distribution and likelihood function. For each subject, we fit the Bayesian observer model (**Fig. 3a**) to the full set of speed discrimination data by maximizing the likelihood of the data according to equation (4). This procedure requires a local parametric description of the likelihood and the prior. For this reason, we make the following assumptions. (i) We assume the prior is smooth relative to the width of the likelihood. Specifically, we assume that the logarithm of the prior is well approximated by a straight line over the range of velocities corresponding to the width of the likelihood function. (ii) We assume the likelihood, $p(\vec{m}|v)$, is well approximated by a Gaussian centered at a peak value, m_v, that can be considered as the scalar representation of the visual speed measurement (that is, a read-out of $\vec{m}$). Constraints on the noise distribution relate only to the projected value m_v. We further assume the expected value of m_v to be equal to the actual stimulus speed. (iii) We assume that the width of the likelihood function is separable in stimulus speed and contrast, $\sigma(c,v) = g(v)h(c)$, and that it varies slowly with speed.

The assumptions above allow us to relate the psychophysical data to the likelihood and prior of our probabilistic model. We write the logarithm of the prior as $\ln(p(v)) = av + b$, derive the posterior based on this local approximation of the prior and define the perceived speed $\hat{v}(m_v)$ as its mode. The posterior is

$$p(v|m_v) = \frac{1}{\alpha} p(m_v|v)p(v) = \frac{1}{\alpha} \exp \underbrace{\left[-\frac{(v - m_v)^2}{2\sigma^2(c, m_v)} + a(m_v)v + b(m_v) \right]}_{K(v)}$$

where α is a normalization constant independent of v. Note that the parameters $\{\sigma, a, b\}$ are functions of the measurement m_v rather than the true stimulus

speed v. The posterior is maximal when the exponent $K(v)$ is maximal. Thus, we differentiate $K(v)$ with respect to v, set it to zero and solve for v to find the following expression for the perceived speed:

$$\hat{v}(m_v) = m_v + \underbrace{a(m_v)\sigma^2(c, m_v)}_{\Delta(m_v)} \qquad (1)$$

where $\Delta(m_v)$ represents the relative perceptual bias. Equation (1) describes the perceived speed for a single measurement m_v, which we assume is acquired during a single trial of our experiment. Over many trials, the expected value of the perceived speed for a given stimulus with speed v_{stim} and contrast c_{stim} is equal to the expected value of m_v (which we assume is the stimulus velocity v_{stim}) plus the value of $\Delta(m_v)$ evaluated at that expected value: hence

$$E\langle \hat{v}(m_v)|v_{stim}\rangle = v_{stim} + \Delta(m_v)|_{m_v = v_{stim}} \qquad (2)$$
$$= v_{stim} + a(v_{stim})\sigma^2(c_{stim}, v_{stim}).$$

That is, the bias is a product of the slope of the logarithm of the prior and the squared width of the likelihood.

Similarly, we derive the variance of the perceived speed. Because the estimator is a deterministic function of the measurement, the variance of the estimate only depends on the variance of the measurement. For a given stimulus, we can linearize the estimator by a first-order Taylor approximation, and can approximate the variance as the variance of the linearized estimator: thus

$$var\langle \hat{v}(m_v)|v_{stim}\rangle \approx var\langle m_v|v_{stim}\rangle \left(\frac{\partial \hat{v}(m_v)}{\partial m_v}\Big|_{m_v = v_{stim}}\right)^2$$
$$\approx var\langle m_v|v_{stim}\rangle \left(1 + \frac{\partial \Delta(m_v)}{\partial m_v}\Big|_{m_v = v_{stim}}\right)^2.$$

Under assumptions (i) and (iii) (smooth prior and mild speed dependence of likelihood width, respectively), the perceived speed bias $\Delta(m_v)$ remains locally constant. Thus, the variance of the perceived speed $\hat{v}$ is approximately equal to the variance of the measurement expressed in the speed domain m_v, which is approximately the squared width of the likelihood

$$var\langle \hat{v}(m_v)|v_{stim}\rangle \approx var\langle m_v|v_{stim}\rangle \approx \sigma^2(c_{stim}, v_{stim}). \qquad (3)$$

Accordingly, the shape of the distribution of the estimate $p(\hat{v}(m_v)|v_{stim})$ matches the shape of the likelihood function, which we assumed to be Gaussian. Thus, the analysis above defines the distribution of the speed estimate for a given stimulus as a function of the local parameters of the likelihood function and the prior distribution of our Bayesian observer model. Namely, $p(\hat{v}(m_v)|v_{stim})$ is a Gaussian with mean and variance given by equations (2) and (3), respectively.

Signal detection theory. For any given prior distribution and likelihood function, the model simulates the trial-to-trial behavior in the 2AFC speed discrimination task by sampling the speed distribution $p(\hat{v}(m_v)|v_{stim})$ of each stimulus and choosing the stimulus whose sample has a higher speed value. Over a large number of simulated trials, the decision probability will follow a psychometric function according to the cumulative probability function[45,46]

$$P(\hat{v}_2 > \hat{v}_1) = \int_0^\infty p(\hat{v}_2(m_2)|v_2) \int_0^{\hat{v}_2} p(\hat{v}_1(m_1)|v_1)d\hat{v}_1 d\hat{v}_2 \qquad (4)$$

If the prior distribution and likelihood function are correct, then equation (4) should fit the experimentally measured points on the psychometric function. To extract the prior distribution and the speed and contrast dependence of the likelihood, we discretize these functions over speed and contrast and perform a maximum likelihood fit against all recorded data. The prior distribution is reconstructed by numerical integration of the fitted local slope values $a(v)$.

Contrast-dependent likelihood width. The functional form of $h(c)$ (see **Fig. 4**) is motivated by assuming that the measurements $\vec{m}$ are the responses of a set of spatiotemporally tuned cortical neurons involved in the perception of visual

speed, and thus that the variability in $\vec{m}$ and consequently the likelihood width are determined by the response behavior of these neurons. The average firing rate of cortical neurons as a function of contrast is well described by $r(c) = r_{max} c^q/(c^q + c_{50}^q) + r_{base}$, where r_{max} and r_{base} are maximum and baseline firing rate, respectively, and q and c_{50} specify the slope and the semisaturation point of the contrast response function[19]. The variability of cortical responses approximately follows a Poisson distribution; that is, the variance of response grows proportionally with the mean firing rate. This implies that the relative variability in the measurement and therefore the likelihood width decrease in inverse proportion to the square-root of the firing rate. Combining these two descriptions gives the parametric form used in describing the likelihood width (**Fig. 4**) as

$$h(c) = \frac{1}{\sqrt{(r_{max}c^q/(c^q + c_{50}^q) + r_{base})}}.$$

Fitted values for slope and semisaturation point vary across subjects in the range of $q = 1.6\ldots2.5$ and $c_{50} = 0.15\ldots0.3$ (for subject 2, c_{50} was not well constrained by the data and so was restricted to lie in this range). Note that these parameters depend on our definition of contrast.

Logarithmic speed representation. The analysis above is written in terms of the speed v but can be applied to any monotonic function of speed. We would like to choose a representation such that the approximation in equation (3) is valid: that is, so that assumption (iii) holds (slowly varying likelihood width). Several results in the psychophysics literature suggest that visual speed discrimination approximately follows a Weber-Fechner law and thus is roughly proportional to speed[21,22]. This is consistent with a log-Normal likelihood function (Gaussian in the logarithmic speed domain—assumption (ii)). But to account for the deviation from the Weber-Fechner law at low speeds, we use a modified logarithmic transformation $\tilde{v} = \ln(1 + v/v_0)$, where v_0 is a small constant. Throughout our analysis, we choose a fixed value $v_0 = 0.3$ deg s^{-1}, which results in an approximately constant $g(v)$ (see **Fig. 4**). Other choices for v_0 necessarily lead to a change in $g(v)$ because $g(v)$ expresses the speed dependence likelihood width in the $\tilde{v}$ domain. However, they do not affect the likelihood function in the linear speed domain. We have also verified that neither the fitting results nor the extracted prior are substantially changed when v_0 is varied by an order of magnitude in either direction. Notably, it has been reported that neurons in the medial temporal area (area MT) of macaque monkeys have speed-tuning curves that are approximately log-Normal in visual speed according to above modified logarithmic representation[41]. These neurons are known to play a central role in the representation of motion, and it seems natural to assume that they are involved in tasks such as our psychophysical experiments.

Note that although the Bayesian estimation is described in a logarithmic speed domain, it is computed with reference to the world representation of visual object speed. Thus, estimation is performed by transforming the posterior probability to the linear domain, selecting the estimate and transforming it back to the logarithmic speed domain.

Note: Supplementary information is available on the Nature Neuroscience website.

COMPETING INTERESTS STATEMENT
The authors declare that they have no competing financial interests.

1. Thompson, P. Perceived rate of movement depends on contrast. *Vision Res.* **22**, 377–380 (1982).
2. Stone, L. & Thompson, P. Human speed perception is contrast dependent. *Vision Res.* **32**, 1535–1549 (1992).
3. Simoncelli, E. *Distributed Analysis and Representation of Visual Motion*. Thesis, Massachusetts Institute of Technology (1993).

4. Weiss, Y., Simoncelli, E. & Adelson, E. Motion illusions as optimal percept. *Nat. Neurosci.* **5**, 598–604 (2002).
5. Helmholtz, H. *Treatise on Physiological Optics* (Thoemmes Press, Bristol, UK, 2000). Original publication 1866.
6. Betsch, B., Einhäuser, W., Körding, K. & König, P. The world from a cat's perspective - statistics of natural videos. *Biol. Cybern.* **90**, 41–50 (2004).
7. Roth, S. & Black, M. On the spatial statistics of optical flow. *International Conference on Computer Vision ICCV*, 42–49 (2005).
8. Dong, D. & Atick, J. Statistics of natural time-varying images. *Network: Comput. Neural Syst.* **6**, 345–358 (1995).
9. Brenner, N., Bialek, W. & de Ruyter van Stevenick, R. Adaptive rescaling maximizes information transmission. *Neuron* **26**, 695–702 (2000).
10. Fennema, C. & Thompson, W. Velocity determination in scenes containing several moving objects. *Comput. Graph. Image Process.* **9**, 301–315 (1979).
11. Horn, B. & Schunck, B. Determining optical flow. *Artif. Intell.* **17**, 185–203 (1981).
12. Simoncelli, E., Adelson, E. & Heeger, D. Probability distributions of optical flow. *IEEE Conference on Computer Vision and Pattern Recognition*, 310–313 (IEEE, 1991).
13. Weber, J. & Malik, J. Robust computation of optical flow in a multi-scale differential framework. *Int. J. Comput. Vis.* **14**, 67–81 (1995).
14. Weiss, Y. & Fleet, D. Velocity likelihoods in biological and machine vision. in *Probabilistic Models of the Brain* (Bradford Book, MIT Press, Cambridge, Massachusetts, 2002).
15. Hürlimann, F., Kiper, D. & Carandini, M. Testing the Bayesian model of perceived speed. *Vision Res.* **42**, 2253–2257 (2002).
16. Stocker, A.A. *Analog VLSI Circuits for the Perception of Visual Motion*. (John Wiley & Sons, Chichester, UK, 2006).
17. Ascher, D. & Grzywacz, N. A Bayesian model for the measurement of visual velocity. *Vision Res.* **40**, 3427–3434 (2000).
18. Stocker, A. & Simoncelli, E. Constraining a Bayesian model of human visual speed perception. in *Advances in Neural Information Processing Systems NIPS* Vol. 17 (eds. Saul, L.K., Weiss, Y. & Bottou, L.) (MIT Press, Cambridge, Massachusetts, 2005).
19. Sclar, G., Maunsell, J. & Lennie, P. Coding of image contrast in central visual pathways of the macaque monkey. *Vision Res.* **30**, 1–10 (1990).
20. McKee, S. & Nakayama, K. The detection of motion in the peripheral visual field. *Vision Res.* **24**, 25–32 (1984).
21. McKee, S., Silvermann, G. & Nakayama, K. Precise velocity discrimintation despite random variations in temporal frequency and contrast. *Vision Res.* **26**, 609–619 (1986).
22. Welch, L. The perception of moving plaids reveals two motion-processing stages. *Nature* **337**, 734–736 (1989).
23. Stocker, A. Analog integrated 2-D optical flow sensor. *Analog Integr. Circuits Signal Process.* **46**, 121–138 (2006).
24. Yuille, A. & Grzywacz, N. A computational theory for the perception of coherent visual motion. *Nature* **333**, 71–74 (1988).
25. Heeger, D. & Simoncelli, E. Model of visual motion sensing. in *Spatial Vision in Humans and Robots* (Cambridge Univ. Press, Cambridge, UK, 1994).
26. Knill, D.C. & Richards, W. (eds.). *Perception as Bayesian Inference* (Cambridge Univ. Press, Cambridge, UK, 1996).
27. Ernst, M. & Banks, M. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**, 429ff (2002).
28. Hillis, J., Ernst, M., Banks, M. & Landy, M. Combining sensory information: mandatory fusion within, but not between senses. *Science* **298**, 1627ff (2002).
29. Mamassian, P., Landy, M. & Maloney, L. Bayesian modelling of visual perception. in *Probabilistic Models of the Brain* (MIT Press, Cambridge, Massachusetts, 2002).
30. Knill, D. & Saunders, J. Do humans optimally integrate stereo and texture information for judgements of surface slant? *Vision Res.* **43**, 2539–2558 (2003).
31. Alais, D. & Burr, D. The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.* **14**, 257–262 (2004).
32. Körding, K. & Wolpert, D. Bayesian integration in sensorimotor learning. *Nature* **427**, 244–247 (2004).
33. Yuille, A., Fang, F., Schrater, P. & Kersten, D. Human and ideal observers for detecting image curves. *Advances in Neural Information Processing Systems NIPS* Vol. 16 (eds. Thrun, S., Saul, L. & Schölkopf, B.) (MIT Press, Cambridge, Massachusetts, 2004).
34. Thompson, P., Brooks, K. & Hammett, S. Speed can go up as well as down at low contrast: implications for models of motion perception. *Vision Res.* **46**, 782–786 (2006).
35. Albright, T. Direction and orientation selectivity of neurons in visual area MT of the macaque. *J. Neurophysiol.* **52**, 1106–1130 (1984).
36. Movshon, J., Adelson, E., Gizzi, M. & Newsome, W. The analysis of moving visual patterns. *Exp. Brain Res. Suppl.* **11**, 117–151 (1985).
37. Britten, K., Shadlen, M., Newsome, W. & Movshon, A. The analysis of visual motion: a comparions of neuronal and psychophysical performance. *J. Neurosci.* **12**, 4745–4765 (1992).
38. Priebe, N. & Lisberger, S. Estimating target speed from the population response in visual area MT. *J. Neurosci.* **24**, 1907–1916 (2004).
39. Pouget, A., Dayan, P. & Zemel, R. Inference and computation with population codes. *Annu. Rev. Neurosci.* **26**, 381–410 (2003).
40. Shadlen, M., Britten, K., Newsome, W. & Movshon, J. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J. Neurosci.* **16**, 1486–1510 (1996).
41. Nover, H., Anderson, C. & DeAngelis, G. A logarithmic, scale-invariant representation of speed in macaque middle temporal area accounts for speed discrimination performance. *J. Neurosci.* **25**, 10049–10060 (2005).
42. Simoncelli, E.P. Local analysis of visual motion. in *The Visual Neurosciences* (MIT Press, Cambridge, Massachusetts, 2003).
43. Pack, C., Hunter, J. & Born, R. Contrast dependence of suppressive influences in cortical area MT of alert macaque. *J. Neurophysiol.* **93**, 1809–1815 (2005).
44. Stocker, A. & Simoncelli, E. Sensory adaptation within a Bayesian framework for perception. in *Advances in Neural Information Processing Systems NIPS* Vol. 18 (MIT Press, Vancouver, 2006).
45. Green, D. & Swets, J. *Signal Detection Theory and Psychophysics* (Wiley, New York, 1966).
46. Wickens, T.D. *Elementary Signal Detection Theory* (Oxford University Press, Oxford, 2001).