Characterizing the spatial-frequency sensitivity of perceptual templates

Zhong-Lin Lu

Laboratory of Brain Processes, Department of Psychology, University of Southern California, Los Angeles, California 90089-1061

Barbara Anne Dosher

Memory, Attention, and Perception Laboratory, Department of Cognitive Sciences and Institute of Mathematical Behavioral Sciences, University of California, Irvine, California 92697

Received July 13, 2000; accepted January 10, 2001; revised manuscript received February 13, 2001

Filtered external noise has been an important tool in characterizing the spatial-frequency sensitivity of perceptual templates. Typically, low-pass- and/or high-pass-filtered external noise is added to the signal stimulus. Thresholds, the signal energy necessary to maintain given criterion performance levels, are measured as functions of the spatial-frequency passband of the external noise. An observer model is postulated to segregate the impact of the external noise and the internal noise. The spatial-frequency sensitivity of the perceptual template is determined by the relative impact exerted by external noise in each frequency band. The perceptual template model (PTM) is a general observer model that provides an excellent account of human performance in white external noise [Vision Res. **38**, 1183 (1998); J. Opt. Soc. Am. A **16**, 764 (1999)]. We further develop the PTM for filtered external noise and apply it to derive the spatial-frequency sensitivity of perceptual templates. © 2001 Optical Society of America

OCIS codes: 330.4060, 330.6110, 330.5510, 330.5000, 330.1880, 330.6100.

1. INTRODUCTION

What is the spatial-frequency sensitivity of perceptual templates used by the visual system when we recognize objects?

A. Channel Theory

Over 30 years of research in both neurophysiology and visual psychophysics has led to the view that the early visual system consists of spatial-frequency channels.[1–10] Retinal images of objects are decomposed into spatial-frequency components represented as channel activities. Object recognition is based on the further processing of this representation by later stages in the visual system. Even though the channel architecture of the early visual system is an important organizational principle in spatial vision, it is concerned only with the early stages of visual processing. Many important details have not been specified. For example, the relative weights of the visual channels involved in performing complex object-recognition tasks (e.g., the reading of alphanumerical characters, face recognition) cannot be predicted directly from the channel theory.

B. Existing Methods

Many methods have been developed to characterize the spatial-frequency sensitivity of perceptual templates involved in object recognition. One type of method is based on interaction phenomena: summation at threshold,[2,5,11–15] visual masking,[6,16–18] and selective adaptation.[19–22] A second type is based on visual discrimination between stimuli that are at or near the detection threshold.[8,23–27] Another method is based on the analysis of the correlation between the noise samples and human performance on a single-trial basis.[28–30] Whereas most of these methods only offer qualitative answers or have only been applied to simple visual tasks, visual masking and the correlation method have been the most widely used quantitative methods. In this article, we concentrate on the method that is based on a particular form of visual masking, i.e., masking by spatial-frequency-filtered external noise. The aim is to estimate the sensitivity of the observer to different spatial frequencies in the stimulus.

C. Filtered Noise Masking

Filtered-external-noise masking experiments have been widely used to reveal the spatial-frequency sensitivities of perceptual templates involved in object recognition.[6,17,31–48] The principle behind the filtered-external-noise method is rather simple: Only noise energy at frequencies inside the tuning of the perceptual template affects performance, whereas noise energy outside the tuning of the perceptual template does not affect performance.

D. Role of Observer Models

If human performance were only limited by the external noise, the spatial-frequency sensitivity of perceptual templates would be observed simply by measuring human performance as a function of the spatial-frequency passband of the filtered external noise. However, human performance is also limited by internal noise, owing to intrinsic stimulus variability, receptor sampling errors, filter location jitter, randomness of neural responses, and loss of information during neural transmission.[49] Thus it is

necessary to separate the contribution of external noise from that of the internal noise by applying the filtered external noise method to quantitatively characterize the spatial-frequency sensitivity of perceptual templates. For the purpose of precise estimation, an accurate observer model is necessary.

E. Perceptual Template Model

To our knowledge, the previous quantitative applications[17,44,47,48] of the filtered-external-noise method were based on a simple noisy linear amplification model (LAM), an observer model consisting of a linear amplification stage and an additive internal-noise source, whose amplitude is independent of the input.[50–53] However, it has long been recognized that the simple LAM is an incomplete model of human performance.[54,55] Recently several alternative models[56–59] have been proposed. These models provide better accounts of human performance in detecting or identifying signals embedded in various manipulated amounts of white Gaussian external noise. Although conceptually very different, these models also have very similar mathematical properties and often provide equally good accounts of human performance.[58]

The perceptual template model (PTM)[57–59] was recently developed by us as an observer model at the overall visual-system level. It is an elaborated version of the simple LAM model with two additional components: a nonlinear transducer function and a multiplicative noise source whose amplitude depends on the energy in the input. In the decision stage, as in the simple LAM model, the PTM does not assume any decisional uncertainty for tasks involving a single stimulus location and known signals.[56,60] This approach follows Burgess's demonstration that identification performance made nearly ideal use of matched filters,[60] as well as the conclusion from a direct comparison of the PTM and a simple LAM model with decision uncertainty (the EAW model by Eckstein, Ahumada, and Watson[56]).

The PTM provides at least as good of an account of a range of data as uncertainty-based elaborations of LAM approaches. We[58] found that the mathematical properties of the nonlinear transducer functions in the PTM are very similar to those of the uncertainty in the EAW model over a large range of d' levels. The PTM is preferred because: (1) Nonlinear transducer functions are a key component of models in pattern masking.[61,62] (2) The concept of a nonlinear transducer function is consistent with nonlinear properties of visual neurons.[63] (3) In stimulus identification by well-practiced observers, previous evidence suggests that stimulus uncertainty does not appear to play a major role.[60] In a recent perceptual-learning experiment studying Gabor orientation identification in peripheral vision, Dosher and Lu[59] demonstrated that the threshold ratio between two d' levels at all the external-noise levels for each observer is constant across days, even though the thresholds themselves were improved by a factor of almost three. This indicates that any hypothetical-uncertainty effects, counter to expectations, were unchanged over substantial improvements in performance. In contrast, it is reasonable to assume that nonlinear transducer functions (as in the PTM) may be unaffected by practice. (4) To account for their data, EAW had to vary the degree of uncertainty for different external noise levels for the same observer in the same experiment in nonsystematic ways. Furthermore, the EAW model fits are essentially equivalent for a wide range of values of uncertainty. (5) In accounting for the data in Lu and Dosher,[58] the best-fitting EAW models require a fairly large number of hidden channels, in the range of 20 to 200. Current models of early visual systems specify fewer visual channels.[4,12,15,19,20,64–67]

In the case of multiple external locations or of an increase in the number of possible signals, the PTM would require uncertainty extensions consistent with an ideal observer.[68] Previously, we[58] rejected the simple LAM in favor of the PTM in two-alternative forced-choice detection and identification tasks in which signal stimuli were embedded in varying amounts of external noise. In this paper, we extend our development of the PTM model to derive the tuning characteristics of the perceptual template in a two-interval forced-choice detection task in which the signal stimulus is embedded in external noise filtered in various ways. Again, we directly compare the PTM and the LAM by using statistical testing procedures. We demonstrate that the LAM provides not only a statistically poorer account for the data, but also imprecise estimates of the spatial-frequency characteristics of the perceptual template. This suggests a reinterpretation of previous LAM-based estimates of the spatial-frequency tuning of perceptual templates.

2. THEORY

This section describes the PTM. Accuracy of task performance as measured by d' is derived as a function of signal contrast c, the cutoff frequency ($f_{\text{cutoff}}^{\text{lowpass}}$ or $f_{\text{cutoff}}^{\text{highpass}}$), and the template function $T(f)$. In the experiment presented here, the perceptual task is detection, but the model is quite general.

A. Perceptual Template Model

At the overall system level, the PTM[57–59,72] has successfully characterized human behavior in detecting or identifying a signal embedded in random white Gaussian external noise. A PTM (see Fig. 1a) consists of five components: (1) a perceptual template $T(.)$; (2) a nonlinear transducer function, which raises the absolute value of the input to the γth power but retains the sign $[\text{sgn}(.)|.|^{\gamma}]$; (3) a multiplicative internal noise source of Gaussian distribution with mean zero and standard deviation proportional (with a factor N_{mul}) to the total amount of energy in the output of the perceptual template; (4) an additive internal-noise source, also of Gaussian distribution, with mean zero and standard deviation N_{add}; and (5) a decision process based on the noisy input that, depending on the task,[73] reflects either detection or discrimination.

For applications to the spatial-frequency domain, the perceptual template is assumed to be space–time separable:[74]

$$T(.) = T(f_x, f_y, t) = T_s(f_x, f_y)T_t(t), \qquad (1)$$

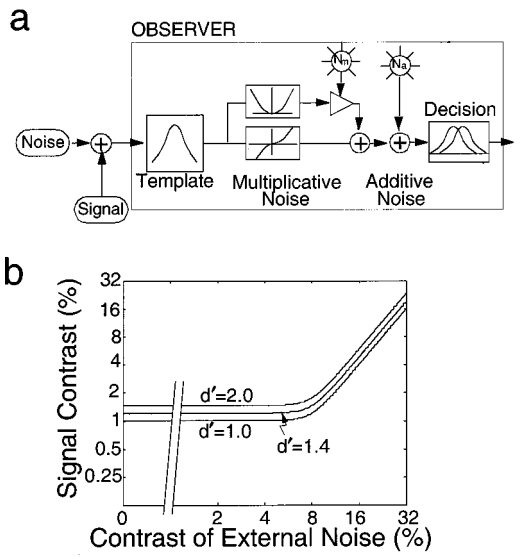

Fig. 1. a, Noisy PTM. There are five major components: (1) a perceptual template, (2) nonlinear transducer functions, (3) a multiplicative internal-noise source, (4) an additive internal-noise source, and (5) a decision process. A good example of a perceptual template is a spatial-frequency filter $F(f)$, with a center frequency and a bandwidth such that a range of frequencies adjacent to the center frequency pass through with smaller gains. The nonlinear transducer function takes the form of an expansive power function. Limitations of human observers are modeled as equivalent internal noise. Multiplicative noise is an independent noise source whose amplitude is proportional to the (average) amplitude of the output from the perceptual template. Additive internal noise is another noise source whose amplitude does not vary with signal strength. Both multiplicative and additive noise is added to the output from template matching, and the noisy signal is the input to a task-appropriate decision process. b, threshold versus contrast functions from a PTM model. The signal contrast necessary to achieve different criteria ($d' = 1.0, 1.4, 2.0$) is a function of the contrast of external noise.

where f_x, f_y are indexes of spatial frequencies. Without losing generality, the spatial component of the perceptual template is scaled such that:

$$\int\int T_s^2(f_x, f_y)\mathrm{d}f_x\mathrm{d}f_y = 1.0. \quad (2)$$

The nonlinear transducer function and the noise sources in a PTM model may be characterized by systematically manipulating the amount (or contrast) of white external noise in the stimulus and measuring the signal contrast (threshold) required to reach certain criterion performance levels (threshold versus contrast functions, Fig. 1b). In fact, Lu and Dosher[58] showed that TVC functions at three different criterion levels are necessary to fully characterize the nonlinearity and the multiplicative noise in a PTM. Three criterion levels are also sufficient to localize changes in the state of the observer to different aspects of the PTM.[59]

The spatial-frequency sensitivity of the template in a PTM may be characterized by systematically manipulating the spatial-frequency composition of the external noise and measuring the signal contrast required to reach certain criterion performance levels [threshold-versus-frequency (TVF) functions, Fig. 2i]. This approach uses noise filtering to specify the spatial-frequency characteristics of the template and three criterion performance levels to specify the nonlinearity and the noise sources in the PTM. This article develops the filtered-noise approach within the PTM and provides an empirical application.

B. d' Function

1. Signal and External Noise in the Input Stimulus

Both the signal (Figs. 2e and 2f) and the external noise (Figs. 2a–2d) in the input stimulus can be expressed as functions of space and time or, equivalently, functions of spatial frequency and time. In most applications, signal in the stimulus image is space–time separable:

$$S(f_x, f_y, t) = cS_s(f_x, f_y)S_t(t), \quad (3)$$

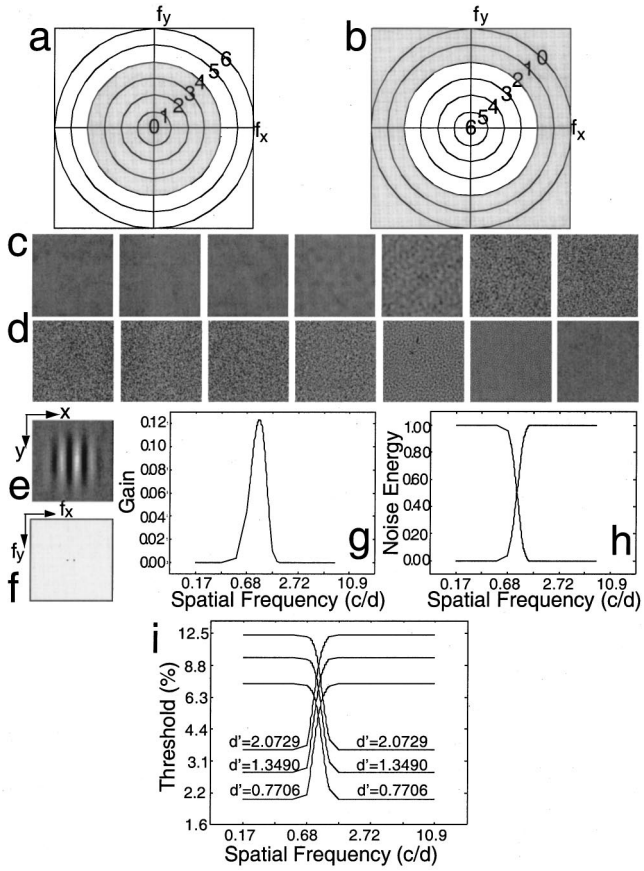

Fig. 2. a, Two-dimensional low-pass spatial-frequency filters with seven different passbands. b, two-dimensional high-pass spatial-frequency filters with seven different passbands. c, examples of low-pass-filtered noise. From left to right, the examples resulted from filtering a white Gaussian noise image through the seven filters in a. d, examples of high-pass-filtered noise. From left to right, the examples result from filtering white Gaussian noise images through the seven filters in b. e, a Gabor template in real space. f, the spatial-frequency power spectrum of the Gabor template in e. g, the gain (amplification factor applied to each spatial frequency) of the template in e as a function of spatial-frequency (distance from the origin in f). h, the amount of noise energy passed through the perceptual template in e, f, and g as a function of the passband of the low-pass filters (increasing with cutoff frequency) in a and the passband of the high-pass filters (decreasing with cutoff frequency) in b. i, the predicted contrast thresholds as functions of the passband of low-pass filters (increasing with cutoff frequency) in a and the passband of the high-pass filters (decreasing with cutoff frequency) in b at three performance levels for a perceptual template model with its template described in e, f, and g.

where c is the contrast of the signal, $S_t(t)$ is the time course of the signal stimulus, $S_s(f_x, f_y)$ is the Fourier transformation of the signal. $\int\int |S_s(f_x, f_y)|^2 \mathrm{d}f_x \mathrm{d}f_y = 1.0$ because $S_s(x, y)$ is normalized. This normalization allows the definition of an effective contrast c.

For space–time-separable filtered Gaussian external noise, the external noise in the stimulus image is:

$$N_{\text{ext}}(f_x, f_y, t) = N_{\text{ext},s}(f_x, f_y) N_{\text{ext},t}(t) = \sigma_{\text{ext}} \mathbf{G}(f_x, f_y) F(f_x, f_y) N_{\text{ext},t}(t), \quad (4)$$

where σ_{ext} is the standard deviation of the Gaussian external noise before filtering. $\mathbf{G}(f_x, f_y)$ is the Fourier transformation of an image whose pixel contrasts are identically distributed, independent standard Gaussian random variables. Therefore at every point (f_x, f_y), $\mathbf{G}(f_x, f_y)$ is a complex number whose real and imaginary parts are both identically distributed, independent Gaussian random variables. The expected value of $|G(f_x, f_y)|^2$ is 1.0. $F(f_x, f_y)$ is the filter applied to the external noise by the experimenter. $N_{\text{ext},t}(t)$ is the time course of the external noise. In the following application, σ_{ext}, which determines the unfiltered external noise contrast, will be a constant.

2. Template Matching

The signal in the stimulus is processed through the perceptual template to yield a template response to the (space–time separable) signal stimulus S_1:

$$S_1 = \int\int\int T(f_x, f_y, t) S(f_x, f_y, t) \mathrm{d}f_x \mathrm{d}f_y \mathrm{d}t = c \int\int T_s(f_x, f_y) S_s(f_x, f_y) \mathrm{d}f_x \mathrm{d}f_y \int T_t(t) S_t(t) \mathrm{d}t. \quad (5)$$

Let $f = (f_x^2 + f_y^2)^{1/2}$, let $T_s(f)$ be the average gain of the template at radius f and let $S_s(f)$ be the average magnitude of the stimulus at radius f. Then, under certain conditions,[75] the output of the template to the stimulus is

$$S_1 \approx 2\pi c \int T_s(f) S_s(f) f \mathrm{d}f \int T_t(t) S_t(t) \mathrm{d}t. \quad (6)$$

The (space–time separable) external noise in the stimulus is also processed through the perceptual template:

$$N_1 = \int\int\int T(f_x, f_y, t) N_{\text{ext}}(f_x, f_y, t) \mathrm{d}f_x \mathrm{d}f_y \mathrm{d}t = \int\int T_s(f_x, f_y) N_{\text{ext},s}(f_x, f_y) \mathrm{d}f_x \mathrm{d}f_y \int T_t(t) N_{\text{ext},t}(t) \mathrm{d}t = \sigma_{\text{ext}} \int T_s(f_x, f_y) \mathbf{G}(f_x, f_y) F(f_x, f_y) \mathrm{d}f_x \mathrm{d}f_y \times \int T_t(t) N_{\text{ext},t}(t) \mathrm{d}t. \quad (7)$$

Because $\mathbf{G}(f_x, f_y)$ is made of identically distributed, independent standard Gaussian random variables, the expected value of N_1 is 0. The variance of N_1 is the sum of the variance at every point in the Fourier space:

$$\sigma_{N_1}^2 = \sigma_{\text{ext}}^2 \int\int T_s^2(f_x, f_y) F^2(f_x, f_y) \mathrm{d}f_x \mathrm{d}f_y \times \left[\int T_t(t) N_{\text{ext},t}(t) \mathrm{d}t \right]^2. \quad (8a)$$

If we define $f = (f_x^2 + f_y^2)^{1/2}$, $T_s(f)$ as the average gain of the template at radius f, and $F(f)$ as the amplitude of the external noise at radial frequency f, we can express $\sigma_{N_1}^2$ in the polar-coordinate system. The variance of N_1 is under certain circumstances:[76]

$$\sigma_{N_1}^2 \approx 2\pi \sigma_{\text{ext}}^2 \int T_s^2(f) F^2(f) f \mathrm{d}f \left[\int T_t(t) N_{\text{ext},t}(t) \mathrm{d}t \right]^2. \quad (8b)$$

Because experimentally S_1 and σ_{N_1} can only be known to a constant, without losing generality, we set:

$$\alpha = 2\pi \int T_t(t) S_t(t) \mathrm{d}t, \quad (9a)$$

$$2\pi \left[\int T_t(t) N_{\text{ext},t}(t) \mathrm{d}t \right]^2 = 1.0. \quad (9b)$$

Thus

$$S_1 = \alpha c \int T_s(f) S_s(f) f \mathrm{d}f, \quad (10a)$$

$$\sigma_{N_1}^2 = \sigma_{\text{ext}}^2 \int T_s^2(f) F^2(f) f \mathrm{d}f. \quad (11)$$

In the application described here, both low-pass and high-pass ideal filters were used in the generation of filtered external noise. The low-pass and high-pass filters can be described by using step functions $A_i(f)$ $(i = 1, 2, \ldots, N)$ defined on consecutive, concentric annular regions starting from the center of the Fourier space (dc), where for each step function, the points on the annulus map into 1.0, and the points outside the annulus map into 0. An ideal low-pass filter (Fig. 2c) can be described as

$$F_i^{\text{lowpass}}(f) = \sum_{j=1}^{i} A_i(f). \quad (12)$$

An ideal high-pass filter (Fig. 2d) can be described as:

$$F_i^{\text{highpass}}(f) = \sum_{j=N-i-1}^{N} A_j(f). \quad (13)$$

The spatial-frequency resolution of the spatial template $T_s(f)$ is limited by the number of low-pass and high-pass filters used in the experiment, i.e., it is only possible to determine the mean $T_s(f)$ over each concentric annulus. We define the average gain of the spatial component the template per unit area in annulus j as

$$T(j) = \frac{\int T_s(f) A_j(f) f \mathrm{d}f}{f_j^2 - f_{j-1}^2}. \quad (14)$$

We further define the average magnitude of the signal stimulus per unit area in annulus j as

$$S(j) = \frac{\int S_s(f) A_j(f) f \mathrm{d}f}{f_j^2 - f_{j-1}^2}. \quad (15)$$

$S(j)$ is completely specified by the signal stimulus (Figs. 2f and 2g).

If we combine Eqs. (10a) and (15), after template matching the signal in the stimulus becomes

$$S_1 = ac\sum_{j=1}^{N} [T(j)S(j)(f_j^2 - f_{j-1}^2)] \quad (10b)$$

If we combine Eqs. (11)–(14), the variance of the ith low-pass external noise after template matching (solid curve in Fig. 2h) is

$$\sigma_{\text{lowpass}}^2(i) = \sigma_{\text{ext}}^2\sum_{j=1}^{i}\left[\int T_s^2(f)A_j^2(f)f\,df\right]$$

$$= \sigma_{\text{ext}}^2\sum_{j=1}^{i}[T^2(j)(f_j^2 - f_{j-1}^2)^2]. \quad (16a)$$

The variance of the ith high-pass external noise after template matching (dotted curve in Fig. 2h) is

$$\sigma_{\text{highpass}}^2(i) = \sigma_{\text{ext}}^2\sum_{j=N-i-1}^{N}\left[\int T_s^2(f)A_j^2(f)f\,df\right]$$

$$= \sigma_{\text{ext}}^2\sum_{j=N-i-1}^{N}[T^2(j)(f_j^2 - f_{j-1}^2)^2]. \quad (16b)$$

To summarize, after template matching, the template response to the signal in the stimulus S_1 is described by Eq. (10a), in which α and $T(j)$ $(j = 1, 2, \ldots, N)$ are unknown; the template response to the external noise in the stimulus has mean zero, and variances $\sigma_{\text{lowpass}}^2(i)$ or $\sigma_{\text{highpass}}^2(i)$ described by Eq. (16), in which $T(j)$ $(j = 1, 2, \ldots, N)$ are unknown.

3. Nonlinear Transducer Function

The nonlinear transducer function in the PTM simply raises its input to its γth power. Thus after the nonlinear transducer stage, the signal in the stimulus becomes

$$S_2 = S_1^{\gamma}, \quad (17)$$

and the variance of the external noise in the stimulus becomes

$$\text{Var}_{\text{ext}} = \sigma_{\text{passband}}^{2\gamma}(i). \quad (18)$$

4. Total Variance of All of the Noise Sources at Decision Stage

The total variance of all of the noise sources at the decision stage[77] is the sum of the variances of all the noise sources: the external noise Var_{ext} [Eq. (18)], the internal multiplicative noise $\text{Var}_{\text{mul}} = N_{\text{mul}}^2(S_2^2 + \text{Var}_{\text{ext}})$, and the internal additive noise $\text{Var}_{\text{add}} = N_{\text{add}}^2$:

$$\text{Var}_{\text{total}} = \text{Var}_{\text{ext}} + \text{Var}_{\text{mul}} + \text{Var}_{\text{add}}$$

$$= \sigma_{\text{passband}}^{2\gamma}(i) + N_{\text{mul}}^2[S_2^2 + \sigma_{\text{passband}}^{2\gamma}(i)] + N_{\text{add}}^2$$

$$= (1 + N_{\text{mul}}^2)\sigma_{\text{passband}}^{2\gamma}(i) + N_{\text{mul}}^2 S_1^{2\gamma} + N_{\text{add}}^2. \quad (19)$$

5. d' Function

Finally, the noisy signal S_2 with variance $\text{Var}_{\text{total}}$ (across trials) is submitted to the decision process. The details of the decision process depend on the particular task, e.g., detection versus identification. These are modeled elsewhere.[63] In this paper, we focus on signal discriminability d' determined by the signal-to-noise ratio:

$$d' = \frac{S_2}{(\text{Var}_{\text{total}})^{1/2}} = \frac{S_1^{\gamma}}{[(1 + N_{\text{mul}}^2)\sigma_{\text{passband}}^{2\gamma}(i) + N_{\text{mul}}^2 S_1^{2\gamma} + N_{\text{add}}^2]^{1/2}}, \quad (20)$$

with S_1 defined in Eq. (10a), and σ_{bandpass} defined in Eq. (16).

In Eq. (20), d' is expressed as a function of $N + 4$ parameters. Of the $N + 4$ parameters, signal stimulus contrast c is controlled by the experimenter; the relative efficiency of the signal stimulus α is controlled by the time course of the signal and the external noise, and the temporal properties of the template in the model; and the other $N + 2$ parameters are additional model parameters: the exponent of the nonlinear transducer function γ, the internal multiplicative noise N_{mul}, the internal additive noise N_{add}, and $T_{s,i}$ for $i = 1, 2, \ldots, N - 1$. So, to completely characterize a PTM model, $N + 4$ parameters are necessary.

6. Threshold-versus-Frequency Functions

One can solve Eq. (20) to express threshold c_τ (signal contrast required for the observer to reach a particular performance criterion level d') as a function of the cutoff spatial frequency.

$$c_r(i) = \frac{1}{\alpha\sum T(j)S(j)(f_j^2 - f_{j-1}^2)} \times \left[\frac{(1 + N_{\text{mul}}^2)\left[\sum T^2(j)(f_j^2 - f_{j-1}^2)^2\right]^{\gamma} + N_{\text{add}}^2}{1/d'^2 - N_{\text{mul}}^2}\right]^{1/2\gamma}. \quad (21)$$

Figure 2i plots $\log(c_\tau)$ required for a sample PTM to reach three criterion performance levels as functions of cutoff frequencies of the low-pass-(solid curves) or high-pass-(dotted curves) filtered external noise. These are called TVF functions. To derive these functions, we used the PTM with $\alpha = 4.0$, $N_{\text{mul}} = 0.2067$, $N_{\text{add}} = 0.0087$, $\gamma = 2.0$, and the perceptual template $T(j)$ plotted in Fig. 2g. The proportional noise energy through each passband $\sigma_{\text{passband}}^2(i)$ is plotted in Fig. 2h. The signal stimulus is a Gabor function whose spatial profile is in Fig. 2e and whose spatial-frequency profile is in Fig. 2f; the standard deviation of the external noise $\sigma_{\text{ext}} = 0.3226$.

The process of deriving the perceptual template $T(j)$ reverses the direction of what is shown in Figs. 2g–2i. Typically, TVF functions (or, equivalently, full psychometric functions) are measured experimentally. A PTM with α, N_{mul}, N_{add}, γ, and $T(j)$ as parameters is fitted to the TVF functions. One can infer the perceptual template

$T(j)$ parameters by searching for the optimal parameters of the PTM. This paper gives an example of an application of the methodology.

7. Psychometric Functions

Another way to evaluate human performance is through the underlying psychometric functions. For an M-alternative forced-choice task, the spatial-frequency characteristics of a perceptual template may be related to a psychometric function in a particular filtered noise (e.g., the ith low-pass external noise) condition using the following equation:

$$\text{Percent Correct} = \int_{-\infty}^{+\infty} g(x - d')G(x)^{M-1}\mathrm{d}x, \quad (22)$$

where $g(x)$ and $G(x)$ are the probability-density and cumulative-probability functions of a standard normal random variable, and d' is described by Eq. (20).

For a known PTM, Eq. (22) may be used to compute the theoretical psychometric function for each filtered external-noise condition. To derive the parameters of a PTM, we can compare theoretical predictions with the measured psychometric functions, and the best fitting model can be estimated. We give an example that uses a maximum likelihood estimation procedure[78] in this paper.

3. EXPERIMENT

We used the method of constant stimuli[79] to measure psychometric functions in a two-interval forced-choice (2IFC) Gabor detection task. The signal Gabor was embedded in several levels of low-pass-and high-pass-filtered external noise. From the psychometric functions, we derived TVF functions for both low-pass and high-pass external noise at three criterion performance levels. The data were then used to fully characterize the perceptual template. Four models, a LAM and a PTM each either with a single template for high-pass and low-pass-filtered noises or with two independent templates for high-pass and low-pass noises, were compared statistically. Another approach, which directly fit the full psychometric functions, was also developed in parallel.

A. Method

1. Apparatus

All of the signal and noise frames were generated and displayed in real time through use of programs based on a C++ version of Video Toolbox[80] on a 7500/100 Macintosh computer. The stimuli were presented on a Nanao Technology FlexScan 6600 monitor with a P4 phosphor and a refresh rate of 120 frames/per second, driven by the internal video card in the Macintosh. A special circuit[80] was used to combine two eight-bit output channels of the video card to produce 6144 distinct gray levels (12.6 bits).

A psychophysical procedure[81] was used to generate a linear lookup table that evenly divides the entire dynamic range of the monitor (from 1.0 cd/m^2 to 53 cd/m^2) into 256 levels. The background luminance was set at 27 cd /m^2. All displays were viewed binocularly with a natural pupil at a viewing distance of approximately 72 cm in a dimly lighted room.

2. Stimulus and Display

The signal stimuli were vertical Gabor patterns with luminance $l(x, y)$ at location (x, y) defined by

$$l(x, y) = l_0\left[1.0 + c_{\text{peak}}\sin(2\pi f x)\exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)\right], \quad (23a)$$

where $f = 0.9$ c/deg, $\sigma = 1.1$ deg, and $l_0 = 27$ cd/m^2. Each Gabor was rendered on a 128×128 pixel grid, subtending 5.9×5.9 deg^2. The peak contrast c_{peak} of the Gabor was controlled by the experimental conditions.

Thus local stimulus contrast, defined as $c(x, y) = l(x, y) - l_0/l_0$, is

$$c_{\text{peak}}\sin(2\pi f x)\exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right). \quad (23b)$$

The external-noise samples were generated in the following way. First, a 128×128 matrix was filled with real numbers, each sampled from a Gaussian random variable with mean 0 and standard deviation 0.3226. The corresponding 128×128 pixel array, viewed from 72 cm, had an expected uniform power spectrum from 0.17 c/deg to 10.88 c/deg. Seven ideal low-pass digital filters (Fig. 2a) were constructed in spatial-frequency space with cutoff frequencies at 0.17, 0.34, 0.68, 1.36, 2.72, 5.44, and 10.88 c/deg, and seven ideal high-pass digital filters (Fig. 2b) were constructed with cutoff frequencies at 10.88, 5.44, 2.72, 1.36, 0.68, 0.34, and 0.17 c/deg. Second, we computed the Fourier transformation of the noise matrix. We then applied a particular filter function to the real and the imaginary components of the Fourier transformation. Finally, we performed an inverse Fourier transformation on the filtered image to convert the external noise back to the real space. Whereas so far all computations had been performed with floating-point precision, the filtered external noise was sampled at 256 equally spaced linear levels. New noise samples were generated in each trial.

3. Design

The experiment consisted of 30 sessions, each with 196 trials, and lasted about 15 minutes. In each session, seven low-pass- and seven high-pass-filtered external noises were used. A psychometric function, sampled at seven signal-contrast levels, was measured in each filtered external-noise condition. The sampled signal contrasts were predetermined through pilot studies.

All conditions were intermixed. For each observer, data from all 30 sessions were combined to generate 14 psychometric functions, one for each low-pass- or high-pass-filtered external-noise condition. Sixty observations were made on each point of the psychometric functions.

Figure 3 depicts a typical trial. Initialized by the observer with a key press, a trial consisted of two intervals (one in which the signal was present, and the other in which the signal was absent, in random order from trial to trial) separated by 500 ms. Each interval of a trial starts with a 250-ms-fixation cross in the center of the

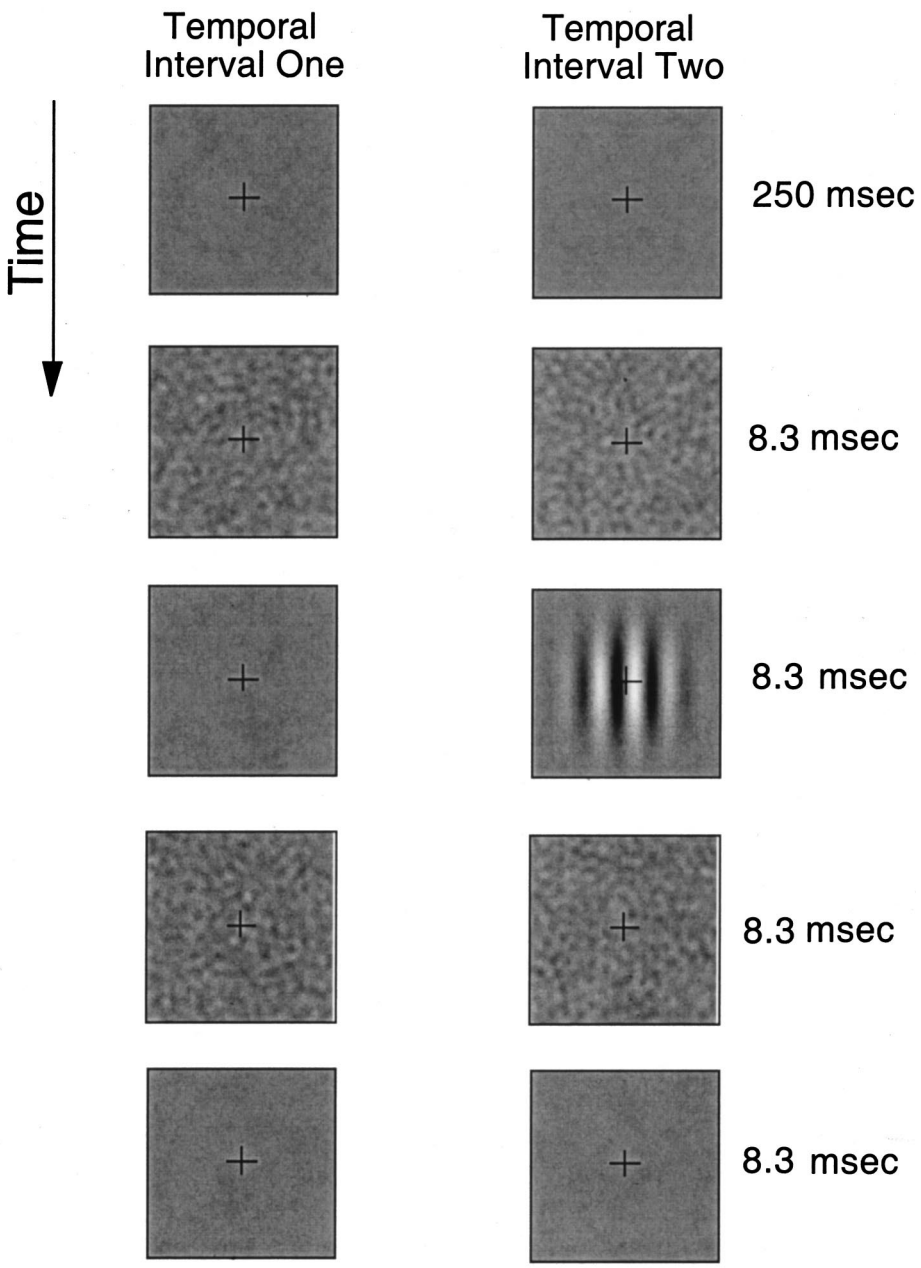

Fig. 3. A trial in the two-interval forced-choice Gabor detection task. A trial consists of two intervals. The first interval starts with a 250-ms fixation cross. In the next three stimulus frames, each lasting 8.3 ms, one external-noise frame, one signal/blank frame, and another noise frame, appear in the center of the display. The second interval starts 500 ms after the end of the first. Only one of the two intervals contains a signal Gabor. The method of constant stimuli was used to measure psychometric functions at seven different contrast levels.

monitor. In the next three refreshes, each lasting 8.3 ms, one noise frame, one signal (if signal is present) or blank (if signal is absent) frame, and another independent noise frame appear in the center of the display. All four noise frames in a given trial were identically distributed, independent random Gaussian noise images filtered through the same low-pass or high-pass filter. The observer identified with a key press the interval that contained the Gabor. A correct response was followed immediately by a brief beep.

4. Observers

Three University of Southern California students (two volunteers and one paid) with normal or corrected-to-normal vision, naive to the purposes of the experiment, served as observers in the experiment.

B. Results

Data were organized as psychometric functions. A total of 14 psychometric functions, one for each filtered external-noise condition, resulted for each observer. Each psychometric function measured the percent correct Pc in detecting the Gabor at seven different signal-contrast levels, for a total of 420 observations. To estimate threshold signal contrast, we first fit a Weibull function[82]

$$\mathrm{Pc} = \max - (\max - 0.5) \times 2^{-(c/\rho)^{\eta}} \qquad (24)$$

to each of the 14 psychometric functions using a maximum likelihood procedure.[78] In Eq. (24), max is less than 1.0; the difference between max and 1.0 reflects the fact that human observers are prone to stimulus-independent errors (or lapse). We computed threshold signal contrasts at three criterion levels: 65%, 75%, and 85% correct detection. These three performance levels correspond to the d' levels of 0.7706, 1.3490, and 2.0729 in the 2IFC tasks used in this experiment.

In the first row of Fig. 4, threshold contrasts at 75% correct detection are plotted against the cutoff spatial frequencies of the external noise for each of the three observers. Error bars[83–85] indicate the standard deviation of each threshold.

For each observer, the TVF functions at the three different performance levels are approximately vertical shifts of each other, suggesting that the threshold ratio between two performance levels for a given filtered external-noise condition is a constant across all the filtered external-noise conditions. The mean threshold ratio between 75% and 65% correct performance levels across all the external-noise conditions is 1.30 ± 0.02 for observer AT, 1.31 ± 0.03 for observer QL, and 1.33 ± 0.02 for observer SM. The mean threshold ratio between 85% and 75% correct performance levels across all the external-noise conditions is 1.24 ± 0.02 for AT, 1.25 ± 0.03 for QL, and 1.37 ± 0.04 for SM. None of the threshold ratios are significantly different from the corresponding mean. Because of the similarity of the TVF functions for different criteria, we only plotted the TVF functions at 75% correct threshold.

As discussed in Lu and Dosher[58] and in Dosher and Lu,[59,72] the LAM predicts that the threshold ratio between two performance levels in the same external noise condition be equal to the corresponding d' ratio. For the 2IFC task in discussion, the corresponding d' ratios between 75% and 65% correct and between 85% and 75% correct are 1.7506 and 1.5366, respectively. Clearly, the observed threshold ratios are inconsistent with the LAM prediction. On the other hand, these ratios are very similar to those observed in Lu and Dosher,[55] where an elaboration of the LAM to a PTM was required to account for the data. In the next section, we directly compare the LAM and the PTM statistically.

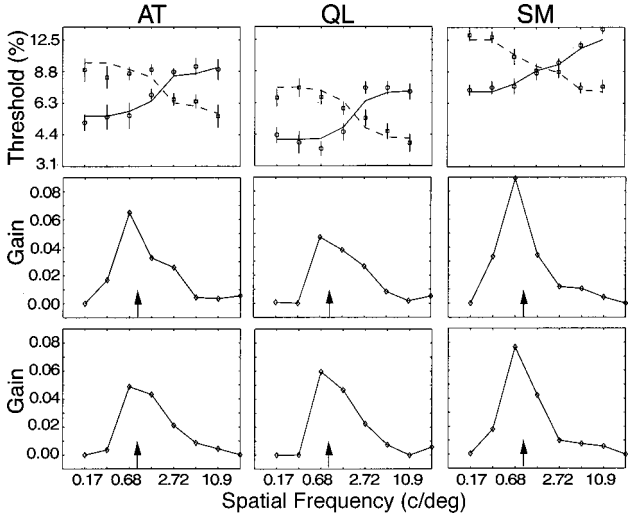

Fig. 4. First row, contrast threshold as function of passband of the low-pass and high-pass filters for three observers at 70% correct performance level. The curves were from fits to the PTM. Second row, the best-fitting spatial-frequency sensitivity of the perceptual templates (estimated from fitting the TVF functions). The arrows on the x axis indicate the center frequency of the Gabor stimulus. Third row, the best-fitting spatial-frequency sensitivity of the perceptual templates (estimated from fitting the full psychometric functions). The arrows on the x axis indicate the center frequency of the Gabor stimulus.

C. Fitting Models

1. Four Models

We considered four different models: a PTM with two independent (uncoupled) templates for high-pass and low-pass external noises (uPTM), a PTM with a single (coupled) template for high-pass and low-pass external noises (cPTM), a LAM with two independent (uncoupled) templates for high-pass and low-pass external noises (uLAM), and a LAM with a single (coupled) template (cLAM).

All four models are described directly or with some modifications by Eq. (21), which expresses threshold-signal contrast as a function of d', the perceptual template, and the other parameters of a PTM. A uPTM has 18 parameters: N_{mul}, N_{add}, α, γ, and two sets of $T(j)$ parameters, one for the high-pass conditions and the other for the low-pass conditions. A uLAM has 16 parameters: N_{add}, α, and two sets of $T(j)$ parameters. The uLAM is a reduced uPTM with $N_{\mathrm{mul}} = 0.0$ and $\gamma = 1.0$. A cPTM has 11 parameters: N_{mul}, N_{add}, α, γ, and one set of $T(j)$ parameters, coupled for the high-pass and low-pass conditions. A cLAM is a reduced cPTM with $N_{\mathrm{mul}} = 0.0$ and $\gamma = 1.0$. It has 9 parameters: N_{mul}, N_{add}, α, γ, and one set of $T(j)$ parameters, coupled for the high-pass conditions and the low-pass conditions. The uPTM and uLAM cover the possibility that the observer was performing off-frequency looking in the high-pass and low-pass conditions.[48,86–88] The four models are within a nested structure. We can use an F ratio test (in fitting TVFs) and a χ^2 test (in fitting full psychometric functions) to statistically compare them (see details below).

2. Fitting Threshold-versus-Frequency Functions

To estimate the perceptual template(s) and compare the models, we fit all four models to the observed TVF functions with the same procedure.[89] Both the uPTM and the cPTM give much better accounts for the data than the uLAM and the cLAM (see the r^2 parameters listed in Table 1.) Statistically, the 18-parameter uPTM does not provide a significantly better fit than the 11-parameter cPTM [No $F(7, 24)$ for any subject was significant ($p > 0.25$)]. The uPTM provided a much better account of the data than the uLAM [$p < 0.00001$ for all the three $F(2, 24)$ parameters.] On the other hand, the 11-parameter cPTM fits the data significantly better than the nine-parameter cLAM [all the $F(2, 31)$ parameters are statistically significant; $p < 0.0001$]. In summary, for all three observers, cPTM is the best model.

The best-fitting parameters of the cPTM are listed in Table 1. The second row of Fig. 4 depicts the spatial-

frequency sensitivity of the perceptual templates from the best-fitting PTM models to the TVF functions of the three observers. The gain of the perceptual templates in each spatial-frequency range is normalized relative to the total area of frequency space covered by that range. The arrows on the horizontal axis denote the center spatial-frequency of the signal Gabor stimulus. As seen in Fig. 4, the peaks of the estimated spatial-frequency sensitivity functions more or less coincide with the center frequency of the signal Gabor. On the other hand, the LAM model provides such a poor fit to the TVF functions that the estimated spatial-frequency-sensitivity functions from the LAM peaked either plus or minus one octave away from the center frequency of the Gabor.

Averaged across observers, the bandwidth of the best-fitting perceptual template model, defined as the ratio of the two spatial frequencies at half height of the spatial-frequency characteristics of the perceptual templates, is ~1.57 octaves (from 0.43 to 1.27 c/deg). In log units, the attenuation on the low-frequency flank of the tuning function can be described as $4\log(f/f_0)$; the attenuation on the high-frequency flank of the tuning function can be described as $-1.1\log(f/f_0)$. The estimated center frequency f_0 of the template was 0.68 c/deg. The actual frequency of the Gabor signal was 0.9 c/deg. These match within the sampling grain of the template estimate, which is limited by the number of filter bands employed in the experiment.

To further validate the method developed here, we applied the same methodology to estimate the spatial-frequency sensitivity of the perceptual templates in a four-alternative forced-choice Gabor-orientation-identification experiment in the context of spatial-attention manipulations,[90] where we concluded that the quantitatively estimated perceptual templates for both attended and unattended locations were the same, and both of them centered around the frequency of the signal Gabor.

Table 1. Parameters of the Perceptual Template Model That Best Fits the TVFs[a]

Parameter	Observer		
	AT	QL	SM
N_{mul}	0.0074	0.0136	0.4137
N_{add}	0.0355	0.0278	0.0043
α	0.2767	0.3288	0.1947
γ	1.9743	2.1296	3.2463
T(1)	0.0000	0.0016	0.0000
T(2)	0.0479	0.0016	0.0941
T(3)	0.3930	0.2849	0.5422
T(4)	0.5589	0.5425	0.6859
T(5)	0.8444	0.8438	0.7445
T(6)	0.8726	0.9363	0.9045
T(7)	0.9424	0.9535	1.0000
r^2 (cPTM)	0.9172	0.9055	0.9466
r^2 (uPTM)	0.9406	0.9255	0.9466
r^2 (cLAM)	0.4623	0.4329	0.5031
r^2 (uLAM)	0.7115	0.4686	0.5234
F_1 (7, 24)	1.3506	0.9204	0.0000
F_2 (2, 24)	46.283**	73.595**	95.101**
F_3 (2, 31)	85.156**	77.516**	128.73**

[a] $F_1(7, 24)$ is the test for significant improvement of uPTM over cPTM; $F_2(2, 24)$ is the test for significant improvement of uPTM over uLAM; $F_3(2, 31)$ is the test for significant improvement of cPTM over cLAM. ** indicates $p < 0.00001$.

3. Fitting Full Psychometric Functions

An alternative approach to testing the LAM and PTM models evaluates their ability to directly predict the full psychometric functions.[91] Again, for all three observers, the uPTM did not provide a significantly better fit to the data than the cPTM [$\chi^2(7) = 3.0, 11.0, 1.4$; $p > 0.85, 0.14, 0.95$ for observers AT, QL, and SM] did, indicating that the observers were using statistically indistinguishable templates in the low-pass and high-pass conditions. In all cases, the uPTM provided a much better fit than uLAM [$\chi^2(2) = 120, 142, 56$; $p < 0.00001$ for all three observers] did. Similarly, the cLAM provided a significantly worse fit to the data [$\chi^2(2) = 131, 86, 77$; $p < 0.00001$ for all three observers] than cPTM did, reaffirming that transducer nonlinearity and multiplicative noise are necessary in accounting for the data. To reiterate, the cPTM provided the best fit to all the data in this alternative method of testing that accounts for full psychometric functions.

The best-fitting spatial-frequency characteristics of the perceptual template derived from the full psychometric functions is plotted in the third row of Fig. 4. The parameters of the best-fitting cPTM in the full-psychometric-function approach corresponded closely to those estimated from the TVF functions at three performance criterion levels. One reason for the slight difference between the two sets of estimates is the different weights placed by the two procedures on the variability in the data in different portions of the psychometric functions.

4. CONCLUSIONS AND DISCUSSION

In a previous article,[58] we described an observer model (the PTM) and its application to the characterization of perceptual inefficiencies through use of white external noise. We examined the noisy LAM,[50–53] widely used to interpret data from equivalent internal-noise experiments, by testing its prediction that the ratio between two thresholds at each given external-noise contrast should be equal to the ratio of the corresponding d' levels for all observers and noise levels. We demonstrated in two experiments that this prediction failed. Direct fits of the LAM were also relatively poor, and the observed psychometric functions were incompatible with the required cumulative Gaussian form. We introduced the PTM as an elaboration of the LAM, with two additional components: a nonlinear transducer function and multiplicative noise. The PTM provides a good account of the data. It accommodates the observed threshold ratios while still predicting the equivalence of threshold ratios over external-noise levels. The PTM fit the three threshold data and yielded a good direct fit to the full psychometric functions.

In this article, we extend the application of the PTM to characterize the spatial-frequency response of the percep-

tual template as well as the perceptual inefficiencies of the human observer. We explicitly derived the functional relationship between the signal threshold required to achieve a certain level of performance in filtered external noise and the frequency characteristics of the perceptual template. Numerical procedures were developed to apply this functional relationship to estimate the perceptual template from experimental data, either by fitting TVF functions or by fitting the full psychometric functions. Application to an experiment allowed us to directly compare the PTM and the LAM statistically. The PTM provided much better accounts of the data. LAM, on the other hand, provided relatively imprecise characterizations of the spatial-frequency tuning of the perceptual template.

We rejected both the uPTM and the uLAM in favor of a cPTM in this study. This suggested that off-frequency looking[48,86–88] did not play a significant role in our experiments. This might be an adaptive response to the randomized experimental procedure. Namely, every experimental trial was drawn randomly from either the low-pass-filtered or high-pass-filtered external-noise family with equal probability. This experimental structure may have discouraged the observers from the active use of an off-frequency-looking strategy. Losada and Muller[48] found that at high-noise spectral densities masking by notched noise was greater than the summed masking of high- and low-pass noise when different noise conditions (high-pass, low-pass, and notched) were run in separate blocks, indicating the presence of off-frequency looking. At low-noise spectral densities, Losada and Muller[48] did not find any evidence of off-frequency looking. Our procedure is probably closer to the notched-noise method, except that the high-pass and low-pass noises were mixed across trials. It would be interesting to study off-frequency looking with the method developed here with use of procedures that encourage off-frequency looking.

The estimated shape and bandwidth of the spatial-frequency tuning of the perceptual templates in this study were in general comparable to those reported in the literature.[6,17,44,48] For example, Henning[44] reported log attenuations of $2.68\log(f/f_0)$ on the low-frequency side of the perceptual template and of $-1.62\log(f/f_0)$ on the high-frequency side of the perceptual template, compared with the $4.0\log(f/f_0)$ and $-1.1\log(f/f_0)$ estimated in this study. The estimated bandwidth of perceptual templates for sine-wave gratings is generally approximately 1–2 octaves.[6,17,44,48] Our estimated bandwidth of ~1.6 octaves falls right in the middle. Although purely empirical measures of bandwidth may be reasonable approximations, we suggest that estimates based on formal quantitative model fits of the PTM to either multiple-threshold levels or full psychometric functions will provide more-precise estimates and a framework for statistical evaluation that will be extremely useful in discriminating current theoretical claims in areas of perception.

The PTM we consider consists of a signal path that is essentially a single channel. The internal noise and nonlinearity in the model could implicitly reflect activities in multiple channels. In our applications, this form of model may be adequate because of the nature of the stimuli. Extensions to multiple-channel information and decision integration may be necessary for other situations.

Recent interest in understanding the mechanisms of object recognition,[92] attention, and perceptual learning[57–59] require the precise characterization of the perceptual template. In our own work, we have demonstrated that attention[71,93] and perceptual learning[59,94] can improve human performance through the exclusion of external noise in many conditions. Although external-noise exclusion might be achieved in a number of different ways, an important question for both psychology and neurophysiology is whether attention and/or perceptual learning can alter the spatial-frequency characteristics of perceptual templates.[95,96] The method described here has an obvious application in answering these questions.

ACKNOWLEDGMENT

This research is supported by the U.S. Air Force Office of Scientific Research, Life Sciences, Visual Information Processing Program.

Address correspondence to Zhong-Lin Lu, Department of Psychology, SGM 501, University of Southern California, Los Angeles, California 90089-101 (e-mail, zhonglin@rcf.usc.edu), or Barbara Anne Dosher, Department of Cognitive Sciences, 3151 SSPA, University of California, Irvine, California 92697 (e-mail, bdosher@uci.edu).

REFERENCES AND NOTES

1. D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's striate cortex," J. Physiol. **160**, 106–154 (1962).
2. H. R. Blackwell, "Neural theories of simple visual discriminations," J. Opt. Soc. Am. **53**, 129–160 (1963).
3. C. Enroth-Cugell and J. G. Robson, "The contrast sensitivity of retinal ganglion cells of the cat," J. Physiol. **258**, 517–552 (1966).
4. F. W. Campbell and J. G. Robson, "Application of Fourier analysis to the visibility of gratings," J. Physiol. **197**, 551–566 (1968).
5. M. B. Sachs, J. Nachmias, and J. G. Robson, "Spatial-frequency channels in human vision," J. Opt. Soc. Am. **61**, 1176–1186 (1971).
6. C. F. Stromeyer and B. Julesz, "Spatial-frequency masking in vision: critical bands and spread of masking," J. Opt. Soc. Am. **64**, 1221–1232 (1972).
7. P. H. Schiller, B. L. Finlay, and S. F. Volman, "Quantitative studies of single cell properties in monkey striate cortex. III. Spatial-frequency," J. Neurophysiol. **39**, 1334–1351 (1976).
8. A. B. Watson and J. G. Robson, "Discrimination at threshold: labelled detectors in human vision," Vision Res. **21**, 1115–1122 (1981).
9. R. L. DeValois, D. G. Albrecht, and L. G. Thorell, "Spatial-frequency selectivity of cells in macaque visual cortex," Vision Res. **22**, 545–559 (1982).
10. N. V. S. Graham, *Visual Pattern Analyzers* (Oxford U. Press, New York, 1989).
11. C. H. Graham, R. H. Brown, and F. A. Mote, "The relation of size stimulus and intensity in the human eye. I. Intensity threshold for white light," J. Exp. Psychol. **24**, 554–573 (1939).
12. J. P. Thomas, "Model of the function of receptive fields in human vision," Psychol. Rev. **77**, 121–134 (1977).

13. P. E. King-Smith and J. J. Kulikowski, "The detection of gratings by independent activation of line detectors," J. Physiol. **247**, 237–271 (1975).
14. N. V. S. Graham, "Visual detection of aperiodic spatial stimuli by probability summation among narrowband channels," Vision Res. **17**, 637–652 (1977).
15. H. R. Wilson and J. R. Bergen, "A four mechanism model for threshold spatial vision," Vision Res. **19**, 19–32 (1979).
16. G. E. Legge and J. M. Foley, "Contrast masking in human vision," J. Opt. Soc. Am. **70**, 1458–1471 (1980).
17. G. B. Henning, B. G. Hertz, and J. L. Hinton, "Effects of different hypothetical detection mechanisms on the shape of spatial-frequency filters inferred from masking experiments: I. Noise masks," J. Opt. Soc. Am. **71**, 574–581 (1981).
18. H. R. Wilson, D. K. McFarlane, and G. C. Phillips, "Spatial-frequency tuning of orientation selective units estimated by oblique masking," Vision Res. **23**, 873–882 (1983).
19. A. Pantle and R. Sekuler, "Size detecting mechanisms in human vision," Science **162**, 1146–1148 (1969).
20. C. B. Blakemore and F. W. Campbell, "On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images," J. Physiol. **203**, 237–260 (1969).
21. C. B. Blakemore and J. Nachmias, "The orientation specificity of two visual after-effects," J. Physiol. **213**, 157–174 (1971).
22. A. S. Gilinsky, "Orientation-specific effects of patterns of adapting light on visual acuity," J. Opt. Soc. Am. **58**, 13–18 (1968).
23. L. Olzak and J. P. Thomas, "Gratings: why frequency discrimination is sometimes better than detection," J. Opt. Soc. Am. **71**, 64–70 (1981).
24. J. Nachmias and A. Weber, "Discrimination of simple and complex gratings," Vision Res. **15**, 217–223 (1975).
25. D. J. Tolhurst and R. S. Dealey, "The detection and identification of lines and edges," Vision Res. **15**, 1367–1372 (1975).
26. J. P. Thomas and J. Gille, "Bandwidths of orientation channels in human vision," J. Opt. Soc. Am. **69**, 652–660 (1979).
27. J. P. Thomas, J. Gille, and R. Barker, "Simultaneous detection and identification: theory and data," J. Opt. Soc. Am. **72**, 1642–1651 (1982).
28. B. L. Beard and A. J. Ahumada, Jr., "Technique to extract relevant image features for visual tasks," in *Human Vision and Electronic Imaging III*, B. E. Rogowitz and T. N. Pappas, eds., Proc. SPIE **3299**, 79–85 (1998).
29. B. L. Beard and A. J. Ahumada, Jr., "Detection in fixed and random noise in foveal and parafoveal vision explained by template learning," J. Opt. Soc. Am. A **16**, 755–763 (1999).
30. J. A. Solomon and M. J. Morgan, "Reverse correlation reveals psychophysical receptive fields," Invest. Ophthalmol. Visual Sci. **40**, S572 (1999).
31. H. Fletcher, "Auditory patterns," Rev. Mod. Phys. **12**, 47–65 (1940).
32. H. B. Barlow, "Incremental thresholds at low intensities considered as signal/noise discrimination," J. Physiol. **136**, 469–488 (1957).
33. J. A. Swets, D. M. Green, and W. P. Tanner, Jr., "On the width of critical bands," J. Acoust. Soc. Am. **34**, 108–113 (1962).
34. U. Greis and R. Rohler, "Untersuchung der subjektiven Detailerkennbarkeit mit Hilfe der Ortsfrequenzfilterung," Opt. Acta **17**, 515–526 (1970).
35. H. Pollehn and H. Roehrig, "Effect of noise on the MTF of the visual channel," J. Opt. Soc. Am. **60**, 842–848 (1970).
36. B. E. Carter and G. B. Henning, "The detection of gratings in narrow-band visual noise," J. Physiol. **219**, 355–365 (1971).
37. A. P. Ginsburg, "Psychological correlates of a model of the human visual system," IEEE Trans. Aerosp. Electron. Syst. **71-C-AES**, 283–290 (1971).
38. L. D. Harmon and B. Julesz, "Masking in visual recognition: effects of two-dimensional filtered noise," Science **180**, 1194–1197 (1973).
39. A. P. Ginsburg, "Visual information processing based on spatial filters constrained by biological data," Ph.D dissertation (University of Cambridge, Cambridge, UK, 1978), Library of Congress 79-600156.
40. A. P. Ginsburg and D. W. Evans, "Predicting visual illusions from filtered images based on biological data," J. Opt. Soc. Am. **69**, 1443 (1979) (abstract).
41. M. C. Morrone, D. C. Burr, and J. Ross, "Added noise restores recognizability of coarse quantized images," Nature **305**, 226–228 (1983).
42. G. E. Legge, D. G. Pelli, G. S. Rubin, and M. M. Schleske, "Psychophysics of reading: I. Normal vision," Vision Res. **25**, 239–252 (1985).
43. M. Pavel, G. Sperling, T. Riedl, and A. Vanderbeek, "Limits of visual communication: the effect of signal-to-noise ratio on the intelligibility of American Sign Language," J. Opt. Soc. Am. A **4**, 2355–2365 (1987).
44. G. B. Henning, "Spatial-frequency tuning as a function of temporal frequency and stimulus motion," J. Opt. Soc. Am. A **5**, 1362–1373 (1988).
45. T. R. Riedl and G. Sperling, "Spatial-frequency bands in complex visual stimuli: American Sign Language," J. Opt. Soc. Am. A **5**, 606–616 (1988).
46. D. H. Parish and G. Sperling, "Object spatial frequencies, retinal spatial frequencies, noise, and the efficiency of letter discrimination," Vision Res. **31**, 1399–1415 (1991).
47. J. A. Solomon and D. G. Pelli, "The visual channel that mediates letter identification," Nature **369**, 395–397 (1994).
48. M. A. Losada and K. T. Muller, "Color and luminance spatial tuning estimated by noise masking in the absence of off-frequency looking," J. Opt. Soc. Am. A **12**, 250–260 (1995).
49. J. Nachmias, "Signal detection theory and its application to problems in vision," in *Handbook of Sensory Physiology*, D. Jameson and L. M. Hurvich, eds. (Springer–Verlag, Berlin, 1972), Vol. 7/4, Chap. 8.
50. N. S. Nagaraja, "Effect of luminance noise on contrast thresholds," J. Opt. Soc. Am. **54**, 950–955 (1964).
51. D. G. Pelli, "Effects of visual noise," Ph.D. dissertation (University of Cambridge, Cambridge, UK, 1981).
52. A. E. Burgess, R. F. Wagner, R. J. Jennings, and H. B. Barlow, "Efficiency of human visual signal discrimination," Science **214**, 93–94 (1981).
53. A. J. Ahumada and A. B. Watson, "Equivalent-noise model for contrast detection and discrimination," J. Opt. Soc. Am. A **2**, 1133–1139 (1985).
54. D. G. Pelli, "Uncertainty explains many aspects of visual contrast detection and discrimination," J. Opt. Soc. Am. A **2**, 1508–1532 (1985).
55. A. E. Burgess and B. Colborne, "Visual signal detection: IV. Observer inconsistency," J. Opt. Soc. Am. A **5**, 617–627 (1988).
56. M. P. Eckstein, A. J. Ahumada, and A. B. Watson, "Visual signal detection in structured backgrounds: II. Effects of contrast gain control, background variations, and white noise," J. Opt. Soc. Am. A **14**, 2406–2419 (1997).
57. Z.-L. Lu and B. A. Dosher, "External noise distinguishes mechanisms of attention," Vision Res. **38**, 1183–1198 (1998).
58. Z.-L. Lu and B. A. Dosher, "Characterizing human perceptual inefficiencies with equivalent internal noise," J. Opt. Soc. Am. A **16**, 764–778 (1999).
59. B. A. Dosher and Z.-L. Lu, "Mechanisms of perceptual learning," Vision Res. **39**, 3197–3221 (1999).
60. A. E. Burgess, "Visual signal detection: III. On Bayesian use of prior knowledge and cross correlation," J. Opt. Soc. Am. A **2**, 1498–1507 (1985).
61. J. Nachmias and R. V. Sansbury, "Grating contrast: Discrimination may be better than detection," Vision Res. **14**, 1039–1042 (1974).
62. J. M. Foley and G. E. Legge, "Contrast detection and near-threshold discrimination in human vision," Vision Res. **21**, 1041–1053 (1981).
63. D. J. Heeger, "Normalization of cell responses in cat striate cortex," Visual Neurosci. **9**, 181–197 (1992).

64. D. H. Hubel and T. N. Wiesel, "Uniformity of monkey striate cortex: a parallel relationship between field size, scatter, and magnification factor," J. Comp. Neurol. **158**, 295–306 (1974).
65. S. P. Mckee and G. Westheimer, "Improvement in vernier acuity with practice," Percept. Psychophys. **24**, 258–262 (1978).
66. N. Graham, "Spatial frequency channels in human vision: detecting edges without edge detectors," in *Visual Coding and Adaptability*, C. S. Harris, ed. (Erlbaum, Hillsdale, N. J., 1980), pp. 215–262.
67. M. A. Webster and R. L. de Valois, "Relationship between spatial-frequency and orientation tuning of striate-cortex cells," J. Opt. Soc. Am. A **2**, 1124–1132 (1985).
68. Alternatively, paradigms such as a partial report (Ref. 69) or a concurrent report (Refs. 70 and 71) may be used to eliminate structural uncertainty.
69. G. Sperling, "The information available in brief visual presentations," Psychol. Monogr. **11**, 1–74 (1960).
70. G. Sperling and B. Dosher, "Strategy and optimization in human information processing," in *Handbook of Perception and Performance*, K. Boff, L. Kaufmon, and J. Thomas, eds. (Wiley, New York, 1986), Vol. 1, Chap. 2, pp. 1–65.
71. B. A. Dosher and Z.-L. Lu, "Noise exclusion in spatial attention," Psychol. Sci. **11**, 139–146 (2000).
72. B. A. Dosher and Z.-L. Lu, "Mechanisms of perceptual attention in precuing of location," Vision Res. **40**, 1269–1292 (2000).
73. N. A. MacMillan and C. D. Creelman, *Detection Theory: A User's Guide* (Cambridge U. Press, New York, 1991).
74. The assumption holds well for intermediate spatial and temporal frequencies. At low spatial frequencies, low temporal frequencies reduce contrast sensitivity; at high spatial frequencies, high temporal frequencies reduce contrast sensitivity [J. G. Robson, "Spatial and temporal contrast-sensitivity functions of the visual system," J. Opt. Soc. Am. **56**, 1141–1142 (1966); D. H. Kelly, "Flickering patterns and lateral inhibition," J. Opt. Soc. Am. **59**, 1361–1370 (1969).] Henning[44] concluded that counterphase flicker at and below 10 Hz has no effect on the shape of spatial-frequency tuning below 4 c/deg, provided that both the masker and the signal have the same temporal characteristics.
75. This approximation becomes exact if (i) $T_s(f_x, f_y) = T_s(f)$ and $S_s(f_x, f_y) = S_s(f)$, that is, the template and the signal are radially symmetric in Fourier space; or (ii) $T_s(f_x, f_y) = kS_s(f_x, f_y)$, $\forall f_x, f_y$; that is, the template is perfectly matched to the signal stimulus; or (iii) either the template or the signal stimulus (or both) are uniform for every radius where the template and the signal overlap in Fourier space. In the current application, condition (iii) holds because the stimulus is a pair of points in Fourier space, and condition (ii) is approximately true because humans tend to use nearly optimal templates in simple stimulus situations.[60] If one or more of these conditions holds approximately, then the equations should provide reasonable approximations.
76. This approximation becomes exact if (i) $T_s(f_x, f_y) = T_s(f)$ and $F(f_x, f_y) = F(f)$; that is, the template and the experimenter-applied filter are radially symmetric in Fourier space; or (ii) $T_s(f_x, f_y) = kF(f_x, f_y)$, $\forall f_x, f_y$; that is, the template is matched to the noise filter; or (iii) either the template or the noise filter (or both) are uniform for every radius where the template and the noise filter overlap in Fourier space. In the current application, $F(f_x, f_y) = F(f)$ = constant, the expected spectrum of the Gaussian noise is uniform, and condition (iii) is met. In most applications it will be possible to construct filters such that condition (iii) holds.
77. In the current development, cross products in the form $(\beta^2 c^2 + N_{ext}^2)^\gamma$ are eliminated in order to yield analytical solutions. The effects of the cross terms have been evaluated in two of our previous publications. In one study,[59] PTMs with full cross-product forms were fit to the data by methods of iterative solution. The results were equivalent in pattern to those from fits of PTMs without cross products, and the cross-product terms were small. In the other study,[72] the analytical PTMs without cross products were compared with full stochastic PTMs. The analytical form was found to be a good approximation of the stochastic model.
78. W. L. Hays, *Statistics*, 3rd ed. (CBS College Publishing, New York, 1981).
79. R. S. Woodworth and H. Schlosberg, *Experimental Psychology*, 2nd ed. (Holt, Rinehart & Winston, New York, 1954).
80. D. G. Pelli and L. Zhang, "Accurate control of contrast on microcomputer displays," Vision Res. **31**, 1337–1350 (1991).
81. Z.-L. Lu and G. Sperling, "Second-order reversed phi," Percept. Psychophys. **61**, 1075–1088 (1999).
82. For an excellent discussion on fitting psychometric functions, see F. A. Wichmann and N. J. Hill, "The psychometric function I: fitting, sampling and goodness-of-fit," Percept. Psychophys. accepted for publication.
83. A resampling method (Refs. 84, 85) was used to compute the standard deviation of each threshold. We assumed that the number of correct responses at each signal contrast level on every psychometric function has a binomial distribution with a single-even probability p, which is the measured percent correct. We then generated a theoretically resampled psychometric function for a given condition by independently replacing the number of correct responses at each signal stimulus contrast on the psychometric function with a sample from the corresponding binomial distribution. Repeating this process 2000 times, we generated 2000 theoretically resampled psychometric functions in every external-noise condition. We estimated the standard deviation for each threshold by fitting Weibull to these theoretically resampled psychometric functions and computing the standard deviation of the 2000 resampled thresholds for each external-noise condition.
84. L. T. Maloney, "Confidence intervals for the parameters of psychometric functions," Percept. Psychophys. **47**, 127–134 (1990).
85. F. A. Wichmann and N. J. Hill, "The psychometric function II: bootstrap based confidence intervals and sampling" Percept. Psychophys. (to be published).
86. R. D. Patterson, "Auditory filter shapes derived with noise stimuli," J. Acoust. Soc. Am. **59**, 640–654 (1976).
87. D. G. Pelli, "Channel properties revealed by noise masking," Invest. Ophthalmol. Visual Sci. **19**, 44A (1980).
88. M. E. Perkins and M. S. Landy, "Nonadditivity of masking by narrow-band noises," Vision Res. **31**, 1053–1065 (1991).
89. Implemented in Matlab 5.3, the procedure for a given model consisted of the following. (1) For a given set of the model parameters, using Eq. (21) to compute $\log(c_\tau^{\text{theory}})$ from the model for each external-noise condition at three different performance levels. (2) Computing the squared difference between the log threshold prediction from the model and the observed $\text{sqdiff} = [\log(c_\tau^{\text{theory}}) - \log(c_\tau)]^2$ for each threshold. The log approximately equates the standard error over large ranges in contrast thresholds, corresponding to weighted least squares, an equivalent to the maximum likelihood solution for continuous data. In the current data set, this assumption is true. (3) Computing L: summation of sqdiff from all the thresholds across all the external-noise conditions. (4) Using a gradient-descent method to adjust the model parameters to find the minimum L. (5) After obtaining the minimum L, computing the r^2 statistic to evaluate the goodness of the model fit:

$$r^2 = 1.0 - \frac{\sum [\log(c_\tau^{\text{theory}}) - \log(c_\tau)]^2}{\sum \{\log(c_\tau) - \text{mean}[\log(c_\tau)]\}^2}, \qquad (25)$$

where Σ and mean( ) run over all the thresholds for a particular observer. An F test for nested models was used to statistically compare the four models. An F is defined:

$$F(df_1, df_2) = [(r_{\text{full}}^2 - r_{\text{reduced}}^2)/df_1]/[(1 - r_{\text{full}}^2)/df^2], \qquad (26)$$

where $df_1 = k_{\text{full}} - k_{\text{reduced}}$, and $df_2 = N - k_{\text{full}}$. The k variables are the number of parameters in each model, and N is the number of predicted data points.

90. B. A. Dosher and Z.-L. Lu, "Perceptual templates in spatial attention," Invest. Ophthalmol. Visual Sci. **41**, S750 (2000).
91. Full sets of maximum-likelihood fits were performed on the psychometric functions for all four models: uPTM, cPTM, uLAM, and cLAM. For an observer who is correct in K_{ij} trials among a total of N_{ij} trials in the jth signal contrast and ith filter condition, the likelihood of a model that predicts a fraction of P_{ij} correct in each condition is defined as

$$\text{likelihood} = \prod_{i=1}^{I} \prod_{j=1}^{J} \frac{N_{ij}!}{K_{ij}!(N_{ij} - K_{ij})!} P_{ij}^{K_{ij}} (1 - P_{ij})^{N_{ij} - K_{ij}}, \tag{27}$$

where P_{ij} is defined by Eq. (22). Asymptotically, $\chi^2(df)$ statistics could be used to compare the proper set of models:

$$\chi^2(df) = 2.0 \times \log\left(\frac{\text{likelihood}_{\text{full}}}{\text{likelihood}_{\text{reduced}}}\right), \tag{28}$$

where df is the difference of the number of parameters between the full and the reduced models.

92. D. G. Pelli, "Close encounters—an artist shows that size affects shape," Science **285**, 844–846 (1999).
93. Z.-L. Lu and B. A. Dosher, "Spatial attention: different mechanisms for central and peripheral temporal precues?" J. Exp. Psychol. **26**, 1534–1548 (2000).
94. B. A. Dosher and Z.-L. Lu, "Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting," in Proc. Natl. Acad. Sci. USA **95**, 13 988–13 993 .
95. Y. Yeshurun, and M. Carrasco, "Attention improves or impairs visual performance by enhancing spatial resolution," Nature **396**, 72–75 (1998).
96. Z.-L. Lu and B. A. Dosher, "Attention fine-tunes perceptual templates in spatial cuing," Bull. Psychonom. Soc. **40**, 52 (1999).