Neuronal Population Decoding Explains the Change in Signal Detection Sensitivity Caused by Task-Irrelevant Perceptual Bias

Satohiro Tajima
tajima@mns.k.u-tokyo.ac.jp
Department of Complexity Science and Engineering, University of Tokyo, Kashiwa City, Chiba 277-8561 Japan, and Nagano Station, Japan Broadcasting Corporation, Nagano City, Nagano 380-8502, Japan

Hiromasa Takemura
hiromasa@fechner.c.u-tokyo.ac.jp
Department of Life Sciences, University of Tokyo, Komaba, Meguro-ku, Tokyo 153-8902, Japan, and Japan Society for the Promotion of Science, Tokyo, Japan

Ikuya Murakami
ikuya@fechner.c.u-tokyo.ac.jp
Department of Life Sciences, University of Tokyo, Komaba, Meguro-ku, Tokyo 153-8902, Japan

Masato Okada
okada@k.u-tokyo.ac.jp
Department of Complexity Science and Engineering, University of Tokyo, Kashiwa City, Chiba 277-8561, Japan, and Brain Science Institute, RIKEN, Wako City, Saitama 351-0198, Japan

Spatiotemporal context in sensory stimulus has profound effects on neural responses and perception, and it sometimes affects task difficulty. Recently reported experimental data suggest that human detection sensitivity to motion in a target stimulus can be enhanced by adding a slow surrounding motion in an orthogonal direction, even though the illusory motion component caused by the surround is not relevant to the task. It is not computationally clear how the task-irrelevant component of motion modulates the subject’s sensitivity to motion detection. In this study, we investigated the effects of encoding biases on detection performance by modeling the stochastic neural population activities. We modeled two types of modulation on the population activity profiles caused by a contextual stimulus: one type is identical to the activity evoked by a physical change in the stimulus, and the other is expressed more simply in terms of response gain modulation. For both encoding schemes, the motion detection performance of the ideal observer is enhanced by

Neural Computation 22, 2586–2614 (2010) © 2010 Massachusetts Institute of Technology
a task-irrelevant, additive motion component, replicating the properties observed for real subjects. The success of these models suggests that human detection sensitivity can be characterized by a noisy neural encoding that limits the resolution of information transmission in the cortical visual processing pathway. On the other hand, analyses of the neuronal contributions to the task predict that the effective cell populations differ between the two encoding schemes, posing a question concerning the decoding schemes that the nervous system used during illusory states.

1 Introduction

Human perception of sensory stimulus is known to be greatly affected by its context, including such factors as motion in the stimulus surround (Duncker, 1929; Marshak & Sekuler, 1979; Gogel, 1979; Gogel & Griffin, 1982; Wallach & Becklen, 1983; Regan, 1986; Reinhardt-Rutland, 1988; Murakami & Shimojo, 1993; Kim & Wilson, 1997), and there are intensive studies on the nature of perceptual biases. The questions of whether and how perceptual biases affect the resolution of human sensory processing, however, have not been thoroughly investigated. Recently, our group (Takemura & Murakami, 2010) reported psychophysical evidence that the motion detection sensitivity of human subjects can be enhanced by an illusory motion, which is induced by presenting a moving stimulus in the contextual surround (known as “induced motion” or “motion contrast”; Duncker, 1929; Marshak & Sekuler, 1979; Reinhardt-Rutland, 1988; Murakami & Shimojo, 1993). This phenomenon indicates that the resolution of human perception can be indirectly modulated, and sometimes even heightened, under the existence of an illusion. How the addition of illusory motion is related to the subject’s sensitivity to motion detection, however, has not yet been computationally clarified. While neural representations in the cortical area, such as MT/V5 in primates, have been related to both motion processing abilities (e.g., discrimination, detection) and perceptual biases in motion illusions (e.g., motion contrast, motion aftereffects), the neural mechanisms concerning motion processing abilities under illusory conditions have rarely been studied. The work presented in this letter is motivated by this existing gap. We aim to link neural processing to task performance through models of population encoding and decoding and to provide clues to revealing the neurophysiological basis for a motion illusion modulating a subject’s sensitivity. We show that neuronal population (de)coding provides an explanation for the enhancement of motion detection sensitivity caused by the additional illusory motion. The success of the current approach implicates a strong dependence of human performance on the resolution of encoding by a neuronal population at a particular sensory processing stage.
(a) Visual stimulus for motion induction (Takemura & Murakami, 2010). (b) Schematic illustration of motion induction by the surrounding stimulus, with the following components: $s_t$, the physical motion speed of the target in a horizontal direction; $s_s$, the surrounding motion speed; $s_m$, the speed of the illusory motion component in a vertical direction, induced by the surrounding stimulus; and $\theta$ and $s$, the direction and speed of the perceived motion, respectively. In the experiment, the subjects viewed the stimulus under two conditions of the target temporal frequency, with a value of 0.0125 or 0.025 cycles/s, whereas the surround speed was varied from 0 to 2.34 cycles/s. The spatial frequencies of the target and the surround stimulus were both 0.53 cycles/deg.

Figure 1: (a) Visual stimulus for motion induction (Takemura & Murakami, 2010). (b) Schematic illustration of motion induction by the surrounding stimulus, with the following components: $s_t$, the physical motion speed of the target in a horizontal direction; $s_s$, the surrounding motion speed; $s_m$, the speed of the illusory motion component in a vertical direction, induced by the surrounding stimulus; and $\theta$ and $s$, the direction and speed of the perceived motion, respectively. In the experiment, the subjects viewed the stimulus under two conditions of the target temporal frequency, with a value of 0.0125 or 0.025 cycles/s, whereas the surround speed was varied from 0 to 2.34 cycles/s. The spatial frequencies of the target and the surround stimulus were both 0.53 cycles/deg.

Here, we briefly review the experimental stimuli used in our previous study (Takemura & Murakami, 2010), as well as our main findings. Figures 1a and 1b show the visual stimulus and a schematic view of the subject’s perception, respectively. In the experiment, we measured subjects’ motion detection sensitivity by asking them to judge the direction of horizontal motion of a central Gabor patch, which was the target stimulus. The target Gabor patch was surrounded by an annulus displaying vertical motion as a contextual stimulus. The subjects were asked to judge the direction of the target motion from two alternatives: leftward or rightward. The carrier inside the target stimulus was a vertical sinusoidal grating barely drifting to the left or right (shown as the dark rightward-pointing arrow in Figure 1b). The surrounding annulus contained a horizontal grating and moved vertically (the dark downward-pointing arrow). The surround motion induces in the central target stimulus an illusory motion that is opposite to the motion of the surround itself (the gray dashed arrow). In the subject’s perception, as a result of integration of the physical horizontal motion component and the illusory vertical motion component, the target motion is identified as obliquely upward or downward (the gray solid arrow): the faster the surround moves, the more obliquely the subject perceives the target motion (i.e., the smaller the value of $\theta$ in Figure 1b).

Notably, the illusory vertical motion component was irrelevant to the task, since the subject’s task in the experiment was to judge the horizontal motion component of the target as leftward or rightward. The results of
the experiment showed, however, that a subject’s correct response rate for
the task increased or decreased depending on the speed of the surround-
ing stimulus, as in the data shown later in Figure 5b. When the surround
motion had a moderate speed, the subjects performed better than when
the surrounding stimulus was static. The performance began to deterio-
rate when the surround speed exceeded a critical value, and it sometimes
became worse than the baseline performance (i.e., that with the static sur-
round). These results show that a subject’s detection sensitivity to horizontal
motion can be modulated by adding a task-irrelevant motion component.
The enhancement at adequate surround speeds indicates that sensory in-
put signals are implicitly retained at early visual processing stages at a
finer precision than our normal perceptual resolution and that an optimal
testing condition helps get access to this intrinsic information. In this let-
ter, we focus on what neural computation could enable the task-irrelevant
component to modulate detection sensitivity.

For enhancement of sensitivity, integration of the horizontal and vertical
motion components seems critical, considering that their resultant vector
length (s in Figure 1b) would be longer than the vector length of the physical
movement of the target itself (|s|). It is arguable, however, whether we
can discuss such integration of illusory and physical motion components
in the same way as for the processing of a physically oblique movement
(that is produced, for example, by a plaid pattern; Adelson & Movshon,
1982; Derrington & Badcock, 1992). We therefore consider two different
models for motion integration: one model assumes that the neural response
to the illusory oblique motion is identical to that for a physically oblique
motion, while the other model assumes that the neural responses are not
exactly identical, but the suppressive surround causes a biased neural re-
response pattern, which roughly mimics the physical changes in the direction
and speed of motion. Figures 2b and 2c depict the two different cases of
hypothetical population activity given by these two encoding schemes. In
Figure 2b, the modulations in the activation pattern are identical to those
caused by physical changes in the direction and speed of motion, while in
Figure 2c, the neuronal activities are modulated by a simple gain control
process through suppressive surround-to-center interaction. From now on,
we term these two encoding schemes physical modulation and gain modula-
tion, respectively. The gain modulation is directly related to the well-known
notion of an antagonistic surround (Allman, Miezin, & McGuinness, 1985;
Tanaka et al., 1986; Komatsu & Wurtz, 1988; Born & Tootell, 1992; Raiguel,
Van Hulle, Xiao, Marcar, & Orban, 1995; Eifuku & Wurtz, 1998; Born, Groh,
Zhao, & Lukasewycz, 2000) and would be more commonly accepted as a
model of the surround effect, but it has not, to our knowledge, been ex-
PLICITLY established which model truly provides better approximation for
the real activation pattern of a whole neural population. Thus, in this let-
ter, we do not discard physical modulation as a possible model of the
neuronal activity while seeing an induced motion. In addition, the physical
modulation model also predicts the human performance tested with a physically oblique motion (not an illusory induced motion), which remains to be examined by psychophysical experiments. The subsequent sections provide detailed formulations of these two models and analyze their encoding accuracies in a consistent methodological framework. The model plausibility, relation to other work, and experimental predictions are discussed at the end of this letter.
In this section we construct a model of the neural population response to a visual stimulus. We consider a population of motion-sensitive neurons, which are found in the higher visual cortex of primates. Neurons in cortical area MT/V5 have response tuning over visual motion direction and speed at each retinal location. Previous intensive investigation in area MT has revealed strong relationships between MT neural activities and the subject's perceptual performances in visual motion detection or discrimination tasks (Britten, Shadlen, Newsome, & Movshon, 1992; Celebrini & Newsome, 1995). The activities of those neurons are known to be suppressed by the presence of a surrounding stimulus moving in their preferred directions (Allman et al., 1985; Komatsu & Wurtz, 1988; Born & Bradley, 2005), and this suppression is considered as a possible physiological substrate for the induced motion. Furthermore, the typical size of the classical receptive field is consistent with the size estimated by psychophysical experiments (Murakami & Shimojo, 1993).

As already noted, we consider two encoding schemes. In both cases, the stimulus is defined by two variables: the horizontal motion speed \( s_t \) and the vertical motion speed \( s_m \). For convenience, and without loss of generality, we assume that the vertical component is always in the upward direction. In the gain modulation model, this corresponds to assuming that motion in the surround \( s_s \), and thus the suppressive signal, is in the downward direction.

### 2.1 Physical Modulation Model

For the physical modulation model, we consider the neural population that codes the two-dimensional motion vector \((s_t, s_m)\). It is reparameterized in polar coordinates as

\[
(s, \theta) = \left( \sqrt{s_t^2 + s_m^2}, \arctan(s_t/s_m) \right),
\]

where \( \theta \) and \( s \) are the motion direction and speed, respectively \((-\pi \leq \theta < \pi, s > 0)\). Here, \( \theta > 0 \) when the motion is oblique in a clockwise direction from upward, \( s_t > 0 \) when the target moves rightward, and \( s_m > 0 \) when the induced motion is upward (see Figure 1a). In this study we assume that the strength of the illusory motion \( s_m \) is equal to the surround speed parameter \( s_s \), but it is a straightforward manipulation to replace this relation by \( s_m = F(s_s) \), where \( F \) is an arbitrary function obtained in a complementary experiment (such as a direction-matching task).

In the subsequent part of this section, we formalize the neurons’ population response at motion direction \( \theta \) and speed \( s \). These stimulus parameters are transformed into the responses of \( N \) neurons with different preferred directions, covering all possible directions. The mean firing rate of the \( i \)th neuron, \( \lambda_{i,\theta} \), is determined by the stimulus parameters and the preferred
motion direction of the cell, $\phi_i$:

$$\lambda_i = \lambda(\phi_i, \theta, s) \equiv \frac{\lambda_b (1 + g(s) f(\theta - \phi_i))}{Z(s)}, \quad (2.2)$$

$$Z(s) = \frac{1}{N} \sum_{i=1}^{N} (1 + g(s) f(\theta - \phi_i)), \quad (2.3)$$

where $Z(s)$ is a normalization constant (Heeger, 1992; Simoncelli & Heeger, 1998). $\lambda_b$ determines the spontaneous firing rate, which we set at $\lambda_b = 20$ spikes/s (Heeger, Simoncelli, & Movshon, 1996; Britten, Shadlen, Newsome, & Movshon, 1993). The function $g$ determines the neuronal response to different speeds of motion, satisfying $g(0) = 0$; here, we model it as $g(s) \equiv a \ln(1 + s)$, where $a = 0.8$ (the order of the resulting response is roughly consistent with previous experimental data; Perrone & Thiele, 2001; Mikami, Newsome, & Wurtz, 1986). Although MT cells in a real cortex are tuned with respect to motion speeds and $g(s)$ is not monotonic (Perrone & Thiele, 2001; Mikami et al., 1986; Priebe, Cassanello, & Lisberger, 2003), in this study we consider motion limited to a regime of very slow speeds, and we approximate speed tuning by a monotonic function. The function $f$ represents the direction tuning, which we model using a circular gaussian (von Mises) function,

$$f(\theta - \phi_i) \equiv e^{\kappa(-1 + \cos(\theta - \phi))}, \quad (2.4)$$

where we take $\kappa = 3$ (Britten & Newsome, 1998; Jazayeri & Movshon, 2006). From the above equations, we can extract the neuronal response surfaces for two-dimensional motion, as shown in Figure 2a. At a sufficiently large population size $N$, the summation over the cells in equation 2.3 is replaced by integration over the cells’ preferred directions. If we assume that the preferred directions are distributed uniformly, the mean firing rate of a cell with preferred direction $\phi$ is given by

$$\lambda(\phi, \theta, s) = \lambda_b \frac{(1 + g(s) f(\theta - \phi))}{1 + g(s) e^{-\kappa} I_0(\kappa)}, \quad (2.5)$$

where $I_0(\kappa)$ is the modified Bessel function of first order:

$$I_0(\kappa) \equiv \frac{1}{2\pi} \int_{-\pi}^{\pi} d\phi \ e^{\kappa \cos \phi}. \quad (2.6)$$

Figure 2b shows the resulting population activity over the entire range of preferred directions. The figure illustrates that an increase in the surround motion causes a heightening of the peak activity, as well as a shift in its
Figure 3: Schematic of the gain modulation model. The model is described by suppressive gain modulation of the bottom-up input signal and subsequent response normalization that maintains a constant population-average response. The figure shows the case in which the target motion is downward and the surround motion is rightward, with three cases of the surround motion speed (static, slow, and fast). The surround suppression and response normalization can be integrated into a single process, although the two factors are presented separately for an illustrative purpose.

2.2 Gain Modulation Model. For gain modulation, the mean firing rate $\lambda_i$ is defined with a different equation from the one introduced for physical modulation. We consider suppressive gain modulation that is directionally tuned. Such suppression can be attributed to lateral inhibitions within area MT (Born et al., 2000) or to inhibitory surround signals carried by the bottom-up inputs from an earlier processing stage (Rust, Mante, Simoncelli, & Movshon, 2006). Figure 3 provides a schematic illustration of the model. In this model, we introduce a second term that modulates the gain of the original population activity:

$$\lambda_i = \lambda_D(\phi_i, s_t, s_s) \equiv \frac{\lambda_b(1 + g(|s_t|) f(\pi/2 - \phi_i))/(1 + g_D(|s_s|) f_D(\pi - \phi_i))}{Z_D(s_t, s_s)}.$$  

(2.7)

We slightly modified the gain functions ($g$ and $g_D$) to take the absolute values of speeds ($s_t$ and $s_s$, respectively) as their variables. The normalization term is defined by

$$Z_D(s_t, s_s) \equiv \frac{1}{2\pi} \int_{-\pi}^{\pi} d\phi \frac{1 + g(|s_t|) f(\pi/2 - \phi_i)}{1 + g_D(|s_s|) f_D(\pi - \phi_i)}.$$  

(2.8)
The second term in the numerator on the right-hand side of equation 2.7 represents the divisive gain suppression. The functions $g_D$ and $f_D$ have the same forms as $g$ and $f$, respectively, but they can have different parameters. It should be noted that the surround suppression and response normalization can be integrated into a single process, although they are shown as two separate steps in Figure 3 for convenience of illustration. Figure 2c shows the resulting pattern of the population response under the gain modulation model. As a consequence of the normalization process, the gain suppression by the surround causes a shift and heightening of the peak, as seen under the physical modulation model. These changes in the activation pattern can be related to the subject's perception, in which the perceived target movement is oblique and faster than the physical motion.

Now our question is how these modulations of neural encoding can affect the subject's perceptual ability to detect the horizontal target motion. In this study, we answer this question by analyzing how accurately the stimulus values can be discriminated according to the neural population activity. Before proceeding to the analytical details, we use schematic drawings, shown in Figure 4, to provide an intuitive description of the relation between the task difficulty and the amount of bias in the neural encoding. To tell in which direction the target moved, the subject has to discriminate the population activities evoked by leftward and rightward targets. If the two activation patterns have similar shapes, the discrimination is vulnerable to noise in the neural responses; as a result, the subject is more likely to misjudge the leftward and rightward targets. For both the physical modulation model and the gain modulation models, a faster surround causes a greater shift and heightening of the activation peaks (see Figures 2b and 2c). These effects are schematically depicted on the left side of Figure 4, with icons indicating the three representative surround speed conditions. When the surround is static, the leftward and rightward targets respectively evoke neural population activities with small peaks at 90 and $-90$ degrees. Since the motion signals are weak for barely moving targets, both activation patterns are nearly flat and therefore do not differ widely from each other, meaning that the task would not be easy for the subject. When the surround moves at a moderate speed, the difference between the activities evoked by the leftward and rightward targets becomes more significant because of the heightening of the activation peaks; this can lead to better subject performance than that in the static surround condition. For an extremely fast surround, however, we expect the discrimination performance to deteriorate again, because a strong surround modulation causes both the leftward and rightward targets to evoke similar activation patterns, whose peaks are aggregated at the direction opposite that of the surround motion. The intuitive insights in these illustrations suggest that the opposing effects of enhancement and degradation in the subject's motion detection performance can be qualitatively
explained for both the physical modulation and gain modulation models by considering a neural population encoding bias that leads to a trade-off between the shift and the heightening of the activation peaks. Note that the scale difference in the neural firing rate is critical because magnifying the mean firing rates leads to an increase in the signal-to-noise ratio in discriminating neural activities. To characterize the quantitative aspects of the surround effects on the detection sensitivity, we need to assume a probabilistic model of the neural response noise as well as a concrete model of the decoding process, in which the stimulus values are estimated from trial-by-trial neural population responses. The remaining part of this section describes our population decoding model and how we quantify its performance.
2.3 Decoding. In the decoding stage, the two-dimensional stimulus value \((s_t, s_s)\) is estimated from the spike counts emitted by the neurons. We have a set of firing rates \(r\) for the \(N\) neurons in each trial \((r = [r_1, \ldots, r_N])\). In the model, we assume that each neuron fires independently, and then the joint probability for the whole population is given by

\[
P(r \mid s_t, s_s) = \prod_{i=1}^{N} P(r_i \mid s_t, s_s). \tag{2.9}
\]

The number of spikes for each neuron varies across trials. We assume that this variability is described by a Poisson process (Seung & Sompolinsky, 1993; Jazayeri & Movshon, 2006):

\[
P(r_i \mid s_t, s_s) = \left(\frac{\lambda_i T}{r_i T}\right)^{r_i} \frac{e^{-\lambda_i T}}{(r_i T)!}, \tag{2.10}
\]

where \(T\) is the length of time during which spikes are sampled. The log likelihood of \((s_t, s_s)\) for a given set of neural firing rates, \(r\), can be written in a factorized form as

\[
L \equiv \ln P(r \mid s_t, s_s) = \sum_{i=1}^{N} \ln P(r_i \mid s_t, s_s). \tag{2.11}
\]

An explicit expression of the log likelihood for the physical modulation model is given by

\[
L = T \sum_{i=1}^{N} r_i \left\{ \ln (1 + g(s) f(\theta - \phi_i)) - \ln Z(s) \right\} + \text{const}. , \tag{2.12}
\]

while for the gain modulation model, we have

\[
L = T \sum_{i=1}^{N} r_i \left\{ \ln (1 + g(|s_i|) f(\pi/2 - \phi_i)) - \ln (1 + g_D(|s_i|) f_D(\pi - \phi_i)) - \ln Z_D(s) \right\} + \text{const}. , \tag{2.13}
\]

where const. represents the terms that do not depend on either \(s_t\) or \(s_s\). Now we assume an ideal observer, which maximizes the log likelihood:

\[
(\hat{s}_t, \hat{s}_s) \equiv \arg \max_{(s_t, s_s)} L . \tag{2.14}
\]
Here, \((\hat{s}_t, \hat{s}_s)\) gives an unbiased estimate of \((s_t, s_s)\). In the case of the physical modulation model, the vector \((\hat{s}_t, \hat{s}_s)\) is mapped through equation 2.1 to \((\hat{\theta}, \hat{s})\), which is defined as

\[
(\hat{\theta}, \hat{s}) \equiv \arg \max_{(\theta, s)} L.
\] (2.15)

Note that the decoding in our model includes estimation of the strength of the stimulus (the speed of motion in this case), in addition to the direction.

2.4 Quantification of Decoding Performance. Although the decoding process is deterministic, the results of estimation vary from trial to trial because of the stochasticity of each neuronal firing in the encoding stage. The degree of variability determines subject performance in a task: larger variability produces more failed trials, resulting in a lower success rate.

As a measure of the variability in the estimate of \((\hat{s}_t, \hat{s}_s)\), we can use the Fisher information of the neural population code (Paradiso, 1988; Seung & Sompolinsky, 1993; Dayan & Abbott, 2001). In our problem setting, we are interested only in the distribution \(P(\hat{s}_t \mid s_t, s_s)\), which is obtained by marginalization over the task-irrelevant parameter \(s_s\). Since the variability of \(\hat{s}_s\) does not influence that distribution, we need to estimate only the variability of \(\hat{s}_t\). We denote the Fisher information of \(\hat{s}_t\) by \(J_{s_t}\). Assuming unbiased estimation, its inverse gives a lower bound for the variability of \(\hat{s}_t\) through the Cramér-Rao inequality:

\[
\text{Var}[\hat{s}_t] \geq J_{s_t}^{-1},
\] (2.16)

where

\[
J_{s_t}(s_t, s_s) = \frac{NT}{2\pi} \int_{-\pi}^{\pi} d\phi \ J_{s_t}^{\text{Local}}(\phi; s_t, s_s),
\] (2.17)

\[
J_{s_t}^{\text{Local}}(\phi; s_t, s_s) \equiv \mathbb{E} \left[ -\frac{1}{T} \frac{\partial^2 \ln P(r(\phi) \mid s_t, s_s)}{\partial s_t^2} \right]_{s_t, s_s}
\] (2.18)

\[
= \frac{1}{\lambda(\phi, \theta, s)} \left( \frac{\partial \lambda(\phi, \theta, s)}{\partial s_t} \right)^2.
\] (2.19)

for the physical modulation model. The explicit forms of \(J_{s_t}^{\text{Local}}\) and \(J_{s_t}\) are analytically derived in appendix A, although the simulation results in the following section were obtained by numerically computing the integrals (through discretization over the neuronal preferred speeds and directions). The Fisher information for the case of the gain modulation model is obtained by simply substituting \(\lambda_D(\phi, s_t, s_s)\) for \(\lambda(\phi, \theta, s)\) in the above equation.
For a sufficiently large number of cells, the distribution of the maximum likelihood estimator asymptotically approaches a normal distribution, with the estimation variability matching the limit given by equation 2.16 (Dayan & Abbott, 2001):

\[ P(\hat{s}_t \mid s_t, s_s) = N \left( \hat{s}_t; s_t, \mathcal{J}^{-1}_{s_t} \right) . \] (2.20)

When the task is a single-interval, two-alternative forced choice, the predicted correct response rate for the human subject is obtained by using \( \mathcal{J}_{s_t} \) as follows:

\[ \text{CR}(s_t, s_s) = \Phi(s_t\sqrt{\mathcal{J}_{s_t}}), \] (2.21)

where \( \Phi \) is the cumulative distribution function of the standard normal distribution:

\[ \Phi(x) \equiv \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} \exp \left( -\frac{\xi^2}{2\sigma^2} \right) . \] (2.22)

We note, again, that in the case of the gain modulation model, the values \((s_t, s_s)\) are not directly related to the motion vector perceived by a real subject. From a biological viewpoint, it is unlikely that the human brain decodes the target physical motion \( s_t \) with full knowledge of the encoding model given by equation 2.7. Nevertheless, the Fisher information still provides a useful upper-bound measure for the accuracy of neuronal coding, which then yields a theoretical upper limit for subject performance. Later, we discuss a more biologically plausible decoding scheme and its consequences.

3 Results

Here we show the model predictions on the subject’s performances at the motion detection task as in the psychophysical experiments in Takemura and Murakami (2010). First, the results for the physical modulation model and the gain modulation model are shown in sections 3.1 and 3.2, respectively. We also examine the predictability of the model from the viewpoint of the conventional model of motion direction decoding in section 3.3.

3.1 Physical Modulation Model. Figure 5c shows the Fisher information for the physical modulation model, calculated with equations 2.1, 2.17, and 2.19 at different vertical speeds. Following Takemura and Murakami (2010), here we continuously vary the speed of the vertical motion component from \( s_s = 0 \), with the task-relevant speed \( s_t \) fixed at 0.0125 or 0.025 cycles/s. Under both target speed conditions, the Fisher information increases
Figure 5: Psychophysical data and predictions under the physical modulation model. (a) Fisher information estimated from the data shown in b (see the text for details). (b) Correct response rates obtained by a psychophysical experiment (Takemura & Murakami, 2010), where the leftmost markers represent the values for the condition in which \( s_2 = 0 \). (c) Fisher information per cell, averaged over all directions of motion preferred by the cells. (d) Correct response rates for the task, computed with \( N = 1000 \) for the number of cells and \( T = 0.5 \) s for the duration of observation. To simplify the presentation in this and subsequent figures, we omit the stimulus-dependent dimension \( (s^2/cycles^2) \) from the label for the Fisher information. (e) Local contributions of direction-selective cells to the Fisher information, shown for three typical conditions of the surround speed (0, 0.1, and 2 cycles/s). The arrows located at the rightmost end of the graph represent the mean values of the Fisher information over the cells, and the number to the right of each arrow indicates the surround speed. The target speed was \( s_1 = 0.025 \) cycles/s.

until it reaches a peak, and then it decreases. Substituting the calculated values into equation 2.19 yields a prediction on the subject’s performance of the task, as shown in Figure 5d. Clearly the predicted performance is consistent with the experimental data shown in Figure 5b. In addition, we can also reverse-engineer the Fisher information from the real data by using equation 2.19; note that here, we assume that the subject’s perception is determined by an unbiased estimator with the minimum variance and a gaussian distribution. Figure 5a shows the estimated values of the Fisher information. Comparing this figure and Figure 5c confirms the consistency
between the model prediction and the experimental data in the domain of
the Fisher information as well.

What causes enhancement and degradation in subject performance? We
have given the results so far in terms of the Fisher information carried by
the full population of neurons with different preferred directions. Revisit-
ing the step before taking the integral in equation 2.19 yields a profile of the
contribution of each neuron to the total Fisher information. Figure 5e shows
the distribution of the local contributions, $J_{\text{Local}}(\phi; s_t, s_s)$, of the units with
different preferred directions, plotted at three typical values of the vertical
speed. This figure demonstrates that the locus of the effective cell popula-
tion strongly depends on the speed of the vertical component: units tuned
to the horizontal direction of motion are the most informative when the
stimulus moves in a purely horizontal direction ($s_s = 0$), but adding ver-
tical motion shifts the peak loci toward the upward direction ($s_s = 0.1, 2$).
While the widths of the hills remain roughly constant, the heights of the
peaks are maximized at a particular value of the vertical speed (e.g., $s_s =
0.1$, the thin solid line in Figure 5e), resulting in the maximal increase in
both the total Fisher information and the subject performance. The char-
acteristics of the unit-wise contributions to behavior can be qualitatively
tested by psychophysical (Hol & Treue, 2001; Jazayeri & Movshon, 2007)
or neurophysiological (Shadlen, Britten, Newsome, & Movshon, 1996; Pu-
rushothaman & Bradley, 2005) means, as we will discuss later.

Figures 6a and 6b, respectively, show the Fisher information and the
correct response rate in two-dimensional space derived from the physical
modulation model, over a wide range of horizontal motion components.
Figure 6a shows that in our model, an increase in the horizontal motion
speed $s_t$ causes an increase in the value of $s_s$ that maximizes the Fisher in-
formation. Figure 6c represents the change in the predicted subject perfor-
mance shown in Figure 6b as compared to the baseline (i.e., the performance
with no vertical motion, $s_s = 0$). This figure also shows that in our model,
significant enhancement and degradation of task performance are observed
within a particular range of the target horizontal speed $s_t$, roughly $0.005 <
s_t < 0.05$. This is the very range in which the performance enhancement and
degradation were originally observed in the psychophysical experiments.

3.2 Gain Modulation Model. Figure 7 shows the results for the gain
modulation model. In the computation, we assume a maximum likelihood
estimator for the horizontal and vertical speed parameters ($s_t, s_s$), as de-
scribed in the previous section. The tuning functions, $f$ and $g$, in the excita-
tory term in equation 2.7 are the same as in the physical modulation model.
For the suppressive term, we set a slightly broader directional tuning ($f_D(\theta)$
$= \exp(\kappa_D(-1 + \cos \theta))$, where $\kappa_D = 2.5$) and a larger gain for the speed
response ($g_D(|s_s|) = a_D \ln(1 + |s_s|)$, where $a_D = 30$); the general properties
of the predicted performance are independent of the parameter changes.
Figures 7a and 7b, respectively, show the Fisher information and the correct
Figure 6: Fisher information and the correct response rate following the physical modulation model over a wide range of vertical and horizontal speeds. (a) Two-dimensional representation of the mean Fisher information, which was partially represented in Figure 5c. (b) Two-dimensional representation of the correct response rate, which was partially represented in Figure 5d. (c) Relative increase and decrease in the correct response rate as compared to when the vertical speed was \( s_v = 0 \).
response rate derived from it as functions of the target speed. Clearly the model captures the qualitative features of the experimental data shown in Figures 5a and 5b, as in the case of the physical modulation model. Figures 7d to 7f illustrate that the model results here also agree with the physical modulation results, in that significant enhancement and degradation of the correct response rate are observed when the target speed is set within a particular range \((0.005 < s_t < 0.05)\).

A main difference from the physical modulation model arises in the domain of the local contributions of units to task performance. Figure 7c shows the distribution of the unit-wise contributions to the total Fisher information in the gain modulation model. Although the directions of the peak loci shifts are consistent with those seen in Figure 5e, the amplitudes of the shifts are much more moderate in this case. In particular, units making no contribution never depend on the vertical speed, which is not the case in the physical modulation model. This is natural, considering that gain modulation affects only the gain of each unit response, not the tuning properties of each unit response, such as direction preference. To show this explicitly, we rewrite equation 2.7 as

\[
\lambda_D(\phi, s_t, s_s) = v(\phi, s_s)\lambda(\phi_i, s_t, 0), \tag{3.1}
\]
where

\[ \nu(\phi, s_s) \equiv \frac{Z(s_t)/Z_D(s_t, s_s)}{1 + g_D(|s_s|) f_D(\pi - \phi)} \]  

(3.2)

is interpreted as the effective gain modulation by the addition of a surrounding contextual stimulus with speed \( s_s \). Note that \( \lambda_D(\phi_i, s_t, 0) = \lambda(\phi_i, s_t, 0) \) and \( Z_D(s_t, 0) = Z(s_t) \). From equations 2.19 and 3.1, the local Fisher information in the gain modulation model can be represented in a variable-separated form (for \( s_s \) and \( s_t \)) as

\[ J_{\text{Local}}(\phi; s_t, s_s) = \nu(\phi, s_s) J_{\text{Local}}(\phi; s_t, 0). \]  

(3.3)

This equation clearly shows that in gain modulation, the loci of ineffective units are pinned to fixed values, and the effective population does not shift as dramatically as in the physical modulation model, regardless of the contextual speed.

3.3 Directional Decoding. In the previous sections, we assumed an ideal observer who infers the speed of the horizontal component of motion in a Cartesian coordinate system from a given neural response with maximum likelihood estimation. In this section, we consider how the current model is related to conventional decoder models that use a polar coordinate system. Several models for biologically plausible implementations of maximum likelihood estimation have been proposed (Jazayeri & Movshon, 2006; Pouget, Zhang, Deneve, & Latham, 1998; Deneve, Latham, & Pouget, 1999). Conventionally, many of those models have demonstrated estimation of stimulus direction (or orientation). In our problem setting, we can also assume that a subject performs maximum likelihood estimation in the space of the direction and speed \((\theta, s)\), not in terms of the vertical and horizontal velocity components in a Cartesian coordinate system. In this case, the subject makes a judgment based on the perceived direction of motion. In this section, we examine whether the predictability of the model holds in that situation.

For the physical modulation model, the estimation is mathematically the same as the estimation of \((s_t, s_s)\), as already discussed; thus, the estimation of \((\theta, s)\) gives the same predictions on subject task performance. Here, we consider a subject using the perceived direction \( \hat{\theta} \) for the task, where \((\hat{\theta}, \hat{s})\) are the direction and speed components giving the maximum likelihood for the evoked neural population activity. The subject reports the target motion as rightward if \( 0 < \hat{\theta} < \pi \) and leftward if \( -\pi < \hat{\theta} < 0; \hat{\theta} = 0, \pi \) defines the decision boundaries. Now suppose that the target physically moves rightward. On average, over a large number of trials, \( \hat{\theta} \) is expected to be positive since it is an unbiased estimator. There will, however, be a proportion of trials in which the direction is estimated as \( \pi < \hat{\theta} < 0 \) because of trial-to-trial
Figure 8: (a) Trade-off between the bias (shift from the purely horizontal direction) and the variance in motion estimation for the physical modulation model. The dotted lines represent the shift in direction from purely horizontal. The solid lines represent the standard deviation of the estimated direction of motion ($\hat{\theta}$) given by the Fisher information for the direction component. The light and dark colors, respectively, correspond to two different cases of the target horizontal speed: $s_t = 0.0125$ and 0.025 cycles/s. (b) Distributions of the direction of estimated motion for $s_t = 0.025$ cycles/s, obtained by a Monte Carlo method. The figure presents three typical cases of the surround speed ($s_s = 0, 0.1, \text{and } 2$ cycles/s). The dark and light bars represent the cases in which the target stimulus moves rightward and leftward, respectively.

variability; in such trials, the subject reports incorrectly. The same applies when the target moves leftward. The subject’s correct response rate reaches a high level when $\hat{\theta}$ is expected to be comparatively less variable, while the correct response rate is degraded when $\hat{\theta}$ has large variability or is expected to be biased toward the boundaries $\hat{\theta} = 0, \pi$. Figure 8 depicts the relation
between the variability and the bias for varying strengths of the vertical motion component. The figure shows that adding vertical motion causes a kind of trade-off: an increase in motion speed makes the estimated direction less variable (the solid lines, shown as standard deviations), but at the same time, the expected value of the estimator (the dashed lines, shown as biases from the physical direction of motion) shifts toward the decision boundary of $\hat{\theta} = 0$. As a result of these two opposing factors, both increases and decreases in the correct response rate occur under specific conditions. Here, we again use the Fisher information and the Cramér-Rao inequality to estimate the variability. The Fisher information for the direction of motion $\theta$ is given as

$$ J_{\theta}(s_1, s_2) \equiv E \left[ -\frac{\partial^2 L}{\partial \theta^2} \bigg| s_1, s_2 \right] $$

$$ = \frac{NT}{2\pi} \int_{-\pi}^{\pi} d\phi \frac{\lambda_b \kappa^2 g(s)^2}{Z(s)} \frac{f(\theta - \phi)^2 \sin(\theta - \phi)^2}{1 + g(s)f(\theta - \phi)}. $$

(3.5)

In this case, the subject’s correct response rate can be calculated by considering a wrapped normal distribution, but we should note that this approximation would be inaccurate in the regime with slow speeds of motion. This is why we considered not the directional decoding but the decoding with Cartesian coordinates in the previous sections.

We can also apply exactly the same decoding scheme with the gain modulation model. Here, the stimulus was encoded according to the gain modulation model but decoded with the same algorithm used for physical modulation. In other words, the surrounding context of the visual stimulus distorts the pattern of the population response to the target motion, but the decoding process is not “aware of” this distortion, thus causing the illusory percept of a motion direction. That is, the decoder uses a mismatched model of the encoding process under the illusory motion, and the situation is not precisely maximum likelihood estimation but readout with a mismatched model. Figure 9a shows the results of directional estimation for the gain modulation model. This model also predicts the trade-off between the bias and variance, as seen in the physical modulation model. Figure 9b shows the distributions over the decoded directions of motion. The general properties are quite similar to those seen for the physical modulation model, except that in this case, the shapes of the distributions are skewed. As the vertical speed increases, because of the gain modulation distorting the population activation patterns, the mean of the distribution shifts from purely horizontal directions to obliquely upward (in this case). At the same time, the variance of the decoded direction decreases because of the heightening of the peak of activation (see Figure 2c). As a consequence, the above-noted trade-off between the bias and the variance also occurs in this model. Figure 9a summarizes this result. Figure 9c shows the final results in terms of the
Figure 9: (a) Monte Carlo simulation of the mismatched decoding for the “divisive” model. The decoder estimates the direction and speed of motion with the same algorithm as for the physical modulation model (see the text for the details). The dotted line with open symbols represents the circular mean (Zar, 1974) of the estimated direction of motion, shown as shifts in direction from purely horizontal. The solid line with filled symbols represents the square-root circular variance (Zar, 1974) of the estimated direction of motion. These values were computed from 500,000 trials. The light and dark colors, respectively, correspond to two different target horizontal speeds: \( s_t = 0.0125 \) and \( 0.025 \) cycles/s. For the simulation, we used the same parameters as for our analytical calculations. (b) Distributions of the estimated motion direction for three typical cases of the vertical speed \( (s_v = 0, 0.1, \text{and } 2 \text{ cycles/s}) \). The dark and light bars, respectively, represent the cases in which the target stimulus moved rightward and leftward, as in Figure 8. (c) Predicted correct response rates. The horizontal dashed lines represent the correct response rates at \( s_v = 0 \).

success rates for the detection task. Corresponding with the experimental data, the model predicts both the enhancement and the degradation of the detection sensitivity.

4 Discussion

We have proposed a theoretical explanation of how human motion sensitivity is enhanced and degraded by an illusion that adds a task-irrelevant
motion component. We modeled two types of population activities: one exactly matches physical changes in the stimulus direction and speed, while the other is more simply caused by a divisive gain modulation. We found that the motion detection performance of the ideal observer, who optimally decodes input stimulus parameters from a given neural population response, replicates the properties observed in human subjects for both the physical and gain modulations.

To analyze the psychophysical performance under the motion illusion, we have modeled the neuronal properties in terms of insights in the motion processing area, including motion direction tuning, speed tuning, center-surround antagonism, response normalization, and stochasticity in neuronal firing. The situation is more complex for a real cortical network, which contains elements that have been simplified in this study, such as speed tuning, pattern selectivity of MT cells, and noise correlations among neuronal firings. Nevertheless, we believe that the model still provides a basic picture of neuronal representation under illusory stimulus conditions.

The experimental results suggest that human perception is not an outcome of the optimal decoding. Since the surround stimulus is irrelevant to the task and contains no information about the center, the theoretically optimal strategy for discriminating central motion direction is to decode it at the early local processing stage, which is advantageous from the viewpoint of data processing inequality. The psychophysical experiments by Takemura and Murakami (2010), however, have shown that the subject’s perceptual ability to discriminate central motion direction strongly depends on the surround stimulus conditions, indicating that human subjects cannot ignore the surround motion even when it is completely uninformative for the task since their performances should be independent of the surround speed if they ignored the surround motion. Moreover, it is also shown that the subject’s discrimination ability is enhanced at adequate speed settings. This result further confirms that the human subjects never have direct access to the raw representation at the early processing stage but ought to make decisions based on a processed representation at the later stage even though some information about the center would be lost in the later stage. In this letter, we aimed at proposing a biologically plausible model explaining the experimental data rather than modeling the theoretically optimal decoding scheme. With our model, we found that the surround-to-center suppression and response normalization in the later motion processing stage (such as area MT in the primates) can account for the change in motion discrimination performance, considering the surround modulation causing a trade-off between shift and heightening of the peak in the evoked neural population activity.

We should also comment on the aperture problem contained in the experimental stimuli that Takemura and Murakami (2010) used. Since they used grating stimuli, the motion direction of the target stimulus is inherently ambiguous. For such stimuli, the stimulus parameters \( s_1 \) and \( s_2 \) should...
be considered as reflecting the raw information of the stimulus temporal frequencies rather than velocities of the true two-dimensional motions. The individual model units respond to the grating stimulus depending on the amounts of motion energies within their spatiotemporal frequency passbands. In other words, each unit does not represent the information more than one-dimensional motion signals, regardless of whether the modulation type is physical modulation or gain modulation (note that the “physical modulation” does not mean that each unit responds based on the a priori knowledge of the recovered two-dimensional motion vector, but means that the whole distributional shape of unit activation emulates the physical change of stimulus motion, as shown in Figure 2b). The current model assumes that the two-dimensional motion vectors are estimated from the population response of those units, as in the previous models (Simoncelli & Heeger, 1998).

Many researchers have studied models of population coding. The main contribution of our model is its success in demonstrating that the neuronal encoding bias caused by a contextual stimulus can enhance, as well as degrade, the subject’s task performance. Tzvetanov and Womelsdorf (2008) reported success in predicting human performance, particularly threshold elevation, in fine discrimination of directions of motion under the illusion of induced motion by a surrounding stimulus. For such fine discrimination, no enhancement of sensitivity was observed, which differs from the experimental data shown here. As we discuss later in this section, our problem setting is not purely defined by a fine discrimination task but has aspects of both fine and coarse discrimination, as well as intermediate situations between these end points. Tzvetanov and Womelsdorf used a more simplified model in which no normalization of neural activity is taken into account. Obviously suppression without a subsequent normalization process causes only a decrease in neuronal information, not an increase, and it does not reproduce the enhancement of subject performance shown in our data. This indirectly supports the idea of normalization over the cell population (Heeger, 1992; Heeger et al., 1996; Simoncelli & Heeger, 1998). Although the response normalization in our model is inspired by Heeger et al.’s (1996) model, their original model does not take into account the surround-to-center interaction and cannot be applied as it is to our current problem of induced motion. For a recent model of area MT, Rust et al. (2006) reported that a wide variety of MT neural responses is explained by an extension of the linear-nonlinear model, but their model does not include the response normalization or any neural interaction after the nonlinear process, whereas in our model, the response normalization is the critical process that causes the heightening of the peak activation in the neuronal population response. In addition, the stochasticity of the neuronal response, one of the key factors to explain the sensitivity changes, was not modeled in the previous studies.

We also confirmed that the proposed model can explain some previously known psychophysical phenomena other than finding of Takemura and
Figure 10: (a) Replication of the motion direction shift reported in Kim and Wilson (1997). The abscissa shows the direction of motion in the surround, and the ordinate shows the perceived direction of motion in the center, where the physical motion direction of the center was always set at 45 degrees. We used the gain modulation model and directional decoding, described in section 3.3, to replicate their result. (b) Replication of the modulation in direction discrimination threshold reported in Tzvetanov and Womelsdorf’s (2008) study. The abscissa shows the direction of motion in the surround, and the ordinate indicates the (normalized) discrimination threshold. We applied the same model as the gain modulation model and Fisher information analysis described in this letter. In making both plots in panels a and b, the model parameters were set as the same as the main simulation. The dashed line in each plot shows the performance with no motion in the surround.

Murakami (2010). Two simulation examples are shown in Figure 10. In Figure 10a, the model reproduces the result of motion direction shift reported in Kim and Wilson’s (1997) study, where the subject perceives an illusory shift of motion direction in the stimulus center depending on the surround motion direction; Figure 10b shows a reproduction of the data in Tzvetanov and Womelsdorf’s (2008) study, which measured the subject’s ability to discriminate slightly different directions of motion under the presentation of contextual motion in the surround, with the highest discrimination threshold obtained when the surround moved in a direction moderately deviating from the central motion. In both examples, the model captures the qualitative characteristics of the experimental data. In addition, recently our group found that adaptation to vertical motion causes sensitivity enhancement for horizontal motion, just as they found in the induced motion experiment (Takemura & Murakami, 2009). Since the adaptation is related to the gain modulation in the neural responses, the sensitivity changes after motion adaptation are expected to be explained by a model similar to the currently proposed one.

The results of the physical modulation model predict that the human ability of motion detection is also enhanced by adding a physical motion
component, not an illusory motion induced by its contextual surround, in the orthogonal direction. For example, we can use plaid motions to examine this prediction. Let us consider measuring the subject’s sensitivity at detecting a slow Gabor motion that is moving horizontally left or right. Superimposing on it another Gabor patch that has the same size and moves in an orthogonal direction (e.g., upward), the subject perceives a plaid pattern moving in an oblique direction (e.g., oblique upward from the pure leftward or rightward). In such a situation, our model predicts that the subject’s judgment on whether the target Gabor moved left or right should be more accurate when the added Gabor patch moved vertically at a moderate speed, even though the additional vertical motion is irrelevant to the task. Although Derrington and Badcock (1992) previously reported that subjects are more sensitive at discriminating the motion of a plaid (components oriented +60 and −60 degrees) than its components alone, it remains to be examined whether their results hold when an orthogonal plaid is used.

Our models also make predictions on the contributions of individual units to task performance. These predictions can be tested in various frameworks proposed in previous experimental work. Psychophysically, for example, motion adaptation (Hol & Treue, 2001) or addition of a subthreshold visual motion signal (Jazayeri & Movshon, 2007) modulates the activity of a particular cell population, leading to changes in the subject’s performance. The amount of performance change reflects the importance of a unit in the perceptual task. Alternatively, unit contributions to the task can be more directly tested by electrophysiological measurements of neurometric precision in discriminating stimuli (Shadlen et al., 1996; Purushothaman & Bradley, 2005). Other electrophysiological techniques make it possible to obtain direct measures of neuronal contributions to behavior, such as choice probability or mutual information (Shadlen et al., 1996; Purushothaman & Bradley, 2005).

In the physical modulation model, the two extreme situations of $s_s = 0$ and $s_s \rightarrow \infty$ respectively correspond to coarse and fine discrimination tasks. Our results for each situation agree with previous experimental findings (Hol & Treue, 2001; Shadlen et al., 1996; Purushothaman & Bradley, 2005). For example, in fine discrimination, the most informative neurons are not those tuned to the stimulus direction itself but those whose tuning peaks are located moderately away from the stimulus direction (Purushothaman & Bradley, 2005), which is consistent with the theoretically derived information distribution shown by the dotted line in Figure 5e. For the intermediate situation (e.g., the thin solid line in Figure 5e), however, profiles of the information distribution have not yet been explicitly studied. Coarse and fine discriminations have conventionally been studied through different tasks. Technically as well, the visual stimuli used in those tests tend to differ from each other (e.g., the coherence of randomly moving dots is often manipulated in a coarse discrimination paradigm, whereas the direction
of motion is the variable parameter in fine discrimination). This makes it
difficult to compare these tasks directly. The current problem setting would
work as a tool for bridging this gap. Applying the theoretical predictions
provided here, we can explore the neural computation underlying coarse
and fine discrimination, as well as intermediate situations, from a unified
viewpoint.

In regard to this issue, the gain modulation model gives some
paradoxical predictions. It predicts that subjects will perceive an oblique
motion just as they do in the physical modulation model. Therefore, with
a very fast surround, for example, subjects should feel as if they are per-
forming a fine discrimination task (because they perceive the target as
moving almost vertically and slightly leftward or rightward in this situ-
ation). In contrast, with a static surround they should feel that they are
performing coarse discrimination. Thus, the subjective task impression is
expected to differ depending on the task-irrelevant speed component. In
this model, however, the informative cell population has almost no de-
pendence on the stimulus conditions (see Figure 7c). A naturally arising
question is how the information is decoded by the later processing stages
in this situation. The information could be decoded the same ways as for
a physically oblique stimulus motion (as we have previously seen for di-
rectional decoding under the gain modulation model). In this case, instead
of the subject obtaining the benefit of not needing to switch the decoder
for different stimulus conditions, the decoding is not optimal because the
model that the decoder uses does not perfectly match the actual encoding
model. Alternatively, the stimulus information under the illusory condition
could be decoded differently from the case of a stimulus with no illusion, so
that the information loss in the decoding is minimized. Physiological mea-
surement linking the neuronal precision and behavior (Shadlen et al., 1996;
Purushothaman & Bradley, 2005) would provide clues to understand this
question.

Appendix: Explicit Forms of Fisher Information

The explicit forms of $J_{\text{Local}}$ and $J_{\text{st}}$ in equations 2.17 and 2.19 for the physical
modulation model are derived as follows. First,

$$\frac{\partial \lambda}{\partial s_t} = \frac{\partial \lambda}{\partial \theta} \frac{\partial \theta}{\partial s_t} + \frac{\partial \lambda}{\partial s} \frac{\partial s}{\partial s_t}, \quad (A.1)$$

$$\frac{\partial \lambda}{\partial \theta} = \lambda bg(s)f(\theta - \phi) \kappa \sin(\theta - \phi) Z(s), \quad (A.2)$$

$$\frac{\partial \lambda}{\partial s} = \lambda bg'(s)f(\theta - \phi) \frac{Z(s)}{Z(s)^2}, \quad (A.3)$$
\[ \frac{\partial \theta}{\partial s_t} = -\frac{1}{c} \cos \theta, \quad (A.4) \]
\[ \frac{\partial s}{\partial s_t} = \sin \theta. \quad (A.5) \]

Substituting these equations into equation 2.19 yields

\[ J_{st}^{\text{Local}}(\phi; s_t, s_s) = \frac{\lambda_b f(\theta - \phi)^2}{Z(s)(1 + g(s)f(\theta - \phi))} \left( -\frac{\kappa g(s)}{c} \cos \theta \sin(\theta - \phi) + \frac{g'(s)}{Z(s)} \sin \theta \right)^2. \quad (A.6) \]

Then, from equation 2.17,

\[ J_{s_t}(s_t, s_s) = \frac{N T \lambda_b}{Z(s)} \int_{-\pi}^{\pi} d\phi \frac{f(\phi)^2}{(1 + g(s)f(\phi))} \left( -\frac{\kappa g(s)}{c} \cos \theta \sin \phi + \frac{g'(s)}{Z(s)} \sin \theta \right)^2. \quad (A.7) \]

Acknowledgments

H.T. is funded by the Japan Society for the Promotion of Science. I.M. is funded by the Nissan Science Foundation. M.O. is funded by the Ministry of Education, Culture, Sports, Science and Technology of Japan (grant nos. 20650019, 20240020, and 18079003).

References


Received December 14, 2009; accepted March 29, 2010.
This article has been cited by: