# A standard model for foveal detection of spatial contrast

**Andrew B. Watson** NASA Ames Research Center, Moffett Field, CA, USA

**Albert J. Ahumada, Jr.** NASA Ames Research Center, Moffett Field, CA, USA

The ModelFest data set was created to provide a public source of data to test and calibrate models of foveal spatial contrast detection. It consists of contrast thresholds for 43 foveal achromatic contrast stimuli collected from each of 16 observers. We have fit these data with a variety of simple models that include one of several contrast sensitivity functions, an oblique effect, a spatial sensitivity aperture, spatial frequency channels, and nonlinear Minkowski summation. While we are able to identify one model, with particular parameters, as providing the lowest overall residual error, we also note that the differences among several good-fitting models are small. We find a strong reciprocity between the size of the spatial aperture and the value of the summation exponent: both are effective means of limiting the extent of spatial summation. The results demonstrate the power of simple models to account for the visibility of a wide variety of spatial stimuli and suggest that special mechanisms to deal with special classes of stimuli are not needed. But the results also illustrate the limited power of even this large data set to distinguish among similar competing models. We identify one model as a possible standard, suitable for simple theoretical and applied predictions.

Keywords: vision, spatial, pattern, detection, threshold, contrast, contrast sensitivity, model, ModelFest

## Introduction

### Models of spatial sensitivity

Spatial pattern is one of the primary effective elements of visual stimulation. Pattern vision begins with the ability to sense variations over space in the intensity of the light image. Development of models of this ability has therefore been, and continues to be, a goal of much of vision research. Early treatments of spatial sensitivity emphasized the role of summation within a fixed area, exemplified in such formulations as Ricco's Law (Graham & Margaria, 1935), and resolution, exemplified in acuity measurements (Shlaer, 1937). Introduction of the contrast sensitivity function (Campbell & Robson, 1968) lead to a somewhat more general model embodied in a spatial filter (Campbell, Carpenter, & Levinson, 1969), and later developments led to the idea of multiple spatial filters (Blakemore & Campbell, 1969). Separately, there have been advances in our understanding of how sensitivity varies with eccentricity (Robson & Graham, 1981), orientation (Berkley, Kitterle, & Watkins, 1975; Campbell, Kulikowski, & Levinson, 1966), and pattern size (Robson & Graham, 1981).

Many of the studies in this area have, with good reason, concentrated on a single dimension of stimulus variation. But it is desirable to have a model that is sufficiently general to accommodate variation in all of the relevant dimensions. Apart from the theoretical desire for generality, there are also important practical applications in which such a model would be useful.

One challenge for those seeking such a general model is the fact that much of the data to be modeled come from different labs, and the reports are frequently lacking in details that would allow combination of data across labs. This difficulty led to the creation of the ModelFest data set.

### ModelFest

The ModelFest experiment was a collaboration among several laboratories to collect a single set of common data for testing and calibration of contrast detection models. The ModelFest data set consists of a collection of contrast thresholds for 43 stimuli from 16 observers in 10 labs (Carney et al., 1999; Carney et al., 2000; Watson, 1999). In Phase 1 of that effort, extending through 1999, data were collected from nine observers. In Phase 2, data were collected from an additional seven observers.

### Previous analyses

Previously, one of us examined the fit of various models to the ModelFest Phase 1 data (Watson, 2000). The data were found to be consistent with a simple model composed of a contrast sensitivity filter (CSF) followed by Minkowski summation with an exponent of about 2.5. Augmenting the model with multiple frequency channels yielded a slightly improved fit and a higher exponent of about 3.8.

The ModelFest Phase 1 data have also been examined in a number of other reports. Chen & Tyler (2000) applied principal components analysis to derive the receptive

doi: 10.1167/5.9.6 Received February 18, 2005; published October 26, 2005 ISSN 1534-7362 

fields of putative detectors and arrived at three, which are the following: a spot detector, a bar detector, and a grating detector (Walker, Klein, & Carney, 1999). Carney et al. (2000) examined relationships among subsets of thresholds to address questions regarding spatial summation and mechanism bandwidths. None of these reports attempted to fit the entire data set with a single model.

### Present analyses

In this report, as in Watson (2000), we fit various models to the entire set of 43 thresholds. This paper extends the earlier report in the following ways. First, as noted above, additional data from seven new observers have been collected. Second, we have introduced and evaluated new elements to the model, notably an oblique effect and a spatial aperture. And lastly, in this report we consider a large number of specific functional forms for the CSF. The fits here provide a reasonably definitive evaluation of a number of candidate forms for the CSF.

Following Watson (2000), we have used a component model, consisting of a cascade of elements that may be introduced or removed and whose parameters may be fixed or allowed to vary. In the latter case, we create what are called nested models, with one being a more constrained version of the other. This component model allows some insight into which components are most crucial to accurate predictions, and more generally it indicates the relative contribution to accuracy of each component. The nested cases permit some simple statistical tests.

One result of these analyses is the specification of a standard model for foveal contrast detection. This model is not the best-fitting model of all we have considered, but it provides an excellent fit with very few assumptions, parameters, and calculations. We believe it may be useful in a variety of theoretical and applied contexts. We also believe it provides a valuable benchmark against which more complicated models may be compared.

## ModelFest experiment

### Stimuli

The ModelFest stimuli have been described elsewhere (Carney et al., 1999, 2000; Watson, 2000), but we provide a brief summary here. The stimuli, shown in Figure 1, and described in Table 1, consisted of 43 grayscale images, each 256 × 256 pixels in size. Each stimulus is identified by an index number between 1 and 43. A file containing all of the images is included as a supplement to this paper in the file modelfest-stimuli, which is described more completely in Appendix A.

Each pixel was represented by an eight-bit number between 1 and 255. The stimuli were rendered, using a variety of hardware and software techniques, so that pixel graylevel $g$ in the image was converted to luminance $L$ on the display according to the formula

$$L(g) = L_0\left(1 + \frac{c}{127}(g - 128)\right) \quad (1)$$

where $c$ is the contrast of the stimulus and $L_0$ is the mean luminance. In each lab, $L_0$ was fixed to a value in the range 30 ± 5 cd m$^{-2}$. The mathematical notation used in this paper is summarized in Appendix C.

The viewing distance was set so that each pixel subtended 1/120th of a degree, and the entire image subtended 256/120 = 2.133 degrees. Viewing was binocular with natural pupils.

In the time dimension, the stimulus followed a Gaussian time course with a standard deviation of 0.125 s. The display frame rate was at least 60 Hz.

The stimuli were presented at the center of an otherwise uniform screen whose luminance matched the mean luminance of the stimulus ($L_0$). Fixation guides were presented continuously in the form of "L"-shaped marks at the four corners of the stimulus image.

### Methods

Contrast detection thresholds for the 43 stimuli were collected for 16 observers in 10 labs. The labs differed somewhat in details of procedure, but all adhered to the following methods. Thresholds were measured using a two-interval forced-choice method with feedback. Each threshold was based on at least 32 trials, and measurement of each threshold was repeated at least four times.

### Data

To exclude any ambiguity regarding the data set we have analyzed and modeled, we define a "ModelFest Baseline Dataset." This consists of the first four thresholds reported for each of the 16 observers for each of the 43 stimuli. Each threshold has been expressed as $\log_{10}(c)$, where $c$ is contrast as defined in Equation 1. Each value has been rounded to three decimal places. This data set is provided as a supplement to this paper, as the text file modelfestbaselinedata.csv, described more completely in Appendix B.

### Descriptive statistics

Results in this paper are primarily expressed in decibels (dB = 20 $\log_{10}$ $c$). In those units, each threshold is $t_{s,o,r}$, where the indices refer to stimulus ($s$ = 1,…,$S$), observer

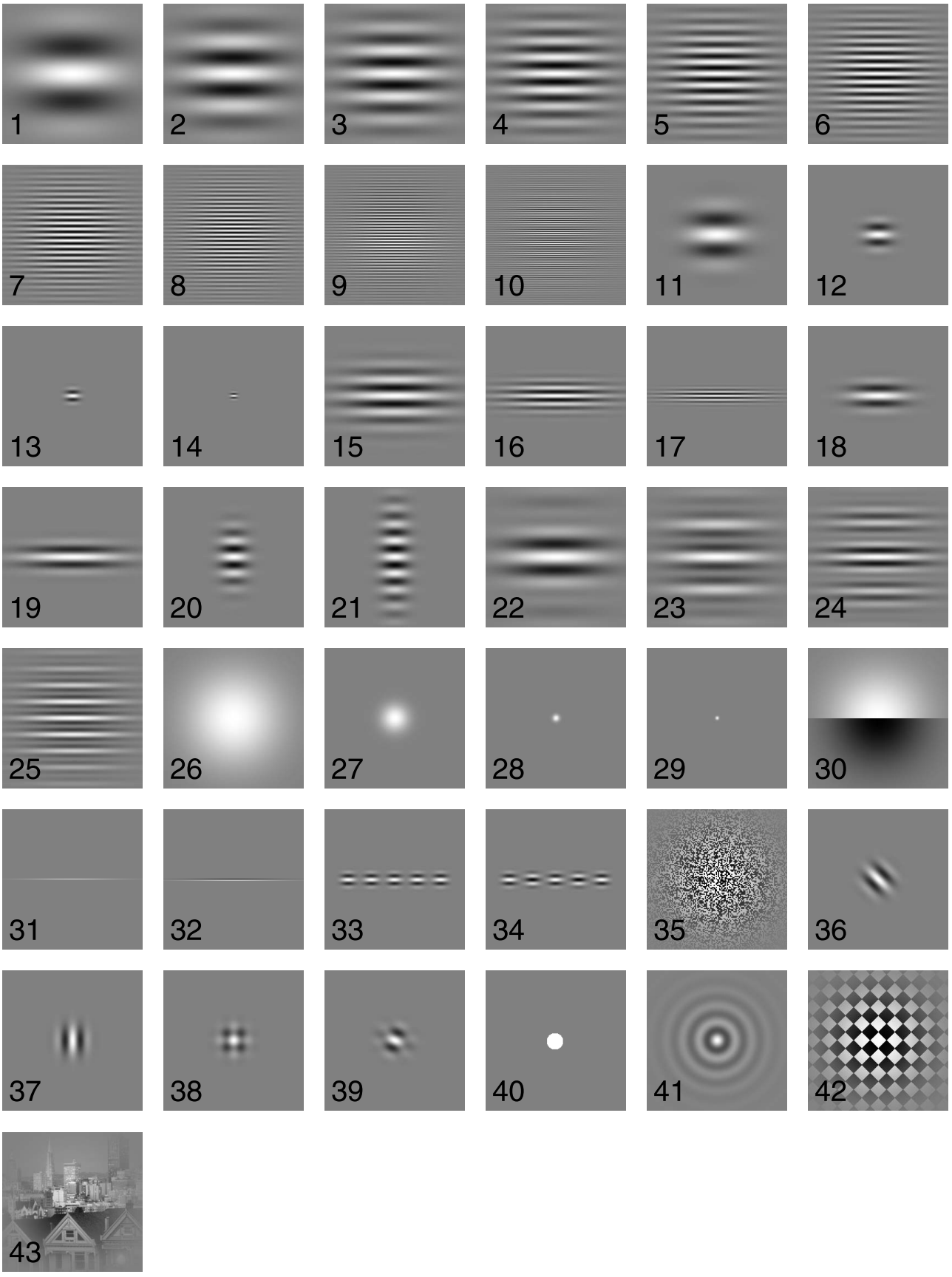


Figure 1. ModelFest stimuli. Each is a monochrome image subtending 2.133 × 2.133 degrees. The index numbers have been added for identification and were not present in the stimuli.

| Index | Type | Parameters |
|---|---|---|
| 1 | Gabor, fixed size | 1.12 cycles/degree |
| 2 | Gabor, fixed size | 2 cycles/degree |
| 3 | Gabor, fixed size | 2.83 cycles/degree |
| 4 | Gabor, fixed size | 4 cycles/degree |
| 5 | Gabor, fixed size | 5.66 cycles/degree |
| 6 | Gabor, fixed size | 8 cycles/degree |
| 7 | Gabor, fixed size | 11.3 cycles/degree |
| 8 | Gabor, fixed size | 16 cycles/degree |
| 9 | Gabor, fixed size | 22.6 cycles/degree |
| 10 | Gabor, fixed size | 30 cycles/degree |
| 11 | Gabor, fixed cycles | 2 cycles/degree, $b_x = b_y = 1$ octave |
| 12 | Gabor, fixed cycles | 4 cycles/degree, $b_x = b_y = 1$ octave |
| 13 | Gabor, fixed cycles | 8 cycles/degree, $b_x = b_y = 1$ octave |
| 14 | Gabor, fixed cycles | 16 cycles/degree, $b_x = b_y = 1$ octave |
| 15 | Gabor, elongated | 4 cycles/degree, $\sigma_x = 0.05°$, $b_y = 0.5$ octave |
| 16 | Gabor, elongated | 8 cycles/degree, $\sigma_x = 0.05°$, $b_y = 0.5$ octave |
| 17 | Gabor, elongated | 16 cycles/degree, $\sigma_x = 0.05°$, $b_y = 0.5$ octave |
| 18 | Gabor, elongated | 4 cycles/degree, $b_x = 2$ octave, $b_y = 1$ octave |
| 19 | Gabor, elongated | 4 cycles/degree, $\sigma_x = 0.05°$, $b_y = 1$ octave |
| 20 | Gabor, elongated | 4 cycles/degree, $b_x = 1$ octave, $b_y = 2$ octave |
| 21 | Gabor, elongated | 4 cycles/degree, $b_x = 1$ octave, $\sigma_y = 0.5°$ |
| 22 | Compound Gabor | 2 and $2\sqrt{2}$ cycles/degree |
| 23 | Compound Gabor | 2 and 4 cycles/degree |
| 24 | Compound Gabor | 4 and $4\sqrt{2}$ cycles/degree |
| 25 | Compound Gabor | 4 and 8 cycles/degree |
| 26 | Gaussian | $\sigma_x = \sigma_y = 30$ min |
| 27 | Gaussian | $\sigma_x = \sigma_y = 8.43$ min |
| 28 | Gaussian | $\sigma_x = \sigma_y = 2.106$ min |
| 29 | Gaussian | $\sigma_x = \sigma_y = 1.05$ min |
| 30 | Edge × Gaussian | |
| 31 | Line × Gaussian | 0.5 min (1 pixel) wide horizontal line |
| 32 | Dipole × Gaussian | 3 pixels wide |
| 33 | 5 collinear Gabors | 8 cycles/degree, in phase, $b_x = b_y = 1$ octave, separation = 5 $\sigma_x$ |
| 34 | 5 collinear Gabors | 8 cycles/degree, out of phase, $b_x = b_y = 1$ octave, separation = 5 $\sigma_x$ |
| 35 | Binary noise | 1 × 1 min samples |
| 36 | Oriented Gabor | 4 cycles/degree, 45°, $b_x = b_y = 1$ octave |
| 37 | Oriented Gabor | 4 cycles/degree, 0°, $b_x = b_y = 1$ octave |
| 38 | Compound Gabor | 4 cycles/degree, 0° and 90°, $b_x = b_y = 1$ octave |
| 39 | Compound Gabor | 4 cycles/degree, 45° and 90°, $b_x = b_y = 1$ octave |
| 40 | Disk | 1/4° diameter |
| 41 | Bessel × Gaussian | 4 cycles/degree |
| 42 | Checkerboard | 4 cycles/degree fundamental |
| 43 | Natural image | Image of San Francisco |

Table 1. Definition and parameters of each of the 43 ModelFest stimuli. Parameters $\sigma_x$ and $\sigma_y$ are the Gaussian standard deviations in horizontal and vertical dimensions; $b_x$ and $b_y$ are the half-amplitude full bandwidths in horizontal and vertical frequency dimensions. Unless stated otherwise, $\sigma_x = \sigma_y = 0.05$ degrees, sinusoids were modulated vertically (90° orientation) and were in cosine phase.

($o = 1,\ldots,O$), and replication ($r = 1,\ldots,R$). The mean for each observer over replications can be written $t_{s,o}$, and these are shown for all 16 observers in Figure 2, plotted as a function of the arbitrary index number. Each observer is represented by a different color. We write $t_o$ for the mean of $t_{s,o}$ over stimuli for each observer, and $t_s$ for the mean over observers for each stimulus. The variability among observers can be represented by

$$\mathrm{RMS}_0 = \sqrt{\frac{1}{SO}\sum_{s=1}^{S}\sum_{o=1}^{O}(t_{s,o} - t_s)^2}. \quad (2)$$

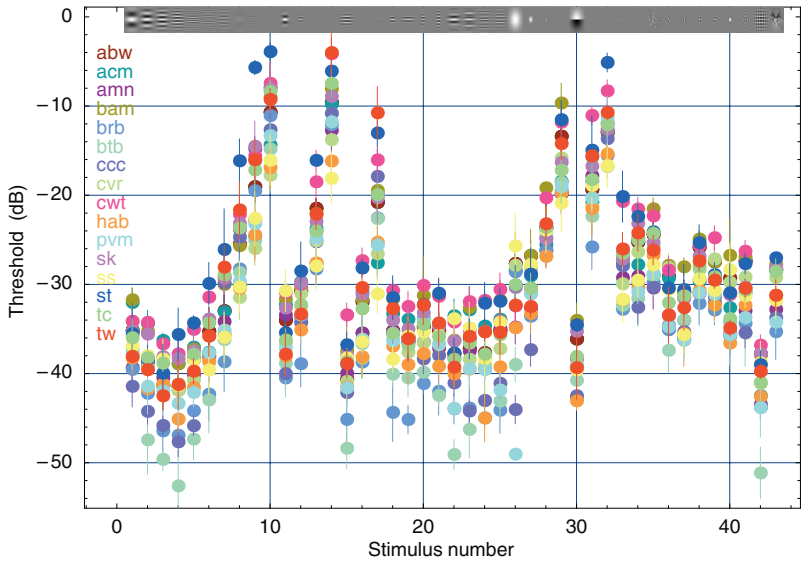


Figure 2. Data from the ModelFest experiment. Each point is the mean for one observer for one stimulus, and the error bars indicate ±2 *SE*. Each observer is represented by a distinct color. The small pictures at the top illustrate the stimuli.

This is the RMS error of a model in which threshold for each stimulus is given by the mean over observers. It is also the maximum likelihood estimate of the standard deviation of a normal distribution underlying such a model. The value of $\mathrm{RMS}_0$ for these data is 3.46 dB, indicating considerable variation among observers. Some of this variance is accounted for by the different mean sensitivities of the observers. We can construct a second measure of error,

$$\mathrm{RMS}_1 = \sqrt{\frac{1}{SO}\sum_{s=1}^{S}\sum_{o=1}^{O}((t_{s,o}-t_o)-(t_s-t_0))^2} \qquad (3)$$

in which we subtract the observer means $t_o$ from each threshold $t_{s,o}$, and the grand mean $t_0$ from each stimulus mean $t_s$. This error has a value of 2.29 dB. The RMS error associated with the observers,

$$\mathrm{RMS}_o = \sqrt{\mathrm{RMS}_0^2 - \mathrm{RMS}_1^2} = 2.59\,\mathrm{dB} \qquad (4)$$

can be regarded as the standard deviation of the observer sensitivities in dB.

When this standard deviation estimate is divided by the square root of the number of observers, the result, 0.56 dB, can be regarded as an estimate of the standard deviation of the $t_s - t_0 - (\tau_s - \tau_s)$, where $\tau$ is the corresponding true value. If the models were correct, the models' predictions would be that the $\tau$ and the RMS error of the model would be another estimate of this same standard deviation. Our best possible model RMS error is thus 0.56 dB.

The average thresholds over all observers are shown in Figure 3. The averages are shown both in units of dB

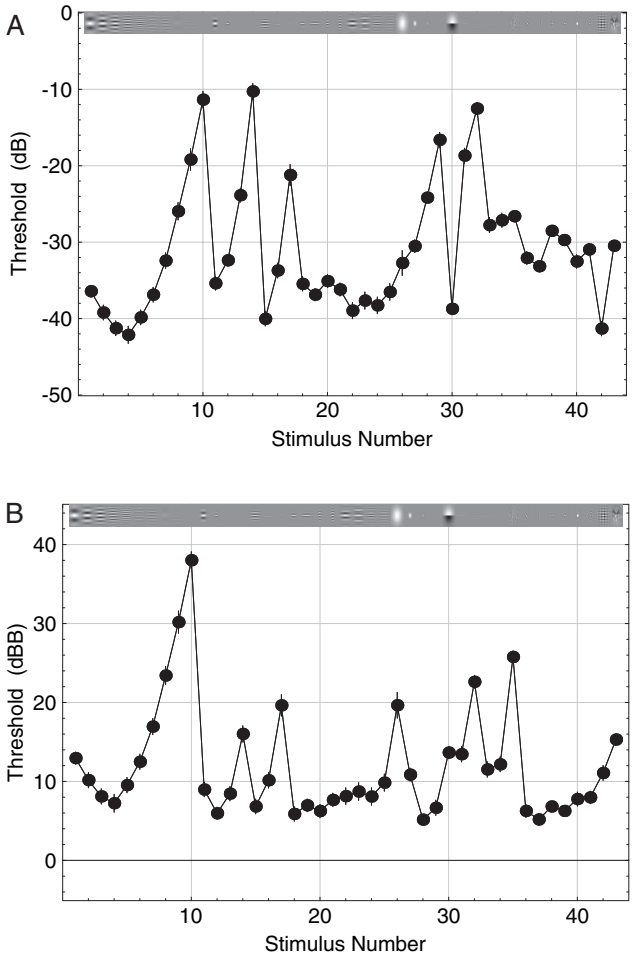


Figure 3. Average ModelFest thresholds. Each point is the mean of 16 observers, and the error bars indicate ±2 *SE*. The small pictures at the top illustrate the stimuli. (A) Thresholds in dB; (B) thresholds in dBB.

(Figure 3A) and in units of dBB (Figure 3B). The dBB is a measure of the contrast energy of a stimulus, normalized by a nominal minimum threshold of $10^{-6}$ deg$^{-2}$ s$^{-1}$. Zero dBB is defined so as to approximate the minimum visible contrast energy for a sensitive human observer (Watson, 2000; Watson, Barlow, & Robson, 1983; Watson, Borthwick, & Taylor, 1997). A virtue of the dBB unit is that it takes into account the contrast energy of the stimulus. One quick observation we may make from Figure 3B is that for the average observer, the best thresholds are about 7 dBB above (less sensitive than) the canonical "sensitive human observer." Curiously, the ModelFest stimulus that the eye sees best is not a Gabor but a small Gaussian (stimulus 28). This differs from the classical result (Watson, Barlow, & Robson, 1983), but that result was obtained with moving rather than stationary targets.

The first 10 thresholds constitute a CSF as measured with Gabor functions of fixed size. It resembles similar data collected previously and shows the typical bandpass shape with a minimum of −42.13 dB (7.24 dBB) at about 4 cycles/degree. The following four stimuli (11–14) form a CSF for Gabor functions with a fixed number of cycles, or equivalently a fixed 1 octave bandwidth. The latter thresholds resemble similar data collected previously by Watson (1987).

# Model structure

In this paper, we investigate a class of models that incorporate a set of sequential operations, several of which may be inserted or removed or whose parameters may be fixed or allowed to vary. In this section, we define the sequence of elements and the individual elements. The overall structure and sequence of elements of the model is shown in Figure 4.

## Input and output

The input to the model was one of the digital stimulus images, as provided in the file modelfest-stimuli described in Appendix A. Each image has $N_y$ = 256 rows and $N_x$ = 256 columns. The output of the model was a contrast threshold.

## Contrast

The stimulus grayscale image was first converted to a luminance contrast image, defined as the luminance image, minus the nominal mean luminance, divided by that mean. Rearranging Equation 1 shows that this is accomplished by subtracting the nominal mean graylevel of 128 and dividing by 127.

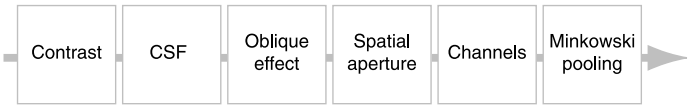


Figure 4. Elements of the component model.

## Contrast sensitivity filter

The contrast image is then filtered by a radially symmetric CSF. The filter is implemented as a discrete digital finite impulse response (FIR) filter created by sampling a one-dimensional CSF in the two-dimensional discrete Fourier transform (DFT) domain. We consider a number of different versions of the CSF, as described below. An example digital CSF is shown in Figure 5, depicted as log gain versus spatial frequency. Note the hole in the center, corresponding to the decline at low frequencies, and the decline toward the edges, corresponding to the decline at high frequencies.

## Oblique effect filter (OEF)

The oblique effect is the well-known decline in contrast sensitivity at oblique orientations (Campbell et al., 1966; MacMahon & MacLeod, 2003). The ModelFest data do not contain sufficient oblique patterns at varying frequencies to effectively constrain this effect, so we have based our oblique effect model on data from (Berkley et al., 1975). These data are shown in Figure 6 as the $\log_{10}$ ratio of thresholds for 0° and 45° oriented gratings at various spatial frequencies.

In these linear-log coordinates, sensitivity at the oblique orientation declines linearly with frequency, reaching a value of about 1.6 $\log_{10}$ units at 25 cycles/degree. We fit a linear function (red line) to these data but truncate it when

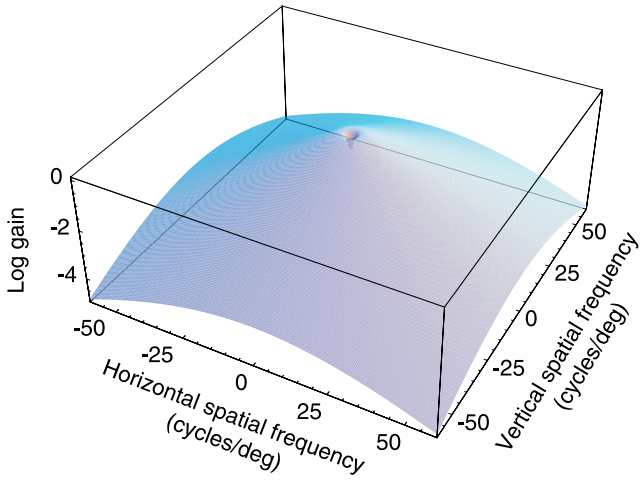


Figure 5. Contrast sensitivity filter (CSF). This example is for the HPmH function, described below. In this picture the peak gain has been arbitrarily set at unity.

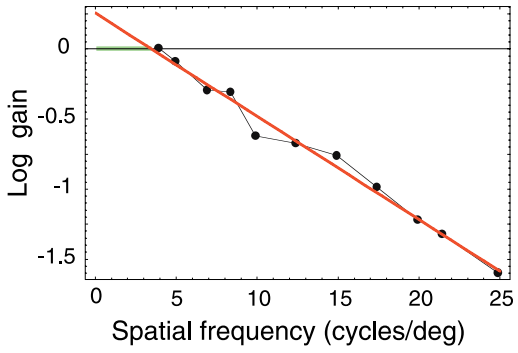


Figure 6. Data and model for the oblique effect. The points are reductions in sensitivity for targets at 45° orientation, relative to that at 0°, as a function of spatial frequency (Berkley et al., 1975). The red line is a linear fit in these linear-log coordinates. The green line is at zero attenuation. The lower envelope of the two lines is the relative attenuation prescribed by the model for patterns at an orientation of 45°.

the function goes above 0 log attenuation (green line). These two lines form the frequency-dependent part of our oblique effect model. We assume in addition that at any given spatial frequency, sensitivity varies as a sinusoidal function of orientation. The resulting model for the oblique effect is then given by

$$O(f,\theta) = 1 - \left(1 - \exp\left(-\frac{f-\gamma}{\lambda}\right)\right)\sin^2(2\theta) \text{ if } f > \gamma$$
$$= 1 \text{ if } f \le \gamma \quad (5)$$

where $\gamma$ = 3.48 cycles/degree, $\lambda$ = 13.57 cycles/degree. This function has two parameters corresponding to the frequency at which sensitivity begins to decline ($\gamma$) and the slope of the linear-log decline ($\lambda$). From this function we can create a discrete FIR digital oblique effect filter (OEF), as shown in Figure 7.

Because both the CSF and the OEF are applied in sequence to the image, they may be combined to form a single contrast sensitivity and oblique effect filter (CSOEF) as shown in Figure 8.

## Aperture

Contrast sensitivity declines rapidly with eccentricity, and the rate of decline increases strongly with spatial frequency (Robson & Graham, 1981). However, in this modeling exercise we have chosen to test only a frequency-independent decline (an aperture). Our rationale was that the region under consideration (2.133 × 2.133 degrees) is relatively small, and we were interested in testing simple models. The form we have chosen for the decline in sensitivity with eccentricity is a Gaussian,

$$A(r) = \exp\left(-\frac{r^2}{2\sigma^2}\right) \quad (6)$$

where $r$ is the distance from fixation in degrees, and $\sigma$ is the standard deviation of the Gaussian (which we also refer to as its size), also in degrees. This function was chosen primarily for mathematical convenience: its rate of decline is easily controlled and it never goes to zero. The peak value of the Gaussian is 1, so that the aperture defines the attenuation of sensitivity relative to that at the point of fixation. The Gaussian aperture multiplies the image produced by the CSF and OEF elements of the model. The aperture was centered on the image, which corresponds to an assumption that the observer fixated the center of each target.

## Channels

There is considerable physiological and psychophysical evidence that the visual system partitions spatial informa-

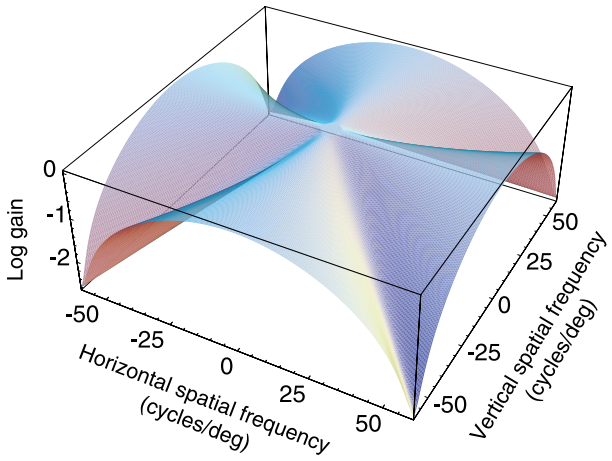


Figure 7. Oblique effect filter (OEF) with parameters $\gamma$ = 3.48 cycles/degree and $\lambda$ = 13.57 cycles/degree.

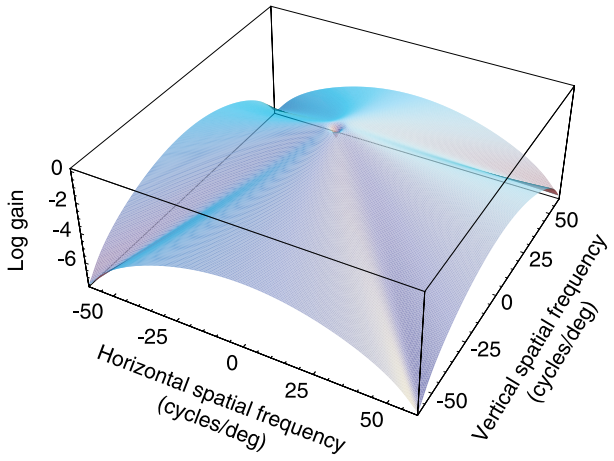


Figure 8. Combined contrast sensitivity and oblique effect filter (CSOEF).

tion into a number of parallel channels, each selective for a band of spatial frequency and orientation. Consequently, spatial frequency channels are a common feature of modern models of spatial vision (Watson & Solomon, 1997). As in (Watson, 2000), we implement a set of channels with Gabor receptive fields. We reproduce Table 2 from that paper to specify the parameters of the channel stage of the model. These are implemented as a set of digital FIR filters in the DFT domain. Channel responses are down-sampled in proportion to frequency (pyramid sampling). For a given value of the Minkowski summation parameter $\beta$ (see below), the channel gains were adjusted to yield approximately flat contrast sensitivity over frequency. This means that variations in contrast sensitivity over frequency are controlled primarily by the CSF. In this report, we did not vary any of the parameters of the Gabor channel component.

## Pooling

The final stage in the component model is a pooling over space and, if present, over channels. Following long precedent, we implement this pooling as a Minkowski metric (Graham, 1977; Quick, 1974; Robson & Graham, 1981; Watson, 1979). Note that all stages in this model prior to pooling are linear, and that the pooled response is assumed to equal 1 at threshold, so we write this as

$$1 = \left[ \sum_{y=1}^{N_y} \sum_{x=1}^{N_x} p_x p_y |c_T r_{x,y}|^\beta \right]^{1/\beta} \quad (7)$$

or

$$c_T = \left[ \sum_{y=1}^{N_y} \sum_{x=1}^{N_x} p_x p_y |r_{x,y}|^\beta \right]^{-1/\beta} \quad (8)$$

where $c_T$ is the contrast threshold, $r_{x,y}$ are the processed pixel values, prior to pooling, to a stimulus of unit peak contrast, and $p_x$ and $p_y$ are the width and height of each pixel in degrees. These latter terms are introduced to make the result independent of the specific resolution at which the calculation is performed.

| | |
|---|---|
| Number of frequencies | 11 |
| Number of orientations | 4 |
| Number of phases | 2 (odd and even) |
| Bandwidth | 1.4 octaves |
| Highest center frequency | 30 cycles/degree |
| Lowest center frequency | 0.9375 cycles/degree |
| Frequency spacing | 1/2 octave |
| Orientation spacing | 45° |
| Pyramid sampling | Yes |

Table 2. Gabor channel model parameters.

The Minkowski formulation is useful because it encompasses a number of pooling models, including energy summation ($\beta = 2$) probability summation ($\beta \sim 3$), and peak detection ($\beta = \infty$) (Watson, 1979).

When channels are present, the calculation of Equation 8 is performed within each channel $q$, and the results are combined over the $Q$ channels,

$$c_T = \left[ \sum_{q=1}^{Q} c_{T,q}^{-\beta} \right]^{-1/\beta}. \quad (9)$$

# Contrast sensitivity filters

The CSF element of the model has been described above. Here we describe the various forms of this element that we considered. Because the CSF is bandpass in form, many of the candidate functions are composed of a high-frequency lobe minus a low-frequency lobe. We have avoided a profusion of symbols by using the same parameter names in different functions. Many of the functions share parameters playing approximately the same role; for example a parameter $f_0$ that scales frequency in the high-frequency lobe, a parameter $f_1$ that scales frequency in the low-frequency lobe, and a parameter $a$ that determines the weight of the low-frequency lobe. Note that $f_0$ and $f_1$ may also be thought of as specifying widths of subtractive center and surround components of the space domain convolution kernel corresponding to the filter (which is in turn sometimes thought of as the receptive field corresponding to the filter). Because many of the "lobe" functions (exp, Gaussian, sech) have a value of 1 at $f = 0$, the DC gain is in these cases equal to $1 - a$.

Each CSF also has a multiplicative gain parameter that is not shown. Each CSF is identified by a symbolic name (DoG, HSmG, etc.) that we use in the remainder of the paper. In the descriptions and in Table 3, we indicate the number of parameters embodied in each function.

## Log-sensitivity interpolation (LSI)

The LSI function is constructed by linear interpolation between log-sensitivity values at each of the 10 spatial frequencies used in the Gabor function stimuli 1–10. In addition, a parameter value is assigned at 0 cycles/degree. (A further fixed value of −50 dB is assigned at a frequency of 256 cycles/image to bound the interpolation.) This function thus has 11 parameters. It is the least constrained of all the CSF functions considered here. It is included to provide a CSF that embodies few assumptions about functional form.

## Constant

This function is a constant at all spatial frequencies. It is introduced to illustrate the effect of the presence or absence of the CSF. The function has only one parameter (gain), and is written

$$S_{\text{constant}}(f) = 1. \tag{10}$$

## DoG

This function is a difference of Gaussians. It is a good description of the sensitivity of individual retinal ganglion cell receptive fields (Enroth-Cugell & Robson, 1966; Enroth-Cugell, Robson, Schweitzer-Tong, & Watson, 1983; Rodieck, 1965). Including gain, it has four parameters:

$$S_{\text{DoG}}(f;f_0,f_1,a) = \exp\left[-(f/f_0)^2\right] - a\,\exp\left[-(f/f_1)^2\right]. \tag{11}$$

## EmG

This consists of an exponential minus a Gaussian:

$$S_{\text{EmG}}(f;f_0,f_1,a) = \exp[-f/f_0] - a\,\exp\left[-(f/f_1)^2\right]. \tag{12}$$

The exponential is suggested by the nearly linear decline in sensitivity at high frequencies on a log-linear plot (Campbell et al., 1966). This CSF was earlier suggested as a good fit to the fixed size Gabor ModelFest stimuli (Carney et al., 2000). Including gain, it has four parameters.

## HmG

This function consists of a hyperbolic secant minus a Gaussian:

$$S_{\text{HmG}}(f;f_0,f_1,a) = \operatorname{sech}[f/f_0] - a\,\exp\left[-(f/f_1)^2\right]. \tag{13}$$

This function does not appear to have been used previously to model the CSF. Including gain, it has four parameters.

## HPmG

This is the same as HmG, except that the scaled frequency argument of the hyperbolic secant is raised to a power:

$$S_{\text{HPmG}}(f;f_0,f_1,a,p) = \operatorname{sech}\left[(f/f_0)^p\right] - a\,\exp\left[-(f/f_1)^2\right]. \tag{14}$$

This function was suggested by Christopher W. Tyler (personal communication, March 12, 2004). Including gain, it has five parameters.

## HmH

This function is a difference of hyperbolic secants:

$$S_{\text{HmH}}(f;f_0,f_1,a) = \operatorname{sech}[f/f_0] - a\,\operatorname{sech}[f/f_1]. \tag{15}$$

This function does not appear to have been used previously to model the CSF. This is the same as HmG, with the Gaussian replaced by a hyperbolic secant. Including gain, it has four parameters.

## HPmH

This is a hyperbolic secant whose scaled frequency is raised to the power $p$ minus a hyperbolic secant:

$$S_{\text{HPmH}}(f;f_0,f_1,a,p) = \operatorname{sech}\left[(f/f_0)^p\right] - a\,\operatorname{sech}\left[f/f_1\right]. \tag{16}$$

This is the same as HPmG, with the Gaussian replaced by a hyperbolic secant. This function does not appear to have been used previously to model the CSF. Including gain, it has five parameters.

## LP

This function is a parabola in a graph of log-sensitivity versus log-frequency (Ahumada & Peterson, 1992; Rohaly & Owsley, 1993). Including gain, it has four parameters. On the low-frequency side, it is truncated at a value of $a$:

$$S_{\text{LP}}(f;f_0,b,a) = 10^{-\left(\frac{\log_{10}(f/f_0)}{b}\right)^2} = 1-a \qquad f < f_0 \text{ and } s_{\text{LP}} < 1-a. \tag{17}$$

## MS

This function was introduced by Mannos & Sakrison (1974) in their pioneering work on image quality. It is the product of what might be called a generalized Gaussian (with exponent other than two) and a linear function of frequency, which serves to enhance high frequencies relative to low:

$$S_{\text{MS}}(f;f_0,a,p) = \left(1-a+\frac{f}{f_0}\right)\exp\left[-(f/f_0)^p\right]. \tag{18}$$

Including gain, it has four parameters.

### YQM

This function was derived from a model of contrast sensitivity by Yang, Qi, & Makous (1995). Like EmG, it includes an exponential decline at high frequencies, with an additional divisive term to attenuate low frequencies:

$$S_{\mathrm{YQM}}(f;f_0,f_1,a) = \frac{\exp[-f/f_0]}{1+\frac{a}{1+(f/f_1)^2}}. \quad (19)$$

Including gain, it has four parameters.

## Model implementation and optimization

The model was implemented in the Mathematica programming language (Wolfram, 2003). Parameters of each version of the model were estimated by means of general optimization routines. To insure the accuracy of the results, we occasionally used three different optimization routines. These were the built-in Mathematica functions FindMinimum and NMinimize and the GlobalSearch function provided by Loehle Enterprises (2004). Note that no optimization procedure is guaranteed to yield the absolute minimum of an arbitrary function; consequently, all errors reported must be regarded as provision upper bounds on the minimum achievable error.

Filtering operations, such as those employed by the CSF, the oblique effect, and the channels, were implemented by cyclic convolution in the frequency domain. Border effects were minimized by the Gaussian apertures used by all ModelFest stimuli.

The measure of error that we use is RMS error in dB. If both mean thresholds $t_j$ and model predictions $m_j$ are specified in dB, and the number of stimuli is $J$, this is given by

$$\mathrm{RMS} = \sqrt{\frac{1}{J}\Sigma(t_j - m_j)^2}. \quad (20)$$

## Model fits

### Effect of components

The component model presents a very large number (264) of configurations to be tested, depending upon the choices made regarding the CSF (11), the OEF (2: present or not), the aperture (2: present or not), channels (2: present or not), and pooling (3: $\beta = 2$, $\beta =$ free, $\beta = \infty$). For each tested configuration, we estimated all free parameters and recorded the residual error. We have not evaluated the fit of every possible configuration but have rather tried to understand the contribution of each component, the best version of each option, and the best obtainable overall fit.

Figure 9 illustrates one trajectory through the error space of the component model. Moving from left to right, each point shows the error as we add one additional component to the model. At each point, all free parameters are re-optimized.

The first point shows the error that results from a model in which all of the model components have been turned off. This consists of a constant CSF followed by peak detection ($\beta = \infty$). Alternatively, we may say that it has no CSF, no pooling, no oblique effect, no aperture, and no channels. It has a single parameter (sensitivity) and assumes a target is detected whenever its peak contrast equals a certain value. While not a reasonable model, it provides a useful error benchmark, of about 8 dB, against which other fits may be compared.

The second point shows the error of a model consisting of the LSI CSF, followed by a peak detector. The 12 parameters of the LSI function have been optimized. Addition of a CSF thus reduces the error by almost a factor of two but still leaves a poor fit with an RMS error of about 4.5 dB.

When peak detection ($\beta = \infty$) is replaced with energy detection ($\beta = 2$), the error is again reduced by about a factor of two to a value of about 2 dB, as shown by the third point in the series. Allowing the pooling exponent to vary ($\beta =$ free), which we call generalized energy, results in yet another drop in error by about a factor of two, as shown by the fourth point. The RMS error at this point is in the neighborhood of 1, which we can characterize as a "good" fit (see below).

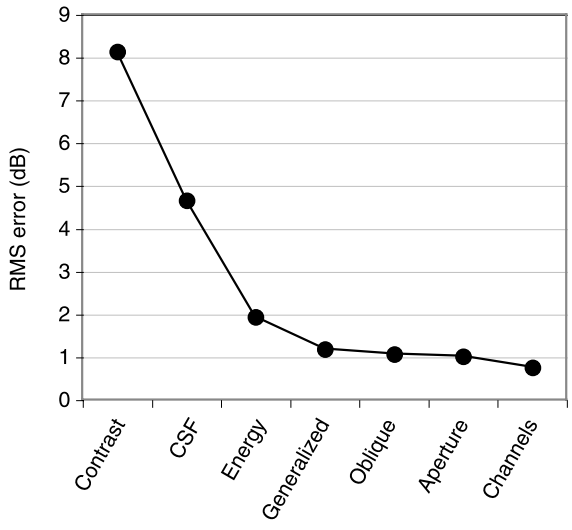


Figure 9. Decline in RMS error as individual model components is added.

The fifth point shows a further small reduction in error as a result of adding the oblique effect, and adding the aperture reduces error by a similarly small amount, yielding an error of less than 1, as shown by the sixth point.

The final point shows the further reduction in error due to the addition of channels. The reduction is substantial but modest relative to the contribution of elements such as the CSF or energy detection. We will return to this point below.

This trajectory is just one of many we might have taken, but it serves to illustrate the relative magnitude of the contribution to the reduction of error yielded by the various model components. In the following sections, we consider in more detail the effect of several of the individual components on the overall error.

## Predictions of best-fitting models

To illustrate the quality of fit of the best-fitting models, we plot in Figure 10 the predictions for each stimulus along with the corresponding average thresholds, for the best channel model (point labeled "channels" in Figure 9) and the best no-channel model (point labeled "Aperture" in Figure 9). The greatest difference between the two fits and the largest error for the no-channel model occur at stimuli 35 and 43, which are the noise sample and the natural image, respectively. We will return to this observation below.

### *Graphic conventions*

To assist the reader in comprehending the results in the remainder of this paper, we have adopted some graphics symbol conventions. An open symbol indicates the use of an aperture, while a filled symbol indicates the absence of an aperture. A square symbol indicates that $\beta = 2$, while a circular symbol indicates that $\beta$ was free to vary. Finally, when channels are included, a dashed line is use to connect the points.

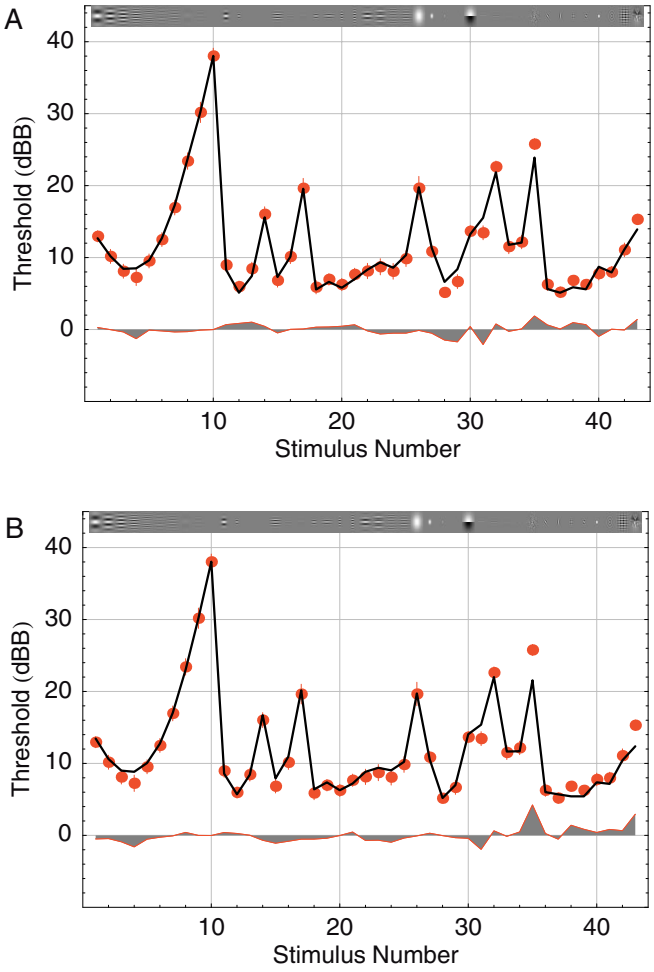


Figure 10. Plot of average thresholds (red points), best-fitting model (black line), and residual error (red line and gray area). (A) The best channel model has an RMS error of 0.76 dB. (B) The best no-channel model has an RMS error of 1.02 dB. The vertical axis is in units of dBB, which are a measure of the contrast energy of the stimulus, normalized by a nominal minimum threshold of $10^{-6}$ deg$^{-2}$ s$^{-1}$ (Watson, 2000; Watson et al., 1997).

## Contrast sensitivity functions

All of the results in Figure 9 were obtained with the LSI CSF, which consists of a linear interpolation between points on a graph of log sensitivity versus spatial frequency. An example of this function is shown in Figure 11. For comparison, the mean empirical sensitivity (inverse of threshold) at each frequency is also shown in red. The close agreement shows that although the LSI CSF is optimized relative to the entire data set, it nevertheless provides a very close fit to the subset of Gabor data.

Note that the LSI function was designed to be a "model-free" CSF, whose shape is free to vary to best match the data. It has 11 parameters, one for each Gabor spatial frequency and one for 0 cycles/degree. Because it embodies few constraints, we expect it to be the best-fitting (lowest error) CSF, and thus a useful benchmark of the achievable fit, and a useful comparison with the other CSF functions.

In addition to the LSI CSF and the constant CSF, we have considered the nine specific CSFs defined in Contrast Sensitivity Filters. Here we assess the performance of these functions in the context of the no-channel model. The result of adopting each variant CSF into this condition is shown by the black symbols in Figure 12. The functions, their numbers of parameters and corresponding RMS errors are also enumerated in Table 3.

With the exception of the DoG and constant, all the functions fit reasonably well and differ in their fit by less than two tenths of a decibel. The best-fitting formula is HPmH, pictured in Figure 13. We plot it along with the parameter points from the LSI function (black circles

| CSF | Parameters | RMS error (dB) |
|---|---|---|
| LSI | 11 | 1.0243 |
| HPmH | 5 | 1.0329 |
| HPmG | 5 | 1.0468 |
| YQM | 4 | 1.0694 |
| EmG | 4 | 1.0755 |
| LP | 4 | 1.0916 |
| HmG | 4 | 1.0959 |
| HmH | 4 | 1.1104 |
| MS | 4 | 1.2009 |
| DoG | 4 | 1.7830 |
| Constant | 1 | 5.8607 |

Table 3. Contrast sensitivity filter (CSF) functions. For each function, we indicate the number of parameters and the residual error. Other conditions: no channels, fixed oblique effect, Gaussian aperture, $\beta$ = free.

from Figure 11) to illustrate that the continuous, analytic five-parameter HPmH function is a close match to the unconstrained 11-parameter LSI function.

In Figure 14 we show all nine functions. The purpose of this figure is to show that all of the functions are in close agreement, with the possible exception of DoG and MS. The latter falls rapidly at low frequencies, while the former is much more "flat topped" than the best-fitting curves. Given the roughly equal performance of the functions, other considerations may influence selection of a function for either applied or theoretical purposes. The two best-fitting curves have five parameters, but some functions with only four parameters perform almost as well. The functions YQM and EmG have inflections near to zero, which may be a concern in some applications. LP has a sharp corner on the low-frequency side, which may also be objectionable, and MS is not well behaved at low frequencies. Some of these attributes are evident in a plot of the derivative of each function in the vicinity of zero, as shown in Figure 15. Both EmG and YQM are negative at zero, and MS climbs rapidly as it approaches zero.

We should note, however, that in applications or simulations in which the CSF is applied to a digital image, the lowest frequency in the image (apart from zero) is 1 cycle/image. For an image subtending D degrees, this lowest fre-

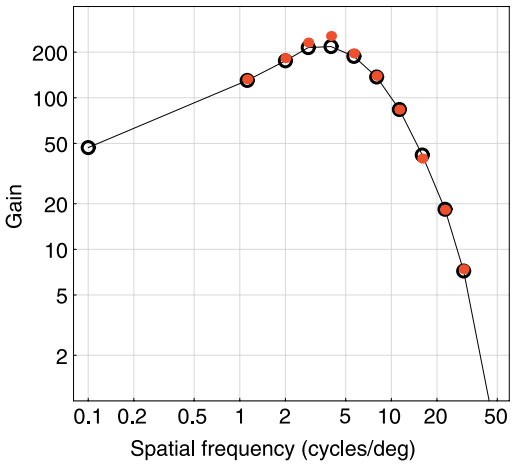


Figure 11. LSI contrast sensitivity filter (CSF). The black circles show estimated sensitivity values at the spatial frequencies employed in the ModelFest fixed size Gabor targets (stimuli 1–10), plus a value for 0 cycles/degree, which we plot arbitrarily at 0.1 cycles/degree. The LSI sensitivity function is linearly interpolated between these points. This example is the best-fitting version for the case of no channels, an aperture, an oblique effect, and $\beta$ free to vary. This corresponds to the point labeled "aperture" in Figure 9, and the overall fit shown in Figure 10B. The red points are the mean sensitivities (inverse thresholds, shifted vertically by an arbitrary factor of 2) from the ModelFest data set for fixed size Gabor targets (stimuli 1–10).

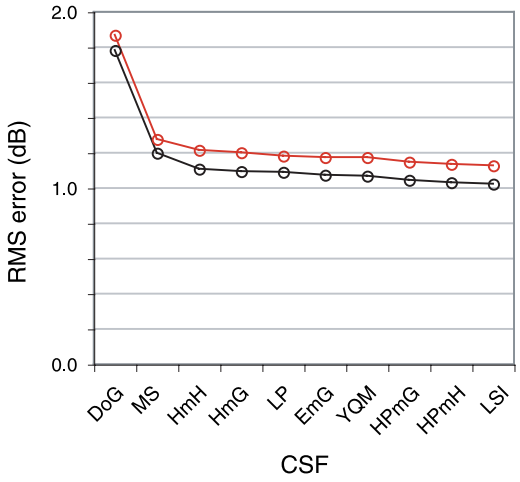


Figure 12. Fit of various contrast sensitivity filter (CSF) functions. The black points are for a fixed oblique effect, the red points are for no oblique effect. Other conditions: no channels, Gaussian aperture, $\beta$ = free.

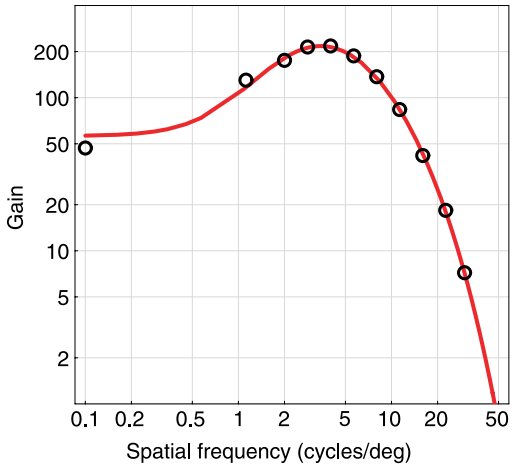


Figure 13. Plot of the HPmH contrast sensitivity filter (CSF). The black points are the estimated parameters of the LSI CSF for comparison. Other conditions: no channels, fixed oblique effect, Gaussian aperture, $\beta$ = free.

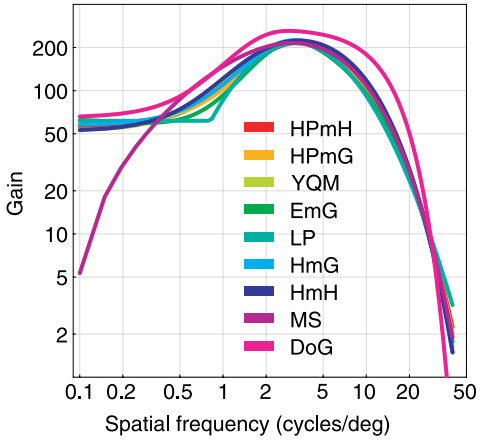


Figure 14. Best-fitting version of each contrast sensitivity filter (CSF). Other conditions: no channels, fixed oblique effect, Gaussian aperture, $\beta$ = free.

quency will be 1/D cycles/degree. If the image is small, for example the 2.13 degrees used in the ModelFest experiment, then this lowest frequency will be 0.47 cycles/degree, and what happens to the function between 0 and 0.47 will not be manifest in the digital filtering.

## Oblique effect

As noted above, the ModelFest data set does not contain enough oblique signals at varying frequencies to allow us to use it to estimate the oblique effect, and we have therefore derived parameters for the effect from prior data. Here we examine the effect of including or excluding that fixed oblique effect. The black points in Figure 12 are for the no-channel model that includes the oblique effect, while the red points are for the same model when the oblique effect is removed. As elsewhere in this paper, each point reflects a re-estimation of all parameters. The figure shows that the oblique effect reduces error uniformly over CSFs by about 0.1 dB. Of course, we might expect that inclusion of stimuli at oblique orientations at high spatial frequencies (where the effect is strongest) would yield much larger differences.

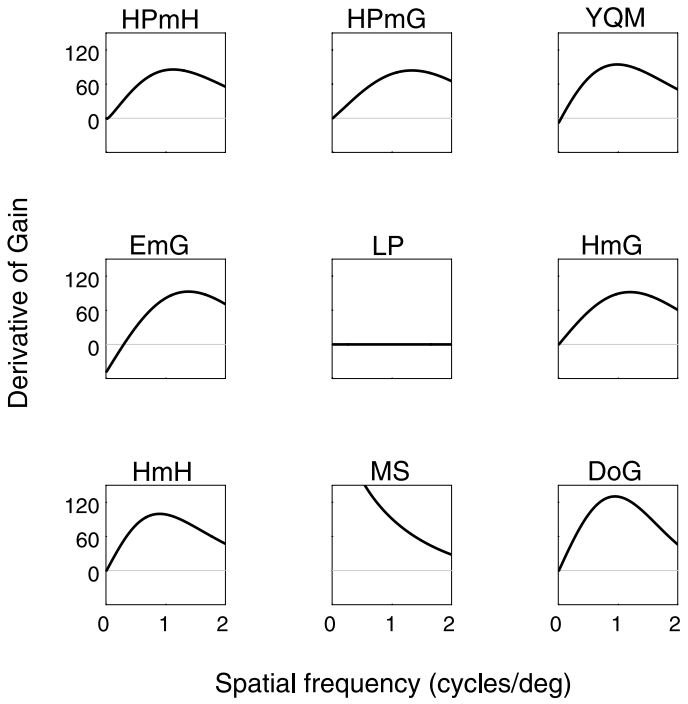


Figure 15. Derivatives of the nine contrast sensitivity filter (CSF) functions in the neighborhood of zero.

## Pooling exponent $\beta$

For the no-channel model, estimated values of the pooling exponent $\beta$ ranged between 2 and 3. Among high quality fits (error < 1.2 dB), the mean $\beta$ was 2.58 ($SD$ = 0.18, $n$ = 21). As we will see in greater detail below, estimates of $\beta$ interact with the presence and size of the spatial aperture. Without an aperture, high quality fits of $\beta$ average 2.7 ($SD$ = 0.02, $n$ = 7) while with an aperture the average was 2.52 ($SD$ = 0.19, $n$ = 14).

In their early study of spatial summation, Robson & Graham (1981) found that both foveal and peripheral results were predicted best with an exponent of 3.5. The reason for the discrepancy between their result and ours is not clear; although they did not include an aperture, they used empirical estimates of the decline in sensitivity with eccentricity, which have a similar effect.

The estimates of $\beta$ do vary somewhat with the CSF. In Figure 16 we show the summation exponents $\beta$ estimated for the no-channel model for each CSF, plotted versus the estimated size, $\sigma$. For all except the poorly fitting DoG, the variations in $\beta$ are modest. But the reciprocity be-

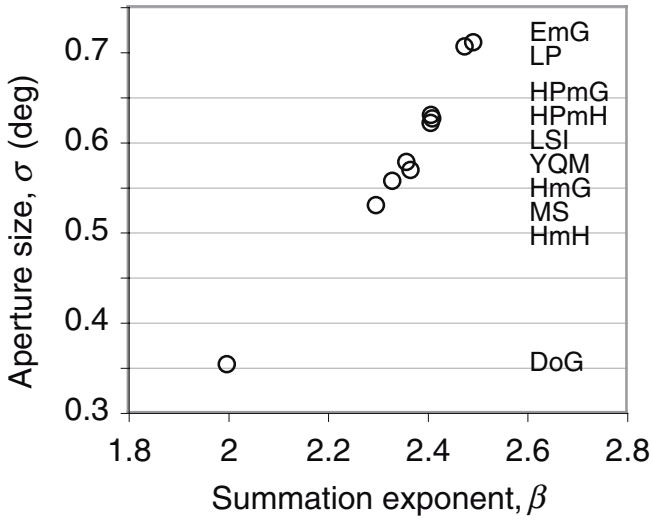


Figure 16. Aperture size $\sigma$ versus summation exponent $\beta$ for the no-channel model. Other conditions: fixed oblique effect.

tween these two parameters $\beta$ and $\sigma$ is striking. We return to this reciprocity in the following section.

## Spatial aperture

The best fit of the no-channel model is obtained when an aperture is included, as shown by the black open circles in Figure 17. However, the fit is only slightly degraded when the aperture is removed (filled red circles). However, the removal of the aperture results in a change in estimate of $\beta$ from 2.39 to 2.70 (averaged over all CSF functions except DoG; $SD$ = 0.06 and 0.02, respectively). This suggests that the aperture and a higher $\beta$ both serve to reduce the efficiency of spatial summation. This notion is confirmed when $\beta$ is fixed at 2. The absence of an aperture now causes a marked increase in error (red filled squares), while the presence of an aperture yields a fit which is only slightly poorer than when $\beta$ is free to vary (black open squares).

This observation is also consistent with the behavior of the estimated values of the aperture size $\sigma$, when $\beta$ is free or when it is fixed at 2. In the latter case, inefficient summation must rely on the aperture, so a relatively small size is estimated ($\sigma$ = 0.364 degrees, $SD$ = 0.003), while in the former case a $\beta$ greater than 2 can do the same job, so a larger aperture is found ($\sigma$ = 0.615 degrees, $SD$ = 0.06) (in both cases, averaged over all CSFs except DoG).

A final observation on the trade-off between $\beta$ and $\sigma$ is provided in Figure 18. Here we have fit the standard model, but fixed $\beta$ at a particular value between 2 and 3, and re-estimated all remaining parameters. We plot the parameter $\sigma$ and the error. This shows that as $\beta$ increases,

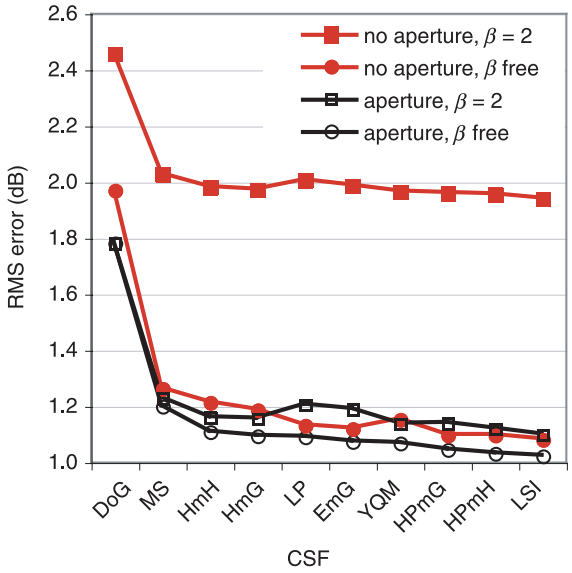


Figure 17. Effect of contrast sensitivity filter (CSF), aperture, and $\beta$ on RMS error for the no-channel model. Other conditions: fixed oblique effect.

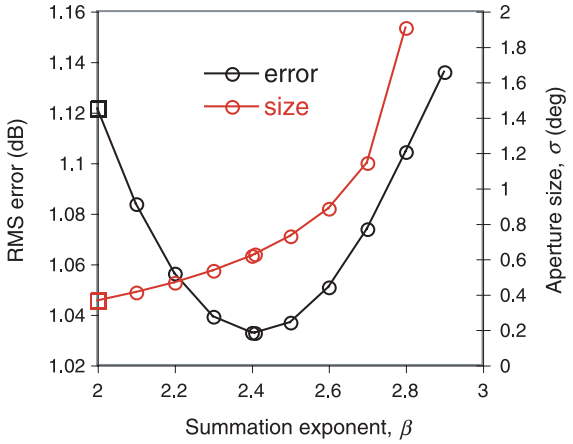


Figure 18. Trade-off between summation exponent $\beta$ and the aperture size, $\sigma$. The value of $\beta$ was fixed and other parameters re-estimated. The estimated value of aperture size $\sigma$ is plotted against the fixed value of $\beta$. Other conditions: HPmH contrast sensitivity filter (CSF), fixed oblique effect, no channels.

the estimated value of size $\sigma$ also increases, so that at a $\beta$ of 3 the aperture is effectively absent. This is further evidence that $\beta$ and $\sigma$ both act to limit the efficiency of spatial summation.

To reiterate, the results show that the visibility of large targets relative to small is less than would be predicted by simple energy summation. This discrepancy can be corrected in two ways: either by using a summation exponent larger than 2 or by introducing a spatial aperture. This observation has both theoretical and practical implications. From a theoretical point of view, it suggests that at least some of the theoretical justification for higher exponents may have been misplaced, and that consequently models (such as template matching) that assume an exponent of 2 may be more tenable than previously supposed. We will return to this point in the discussion.

From a practical point of view, an exponent of 2 lends itself to mathematical and computational efficiencies, and these results suggest it can work almost as well as a higher exponent, provided that a smaller aperture is used.

Although the spatial aperture yields the best model fits, we may ask how it compares to prior estimates of the decline in sensitivity with eccentricity. As noted earlier, this decline is highly dependent upon the spatial frequency of the target. Robson & Graham (1981) show that the decline is approximately 0.5 dB per cycle, independent of frequency. When $\beta$ is free to vary, the average size of the aperture is 0.615 degrees. This corresponds to a decline by a factor of 2 in 0.724 degrees. This rate of decline is consistent with Robson and Graham's rule at a spatial frequency of 16.6 cycles/degree. This is well within the range of ModelFest frequencies, which suggests a compromise between a larger aperture (suitable for lower frequencies) and a smaller one (suitable for higher frequencies).

This outcome may be the result of the absence in the ModelFest data set of any large stimuli. They therefore cannot constrain the summation behavior outside of a degree or two. Combined with the evident reciprocity between $\beta$ and size $\sigma$, a reasonable conclusion is that parameter estimates of either $\beta$ or $\sigma$ should be adopted with caution. Further research will be required to constrain better these two mechanisms for restricting foveal summation.

## Channels

In the preceding sections we considered aspects of the fit of the no-channel model; here we consider models that include channels. Figure 19 shows the RMS error for various combinations of channels, an aperture, and the pooling exponent $\beta$. Before discussing this figure further, we note that if the channels consist of an orthonormal transform, whose individual kernels were orthogonal and whose joint effect has no influence on contrast energy, then when $\beta = 2$, introduction of channels can have no effect. The Gabor channels that we use do not quite meet these conditions, but approximate them, so we should expect little effect of channels when $\beta = 2$. And indeed, the square symbols in Figure 19 confirm this expectation.

The square symbols for $\beta = 2$ also reaffirm the observation made above regarding the trade-off between $\beta$ and the aperture: either an aperture or $\beta > 2$ is required to produce a good fit. If both are absent (solid red squares), the error doubles from about 1 to 2 dB.

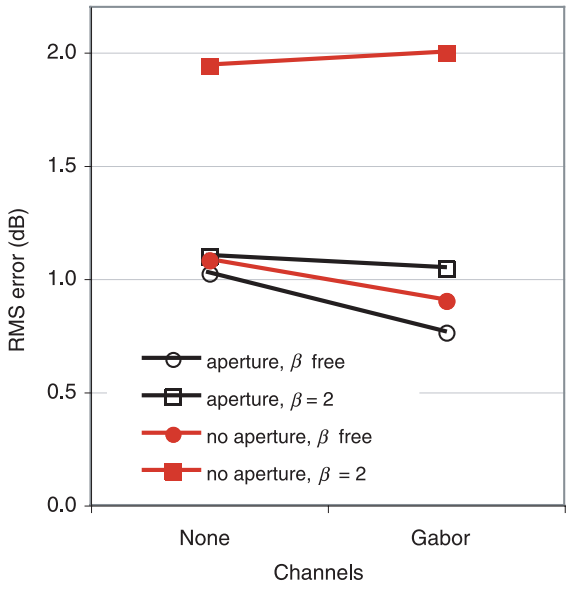


Figure 19. The role of channels, aperture, and pooling exponent $\beta$ in fit of models. Other conditions: fixed oblique effect and LSI contrast sensitivity filter (CSF). All effects of $\beta$, and all differences at $\beta$ = free, are significant at the 0.005 level (Appendix E).

When $\beta$ is free to vary (circular symbols), the addition of channels produces a modest but significant improvement in the fit. The error declines by 0.26 dB when the aperture is present (open red circles), and about 0.18 dB when it is not (solid red circles). The estimated values of $\beta$ are higher when channels are present (2.87 with aperture, 3.40 without) than when they are absent (2.40 with aperture, 2.71 without) and also show some of the trade-off between $\beta$ and aperture.

The results in Figure 19 are for the LSI CSF. The same general pattern is observed for the other CSF functions, although the advantage provided by the channels depends somewhat on the CSF. Error as a function of CSF is plotted in Figure 20 for four models: the channel model with $\beta$ free or fixed at 2, and the no-channel model with $\beta$ free or fixed at 2, in all cases with an aperture. As noted above, when $\beta = 2$, we expect little difference between channel and no-channel models (red and black squares) and this is borne out here.

When $\beta$ is free to vary, addition of channels results in a reduction of error for the best CSFs of about 0.25 dB (black versus red circles).

Some insight into the role of channels in reducing the error is gained from Figure 10, where it can be seen that the biggest change is for stimuli 35 (noise) and 43 (natural image). The advantage of channels for these two stimuli may be that they are broadband, and that channel models correctly exhibit inefficient summation over frequency.

We should note that no effort has been made to optimize the channel parameters of the channel model; the parameters used were consensus values drawn from the literature (Table 2). One aspect of the particular channel model used should also be noted. Although channels are implemented that extend as low as 0.9375 cycles/degree, there is no channel at 0 cycles/degree. Thus, targets such as the Gaussian blobs must be detected by channels centered

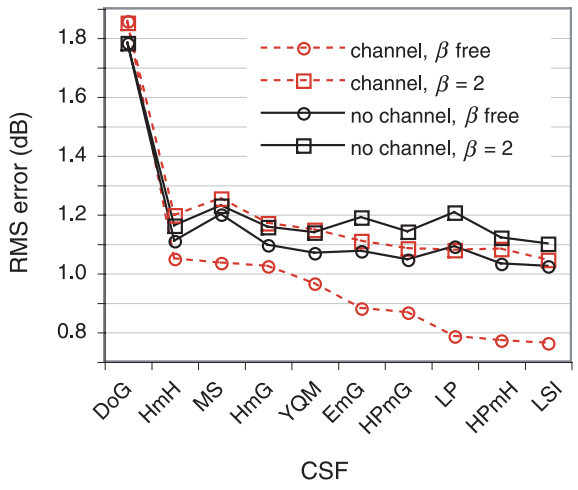


Figure 20. Error for channel and no-channel models. Other conditions: fixed oblique effect and Gaussian aperture.

at nonzero frequencies. It is possible that a channel model with a channel at zero frequency would provide a still better fit.

### Normalized RMS error

To this point we have expressed performance of each model in terms of RMS error. A measure which takes into account the number of parameters is given by the normalized RMS error, defined here as

$$\mathrm{NRMS} = \sqrt{\frac{1}{J-N}\Sigma(t_j - m_j)^2}. \quad (21)$$

where $N$ is the number of parameters of a model. In using this measure, we do not treat the addition of channels or the oblique effect as adding a parameter because no parameters were estimated in those cases. When comparisons are based on this measure, the best-fitting models are generally those with channels, a fixed oblique effect, a Gaussian aperture, and $\beta > 2$. To allow additional comparisons, the fifty conditions yielding the lowest NRMS values are shown in Table 4.

## Discussion

### Theoretical issues

In this discussion we draw a distinction between a metric—by which we mean a particular computational formula for predicting target thresholds—and a model—by which we mean a theoretical conjecture about particular mechanism or mechanisms that play a role in detecting the targets. To this point we have focused exclusively on metrics. Now we attempt to draw some connections between models and metrics.

Models for visual detection have proposed a great variety of mechanisms: optical blurring, cone sampling, transducer noise, transducer nonlinearities, multiple channels of precortical filtering (e.g., magno- versus parvo-cellular and off- versus on-center), and oriented, narrow-band receptive fields at the cortical level. Mechanisms may vary according to eccentricity, and they may include noise with a nonwhite spectrum and noise that is stimulus related. Models have also adopted various mechanisms for categorizing the stimuli into those that contain a signal and those that do not. For example, the observer may make the optimal decision based on the corrupted sensory information available, or may have uncertainties about the stimuli or have other nonoptimal decision processes, such as template noise, noisy category boundaries, and suboptimal summation. In most models, particular mechanisms have been proposed to account for particular empirical results, but the need for the mechanism in the presence of other mechanisms has rarely been demonstrated.

The image-based metrics we have tested can say little about the need for any of the above mechanisms, because other mechanisms can substitute in particular situations, as demonstrated by the trade-off between aperture and spatial summation exponent shown above.

On the other hand, some proposed models do predict that one of the metrics we have tested will predict contrast thresholds. Such models must have contrast sensitivity frequency effects represented by an initial linear filter but may differ as to whether they need channels and what summation rule is required. The latter is usually a variant of four general types: peak detectors, probability summation, energy detection, and template matching.

#### *Channels*

Of the metrics we have evaluated, the best fit is provided by the Gabor channel metric, especially when combined with the Gaussian aperture. We speculate that an even better fit might be provided by an aperture size that differs for each channel (Robson & Graham, 1981). We have noted above that channels may improve the fit by reducing the efficiency of summation over frequency for broadband targets, such as the noise image and natural image.

The channels here may also be helping to account for the effects of unrelated mechanisms. Position uncertainty, for example, causes low-frequency Gabor images to be detected more efficiently than high-frequency Gabor targets of equal size (Ahumada, 2002; Burgess & Ghandeharian, 1984). The channel model can account for this effect through linear summation within a small, high-frequency mechanism and weaker summation across several such mechanisms.

The Gabor channel metric might also mimic the effects of other types of channels, such as line or local edge detectors. Note that while both models of Figure 10 predict that the edge (30) should be more detectable than the line (31), the actual thresholds are nearly identical.

#### *Energy detection*

A metric in which $\beta = 2$ is generically described as an energy model. Energy models predict that at threshold all targets have the same filtered contrast energy. Such models can arise from several different mechanisms. In the energy-only model (Manahilov & Simpson, 2001), targets are filtered, their energy collected, and noise added to account for the variability of detection. Manahilov & Simpson (2001) have shown that the energy-only model is consistent with their data on summation between Gabor patches with frequencies a factor of three apart. As we have shown, the energy metric (without an aperture) is not consistent with the ModelFest data. The energy metric

| Oblique | Aperture | Channels | $\beta$ | CSF | RMS | $N$ | NRMS |
|---|---|---|---|---|---|---|---|
| Fixed | Gaussian | Gabor | Free | HPmH | 0.772 | 7 | 0.844 |
| Fixed | Gaussian | Gabor | Free | LP | 0.787 | 6 | 0.848 |
| Fixed | Gaussian | Gabor | Free | LSI | 0.763 | 13 | 0.914 |
| Fixed | Gaussian | Gabor | Free | HPmG | 0.867 | 7 | 0.948 |
| Fixed | Gaussian | Gabor | Free | EmG | 0.881 | 6 | 0.950 |
| Fixed | Gaussian | Gabor | Free | YQM | 0.966 | 6 | 1.041 |
| Fixed | | Gabor | Free | LSI | 0.903 | 12 | 1.064 |
| Fixed | Gaussian | Gabor | Free | HSmG | 1.024 | 6 | 1.104 |
| Fixed | Gaussian | Gabor | Free | MS | 1.036 | 6 | 1.117 |
| Fixed | Gaussian | | Free | HPmH | 1.033 | 7 | 1.129 |
| Fixed | Gaussian | Gabor | Free | HmH | 1.050 | 6 | 1.132 |
| Fixed | Gaussian | | Free | HPmG | 1.047 | 7 | 1.144 |
| Fixed | Gaussian | Gabor | 2 | LP | 1.080 | 5 | 1.148 |
| Fixed | Gaussian | | Free | YQM | 1.069 | 6 | 1.153 |
| Fixed | Gaussian | | Free | EmG | 1.075 | 6 | 1.159 |
| Fixed | Gaussian | Gabor | 2 | HPmH | 1.083 | 6 | 1.168 |
| Fixed | Gaussian | | Free | LP | 1.092 | 6 | 1.177 |
| Fixed | Gaussian | | Free | HSmG | 1.096 | 6 | 1.181 |
| Fixed | Gaussian | Gabor | 2 | EmG | 1.111 | 5 | 1.181 |
| Fixed | | | Free | HPmH | 1.098 | 6 | 1.183 |
| Fixed | | | Free | HPmG | 1.098 | 6 | 1.184 |
| Fixed | | | Free | EmG | 1.122 | 5 | 1.193 |
| Fixed | Gaussian | | Free | HmH | 1.110 | 6 | 1.197 |
| Fixed | | | Free | LP | 1.133 | 5 | 1.206 |
| Fixed | Gaussian | | 2 | HPmH | 1.122 | 6 | 1.209 |
| Fixed | Gaussian | | 2 | YQM | 1.139 | 5 | 1.211 |
| Fixed | Gaussian | Gabor | 2 | YQM | 1.149 | 5 | 1.222 |
| Fixed | Gaussian | | Free | LSI | 1.024 | 13 | 1.226 |
| Fixed | | | Free | YQM | 1.154 | 5 | 1.227 |
| Fixed | Gaussian | | 2 | HPmG | 1.142 | 6 | 1.231 |
| Fixed | Gaussian | | 2 | HSmG | 1.157 | 5 | 1.231 |
| Fixed | Gaussian | Gabor | 2 | LSI | 1.046 | 12 | 1.232 |
| Fixed | Gaussian | | 2 | HmH | 1.162 | 5 | 1.236 |
| | Gaussian | | Free | HPmH | 1.137 | 7 | 1.242 |
| Fixed | Gaussian | Gabor | 2 | HSmG | 1.171 | 5 | 1.246 |
| Fixed | | | Free | HSmG | 1.189 | 5 | 1.264 |
| Fixed | Gaussian | | 2 | EmG | 1.191 | 5 | 1.266 |
| | Gaussian | | Free | EmG | 1.175 | 6 | 1.267 |
| | Gaussian | | Free | YQM | 1.176 | 6 | 1.267 |
| Fixed | Gaussian | Gabor | 2 | HmH | 1.198 | 5 | 1.274 |
| Fixed | | | Free | LSI | 1.083 | 12 | 1.275 |
| | Gaussian | | Free | LP | 1.184 | 6 | 1.276 |
| Fixed | Gaussian | | 2 | LP | 1.207 | 5 | 1.283 |
| Fixed | | | Free | HmH | 1.215 | 5 | 1.293 |
| Fixed | Gaussian | | Free | MS | 1.201 | 6 | 1.295 |
| Fixed | Gaussian | | 2 | LSI | 1.100 | 12 | 1.296 |
| | Gaussian | | Free | HSmG | 1.203 | 6 | 1.297 |
| Fixed | Gaussian | | 2 | MS | 1.230 | 5 | 1.309 |
| | Gaussian | | Free | HmH | 1.217 | 6 | 1.312 |

Table 4. The fifty conditions yielding the lowest values of NMRS. Empty cells indicate that a component was absent. CSF = contrast sensitivity filter.

with an aperture, while not as good a fit as the channel model, is still consistent with the data.

Manahilov & Simpson's (2001) results were for patches placed 7.5 degrees above fixation and are in disagreement with those from foveal studies (Graham & Nachmias, 1971; Graham & Robson, 1987; Watson, 1982; Watson & Nachmias, 1980). It is possible that their peripheral results, because they are from a more homogeneous region of retina, do not require an aperture.

### Probability summation

The generalized energy metric, without channels but with $\beta > 2$, may be regarded as the prediction of a probability-summation-only model, where the probability of detection is the probability that any of the noisy outputs of the filter is greater than a constant (Quick, 1974; Robson & Graham, 1981). Note that the probability summation model uses the maximum rule (summation with $\beta = \infty$) but the predictive metric for the model has $\beta < \infty$.

With an aperture, this is the best-fitting no-channel metric. However, its advantage over the energy metric with an aperture is modest (Figure 17). And we have noted that this advantage may arise due to the inclusion of complex stimuli for which a template cannot be formed.

### Peak detection

These data provide conclusive evidence that a simple peak detector on the filter output is not a tenable metric for foveal contrast detection. This is evident in the point labeled "CSF" in Figure 9, which shows that CSF filter followed by a peak detector yields an error of about 4.5 dB, over five times the error of the best-fitting model. The Gabor channel metric with a peak detector has an RMS error of 2.43 dB, indicating that even the addition of these other elements cannot redeem the peak detector. Models with peak detectors must include noise before this stage so that the metric beta is less than infinity.

### Template models

Models that are indifferent to the particular target presented have been previously rejected because they do not perform as well as human observers in the presence of noise (Eckstein, Ahumada, & Watson, 1997) and they do not predict the classification images that result (Shimozaki, Eckstein, & Abbey, 2005).

In contrast, template models assume that the observer constructs one or more templates representing the signal and compares the internal representation of the stimulus with the templates.

If the matching rule is computing the dot product of stimulus and a single template and comparing the result with a criterion, and if the noise is additive white Gaussian and if the templates all correlate equally well with their associated signals, then the stimulus energy will predict the detectability. (When the correlations are one, the template model procedure is the ideal observer for a signal known exactly.) That is, like the energy-only model, a template model can predict that thresholds will be at a constant filtered contrast energy.

Perhaps the first explicit description of the template or matched filter as a model of contrast detection is due to Hauske (1974; Hauske, Wolf, & Lupp, 1976), who noted that it could explain results of subthreshold summation of lines, edges, and bars. Recent work on classification images has given additional credence to the template model (Abbey & Eckstein, 2002; Eckstein & Ahumada, 2002; Levi & Klein, 2002; Murray, Bennett, & Sekuler, 2002; Solomon, 2002), as have experiments on perceptual learning (Beard & Ahumada, 1999; Lu & Dosher, 2004).

The template models suggest that there should be divergences from the energy metric in the ModelFest data. As mentioned above, position uncertainty for high-frequency narrow band stimuli essentially means that both the sine and cosine phase Gabor templates are needed and that there will be a drop in performance for these stimuli relative to lower frequency stimuli. It also seems unlikely that the observer would construct a template for the noise stimulus that correlates as well with the noise itself as other templates do with their images, so worse performance would also be expected for the noise stimulus.

When no aperture is included, the energy metric has a significantly worse fit than either channel or generalized energy metrics (Figure 17). When both the energy and the probability summation metrics have an aperture, the energy metric fits slightly worse (see Table 6), and the best-fitting exponent is still greater than 2. Within the template framework, multiple explanations are possible for this result: more templates could be used, resulting in increased uncertainty (Eckstein et al., 1997); the template could correlate worse with the signal image, an effect known by the term sampling efficiency (Legge, Kersten, & Burgess, 1987); or the template might be noisier (McIlhagga & Paakkonen, 1999), possibly from slower learning (Beard & Ahumada, 1999). We conclude that a template model, with an aperture and with caveats regarding the perfection of the templates, is a plausible explanation of the ModelFest results.

Channels, which are supported by physiological data (De Valois, Albrecht, & Thorell, 1982; Ringach, Hawken, & Shapley, 2002), can coexist with a template model: the template is formed through suitable weighting of channel outputs. In this case, the channels would play no role in the detection except perhaps in the template sampling efficiency. One nice feature of the contrast energy metrics is that they can be computed in other domains, such as Fourier, DCT, or wavelets. Even nonorthogonal domains like the Gabor channels give essentially the same results (note the similarity of the $\beta = 2$, channel and no-channel curves in Figure 20).

A template model can also coexist with a probability summation model: for relatively compact and simple stimuli, a template is formed and energy summation is observed, while for complex or dispersed stimuli (over time or space) for which a template cannot be formed, probability summation is observed. This raises the question of limits on template formation, which are a subject for further study.

### *Limitations of ModelFest data set*

It is important to remember that the analyses presented here depend on the selection of stimuli adopted in the ModelFest experiment. For example, we have noted that broadband stimuli are difficult to fit without a channel model; removing these stimuli, or adding many more, could have altered our conclusions. Likewise we have noted that a template model is likely to be less sensitive to complex targets (such as a noise sample); including fewer or more of these would bias conclusions towards or away from such a model.

In addition, the ModelFest stimuli do not adequately test or constrain certain aspects of a complete model of contrast detection. For one, the stimuli were confined to 2.13 × 2.13 degrees and thus do not test spatial summation (or the nature of the aperture) beyond a narrow foveal range. Second, as we have noted above, the data do not constrain the oblique effect very well.

Another drawback of the ModelFest stimuli is that they largely confounded size and eccentricity: stimuli were always centered on fixation, and when enlarged, they grew into the periphery. Furthermore, except for the disk, they were always windowed by a Gaussian aperture. This led to the confounding of the spatial aperture of the metric and the summation exponent in our results. This confound might be removed by, for example, exploring summation among annular rings of a given frequency.

A further omission of the ModelFest group was the failure to report individual trial data. The shape of the psychometric function may be diagnostic for some models. For example, uncertainty is associated with a steepening of the psychometric function. Although single threshold runs do not usually have enough data to estimate psychometric function slopes, good estimates might have been obtained from multiple runs.

Because there are so many mechanisms involved in the possible models, and so few dimensions of variation in the stimuli, measurements of particular mechanism properties must be confounded. The addition of a small amount of background noise would have helped unconfound these measurements (Pelli, Levi, & Chung, 2004). A striking aspect of the data is the 10-dB range of observer sensitivities. We cannot say whether this is mainly due to internal noise variations, sampling efficiency variations, uncertainty variations, or variations in other mechanisms.

Finally, we note that these data do not address visibility of moving or rapidly varying targets, whose sensitivity is known to differ systematically from stationary targets (van Nes & Bouman, 1967), nor do they address variations in sensitivity with changes in background illumination (van Nes, Koenderink, Nas, & Bouman, 1967).

## Contrast sensitivity functions

We have found that a range of specific CSF formulae is about equally good in fitting the average ModelFest data. For well-fitting no-channel models, these functions have a peak value of about 220 at 3.34 cycles/degree ($\beta$ free) or 290 at 3.44 cycles/degree ($\beta = 2$) (Table 5). Note that these values, especially the peak gain, are not independent of the other parameters used in the metric.

Because we have only considered fits to the average data, we make no claims about fits to individual observers. Elsewhere, a curve that fits the average well has been found to be poor at fitting many individuals in a population of older observers (Rohaly & Owsley, 1993).

We also note that the CSF employed here is used to construct a two-dimensional filter that is then used to filter the stimulus image. In addition, it operates in conjunction with other model elements, such as the oblique effect, the spatial aperture, and nonlinear pooling. As such, it cannot be directly compared to previous curve fits to grating or Gabor thresholds on a plot of contrast versus spatial frequency. Nonetheless, we note that the resulting functions (e.g., HPmH) clearly do provide a good curve fit to fixed size Gabor thresholds (Figures 11 and 13).

Some caution is warranted in using the parameters derived here to describe a larger population of observers. The collection of 16 ModelFest observers were not selected to be representative of any particular population. Examination of Figure 2 and of the $RMS_0$ value of 3.46 dB (Equation 2) shows that there is considerable variation with the ModelFest population.

## Standard metrics

A simple metric of the visibility of foveal contrast patterns would be valuable in a wide variety of applications. In such a metric, simplicity of implementation and application must be balanced with accuracy. While the channel metric provides the most accurate results, the magnitude of improvement, especially when compared to the variability among observers (Figure 2), is modest. Further, the channel metric involves considerably more computation (about 3×) and is less robust, in the sense that its performance may depend upon detailed aspects of the implementation—for example, the treatment of borders and the zero frequency signal.

For these reasons we propose a no-channel metric as a standard. In particular, we propose the no-channel metric that includes a Gaussian aperture, the fixed oblique effect, the HPmH CSF, and $\beta > 2$. The parameters of this metric are shown in one line of Table 5 and are la-

beled Standard A. Because it provides nearly as good a fit, and is computationally much simpler, we also provide second Standard B in which $\beta = 2$ (orange highlighting). The user's particular application will determine which of these two standards is appropriate. We hope that these standards will provide useful benchmarks for both future theoretical modeling as well as practical calculations of foveal spatial pattern thresholds.

## Conclusions

We have explored the quality of fit yielded by a wide range of models and parameter settings. We have shown that the ModelFest data set can be fit quite well with relatively simple models. Over the entire set, the residual error for the best model is 0.76 dB. This best model includes a spatial aperture, an oblique effect, a summation exponent of about 2.9, Gabor channels, and a CSF that is a linear interpolation between 11 discrete values.

We have found that a range of specific formulae for the CSF fits nearly as well as the linear interpolation CSF, but with many fewer estimated parameters.

We have found that models which lack channels fit less well, but that the increase in error is small. This suggests that practical models may dispense with the added complexity of channels.

We have found a profound trade-off between the summation exponent $\beta$ and the sizes of the spatial aperture. About the same fit is obtained by including a spatial aperture, or an exponent grater than 2. We interpret this as a consequence of both serving to restrict the efficiency of spatial summation.

The success of models with exponents at or near 2 (when combined with an aperture) provides support for a template model of pattern detection.

We have proposed a particular standard metric for foveal contrast detection, which includes an oblique effect, a specific CSF formula, a Gaussian aperture, a spatial pooling exponent greater than 2, but no channels. We also propose a second metric in which the spatial pooling exponent is equal to 2. We hope these standards will provide useful measures in practical applications and an informative benchmark for theoretical analyses.

## Appendix A: ModelFest stimuli

We have provided a file modelfest-stimuli containing all of the stimuli used in this experiment. The images are provided in a single binary file of $43 \times 256 \times 256 = 2{,}818{,}048$ bytes. The ordering of bytes with respect to pixels is left to right, top to bottom, and image 1 to image 43.

## Appendix B: ModelFest data

The file modelfestbaselinedata.csv contains the ModelFest Baseline Dataset. The file is in text form, in comma-separated-value (CSV) format. This format is easily read by many applications such as Microsoft Excel. The structure of the file is one line per subject. The first value in each line is the observer initials; the remainder of line is 172 numbers, corresponding to four thresholds for each of the 43 stimuli. The numbers indicate the $\log_{10}$ of the threshold contrast and are rounded to three decimal places.

## Appendix C: Notation

Here we provide a summary of the notation used in this paper.

| Term | Definition | Units |
|---|---|---|
| $L(g)$ | Luminance as a function of gray level $g$ | cd/m$^2$ |
| $L_0$ | Mean luminance | cd/m$^2$ |
| $c$ | Contrast | |
| $g$ | Digital graylevel | |
| $t_{s,o,r}$ | Threshold for stimulus $s$, observer $o$, and replication $r$ | dB contrast |
| $O(f,\theta)$ | Oblique effect filter | |
| $f$ | Spatial frequency | cycles/degree |
| $\theta$ | Orientation | radians |
| $\gamma$ | Oblique effect parameter | cycles/degree |
| $\lambda$ | Oblique effect parameter | cycles/degree |
| $A(r)$ | Aperture | |
| $r$ | Radial distance from fixation | degree |
| $\sigma$ | Size (standard deviation) of Gaussian aperture | degree |
| $\beta$ | Summation exponent | |
| $p_x$, $p_y$ | Width and height of each pixel in the input image | degree |
| $N_x$, $N_y$ | Number of rows and columns in the input image | pixels |
| $r_{x,y}$ | Processed pixel value | |
| $S(f)$ | CSF | |
| $f_0$ | High-frequency scale, used in several CSFs | cycles/degree |
| $f_1$ | Low-frequency scale, used in several CSFs | cycles/degree |
| $a$ | Parameters used in several CSFs, usually attenuation at low frequencies | |
| $p$ | Parameter used in several CSFs | |

# Appendix D: Model parameters

Here we provide parameters for the no-channel model with various CSF formulae. We show both metrics in which $\beta$ was estimated and in which it was fixed at 2. Within these two categories, conditions are sorted by NRMS. The proposed standards A and B are highlighted. See the definition of each CSF for a definition of the parameters. We also show two derived parameters for each function: the peak gain (max) and the peak frequency ($f_{max}$).

| Standard | CSF | RMS | NP | NRMS | Gain | $f_0$ | $f_1$ | $a$ | $\beta$ | $p$ | $\sigma$ | $f_{max}$ | max |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | HPmH | 1.0329 | 7 | 1.1288 | 373.08 | 4.1726 | 1.3625 | 0.8493 | 2.4081 | 0.7786 | 0.6273 | 3.45 | 217.3 |
| | HPmG | 1.0468 | 7 | 1.1440 | 289.45 | 5.3459 | 1.9793 | 0.7983 | 2.4054 | 0.8609 | 0.6311 | 3.32 | 221.8 |
| | YQM | 1.0694 | 6 | 1.1529 | 466.38 | 7.0629 | 0.6951 | 7.7712 | 2.3557 | | 0.5790 | 3.32 | 219.8 |
| | EmG | 1.0755 | 6 | 1.1594 | 360.24 | 7.5237 | 1.8972 | 0.8155 | 2.4725 | | 0.7071 | 3.18 | 218.4 |
| | LP | 1.0916 | 6 | 1.1767 | 214.46 | 3.2316 | | 0.7127 | 2.4902 | 0.8081 | 0.7118 | 3.50 | 213.6 |
| | HmG | 1.0959 | 6 | 1.1814 | 258.17 | 6.8432 | 1.7483 | 0.7778 | 2.3277 | | 0.5579 | 3.20 | 225.3 |
| | HmH | 1.1104 | 6 | 1.1970 | 271.71 | 6.7770 | 1.0461 | 0.8082 | 2.2950 | | 0.5311 | 3.39 | 223.8 |
| | MS | 1.2009 | 6 | 1.2946 | 551.29 | 1.7377 | | 1.0465 | 2.3643 | 0.6937 | 0.5702 | 3.06 | 215.0 |
| | DoG | 1.7830 | 6 | 1.9222 | 272.74 | 15.3870 | 1.3456 | 0.7622 | 1.9960 | | 0.3548 | 2.90 | 261.2 |
| B | HPmH | 1.1216 | 6 | 1.2091 | 501.20 | 4.3469 | 1.4476 | 0.8514 | 2 | 0.7929 | 0.3652 | 3.62 | 289.0 |
| | YQM | 1.1387 | 5 | 1.2113 | 621.38 | 7.0856 | 0.7285 | 8.0721 | 2 | | 0.3656 | 3.46 | 284.0 |
| | HPmG | 1.1416 | 6 | 1.2307 | 359.87 | 6.0728 | 1.9505 | 0.7931 | 2 | 0.9186 | 0.3655 | 3.37 | 292.0 |
| | HmG | 1.1572 | 5 | 1.2310 | 329.93 | 6.9248 | 1.8045 | 0.7827 | 2 | | 0.3662 | 3.29 | 286.6 |
| | HmH | 1.1620 | 5 | 1.2361 | 345.78 | 6.7581 | 1.1210 | 0.8128 | 2 | | 0.3688 | 3.52 | 279.4 |
| | EmG | 1.1905 | 5 | 1.2664 | 504.43 | 7.6399 | 1.9788 | 0.8163 | 2 | | 0.3635 | 3.30 | 302.0 |
| | LP | 1.2065 | 5 | 1.2835 | 299.21 | 3.3578 | | 0.7193 | 2 | 0.8009 | 0.3612 | 3.50 | 298.9 |
| | MS | 1.2301 | 5 | 1.3085 | 707.51 | 2.4887 | | 0.9846 | 2 | 0.7748 | 0.3596 | 3.41 | 273.6 |
| | DoG | 1.7830 | 5 | 1.8967 | 271.70 | 15.3852 | 1.3412 | 0.7615 | 2 | | 0.3563 | 2.89 | 260.3 |

Table 5. Parameters and error for the no-channel model with nine contrast sensitivity filter (CSF) functions. Other conditions: fixed oblique effect, Gaussian aperture. The proposed standards A and B are highlighted. NP is number of parameters, max is the maximum value of the CSF, and $f_{max}$ is the frequency at which the peak occurs.

# Appendix E: Statistics

Where two variants of the component model differ only in that in one ($M_1$), a parameter is free to vary while in the other ($M_0$) it is fixed, we say that the two models are nested, in that $M_1$ is a more general version of (and includes) $M_0$. In such cases, it is possible to construct simple statistical tests.

If we write SS for the sum of squares of the residual error for each hypothesis, then the statistic

$$q = \left(\frac{SS_0 - SS_1}{df_0 - df_1}\right) \Big/ \frac{SS_1}{df_1} \tag{A1}$$

will have the $F$ ratio distribution with 1 and $df_1$ degrees of freedom. In the table, we provide $1 - p$ values for various nested comparisons. All except the final test are significant at the .05 level.

| $M_0$ | $M_1$ | $1 - p$ | Channels | Aperture | $\beta$ | CSF | Figure |
|---|---|---|---|---|---|---|---|
| $\beta$ = 2 | $\beta$ free | .039 | No | Yes | | LSI | 17 |
| $\beta$ = 2 | $\beta$ free | .015 | No | Yes | | HPmH | 17 |
| $\beta$ = 2 | $\beta$ free | 0 | Yes | Yes | | LSI | 19 |
| $\beta$ = 2 | $\beta$ free | .039 | No | Yes | | LSI | 19 |
| $\beta$ = 2 | $\beta$ free | 0 | No | No | | LSI | 19 |
| $\beta$ = 2 | $\beta$ free | 0 | Yes | No | | LSI | 19 |
| $\sigma$ = inf | $\sigma$ free | 0 | No | | 2 | LSI | 17 |
| $\sigma$ = inf | $\sigma$ free | .002 | Yes | | Free | LSI | 19 |
| $\sigma$ = inf | $\sigma$ free | .07 | No | | Free | LSI | 19 |

Table 6. Statistical tests of nested comparisons. Columns 4–7 indicate features shared by both $M_0$ and $M_1$. Column 8 indicates the figure in which the comparison may be found. CSF = contrast sensitivity filter.

## Acknowledgments

We thank the members of the ModelFest group for collection of the data and for useful discussions. We thank Ellen Salud for expert assistance in production of this paper. This research was supported by NASA Research Grants 711-80-03 and 131-20-30-00.

Commercial relationships: ABW has an intellectual property interest in standard observer metrics.
Corresponding author: Andrew B. Watson.
Email: andrew.b.watson@nasa.gov.
Address: MS 262-2 NASA Ames Research Center, Moffett Field, CA 94035-1000, USA.

## References

Abbey, C. K., & Eckstein, M. P. (2002). Classification image analysis: Estimation and statistical inference for two-alternative forced-choice experiments. *Journal of Vision, 2*(1), 66–78, http://journalofvision.org/2/1/5/, doi:10.1167/2.1.5. [PubMed] [Article]

Ahumada, A. J., Jr. (2002). Classification image weights and internal noise level estimation. *Journal of Vision, 2*(1), 121–131, http://journalofvision.org/2/1/8/, doi:10.1167/2.1.8. [PubMed] [Article]

Ahumada, A. J., Jr., & Peterson, H. A. (1992). Luminance-model-based DCT quantization for color image compression [Abstract]. In B. E. Rogowitz (Ed.), *Human vision, visual processing, and digital display III* (Vol. 1666, pp. 365–374): Proceedings of the SPIE.

Beard, B. L., & Ahumada, A. J., Jr. (1999). Detection in fixed and random noise in foveal and parafoveal vision explained by template learning. *Journal of the Optical Society of America A, 16*(3), 755–763. [PubMed]

Berkley, M. A., Kitterle, F., & Watkins, D. W. (1975). Grating visibility as a function of orientation and retinal eccentricity. *Vision Research, 15,* 239–244. [PubMed]

Blakemore, C., & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology, London, 203,* 237–260. [PubMed]

Burgess, A., & Ghandeharian, H. (1984). Visual signal detection: I. Ability to use phase information. *Journal of the Optical Society of America A, 1*(8), 900–905. [PubMed]

Campbell, F. W., Carpenter, R. H. S., & Levinson, J. Z. (1969). Visibility of aperiodic patterns compared with that of sinusoidal gratings. *Journal of Physiology, 204,* 283–298. [PubMed]

Campbell, F. W., Kulikowski, J. J., & Levinson, J. Z. (1966). The effect of orientation on the visual resolution of gratings. *Journal of Physiology, 187,* 427–436. [PubMed]

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology, London, 197,* 551–566. [PubMed]

Carney, T., Klein, S. A., Tyler, C. W., Silverstein, A. D., Beutter, B., Levi, D., et al. (1999). The development of an image/threshold database for designing and testing human vision models [Abstract]. In B. Rogowitz & T. Pappas (Eds.), *Human vision, visual processing, and digital display IV* (Vol. 3644, pp. 542–551): SPIE, Bellingham, WA.

Carney, T., Tyler, C. W., Watson, A. B., Makous, W., Beutter, B., Chen, C.-C., et al. (2000). ModelFest: Year one results and plans for future years [Abstract]. In B. Rogowitz & T. Pappas (Eds.), *Human vision, visual processing, and digital display V* (Vol. 3959, pp. 140–151): SPIE, Bellingham, WA.

Chen, C.-C., & Tyler, C. W. (2000). ModelFest: Imaging the underlying channel structure [Abstract]. In B. Rogowitz & T. Pappas (Eds.), *Human vision, visual processing, and digital display V* (Vol. 3959, pp. 152–159): SPIE, Bellingham, WA.

De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in Macaque visual cortex. *Vision Research, 22,* 545–559. [PubMed]

Eckstein, M. P., & Ahumada, A. J., Jr. (2002). Classification images: A tool to analyze visual strategies. *Journal of Vision, 2*(1), 1x, http://journalofvision.org/2/1/i/, doi:10.1167/2.1.i. [PubMed] [Article]

Eckstein, M. P., Ahumada, A. J., Jr., & Watson, A. B. (1997). Visual signal detection in structured backgrounds: II. Effects of contrast gain control, background variations and white noise. *Journal of the Optical Society A, 14*(9), 2406–2419. [PubMed]

Enroth-Cugell, C., & Robson, J. G. (1966). The contrast sensitivity of retinal ganglion cells of the cat. *Journal of Physiology, (London), 187,* 517–552.

Enroth-Cugell, C., Robson, J. G., Schweitzer-Tong, D., & Watson, A. B. (1983). Spatio-temporal interactions in cat retinal ganglion cells showing linear spatial summation. *Journal of Physiology, (London), 341,* 279–307.

Loehle Enterprises. (2004). Global Optimization 4.3 for Mathematica. [Link]

Graham, C. H., & Margaria, R. (1935). Area and intensity-time relation in the peripheral retina. *American Journal of Physiology, 113,* 299–305.

Graham, N. (1977). Visual detection of aperiodic spatial stimuli by probability summation among narrowband channels. *Vision Research, 17*(5), 637–652. [PubMed]

Graham, N., & Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: A comparison of single-channel and multiple-channels models. *Vision Research, 11*(3), 251–259. [PubMed]

Graham, N., & Robson, J. G. (1987). Summation of very close spatial frequencies: The importance of spatial probability summation. *Vision Research, 27*(11), 1997–2007. [PubMed]

Hauske, G. (1974). Adaptive filter mechanisms in human vision. *Kybernetik, 16*(4), 227–237. [PubMed]

Hauske, G., Wolf, W., & Lupp, U. (1976). Matched filters in human vision. *Biological Cybernetics, 22*(4), 181–188. [PubMed]

Legge, G. E., Kersten, D., & Burgess, A. E. (1987). Contrast discrimination in noise. *Journal of the Optical Society of America A, 4,* 391–404. [PubMed]

Levi, D. M., & Klein, S. A. (2002). Classification images for detection and position discrimination in the fovea and parafovea. *Journal of Vision, 2*(1), 46–65, http://journalofvision.org/2/1/4/, doi:10.1167/2.1.4. [PubMed] [Article]

Lu, Z. L., & Dosher, B. A. (2004). Perceptual learning retunes the perceptual template in foveal orientation identification. *Journal of Vision, 4*(1), 44–56, http://journalofvision.org/4/1/5/, doi:10.1167/4.1.5. [PubMed] [Article]

Manahilov, V., & Simpson, W. A. (2001). Energy model for contrast detection: Spatial-frequency and orientation selectivity in grating summation. *Vision Research, 41*(12), 1547–1560. [PubMed]

Mannos, J. L., & Sakrison, D. J. (1974). The effects of a visual fidelity criterion on the encoding of images. *IEEE Transactions on Information Theory, IT-20*(4), 525–526.

McIlhagga, W., & Paakkonen, A. (1999). Noisy templates explain area summation. *Vision Research, 39*(2), 367–372. [PubMed]

McMahon, M. J., & MacLeod, D. I. (2003). The origin of the oblique effect examined with pattern adaptation and masking. *Journal of Vision, 3*(3), 230–239, http://journalofvision.org/3/3/4/, doi:10.1167/3.3.4. [PubMed] [Article]

Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2002). Optimal methods for calculating classification images: Weighted sums. *Journal of Vision, 2*(1), 79–104, http://journalofvision.org/2/1/6/, doi:10.1167/2.1.6. [PubMed] [Article]

Pelli, D. G., Levi, D. M., & Chung, S. T. (2004). Using visual noise to characterize amblyopic letter identification. *Journal of Vision, 4*(10), 904–920, http://journalofvision.org/4/10/6/, doi:10.1167/4.10.6. [PubMed] [Article]

Quick, R. F. (1974). A vector magnitude model of contrast detection. *Kybernetik, 16,* 65–67. [PubMed]

Ringach, D. L., Hawken, M. J., & Shapley, R. (2002). Receptive field structure of neurons in monkey primary visual cortex revealed by stimulation with natural image sequences. *Journal of Vision, 2*(1), 12–24, http://journalofvision.org/2/1/2/, doi:10.1167/2.1.2. [PubMed] [Article]

Robson, J. G., & Graham, N. (1981). Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Research, 21*(3), 409–418. [PubMed]

Rodieck, R. W. (1965). Quantitative analysis of cat retinal ganglion cell responses to visual stimuli. *Vision Research, 5,* 583–601. [PubMed]

Rohaly, A. M., & Owsley, C. (1993). Modeling the contrast-sensitivity functions of older adults. *Journal of the Optical Society of America A, 10*(7), 1591–1599. [PubMed]

Shimozaki, S. S., Eckstein, M. P., & Abbey, C. K. (2005). Spatial profiles of local and nonlocal effects upon contrast detection/discrimination from classification images. *Journal of Vision, 5*(1), 45–57, http://journalofvision.org/5/1/5/, doi:10.1167/5.1.5. [PubMed] [Article]

Shlaer, S. (1937). The relation between visual acuity and illumination. *Journal of General Physiology, 21,* 165–188.

Solomon, J. A. (2002). Noise reveals visual mechanisms of detection and discrimination. *Journal of Vision, 2*(1), 105–120, http://journalofvision.org/2/1/2/, doi:10.1167/2.1.2. [PubMed] [Article]

van Nes, F. L., & Bouman, M. A. (1967). Spatial modulation transfer in the human eye. *Journal of the Optical Society of America, 57,* 401–406.

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

Walker, L., Klein, S. A., & Carney, T. (1999). *Modeling the ModelFest data: Decoupling probability summation.* Paper presented at the Optical Society of America Annual Meeting, Santa Clara, CA.

Watson, A. B. (1979). Probability summation over time. *Vision Research, 19,* 515–522. [PubMed]

Watson, A. B. (1982). Summation of grating patches indicates many types of detector at one retinal location. *Vision Research, 22,* 17–25. [PubMed]

Watson, A. B. (1987). Estimation of local spatial scale. *Journal of the Optical Society A, 4,* 1579–1582. [PubMed] [Article]

Watson, A. B. (1999). ModelFest Web Site. Retrieved from http://vision.arc.nasa.gov/modelfest/.

Watson, A. B. (2000). Visual detection of spatial contrast patterns: Evaluation of five simple models. *Optics Express, 6*(1), 12–33. [PubMed] [Article]

Watson, A. B., Barlow, H. B., & Robson, J. G. (1983). What does the eye see best? *Nature, 302*(5907), 419–422. [PubMed] [Article]

Watson, A. B., & Nachmias, J. (1980). Summation of asynchronous gratings. *Vision Research, 20,* 91–94. [PubMed]

Watson, A. B., & Solomon, J. A. (1997). Model of visual contrast gain control and pattern masking. *Journal of the Optical Society A, 14,* 2379–2391. [PubMed] [Article]

Watson, A. B., Borthwick, R., & Taylor, M. (1997). Image quality and entropy masking [Abstract]. In B. Rogowitz & T. Pappas (Eds.), *Human vision and electronic imaging II* (Vol. 3016, pp. 2–12): SPIE, Bellingham, WA.

Wolfram, S. (2003). *The mathematica book,* 5th ed. Champaign, IL: Wolfram Media.

Yang, J., Qi, X., & Makous, W. (1995). Zero frequency masking and a model of contrast sensitivity. *Vision Research, 35*(14), 1965–1978. [PubMed]