Motion illusions as optimal percepts

Yair Weiss[1], Eero P. Simoncelli[2] and Edward H. Adelson[3]

[1] School of Computer Science and Engineering, Hebrew University of Jerusalem, Givat Ram Campus, Jerusalem 91904, Israel

[2] Howard Hughes Medical Institute, Center for Neural Science and Courant Institute of Mathematical Sciences, New York University,
 4 Washington Place, New York, New York 10003, USA

[3] Brain and Cognitive Sciences Department, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, Massachusetts 02139, USA

Correspondence should be addressed to Y.W. (yweiss@cs.huji.ac.il)

The pattern of local image velocities on the retina encodes important environmental information. Although humans are generally able to extract this information, they can easily be deceived into seeing incorrect velocities. We show that these 'illusions' arise naturally in a system that attempts to estimate local image velocity. We formulated a model of visual motion perception using standard estimation theory, under the assumptions that (i) there is noise in the initial measurements and (ii) slower motions are more likely to occur than faster ones. We found that specific instantiation of such a velocity estimator can account for a wide variety of psychophysical phenomena.

The human ability to analyze visual motion in general scenes far exceeds the capabilities of the most sophisticated computer vision algorithms. Yet psychophysical experiments show that humans also make some puzzling mistakes, misjudging speed or direction of very simple stimuli. In this paper, we propose that such mistakes of human motion perception represent the best solution of a rational system designed to operate in the presence of uncertainty.

In both biological and artificial vision systems, motion analysis begins with local measurements such as the output of direction-selective cells in primary visual cortex[1], or of spatial and temporal derivative operators in artificial systems[2,3]. These are then integrated to generate larger, more global motion descriptions. The integration process is essential because the initial local motion measurements are ambiguous. For example, in the vicinity of a contour, only the motion component perpendicular to the contour can be determined (a phenomenon referred to as the 'aperture problem')[2,4–7]. Such an integration stage seems to be consistent with much of the psychophysical[8–11] and physiological[8,12–14] data.

Despite the vast amount of psychophysical data published over the past two decades, the nature of the integration scheme underlying human motion perception remains unclear. This is true even for the simple and widely studied 'plaid' stimulus, in which two superimposed oriented gratings translate (move without changing shape, size or orientation) in the image plane (**Fig. 1a**). Due to the aperture problem, each grating's motion is consistent with an infinite number of possible translational velocities lying on a constraint line in the space of all velocities (**Fig. 1b**). When viewing a single drifting grating in isolation, subjects typically perceive it as translating in a direction normal to its contours (**Fig. 1b**). When two gratings are presented simultaneously, subjects often perceive them as a coherent pattern translating with a single motion[5,7].

How is this coherent pattern motion estimated? Most explanations are based on one of three rules[7]: intersection of constraints (IOC), vector average (VA), or feature tracking (FT). The IOC solution is the unique translation vector consistent with the information of both gratings. Graphically, this corresponds to the point in velocity space that lies at the intersection of both constraint lines (**Fig. 1b**, circle). The VA solution is the average of the two normal velocities. Graphically, this corresponds to the point in velocity space that lies halfway between the two normal velocities (**Fig. 1b**, square). An FT solution corresponds to the velocity of some feature of the plaid intensity pattern (for example, the locations of maximum luminance at the grating intersections)[15,16]. For plaids, the FT and IOC solutions both correspond to the veridical (true) pattern motion.

Which of the three rules best describes human perception? The answer is not clear: depending on the stimulus, the perceived pattern motion can be nearly veridical (consistent with IOC or FT) or closer to the VA solution. The relevant stimulus features include relative grating orientation and speed[17–19], contrast[20], presentation time[17] and retinal location[17].

Similar effects have been reported with stimuli that appear quite different from plaids[16,21]. For a moving rhombus (**Fig. 2**), as for a plaid pattern, the motion of each opposing pair of sides is consistent with a constraint line in the space of velocities. As shown in the velocity space diagrams (**Fig. 2c** and **f**), IOC or FT predicts horizontal motion, whereas VA predicts diagonal motion. Perceptually, however, the rhombus appears to move horizontally at high contrast and diagonally at low contrast. To further complicate the situation, the percept depends on the shape. If the rhombus is fattened (**Fig. 2d**), it appears to move horizontally at both contrasts. To view these moving stimuli, see http://www.cs.huji.ac.il/~yweiss/Rhombus.

One might reason that the visual system uses VA for a thin, low-contrast rhombus, and IOC/FT for a thin, high-contrast rhombus and for a fat rhombus. Although a model based on this *ad hoc* combination of rules certainly fits the data, it is clearly not a parsimonious explanation. Furthermore, each of the idealized rules is limited to stimuli containing straight structures at only two orientations, and does not offer a method for computing the normal velocities of those structures. One would prefer a single, coherent model that could predict the perceived
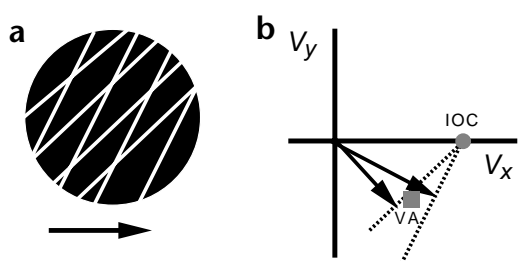
Fig. 1. Intersection of constraints. (**a**) Drifting gratings superimposed in the image plane produce a translating 'plaid' pattern. (**b**) Dotted lines indicate constraint lines; arrows indicate perceived direction of grating viewed in isolation. The IOC solution (circle) is the unique velocity consistent with the constraint lines of both gratings. The VA solution (square) is the average of the two normal velocities. There is experimental evidence for both types of combination rule.

velocity of any arbitrary spatiotemporal stimulus that appears to be translating. We have developed such a model based on a simple formulation of the problem of velocity estimation and on a few reasonable assumptions.

In Helmholtz's view, our percepts are our best guess as to what is in the world, given both sensory data and prior experience[22]. To make this definition more quantitative, one must specify (i) what is 'best' about a best guess, and (ii) the way in which prior experience should influence that guess. In the engineering literature, the theory of estimation formalizes these concepts. The simplest and most widely known estimation framework is based on Bayes' rule (see ref. 23 for examples of Bayesian models in perception and refs. 24 and 25 for Bayesian motion models). Following an approach described in previous work[26–29], we developed an optimal Bayesian estimator (known as an 'ideal observer' in the psychophysics literature) for two-dimensional velocity. Here, as in most studies of the aperture problem, we considered only cases in which humans see a single global translational motion (no deformation, rotation, occlusion boundaries, transparency, or the like). Elsewhere, we have developed extensions of this model that can handle more complicated scenes[29].

Our model begins with the standard principle of intensity conservation: it assumes that any changes in image intensity over time are due entirely to translational motion of the intensity pattern. We then made two basic assumptions: (i) local image measurements are noisy and (ii) image velocities tend to be slow. We formulated these assumptions using probability distributions (see below), and used Bayes' rule to derive the ideal observer (for further mathematical details, see Methods).

We instantiated the first assumption using a noise model commonly used in engineering because of the tractability of the solution: measurements are contaminated with additive, independent, Gaussian noise with a known standard deviation (σ). Although this simple noise model is unlikely to be correct in detail, we show that it is sufficient to account for much of the data. This noise model provides a functional form for the local likelihood: a distribution over the space of velocities that is based on measurements made in a local image patch. We depicted this likelihood as a gray-level image (**Fig. 3**) in which intensity corresponds to probability. For patches containing a single edge, the likelihood function is similar

to a 'fuzzy' constraint line— velocities on the constraint line have the highest likelihood, and the likelihood decreases with distance from the line. The 'fuzziness' of the constraint line is governed by σ, the standard deviation of the assumed noise. At corners, where local motion measurements are less ambiguous, the likelihood no longer has the elongated shape of a constraint line but becomes tightly clustered around the veridical velocity.

This model of additive Gaussian noise also resulted in a dependence of the likelihood on contrast. For a fixed value of σ, the likelihoods were broader at low contrast (**Fig. 3**, bottom). This makes intuitive sense: at low contrast there is less information about the exact speed of the stimulus, and therefore more local uncertainty, so the likelihood is more spread out. In the extreme case of zero contrast, the uncertainty is infinite.

The second assumption underlying our ideal observer model is that velocities tend to be slow. Suggestions that human observers prefer the 'shortest path', or slowest motion consistent with the visual input, date back to the beginning of the 20th century (see ref. 30 and references therein). In particular, Wallach suggested that humans prefer to see the normal velocity for a single line segment because that is the slowest velocity consistent with the image data[5]. Likewise in apparent motion displays, humans tend to choose the shortest path or slowest motion that would explain the incoming information.

We formalized this preference for slow speeds using a prior probability distribution on the two-dimensional space of velocities that is Gaussian and centered on the origin. According to this 'prior', in the absence of any image data, the most probable velocity is zero (no motion), and slower velocities are generally more likely to occur than fast ones. As with the noise model, we have no direct evidence (either from first principles or from empirical measurements) that this assumption is correct. We will show, however, that it is sufficient to account qualitatively for much of the perceptual data.

Under the Bayesian framework, the percept of the ideal observer is based on the posterior probability (the probability of a velocity given the image measurements), which is com-

Fig. 2. Insufficiency of either VA, IOC or FT rules as an explanation for human perception of a horizontally moving rhombus. (**a**) A 'narrow' rhombus at high contrast appears to move horizontally (consistent with IOC/FT). (**b**) A narrow rhombus at low contrast appears to move diagonally (consistent with VA). (**c**) Velocity space constraints for a narrow rhombus. (**d**,**e**) A 'fat' rhombus at low or high contrast appears to move horizontally (consistent with IOC/FT). (**f**) Velocity space constraints for a fat rhombus.

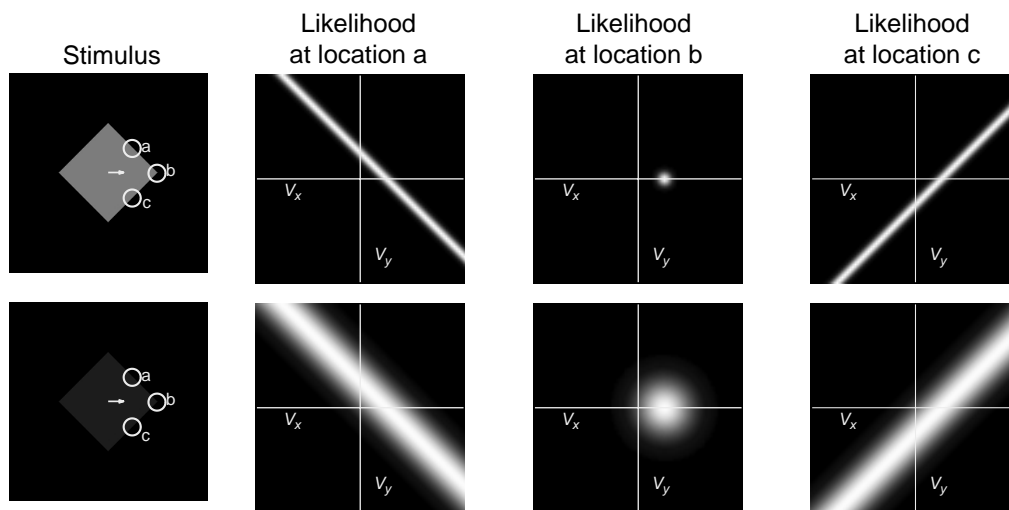|	Likelihood at location a	Likelihood at location b	Likelihood at location c
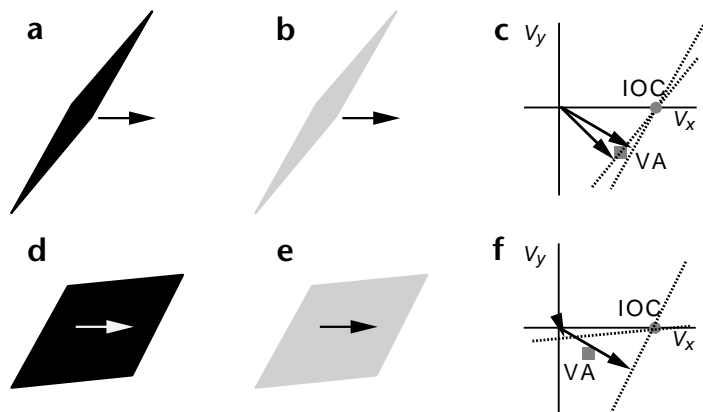
Fig. 3. Likelihood functions for three local patches of a horizontally translating diamond stimulus, computed using equation (4). Intensity corresponds to probability. Top, high-contrast sequence. Bottom, low-contrast sequence, with the same parameter σ. At edges, the local likelihood is a 'fuzzy' constraint line; at corners, the local likelihood peaks around the veridical velocity. The sharpness of the likelihood decreases with decreasing contrast.

puted from the likelihood and prior using Bayes' rule (see Methods). We formulated the posterior distribution by multiplying the prior and the likelihoods at all image locations. This is correct under the assumptions that the noise in the measurements is statistically independent, and that the likelihoods being multiplied correspond to image locations that are moving at the same velocity.

One can calculate the velocity estimate (v^*) of the ideal observer as the mean or maximum of the posterior distribution. Our posterior distribution is Gaussian, and the mean (which is also the most likely) velocity was computed analytically using the following matrix equation:

$$v^* = - \begin{pmatrix} \sum I_x^2 + \dfrac{\sigma^2}{\sigma_p^2} & \sum I_x I_y \\[2em] \sum I_x I_y & \sum I_y^2 + \dfrac{\sigma^2}{\sigma_p^2} \end{pmatrix}^{-1} \begin{pmatrix} \sum I_x I_t \\[2em] \sum I_y I_t \end{pmatrix}$$
(1)

where I_x, I_y, I_t refer to the spatial (two dimensions) and temporal derivatives of the image sequence. The sums were taken over all locations that translate together (here, we assumed this included the entire image). This equation allowed us to predict the ideal observer's velocity estimate for any image sequence. The solution of equation(1) has only one free parameter: the ratio of σ to σ_p. Changing both of these while holding the ratio constant changes the width, but not the peak, of the posterior.

We calculated the posterior for the moving rhombus stimuli (**Fig. 4a–c**), holding the free parameter (σ/σ_p) constant. Consistent with human data, the ideal observer predicts horizontal motion for a narrow, high-contrast rhombus, diagonal motion for a narrow, low-contrast rhombus and nearly horizontal motion for a fat, low-contrast rhombus. For a more quantitative comparison of the ideal observer and human perception, we showed three subjects a continuum of low-contrast rhombuses that varied between the extremes of 'thin' and 'fat', and asked them to report the perceived direction by positioning the cursor

of a computer mouse. The predictions of equation (1) provide an excellent fit to the human experimental data (**Fig. 4d**). In addition, the qualitative predictions remained unchanged while the free parameter was varied over two orders of magnitude (**Fig. 4d**). In fact, no setting of the free parameter could make the perception of narrow rhombuses more veridical than that of fat ones. Similarly, there is no setting that would make the perception of low-contrast rhombuses more veridical than that of high-contrast rhombuses.

RESULTS
We compared the predictions of the ideal observer (the solution of equation (1)) to previously published psychophysical data[17–20,31,32]. The free parameter was adjusted manually for each experiment but held constant for all conditions within each experiment. Different observers probably make different 'assumptions' regarding noise, and indeed, substantial individual differences for these illusions have been reported[17]. As with the rhombus example, the value of the free parameter did not change the qualitative predictions of the model for any of the stimuli discussed here.

Influence of contrast on perceived grating speed
The perceived speed of a single grating depends on contrast[31,33–35], with lower-contrast patterns consistently appearing slower than higher-contrast patterns[34]. This may underlie the tendency of automobile drivers to speed up in the fog[36]. In a psychophysical experiment quantifying this effect[31], subjects were asked to compare the apparent speed of two gratings of different contrast (**Fig. 5a**). The low-contrast grating was consistently perceived to be moving slower. This illusion depended primarily on the ratio of contrasts of the two gratings: the perceived speed was an approximately linear function of the contrast ratio, and was approximately independent of absolute contrast. The ideal observer shows a qualitatively similar contrast dependence. At low contrasts, the likelihood is broader and the prior has a stronger influence on the estimate. Consistent with human perception, the ideal observer also estimates the low-contrast grating as moving slower (**Fig. 5a**).
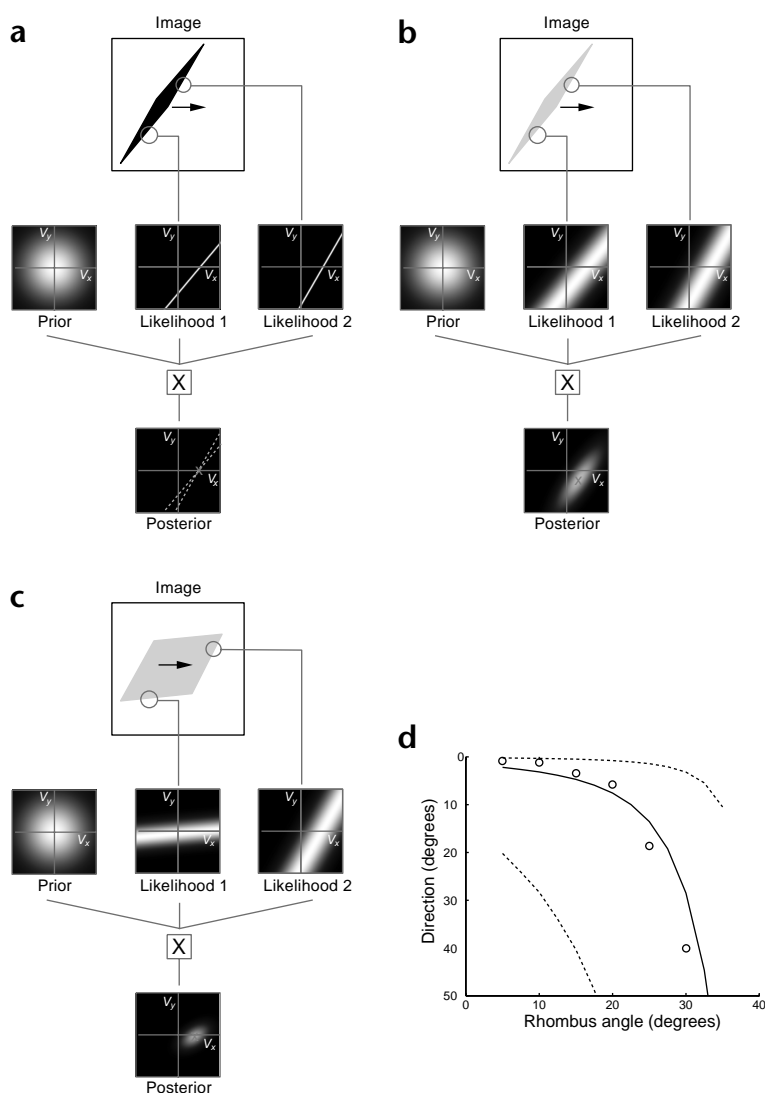
Fig. 4. Predictions of ideal observer for rhombus stimuli. (**a–c**) Construction of the posterior distribution for the rhombus stimuli. For clarity, likelihood functions for only two locations are shown; the estimator used in our study incorporated likelihoods from all locations. (**d**) Circles show perceived direction for a single human subject as rhombus angle was shifted gradually from thin to fat rhombuses (all three subjects showed a similar effect, and all gave informed consent to participate in the study). Each subject was given 100 presentations. Solid line shows the predictions of the Bayesian estimator computed using equation (1), where the free parameter was varied manually to fit the data. Dotted lines indicate the predictions when the free parameter was decreased by a factor of 10 (top dotted line) or increased by a factor of 10 (bottom line).

higher uncertainty and hence the low-contrast grating has less influence on the estimate.

Contrast influence on perceived line direction
Subjects tend to misperceive the direction of a moving line at low contrasts, even when its endpoints are visible[32]. We replotted data from an experiment in which subjects reported the perceived direction of a 'matrix' of lines (**Fig. 5c**). The matrix was constructed by replicating a single line at multiple locations in the visual field. The line was oriented such that its normal velocity was downward even when the line was moving upward. At low contrasts, subjects performed far below chance, indicating that they perceived upward motion while the line actually moved downward. The authors proposed two separate mechanisms to explain this finding, one dealing with terminator (line endpoint) motion and other with line motion. The terminator mechanism was assumed to be active primarily at high contrasts and the line strategy primarily at low contrasts.

We found that at low contrast, the ideal observer also misperceived the direction of motion because the likelihoods are broader and the estimator prefers the normal velocity (which is slower than the true velocity). To obtain a percentage of correct responses for the ideal observer, we assumed that v^* was corrupted by decision noise, and we calculated the probability that the corrupted v^* was in the upward direction. The decision noise was Gaussian in velocity space. The standard deviation of the decision noise determines the sharpness of the psychometric function and was adjusted manually. The predicted percentage correct for the ideal observer was in accordance with human perception (**Fig. 5c**, solid line).

Type I versus type II plaids: perceived direction
In the plaid literature, a distinction is often made between two types of configuration: for a 'type I' plaid, the direction of the veridical velocity lies between that of the two normal velocities; for a 'type II' plaid, the veridical direction lies outside the two normals[17]. In the latter case, the vector average is quite different from the veridical velocity.

At low contrast, the perceived direction for type II plaids is strongly biased in the direction of the vector average, and the perceived direction of type I plaids is largely veridical. We replotted data from a single subject who reported the perceived direction of a plaid under five different conditions[17] (**Fig. 5d**, circles).

The simple ideal observer presented here does not predict the quasilinear shape of the perceived relative speeds, nor does it predict the lack of dependence on total contrast (it makes slightly different predictions for maximum contrasts of 40% and 70%, **Fig. 5a**). We also constructed a slightly more elaborate model that can account for these effects in a more quantitative manner (see Discussion).

Influence of contrast on perceived plaid direction
The perceived direction of a plaid depends on the relative contrast of the two constituent gratings[20]. We replotted data from an experiment in which subjects reported the perceived direction of motion of symmetric plaids while the contrast ratio of the two components was varied (**Fig. 5b**). Perceived direction was always biased toward the normal direction of the higher-contrast grating. The magnitude of the bias changed as a function of the total contrast of the plaid (the sum of the contrasts of the two gratings). Increasing the contrast of both gratings (while the ratio of contrasts is held fixed) resulted in a smaller bias. The ideal observer shows a similar effect (E. P. Simoncelli & D. J. Heeger, *Invest. Opthal. Vis. Sci. Suppl. Abstr.* **33**, 954, 1992), which again follows from the fact that at low contrast, there is
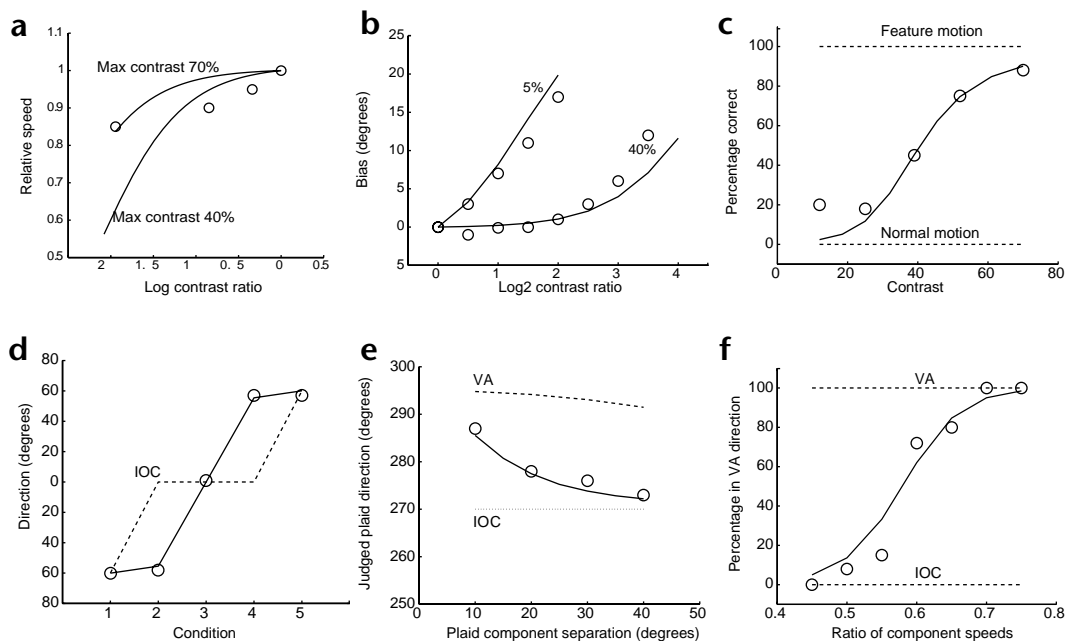
Fig. 5. Comparison of ideal observer (solid lines) to a variety of published psychophysical data (circles). (**a**) Contrast influence on perceived grating speed. Circles indicate the perceived speed of the lower-contrast grating relative to the higher-contrast grating, as a function of the contrast ratio. Solid lines show the predictions of the ideal observer for two different maximal contrasts (data from ref. 31). (**b**) Relative contrast influence on perceived plaid direction (data from ref. 20). (**c**) Contrast influence on perceived line direction (data from ref. 32). (**d**) Perceived direction of type I (conditions 1, 3, 5) versus type II (conditions 2, 4) plaids (data from ref. 17). Dotted line shows the IOC prediction. (**e**) Influence of relative orientation on perception of type II plaid motion (data from ref. 18). (**f**) Influence of relative speed on perception of type II plaids (data from ref. 19).

In all five conditions, the angular separation between the two gratings was 22.3°. In some conditions the two normal velocities were on different sides of the veridical motion (type I), whereas in others they were on the same side of the veridical motion (type II). Subjects saw type I plaids moving in the IOC direction and type II plaids moving in approximately the VA direction (~55° away from veridical direction). The authors of the original study explained their findings using a contrast-dependent combination of first-order and second-order motion analyzers[37].

The ideal observer also predicted different directions of motion for the two types of plaids at low contrast (**Fig. 5d**, solid line). The 'misperception' of type II plaids is similar to the perception of the narrow rhombus: the VA velocity is much slower than the IOC solution and hence it is favored at low contrasts. In the ideal observer, this bias toward the VA solution weakens with increasing contrast, as the likelihoods become narrower.

It has also been reported that the VA bias is more pronounced with shorter presentation durations[17]. We based our ideal observer on instantaneous measurements, so it is not affected by display duration. The formulation can easily be extended so that the ideal observer integrates information over time. This way, increased duration acts in a similar fashion to increased contrast: the longer the duration, the narrower the likelihood. Such an extended formulation predicts that the VA bias would decrease with increased duration. A similar effect of duration has been reported elsewhere[32], which would also be predicted by this extension of our model.

Influence of relative orientation on type II plaids
The perceived direction of a type II plaid depends strongly on the angle between the components. We replotted data from an experiment in which subjects reported the perceived direction

of plaids as a function of this angle, while pattern velocity was held constant[18] (**Fig. 5e**). The perceived direction is not consistent with a pure VA mechanism or a pure IOC mechanism. Instead, it shows a gradual shift from the VA to the IOC solution as the angle between the components increases. The solid line shows the prediction of the ideal observer (the direction of v^* in equation (1)). This situation is similar to the 'narrow' versus 'fat' rhombuses (**Fig. 4**). When two likelihoods whose constraint lines are nearly identical are multiplied, their product will be broad and hence have less of an influence on the posterior. By contrast, when two likelihoods have widely differing constraint lines, their product will be narrow and hence have greater influence on the posterior.

Influence of relative speed on type II plaids
The perceived direction of a plaid also depends on the relative speeds of the components. We plotted data from a single subject[19] who viewed a plaid with IOC and VA directions on opposite sides of upward, and reported whether the motion appeared to be more leftward or rightward (**Fig. 5f**). When the speeds of the two components were similar, the subject answered rightwards (consistent with the VA solution), but when the speeds were dissimilar, the subject answered leftwards (consistent with the IOC solution). We found that the ideal observer described by equation (1) shows a similar shift from leftward to rightward velocities. We again calculated a 'percentage correct' value for the ideal observer by assuming decision noise (**Fig. 5f**, solid line).

DISCUSSION
Research on visual motion analysis has yielded a tremendous amount of experimental data. When viewed in the context of existing rules such as IOC and VA, these data seem contradic-

tory, requiring an arbitrary combination scheme that applies the right rule in the right conditions. Such an approach can successfully fit the data, but is typically lacking in predictive power: with a complicated enough combination scheme one can model any experiment. More importantly, because these rules are not formulated directly on image measurements, it is not clear how one should generalize them for application to arbitrary spatiotemporal stimuli.

Here we have taken an alternative approach. We derived an optimal estimator for local image velocity using the standard assumption of intensity constancy and two additional assumptions: measurement noise and an *a priori* preference for slower velocities. We found, consistent with results in humans, that the motion estimates of this model include apparent biases and illusions. Moreover, the predicted non-veridical percept is quite similar to that exhibited by humans under the same circumstances. Although the model does not account for all of the existing data quantitatively, it correctly predicted a wide range of effects.

Our model does not provide a good quantitative fit to the data of **Fig. 5a** (see Results), which suggest a quasilinear dependence of perceived grating speed on contrast, and minimal dependence on total contrast. Our model has been extended by including a nonlinear 'gain control' function to map stimulus contrast into perceived contrast (F. Hurlimann, D. Kiper & M. Carandini, *Invest. Opthal. Vis. Sci. Suppl. Abstr.* **40**, 794, 2000). For each subject in that study, the authors measured a gain control function from contrast-discrimination experiments. They then used the perceived contrast rather than stimulus contrast as input to our model, and found that when these realistic representations of contrast were used, the quantitative predictions of the Bayesian model were in general agreement with the data. We also found, using a numerical search procedure, that a monotonic nonlinear gain control function enabled our model to better fit the results reported here and in ref. 31 (see **Supplementary Results** online).

One result[33] that is not predicted by our model is the finding that low-contrast gratings actually appear to move faster than high-contrast gratings for temporal frequencies above 8 Hz. However, the same author later was unable to reproduce this result using a forced-choice task[31], and concluded that the original finding was probably "an artifact of the experimental method with subjects making 'speed' matches based on some other criterion".

Our Bayesian estimator is meant as a perceptual model, and does not specify a particular implementation. Nevertheless, the solution can be instantiated using so-called motion energy mechanisms[28,38], and detailed models of the physiology of the motion pathway[24,25,28,39–41] suggest that a population of MT cells may be forming a representation of the local likelihood of velocity. In addition, we believe it should be possible to refine and justify the assumptions we have made. In particular, the prior distribution on velocity could be estimated empirically from the statistics of motion in the world. In a physiological implementation, the noise model should be replaced by one that more accurately reflects the uncertainties of neural responses.

Our model also suggests some future experiments. First, if the single free parameter is observer dependent (but otherwise constant), the magnitude of different illusions for the same subject should be correlated. For example, observers who greatly underestimate the speed of low-contrast gratings should also show a larger bias towards VA in type II plaids. Second, in all of our simulations we used only the maximum (or mean) of the posterior distribution. It would be interesting to test whether human per-

cepts reflect the shape of the full posterior distribution.

We have focused on an ideal observer for estimating a single two-dimensional translation. This model cannot estimate more complicated motions such as rotations and expansions, nor can it handle scenes containing multiple motions. Elsewhere, we describe an extended ideal observer for more general scenes with multiple motions[29]. We show that an ideal observer that assumes that velocity fields are 'slow and smooth'[42] can explain an even wider range of motion phenomena. In particular, the bias toward slower motions can sometimes account for one of the most critical issues in motion perception: the question of whether to combine measurements into a single coherent motion or assume that there are actually multiple motions (H. Farid & E. P. Simoncelli, *Invest. Opthal. Vis. Sci. Suppl. Abstr.* **35**, 1271, 1994).

Although the details of our model should certainly be refined and extended to handle more complicated phenomena, we believe the underlying principle will continue to hold: that many motion 'illusions' are not the result of sloppy computation by various components in the visual system, but rather a result of a coherent computational strategy that is optimal under reasonable assumptions.

METHODS

Most models of early motion extraction rely on an assumption of 'intensity conservation'. Under this assumption, the points in the world, as measured in the image, move but do not change their intensity over time. Mathematically, this is expressed as:

$$I(x,y,t) = I(x + v_x \Delta t, \ y + v_y \Delta t, \ t + \Delta t) \qquad (2)$$

where v_x and v_y are the components of the vector, v, describing the image velocity. If we assume that the observed image is noisy, then intensity is not conserved exactly. Thus, equation (2) becomes

$$I(x,y,t) = I(x + v_x \Delta t, \ y + v_y \Delta t, \ t + \Delta t) + \eta \qquad (3)$$

where η is a random variable representing noise.

We used equation (3) to derive the likelihood at location i, $P(I(x_i,y_i,t)|v_i)$. This required additional assumptions. We assumed the noise, η, is Gaussian with standard deviation σ. We further assumed that the velocity is constant in a small window around x_i,y_i and that the intensity surface $I(x,y,t)$ is sufficiently smooth that it can be approximated by a linear function for small temporal durations. We thus replaced $I(x + v_x \Delta t, \ y + v_y \Delta t, \ t + \Delta t)$ with its first-order Taylor series expansion, which gives:

$$P(I(x_i,y_i,t)|v_i) \propto$$
$$\exp\left(-\frac{1}{2\sigma^2} \int_{x,y} w_i(x,y) \ (I_x(x,y,t)v_x + I_y(x,y,t)v_y + I_t(x,y,t))^2 \ dx \, dy \right) \qquad (4)$$

where $\{I_x, I_y, I_t\}$ denote the spatial and temporal derivatives of the intensity function I, and $w_i(x,y)$ is a window centered on (x_i, y_i). The likelihoods shown in **Fig. 3** and **Fig. 4** are computed from equation (4) with $w(x,y)$ a small Gaussian window.

Finally, we assumed a prior favoring slow speeds:

$$P(v) \propto \exp(-\|v\|^2 / 2\sigma_p^2). \qquad (5)$$

The posterior probability of a velocity was computed by combining the likelihood and prior using Bayes' rule. Because we assumed that the noise is independent over spatial location, the total likelihood function is just a product of likelihoods:

$$P(v|I) \propto P(v) \prod_i P(I(x_i,y_i,t) \ |v), \qquad (6)$$

where the product is taken over all locations i that are moving with a common velocity ($v_i = v$). Substituting equations (4) and (5) into equaion (6),

$$P(v|I) \propto \exp\left(-\|v\|^2/2\sigma_p^2 - \frac{1}{2\sigma^2} \int_{x,y} \sum_i w_i(x,y) \left(I_x(x,y)v_x + I_y(x,y)v_y + I_t\right)^2 dx\, dy \right).$$

Here we assumed the entire image moves according to a single translational velocity, and so summed over all spatial positions. In this case, $\sum_i w_i(x,y)$ is a constant, so the posterior probability is given by:

$$P(v|I) \propto \exp\left(-\|v\|^2/2\sigma_p^2 - \frac{1}{2\sigma^2} \int_{x,y} \left(I(x,y)\, v_x + I_y(x,y)v_y + I_t\right)^2 dx\, dy \right)$$

To find the most probable velocity, we replaced the integral with a discrete sum, took the logarithm of the posterior, differentiated it with respect to v and set the derivative equal to zero. The logarithm of the posterior is quadratic in v so that the solution can be written in closed form using standard linear algebra. The result is given in equation (1).

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

1. Nakayama, K. Biological image motion processing: a review. *Vision Res.* **25**, 625–660 (1985).
2. Horn, B. K. P. & Schunck, B. G. Determining optical flow. *Artif. Intell.* **17**(1–3), 185–203 (1981).
3. Lucas, B. D. & Kanade, T. An iterative image registration technique with an application to stereo vision. in *Proceedings of the 7th International Joint Conference on Artificial Intelligence* 674–679 (Morgan-Kaufmann, San Fransisco, 1981).
4. Wuerger, S., Shapley, R. & Rubin, N. On the visually perceived direction of motion by Hans Wallach: 60 years later. *Perception* **25**, 1317–1367 (1996).
5. Wallach, H. Ueber visuell whargenommene bewegungrichtung. *Psychol. Forsch.* **20**, 325–380 (1935).
6. Marr, D. & Ullman, S. Directional selectivity and its use in early visual processing. *Proc. R. Soc. Lond. B Biol. Sci.* **211**, 151–180 (1981).
7. Adelson, E. & Movshon, J. Phenomenal coherence of moving visual patterns. *Nature* **300**, 523–525 (1982).
8. Movshon, A., Adelson, E., Gizzi, M. & Newsome, W. The analysis of moving visual patterns. *Exp. Brain Res.* **11**, 117–152 (1986).
9. Welch, L. The perception of moving plaids reveals two processing stages. *Nature* **337**, 734–736 (1989).
10. Morgan, M. Spatial filtering precedes motion detection. *Nature* **355**, 344–346 (1992).
11. Schrater, P., Knill, D. & Simoncelli, E. Mechanisms of visual motion detection. *Nat. Neurosci.* **3**, 64–68 (2000).
12. Rodman, H. & Albright, T. Single-unit analysis of pattern motion selective properties in the middle temporal visual area MT. *Exp. Brain Res.* **75**, 53–64 (1989).
13. Movshon, J. A. & Newsome, W. T. Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *Vis. Neurosci.* **16**, 7733–7741 (1996).
14. Okamoto, H. *et al.* MT neurons in the macaque exhibited two types of bimodal direction tuning as predicted by a model for visual motion detection. *Vision Res.* **39**, 3465–3479 (1999).
15. Ferrera, V. & Wilson, H. Perceived direction of moving two-dimensional patterns. *Vision Res.* **30**, 273–287 (1990).
16. Mingolla, E., Todd, J. & Norman, J. The perception of globally coherent motion. *Vision Res.* **32**, 1015–1031 (1992).
17. Yo, C. & Wilson, H. Perceived direction of moving two-dimensional patterns depends on duration, contrast, and eccentricity. *Vision Res.* **32**, 135–147 (1992).
18. Burke, D. & Wenderoth, P. The effect of interactions between one-dimensional component gratings on two dimensional motion perception. *Vision Res.* **33**, 343–350 (1993).
19. Bowns, L. Evidence for a feature tracking explanation of why type II plaids move in the vector sum directions at short durations. *Vision Res.* **36**, 3685–3694 (1996).
20. Stone, L., Watson, A. & Mulligan, J. Effect of contrast on the perceived direction of a moving plaid. *Vision Res.* **30**, 1049–1067 (1990).
21. Rubin, N. & Hochstein, S. Isolating the effect of one-dimensional motion signals on the perceived direction of moving two-dimensional objects. *Vision Res.* **33**, 1385–1396 (1993).
22. Helmholtz, H. *Treatise on Physiological Optics* (Thoemmes, Bristol, UK, 2000; original publication 1866).
23. Knill, D. & Richards, W. *Perception as Bayesian Inference* (Cambridge Univ. Press, Cambridge, 1996).
24. Ascher, D. & Grzywacz, N. A Bayesian model for the measurement of visual velocity. *Vision Res.* **40**, 3427–3434 (2000).
25. Koechlin, E., Anton, J. L. & Burnod, Y. Bayesian inference in populations of cortical neurons: a model of motion integration and segmentation in area MT. *Biol. Cybern.* **80**, 25–44 (1999).
26. Simoncelli, E., Adelson, E. & Heeger, D. in *Proc. IEEE Conf. Comput. Vision Pattern Recog.* 310–315 (IEEE, Washington DC, 1991).
27. Heeger, D. J. & Simoncelli, E. P. in *Spatial Vision in Humans and Robots* Ch. 19 (eds. Harris, L. & Jenkin, M.) 367–392 (Cambridge Univ. Press, 1994).
28. Simoncelli, E. P. *Distributed Representation and Analysis of Visual Motion.* Thesis, Massachusetts Institute of Technology (1993).
29. Weiss, Y. *Bayesian Motion Estimation and Segmentation.* Thesis, Massachusetts Institute of Technology (1998).
30. Ullman, S. *The Interpretation of Visual Motion* (MIT Press, Cambridge, Massachusetts, 1979).
31. Stone, L. & Thompson, P. Human speed perception is contrast dependent. *Vision Res.* **32**, 1535–1549 (1990).
32. Lorenceau, J., Shiffrar, M., Wells, N. & Castet, E. Different motion sensitive units are involved in recovering the direction of moving lines. *Vision Res.* **33**, 1207–1217 (1992).
33. Thompson, P. Perceived rate of movement depends on contrast. *Vision Res.* **22**, 377–380 (1982).
34. Thompson, P., Stone, L. & Swash, S. Speed estimates from grating patches are not contrast normalized. *Vision Res.* **36**, 667–674 (1996).
35. Blakemore, M. & Snowden, R. The effect of contrast upon perceived speed: a general phenomenon? *Perception* **28**, 33–48 (1999).
36. Snowden, R. N., Stimpson, N. & Ruddle, S. Speed perception fogs up as visibility drops. *Nature* **392**, 450 (1998).
37. Wilson, H., Ferrera, V. & Yo, C. A psychophysically motivated model for two-dimensional motion perception. *Vis. Neurosci.* **9**, 79–97 (1992).
38. Weiss, Y. & Fleet, D. in *Probabilistic Models of the Brain* Ch. 4 (eds. Rao, R., Olshausen, B. & Lewicki, M.) 77–96 (MIT Press, Cambridge, Massachusetts, 2002).
39. Nowlan, S. J. & Sejnowski, T. J. A selection model for motion processing in area MT of primates. *J. Neurosci.* **15**, 1195–1214 (1995).
40. Simoncelli, E. & Heeger, D. A model of neuronal responses in visual area MT. *Vision Res.* **38**, 743–761 (1998).
41. Pouget, A., Dayan, P. & Zemel, R. Information processing with population codes. *Nat. Rev. Neurosci.* **1**, 125–32 (2000).
42. Grzywacz, N. & Yuille, A. Theories for the visual perception of local velocity and coherent motion. in *Computational Models of Visual Processing* (eds. Landy, J. & Movshon, J.) 231–252 (MIT Press, Cambridge, Massachusetts, 1991).