

# Spatial summation of peripheral Gabor patches

Velitchko Manahilov, William A. Simpson, and Daphne L. McCulloch

*Vision Sciences Department, Glasgow Caledonian University, City Campus, Cowcaddens Road, Glasgow G4 0BA, Scotland*

Received October 15, 1999; revised manuscript received May 31, 2000; accepted September 7, 2000

Previous studies have specified the foveal pattern that is seen most efficiently, with the assumption that the waveform of the best pattern matches the impulse response of the most sensitive visual filter. We measured the threshold contrast for circular, collinear, and orthogonal Gabor stimuli of 6 Hz temporal frequency presented 7 deg above the fixation point. We found that the threshold contrast energy is minimal for a class of stimuli whose Fourier-spectra bandwidth is less than  $\sim 1$  octave. These findings suggest that an energy algorithm might underlie spatial summation of peripheral Gabor patches. The different behavior of spatial summation in fovea and periphery might reflect the differences in pattern detectability across space in the central and peripheral visual fields. It is also possible that a coherent (cross-correlation) algorithm is employed in detection of foveal stimuli and that an incoherent (energy) algorithm is employed in detection of peripheral stimuli. © 2001 Optical Society of America

OCIS codes: 330.7310, 330.1880, 330.5000, 330.1800, 330.6110, 330.4060.

## 1. INTRODUCTION

There have been numerous studies on how the contrast for detection of a stimulus depends on the stimulus area. In most of the classical work,<sup>1,2</sup> the stimuli were circular disks against a background of uniform luminance. In the past few decades, motivated by the hypothesis that spatial-frequency-selective filters underlie early visual processing,<sup>3</sup> researchers have also used sinusoidal gratings to study spatial summation.<sup>4–8</sup> Some studies have been aimed at specifying the pattern that is seen most efficiently, assuming that the waveform of the best pattern matches the impulse response of the most sensitive filter of early visual processing.<sup>6–8</sup>

The efficiency of a sensory system has been defined in the framework of the ideal-observer concept.<sup>9–11</sup> The ideal observer refers to the best possible performance in detecting signals under specific conditions. Suppose that the observer knows the signal waveform exactly and that the signal is embedded in uncorrelated noise. In such a case the most optimal detection algorithm is cross correlation of the expected signal with the presented noise or signal plus noise.<sup>10,12</sup> The statistical efficiency ( $F$ ) for detection of a signal known exactly at a fixed performance level is defined as the ratio of the signal energy required by an ideal observer ( $E_{id}$ ) to the signal energy required by a human observer under study ( $E_o$ )<sup>9</sup>:

$$F = E_{id}/E_o. \quad (1)$$

According to Eq. (1) the efficiency for detecting any visual signal on a fixed background is inversely proportional to the threshold stimulus energy. Let us assume that the observer consists only of a limited set of filters whose impulse responses have a fixed spatial extent. Such an observer will detect a stimulus most efficiently only when the stimulus waveform matches the impulse response of the most sensitive visual filter. With this approach, the best foveal stimulus (detected with lowest stimulus energy) was found to be a 7-cycles/degree ( $c$ /

deg) sinusoidal grating that was modulated by a circular Gaussian function with width and height of  $\sim 3$  grating cycles and drifting at 4 Hz.<sup>6</sup> For discriminating the motion direction, the best pattern was established to be a Gabor patch of 3  $c$ /deg and 5 Hz, with width and height of 0.44 deg and duration of 0.133 s.<sup>13</sup>

Kersten<sup>7</sup> studied spatial summation of one-dimensional foveal Gabor patches in the presence and absence of dynamic white noise. He found that the efficiency for gratings in noise was quite high (8–30%) for a stimulus width less than 1 cycle (between  $1/e$  points) and decreased for wider stimuli. He suggested that this result resembles the performance of a (mismatched) cross correlator whose impulse response matches the waveform of the most efficiently detected stimulus. However, when patterns were presented in the absence of noise, a single cross correlator could not explain the slow decline of efficiency as the grating width was extended beyond one cycle.

Recently, Polat and Tyler<sup>8</sup> studied spatial summation of foveal circular and elongated Gabor patches with a static vertical-grating carrier and with patch envelope orientation in either the vertical (collinear) or horizontal (orthogonal) direction. The log threshold contrast was plotted as a function of the log number of cycles in the patch elongation. The stimulus size at which the slope of this function reached  $-0.5$  was regarded as representing the size of the most efficient filter in the set.<sup>7</sup> Maximal efficiency was obtained for a collinear elongated pattern of spatial frequency of 4  $c$ /deg, length of approximately four grating cycles, and width of one cycle.

Robson and Graham<sup>5</sup> measured spatial summation of static grating patches with short bars presented at different positions in the visual field. They found that the contrast sensitivity was maximal at the fixation point and declined steadily with increasing the distance from the fixation point. When the stimuli were displayed 42 grating periods above the fixation point, the contrast sensitiv-

ity along a horizontal line was approximately constant. The data obtained were described satisfactorily by assuming probability summation between responding filters in different spatial locations and by taking into account the variations in sensitivity across the visual field.

These findings have shown that the location of the stimuli in the visual field is an important factor that might affect the detection of extended patterns. The studies attempting to find the most efficiently detected stimulus have used stimuli presented in fovea.<sup>6-8</sup> A question arises of what is the most efficiently detected stimulus presented in the visual-field periphery, where the contrast sensitivity across space is essentially constant. To this end, we measured the threshold contrast for detection of spatiotemporal Gabor patches of various spatial configurations (circular, orthogonal, and collinear), which were presented 7 deg above the fixation point. We established that the threshold contrast energy was minimal not for a unique stimulus but for a class of peripheral stimuli. We compared the data obtained with a mismatched cross-correlation model<sup>7</sup> and a probability-summation model.<sup>5</sup> Furthermore, we used an energy model to describe the present data. The energy algorithm is the most optimal algorithm when the observers do not use phase information.<sup>10</sup> The main point of this type of model is the assumption that contrast detection is determined by the energy of the output of the linear filter(s) approximating the early visual stages. Energy models have been used to describe detection of temporal stimuli,<sup>14,15</sup> temporal summation of disks,<sup>14</sup> spatial summation of elliptical Gaussian blobs,<sup>16</sup> and detection of spatiotemporal contrast stimuli.<sup>17</sup> Our experimental data suggest that an energy algorithm might be used in the detection of peripheral Gabor patches. We now describe the models, whose performance we will compare with our experimental results.

## 2. MODELS

### A. Mismatched Cross-Correlation Model

The ideal observer has the best performance in a detection task under the assumptions that it knows *a priori* the signal parameters exactly and that the noise in which the signal is embedded is uncorrelated.<sup>9,10</sup> The ideal observer detects the stimulus when the cross correlation of the expected signal with the presented noise or signal plus noise reaches or exceeds a fixed criterion level. A requirement of the ideal-observer model is that the observer has access to a filter that matches the input waveform, whatever the spatiotemporal characteristics of the input may be. This assumption, however, is not in line with the contemporary understanding of the organization of early visual stages. In the present study we consider a more realistic version of this model (mismatched cross-correlation model), assuming that the visual system consists only of a limited set of filters whose impulse responses have a fixed spatial extent, phase, and spatial frequency. Such filters are optimal only for stimuli having the same spatial extents, phases, and spatial frequencies. Other stimuli, whose waveforms are mismatched to the existing filters, are detected less efficiently.

In the main experiment of this study we measured the contrast threshold for detection of a Gabor patch of 2 c/deg and various spatial spreads. We compared the data obtained with the predictions of a mismatched cross-correlation model,<sup>7</sup> assuming that the mismatched cross correlator has one filter whose impulse response  $h$  is tuned to the spatial frequency  $v_y$  of the stimulus carrier and has a fixed circular spatial spread  $s_0$ :

$$h(x, y) = \cos(2\pi y v_y) \exp(-x^2/s_0^2 - y^2/s_0^2). \quad (2)$$

In this model we did not consider the time domain, because this dimension was not varied in the experiments reported. The response output  $R$  of the mismatched crosscorrelator is given by the cross correlation of the filter impulse response with the stimulus waveform,

$$R = C \iint g(x, y)h(x, y) dx dy, \quad (3)$$

where  $C$  is the contrast of the stimulus and  $g$  describes the unit-amplitude spatial waveform of the stimulus.

At threshold, the mismatched crosscorrelator response is constant  $K$  and with Eq. (3) the threshold contrast may be written as

$$C = K \bigg/ \iint g(x, y)h(x, y) dx dy. \quad (4)$$

The contrast energy of a threshold stimulus in the  $xy$  domain may be written as

$$\begin{aligned} E_s &= C^2 \iint [g(x, y)]^2 dx dy \\ &= K^2 \iint [g(x, y)]^2 dx dy / \\ &\quad \left[ \iint g(x, y)h(x, y) dx dy \right]^2. \end{aligned} \quad (5)$$

Using oblique masking of stimuli presented at 8-deg eccentricity, Swanson and Wilson<sup>18</sup> have found that the spatial frequency bandwidth of the spatial channel centered on 2.8 c/deg is  $\sim 1$  octave. In the mismatched cross-correlation model we used filters whose spatial spread was 0.6 deg (width and height of 1.2 periods) or 1.2 deg (width and height of 2.4 periods). The spatial frequency bandwidth of these filters is 1.36 and 0.68 octaves, respectively, which is close to the bandwidth measured by Swanson and Wilson.<sup>18</sup> The carrier spatial frequency of these filters coincided with the spatial frequency of the stimuli used in the present study (2 c/deg).

The solid and open diamonds in Fig. 1 show the threshold contrast of a mismatched cross correlator calculated by Eq. (4) as a function of the area of circular Gabor patches by using filters whose width and height have 1.2 and 2.4 periods, respectively. This model predicts that the log-log slope changes smoothly from approximately  $-1$  to  $0$  with no extended region of  $-0.5$  slope. According to Eq. (5), the threshold stimulus energy has a minimum for the stimulus whose spatial waveform matches the impulse response of the assumed filter. The solid and open diamonds in Fig. 2 illustrate the threshold stimulus energy as predicted by the mismatched cross-correlation

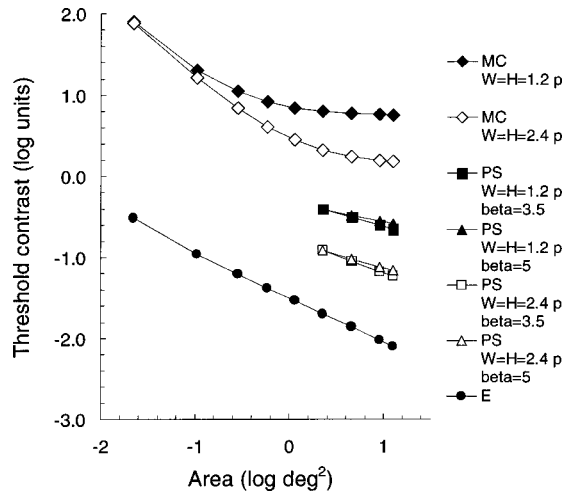


Fig. 1. Model predictions of threshold contrast for detection of circular Gabor stimuli as a function of stimulus area (discussed in the text). Diamonds, predictions of a mismatched cross-correlation model (MC); squares and triangles, predictions of a probability-summation model (PS) with  $\beta = 3.5$  and  $\beta = 5$ , respectively; circles, predictions of an energy model (E); solid symbols, model calculations with use of a filter whose impulse response has width and height of 1.2 cycles; open symbols, model calculations with use of a filter whose impulse response has width and height of 2.4 cycles. The predictions of the mismatched cross-correlation and probability-summation models are vertically shifted.

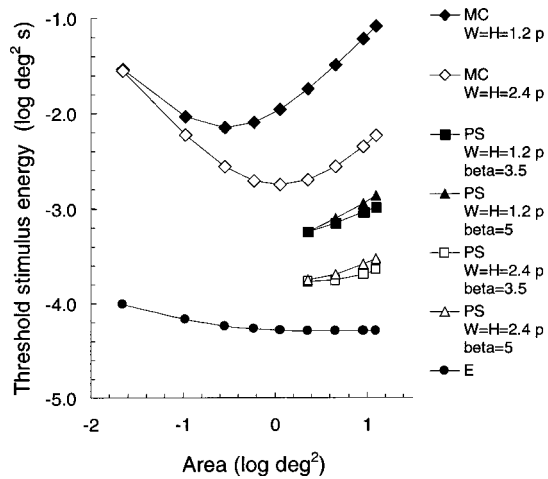


Fig. 2. Model predictions of threshold stimulus energy for detection of circular Gabor stimuli as a function of stimulus area. The other designations are as in Fig. 1.

model. It should be noted that these model predictions can also be calculated analytically by using the equations published by Kersten.<sup>7</sup>

### B. Probability Summation Model

Improvement in detectability of a grating as the stimulus area increases has been regarded as resulting from probability summation across space of spatially distributed filters that are spatial frequency narrowly tuned.<sup>5</sup> We used a simplified version of this model having a single filter whose impulse response is given by expression (2) in order to describe the spatial summation of Gabor patches of 2 c/deg. If noise fluctuations are independent from point to point, the probability of seeing the stimulus  $P$

may be expressed by the probability of guessing  $\gamma$  and the probability that one sample in space of the noisy response  $P_i$  has exceeded the criterion level at least once in the following equation:

$$P = 1 - (1 - \gamma) \prod_i (1 - P_i). \quad (6)$$

The probability that filter  $i$  will signal the occurrence of a stimulus is related to stimulus contrast  $C$  by the following equation<sup>5,19</sup>:

$$P_i = 1 - \exp[-|Cg(x, y) * h(x, y)|^\beta], \quad (7)$$

where  $*$  means convolution of the unit-amplitude stimulus waveform  $g$  and the impulse response  $h$  and the exponent  $\beta$  describes the steepness of the psychometric function.

When Eqs. (6) and (7) are combined, the probability of seeing a stimulus may be written as

$$P = 1 - (1 - \gamma) \exp\left[-\int \int |Cg(x, y) * h(x, y)|^\beta dx dy\right]. \quad (8)$$

Assuming that the observer makes a correct forced-choice response when at least one detection occurs or when the observer guesses correctly, at some fixed probability for stimulus detection we have

$$C^\beta \int \int |g(x, y) * h(x, y)|^\beta dx dy = L, \quad (9)$$

where  $L$  is a constant.

At threshold, the detection contrast may be expressed as

$$C = \left[ L / \int \int |h(x, y) * g(x, y)|^\beta dx dy \right]^{1/\beta}. \quad (10)$$

With Eq. (10), the threshold stimulus energy might be written as

$$E_s = L^{2/\beta} \int \int [g(x, y)]^2 dx dy / \left[ \int \int |h(x, y) * g(x, y)|^\beta dx dy \right]^{2/\beta}. \quad (11)$$

We calculated the predictions of the probability summation model by assuming that the impulse response of the filter in Eq. (2) had a 2-c/deg carrier and that its width and height consisted of 1.2 or 2.4 cycles. We performed the calculations for extended stimuli (width and height within the range of 3.4–8 grating cycles), because for smaller stimuli the model should also take into account the probability summation across filters sensitive to different spatial frequencies. Figure 1 shows the threshold contrast predicted by Eq. (10) as a function of stimulus area using  $\beta = 3.5$  (squares) and  $\beta = 5$  (triangles). These values of the exponent  $\beta$  are consistent with various studies<sup>20–22</sup> that have shown that the estimates of the exponent  $\beta$  are within the range of 2.5–7. The solid squares and triangles denote the calculations with a spatial filter whose width and height consist of 1.2 periods, and the open squares and triangles illustrate the calculations with a spatial filter whose width and height consist of 2.4 periods.

With a filter whose width and height have 1.2 periods, the log–log slope of these functions approximates  $-0.32$  and  $-0.22$  for  $\beta = 3.5$  and  $\beta = 5$ , respectively. When the filter has width and height of 2.4 periods, the log–log slope is  $-0.42$  for  $\beta = 3.5$  and  $-0.33$  for  $\beta = 5$ . Robson and Graham<sup>5</sup> have suggested that this slope is equal to  $-1/\beta$ , that is,  $-0.28$  for  $\beta = 3.5$  and  $-0.2$  for  $\beta = 5$ . We found that the slope depends on both the spatial spread of the filter impulse response and the exponent  $\beta$ . The predicted threshold stimulus energies (Fig. 2, squares and triangles) increase as the stimulus area increases.

### C. Energy Model

The data obtained in the present study were also analyzed by an energy model for contrast detection.<sup>17</sup> This model has three main stages. First, a single spatiotemporal linear filter transforms the retinal image into a neural representation. Second, the filter output ( $F$ ) may be expressed in the frequency domain by convolution (multiplication) of the Fourier spectrum of the stimulus waveform ( $S$ ) with the transfer function of the linear filter ( $H$ ):

$$F(f_x, f_y, f_t) = S(f_x, f_y, f_t) * H(f_x, f_y, f_t). \quad (12)$$

In the third stage, the filter is followed by an energy device that computes the energy of the response ( $E_r$ ) to a stimulus. In the frequency domain this quantity may be calculated by integrating the squared filter output over spatial and temporal frequencies:

$$E_r = \iiint |S(f_x, f_y, f_t)H(f_x, f_y, f_t)|^2 df_x df_y df_t. \quad (13)$$

Denoting the unit-amplitude waveform of the stimulus with  $G$  and the stimulus contrast with  $C$ , we may rewrite Eq. (13) as follows:

$$E_r = C^2 \iiint |G(f_x, f_y, f_t)H(f_x, f_y, f_t)|^2 df_x df_y df_t. \quad (14)$$

It should be noted that this model assumes that the energy device integrates the squared neural response over space and time within a limited temporal epoch. Previous studies have shown that the temporal integration epoch is within 0.2–0.5 s.<sup>14,15,17</sup> In the present study we used flickering Gabor patches, modulated in time by a Gaussian function, whose duration was 0.32 s, and they were within the range of the estimated values of the temporal integration epoch. It is reasonable to suggest that the integration in space is also not infinite. However, we do not have estimates of the limits of spatial integration and have assumed that the energy device integrates the response energy in an area equal to that covering the area of the largest stimulus used in the present study.

A threshold device signals the presence of the stimulus when the response energy reaches or exceeds a fixed criterion level. We assumed that the intrinsic noise distribution was such that the probability of stimulus detection could be approximated by a Weibull function of the response energy,

$$P = 1 - (1 - \gamma)\exp(-E_r^\eta),$$

where  $\eta$  is a constant and  $\gamma$  is the probability of guessing.

At some fixed probability of stimulus detection, the energy of the response to the stimulus is a constant ( $E_0$ ). With Eq. (14), the threshold contrast might be expressed as

$$C = E_0^{1/2} \left/ \left[ \iiint |G(f_x, f_y, f_t)|^2 \times H(f_x, f_y, f_t)|^2 df_x df_y df_t \right]^{1/2} \right. . \quad (15)$$

This model will be used to calculate the threshold contrast for detection of stimuli of a fixed temporal frequency and duration. The magnitude of the transfer function at temporal frequency  $\nu_t$  is proportional to the contrast sensitivity function (CSF) measured at the same temporal frequency.<sup>17</sup> Assuming a radial symmetry of this function in the  $f_x f_y$  domain (similar contrast sensitivity functions to gratings of different orientation), we may write

$$|H(f_x, f_y, \nu_t)| = E_0^{1/2} \text{CSF}(f_x, f_y, \nu_t), \quad (16)$$

and substituting  $H$  in Eq. (15), the threshold contrast becomes

$$C = 1 \left/ \left[ \iiint |G(f_x, f_y, f_t)|^2 \times \text{CSF}(f_x, f_y, \nu_t)|^2 df_x df_y df_t \right]^{1/2} \right. . \quad (17)$$

The threshold stimulus energy is calculated in frequency domain as follows:

$$E_s = C^2 \iiint |G(f_s, f_y, f_t)|^2 df_x df_y df_t,$$

and substituting  $C$  with Eq. (17) we may write

$$E_s = \iiint |G(f_s, f_y, f_t)|^2 df_x df_y df_t \left/ \left[ \iiint |G(f_x, f_y, f_t)|^2 \text{CSF}(f_x, f_y, \nu_t)|^2 df_x df_y df_t \right] \right. . \quad (18)$$

Note that both threshold contrast and threshold stimulus energy have no free parameters when the CSF is known.

The circles in Fig. 1 illustrate the predictions of the threshold contrast by Eq. (17) as a function of stimulus area. The CSF was approximated by a two-dimensional parabola of the form suggested by Watson,<sup>23</sup> the parabola's maximum being at 2  $c/\text{deg}$ :

$$\text{CSF} = 1.8 - 2.5\{\log[(f_x^2 + f_y^2)^{1/2}] - \log(2)\}^2.$$

The spatial-summation function decreases as the stimulus area increases, and the log–log slope approaches  $-0.5$ , which means that the threshold contrast is inversely proportional to the square root of the stimulus area. The threshold stimulus energy (Fig. 2, circles) predicted by Eq. (18) decreases at small stimuli and approaches a lower limit at more extended stimuli.

### 3. METHODS

#### A. Observers

Two of the authors (W.S. and V.M.) and three undergraduate students from the Department of Vision Science at Glasgow Caledonian University took part in the experiments. The students were unfamiliar with the aims of the experiment. All subjects had no practice with elongated stimuli.

#### B. Apparatus

The stimuli were generated with a Dell PC computer ( $1024 \times 768$  pixels, 75 Hz) with use of 256 gray levels and 12-bit luminance resolution.<sup>24</sup> The monitor had a mean luminance of 30 candelas per square meter ( $\text{cd/m}^2$ ) and was surrounded by a large screen ( $100 \times 100$  cm) illuminated to approximate the monitor in luminance and hue. The observers viewed the stimulation field binocularly through the natural pupil from a chin and forehead rest.

#### C. Stimuli

We used three types of stimuli. The circular and collinear stimuli were horizontal cosine gratings (Figs. 3A and 3B) enveloped in space and time by Gaussian functions whose unit-contrast spatiotemporal waveform may be written

$$g(x, y, t) = \cos(2\pi y v_y) \exp[-(x - x_c)^2/s_x^2 - (y - y_c)^2/s_y^2] \cos(2\pi t v_t) \times \exp[-(t - 2.5s_t)^2/s_t^2], \quad (19)$$

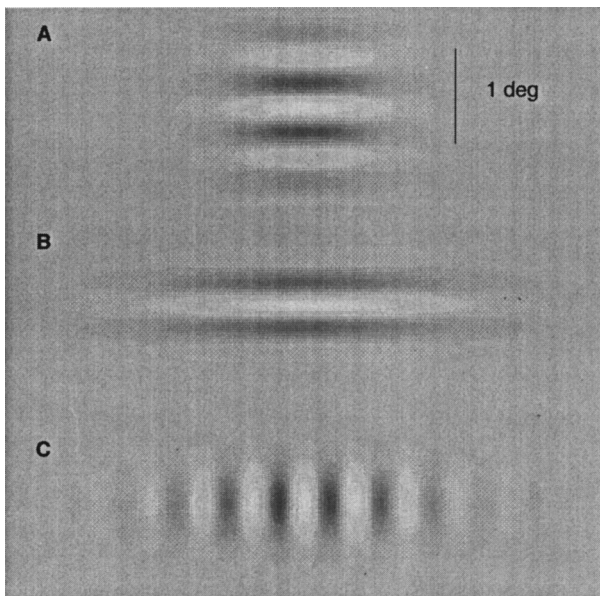


Fig. 3. Examples of static Gabor stimuli used in the present experiments. The stimuli are cosine gratings of 2 c/deg enveloped by a two-dimensional Gaussian function. A, circular Gabor patch with width and height of 1.2 deg (2.4 grating cycles); B, collinear Gabor patch, horizontal grating with width of 2.4 deg (4.8 grating cycles) and height of 0.6 deg (1.2 grating cycles); C, orthogonal Gabor patch, vertical grating, with width and height as in B. The width and height of the stimuli are defined as twice the relevant Gaussian half-width. The stimuli shown have equal area.

where  $v_y$  is the spatial frequency of the horizontal grating,  $v_t$  is the temporal frequency and  $s_x$ ,  $s_y$ , and  $s_t$  are the spatial and temporal Gaussian half-widths, respectively. The midpoint  $(x_c, y_c)$  of the stimulus was centered 7 deg above the fixation point. The width, height, and duration of a stimulus are defined as twice the relevant half-width at  $1/e$  amplitude.<sup>6</sup> The carrier of the orthogonal Gabor patch (Fig. 3C) was a vertical cosine grating, and the patch elongation was in the horizontal direction.

#### D. Procedure

Contrast thresholds for stimulus detection were measured by a two-interval forced-choice adaptive staircase procedure converging to 79% correct responses.<sup>25</sup> A static suprathreshold sample of the signal was presented 1 s before each block of trials to reduce stimulus uncertainty. The intervals were separated by 500 ms. Each interval was marked by an auditory tone. A feedback tone indicated to the subject if he was incorrect. Each trial began when the subject pressed an appropriate button; after 500 ms the stimulus was displayed. Two buttons enabled subjects to answer with "Stimulus was presented during the first interval" or "Stimulus was presented during the second interval." The computer recorded which button was pressed and automatically adjusted the step size for the next trial. After each staircase reversal, the step size was halved. This process started at suprathreshold contrast levels with a contrast step of 0.2 log units and continued until the step size became 0.05 log unit. The subsequent eight staircase reversals were collected as data points. The mean values of the threshold contrasts for every experimental condition were calculated by averaging data collected in four experimental sessions.

### 4. RESULTS

In preliminary experiments, we found that the contrast sensitivity (the reciprocal of the detection threshold) for a circular Gabor stimulus with width and height of 4 deg, and duration of 0.32 s was maximal for a spatial frequency of 2 c/deg and a temporal frequency of 6 Hz. We then used this temporal frequency to measure the spatial CSF. Figure 4 shows the contrast sensitivity for circular Gabor patches (width and height of 4 deg, duration of 0.32 s) as a function of the spatial frequency of the stimulus carrier. The data points represent the contrast sensitivity function for each observer tested. These functions have similar form and are only vertically shifted. The sensitivity for each observer is maximal for a stimulus of 2 c/deg.

We measured the contrast threshold for detection of circular and elongated Gabor patches having a spatial frequency of 2 c/deg and a temporal frequency of 6 Hz. The width and height of the circular Gabor patch varied in the range of 0.13 deg (0.26 grating cycles) to 4 deg (8 grating cycles). The height of the collinear and orthogonal Gabor stimuli was 0.6 deg (1.2 grating cycles), and the width varied in the range of 1.3 deg (2.6 grating cycles) to 6.6 deg (13.3 grating cycles).

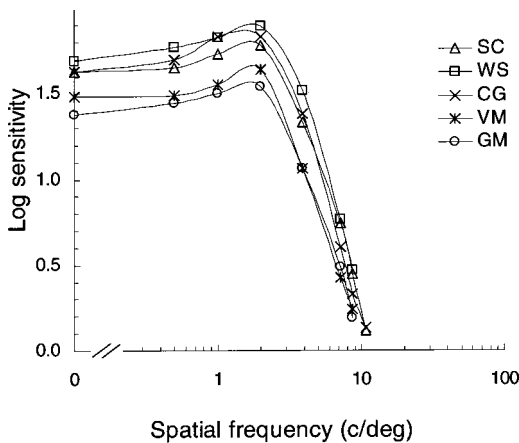


Fig. 4. Mean contrast sensitivity to circular Gabor stimuli as a function of carrier spatial frequency. Stimulus width and height were 4 deg, duration was 0.32 s, and temporal frequency was 6 Hz. The data are from the five observers tested.

It should be noted that the stimuli were described by even functions that have a dc (0 c/deg) component for collinear Gabor patches of 0.6-deg height. This dc component is only 5.7% of the maximal (2-c/deg) component of collinear patches with width of 6.6 deg. The sensitivity to Gabor patches of zero spatial frequency is lower (~0.2 log units) than the sensitivity to 2-c/deg stimuli (see Fig. 4). Therefore, one may assume that the dc component of the collinear Gabor patches would have a negligible effect on the threshold for detection of these stimuli.

**A. Threshold Contrast for Gabor Patches as a Function of Stimulus Area**

Let us analyze the data for elongated Gabor patches. Figure 5 shows for all subjects tested the means of four estimations of the threshold contrast for collinear (gray circles) and orthogonal (black circles) patches as a function of stimulus area. Using a paired *t* test for each subject, we found that the threshold contrasts for the patterns of equal area were not significantly different at 95% confidence level. We fitted the data points shown in Fig. 5 by a linear function (solid lines in Fig. 5) by using the least-squares method. The fitted log-log slopes of the threshold improvement for both collinear and orthogonal configurations were -0.5. The mean values and 95% confidence intervals across the subjects were  $-0.51 \pm 0.03$  for collinear patterns and  $-0.5 \pm 0.05$  for orthogonal patterns.

Figure 6 represents, for each subject, the means of four estimations of the threshold contrast for the three types of stimuli: circular (open circles), collinear (gray circles), and orthogonal (black circles). In the mismatched cross-correlation and probability summation models we approximated the impulse response of the linear filter by a circular function [Eq. (2)] whose width and height consisted of 1.2 or 2.4 periods. The predictions of each model [Eq. (4) and Eq. (10)] are proportional to a constant [*K* in Eq. (4) and *L* in Eq. (10)] that was used to fit the model calculations with the data points for circular Gabor patches by using the least-squares method. For each subject we calculated, with Eq. (4), the ratio between the component of variance due to the lack of fit and the com-

ponent of variance of the experimental error and compared it with the *F* value needed to reject the adequacy of fitting at 95% confidence level.<sup>26</sup> We found that the predictions of the mismatched cross-correlation model using filters with width and height of 1.2 or 2.4 periods (diamonds in Fig. 1) were not in agreement with the data obtained. Using the same filters and  $\beta = 3.5$  and  $\beta = 5$ , we established also that the predictions of the probability-summation model for extended stimuli (triangles and squares in Fig. 1) differed significantly at 95% confidence level from the measured threshold contrasts.

The solid curves in Fig. 6 show the threshold contrast for circular and elongated Gabor patches as a function of stimulus area, calculated by the energy model [Eq. (10)].

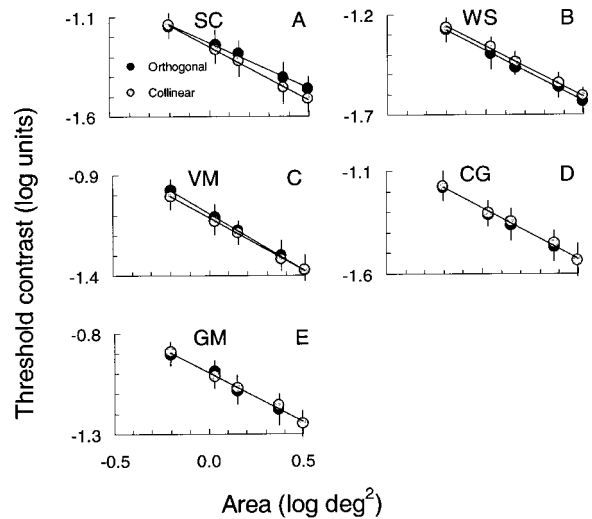


Fig. 5. Mean threshold contrast for detection of collinear (gray circles) and orthogonal (black circles) Gabor stimuli as a function of stimulus area. All stimuli had a height of 0.6 deg. Vertical bars, 95% confidence intervals.

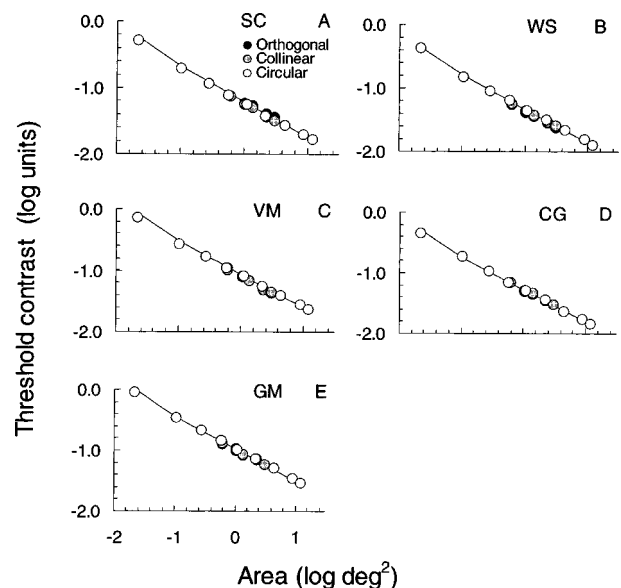


Fig. 6. Mean threshold contrast for detection of circular (open circles) Gabor stimuli as a function of stimulus area. Gray and black circles, data for collinear and orthogonal Gabor stimuli, respectively, for each observer as in Fig. 5; solid curves, predictions of the energy model as discussed in the text.

We approximated the transfer function of each observer by the corresponding CSF (Fig. 4). The analysis of variance showed that the energy model accounts satisfactorily at 95% confidence level for the spatial summation of peripheral Gabor patches of circular and elongated shapes.

We also analyzed the log–log slope of the experimental data for circular patches. It was found that the data points across the subjects tested were fitted with a linear function with a log–log slope of approximately  $-0.5$  ( $-0.52 \pm 0.04$ ) if the analysis did not include the smallest stimuli, which had width and height of 0.26 and 0.73 grating cycle. These findings indicate that our stimuli, excluding the smallest ones, are detected with equal efficiency.<sup>7,8</sup>

## B. Threshold Energy of Gabor Patches as a Function of Stimulus Area

We found that as the area of the Gabor patch increases, regardless of the stimulus shape, the threshold stimulus energy approaches a lower asymptote (Fig. 7). According to the mismatched cross-correlation model, the threshold stimulus energy is minimal for the stimulus whose spatial waveform matches the impulse response of the assumed filter (Fig. 2). We compared the predictions of the mismatched cross-correlation model [Eq. (5)] with the data obtained by means of the analysis of variance and found that they differed significantly at 95% confidence level. The probability-summation model [Eq. (11)] also was not able to account for the data obtained.

The solid curves in Fig. 7 show the energy-model predictions for threshold stimulus energy as a function of stimulus area [Eq. (18)]. These predictions were in agreement at 95% confidence level with the data obtained.

We analyzed these data in the spatial frequency domain. Watson<sup>27</sup> was the first to use a similar approach to model temporal summation. At threshold, the amplitude spectra of the responses of a linear temporal filter to stimuli of short duration are identical, which means that the response energies are constant. Watson has shown that the reciprocity between threshold duration and intensity (Bloch's law) is a property of any linear filter that passes temporal frequencies below some cutoff. Consider the  $f_y$  power spectra of some stimuli used in the present experiments and their responses integrated over the  $f_x f_t$  domain (Fig. 8). We calculated the stimulus power spectra ( $P_s$ ) as follows:

$$P_s(f_y) = C^2 \iint |G(f_x, f_y, f_t)|^2 df_x df_t, \quad (20)$$

The response power spectrum ( $P_r$ ) can be expressed as

$$P_r(f_y) = C^2 \iint |G(f_x, f_y, f_t)H(f_x, f_y, f_t)|^2 df_x df_t, \quad (21)$$

and with Eq. (16), it becomes

$$P_r(f_y) = E_0 C^2 \iint |G(f_x, f_y, f_t)CSF(f_x, f_y, v_t)|^2 df_x df_t. \quad (22)$$

The coefficient of proportionality  $E_0$  was set to an arbitrary value of  $1.5 \times 10^{-4}$ .

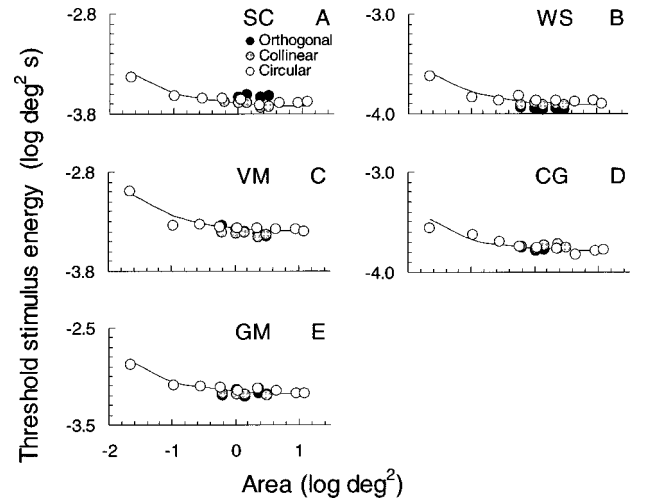


Fig. 7. Mean threshold stimulus energy for detection of circular, collinear, and orthogonal Gabor stimuli as a function of stimulus area. The other designations are as in Fig. 6.

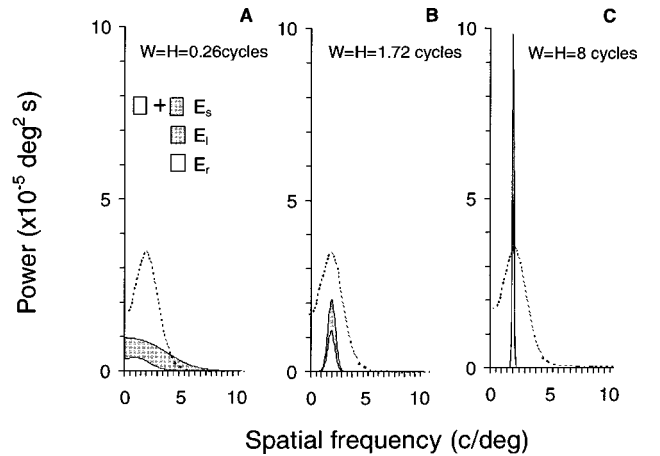


Fig. 8. Stimulus (outlines of the gray areas) and response (outlines of the white areas) spatial frequency power spectra for three threshold circular Gabor stimuli calculated by Eqs. (20) and (22), respectively. Both stimulus width and stimulus height had 0.26 (A), 1.72 (B), and 8 grating cycles (C). Dotted curves are the power spectra of the transfer function in relative units.

The outlines of the white areas in Fig. 8 represent the power spectra of the responses to threshold circular Gabor stimuli with width and height of 0.13 deg (0.26 grating cycles), 0.86 deg (1.7 grating cycles), and 4 deg (8 grating cycles). The response energies (the integral in the frequency domain of the stimulus power spectrum) are similar despite the different shapes of the power spectra. The stimulus energy (the sum of the white and gray areas) equals the sum of the response energy (white area) and the lost energy ( $E_l$ , gray area):

$$E_s = E_r + E_l. \quad (23)$$

According to this balance of energy, the threshold stimulus energy is proportional to the lost energy with the accuracy of an additive constant ( $E_r$ ). It should be noted that the proposed energy balance [Eq. (23)] is based on a simplification of the early visual stages if we regard them as a passive spatiotemporal filter. This could be useful in describing the transformation of threshold im-

ages in early vision. The power spectrum of the smallest Gabor patch (Fig. 8A) is wide and consists of components at higher spatial frequencies, where the sensitivity of the visual system is low. Much of the stimulus energy is lost, and therefore the detection efficiency is lower. For comparison, the dotted curves in Fig. 8 represent the power of the transfer function in arbitrary units. Gabor stimuli whose width and height consist of 1.7 and 8 grating cycles (Figs. 8B and C) have narrow spatial-frequency spectra located within the spatial frequency range, where the visual system is most sensitive. These stimuli are transformed by the visual system most efficiently with minimal lost energy (compare the gray areas in Fig. 8).

As mentioned in Section 1, the detectability across space in the central and peripheral parts of the visual field is inhomogeneous.<sup>5</sup> We presented the stimuli 7 deg above the fixation point, where the pattern sensitivity is approximately constant. Thus the energy model assumed that the contrast sensitivity across space is a constant. However, spatial variations of contrast sensitivity might change the behavior of the threshold stimulus energy as a function of stimulus area. We modeled this situation by assuming that the spatial waveform of the stimuli is modulated by a Gaussian function:

$$W(x, y) = \exp[-(x^2 - y^2)/s^2] \quad (24)$$

A similar function was used by Quick *et al.*<sup>28</sup> to approximate the decrease in contrast sensitivity with eccentricity. With a 2-deg effective width of  $W(x, y)$ , they satisfactorily described the summation of two foveal gratings having close spatial frequencies. Figure 9 shows some model calculations of threshold energy of extended stimuli in which the stimulus waveform was multiplied by  $W(x, y)$ . The solid curves represent the predictions for  $s = \infty$  (constant sensitivity across space),  $s = 4$  deg,  $s = 2$  deg, and  $s = 1$  deg. These model calculations show that when the variations of the contrast sensitivity across space are accounted for, an energy model can mimic the predictions of the mismatched cross-correlation and probability-summation models for the spatial-summation function (compare squares and triangles in Fig. 9 and Fig. 2). It should be noted that it is difficult to estimate the

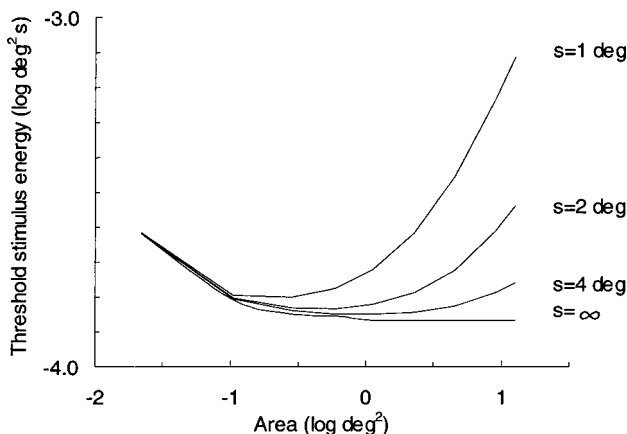


Fig. 9. Effect of the variations in contrast sensitivity in space on the threshold energy of Gabor patches, calculated by the energy model. The spatial waveform of the stimuli was multiplied by a two-dimensional Gaussian whose spatial constant  $s$  was  $\infty$  (constant sensitivity across space), 4, 2, and 1 deg.

true form of sensitivity decline with distance from the foveal center. This is because the local measurements of sensitivity to gratings should be done with spatially localized stimuli, which would have too wide a bandwidth to be considered as a single sinusoidal harmonic. Robson and Graham<sup>5</sup> and Pointer and Hess<sup>29</sup> suggested that sensitivity to gratings of spatial frequencies above 1 c/deg declines  $\sim 0.025$  log unit per grating period. This estimate corresponds approximately to the model prediction for  $s = 4$  deg. These model calculations suggest that the differences in detectability across space in the central and peripheral part of visual field might explain, at least in part, the differences between spatial summation of foveal and peripheral Gabor patches.

## 5. DISCUSSION

In the present study we looked for the most effectively detected stimulus among peripheral Gabor patches having a temporal frequency of 6 Hz and different spatial configurations (circular, collinear, and orthogonal). We found that the threshold contrasts for circular and elongated Gabor patches did not differ significantly. The threshold contrast of stimuli whose width and height consisted of more than 1.2 cycles decreased as a function of stimulus area with a log-log slope of  $-0.5$ . In other words, the threshold contrast for these stimuli was proportional to the square root of the stimulus area. The threshold contrast energy was maximal for the smallest circular Gabor patches and approached a lower limit for more extended stimuli. The mismatched cross-correlation and probability-summation models failed to account for the data obtained. It should be noted that an ideal observer with filters that exactly match all stimuli used in the present study will exhibit a square-root relation between detection threshold and stimulus area.<sup>7</sup> However, an observer with an infinite number of available filters is unrealistic. Moreover, we found a square-root relation for most of the stimuli used but not for very-small-sized stimuli.

We explained the data obtained with an energy model for contrast detection, assuming that the early visual stages consist of a single filter. The two-dimensional filter transfer function was approximated by the CSF measured experimentally, assuming a radial symmetry of this function in the  $f_x f_y$  domain. It is known that in central vision, grating resolution is better for horizontal and vertical than for oblique orientations. At eccentricities larger than  $\sim 20$  deg, resolution is best for meridionally oriented gratings and worst for gratings perpendicular to visual-field meridians. However, at eccentricities of 5–10 deg, resolution becomes similar for all orientations.<sup>30</sup> Similar results were reported by measuring orientation identification of suprathreshold gratings.<sup>31</sup> In the present study the stimuli were centered 7 deg above the fixation point, and therefore one could assume that the contrast sensitivity functions for gratings of different orientations are similar. However, precautions should be taken into account if a similar approach is used for stimuli presented in the fovea or the far periphery.

Inefficiency of the visual system is usually attributed to insufficient sampling and filtering, intrinsic noise of the system, and signal parameter uncertainty.<sup>32</sup> In the framework of the energy model, the filtering properties of the early visual filter might be regarded as a source of detection inefficiency. The power spectrum of the stimulus with the smallest area (Fig. 8A) consists of components at higher spatial frequencies where the visual system is less sensitive, which results in much of the stimulus energy being lost by the linear filter. Stimuli whose width and height consist of more than one grating cycle have narrow power spectra that coincide with the maximum of the transfer function of the system. For these stimuli, much of the stimulus energy passes through the linear filter, and the lost energy is minimal. Thus, the visual filter operates most efficiently for the class of peripheral stimuli whose Fourier spectrum is narrow (bandwidth smaller than 1.37 octaves) with a maximum coinciding with the maximum of the filter's transfer function.

Our results differ from the data obtained in studies on spatial summation of gratings presented in the fovea.<sup>6–8</sup> Some of these studies showed that threshold contrast improved for collinear Gabor patches significantly more than for orthogonal and circular patches.<sup>8</sup> Other studies found that threshold stimulus energy was minimal for a circular stimulus whose waveform consisted of  $\sim 2$ – $3$  cycles.<sup>6</sup> These differences might be due, at least in part, to the different detectability across space in the central and peripheral parts of the visual field. Model calculations have shown that when the variations in sensitivity across the visual field are taken into account, the threshold energy of spatially extended stimuli can resemble the predictions of the mismatched cross-correlation and probability-summation models (Fig. 9).

Spatial summation of circular Gabor patches has also been measured with stimuli presented in the nasal visual field on a background of luminance of  $762 \text{ cd/m}^2$ .<sup>33</sup> It was found that the slope of contrast sensitivity as a function of the number of grating cycles was 1 over a moderate range of grating cycles and decreased beyond it. If these results are expressed as the detection threshold as a function of stimulus area in log-log units, the slope of the summation functions will be  $-0.5$  up to a certain stimulus area corresponding to the energy-model predictions. With approximation of the data presented in their study, the energy model could explain the summation functions for  $2\text{-c/deg}$  Gabor patches presented at  $7\text{-deg}$  eccentricity up to  $\sim 2.5$  grating cycles. Savage and Banks<sup>34</sup> studied spatial summation in scotopic conditions (background luminance of approximately  $-1 \text{ log}$  scotopic td) for square patches of gratings presented at  $20 \text{ deg}$  in the nasal visual field. Their data agreed with the energy-model predictions for stimuli of  $2 \text{ c/deg}$  that have up to  $\sim 12$  grating cycles. In the present study, the stimuli were presented on a  $30\text{-cd/m}^2$  background, and the energy model was able to explain satisfactorily the spatial summation functions up to 8 grating cycles (the largest Gabor patch used). The comparison of these studies suggests that the stimulus size at which the energy model is valid increases as the background luminance decreases. This suggestion, however, needs to be verified in further experiments with

the same subjects and with experimental procedures that use different background luminancies.

It should be noted that human observers could use different sorts of information in contrast detection and could apply different detection algorithms.<sup>10</sup> For coherent detection, when the observers use phase information, the optimal algorithm is cross correlation. On the other hand, if the observers do not have phase information, the incoherent energy algorithm would detect the stimuli most optimally. In our experiments the observers knew exactly the stimulus waveform they had to detect. However, the spatial-summation data of peripheral stimuli were explained by the energy algorithm, which means that the observers likely did not use phase information in contrast detection. One might suggest that the optimal strategy in contrast detection depends on the location of the stimuli. When the stimuli are presented in the central part of the visual field, the optimal algorithm could be cross correlation. The fovea is specialized for processing of fine patterns, and observers may be able to monitor phase information effectively. On the other hand, the observers could use the energy algorithm for detection of stimuli displayed outside of the foveal area. Some studies have shown that discrimination of compound patterns is poorer in the periphery than in the fovea.<sup>35,36</sup> The observer's performance in this task depends on the relative phase of stimulus components, which indicates that phase-information processing might be limited in the periphery. This suggestion needs additional studies applying the approach used by Burgess and Ghandeharian,<sup>37</sup> i.e., exploring the effects of uncertainty about the phase, size, and shape of the stimulus's spatial waveform on detection of noisy peripheral stimuli.

## ACKNOWLEDGMENTS

This work was supported by Centre of Excellence Grant from the Glasgow Caledonian University to the Department of Vision Sciences. We thank G. N. Dutton for his helpful comments on an earlier version of this paper. We are grateful to the anonymous referees for the valuable suggestions and comments on a previous version of the manuscript.

## REFERENCES

1. C. H. Graham and R. Margaria, "Area and the intensity time relation in the peripheral retina," *Am. J. Physiol.* **113**, 299–305 (1935).
2. H. B. Barlow, "Temporal and spatial summation in human vision at different background intensities," *J. Physiol. (London)* **141**, 337–350 (1958).
3. F. W. Campbell and J. G. Robson, "Application of Fourier analysis to the visibility of gratings," *J. Physiol. (London)* **197**, 551–566 (1968).
4. E. R. Howell and R. F. Hess, "The functional area for summation to threshold for sinusoidal gratings," *Vision Res.* **18**, 369–374 (1978).
5. J. G. Robson and N. Graham, "Probability summation and regional variation in contrast sensitivity across the visual field," *Vision Res.* **21**, 409–418 (1981).
6. A. B. Watson, H. B. Barlow, and J. G. Robson, "What does the eye see best?" *Nature* **302**, 419–422 (1983).
7. D. Kersten, "Spatial summation in visual noise," *Vision Res.* **24**, 1977–1990 (1984).

8. U. Polat and C. W. Tyler, "What pattern the eye sees best," *Vision Res.* **39**, 887–895 (1999).
9. W. P. Tanner and T. G. Birdsall, "Definition of *d'* and *h* as psychophysical measures," *J. Acoust. Soc. Am.* **30**, 922–928 (1958).
10. D. B. Green and J. A. Swets, *Signal Detection Theory and Psychophysics* (Wiley, New York, 1974).
11. H. B. Barlow, "The efficiency of detecting changes of density in random dot patterns," *Vision Res.* **18**, 637–650 (1978).
12. A. E. Burgess, "The Rose model, revisited," *J. Opt. Soc. Am. A* **16**, 633–646 (1999).
13. A. B. Watson and K. Turano, "The optimal motion stimulus," *Vision Res.* **35**, 325–336 (1995).
14. C. Rashbass, "The visibility of transient changes of luminance," *J. Physiol. (London)* **210**, 165–186 (1970).
15. J. J. Koenderink and A. J. van Doorn, "Detectability of power fluctuations of temporal visual noise," *Vision Res.* **18**, 191–195 (1978).
16. P. Bijl and J. J. Koenderink, "Visibility of elliptical Gaussian blobs," *Vision Res.* **33**, 243–255 (1993).
17. V. Manahilov and W. Simpson, "Energy model for contrast detection: spatiotemporal characteristics of threshold vision," *Biol. Cybern.* **81**, 61–71 (1999).
18. W. H. Swanson and H. R. Wilson, "Eccentricity dependence of contrast matching and oblique masking," *Vision Res.* **25**, 1285–1295 (1985).
19. R. F. Quick, Jr., "A vector-magnitude model of contrast detection," *Kybernetik* **16**, 65–67 (1974).
20. A. B. Watson, "Probability summation over time," *Vision Res.* **19**, 515–522 (1979).
21. A. B. Watson, "Summation of grating patches indicates many types of detector at one retinal location," *Vision Res.* **22**, 17–25 (1982).
22. N. Graham and J. G. Robson, "Summation of very close spatial frequencies: the importance of spatial probability summation," *Vision Res.* **27**, 1997–2007 (1987).
23. A. B. Watson, "Visual detection of spatial contrast patterns: evaluation of five simple models," *Opt. Express* **6**, 12–33 (2000), <http://epubs.osa.org/opticsexpress/topbiframe.htm>.
24. G. D. Pelli and L. Zhang, "Accurate control of contrast on microcomputer displays," *Vision Res.* **31**, 1337–1350 (1991).
25. H. Levitt, "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477 (1970).
26. V. Manahilov, "Triphasic temporal impulse responses and Mach bands in time," *Vision Res.* **38**, 447–458 (1998).
27. A. B. Watson, "Temporal sensitivity," in *Handbook of Perception and Human Performance I: Sensory Processes and Perception*, K. R. Boff, L. Kaufman, and J. P. Thomas, eds. (Wiley, New York, 1986), pp. 6.1–6.41.
28. R. F. Quick, Jr., W. W. Mullins, and T. A. Reichert, "Spatial summation effects on two-component grating thresholds," *J. Opt. Soc. Am.* **68**, 116–124 (1978).
29. J. S. Pointer and R. F. Hess, "The contrast sensitivity gradient across the human visual field with emphasis on the low spatial frequency range," *Vision Res.* **29**, 1133–1151 (1989).
30. J. Rovamo, V. Virsu, P. Laurinen, and L. Hyvarinen, "Resolution of gratings oriented along and across meridians in peripheral vision," *Invest. Ophthalmol. Visual Sci.* **23**, 666–670 (1982).
31. L. A. Temme, L. Malcus, and W. K. Noell, "Peripheral visual field is radially organized," *Am. J. Optom. Physiol. Opt.* **62**, 545–554 (1985).
32. A. E. Burgess, "High level visual decision efficiencies," in *Vision: Coddling and Efficiency*, C. B. Blakemore, ed. (Cambridge U. Press, New York, 1990), pp. 431–440.
33. M. S. Banks, A. B. Sekuler, and S. J. Anderson, "Peripheral spatial vision: limits imposed by optics, photoreceptors, and receptor pooling," *J. Opt. Soc. Am. A* **8**, 1775–1787 (1991).
34. G. L. Savage and M. S. Banks, "Scotopic visual efficiency: constraints by optics, receptor properties, and rod pooling," *Vision Res.* **32**, 645–656 (1992).
35. O. Braddick, "Is spatial phase degraded in peripheral vision and visual pathology?" in *Documenta Ophthalmologica Proceedings Series*, L. Maffei, ed. (W. Junk, The Hague, 1981), pp. 255–262.
36. H. S. Orbach and H. R. Wilson, "Factors limiting peripheral pattern discrimination," *Spatial Vis.* **12**, 83–106 (1999).
37. A. Burgess and H. Ghandeharian, "Visual signal detection. I. Ability to use phase information," *J. Opt. Soc. Am. A* **1**, 900–905 (1984).