Vision Research 81 (2013) 58–68

Contents lists available at SciVerse ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Inconsistent channel bandwidth estimates suggest winner-take-all nonlinearity in second-order vision

Zachary M. Westrick [a], Christopher A. Henry [b], Michael S. Landy [a,b,*]

[a] Department of Psychology, New York University, USA
[b] Center for Neural Science, New York University, USA

ARTICLE INFO

Article history:
Received 7 November 2012
Received in revised form 23 January 2013
Available online 14 February 2013

Keywords:
Second-order vision
Texture

ABSTRACT

The processing of texture patterns has been characterized by a model that postulates a first-stage linear filter to highlight a component texture, a pointwise rectification stage to convert contrast for the highlighted texture into mean response strength, followed by a second-stage linear filter to detect the texture-defined pattern. We estimated the spatial-frequency bandwidth of the second-stage filter mediating orientation discrimination of orientation-modulated second-order gratings by measuring threshold elevation in the presence of filtered noise added to the modulation signal. This experiment yielded no evidence for frequency tuning. A second experiment, in which subjects had to detect similar second-order gratings while judging their modulation frequency, produced bandwidth estimates of 1–1.5 octaves, similar to estimated bandwidths of first-order channels. We propose that an additional dominant-response-selection nonlinearity can account for these apparently contradictory results.

1. Introduction

The detection and discrimination of luminance-defined patterns has been extensively studied and is well characterized by models that postulate a set of channels tuned for orientation and spatial frequency (De Valois & De Valois, 1988; Graham, 1989). These channels operate essentially as linear filters, and are therefore suited to analysis in terms of their filter properties. This linear account, while broadly applicable, is unable to explain certain aspects of vision. Second-order images, stimuli in which information is encoded by the modulation of something other than average local luminance, are invisible to any purely linear mechanism.

Consider the case of an image that is defined by regions of high and low contrast (Fig. 1A), rather than high and low luminance. A linear filter cannot respond to these contrast modulations because the average luminance is identical in high- and low-contrast image regions. On the other hand, rectified responses of a linear filter tuned to the carrier texture will differ between high- and low-contrast image regions. This observation has led to a now-standard model consisting of a first-stage linear filter, a nonlinear stage (rectification), and a second-stage filter sensitive to the second-order image modulations (Fig. 1B). Thus, one can characterize the filter-rectify-filter (FRF) model in terms of the tuning of the second-stage filter (Bergen & Adelson, 1988; Hallum, Landy, & Heeger, 2011; Kwan & Regan, 1998; Landy & Bergen, 1991; Landy & Oruç, 2002; Larsson, Landy, & Heeger, 2006; Sutter, Beck, & Graham, 1989; Sutter, Sperling, & Chubb, 1995) just as the study of spatial frequency channels concentrated on the tuning of first-stage filters (De Valois & De Valois, 1988; Graham, 1989).

Here, we measure the spatial-frequency bandwidth of second-stage filters using two different techniques. First, we use critical-band-masking, in which threshold for detection of second-order modulation is measured in the presence of low- or high-pass filtered second-order noise (Oruç, Landy, & Pelli, 2006; Solomon & Pelli, 1994). Second, with a task modeled after Watson and Robson's (1981) labelled-lines experiment, thresholds were measured and compared for detection vs. discrimination of two gratings differing in spatial frequency. While this second experiment yielded channel bandwidths on the order of an octave, consistent with previous estimates of both second- and first-stage filters, the critical-band-masking results suggest that the second-stage filter bandwidths are far wider. We propose an extension of the FRF model that includes an additional nonlinear image-segmentation stage, based on second-order vision models proposed by Malik and Perona (1990) and Motoyoshi and Nishida (2004), to resolve the apparent conflict between these results.

2. Experiment 1

2.1. Introduction

In this experiment we measured second-order channel bandwidth using critical-band masking, in which thresholds for the

* Corresponding author. Address: 6 Washington Place, 9th Floor, New York, NY 10003, USA. Fax: +1 212 995 4018.
E-mail address: landy@nyu.edu (M.S. Landy).

0042-6989/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.visres.2013.01.010

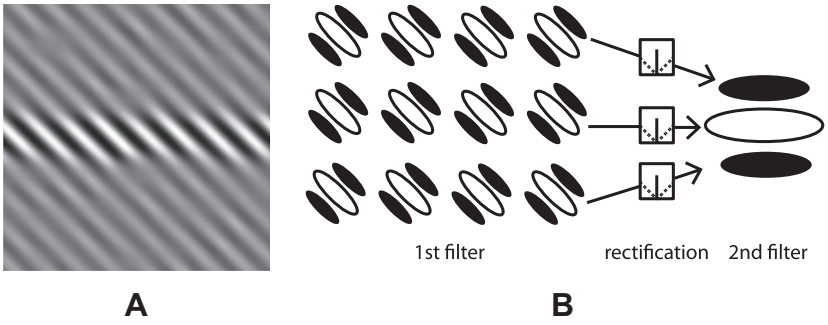

Fig. 1. Second-order processing. (A) Example contrast-defined image. (B) Typical filter-rectify-filter (FRF) model. The first stage consists of processing by linear filters across space sensitive to one of the image's carrier textures. These responses are then rectified, creating a texture-intensity image. Finally, this texture-intensity image is processed by typical spatial-frequency and orientation tuned linear filters to detect the modulator.

detection of a second-order grating are measured both in the absence of noise and in the presence of either high- or low-pass noise with a range of cutoff frequencies. As the noise cutoff approaches the center of the channel being measured, more of it is passed through, increasing threshold. The results are plotted as threshold elevation, relative to a no-noise condition, as a function of noise cut-off. Channel gain is proportional to the slope of the threshold-elevation curve (Solomon & Pelli, 1994). High- and low-pass noise maskers provide independent measures of channel shape. This technique has been used to estimate the tuning of both first- (Henning, Hertz, & Hinton, 1981; Majaj et al., 2002; Solomon & Pelli, 1994) and second-stage visual filters (Oruç, Landy, & Pelli, 2006).

For grating detection, critical-band masking makes a fairly unambiguous prediction. Because useful information only exists at a single spatial frequency, the optimal computation is to use a matched filter, cross correlating the noisy stimulus with the expected grating stimulus (Green & Swets, 1966). If the observer is constrained to use built-in, octave-wide channels, the channel centered on the stimulus spatial frequency will be the most useful, as it is most sensitive to the signal. Thus, we predict that the channels estimated by critical-band masking will be centered on the modulator spatial frequency with bandwidth equal to the bandwidth of the underlying second-stage filters, or perhaps slightly wider if the observers pool off-frequency, less-sensitive channels as well.

2.2. Methods

2.2.1. Stimuli

Stimuli consisted of orientation–modulated second-order sine wave gratings oriented at 0° or 90°. Orientation modulation was produced by adding two opposite-phase contrast-modulated gratings with orthogonal carriers, so that there was no overall contrast-modulation signal when averaged across carrier orientation. This stimulus design prevents an early pointwise contrast nonlinearity, preceding filtering for orientation and spatial-frequency, from producing 1st-order artifacts visible to typical V1 filters (Graham, 1994; Scott-Samuel & Georgeson, 1999). Carriers $C_{45}(x,y)$ and $C_{135}(x,y)$ were oblique sine wave gratings oriented at 45° or 135° with spatial frequency f_c (Fig. 2). Modulators $M_0(x,y)$ and $M_{90}(x,y)$ were horizontal or vertical sine wave gratings of spatial frequency f_m masked by additive "plaid" noise $N_{(l,h)}(x,y)$. $N_{(l,h)}(x,y)$ consisted of Gaussian noise bandpass filtered between cutoffs l and h and with all components removed except those oriented at 0° and 90°. This noisy plaid was used as masker so as to get the maximum masking effect for the minimum masker energy. Fig. 2 shows a schematic example of stimulus construction. "High-pass" noise was actually band-pass, ranging from the low cutoff to half the carrier frequency ($f_c/2$) to avoid first-order artifacts that arise when the modulator is too high in spatial frequency relative to the carrier. Because this cutoff was higher than the highest tested modulator spatial frequency by a factor of two, we assume the highest frequency noise will fall outside of any reasonable second-order channel so that this band-pass noise is effectively high-pass relative to a band-pass second-order channel centered on the target.

Noise spectral density was held constant across each experiment, so that total noise power for $N_{(l,h)}$ was $\frac{h-l}{H-L}c_n$ for full-bandwidth noise power c_n where H and L are the high- and low-frequency cutoffs of the broadest noise condition. After the addition of noise it was necessary to clip a small number of modulator pixels that fell outside the range $(-1,1)$. Subjects first ran a pilot experiment to determine a full-bandwidth noise power that maintained this clipping rate below 5% on average. On trials near threshold the clipping rate was much lower. Stimulus $S(x,y)$ was defined as follows:

$$\begin{aligned} M(x,y) &= N_{(l,h)} + \begin{cases} c_m sin(f_m x + \phi) & \text{if } M = M_0 \\ c_m sin(f_m y + \phi) & \text{if } M = M_{90} \end{cases} \\ m_1(x,y) &= C_{45}(x,y)\sqrt{\frac{1}{2}(1 + M(x,y))} \\ m_2(x,y) &= C_{135}(x,y)\sqrt{\frac{1}{2}(1 - M(x,y))} \\ S(x,y) &= 1 + m_1(x,y) + m_2(x,y) \end{aligned} \tag{1}$$

Modulator phase ϕ was chosen randomly on each trial, and modulator contrast c_m was varied to determine modulation threshold. The stimulus was linearly scaled to occupy the luminance range of our monitor. The square root in Eq. (1) ensured that local modulator contrast power was constant across the image. This square root produces distortions that could potentially become visible as first-order artifacts, but we have shown, using phase randomized stimuli, that any first-order artifact is not detectable for modulator spatial frequencies at or below one-half the carrier spatial frequency (Landy & Oruç, 2002). For this reason we constrain ourselves to modulation targets at those frequencies for this experiment.

Many experiments using orientation modulation have used bandpass-filtered noise carriers (e.g., Landy & Oruç, 2002). Here, pure sine waves were used as carriers instead of filtered noise because bandpass filtered noise contains contrast modulation signals with bandwidth inversely proportional to the bandwidth of the filter used to produce the noise (Kovács & Fehér, 1997). These modulations, akin to beat frequencies, do not show up in the spectrum of the image but have been shown to mask the detection of first-order signals (Derrington & Badcock, 1986; Nachmias & Rogowitz, 1983). We were concerned that these contrast modulations would act as an additional mask for the second-order modulator signal, and hence used sine wave carriers instead (also used by Oruç, Landy, and Pelli (2006)).

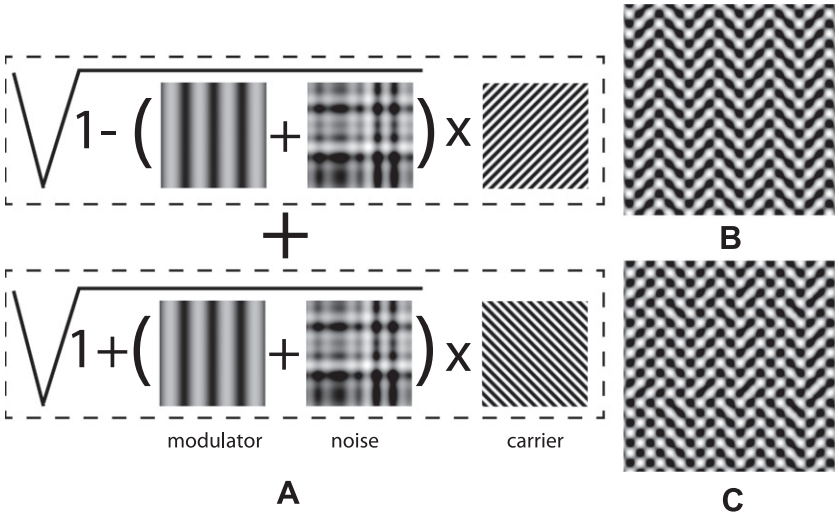

Fig. 2. Stimulus construction diagram (A). The sinusoidal target modulation and plaid noise used to define the stimuli are added to produce two out of phase modulation signals, multiplied by each carrier, and the results are added to produce an orientation-modulated grating. Example stimuli with no modulation noise (B) and with broadband modulation noise (C).

All stimuli were displayed at 85 Hz on a computer using a Diamond Pro 900u CRT monitor with a linearized lookup table and average luminance of 17.5 cd/m^2. Stimuli were 500 × 500 pixels viewed from a distance of 64.5 cm so that stimuli subtended 15° × 15 °.

2.2.2. Procedure

Subjects were required to judge the orientation of the second-order grating in a 2-alternative forced-choice procedure. Each trial consisted of a 500 ms fixation display consisting of a small cross on a mean-luminance background followed by a 250 ms stimulus presentation, after which the subject indicated the orientation (vertical or horizontal) of the second-order grating with a keypress. Feedback was provided on each trial, and subject's responses triggered the onset of the next trial.

Blocks consisted of 100 trials at a single noise cutoff for either high- or low-pass noise. Modulation contrast c_m was varied using two interleaved staircases of 50 trials each. One staircase was 1-up-2-down, converging to 71% correct, while the other was 1-up-3-down, converging to 79% correct. One in five trials was a reminder trial run at high modulation contrast without noise to ensure that the subject remained aware of the second-order grating's appearance. A single session consisted of between 1 and 3 blocks at each noise cutoff level for both high- and low-pass masking noise. For each subject and condition, data were pooled across sessions, so that each dataset consisted of at least two repetitions of each noise cutoff.

Psychometric functions—percentage correct as a function of modulation contrast—were fit with a Weibull function by maximum likelihood to each condition pooled across blocks. Threshold was defined as the modulation contrast at which the psychometric functions reached 75% correct. We then plotted threshold elevations, relative to the no-noise condition, as a function of cutoff frequency. These two threshold-elevation curves were fit simultaneously with a cumulative Gaussian function. Under the critical-band-masking paradigm the derivative of this cumulative Gaussian should yield the second-order channel tuning function.

2.2.3. Subjects

Three subjects (two authors) took part in this experiment. All had normal or corrected-to-normal vision. The Institutional Review Board at New York University approved the experimental procedures and all participants gave informed consent. Table 1 shows the conditions in which each subject took part.

2.3. Results

Previous studies suggest that there are multiple second-order channels and that these channels are tuned for spatial frequency with bandwidths of around one octave (Arsenault & Kingdom, 1999; Ellemberg, Allen, & Hess, 2006; Landy & Oruç, 2002; Oruç, Landy, & Pelli, 2006). Our data are inconsistent with this observation. In Fig. 3 we show data from one subject in three different modulator frequency conditions, all with carrier frequency of 4 cycle/deg. (Note that some of these data have been presented previously and published in abstract form (Landy & Henry, 2007).) At first glance the derived channels, as fit by a cumulative Gaussian, seem to be very broad and are centered in the middle of the range of noise cutoff frequencies, independent of modulator spatial frequency. The data, however, look like they could be fit just as well by an upward sloping straight line, implying a lack of frequency tuning. Fig. 4 shows average data for conditions in which the modulator was one eighth the carrier frequency (e.g., f_c = 4 cycle/deg, f_m = 0.5 cycle/deg). These pooled data also appear linear, with thresholds continuing to rise for noise cutoffs well above the modulator frequency.

Table 1
Results of model comparison between the simple FRF model and the labelling FRF model for all subjects and conditions. The second and third columns show the r^2 ($1 - SS_{model}/SS_{tot}$) values computed for the simple FRF and labelling model respectively, while the fourth column shows the likelihood ratio computed for each model. Negative r^2 values reflect that, for those conditions, the Gaussian channel model predicted by simple FRF fit exceptionally poorly and was outpredicted by the mean of the data. The likelihood ratios reflect strong evidence for the labelling model.

Subject carrier/modulator spatial frequency	r^2 FRF	r^2 labelling	Likelihood ratio
CAH 4/0.25	0.55	0.83	>1000
CAH 4/0.5	0.51	0.96	>1000
CAH 4/1.0	0.73	0.91	212
CAH 8/2	0.76	0.80	4.6
CAH 2/0.25	0.58	0.91	>1000
SYD 2/0.25	0.27	0.80	>1000
SYD 4/0.5	−0.01	0.61	>1000
ZMW 4/0.25	−0.69	0.68	>1000
ZMW 4/0.5	0.06	0.68	>1000

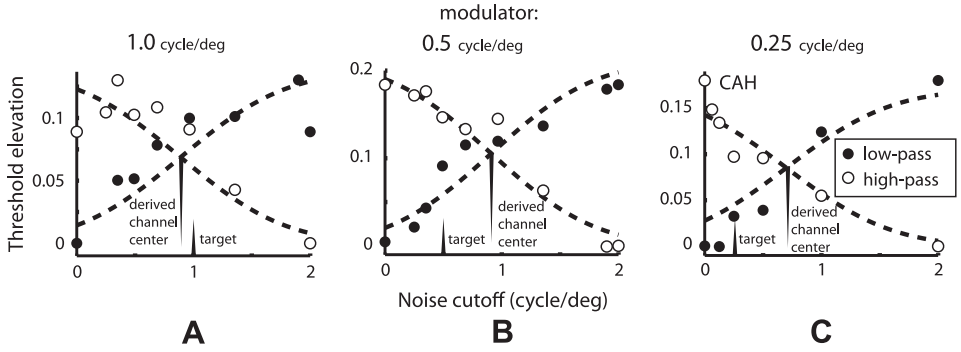

Fig. 3. Threshold elevation as a function of noise cutoff frequency for three different modulation frequencies for subject CAH. Data from other subjects shown in Fig. 4. For a 4 cycle/deg carrier, (A–C) show derived channels for modulation frequencies of 1.0, 0.5, and 0.25 cycle/deg, respectively. All three conditions yield estimated channel centers near one-quarter the carrier frequency, regardless of modulation frequency. The best fitting channel center and bandwidths are (A) 0.89 cycle/deg, 2.3 octaves; (B) 0.91 cycle/deg, 2.2 octaves; (C) 0.7 cycle/deg, 3.6 octaves. The data in (B) and (C) are very poorly described by a channel centered on the modulator target, regardless of bandwidth. These threshold elevation curves could be fit by straight lines, indicating a lack of frequency tuning.

Our data represent a departure from the predictions of the FRF model assuming second-stage filters centered on the modulator target. To demonstrate this for all of our subjects, we apply the following simple test. Under the FRF model, as long as the second-stage filter is centered on the modulation target and has reasonable bandwidth, noise cutoffs more than twice the frequency of the modulator should fall well outside the channel and therefore noise added at those cutoffs should not elevate threshold. Flat threshold elevation means that the correlation between noise cutoff and threshold should be zero for these outside-channel datapoints. Fig. 5 shows estimates with bootstrapped error bars of the correlation between outside-channel noise cutoffs (more than 1.9× modulator frequency) and threshold elevation. Across all subjects and for all relevant (>2 outside-channel cutoffs) conditions the outside-channel correlations are very high and are significantly different from zero, demonstrating that subjects are sensitive to noise outside of even a very broad channel, which is inconsistent with the FRF model.

Why are the critical-band masking data so much in conflict with our predictions? Although first-order patterns can be masked by very distant spatial frequencies, this only occurs when the masker power is much greater than the power of the signal to be detected (Blackwell, 1998), presumably due to cross-frequency normalization. This was not the case in our masking experiment. We consider the following possibilities:

1. The second-order critical-band masking technique is ineffective at measuring second-order channels for an observer using FRF to discriminate second-order orientations.
2. Second-stage filters are not tuned for spatial frequency at all.
3. Second-stage filters are very broadly tuned and centered on approximately one-quarter the carrier frequency.
4. The FRF model's prediction of quasi-linear response to second-order contrast is incorrect, and additional stages of nonlinear processing need to be considered.

The first possibility we consider is that the critical-band masking technique cannot estimate the second-stage filter even if the FRF model accurately describes the computation carried out by the observer. This experimental method could fail for at least two reasons. First, the method assumes the observer makes the detection decision based on the output of a single second-order channel centered on the target spatial frequency independent of the masking condition. For critical-band masking, it is in the observer's best interests to shift the channel they use to detect the target grating, moving it away from the noise to achieve a higher signal-to-noise ratio, thus shifting to a channel centered on a higher spatial frequency with low-pass noise, and a lower spatial frequency

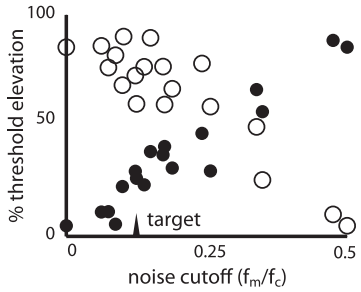

Fig. 4. Threshold elevation vs. noise cutoff. This plot comprises five sets of threshold-elevation data from three subjects, and consists of all conditions collected with a modulator spatial-frequency of one-eighth the carrier spatial frequency. Threshold elevations for each condition were normalized by dividing by the maximal threshold elevation. Noise cutoff frequencies were normalized by the carrier frequency, so that the modulation target is at a normalized spatial frequency of 0.125. Finally, identical conditions were averaged. Subjects were sensitive to noise with much higher frequencies than the target modulation frequency, and the data seem to be better described by a linear fit than a cumulative Gaussian. See Table 1 for a full listing of subjects and conditions.

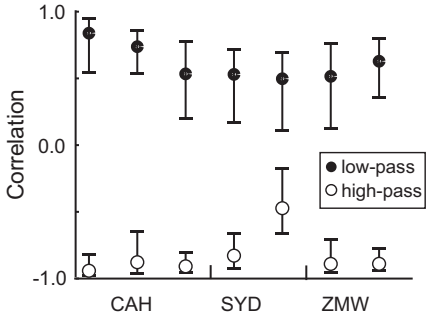

Fig. 5. Estimates with bootstrapped error bars of correlation between noise cutoff with threshold elevation for noise cutoffs $\geqslant 1.9\times$ the spatial-frequency of the modulator target. We propose that these cutoffs are outside any reasonably narrow second-order channel, so that the FRF model predicts a slope, and therefore correlation, of 0. Data shown for all subjects and for all conditions in which more than two noise cutoffs were available (i.e., for low-frequency modulators).

for high-pass noise. This strategy is known as "off-frequency looking" and has been demonstrated for the detection of first-order gratings (Pelli, 1981; Perkins & Landy, 1991). If the observer uses an off-frequency-looking strategy, the estimated channel will be centered on a higher spatial frequency for the low-pass noise conditions than for the high-pass noise conditions. This effect is not

found in our data nor for an experiment in which critical-band masking was used with second-order letters (Oruç, Landy, & Pelli, 2006). The estimate of the second-order channel tuning also assumes near linearity, i.e., that the modulator is de-modulated by the combination of the first-stage filter and nonlinearity, and then processed linearly by the second-stage filter. Because the rectification stage of the FRF model is indeed a nonlinearity, this assumption may be violated, but it is approximately true if the rectification stage consists of a pointwise square or absolute value. In a subsequent section (Section 4.2) we show by simulation that critical-band masking should precisely reveal the tuning of the second-stage filter despite the nonlinear rectification stage.

The second possibility is that the mechanisms underlying second-order vision are, in fact, not frequency tuned. This hypothesis contradicts a number of studies. Several psychophysical studies show that second-order mechanisms are tuned for spatial frequency (Arsenault & Kingdom, 1999; Ellemberg, Allen, & Hess, 2006; Landy & Oruç, 2002; Sutter, Sperling, & Chubb, 1995). In addition, fMRI experiments measuring adaptation to second-order textures have found evidence for both orientation (Larsson, Landy, & Heeger, 2006) and frequency tuning (Hallum, Landy, & Heeger, 2011) in several cortical areas. Finally, single-cell physiological studies in macaque V1 find frequency-tuned responses to contrast modulation (Tanaka & Ohzawa, 2009).

The third possibility is that the mechanisms are broadly tuned with a fixed preferred spatial frequency relative to the carrier. This is also unlikely given prior work. Much of the same evidence against untuned mechanisms applies to mechanisms with extremely high bandwidths. Moreover, if observers had access only to a single channel we would expect that the measured second-order CSF would look like that channel, but estimates of the CSF are relatively flat (much more flat than a single bandpass channel) over a very large range of spatial-frequencies (Jamar & Koenderink, 1985; Kingdom, Keeble, & Moulden, 1995; Landy & Oruç, 2002; Schofield & Georgeson, 1999, 2003; Sutter, Sperling, & Chubb, 1995).

The fourth possibility is that the FRF model needs to be revised. We suggest that in broad strokes, the FRF model is correct in the sense that second-order patterns are analyzed by first-stage filters (tuned to the carrier pattern) and second-stage filters (tuned to the modulator). However, we suggest that the failure of critical-band masking to estimate the second-stage channel tuning is due to additional nonlinearities beyond those described by the FRF model. Thus, in Expt. 2 we use a different approach (a frequency-discrimination task) that does not require a second-order masker to estimate second-stage filter tuning. Then, we describe an elaboration of the FRF model that can resolve the apparent contradiction between different methods' estimates of second-stage filter tuning.

3. Experiment 2

3.1. Introduction

The results of Expt. 1 suggest that channels in second-order vision lack frequency tuning. This seems unlikely, but it is possible that the mechanisms responsible for processing orientation-modulated textures, which are generally harder to detect than their contrast-modulated equivalents, are weakly or not at all selective for second-order spatial frequency. Here, we employ an independent measure of channel bandwidth, one that does not rely on response to a noise masker, to rule out this possibility.

One hallmark of processing in independent channels is the labelled line. If two stimuli are perfectly discriminable at detection threshold, then a strong argument can be made that the mechanisms responsible for their detection are independent of each other. This logic was applied by Watson and Robson (1981) to count the number of separate spatial and temporal frequency channels underlying first-order vision, and again by Ellemberg, Allen, and Hess (2006), extending this result to second-order contrast-modulated stimuli. These studies suggest that, below 2 cycle/deg, first- and second-order mechanisms have similar bandwidths, based on a count of independent channels. This result, however, is based on contrast-modulated white-noise textures that do not require an orientation- and spatial-frequency-tuned first stage of filtering. We extend this paradigm to orientation-defined gratings, and adopt a signal-detection model to obtain an estimate of second-stage filter bandwidth.

3.2. Methods

3.2.1. Stimuli

Stimulus construction was identical to Expt. 1 except for the omission of the noise masker and the addition of a raised-cosine window (cosine width: 3.0°) to soften the stimulus edges.

3.2.2. Procedure

Subjects performed a simultaneous detection and discrimination task. Each trial consisted of a 500 ms display of a fixation cross, a 250 ms stimulus, a 500 ms blank interval, and a second 250 ms stimulus. After the second stimulus display, a blank screen was displayed until the observer's response followed by feedback. In each trial, one interval contained an orientation-modulated grating while the other contained a zero-modulation stimulus (a plaid, i.e., a sum of the two carriers). In each block of trials, the stimulus could have one of two possible modulation spatial frequencies (low, f_1 or high, f_2). The subject's task was to indicate with a single button press both the interval containing the stimulus and which spatial frequency was presented.

Each block began with displays of the two possible spatial frequencies to be discriminated within that block, f_1 and f_2. Within a single session the mean $f_{center} = (f_1 + f_2)/2$ was fixed across all blocks. Trials were presented at one of six modulation contrasts log spaced between 0.08 and 0.6. Trials were counterbalanced between first and second interval and high and low spatial frequency, with four repeats per combination of amplitude, interval, and frequency, yielding $2 \times 2 \times 4 \times 6 = 96$ trials per block. For each value of f_{center}, four blocks were run for each of three pairs of spatial frequencies in which f_1 and f_2 differed by 0.25, 0.5 or 1 octave. Each subject ran in three sessions in which f_{center} was 0.5, 1, or 1.5 cycle/deg.

3.2.3. Modeling

To derive channel bandwidths from the data, we fit a simple signal-detection model of the task. We assume that observers monitor two noisy channels, one with tuning curve centered on f_1 and one centered on f_2 of the current block, resulting in channel responses $\vec{C} = (C_{L,1}, C_{H,1}, C_{L,2}, C_{H,2})$, one response for each channel during each interval. We assume the channels have Gaussian tuning curves with unity peak gain, bandwidth σ_c, and mean response proportional to the presented modulation amplitude m. Channel responses are corrupted by additive Gaussian noise with SD σ_n. On each trial, the observer chooses the interval I that is most likely given the observed channel responses:

$$\begin{aligned} I &= \arg\max_I P(I|\vec{C}) \\ &= \arg\max_I \sum_{f,m} P(I,f,m|\vec{C}) \\ &= \arg\max_I \sum_{f,m} P(\vec{C}|I,f,m), \end{aligned} \tag{2}$$

with frequency estimate F defined analogously. The quantity $P(\vec{C} \mid I,f,m)$ is a function of the two model parameters σ_c and σ_n.

The model was fit to the data separately for each observer and value of f_{center}. The fit was performed through a dense grid-search of the parameter space rather than an optimization algorithm, using Monte Carlo simulation to estimate performance levels. For confidence intervals on the parameter of interest, channel bandwidth, we resampled the raw data within each condition 400 times and report 95% confidence intervals as the middle 95% of the resulting distribution of maximum-likelihood bandwidth estimates.

3.3. Subjects

Three subjects (one author), all with normal or corrected-to-normal vision, took part in this experiment. The Institutional Review Board at New York University approved the experimental procedures and all participants gave informed consent.

3.4. Results

Fig. 6 shows raw performance data and fits of a Weibull function for one subject in the f_{center} = 1 cycle/deg condition. Detection and discrimination thresholds were nearly identical when f_1 and f_2 differed by one octave (Fig. 6A). Discrimination performance suffered for smaller differences between the two discriminanda (Fig. 6B and C). Similar results were obtained from all three subjects. These results are in agreement with previous application of this paradigm to contrast-modulated gratings (Ellemberg, Allen, & Hess, 2006).

To estimate channel bandwidths, we fit the maximum-likelihood observer model to the data. Fig. 7 shows maximum-likelihood channel bandwidth estimates for each subject and mean frequency f_{center}. Channel bandwidths are all approx. 1–1.5 octave. Although bandwidths varied across subjects, there was no trend across mean spatial frequency within each subject, suggesting that second-order channel bandwidths are relatively constant in octaves. These results are consistent with established measurements of first-order channels (Graham, 1989; De Valois & De Valois, 1988) and with previous estimates of second-order channel bandwidth using summation (Landy & Oruç, 2002).

4. Observer modeling

The two experiments we have described have resulted in an apparent contradiction. In Expt. 1 we used the critical-band masking paradigm to estimate second-order filter bandwidth leading to evidence that second-order filters have extremely wide spatial frequency tuning. In Expt. 2, we used a simultaneous detection–discrimination task, resulting in estimates of bandwidth that were more typical (1–1.5 octave). In this section, we develop an alternative to the standard FRF model in an attempt to resolve this apparent contradiction.

4.1. First-stage filter bandwidth

We begin by describing a technical issue related to first-stage filter bandwidth that apparently has been overlooked in work on the FRF model. Typical FRF models assume that the first stage of filtering is carried out by standard V1-style receptive fields with preferred spatial-frequency and orientation tuned to one of the carriers (e.g., Bergen & Adelson, 1988; Bergen & Landy, 1991; Landy & Bergen, 1991; Landy & Oruç, 2002), and there is some evidence for this (Dakin & Mareschal, 2000). This assumption that the first-stage filter in FRF has orientation and frequency bandwidths of approximately 30° and one octave implies an interaction between the first and second stages of filtering that is not supported by second-order contrast sensitivity data.

Consider, for example, a vertical sine wave with carrier frequency f_c modulated in contrast by (i.e., multiplied by) a second, lower-frequency sine wave with frequency f_m. In the frequency domain, the modulator is manifested by distortion products with spatial frequency $f_d = f_c \pm f_m$. These distortion products have spatial frequencies offset from the carrier frequency by an amount equal to f_m. As f_m increases, these distortion products will be displaced further away from the carrier (Fig. 8). In the FRF model, the first-stage filter is typically assumed to be centered on the carrier frequency and is used to isolate the carrier and these distortion products from other spectral components. Thus, as f_m increases, the distortion products that signal the presence of the modulator will be increasingly attenuated by this first-stage filter. In the spatial domain this can be thought of as follows: as f_m increases, the resulting stripes of high and low contrast become narrower, eventually become so narrow that the first-order filter's receptive field effectively blurs over them, attenuating the contrast modulation in the first-order filter's output.

The sidebands created by modulation are the cues detected by the FRF model. The first two computations (FR) serve to demodulate the stimulus, effectively shifting the distortion products so that they now appear as power at the modulation frequency itself (f_m), which is further isolated by the final second-stage filter (the second F in FRF). Thus, the attenuation of those distortion products with increasing modulation frequency implies that second-order contrast sensitivity should be extremely lowpass. This is in contradiction with psychophysical measurements of second-order contrast sensitivity. Previous measurements of second-order contrast sensitivity (for both contrast- and orientation-modulated gratings) suggest that second-order contrast sensitivity is generally either nearly flat (Landy & Oruç, 2002; Sutter, Sperling, & Chubb, 1995) or modestly lowpass (Jamar & Koenderink, 1985; Kingdom, Keeble, & Moulden, 1995; Schofield & Georgeson, 2003). Using stimuli similar to those of the current study, Landy and Oruç (2002) found second-order contrast sensitivity to be almost completely flat over a five-octave range. Although it is possible that some unknown mechanism compensates for the low-pass behavior of the first

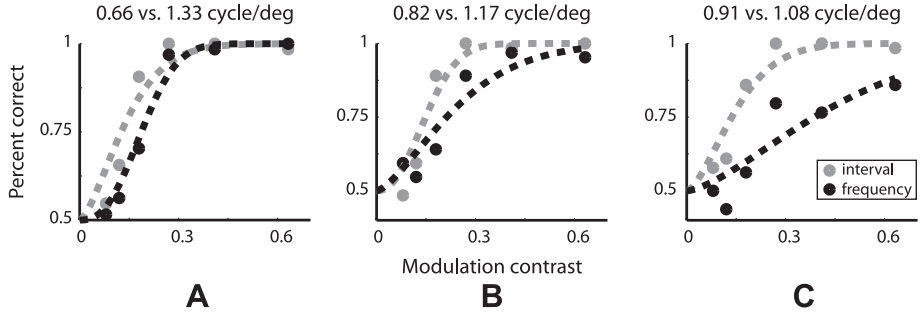

Fig. 6. Detection and discrimination performance for subject ZMW for f_{center} = 1 cycle/deg and separations between f_1 and f_2 of 1 (A), 0.5 (B) and 0.25 (C) octaves. Dotted lines: fits of a Weibull function to the data. As the separation decreases between the two spatial frequencies, discrimination performance becomes much worse than detection performance.

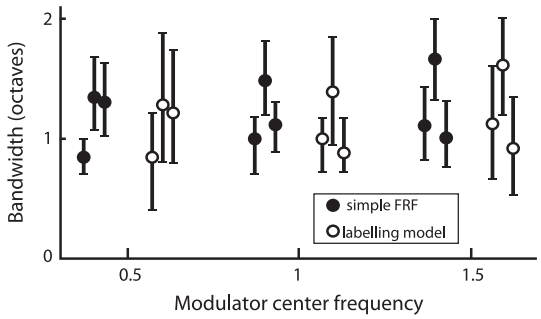

Fig. 7. Bandwidth estimates from Expt. 2. Filled circles: maximum-likelihood bandwidth estimate for each observer and center frequency f_{center}. Error bars: 95% confidence intervals based on bandwidth estimates from 400 bootstrap iterations. Open symbols and error bars show results from the labelling model. Each cluster of points shows data from the nearest modulator center frequency marked below, ordered identically by subject. The leftmost point in each cluster shows subject ZMW, as in Fig. 6.

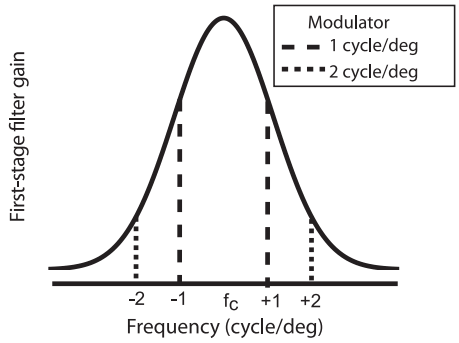

Fig. 8. Low-pass second-order filtering by the first-stage filter in the FRF model. Solid curve: a typical first-stage filter tuned to the carrier frequency f_c (e.g., 4.0 cycle/deg) of a contrast-modulated texture. Modulation at frequency f_m is defined in the Fourier domain by the presence of sidebands offset f_m from the carrier. As f_m increases, these distortion products are increasingly attenuated by the first-stage filter implying a degree of low-pass second-order sensitivity that is not found in human data. We propose that the effective first-stage filter pools over multiple channels and is relatively flat with respect to passing modulation sidebands.

stage, in the modeling we present here, we assume instead that the first-stage filters have significantly wider bandwidth than a typical first-order channel. We use first-stage filters that, in the Fourier domain, are flat around each carrier out past the distortion products generated by any of the modulators we have used. That does not mean we posit the existence of extremely wide-band V1 neurons. Rather, this likely corresponds to pooling over multiple first-order channels before rectification.

4.2. Standard FRF observer model applied to Experiment 1

As a first step, we implemented a straightforward FRF observer to confirm that the second-stage filter can be recovered with critical-band masking and that the peculiarities in our human data were not artifacts of the stimulus generation or the result of the nonlinear rectification stage, and to form the basis for subsequent modeling. We simulated an FRF observer (with late noise added to rectified first-stage filter responses at each location) performing the grating-detection task and analyzed the data in the same way as we analyzed the data from the human observers in Expt. 1.

Our simple observer was implemented by convolving the image with quadrature phase filters and adding the squared magnitudes of the resulting images to obtain a filtered and rectified image. This was performed using filters centered on each of the two carriers, and the resulting images were subtracted to form a demodulated image $I_{FN}(x,y) = E_{45}(x,y) - E_{135}(x,y)$ where $E_{45}(x,y)$ and $E_{135}(x,y)$ represent the energy near carriers C_{45} and C_{135}, respectively. Although I_{FN} is computed as the difference in texture-energy signals associated with the two carriers (as in Landy & Bergen (1991)), this model and the models that follow behave similarly if only one channel is used. Because noise and modulator signals occurred only at 0° and 90°, the energy images were reduced to two 1-D stimuli consisting of averages along either the rows or columns. The model observer then convolved the resulting 1-D signals with a Gabor template with frequency f_m and one-octave bandwidth and chose the orientation with the highest peak response.

The rectification step used here, summing squared responses of quadrature pair filters tuned to each carrier, is equivalent to an energy computation and has the convenient property of producing a demodulated image that does not contain power at multiples of the carrier frequency. This nonlinearity has been previously applied as the rectification stage of other FRF models of second-order vision (Fogel & Sagi, 1989; Landy & Bergen, 1991; Sutter, Beck, & Graham, 1989). We tried other pointwise nonlinearities, such as full-wave or half-wave rectification, but the exact choice of rectification seems to have a relatively minor effect on our observer's behavior. Fig. 9 shows the derived channels from a simulated experiment using this observer. We were able to successfully recover the observer's internal channel using critical-band-masking; the estimate of the model observer's internal channel is nearly perfect (Fig. 9) confirming the validity of critical-band-masking under the assumption of simple FRF but contradicting our experimental results.

4.3. Labelling model

The results of Expt. 2 suggest that second-order channels have fixed bandwidths slightly wider than one octave. This rules out the all-pass channels apparent from the results of Expt. 1, so the challenge is how to account for the highly nonlinear response to noise found in critical-band masking. To resolve this conflict, we propose an additional processing stage in which the locally dominant carrier is dramatically amplified while the locally weak carrier is suppressed, essentially producing an image of dominant texture labels. We apply this model to explain why human observers behave as if they were using extremely wide-band channels in the presence of masking noise and ordinary one-octave-wide channels when discriminating second-order gratings.

There have been previous suggestions that the extraction of dominant response plays a role in second-order processing. Malik and Perona (1990) proposed a nonlinear inhibition stage that locally suppressed weak channel responses as an explanation for the detection of edges between a Gabor field and its luminance-flipped anti-Gabor field counterpart. They termed this mechanism "leaders-take-most", and proposed that it could be computed using only local connections in V1. More recently, Motoyoshi and Nishida (2004), in a study of summation between second-order signals carried by differently oriented carriers, found that isotropic orientation noise affected performance less than would be predicted based on the orientation blur of the Gabor stimuli and the orientation bandwidth of first-order filters. They were able to account for this effect using a similar model that extracted dominant orientation across space and suppressed weak responses. Graham and Sutter (2000) found evidence for normalization of the first-stage channel responses in second-order vision, and attributed it to inhibition between channels that resulted in a similar attenuation of the non-dominant response.

We implement a strong form of dominant-response selection in a model that classifies each location as belonging to a single

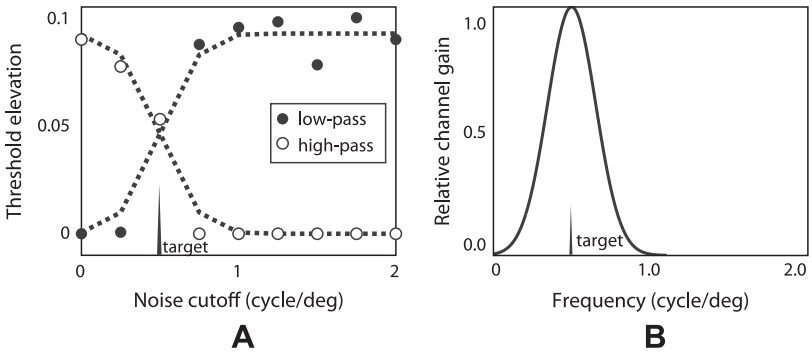

Fig. 9. FRF-observer simulation. (A) Threshold elevation data and cumulative Gaussian fits as in Fig. 3. (B) Estimated channel. Note that the estimated channel is precisely centered on the 0.5 cycle/deg target and has a bandwidth (one octave) identical to that used to simulate the FRF observer.

texture. The input to this model is the filtered and rectified texture energy image I_{FN} from our FRF observer (Fig. 10). We applied the simplest possible form of classification algorithm, the texture-energy image is thresholded so that responses above the mean energy are set to 1 and responses below are set to 0. Although other, more complex models can be devised, such as computing a Bayesian estimate of local texture based on a model of stimulus structure, this simple threshold model fits our critical-band-masking data very well.

Why would the visual system throw away information by coding texture as binary labels, and why does this additional processing stage imply the loss of frequency-tuning for second-order masking but not for second-order discrimination? It could be that our visual system is optimized for second-order borders that arise between adjacent or overlapping textured materials. These texture-defined boundaries, unlike our stimuli, do not blend continuously, so that treating texture energy responses as all or nothing may be appropriate. In the absence of noise, this nonlinearity may cause texture-defined gradients to appear more like sharp edges (see Fig. 2B for a sinusoidal modulator). When noise is added to an existing modulator signal, even at a frequency distant from that of the modulation target, it cannot increase the responses to locally dominant carrier textures because those responses are saturated. Noise can, however, decrease the response to the locally dominant carrier by pushing the previously suppressed carrier to dominance. Masking noise, therefore, has a suppressive effect on the modulator signal regardless of its relative frequency, as demonstrated in our human data.

Once each pixel in the input image is labelled 1 for C_{45} or 0 for C_{135} we average the image along rows or columns to obtain 1D marginals as in the simple FRF observer described above, which can then be processed by filters of various bandwidths to obtain channel responses. To test the model we simulated the model's response to a large collection of images at several low- and high-pass noise cutoffs as well as several modulation contrasts. Each image was filtered and rectified as in the standard FRF model above, then a small amount of white noise was added. A separate early noise parameter—the SD of the pixels in this white noise—was necessary because, in the complete absence of input noise, the model otherwise predicts invariance to modulation contrast. These images (I_{FN} + noise) were then processed by the labelling model with one-octave second-stage filters. Late noise was added to the outputs of the second-stage filters (thus there was a second SD parameter in the model). Performance, defined as the probability that the true orientation had greater energy, was determined based on the noisy second-order channel responses. We computed thresholds and threshold-elevation curves from these performance data as in Expt. 1. The model has two noise parameters: early and late. Because changing the early noise parameter involved resimulating the model on the entire image set, we opted to avoid fitting it and instead chose a small value to roughly match human performance in the no-noise condition.

We simulated the labelling model observer in Expt. 1. The simulated threshold elevation curves match the human data very well (Fig. 11; compare Fig. 3B). As with the human data, model thresholds rise approximately linearly with noise cutoff across the entire range of noise cutoffs. This occurs despite the fact that the simulated observer uses an octave-wide second-stage filter. Table 1 summarizes the performance of the labelling model in describing our human data.

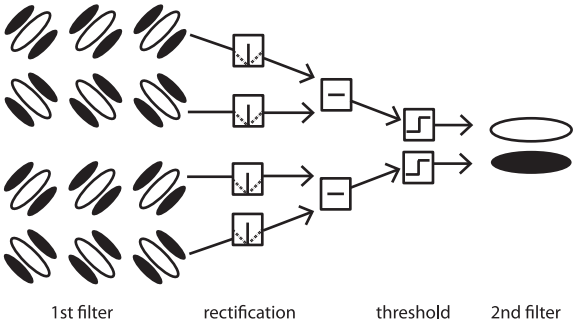

Fig. 10. Labeling model. First, the difference texture energy for each channel is computed at each location in the image. Then, locations exceeding the mean texture energy of the image are labelled 1 and other locations are labelled 0. Finally, the labelled image is processed by a matched second-stage filter.

4.4. Comparison of labelling model and FRF model for Expt. 1

We would like to compare the labelling model to the simple FRF model, but in order to do so we must impose some additional constraints on the simple FRF model. The unconstrained FRF model is only able to fit our threshold-elevation data accurately with an unrealistically broad channel centered near the middle of the noise cutoff range, regardless of modulation frequency. Given the evidence that observers have access to bandpass second-order filters tuned to a variety of modulation frequencies (see Expt. 1 Results and the findings of Expt. 2), we will compare the predictions of the simple FRF model using channels constrained to be centered on the modulation target, but with bandwidth as a free parameter, to the predictions of the labelling model with fixed octave-width (as in results of Expt. 2) channels centered on the target.

Although a least-squares fit of the two parameters in each model produces much better r^2 values for the labelling model (Table 1), we would like to express this model comparison in terms of relative likelihood. Typically one would compute $P(data|model)$

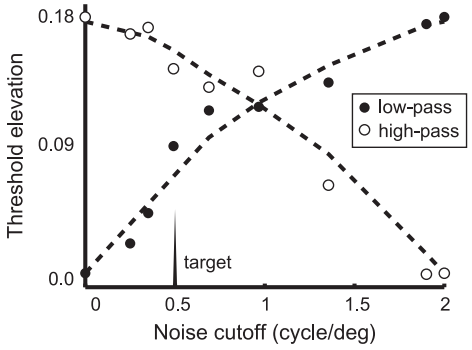

Fig. 11. Performance of labelling model vs. human data for one subject in the f_c = 4 cycle/deg, f_m = 0.5 cycle/deg condition. The labelling model very accurately fits human data even when using an internal channel centered on the target modulation spatial frequency.

for each model. This is difficult, however, because the simple FRF model with the minimal assumption of constant signal-to-noise ratio at threshold makes no specific prediction about human choice data. Instead, it predicts threshold elevation, which is a non-invertible function of the data. We therefore reverse the usual computation, and estimate $P(model|data)$ by generating a bootstrapped distribution of human data thresholds and finding the model parameters that maximize the likelihood of the model predictions of threshold elevation given that distribution. We report the ratio of these likelihoods

$$\max_{\sigma_{early},\sigma_{late}} P\left(M^1_{\sigma_{early},\sigma_{late}}|D\right) \Big/ \max_{bw,scale} P\left(M^2_{bw,scale}|D\right),$$
$$where\ P\left(M^i_p|D\right) = \prod_c D_c\left(M^i_{p,c}\right), \quad (3)$$

where D_c is an estimate of the PDF of human threshold elevations in condition c computed from the smoothed distribution of bootstrap threshold elevation, and $M^i_{p,c}$ is the predicted threshold elevation of model i with parameters p in condition c. The smoothing kernel was a narrow Gaussian, and bootstrapping threshold elevation was performed by a parametric bootstrap as described by Maloney (1990).

By this measure, of the two conditions in which the modulation frequency coincided with the center of the noise range, one showed strong evidence for the labelling model, and one showed moderate evidence (Table 1). This is not very interesting, however, as these are the conditions under which the predictions of the FRF model and the labelling model nearly coincide. For all subjects and all conditions in which the modulator was not in the middle of the noise spectrum, so that a cumulative Gaussian centered on the modulator could not accurately fit nearly straight-line threshold elevation, the evidence in favor of the labelling model was very strong. This analysis agrees with the r^2 values computed via least-squares.

4.5. Effect of adding external noise on labelling model

We were concerned that the labelling model would respond differently to increasing external noise than human observers. We simulated the labelling model performing the same orientation–discrimination task with broadband second-order masking noise as a function of masking-noise contrast. For human observers, threshold modulation contrast is proportional to masking noise contrast both for grating detection (pilot data collected prior to Expt. 1) and for the identification of second-order letters (Oruç, Landy, & Pelli, 2006). For second-order grating detection, the labelling model also exhibits thresholds proportional to masker power (Fig. 12). The appearance of linearity may be counterintuitive for a model that includes a hard threshold, but consider that—in the presence of early noise—addition of noise followed by thresholding is equivalent on average to passing the image through a nonlinearity composed of the CDF of the early noise, and at contrasts near zero a sigmoid is nearly linear.

4.6. Application of labelling model to Expt. 2

Expt. 1 demonstrates that second-order mechanisms behave as though they lack frequency tuning in their response to noise masking. In Expt. 2, however, the performance of subjects discriminating second-order gratings of differing modulation frequency was consistent with target-centered channels of about 1–1.5 octave bandwidth. To verify that our labelling model accounts for both of these very different behaviors, we implemented the same signal-detection-theory model from Expt. 2 with channel outputs computed according to the labelling model with an additional parameter controlling the second-stage filter bandwidth (in the previous model simulations, this was fixed at one octave). The early-noise parameter was fixed as it was in the previous simulations. The resulting 2nd-order channel bandwidth estimates are near 1 octave (Fig. 7) and are very similar to the bandwidths estimated without this additional nonlinearity. Examination of predicted psychometric functions for individual sessions confirms that the labelling model does a reasonably good job of explaining human behavior in this task (Fig. 13). Together, these results suggest that a nonlinear dominant-response-selection operation like the one applied here disrupts the ability of second-order mechanisms to cope with filtered masking noise but does not fundamentally alter their response to preferred stimuli.

5. Discussion

We attempted to measure the spatial-frequency bandwidth of the putative second-stage filter underlying detection of orientation-modulated images using two techniques. Expt. 1, based on critical-band masking (Solomon & Pelli, 1994), measured the threshold elevation in response to bandpass filtered second-order noise in an orientation–discrimination task. According to the theory, noise present within the channel used in the task will raise thresholds while noise outside of this channel should have nearly no effect. Our data revealed no evidence for frequency tuning—noise at any modulation frequency raised thresholds. We verified in modeling that this was not due to the rectification stage of the FRF model thought to underlie second-order vision (Fig. 9), implying that either the FRF model is flawed or that the second-stage filters are not frequency tuned.

Although we had strong reason to suspect that second-stage filters are in fact frequency and orientation tuned (Arsenault & Kingdom, 1999; Ellemberg, Allen, & Hess, 2006; Hallum, Landy, &

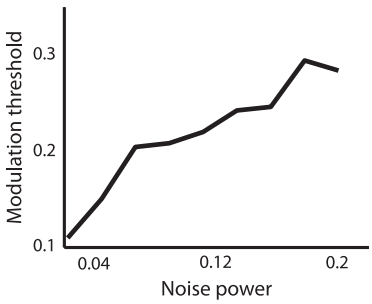

Fig. 12. Labelling model threshold-vs.-noise curve. Threshold on y-axis is modulation contrast threshold, and noise is bandpass from 0 to $f_c/2$ (2.0) cycle/deg. Despite the significant nonlinearity employed by the labelling model, threshold elevation vs. noise power is approximately linear in agreement with human data.

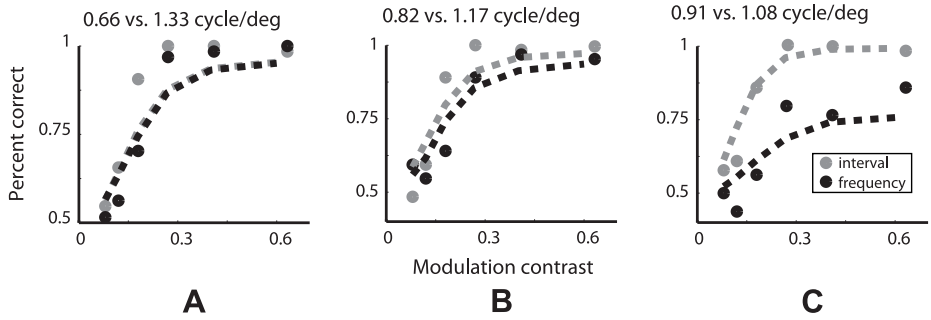

Fig. 13. Example data from Expt. 2 and predicted psychometric functions under the labelling model with best-fitting bandwidth and late-noise parameters. Data shown are taken from Fig. 6, and dashed lines represent the predicted percentage correct of the labelling model with maximum-likelihood bandwidth and late-noise parameters fit to all three conditions.

Heeger, 2011; Kwan & Regan, 1998; Landy & Oruç, 2002; Sutter, Sperling, & Chubb, 1995), we carried out a second experiment to estimate bandwidth without masking noise. Expt. 2 was based on the labelled-lines experiment of Watson and Robson (1981) that showed evidence for independent first-order channels, and follows work by Ellemberg, Allen, and Hess (2006) that applied the technique to contrast-modulated images. This experiment demonstrated that modulation frequency discrimination performance for orientation-modulated images is nearly the same as detection performance at a separation of one octave, strongly supporting independent narrowband channels, and that maximum-likelihood fits of frequency bandwidth reveal channels similar in bandwidth to first-order channels. Moreover, the tuning was similar across a range of modulation frequencies, suggesting that the channels have a fixed bandwidth in octaves.

These conflicting results imply an inadequacy of the FRF model, which we were able to resolve with the addition of a pointwise nonlinearity preceding the second-stage filter that extracts local dominant orientation response. The need for a nonlinearity of this sort has been previously suggested for second-order vision (Malik & Perona, 1990; Motoyoshi & Nishida, 2004), although perhaps not so dramatically as with these masking results. The labelling model is based on inhibition between differing orientation channels, a phenomenon that has been demonstrated in V1 (Carandini, Heeger, & Movshon, 1997; Morrone, Burr, & Maffei, 1982; Somers, Nelson, & Sur, 1995) and psychophysically in second-order vision (Motoyoshi & Kingdom, 2003).

The model has a number of appealing properties, first among which is its simplicity. Although we implemented the nonlinearity as a pointwise step function following demodulation and the addition of early noise, this is equivalent to a steep sigmoidal response function without early noise. As a result this highly nonlinear model behaves under many conditions identically to the nearly linear (in modulation contrast) FRF model. Threshold vs. noise power is nearly linear (Fig. 12), and frequency detection/discrimination data from Expt. 2 were described equally well by the standard FRF model and by the augmented labelling model (Fig. 7). The neural implementation of a labelling nonlinearity could be as simple as a thresholding step following rectification of the first-stage inputs, or could be a consequence of nearly winner-take-all competition, which can emerge as a result of divisive normalization (Carandini & Heeger, 2012). Our model is implemented in terms of winner-take-all competition between first-order channels, but can be implemented with similar results by thresholding within a single channel. An examination of how these findings generalize to other types of second-order stimuli may be necessary in order to further constrain modeling.

A nonlinearity of this sort also resolves an existing problem with the FRF model. Input to the second-stage filters, depending slightly on the choice of rectification operation, should be approximately proportional to the output of the first-stage filters. When measured, however, second-order performance is weakly sensitive (Schofield & Georgeson, 1999) or completely insensitive (Barbot, Landy, & Carrasco, 2011) to carrier contrast over a wide range. This invariance could result from the dominance of early multiplicative noise or from normalization of the first-stage inputs. Our model builds in first-order invariance by normalizing continuous contrast signals to (0, 1) labelling signals.

It is not clear why critical-band-masking applied to second-order letters (Oruç, Landy, & Pelli, 2006) seemed to produce threshold-elevation curves consistent with Gaussian-channels while the same technique applied to much simpler grating stimuli did not. One possibility is that the channels estimated for letters tended to be relatively high in spatial frequency, so that threshold elevation was flat only at frequencies much below the channel observers seemed to use. To avoid first-order confounds, we used low-frequency modulation targets, so we cannot address whether low-frequency (second-order) noise masks high-frequency (second-order) signal. It may be that the nonlinear masking effect of added modulation noise only applies to noise that is much higher frequency than the channel, not to noise at much lower frequency. Although this is not a feature of the model we presented, it is trivial to build in by labelling relative to a local average of the demodulated signal rather than the global average that we used. If a local average is used, then very low frequency modulations behave like DC and do not affect processing.

In the course of modeling the results of these experiments we noticed that the first stage of filtering, if carried out by typical V1 filters, implies a dramatically lowpass operation on the modulator signal (Fig. 8). This lowpass prediction contradicts measurements of the second-order contrast sensitivity function that find it to be relatively flat (Landy & Oruç, 2002; Sutter, Sperling, & Chubb, 1995) or only modestly lowpass (Jamar & Koenderink, 1985; Kingdom, Keeble, & Moulden, 1995; Schofield & Georgeson, 2003). Moreover, it is very difficult to explain how high-frequency modulator noise could mask low-frequency second-order signals if the modulation has effectively been lowpass filtered by the first-stage filters. We assume for our analysis that the effective bandwidth of the first-stage filter is much more broad than a typical first-order channel, a notion supported by evidence of summation across carrier signals with very different orientation (Motoyoshi & Nishida, 2004), and by texture segregation results applying to textures that differ with respect to element arrangement (Graham, Sutter, & Venkatesan, 1993).

Why would the visual system distort second-order images in a way that loses information? Although transforming a continuous measure of relative orientation contrast into a binary output is clearly suboptimal in our task, it may be optimal in the context

of more natural texture-defined images. Texture-defined edges, which frequently co-occur with luminance-defined edges, are often produced by adjacent or occluding objects composed of different materials. If the goal of second-order vision is to process images defined by variations in material properties, which generally do not blend continuously, winner-take-all behavior may reflect an optimal prior favoring sharp edges in texture-defined images.

Acknowledgments

The work reported here was supported by NIH Grants EY08266 and EY16165. We thank Charlie Chubb and Ipek Oruç for comments on the manuscript.

References

Arsenault, A. S., & Kingdom, F. A. A. (1999). Modulation frequency and orientation tuning of second-order texture mechanisms. *Journal of the Optical Society of America A, 16*, 427–435.

Barbot, A., Landy, M. S., & Carrasco, M. (2011). Exogenous attention enhances 2nd-order contrast sensitivity. *Vision Research, 51*, 1086–1098.

Bergen, J. R., & Adelson, E. H. (1988). Early vision and texture perception. *Nature, 333*, 363–364.

Bergen, J. R., & Landy, M. S. (1991). Computational modeling of visual texture segregation. In M. S. Landy & A. J. Movshon (Eds.), *Computational models of visual processing* (pp. 253–271). Cambridge, Mass.: MIT Press.

Blackwell, K. T. (1998). The effect of white and filtered noise on contrast detection thresholds. *Vision Research, 38*, 267–280.

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience, 13*, 51–62.

Carandini, M., Heeger, D. J., & Movshon, A. J. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *Journal of Neuroscience, 17*, 8621–8644.

Dakin, S. C., & Mareschal, I. (2000). Sensitivity to contrast modulation depends on carrier spatial frequency and orientation. *Vision Research, 40*, 311–329.

De Valois, R. L., & De Valois, K. K. (1988). *Spatial vision*. New York: Oxford University Press.

Derrington, A. M., & Badcock, D. R. (1986). Detection of spatial beats: Non-linearity or contrast increment detection? *Vision Research, 26*, 343–348.

Ellemberg, D., Allen, H. A., & Hess, R. F. (2006). Second-order spatial frequency and orientation channels in human vision. *Vision Research, 46*, 2798–2803.

Fogel, I., & Sagi, D. (1989). Gabor filters as texture discriminators. *Biological Cybernetics, 61*, 103–113.

Graham, N. (1994). Non-linearities in texture segregation. In G. R. Bock & J. A. Goode (Eds.). *CIBA foundation symposium* (Vol. 184, pp. 309–329). New York: Wiley.

Graham, N., & Sutter, A. (2000). Normalization: Contrast-gain control in simple (Fourier) and complex (non-Fourier) pathways of pattern vision. *Vision Research, 40*, 2737–2761.

Graham, N., Sutter, A., & Venkatesan, C. (1993). Spatial-frequency- and orientation-selectivity of simple and complex channels in region segregation. *Vision Research, 33*, 1893–1911.

Graham, N. V. S. (1989). *Visual pattern analyzers*. New York: Oxford University Press.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Hallum, L. E., Landy, M. S., & Heeger, D. J. (2011). Human primary visual cortex (v1) is selective for second-order spatial frequency. *Journal of Neurophysiology, 105*, 2121–2131.

Henning, G. B., Hertz, B. G., & Hinton, J. L. (1981). Effects of different hypothetical detection mechanisms on the shape of spatial-frequency filters inferred from masking experiments: I. Noise masks. *Journal of the Optical Society of America, 71*, 574–581.

Jamar, J. H. T., & Koenderink, J. J. (1985). Contrast detection and detection of contrast modulation for noise gratings. *Vision Research, 25*, 511–521.

Kingdom, F. A. A., Keeble, D., & Moulden, B. (1995). Sensitivity to orientation modulation in micropattern-based textures. *Vision Research, 35*, 79–91.

Kovács, I., & Fehér, A. (1997). Non-Fourier information in bandpass noise patterns. *Vision Research, 37*, 1167–1175.

Kwan, L., & Regan, D. (1998). Orientation-tuned spatial filters for texture-defined form. *Vision Research, 38*, 3849–3855.

Landy, M. S., & Bergen, J. R. (1991). Texture segregation and orientation gradient. *Vision Research, 31*, 679–691.

Landy, M. S., & Henry, C. A. (2007). Critical-band masking estimation of 2nd-order filter properties. *Perception, 36*(Suppl.), 61.

Landy, M. S., & Oruç, İ. (2002). Properties of second-order spatial frequency channels. *Vision Research, 42*, 2311–2329.

Larsson, J., Landy, M. S., & Heeger, D. J. (2006). Orientation-selective adaptation to first- and second-order patterns in human visual cortex. *Journal of Neurophysiology, 95*, 862–881.

Majaj, N. J., Pelli, D. G., Kurshan, P., & Palomares, M. (2002). The role of spatial frequency channels in letter identification. *Vision Research, 42*, 1165–1184.

Malik, J., & Perona, P. (1990). Preattentive texture discrimination with early visual mechanisms. *Journal of the Optical Society of America A, 7*, 923–932.

Maloney, L. T. (1990). Confidence intervals for the parameters of psychometric functions. *Perception and Psychophysics, 47*, 127–134.

Morrone, M. C., Burr, D. C., & Maffei, L. (1982). Functional implications of cross-orientation inhibition of cortical visual cells. *Journal of Physiology, 216*, 335–354.

Motoyoshi, I., & Kingdom, F. A. A. (2003). Orientation opponency in human vision revealed by energy-frequency analysis. *Vision Research, 43*, 2197–2205.

Motoyoshi, I., & Nishida, S. (2004). Cross-orientation summation in texture segregation. *Vision Research, 44*, 2567–2576.

Nachmias, J., & Rogowitz, B. E. (1983). Masking by spatially modulated gratings. *Vision Research, 23*, 1621–1629.

Oruç, İ., Landy, M. S., & Pelli, D. G. (2006). Noise masking reveals channels for second-order letters. *Vision Research, 46*, 1493–1506.

Pelli, D. G. (1981). *Effects of visual noise*. PhD thesis, Cambridge University.

Perkins, M. E., & Landy, M. S. (1991). Nonadditivity of masking by narrow-band noises. *Vision Research, 31*, 1053–1065.

Schofield, A. J., & Georgeson, M. A. (1999). Sensitivity to modulations of luminance and contrast in visual white noise: Separate mechanisms with similar behavior. *Vision Research, 39*, 2697–2716.

Schofield, A. J., & Georgeson, M. A. (2003). Sensitivity to contrast modulation: The spatial frequency dependence of second-order vision. *Vision Research, 43*, 243–259.

Scott-Samuel, N. E., & Georgeson, M. A. (1999). Does early non-linearity account for second-order motion? *Vision Research, 39*, 2853–2865.

Solomon, J. A., & Pelli, D. G. (1994). The visual filter mediating letter identification. *Nature, 369*, 395–397.

Somers, D. C., Nelson, S. B., & Sur, M. (1995). An emergent model of orientation selectivity in cat visual cortical simple cells. *Journal of Neuroscience, 15*, 5448–5465.

Sutter, A., Beck, J., & Graham, N. (1989). Contrast and spatial variables in texture segregation: Testing a simple spatial-frequency channels model. *Perception and Psychophysics, 46*, 312–332.

Sutter, A., Sperling, G., & Chubb, C. (1995). Measuring the spatial frequency selectivity of second-order texture mechanisms. *Vision Research, 35*, 915–924.

Tanaka, H., & Ohzawa, I. (2009). Surround suppression of v1 neurons mediates orientation-based representation of high-order visual features. *Journal of Neurophysiology, 101*, 1444–1462.

Watson, A. B., & Robson, J. G. (1981). Discrimination at threshold: Labelled detectors in human vision. *Vision Research, 21*, 1115–1122.