# The economics of motion perception and invariants of visual sensitivity

**Sergei Gepshtein**
Perceptual Dynamics Laboratory, Brain Science Institute, RIKEN, Wakoshi, Saitama, Japan

**Ivan Tyukin**
Perceptual Dynamics Laboratory, Brain Science Institute, RIKEN, Wakoshi, Saitama, Japan, & Department of Mathematics, University of Leicester, Leicester, UK

**Michael Kubovy**
Department of Psychology, University of Virginia, Charlottesville, VA, USA

Neural systems face the challenge of optimizing their performance with limited resources, just as economic systems do. Here, we use tools of neoclassical economic theory to explore how a frugal visual system should use a limited number of neurons to optimize perception of motion. The theory prescribes that vision should allocate its resources to different conditions of stimulation according to the degree of balance between measurement uncertainties and stimulus uncertainties. We find that human vision approximately follows the optimal prescription. The equilibrium theory explains why human visual sensitivity is distributed the way it is and why qualitatively different regimes of apparent motion are observed at different speeds. The theory offers a new normative framework for understanding the mechanisms of visual sensitivity at the threshold of visibility and above the threshold and predicts large-scale changes in visual sensitivity in response to changes in the statistics of stimulation and system goals.

Keywords: uncertainty principle, equilibrium, visual sensitivity, normative model, optimality, apparent motion, motion adaptation, utility

Citation: Gepshtein, S., Tyukin, I., & Kubovy, M. (2007). The economics of motion perception and invariants of visual sensitivity. *Journal of Vision, 7*(8):8, 1–18, http://journalofvision.org/7/8/8/, doi:10.1167/7.8.8.

## Introduction

Consider an organism that uses vision to estimate the speeds of moving objects. With unlimited resources, it could estimate the speeds by having a separate mechanism to measure each speed, resulting in excellent precision. With minimal resources, it might have to make do with a single mechanism to measure all speeds, resulting in poor precision. How can the visual system find the right balance between frugality and precision?

In this article, we find the balance using tools developed in the neoclassical economic theory. Our approach rests on the notion of equilibrium, introduced into the economic theory from mechanics (Pareto, 1906). In economics, this method allows one to find an optimal balance in consumption of incommensurable goods ("apples and oranges") for a customer or a market with a limited budget. The allocation of resources to bundles of goods is optimal when two conditions are met: (1) Satisfaction from one component of the bundle cannot improve without reducing satisfaction from some other component, and (2) satisfaction from all the goods is the highest possible. Our analysis of motion perception using this approach leads to equations of equilibrium very similar to the equations in economics (Table 1). Now, "apples and oranges" correspond to the parameters of optical stimulation, and the "degree of dissatisfaction" corresponds to the errors—or the amount of uncertainty[1]—in estimating the parameters (Gepshtein & Tyukin, 2006; Gepshtein, Tyukin, Kubovy, & van Leeuwen, 2006). Just as a consumer with limited financial resources seeks to minimize his or her dissatisfaction from a basket of fruit, we assume that vision, with a limited pool of speed-tuned neurons, seeks to minimize measurement errors.

The well-known fact that visual sensitivity to the parameters of stimulation varies across the parameters is a manifestation of the differential allocation of visual resources. A comprehensive summary of the visual sensitivity is the spatiotemporal contrast sensitivity function (Kelly, 1979, 1994). We plot this function in the logarithmic space–time distance coordinates in Figure 1A. (See Appendix A for details of its construction.) In this format, the different speeds are represented by the parallel lines, called *speed lines*. In Figure 1A, we show two characteristics of visual sensitivity: the maximal sensitivity set and the isosensitivity sets. The maximal sensitivity set is represented by the grey hyperbolic curve that runs

doi: 10.1167/7.8.8 Received November 28, 2006; published June 21, 2007 ISSN 1534-7362 

| Framework | Role | Form of equilibrium | Equation |
|---|---|---|---|
| Local | Spendthrift extreme | $\partial U_S v + \partial U_T = 0$ | 9 |
| Total integration | Frugal extreme | $\partial U_S v_e + \partial U_T = 0$ | 12 |
| Weighted | Plausible compromise | $\partial U_S g(v) + \partial U_T = 0$ | 18 |

Table 1. Equilibrium in the three optimization frameworks.

across the speeds. The isosensitivity sets are represented by contours in different colors. Each contour connects the parameters of spatiotemporal stimuli that reach the threshold of visibility at the same contrast. Although these data were obtained with drifting sinusoidal gratings at the threshold of visibility, the estimates of maximal sensitivity are consistent with results produced by other methods (Burr & Ross, 1982; Kelly, 1994; Newsome, Mikami, & Wurtz, 1986; van de Grind, Koenderink, & van Doorn, 1986; van Doorn & Koenderink, 1982a, 1982b; reviewed in Nakayama, 1985).

The visual system estimates speed using a pool of neurons whose receptive fields can be tuned to different speeds (Nover, Anderson, & DeAngelis, 2005; Perrone & Thiele, 2001; Priebe, Cassanello, & Lisberger, 2003; Simoncelli & Heeger, 1998) at different locations in the distance plot (Figure 1). The uncertainty associated with this estimation varies across the locations. A system with limited resources should maximize the quality of its measurements per unit of resources and thus allocate more resources where the expected uncertainty is low. In other words, the amount of resources allocated to different conditions ($S$, $T$) should be inversely related to the amount of errors expected at these conditions. In agreement with the neural network theory (Barron, 1993; Park & Sandberg, 1993), we expect that a larger number of neurons tuned to a particular value of a parameter leads to a better ability to discriminate signals around that value and, consequently, to a higher sensitivity in its vicinity. Thus, if we knew how the uncertainties, $U(S, T)$, were distributed across space and time, we could predict how vision should distribute its sensitivity.

In this work, we pursue this normative approach. First, we estimate the uncertainty intrinsic to the measurement

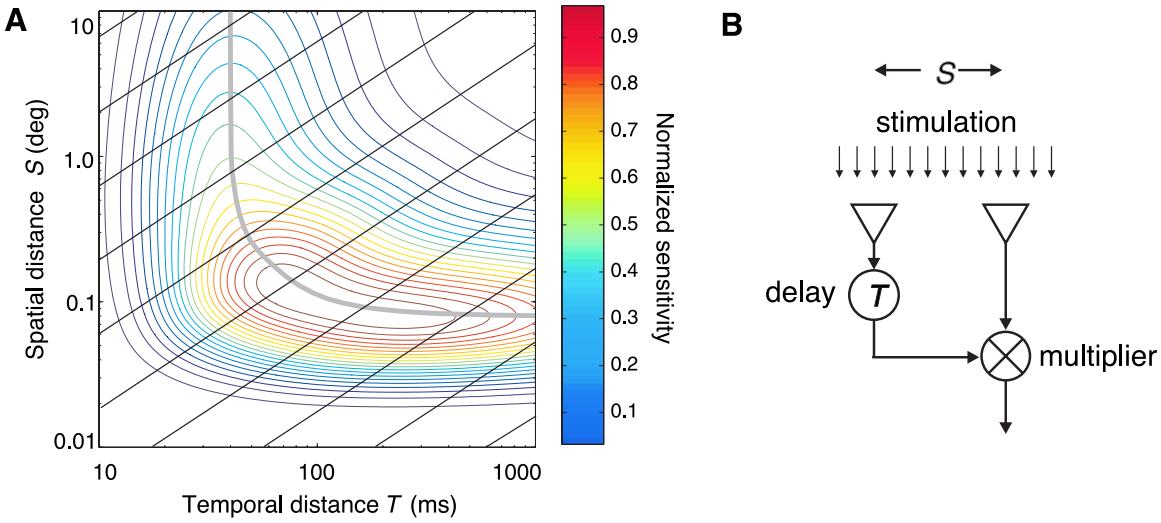


Figure 1. Human spatiotemporal sensitivity. (A) Human spatiotemporal contrast sensitivity as a function of spatial and temporal distance (space–time distance plot). The colored contours represent constant levels of contrast sensitivity; their colors represent normalized sensitivity. The grey hyperbolic curve represents the conditions of maximal sensitivity. The parallel *speeds lines* represent different speeds of motion. We plot sensitivity in space–time (and not in the frequency domain, as it is often done) because we relate sensitivity to perception of apparent motion, which is easier to understand in space–time. The data are from Kelly (1979). The prediction for allocation of sensitivities by our normative theory is plotted in Figure 6. (B) A conceptual scheme of motion detector (Reichardt, 1961, 1969) tuned to spatial and temporal parameters $S$ and $T$. Each point in Panel A corresponds to a pair of these parameters. The detector is stimulated the strongest when a signal travels spatial distance $S$ (from left to right in the illustration) over temporal interval $T$. Assuming a slightly more sophisticated version of the detector—the standard model (Adelson & Bergen, 1985; van Santen & Sperling, 1984)—the detector's tuning characteristics and the results of Kelly in Panel A are converted between the space–time and frequency domains (Nakayama, 1985; Appendix A).

of spatiotemporal signals (*measurement uncertainty*; Gabor, 1946). Second, we predict at which locations in the space–time distance plot the conditions are most favorable for estimating speed, taking into account estimates of uncertainty about the speed of stimulation (*stimulus uncertainty*; Dong & Atick, 1995). We start from constraints that apply to any measurement whatsoever and deduce how the visual system might achieve the compromise between frugality and precision in face of the uncertainties. We find that the invariant properties of the optimal and uniformly suboptimal conditions for motion measurement predicted by the theory are similar to the maximal-sensitivity and isosensitivity conditions found in biological vision.

# Measurement uncertainty

## Uncertainty principle of measurement

Consider receptive fields sensitive to a range of spatial and temporal frequencies, $\Delta_f > 0$, and a range of locations in space and time, $\Delta_x > 0$. These ranges produce uncertainty with respect to the content[2] and location of a stimulus. By the uncertainty principle of measurement (Gabor, 1946; Resnikoff, 1989), the product of these uncertainties (spatial *or* temporal) cannot exceed a positive constant, $C_1$:

$$\Delta_f \times \Delta_x \geq C_1. \tag{1}$$

In a system that is performing at its best (i.e., where $\Delta_f \times \Delta_x = C_1$) one uncertainty cannot decrease without the other increasing. Henceforth, we will express the changes of uncertainty as equations and not as inequalities because our goal is to predict the best performance of the system.

Suppose $\Delta_f$ obeys the invariance of relative uncertainties:

$$\Delta_f / f = C_2, \tag{2}$$

where $C_2 > 0$ is a constant, as it was shown for visual receptive fields (e.g., Kulikowski, Marcelja, & Bishop, 1982). This property represents a scale invariance of measurement error, similar to Weber's law. In contrast to the generally valid uncertainty principle, Equation 2 summarizes empirical observations, whose generality is not confirmed. Thus, Equation 2 might serve as a postulate—or an assumption—from which we derive the function describing how measurement uncertainties change across the conditions of stimulation. (We relax this assumption below.) From Equations 1 and 2, it follows:

$$\begin{aligned} \Delta_x &= C_1/(C_2 \times f), \\ \Delta_f &= C_2 \times f. \end{aligned} \tag{3}$$

Assuming additivity of uncertainties, from Equation 3 we can estimate how uncertainty changes across the entire range of spatial or temporal parameters:

$$U = k_1\Delta_x + k_2\Delta_f = k_1\left(\frac{C_1}{C_2 \times f}\right) + k_2(C_2 \times f), \tag{4}$$

where $\Delta_x$ is an interval of spatial or temporal locations ($\Delta_S$ or $\Delta_T$), $\Delta_f$ is an interval of spatial or temporal frequencies ($\Delta_{f_S}$ or $\Delta_{f_T}$), and $k_i$ are unit coefficients that bring the terms to the same units. Equation 4 indicates that the measurement uncertainty varies as a composition of decaying and growing functions of frequency (Figure 1A).

For a visual system that samples signals using Gabor filters (i.e., filters obtained by multiplication of Gaussian and harmonic functions; e.g., Daugman, 1985; Jones & Palmer, 1987; MacKay, 1981), we can estimate the shape of uncertainty function without the assumption of invariance of relative uncertainties (Equation 2). Gabor (1946) showed that in such a system, there exists a simple relationship between the uncertainty in space–time and in frequency domain (his Equation 1.27). We formulate that relationship in our terms as

$$\Delta_x = \lambda_x/\alpha, \quad \Delta_f = \lambda_f\alpha, \tag{5}$$

where $\alpha$ corresponds to distance (spatial or temporal, against which we plotted sensitivity in Figure 1A) and $\lambda_i$ are the weights (as in Equation 7 below). By adding uncertainties as we did in Equation 4, we again find that the composite uncertainty decays or grows as a function of frequency over different intervals of frequency (which is inversely related to $\alpha$). Thus, these features of the composite uncertainty function do not depend on the assumption of invariance of relative uncertainties.

## Spatiotemporal uncertainty function

By applying Equation 4 separately to spatial and temporal uncertainties, we derive the spatial and temporal uncertainty functions:

$$\begin{aligned} U_S &= k_{1,S} \times \Delta_S + k_{2,S} \times \Delta_{f_S} \\ &= k_{1,S}\left(\frac{C_{1,S}}{C_{2,S} \times f_S}\right) + k_{2,S}\left(C_{2,S} \times f_S\right), \\ U_T &= k_{1,T} \times \Delta_T + k_{2,T} \times \Delta_{f_T} \\ &= k_{1,T}\left(\frac{C_{1,T}}{C_{2,T} \times f_T}\right) + k_{2,T}\left(C_{2,T} \times f_T\right). \end{aligned} \tag{6}$$

By the assumption of additivity, from Equation 6, we obtain the generalized spatiotemporal uncertainty function:

$$\begin{aligned} U_{ST}(S,T) &= U_S + U_T, \quad U_S = \frac{\lambda_1}{S} + \lambda_3 S, \\ U_T &= \frac{\lambda_2}{T} + \lambda_4 T, \end{aligned} \tag{7}$$

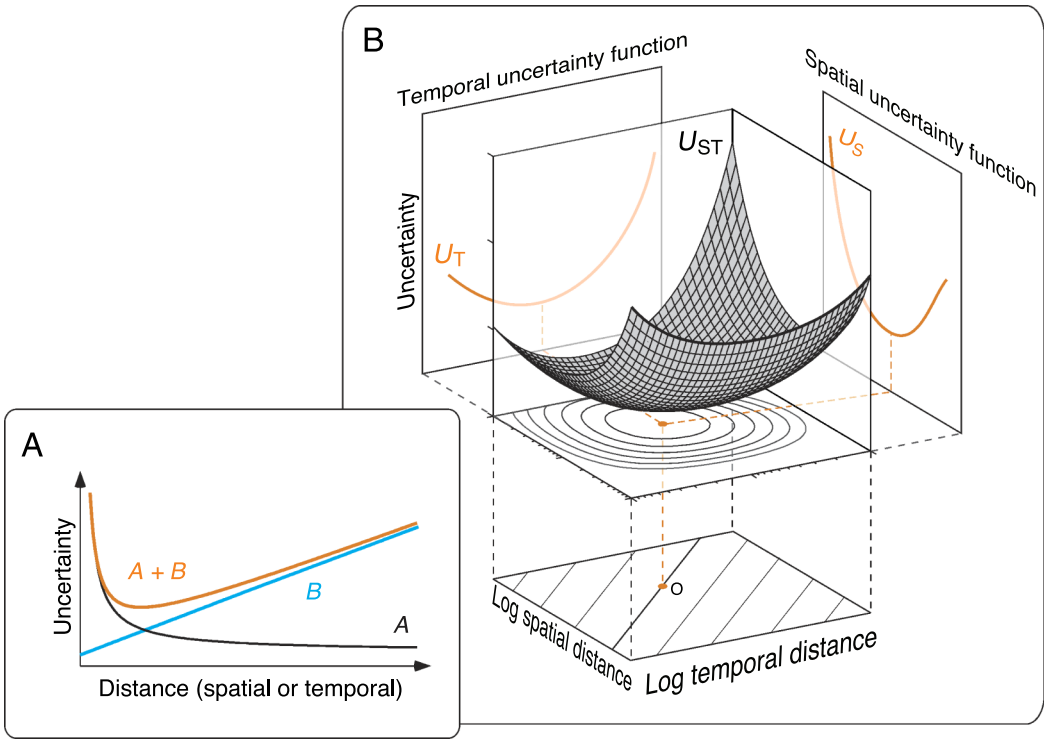


Figure 2. Uncertainty functions. (A) Functions *A* and *B* represent uncertainties associated with measuring the frequency content of signals and localizing signals, respectively. Measurements of large spatial or temporal distances, using large receptive fields, suffer from uncertainty about the location of signals in space or time (large values of function *B* at large distances). Measurements of small spatial or temporal distances, using small receptive fields, suffer from uncertainty about the location of signals in the frequency domain, spatial or temporal (large values of Function *A* at small distances). We describe the growth of uncertainty toward the extremes of spatial and temporal dimensions of stimulation by assuming additivity of the small- and large-scale uncertainties (Function *A* + Function *B*). (B) Adding the spatial and temporal uncertainties ($U_S$ and $U_T$) in space–time yields a spatiotemporal uncertainty function, $U_{ST}$. ($U_S$ and $U_T$ are shown on the side panels; each is a replica of Function *A* + *B* in Panel A.) The minimum of $U_{ST}$—global optimum *O*—and the level curves of $U_{ST}$ are shown in a space–time plot in the upper horizontal plane. Shown in a space–time plot in the lower horizontal plane are speed lines, which are parallel to each other in logarithmic coordinates. The global optimum projects on one of the speed lines. At that point, the estimates of speed are least affected by the uncertainties. There are similarly optimal points on other speed lines; these are the points where spatial and temporal uncertainties are in equilibrium (Figure 3, Table 1).

where we converted $f_S$ and $f_T$ into $S$ and $T$ (Appendix A) and replaced the constants of Equation 6 and the constants of the frequency–distance conversion by constants $\lambda_i$. It is convenient to think of these constants as the weights expressing how much each term of Equation 7 affects system goals.

We found that the shapes of predicted isosensitivity contours, which we derive in Figure 5, are similar to the shape of human isosensitivity contours in Figure 6 when $\lambda_3$ and $\lambda_4$ were roughly two orders of magnitude greater than $\lambda_1$ and $\lambda_2$. The fact that the weights of terms concerning localization in space–time ($\lambda_3$, $\lambda_4$) are large in comparison with the weights of terms concerning frequency identification ($\lambda_1$, $\lambda_2$) suggests that the spatiotemporal sensitivity function in Figure 6 reflects visual behavior whose implicit goal has to do more with signal localization than signal identification.

In Figure 2B, we plot the spatial and temporal uncertainty functions $U_S$ and $U_T$ and the joint spatiotemporal uncertainty function $U_{ST}$. Each one-dimensional uncertainty function on the side panels of Figure 2B shows that uncertainties grow toward the high and low ends of each dimension. The composition of uncertainty functions is illustrated in Figure 2A:

- At the low end of each dimension, small receptive fields allow precise estimation of signal location in space or time (low values of Function *A*), but not in the spatial or temporal frequency domains (high values of Function *B*).
- In contrast, at the high end of each dimension, the receptive fields allow the precise estimation of signal location in the spatial or temporal frequency domains (low values of Function *A*), but not in space or time (high values of Function *B*).

### *Consequences of the additivity assumptions*

In Equation 4, we assumed additivity of uncertainties because for a system at the optimal limit of its performance,

multiplication of uncertainties within a dimension (spatial or temporal) amounts to adding a constant (Equation 1). The role of additivity in Equation 7 is different. Multiplication of spatial and temporal uncertainties in Equation 7 yields a new inseparable space–time term in the uncertainty function; this term does not change qualitatively significant predictions of the theory (page 8).

## Optimal conditions for speed estimation

The visual system would perform optimally if it used small receptive fields to estimate location in space–time and large receptive fields to estimate location in the frequency domain, perhaps using specialized subsystems to accomplish different tasks. But let us consider a frugal system with scarce resources that cannot afford a separation to subsystems. Such a system should compromise and use the same receptive fields to measure several kinds of information: minimizing the aforementioned uncertainties and also trying to minimize the uncertainty of speed estimation.

The measurement uncertainties described by the spatiotemporal uncertainty function constrain the visual system's ability to estimate speed. The effect of measurement uncertainty on speed estimation is smallest at the minimum of $U_{ST}$ (red circle) in Figure 2. This *global optimum* falls on one of the speed lines. Because speed can be measured with least uncertainty at that point, the visual system should allocate more resources for measuring speed at that point than at any other point on this speed line. But where are the similarly optimal conditions for measuring other speeds?

Because a straightforward answer to this question is biologically implausible (as we show next), we answer it in two steps. First, we establish two extreme optimality frameworks for speed estimation: One is overly spendthrift but estimates speed with infinite precision; the other is overly frugal but is imprecise at almost any speed. Second, we find a principled balance between these extremes and thus predict biologically plausible optimal conditions for the estimation of every speed.

### Local optimization: A spendthrift extreme

We find an optimal condition on each speed line, similar to the condition of global optimum, by noting that at the global optimum, spatial and temporal uncertainties are exactly in balance: A change in spatiotemporal uncertainty ($dU_{ST}$; Equation 7) due to changes of $U_S$ and $U_T$ is the total derivative of $U_{ST}$:

$$dU_{ST} = \partial U_S\, dS + \partial U_T\, dT, \tag{8}$$

where $\partial U_S = \partial U_{ST}/\partial S$ and $\partial U_T = \partial U_{ST}/\partial T$ are the spatial and temporal partial derivatives of $U_{ST}$, respectively. Because the spatial and temporal uncertainties balance each other where the total derivative is zero, the equilibrium for every speed $v = S/T$ (which we write as $dS = v\,dT$ for $dT \neq 0$) is

$$\partial U_S v + \partial U_T = 0. \tag{9}$$

By substituting $U_S$ and $U_T$ from Equation 7 into Equation 9, it follows that the conditions of equilibrium are satisfied by a hyperbola in the space–time plot (gray curve in Figure 4):

$$S = \sqrt{\frac{\lambda_1 T^2 v}{(\lambda_3 v + \lambda_4)T^2 - \lambda_2}}. \tag{10}$$

As we show in Figure 3, the hyperbola for every speed passes through two special points: One is the global optimum (the orange circle), common to all the hyperbolas; the other is the local optimum for the speed (circles marked $O_i$ in the insets). The latter point is unique for most speeds. (Only for the speed whose speed line passes through the global optimum do the global and local optima coincide; Figure 3, inset A2.)

The optimal point on each speed line (local optimum) is the intersection of the speed line and the hyperbola representing the solution of the equation of balance of uncertainties for that speed. We find such balance points on each speed line and thus discover the *local optimal set* for estimating all the speeds; it is shown as a thick black curve in Figure 3.

Note that we find the local minima in Equation 8 by varying uncertainty functions in space and time (i.e., balancing $dS$ and $dT$) rather than in the frequency domain (i.e., balancing $df_S$ and $df_T$) because we are interested in the optimal conditions for speed estimation rather than the optimal conditions for measuring signals' frequency contents.

Local optimization is biologically implausible for two reasons:

R1. Local optimization assumes that the visual system can tune its processing units for each speed independent of other speeds. This assumption must be false: Any estimate of speed must involve some spatial and temporal integration because biological receptive fields are extended in space and time.

R2. Local optimization assumes that all speeds are equally important for perception. This assumption must be false for two reasons: (1) The distribution of speeds in the perceptual ecology is not uniform (Dong & Atick, 1995) and (2) some speeds are more important to the organism than others.

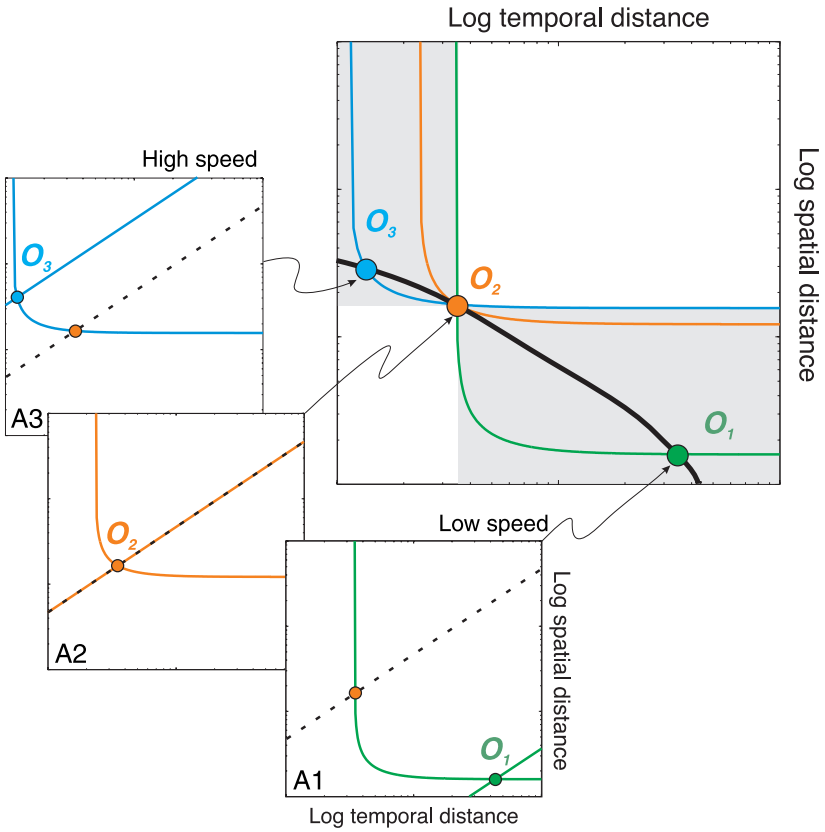


Figure 3. Local optimization. The local optimum for estimating speed $v_i$ is intersection $O_i$ of a hyperbola (which consists of points where the spatial and temporal uncertainties are balanced) and the speed line of $v_i$. Insets A1–A3 show three such optima. The set of all the intersections is the *local optimal set,* shown as a thick black curve. The orange circle is the point of global optimum of uncertainty from Figure 2. The dashed line is the speed line that passes through global optimum.

## Optimization by "total integration": A frugal extreme

To understand how the integration across speeds in biological vision affects the shape of the optimal set for speed estimation, we first consider an extreme case: integration across all speeds. This total-integration framework represents an overly frugal system, whose resources are extremely limited and whose speed resolution is poor. The limiting case of such a system is the one with a single speed-tuned mechanism sensitive to the entire range of speeds.

To find the optimal set predicted for the overly frugal total-integration framework, we integrate the contribution of all speeds into the optimization process and use each speed's prevalence in the perceptual ecology as a weight in the integration. The effect of weighting the contribution of each speed by its prevalence in the perceptual ecology is

$$\begin{aligned} dU_I &= \int_0^\infty p(v)\, dU_{ST}\, dv = 0 \\ \Rightarrow \quad &\int_0^\infty p(v)(\partial U_S v + \partial U_T)\, dv = 0, \end{aligned} \tag{11}$$

where $p(v)$ is the distribution of speeds in the natural stimulation (Dong & Atick, 1995). By rewriting the latter expression as

$$\left(\int_0^\infty p(v)\, v dv\right)\partial U_S + \left(\int_0^\infty p(v) dv\right)\partial U_T = 0, \tag{12}$$

taking into account the fact that $\int_0^\infty p(v)dv = 1$, and introducing *expected speed*

$$v_e = \int_0^\infty p(v)\, v dv, \tag{13}$$

we obtain

$$\partial U_S v_e + \partial U_T = 0. \tag{14}$$

Thus, the integration yields an optimal set whose mathematical form is that of equilibrium (Table 1).

The total-integration framework is not vulnerable to the objections we raised about local optimization because many speeds contribute to the optimization process

(response to R1) and because the contribution of each speed is weighted by the distribution of speeds in the perceptual ecology (response to R2; in spirit of the Bayesian approach: Knill & Richards, 1996; Maloney, 2002; Weiss, Simoncelli, & Adelson, 2002). In the ecological distribution, low speeds are more likely than high ones (Dong & Atick, 1995).

The optimal set according to total integration is invariantly a hyperbola in the distance plot: the grey curve in Figure 4. Why a hyperbola? As we saw in Figure 3, the conditions of equilibrium for *each speed* form a hyperbola in the space–time distance plot (Equations 9 and 10). The local optimization framework can afford finding the condition of equilibrium for each speed. Because multiple equilibria—one per speed—contributed to the local optimal set, the form of that set is generally different from hyperbolic. In contrast, the total integration framework cannot access individual speeds. It can only estimate the weighted mean of all the speeds that contribute to the integration. The weighted mean of speed is the most likely speed in the perceptual ecology: the expected speed, $v_e$. Thus, a single speed dominates optimization in this framework, because of which the optimal set is a single hyperbola.

This result, that a single speed dominates optimization in this framework, is remarkable. It means that the shape of the distribution of speeds in the perceptual ecology matters only inasmuch as it determines the mathematical expectation of that distribution: the expected speed. As we show below, a realistic compromise between the two extreme optimization frameworks yields an optimal set very similar to the set predicted by the total-integration framework. It is therefore the expected speed, and not the shape of the distribution of speeds, that controls the properties of the optimal set predicted by our theory.

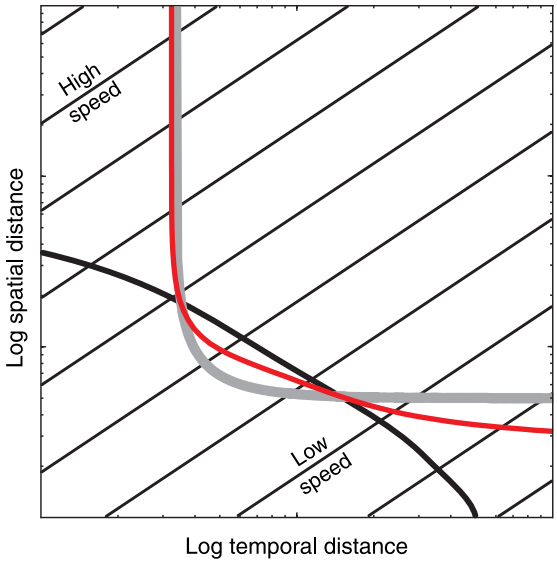


Figure 4. Three optimal sets for speed estimation. The black curve represents the *local optimal set*, as in Figure 3. The grey hyperbola represents the *total-integration optimal set*. The red curve represents the *weighted optimal set,* which is a realistic compromise that falls between the two extreme sets. At low speeds, the realistic optimal set is closer to the local optimal set than to the total-integration optimal set because low speeds prevail in the perceptual ecology.

Notice that the two extreme frameworks—local optimization and total integration—are dominated by different uncertainties: The spendthrift local-optimization framework can afford to allocate as much resources as needed to measure every speed, so its optimal set depends only on $U(S, T)$, that is, only on the measurement uncertainty. By contrast, the frugal total-integration framework must allocate its resources carefully. It must take into account the statistics of stimulation—the stimulus uncertainty—so it can measure speed with optimal precision only at the most likely speed $v_e$. The frugal framework is more sophisticated than the spendthrift one: its optimal set is determined by both the measurement and stimulus uncertainties.

## Weighted optimization: A realistic compromise

A realistic optimal set must lie between the spendthrift and the frugal. The more limited the resources of a system, the more it should rely on stimulus uncertainty and the more closely the optimal set should approach the prescription by frugal total integration. On the other hand, the more speed-tuned mechanisms a system can afford, the more such mechanisms it can allocate to the more common spatiotemporal parameters, where stimulus uncertainty is relatively low. Thus, for speeds prevalent in the stimulation, the optimal set should be closer to the set prescribed by local optimization than for other speeds. Therefore, the optimal strategy for resource allocation across speeds depends on stimulus uncertainty: If a speed is likely (low stimulus uncertainty), its optimal points should be closer to the prediction of local optimization. If a speed is unlikely (high stimulus uncertainty), its optimal points should be closer to the prediction of total integration.

We model this compromise by taking a linear combination of the conditions for optimization for the local ($dU_{ST}$) and total-integration ($dU_I$) frameworks:

$$w(v)\,dU_{ST}(v) + [1 - w(v)]\,dU_I = 0. \quad (15)$$

Remarkably, this method of combination preserves the equilibrium of spatial and temporal uncertainties (Table 1), as we show next.

The effect of distribution of speeds $p(v)$ on the optimization of speed measurement depends on how precisely visual mechanisms are tuned to speed. In our formulation of the maximal-sensitivity set in Equation 15, precision of tuning is represented by the interval of

integration across speeds; the interval is denoted as $\Omega(v)$ below. First, we derive the integration intervals from the principle of uncertainty of measurement (as we did to obtain Equation 3). We find that the boundaries of integration should grow with speed. Then, we find the weight $w(v)$ by integrating the contributions of different speeds over $\Omega(v)$:

$$w(v) = \int_{\Omega(v)} p(v_1)\, dv_1. \tag{16}$$

### Boundaries of integration

We estimate the boundaries of integration intervals the same way we have estimated the uncertainties associated with the spatial and temporal dimensions of stimulation (Equation 3). We rewrite Equation 2 in terms of spatial and temporal distances:

$$\begin{aligned} \Delta_S/S &= C_S, \\ \Delta_T/T &= C_T, \end{aligned} \tag{17}$$

where $C_S$ and $C_T$ are constants. From this, we estimate the interval of uncertainty of speed. Let $|u_a|$ be the absolute uncertainty (uncertainty pedestal), which does not depend on speed. Then, we write the lower ($v_a$) and upper ($v_b$) boundaries of the interval as

$$\begin{aligned} v_a &= -u_a + (S - \Delta_S)/(T + \Delta_T), \\ v_b &= u_a + (S + \Delta_S)/(T - \Delta_T), \end{aligned} \tag{18}$$

where the first equation describes the minimal value of measurement with an error in spatial and temporal estimates, and the second equation describes the maximal value of such measurement. We consider only the positive values of the interval's boundaries. Because speed $v = S/T$,

$$\begin{aligned} v_a &= -u_a + v(1 - C_S)/(1 + C_T), \\ v_b &= u_a + v(1 + C_S)/(1 - C_T); \\ \text{hence, } v_b - v_a &= 2u_a + 2v(C_S + C_T)/(1 - C_T^2). \end{aligned} \tag{19}$$

Thus, interval $\Omega(v) = [v_a, v_b]$ grows linearly with speed.

### Integration

By the integration over the variable interval $\Omega(v)$, we find weights $w(v)$ for Equation 15 and thus determine the realistic optimal set:

$$\partial U_S\, g(v) + \partial U_T = 0, \tag{20}$$

where $g(v) = w(v)\, v + [1 - w(v)]v_e$. This set has a form similar to Equations 9 and 14; that is, it also preserves the equilibrium of spatial and temporal uncertainties (Table 1). But now, factor $g(v)$ that modulates $\partial U_S$ is a function of speed.

The explicit form of this optimal set in the distance plot is:

$$S = \sqrt{\frac{\lambda_1 T^2 g(v)}{(\lambda_3 g(v) + \lambda_4) T^2 - \lambda_2}}, \tag{21}$$

which we plot as a red curve in Figure 4.

The realistic optimal set resembles the maximal-sensitivity set of biological vision (Burr & Ross, 1982; Kelly, 1979; Nakayama, 1985; Newsome et al., 1986; van de Grind et al., 1986; van Doorn & Koenderink, 1982a, 1982b) in three ways:

1. It is roughly a branch of a rectangular hyperbola in the first quadrant, whose equation is of the form $ST$ = const, where $S$ and $T$ are spatial and temporal distances, respectively. The hyperbolic shape implies a trading relation between the spatial and temporal parameters.
2. It approaches a vertical asymptote for low values of $T$.
3. It deviates from a horizontal asymptote for high values of $T$, predicted by local optimization.

These similarities do not depend on the choice of parameters of the spatiotemporal uncertainty function (Equation 6) and on the assumption of additivity of spatial and temporal uncertainties (Equation 7). Changing the parameters in Equation 6 results in changing the location of the optimal set in the distance plot. Abandoning the assumption of additivity amounts to having an inseparable space–time term in the spatiotemporal uncertainty function. The weight of that term affects the curvature of the optimal set in the transition between its vertical and nearly horizontal branches in the distance plot.

## Equally suboptimal conditions for speed estimation

We have characterized the visual system's optimal sensitivity to motion. To characterize its performance when it is not at its best, we derive *equivalence sets* of uncertainty for speed estimation. Just as the realistic optimal set does, these equivalence sets balance measurement uncertainty and stimulus uncertainty.

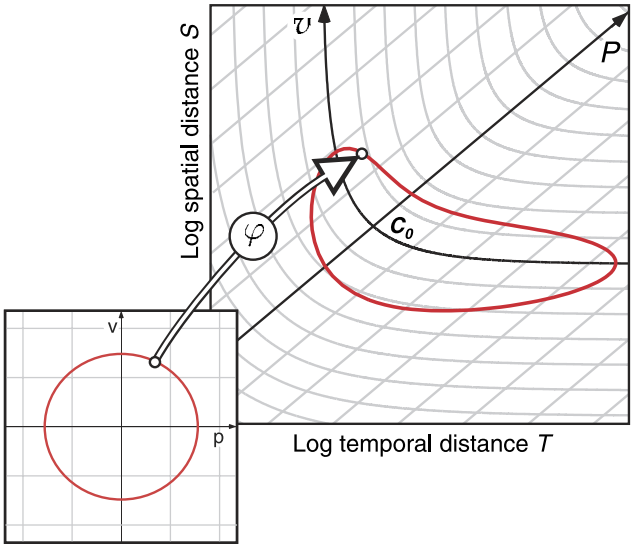


Figure 5. Construction of an equivalence set. Finding equivalence sets in a space–time plot amounts to constructing a circle in the curvilinear coordinate system, shown as a gray net. The coordinate system is the cartesian product of two parametric grids: (1) curvilinear *balance grid,* which consists of curves of equal imbalance of spatial and temporal uncertainties, and (2) linear *speed grid,* which consists of speed lines. The axes are the line of expected speed $v = v_e$ (marked $p$) and the optimal set by weighted optimization (marked $v$). We construct an equivalence set by transforming coordinates $p$ and $v$- which results in a new coordinate system (**p**, **v**), shown in the inset- and then drawing a circle in (**p**, **v**). The shape of the equivalence set is determined by mapping $\varphi$ of the circle: from (**p**, **v**) to the space–time plot (Appendix B).

*Measurement uncertainty*. Because the realistic optimal set represents a perfect balance of measurement uncertainties, any point outside of it represents a degree of imbalance of these uncertainties. In Figure 5, we capture this variation of uncertainties using a curvilinear parametric grid (Struik, 1961). The curves for different degrees of imbalance of measurement uncertainties form a family of nonintersecting grid lines in the distance plot. Each curve has the roughly hyperbolic shape of the realistic optimal set. The heavier curve $v$ is the curve of exact balance. The farther a curve of this grid is from $v$, the greater imbalance of measurement uncertainties it represents.

*Stimulus uncertainty*. Stimulus uncertainty is represented in Figure 5 by the linear part of the grey net formed by the speed lines. The heavier line $p$ is the speed line for the most likely speed $v_e$. The farther a line of this grid is from $p$, the less likely the corresponding speed is in the perceptual ecology.

We find the conditions where the two kinds of uncertainty are equally imbalanced by constructing circles in the curvilinear system of the coordinates (Figure 5). We set a logarithmic scale on the two parametric grids, in agreement with the theoretical and physiological considerations presented by Nover et al. (2005). The details are described in Appendix B. We draw the family of resulting equivalence sets in Figure 6A. Independent of the choice of parameters, the equivalence sets form closed contours in the distance plot. The shapes of contours generally follow the invariantly hyperbolic shape of the maximal sensitivity set. The theoretical equivalence sets are very similar to the observed "bent loaf-of-bread" shape of human isosensitivity contours (Kelly, 1994; our Figure 1A).

To summarize, we have proposed that the visual system should allocate its resources to the conditions of stimulation according to the uncertainty of speed estimation. The similarity between the theoretical optimal and equally suboptimal sets to the empirical maximal sensitivity and isosensitivity sets supports this view and suggests that human vision allocates its resources by the prescription of balance between uncertainties. In agreement with the theory, the visual system seems to allocate (a) more resources to the conditions of stimulation where the different uncertainties are balanced exactly and (b) equal amount of resources to the conditions with the same degree of imbalance of uncertainties.

# Discussion

## Summary

We have developed a normative theory of motion perception. We assumed that the visual system attempts to minimize errors in estimation of motion speed. But

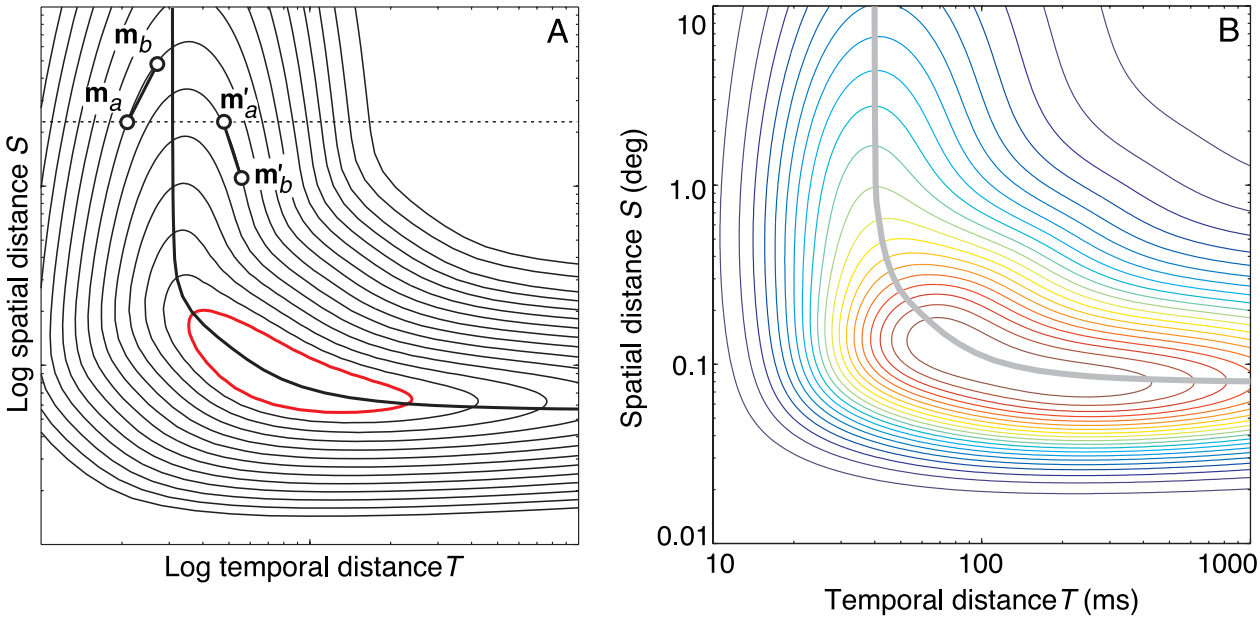


Figure 6. Equivalence contours. (A) Theoretical equivalence sets (derived as explained in Figure 5 and Appendix B). We refer to the graphic representation of the equivalence sets as *equivalence contours*. The equivalence contours reproduce the "bent-loaf-of-bread" shape of human isosensitivity contours (Panel B). The two pairs of connected circles ($\mathbf{m}_a - \mathbf{m}_b$ and $\mathbf{m}'_a - \mathbf{m}'_b$) demonstrate that different regimes of space–time combination are expected under different conditions of stimulation. In the $\mathbf{m}_a - \mathbf{m}_b$ pair, equivalent conditions are obtained by increasing both spatial and temporal distances, from $\mathbf{m}_a$ to $\mathbf{m}_b$ (space–time coupling, supporting Korte's third law of motion). In the $\mathbf{m}'_a - \mathbf{m}'_b$ pair, equivalent conditions are obtained by increasing temporal distance and decreasing spatial distance (space–time tradeoff, contrary to Korte's law). (B) Human isosensitivity contours from Figure 1A.

biological vision with limited resources cannot optimize for every speed. We showed that in such a system the best conditions for the estimation of speed are obtained where uncertainties from different sources are balanced (i.e., are in equilibrium). These conditions form an optimal set for speed estimation; the invariantly hyperbolic shape of this set is similar to the shape of the maximal sensitivity set of human vision. The equally suboptimal conditions for speed estimation are obtained where uncertainties from different sources are imbalanced to the same degree. The conditions of equal imbalance form equivalence sets for speed estimation; the shapes of these sets are similar to the shapes of the isosensitivity sets of human vision.

We used equilibrium analysis of uncertainties to explore the consequences of very basic properties of visual measurement and considerations of biological parsimony for perception of motion. Because our approach rests on very basic considerations, the predictions of the equilibrium theory are qualitative. For example, the theory does not predict the exact shapes of the isosensitivity contours. Instead, the theory allows one to see that the peculiar bent-loaf-of-bread shape of human spatiotemporal sensitivity function plausibly manifests an optimal allocation of limited neural resources. In the following, we illustrate several further consequences of our analysis. We start from the theoretical equivalence sets and show that the invariant shape of these sets helps to reconcile seemingly inconsistent data on apparent motion. Then, we turn to the theoretical optimal set and examine how its properties change in response to changes in the statistics of stimulation and changes in system goals, as well as what testable predictions these changes entail.

## Apparent motion

From the shape of theoretical equivalence sets (Figure 6), it follows that spatial and temporal distances must interact differently under different conditions of stimulation to produce equivalent condition for motion measurement. We illustrate this in Figure 6 using two pairs of conditions shown as two pairs of connected circles: $\mathbf{m}_a - \mathbf{m}_b$ and $\mathbf{m}'_a - \mathbf{m}'_b$.[3] Conditions within a pair correspond to the same degree of imbalance between uncertainties because they belong to the same equivalence set. In the $\mathbf{m}_a - \mathbf{m}_b$ pair, $\mathbf{m}_b$ is longer than $\mathbf{m}_a$ in both space and time (coupling regime), whereas in the $\mathbf{m}'_a - \mathbf{m}'_b$ pair, $\mathbf{m}'_b$ is longer than $\mathbf{m}'_a$ in time but shorter in space (tradeoff regime). Coupling obtains where the equivalence contours have positive slopes, and tradeoff obtains where they have negative slopes.

In fact, both regimes of tradeoff and coupling were observed in studies of apparent motion using suprathreshold stimuli. But the difference between the two was interpreted as a discrepancy between empirical

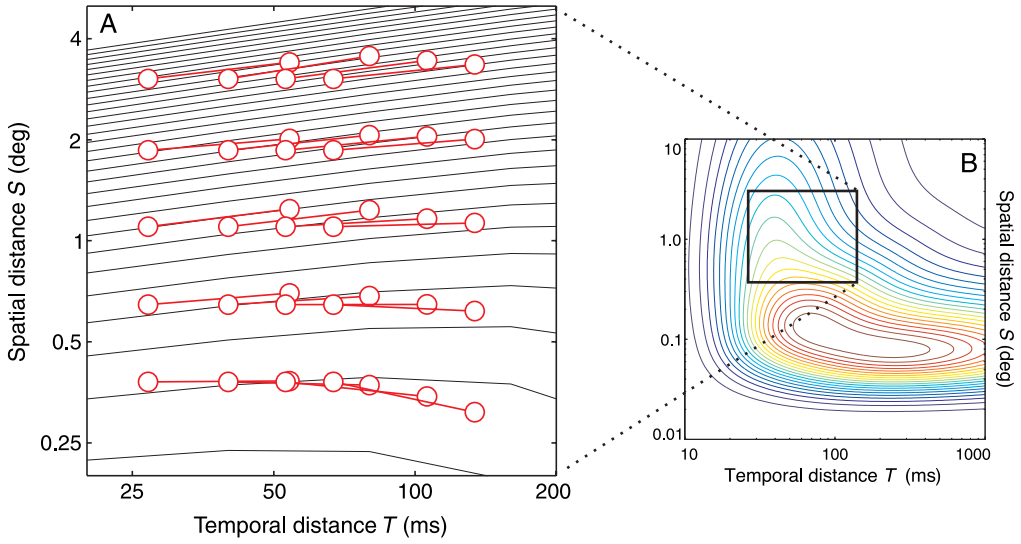


Figure 7. Empirical equivalence sets of apparent motion. (A) The pairs of red connected circles represent the pairs of conditions of apparent motion in perceptual equilibrium: They were experienced equally often (Gepshtein & Kubovy, 2007). The thin black lines on the background are the empirical equivalence contours of apparent motion derived by Gepshtein and Kubovy (2007) from the pairwise equilibria. The slopes of both the empirical equivalence sets and the lines connecting the equilibrium pairs gradually change across the plot, indicating a gradual change from the regime of tradeoff to the regime of coupling, in qualitative agreement with measurements at the threshold (Figure 1A, copied to Panel B) and with the predictions of the equilibrium theory (Figure 6A). (B) Human isosensitivity contours from Figure 1A. The rectangle marks the region of conditions in which Gepshtein and Kubovy could measure the points of equilibrium of apparent motion.

findings rather than a manifestation of an optimal behavior of the visual system.

- Space–time coupling corresponds to Korte's (1915) venerable third law of motion (Koffka, 1935/1963; Lakatos & Shepard, 1997; Neuhaus, 1930). Korte presented his observers with two flashes separated by variable spatial and temporal distances. He first found the distances that gave rise to a compelling experience of motion ("good motion"). But he found that he could not change just the spatial distance or just the temporal distance without reducing the strength of motion. To restore the experience of good motion, he had to increase or decrease both.

- Evidence of space–time tradeoff was found by Burt and Sperling (1981), who used ambiguous apparent motion displays, in which motion could be seen in one of several directions ("paths") in the same stimulus. When the authors varied the spatial and temporal distances in their display, they found that they had to decrease one distance and increase the other to maintain the equilibrium between the perception of different paths.

Gepshtein and Kubovy (2007) reproduced both results—coupling and tradeoff—using the same suprathreshold stimulus. They showed that the qualitatively different regimes of apparent motion are special cases of a general pattern: a smooth transition between the tradeoff and coupling as a function of speed (Figure 7); tradeoff occurs at low speeds and coupling occurs at high speeds, in agreement with predictions of the equilibrium theory. According to the theory, the different regimes are observed under different conditions of stimulation because the frugal visual system balances the measurement and stimulus uncertainties associated with speed estimation.

## Motion adaptation

The equilibrium theory predicts changes in motion perception in response to changes in the statistics of speed in the stimulation, because the properties of the optimal set for speed estimation depend on the statistics Equation 20. When vision is excessively stimulated with a single speed or a narrow band of speeds, as it is often the case in motion adaptation studies, the equilibrium theory predicts that sensitivity should change for the adapting speed(s) and also for speeds very different from the adapting one. Recall that the optimal set for speed estimation predicted by the theory is represented by a nearly hyperbolic curve in the distance plot (Figure 4). The position of this curve in the distance plot depends on the most likely speed (expected speed $v_e$) in the stimulation (Equation 14). If changes in statistics of stimulation change the expected speed, then the optimal set for speed estimation should change its position in the distance plot. For example, in Figure 8 we plot the pre-adaptation optimal set computed

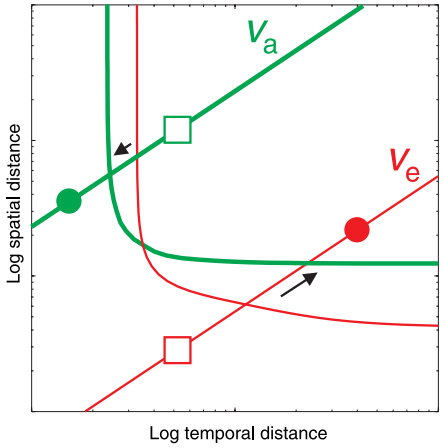


Figure 8. Predicted consequences of motion adaptation. The red and green lines represent expected speeds: $v_e$ of the natural (pre-adaptation) environment (Equation 13) and $v_a$ of the new (adapting) environment, respectively. As a result of adaptation, the optimal set for speed estimation changes from the one represented by the red curve (from Figure 4) to the one represented by the green curve. The arrows indicate the direction of improvement of sensitivity along the two speed lines. The circles and the squares on the speed lines mark examples of conditions where sensitivity improves or deteriorates, respectively. Visual sensitivity is expected to change across a wide range of speeds. Thus, sensitivity is expected to improve for some conditions on the red speed line, away from the adapting speed, but here the improvement is expected at larger temporal distances than on the adapting speed line.

for an expected speed $v_e$ and the post-adaptation optimal set computed for an adapting speed $v_a$. Because of the change in the position of the optimal set, visual sensitivity is expected to change across a wide range of speeds, more so at high speeds, where speed-specific optimization operates across wider range of speeds than at low speeds (Equation 19).

Against the intuition, and contrary to what is commonly expected in the motion adaptation literature (reviewed in Clifford & Wenderoth, 1999, and Krekelberg, Van Wezel, & Albright, 2006), the equilibrium theory predicts that visual performance will both improve and deteriorate at the adapting speed, depending on *where* it is measured on the speed line $v_a$. The point of optimal sensitivity will move along the speed line, toward the prediction by local optimization: Visual sensitivity is expected to improve on the part of the speed line that is close to the local optimum but deteriorate in the vicinity of its preadaptation optimum, for most magnitudes of adapting speeds. For example, in Figure 8, an improvement of sensitivity is expected at the condition marked by the green circle, and a deterioration is expected at the condition marked by the green square. An exception from this pattern of changes is the region of global optimum (the intersection of two curves in Figure 8), where little change is expected independent of the magnitude of adapting speed.

These predictions can be tested by measuring spatiotemporal sensitivity across a wide range of parameters to obtain a comprehensive characteristic of sensitivity (Figure 1A) before and after adaptation. Such measurements will show whether changes in sensitivity occur globally (i.e., at speeds different from the adapting speed) and whether the global changes follow the pattern predicted by the equilibrium theory. To our knowledge, no such comprehensive studies have been undertaken. But data from previous motion adaptation studies suggest that changes in spatiotemporal sensitivity following adaptation are global. For example, Krekelberg et al. (2006) studied changes in the responsiveness of speed-tuned cells in the cortical area MT of behaving monkeys. It was found that the susceptibility of a cell to short-term adaptation depended on whether the adapting speed was at the cell's optimum speed. When the adapting speed was at the cell's optimum, the effect of adaptation was often smaller than when the adapting speed was different from the cell's optimum. Krekelberg et al. also found that, when the adapting speed was at the cell's optimum, discrimination performance sometimes improved and sometimes deteriorated, as we anticipated in Figure 8. To test whether the pattern of changes in discrimination performance follows the pattern predicted in Figure 8, the adapting stimuli should be narrowly localized in the space of parameters. (The adapting stimuli of Krekelberg et al. were broadband.)

## System goals

As we mentioned above, parameters $\lambda_i$ of spatiotemporal uncertainty (Equation 7) can be thought of as weights expressing how much each of the uncertainty terms affects the precision of speed estimation. The visual system can modulate the effect of uncertainties on system behavior by redistributing its resources, which, in the present formulation of the equilibrium theory, is equivalent to changing the weights.

The terms of Equation 7 belong to two groups: One mostly affects the ability to localize signals ($\lambda_3 S$ and $\lambda_4 T$), and the other mostly affects the ability to identify signals ($\lambda_1/S$ and $\lambda_2/T$). Suppose that one of the two tasks—signal localization versus signal identification—becomes more important to the visual system than the other. For example, consider an observer whose task is to discriminate the location of moving targets in the laboratory on one day and categorize moving targets by their spatial features on another day. By the equilibrium theory, the visual system will be able to improve its performance in one task, at the expense of performance in the other task, by redistributing its resources. Then, predictable global changes in visual sensitivity are expected.

Thus, performance in the identification task will improve when the visual system allocates less resources for measuring stimulus location and more resources for measuring its frequency content. This change corresponds to *decreasing* weights $\lambda_3$ and $\lambda_4$ relative to weights $\lambda_1$ and $\lambda_2$ in Equation 7. The changes in weights imply a change in the position of the optimal set in the distance plot (Figure 4). For example, the emphasis on signal identification (decreasing $\lambda_3$ and $\lambda_4$ relative to $\lambda_1$ and $\lambda_2$) will cause the optimal set defined by Equation 20 to move away from the origin of the distance plot. But when the emphasis is on signal identification, and $\lambda_3$ and $\lambda_4$ *increase* relative to $\lambda_1$ and $\lambda_2$, the optimal set will move toward the origin. Changes in the position of the theoretical optimal set imply large-scale changes in visual sensitivity across the distance plot because the theoretical equivalence sets (Figure 5 and Appendix B) will change their positions as well.

Notice that the equilibrium theory predicts distinct changes in the pattern of visual sensitivity in response to changes in statistics of speed (as in motion adaptation) and changes in system goals. When statistics of speed change, the optimal set is expected to move along the direction of maximal speed variation, that is, along the negative diagonal of the distance plot: from the right bottom corner to the left top corner (Figure 8). When system goals change, the optimal set is expected to move along the other diagonal: the imaginary line connecting the left bottom and the right top corners of the distance plot (not shown in Figure 8).

Notice also that by the equilibrium theory, one cannot use different tasks and expect to obtain quantitatively consistent evidence about the same sensitivity characteristic of the visual system. Different tasks will induce changes in the quantitative detail of the spatiotemporal sensitivity function. The amount and time course of such changes are interesting topics for future research.

## Comparison to other normative models of vision

Contemporary decision-theoretic models of perception and behavior have features that appear similar to our approach. We will now review the similarities and differences between the approaches.

### Weak fusion

The "weak fusion" theory of cue combination (Clark & Yuille, 1990; Landy, Maloney, Johnsten, & Young, 1995; Maloney & Landy, 1989; Yuille & Bülthoff, 1996) predicts that the nervous system combines different estimates of a parameter of interest using a linear weighting rule, similar to the linear weighting in our Equation 15. The linear weighting of the weak-fusion framework is derived from the maximum likelihood principle (Yuille & Bülthoff, 1996) and is used to implement the assumption that a goal of cue combination is to maximize the precision (minimize uncertainty) of the combined estimate. In our theory, linear weighting appears for other reasons; it allows us to find a principled compromise between two extreme cases of optimal resource allocation: one in a system with minimal resources and the other in a system with infinite resources. The optimal sets in both cases are characterized by the equilibrium of spatial and temporal uncertainties. We use the linear weighting scheme because it is the simplest one we found that preserves the equilibrium of spatial and temporal uncertainties in the compromise optimal set (Table 1).

### Bayesian inference

The weights that appear in Equation 15 depend on the distribution of speeds in the perceptual ecology. Our use of the statistics of speed resembles the use of expected distributions of parameter values in inferential theories of perception (Knill & Richards, 1996) derived from the Bayesian decision theory (Berger, 1985; Maloney, 2002). In such theories, the probabilities of sensory estimates ("likelihood functions") and the probabilities of corresponding parameter values in the natural stimulation ("prior distributions") are combined by point-by-point multiplication, following Bayes' rule, making the prevalent values in the stimulation more likely to be perceived than the less common values. Thus, the factors that determine predictions are generally defined in the space of estimated parameters. In contrast, our predictions are derived in a space whose dimensions are different from the space of estimated parameters: Our predictions depend on measurement uncertainty in addition to stimulus uncertainty, and the effects of the two kinds of uncertainty are separable. Thus, in Figure 5, the "speed grid" depends on stimulus uncertainty and the "balance grid" depends on measurement uncertainty. The separation of these two grids means that adaptive changes in system behavior are not reducible to changes along the dimension of estimated parameters alone.

### Utility theories

Decision-theoretic models of perception and behavior use the notion of utility (Bernoulli, 1954; Kahneman & Tversky, 2000; Luce & Raiffa, 1957; Stigler, 1950) to account for the fact that different errors in sensory estimation or movement execution affect behavior differently. The "costs" of different errors are usually represented by a utility function (Geisler & Kersten, 2002; Maloney, Trommershäuser, & Landy, 2007; Trommershäuser, Maloney, & Landy, 2003). The use of these functions

resembles the use of prior probability functions in Bayesian inferential models: The probabilities of possible sensory estimates or movements are weighted by the corresponding utilities. In our theory, concepts similar to utility play a role at two places:

1. In our measurement uncertainty function (Equation 7), coefficients $\lambda_i$ represent the weights of the components of uncertainty function. One can think of the coefficients as the costs of measurement errors, just as the values of a utility function can be thought of as costs of errors in sensory estimation. As we argued in the System goals section, the visual system may change the weights of the components of measurement uncertainty to optimize itself for different tasks. This will affect our predictions differently from how changes in utilities affect predictions of decision-theoretic models:

   a. Changing weights in Equation 7 will affect the measurement uncertainty and not the stimulus uncertainty. The predicted changes will occur along a dimension separate from the dimension of stimulus uncertainty. (As we explained in the Bayesian inference section, the effects of measurement uncertainty and stimulus uncertainty in our theory are separable.)

   b. Changing weights in Equation 7 will affect the visual sensitivity of the visual system, whereas changing utility functions in the decision-theoretic models will affect visually guided behavior at a later stage. It is plausible that weighting of errors in the nervous system happens both early (as implied by our theory) and late (as implied by the decision-theoretic models). But the two kinds of weighting are expected to occur in different parameter spaces: The effects of early weighting will occur along the dimension of measurement uncertainty, whereas the effects of late weighting will occur in the space of estimated parameters. These differences will allow one to experimentally separate the effects of early and late weighting of errors.

2. In our discussion of the limitations of the local-optimization approach, we noted that both the speed prevalence and the *importance* of speeds for the organism should affect motion perception (claim R2 in the Local optimization: A spendthrift extreme section). In the present work, we used only the estimates of speed prevalence because, to our knowledge, no estimates of a speed "importance function" exist. The effect of speed importance on equilibrium theory predictions can be studied by modifying the expression for $p(v)$ in Equations 13 and 16 and in the derivation of the equivalence sets (Appendix B).

## Conclusions

The equilibrium theory of speed estimation offers a principled explanation of the distribution of human visual sensitivity and explains why qualitatively different regimes of apparent motion are observed at different speeds. On this view, the shapes of the empirical maximal sensitivity set and the isosensitivity sets (measured at the threshold of visibility) and the different regimes of apparent motion (measured above the threshold) are manifestations of the optimal balance of uncertainties in a visual system that seeks to maximize the precision of its measurements with limited resources. Thus, the equilibrium theory offers a normative framework for understanding motion perception at the threshold of visibility and above the threshold, and predicts how the visual system should adjust its sensitivity in response to changes in the statistics of stimulation and changes in system goals.

# Appendix A

## Construction of Figure 1

To display the isosensitivity contours in the space–time distance plot, we use Equations 5–8 of Kelly (1979), with which he fit the spatiotemporal thresholds for the detection of drifting sinusoidal gratings. Following Nakayama (1985, p. 637), we computed the spatial and temporal *distances* between successive discrete stimuli that correspond to Kelly's spatial and temporal frequencies. Nakayama assumed that motion is detected by pairs of spatial-frequency filters ("quadrature pairs"; Adelson & Bergen, 1985; Gabor, 1946), in agreement with physiological (Marcelja, 1980; Pollen & Ronner, 1981) and computational (Sakitt & Barlow, 1982) evidence. The two parts of such a detector are tuned to the same spatial frequency $f_s$, but their spatial phases differ by $\pi/2$, a quarter of the spatial period of the optimal stimulus (see also van Santen & Sperling, 1984). When such a detector is stimulated by a luminance grating with spatial frequency $f_s$, a spatial shift by $S = 1 / (4f_s)$ will activate the detector optimally. Similarly, the optimal temporal interval $T$ of a detector is equal to the quarter period of its optimal temporal modulation: $T = 1 / (4f_T)$. By this argument, there exists a simple correspondence between the frequency tuning of motion detectors and the spatial and temporal distance between successive stimuli that activate the detectors optimally. Using the above expressions for $S$ and $T$, we mapped Kelly's spatiotemporal threshold surface (his Figure 15) to the logarithmic space–time distance plot in Figure 1A. In the distance plot, the maximal sensitivity set is convex toward the origin. When plotted in the coordinates of spatial and temporal

frequencies, as in Kelly's Figure 15, the maximal sensitivity set is concave toward the origin.

We use the estimates of spatiotemporal sensitivity obtained by Kelly (1979) from measurements using an image stabilization technique that afforded precise control over retinal motion. Those data had the same form as the data obtained with no image stabilization (Kelly, 1969, 1972; Kulikowski, 1971; Robson, 1966; van Nes, Koenderink, Nas, & Bouman, 1967), as Kelly and Burbeck (1984) observed. This allows us to relate the estimates of Kelly (1979) to results from studies that did not use image stabilization.

# Appendix B

## Derivation of the equivalence sets (Figures 5 and 6)

We derived equivalence sets for speed estimation using a curvilinear system of coordinates (the grey net in Figure 5) embedded in the distance plot. The system consists of two parametric grids: a curvilinear grid $p$ that parameterizes measurement uncertainty and a linear grid $v$ that parameterizes stimulus uncertainty. An equivalence set consists of the loci in coordinate system $(p, v)$ that are equidistant from the origin of $(p, v)$. That is, an equivalence set is a circle in the curvilinear coordinates. We illustrated that in the inset of Figure 5, by transforming $(p, v)$ into a rectilinear coordinate system $(\mathbf{p}, \mathbf{v})$.

We constructed circles in $(p, v)$ in two steps:

1. Parameterization. We parameterized $(p, v)$ as follows:

   $v$: For $v$, we set the origin to $v_e$ because by Equation 14, the *a priori* uncertainty about speed is minimal when $v = v_e$. We used a logarithmic scale, as suggested by Nover et al. (2005). The logarithmic scale leads to Weber's law for speed discrimination thresholds (McKee & Watamaniuk, 1994; Nover et al., 2005). Also, the logarithmic scale allowed us to parameterize speeds from the interval $[0, \infty]$ to interval $[-\infty, \infty]$. Thus, we defined the scaling function:

   $$\mathbf{v}(v) = \ln(v/v_e)^{1/k_v}, \tag{B1}$$

   where $k_v$ is a constant.

   $p$: If the optimal set represents perfect balance, an equivalence set shares a *degree of imbalance*. Hence, to construct grid $p$, we generalized Equation 20:

   $$\partial U_S\, g(v) + \partial U_T = p, \tag{B2}$$

   where $p$ is the degree of imbalance. We set the origin of grid of $p$ to zero because at $p = 0$, the solution of Equation B2 is the optimal set. The optimal solutions of Equation B2 are feasible only below a boundary value of $p$ that depends on speed: $p_{max}(v) = \lambda_3\, g(v) + \lambda_4$. Because we had no a natural scale for $p$ (in contrast to the scale for $v$), we assumed a logarithmic scaling function, as we did for $v$:

   $$\mathbf{p}(p) = \ln(1 - p/B)^{-1/k_p}, \tag{B3}$$

   where $k_p$ is a constant, $B = 1$ for $p \in [-\infty, 0]$, and $B = p_{max}(v)$ for $p > 0$. The choice of the logarithmic function did not affect the invariant properties of the equivalence sets (Figure 6)—their closedness and hyperbolic shape—and their similarity to human isosensitivity contours (Figure 1A).

This parametrization results in a rectangular coordinate system $(\mathbf{p}, \mathbf{v})$ shown in the inset of Figure 5.

2. Mapping. An equivalence set is a circle in $(\mathbf{p}, \mathbf{v})$ centered on $(\mathbf{p}(0) = 0, \mathbf{v}(v_e) = 0)$. We illustrate this in Figure 5 for one point, as mapping $\varphi$ of a point from $(\mathbf{p}, \mathbf{v})$ (in the inset) to a point in the space–time plot (in the main panel). From Equations 7 and 20, it follows that the coordinates of a point $(p = p', v = v')$ in $(S, T)$ are

   $$\begin{cases} S = \sqrt{\dfrac{\lambda_1 T^2 g(v')}{(p_{max}(v') - p')T^2 - \lambda_2}} \\ T = S/v'. \end{cases} \tag{B4}$$

   (Note that the first equation is a generalization of Equation 10.) For a point $(p', v')$, we used Equation B4 to find the coordinates of $(p'', v'')$ in $(S, T)$ for all points whose distance from $(\mathbf{p} = 0, \mathbf{v} = 0)$ was constant, such that $\left\| \mathbf{p}(p'), \mathbf{v}(v') \right\| = \left\| \mathbf{p}(p''), \mathbf{v}(v'') \right\|$. Thus, the solution of Equation B4 constitutes the mapping $\varphi$: $(\mathbf{p}, \mathbf{v}) \rightarrow (S, T)$.

# Acknowledgments

We thank Wilson S. Geisler and Michael S. Landy for discussions; Thomas D. Albright, Cees van Leeuwen, and Sergey Savel'ev for comments about an earlier version of this article; and Peter Jurica for help in simulations of the normative theory. Parts of this work were presented at the

5th Vision Sciences Society meeting (May 2006, Sarasota, FL, USA) and the 4th Asian Conference on Vision (July 2006, Matsue, Shimane, Japan). Michael Kubovy was supported by NEI Grant R01 EY12926 and NIDCD Grant R01 DC005636.

Commercial relationships: none.
Corresponding author: Sergei Gepshtein.
Email: sergei@brain.riken.jp.
Address: Perceptual Dynamics Laboratory, Brain Science Institute, RIKEN, Wakoshi, Saitama, Japan.

## Footnotes

[1]Clarification of terms: In measuring parameters of stimulation (such as speed), a measurement *error* is the difference between an estimate and the true value of the parameter of interest. Measurement errors characterize *uncertainty* about the parameter value. Multiple measurements generally yield estimates that differ from one another and from the true value of the parameter. The more dispersed the distribution of these errors, the greater the uncertainty and the lower the *precision* of estimation. In sensory estimation, different parameter values correspond to different conditions of stimulation. We describe this by saying that the uncertainty and precision of estimation vary across the conditions of stimulation.

[2]Analysis of the spatial-frequency content of stimuli is important for perception of motion for two reasons. First, the visual system needs to *identify* moving objects. Performance in the identification task depends on an analysis of spatial-frequency content of stimulation (De Valois & De Valois, 1990). Second, the visual system needs to solve the *motion matching problem* (Hildreth, 1984; Ullman, 1979), and the ability to solve it depends on the ability to measure the spatial-frequency content of stimulation. Motion matching problem is particularly difficult when (spatially) small receptive fields are used, because small regions of retinal image are not unique: They are similar to many other regions of the image, and the matching process produces many spurious matches. A similar argument was proposed by Banks, Gepshtein, and Landy (2004) with respect to the binocular matching problem.

[3]Each point in the distance plot (Figure 1) can represent a narrow-band visual stimulus. In broad-band stimuli, such as the apparent motion displays used by Burt and Sperling (1981) or Gepshtein and Kubovy (2007), a point in the distance plot corresponds to the fundamental spatial and temporal frequencies contained in the stimulus or simply to the spatial and temporal distances between successive dots.

## References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A, Optics and image science, 2,* 284–299. [PubMed]

Banks, M. S., Gepshtein, S., & Landy, M. S., (2004). Why is spatial stereoresolution so low? *Journal of Neuroscience, 24,* 2077–2089. [PubMed] [Article]

Barron, A. R. (1993). Universal approximation bounds for superpositions of a sigmoidal function. *IEEE Transaction on Information Theory, 39,* 930–945.

Berger, J. O. (1985). *Statistical decision theory and bayesian analysis* (2nd ed.). New York: Springer.

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica, 22,* 23–36. (Originally published in 1738. Translated by Louise Sommer).

Burr, D. C., & Ross, J. (1982). Contrast sensitivity at high velocities. *Vision Research, 22,* 479–484. [PubMed]

Burt, P., & Sperling, G. (1981). Time, distance, and feature trade-offs in visual apparent motion. *Psychological Review, 88,* 171–195. [PubMed]

Clark, J. J., & Yuille, A. L. (1990). *Data fusion for sensory information processing systems*. Norwell, MA: Kluwer Academic Publishers.

Clifford, C. W., & Wenderoth, P. (1999). Adaptation to temporal modulation can enhance differential speed sensitivity. *Vision Research, 39,* 4324–4332. [PubMed]

Daugman, J. G. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America A, Optics and image science, 2,* 1160–1169. [PubMed]

De Valois, R. L., & De Valois, K. K. (Eds.). (1990). *Spatial vision*. New York: Oxford University Press.

Dong, D., & Atick, J. (1995). Statistics of natural time-varying images. *Network: Computation in Neural Systems, 6,* 345–358.

Gabor, D. (1946). Theory of communication. *Institution of Electrical Engineers, 93*(Pt. III), 429–457.

Geisler, W. S., & Kersten, D. (2002). Illusions, perception and Bayes. *Nature Neuroscience, 5,* 508–510. [PubMed] [Article]

Gepshtein, S., & Kubovy, M. (2007). The lawful perception of apparent motion. *Journal of Vision, 7*(8):9, 1–15, http://journalofvision.org/7/8/9/, doi:10.1167/7.8.9. [PubMed] [Article]

Gepshtein, S., & Tyukin, I. (2006). Why do moving things look as they do? *Vision. The Journal of the Vision Society of Japan, 18,* 64.

Gepshtein, S., Tyukin, I., Kubovy, M., & van Leeuwen, C. (2006). A Pareto-optimality theory of motion perception [Abstract]. *Journal of Vision, 6*(6):577, 577a, http://journalofvision.org/6/6/577/, doi:10.1167/6.6.577.

Hildreth, E. C. (1984). *Measurement of visual motion*. Cambridge, MA: MIT Press.

Jones, J. P., & Palmer, L. A. (1987). An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology, 58,* 1233–1258. [PubMed]

Kahneman, D., & Tversky, A. (Eds.). (2000). *Choices, values, and frames*. New York: Cambridge University Press.

Kelly, D. H. (1969). Flickering patterns and lateral inhibition. *Journal of the Optical Society of America, 59,* 1361–1370.

Kelly, D. H. (1972). Adaptation effects on spatio-temporal sine-wave thresholds. *Vision Research, 12,* 89–101. [PubMed]

Kelly, D. H. (1979). Motion and vision. II. Stabilized spatio-temporal threshold surface. *Journal of the Optical Society of America, 69,* 1340–1349. [PubMed]

Kelly, D. H. (1994). Eye movements and contrast sensitivity. In D. H. Kelly (Ed.), *Visual science and engineering (Models and applications)* (pp. 93–114). New York: Marcel Dekker, Inc.

Kelly, D. H., & Burbeck, C. A. (1984). Critical problems in spatial vision. *Critical Reviews in Biomedical Engineering, 10,* 125–177. [PubMed]

Knill, D. C., & Richards, W. (Eds.). (1996). *Perception as Bayesian inference*. Cambridge, UK: Cambridge University Press.

Koffka, K. (1935/1963). *Principles of Gestalt psychology*. New York: A Harbinger Book, Harcourt, Brace & World, Inc.

Korte, A. (1915). Kinematoskopische Untersuchungen [Kinematoscopic investigations]. *Zeitschrift für Psychologie, 72,* 194–296.

Krekelberg, B., van Wezel, R. J., & Albright, T. D. (2006). Adaptation in macaque MT reduces perceived speed and improves speed discrimination. *Journal of Neurophysiology, 95,* 255–270. [PubMed] [Article]

Kulikowski, J. J. (1971). Some stimulus parameters affecting spatial and temporal resolution of human vision. *Vision Research, 11,* 83–93. [PubMed]

Kulikowski, J. J., Marcelja, S., & Bishop, P. O. (1982). Theory of spatial position and spatial frequency relations in the receptive fields of simple cells in the visual cortex. *Biological Cybernetics, 43,* 187–198. [PubMed]

Lakatos, S., & Shepard, R. N. (1997). Constraints common to apparent motion in visual, tactile, and auditory space. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 1050–1060. [PubMed]

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research, 35,* 389–412. [PubMed]

Luce, R. D., & Raiffa, H. (1957). *Games and decisions*. New York: Wiley.

MacKay, D. M. (1981). Strife over visual cortical function. *Nature, 289,* 117–118. [PubMed]

Maloney, L. T. (2002). Statistical decision theory and biological vision. In D. Heyer & R. Mausfeld (Eds.), *Perception and the physical world* (pp. 145–189). New York: Wiley.

Maloney, L. T., & Landy, M. S. (1989). A Statistical framework for robust fusion of depth information. In W. A. Pearlman (Ed.), *Visual communications and image processing IV* (vol. 1199, pp. 1154–1163). Proceedings of SPIE.

Maloney, L. T., Trommershäuser, J., & Landy, M. S. (2007). Questions without words: A comparison between decision making under risk and movement planning under risk. In W. Gray (Ed.), *Integrated models of cognitive systems* (pp. 297–313). New York: Oxford University Press.

Marcelja, S. (1980). Mathematical description of the responses of simple cortical cells. *Journal of the Optical Society of America, 70,* 1297–1300. [PubMed]

McKee, S. P., & Watamaniuk, S. N. J. (1994). The psychophysics of motion perception. In A. T. Smith & R. J. Snowden (Eds.), *Visual detection of motion* (pp. 85–114). London, UK: Academic Press.

Nakayama, K. (1985). Biological image motion processing: A review. *Vision Research, 25,* 625–660. [PubMed]

Neuhaus, W. (1930). Experimentelle Untersuchung der Scheinbewegung [Experimental investigation of apparent movement]. *Archiv Für die Gesamte Psychologie, 75,* 315–458.

Newsome, W. T., Mikami, A., & Wurtz, R. H. (1986). Motion selectivity in macaque visual cortex. III. Psychophysics and physiology of apparent motion. *Journal of Neurophysiology, 55,* 1340–1351. [PubMed]

Nover, H., Anderson, C. H., & DeAngelis, G. C. (2005). A logarithmic, scale-invariant representation of speed in macaque middle temporal area accounts for speed

discrimination performance. *Journal of Neuroscience, 25,* 10049–10060. [PubMed] [Article]

Pareto, V. (1906). *Manual of political economy*. New York: Augustus M. Kelley. (1971 translation of 1927 edition.)

Park, J., & Sandberg, I. W. (1993). Approximation and radial-basis-function networks. *Neural Computation, 5,* 305–316.

Perrone, J. A., & Thiele, A. (2001). Speed skills: Measuring the visual speed analyzing properties of primate MT neurons. *Nature Neuroscience, 4,* 526–532. [PubMed] [Article]

Pollen, D. A., & Ronner, S. F. (1981). Phase relationships between adjacent simple cells in the visual cortex. *Science, 212,* 1409–1411. [PubMed]

Priebe, N. J., Cassanello, C. R., & Lisberger, S. G. (2003). The neural representation of speed in Macaque area MT/V5. *Journal of Neuroscience, 23,* 5650–5661. [PubMed] [Article]

Reichardt, W. (1961). Autocorrelation, a principle of evaluation of sensory information by the central nervous system. In W. A. Rosenbluth (Ed.), *Sensory communication* (pp. 303–318). Cambridge, MA: MIT Press.

Reichardt, W. (1969). Movement perception in insects. In W. Reichardt (Ed.), *Processing of optical data by organisms and by machines* (pp. 465–493). London, UK: Academic Press.

Resnikoff, H. L. (1989). *The illusion of reality*. New York: Springer-Verlag New York, Inc.

Robson, J. G. (1966). Spatial and temporal contrast-sensitivity functions of the visual system. *Journal of the Optical Society of America, 56,* 1141–1142.

Sakitt, B., & Barlow, H. B. (1982). A model for the economical encoding of the visual image in cerebral cortex. *Biological Cybernetics, 43,* 97–108. [PubMed]

Simoncelli, E. P., & Heeger, D. J. (1998). A model of neuronal responses in visual area MT. *Vision Research, 38,* 743–761. [PubMed]

Stigler, G. J. (1950). The development of utility theory. *Journal of Political Economy, 58,* 307–327.

Struik, D. J. (1961). *Lectures on classical differential geometry*. New York: Dover.

Trommershäuser, J., Maloney, L. T., & Landy, M. S. (2003). Statistical decision theory and the selection of rapid, goal-directed movements. *Journal of the Optical Society of America A, Optics, image science, and vision, 20,* 1419–1433. [PubMed]

Ullman, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.

van de Grind, W. A., Koenderink, J. J., & van Doorn, A. J. (1986). The distribution of human motion detector properties in the monocular visual field. *Vision Research, 26,* 797–810. [PubMed]

van Doorn, A. J., & Koenderink, J. J. (1982a). Spatial properties of the visual detectability of moving spatial white noise. *Experimental Brain Research, 45,* 189–195. [PubMed]

van Doorn, A. J., & Koenderink, J. J. (1982b). Temporal properties of the visual detectability of moving spatial white noise. *Experimental Brain Research, 45,* 179–188. [PubMed]

van Nes, F. L., Koenderink, J. J., Nas, H., & Bouman, M. A. (1967). Spatiotemporal modulation transfer in the human eye. *Journal of the Optical Society of America, 57,* 1082–1088. [PubMed]

van Santen, J. P., & Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America A, 1,* 451–473. [PubMed]

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience, 5,* 598–604. [PubMed]

Yuille, A. L., & Bülthoff, H. H. (1996). Bayesian decision theory and psychophysics. In D. C. Knill & W. Richards (Eds.), *Perception as Bayesian inference* (pp. 123–161). Cambridge, UK: Cambridge University Press.