Detection and discrimination of texture modulations defined by orientation, spatial frequency, and contrast

Nicolaas Prins and Frederick A. A. Kingdom

McGill Vision Research, McGill University, 687 Pine Avenue West H4-14, Montréal, Québec, Canada, H3A 1A1

Received July 17, 2002; revised manuscript received October 1, 2002; accepted October 2, 2002

We sought to determine whether the detection and the identification of texture modulations are mediated by a common mechanism. On each trial two textures were presented, one of which contained a modulation in orientation (OM), spatial frequency (FM), or contrast (CM). Observers were required to indicate whether the modulated texture was presented in the first or the second interval as well as the nature of the texture modulation. The results showed that for two of the three pairwise matchings (OM–FM and OM–CM) detection and identification performance were nearly identical, suggesting a common underlying mechanism. However, when FM and CM textures were paired, discrimination thresholds were significantly higher than detection thresholds. In the context of the filter–rectify–filter model of texture perception, our results suggest that the mechanisms underlying detection are labeled with respect to their first-order input; i.e., the identities of these mechanisms are available to higher levels of processing. Several possible explanations for the misidentification of FM and CM at detection threshold are considered.

OCIS codes: 330.1880, 330.4060, 330.5000, 330.5510, 330.6100, 330.7310.

1. INTRODUCTION

Modulations of orientation, spatial frequency, or contrast in the visual image provide the observer with important clues regarding the spatial layout of the visual scene. Such modulations are often termed second-order modulations, as they are not defined by luminance variations and hence cannot be detected by first-order (luminance) filters only. Instead, human processing of second-order variations in the visual image has often been modeled by second-order mechanisms (see Refs. 1 and 2 for recent reviews). Although the specific instantiation of the model differs somewhat among researchers, the basic idea behind these models is identical. The model is depicted schematically in Fig. 1. The stimulus is a texture that contains a square-wave spatial-frequency modulation. The mechanism first filters the texture by using orientation- and spatial-frequency-selective filters (such as simple cells). The output from this first stage is then subjected to a nonlinearity (here, full-wave rectification) and consequently integrated by a larger second-stage filter. Any difference in spectral content between the excitatory and the inhibitory regions of the second-stage filter will lead to the activation of appropriate second-stage filters. We shall refer to these mechanisms as filter–rectify–filter (FRF) mechanisms.

Activation of an FRF mechanism may arise from a difference between texture regions in orientation, spatial frequency, or contrast. Hence any individual FRF mechanism will be indiscriminate as to the nature of the texture discontinuity that triggered its response. This inherent flexibility of the FRF mechanism has made it possible to relate it to the perception of texture modulations in contrast,[3,4] orientation,[5–11] spatial frequency,[11,12] and even pattern arrangement.[11,13] On the other hand, it follows that if the *same* FRF mechanism underlies the detection of different types of texture modulation, we would expect observers to have difficulty identifying the *type* of modulation at its detection threshold. This represents one of the possible outcomes of the experiments reported in the present study, whose aim is to establish whether different types of texture modulation can be identified at their individual detection thresholds. However, we show below how a consideration of the likely pattern of activity across the range of FRF mechanisms activated by different types of texture modulation leads to an alternative possibility, namely, that different texture-modulation types will maximally activate *different* FRF mechanisms and that therefore the type of texture modulation might be identifiable at detection threshold.

At the core of the argument is the observation that the distribution of active FRF mechanisms across the orientation and spatial-frequency preferences of their first-order filters will be different for different types of texture modulation, even at very low depths of modulation. This is so because FRF mechanisms, by their very nature, respond to the *difference* in spectral content between different texture regions. In the case of low-depth-of-modulation orientation modulation (OM) and frequency modulation (FM) textures, this difference in spectral content may have peaks that are nowhere near the center orientation and spatial frequency of the texture.[14] This is illustrated in Fig. 2. In Fig. 2a we plot, in Fourier space, the spectral amplitude distribution of an unmodulated texture at the dc spatial frequency of 5 cycles per degree (cpd) and the dc orientation of 0° (horizontal).[15] The texture has a spatial-frequency bandwidth (full-width at half-height) of 1.5 octaves and an orientation bandwidth of 60° (these values correspond to those of the textures used in this study). Figure 2b plots the difference in spectral amplitude between the two texture regions of a

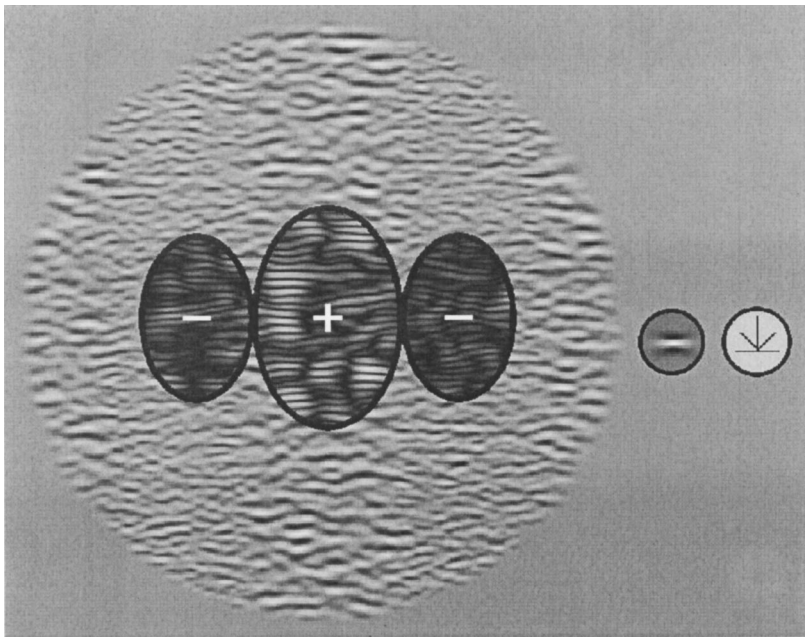


Fig. 1. FRF mechanism. A schematic FRF mechanism is shown atop a frequency-modulated texture. The amplitude of modulation is 0.2 octave. The excitatory center of the FRF shown covers a texture region in which the center spatial frequency is 0.2 octave below the texture's dc spatial frequency (5 cpd) and the inhibitory surrounds cover texture regions in which the center spatial frequency is 0.2 octave above the dc spatial frequency. The texture is first filtered with simple first-order luminance filters selective for orientation and spatial frequency. One example Gabor-shaped filter, maximally tuned to a spatial frequency of 3.5 cpd, is shown on the right. The rectified output from the first-stage filter is consequently integrated by the larger second-stage filter. The center region leads to activation of the FRF mechanism (+), whereas the surround leads to inhibition of the FRF mechanism (−). Shown within the receptive fields of the second-stage filter is the full-wave-rectified output from the first-stage (luminance) filter, on the far right of the figure.

square-wave-modulated OM texture: one texture region is centered at 4° counterclockwise from horizontal (θ_1); the other is centered at 4° clockwise from horizontal (θ_2). From this figure it is clear that the difference in spectral content peaks at orientations around 30° either side from horizontal. We[14] recently measured threshold OM depths at which the spatial frequency of the OM could be discriminated. In the critical conditions, observers were adapted to a simple luminance grating at varying orientations. Thresholds peaked when observers were adapted to a simple luminance grating oriented at 30° (or −30°) from horizontal. This result indicates that the underlying mechanism indeed acts upon the difference in spectral content between different texture regions. Hence two distinct groups of mechanisms will be active in response to an OM. Both groups will have a first-order spatial-frequency preference that lies at the center spatial frequency contained in the texture, but the two groups will differ in their first-order preference for orientation. Note that the responses of the two groups of FRF mechanisms will be in counterphase. Mechanisms tuned to orientations at 30° clockwise from horizontal will be maximally activated when their excitatory centers cover the texture region where orientations peak in the clockwise direction. Mechanisms tuned to orientations at 30° counterclockwise, on the other hand, will be maximally activated when their excitatory centers cover the texture region where orientations peak in the counterclockwise direction.

A similar argument applies to FM textures. In Fig. 2c we plot the difference in spectral content between the two texture regions of a square-wave-modulated FM texture: One texture region is centered at a spatial frequency that lies 0.1 octave below the dc spatial frequency (f_1); the other texture region is centered at a spatial frequency that lies 0.1 octave above the dc spatial frequency (f_2). The peaks in the difference distribution lie at spatial frequencies quite far removed from the dc spatial frequency, and again there will be two distinct groups of active mechanisms that, in the case of an FM texture, differ with respect to their first-order spatial-frequency preferences but are both centered at the dc orientation of the texture. We[14] confirmed that threshold elevation in a modulation spatial-frequency-discrimination task using FM textures peaked when observers were adapted to a simple luminance grating with a spatial frequency approximately 1 octave below the dc spatial frequency of the texture. There was only a slight indication that luminance filters tuned to spatial frequencies above the dc spatial frequency of the texture were involved. This is in all likelihood due to a lower contrast sensitivity at these higher frequencies. The two groups of active FRF mechanisms will again respond in counterphase to the texture modulation in a fashion similar to that in the case of OMs.

In the case of a contrast modulation (CM) texture, the distribution of the spectral difference between two texture regions differing in contrast will be identical (except for a scaling factor) to the spectral distribution of a texture centered at the dc orientation and spatial frequency. In other words, the difference distribution will simply be a scaled version of that in Fig. 2a.

As outlined above and illustrated in Fig. 2, each of the three types of texture modulation used in the present research (OM, FM, and CM) will lead to a unique pattern of activation of mechanisms that differ in the preferences of their first-order (luminance) filters. A discontinuity in orientation between two texture regions will lead to two groups of active FRFs, both of which will be centered at the stimulus spatial frequency but at two distinct orientations. Conversely, a spatial-frequency discontinuity will again lead to two groups of active FRFs; however, in this case both will be centered at the peak orientation of the stimulus, but each group will have different preferences for spatial frequency. A discontinuity in contrast will lead to the activation of only one group of FRFs, centered at the peak orientation and spatial frequency of the texture.

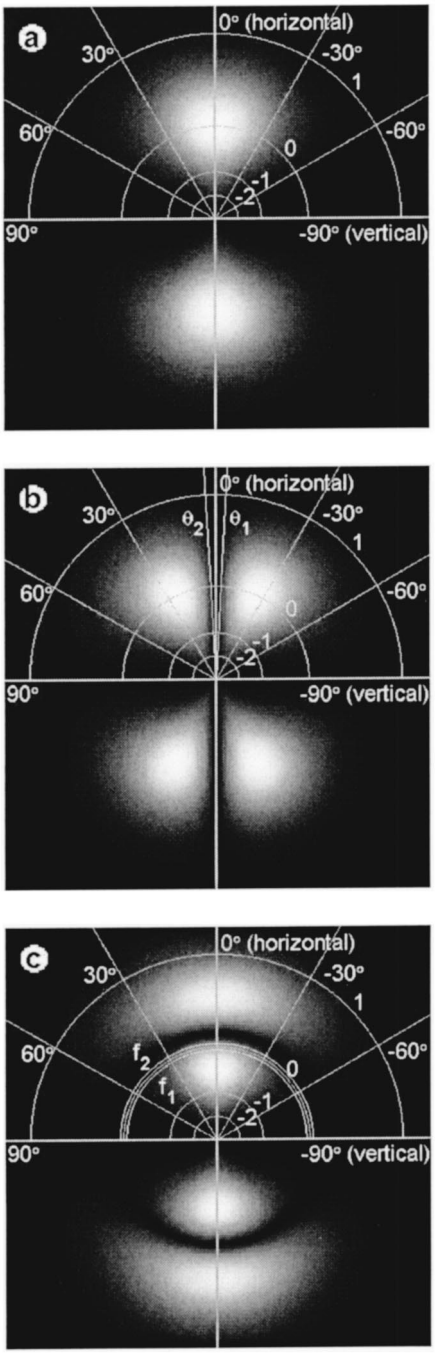


Fig. 2. a, Idealized amplitude spectrum of an unmodulated texture at dc spatial frequency (5 cpd) and dc orientation (horizontal). Besides a scaling factor, this corresponds also to the *difference* between the amplitude spectra of the two regions of a CM texture. b, Absolute difference between the idealized amplitude spectra of the two texture regions of an OM texture at an amplitude of modulation of 4°. c, Absolute difference between the idealized amplitude spectra of the two texture regions of an FM texture at an amplitude of modulation of 0.1 octave. FRF mechanisms tuned to orientations/spatial frequencies at the peaks in the spectral-difference distribution will be maximally responsive to the texture modulation.

It follows that, in theory, the distribution of active FRF mechanisms across the orientation and spatial-frequency preferences of their first-order filters may be used to disambiguate different types of texture modulation. This strategy would require that mechanisms be labeled[16] with regard to the identity of their first-order filters. Recently, Prins and Mussap[10,17] showed that FRF mechanisms are labeled with respect to the location of their receptive fields in the visual field such that second-order modulations can be accurately localized. If FRF mechanisms are, indeed, labeled with respect to the tuning preferences of their front-end filters also, the nature of the texture modulation may be resolved by determining the distribution of active mechanisms across their first-order orientation and spatial-frequency preferences.

If the mechanisms involved are indeed labeled with respect to their front-end input and the distribution of active texture mechanisms is used to determine the nature of the texture modulation, it follows that the detection and the identification of texture modulations are subserved by the same mechanisms. Our present goal is to test this idea. Our strategy is similar to that employed by Watson and Robson.[16] Observers are presented with texture modulations and are asked to perform both a detection task ("Which of two intervals contained a modulated texture?") and a discrimination task ("What type of texture modulation was presented?"). If detection and identification are indeed subserved by the same mechanisms, detection and identification performance should be described by identical psychometric functions.

To pre-empt, we find that the three different types of textures can be identified as well as they can be detected when OM textures are paired with FM textures or with CM textures. These results are compatible with the notion that the identification of texture modulations in these cases is subserved by the same mechanisms that underlie their detection. We have argued above that this is possible because the mechanisms that detect different types of modulation are different for each of the types of modulation, and hence the identification of active mechanisms reveals the type of modulation. However, when CM textures are paired with FM textures, identification performance is significantly worse than detection performance. We discuss several possible reasons why CM and FM textures might be confused at detection threshold.

2. STIMULI

Textures consisted of randomly positioned Gabor micropatterns,

$$L(x, y) = L_0 + L_m \cos\{2\pi f[x \sin(\theta) + y \cos(\theta)] + \phi\} \times \exp[-(x^2 + y^2)/(2\sigma_e{}^2)], \quad (1a)$$

where

$$\sigma_e = (f\pi)^{-1}[0.5 \ln(2)]^{1/2}(2^{1.5} + 1)(2^{1.5} - 1)^{-1}, \quad (1b)$$

L_0 is mean luminance (124.0 cd m^{-2}), L_m is the luminance modulation amplitude (61.9 cd m^{-2}), f is spatial frequency, θ is orientation, ϕ is the phase of the cosine component ($\phi = \pi/2$ or $3\pi/2$), and σ_e is the standard deviation of the Gaussian envelope. The value of σ_e covaried with the value of f so as to keep the spatial-frequency bandwidth (full-width at half-height) of the micropatterns constant at 1.5 octaves across different frequencies.

The stimulus area was circular with a radius of 4.1°. The positioning of the Gabor micropatterns within the stimulus area was random under the constraint that the minimum center-to-center separation between any two micropatterns be equal to the standard deviation of the micropattern envelope [σ_e in formula (1)]. This constraint ensured approximately equal coverage across the stimulus area and avoided excessive luminance summation (which would otherwise occur where several micropatterns happened to overlap). The luminance modulations, but not the dc components, of the Gabor micropatterns were summed where they overlapped. The number of micropatterns per unit area varied as a function of the spatial frequency of the micropatterns such that the number of micropatterns per one degree squared stimulus area was $0.6/\sigma_e{}^2$. For example, at the center spatial frequency of 5 cpd, the number of micropatterns per one degree squared was equal to 97. Implementing this density constraint ensured that coverage and rms contrast between regions of differing spatial frequency were, at least statistically, equal.

The spatial frequency, orientation, or contrast of the individual Gabor micropatterns was square-wave modulated around a center spatial frequency of 5 cpd, a center orientation of 0° (horizontal) and a Michelson luminance contrast of 0.21. In the case of spatial-frequency modulation the two spatial frequencies present in the texture differed from the center spatial frequency by an equal number of log-frequency units. The phase of the texture modulation was randomized across trials. The stimulus area contained two full cycles of the square-wave modulation, corresponding to a bar width of 2.1°. The bars were oriented vertically. Figure 3a displays an example CM texture, with a modulation amplitude of 25%. Figure 3b displays an example of OM texture, with a modulation amplitude of 8° (peak-to-trough difference of 16°). Figure 3c displays an example FM texture, with a modulation amplitude of 0.2 octave.

Textures were generated online in computer memory and were presented on a Clinton Monoray monitor controlled by a Cambridge Research Systems VSG 2/5 graphics board. At the employed viewing distance of 100 cm, the resolution of the monitor was equal to 42.3 pixels/deg.

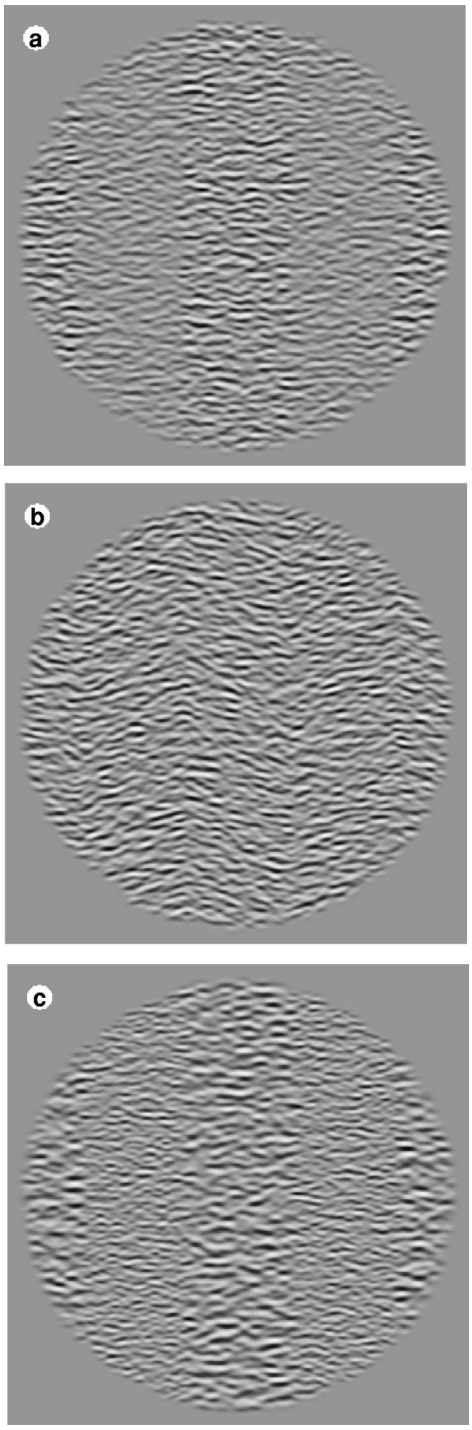


Fig. 3. Example textures. a, CM texture; modulation amplitude is 25%. b, OM texture; modulation amplitude is 8°. c, FM texture; modulation amplitude is 0.2 octave.

3. PROCEDURE

Each trial consisted of the presentation of two textures, one of which was modulated in orientation, spatial frequency, or contrast. The task of the observers (the two authors, NP and FK; and two naïve observers, HW and AW) was twofold. The observers were to indicate whether the modulated texture was presented in the first or the second interval and also to indicate the nature of the texture modulation. Observers indicated their re-

sponses through button presses. For three observers (FK, HW, and AW) the order of the two responses was as above. As a control, the response order was reversed in half the blocks for NP. After both responses were made, feedback was given in the form of two consecutive beeps (the order of which matched the response order, the pitch of which indicated whether the response was correct or incorrect). In any given block of trials only two types of modulation were presented. There were thus three types of blocks, one pairing OM and FM textures, one pairing OM and CM textures, and one pairing FM and CM textures.

The modulation depth on each trial was determined by the method of constant stimuli. Each type of modulation was presented at eight equally spaced depths of modulation. However, owing to an error in the computer program governing stimulus presentation and response collection, responses made to trials in which the depth of modulation of the test texture was 0 were miscoded, and hence discarded, leaving data for seven depths of modulation. An appropriate range of modulation depths was determined for each type of modulation and each observer separately in pilot trials.

Each texture in a trial was presented for 200 ms, separated by an interval in which a blank screen (at mean luminance) was presented for 500 ms. After the observer completed the two required responses, feedback was given and the next trial was initiated. The different pairings of texture types were run in the following order: OM–FM, OM–CM, FM–CM. All observers had normal or corrected-to-normal vision.

4. ANALYSIS

The stimulus levels for the different modalities (OM, FM, and CM) were first normalized into unitless measures so that responses from the different modalities could be meaningfully pooled into combined psychometric curves. To accomplish this we assumed that the functions describing the probabilities that the underlying mechanisms could detect the stimulus were accurately described by Weibull curves for all three types of texture modulation. To normalize the units of measurement, we fitted the detection data for the three different types of texture with individual Weibull curves. The normalized units of measurement were then defined as

$$s_{\mathrm{norm}} = (s/\alpha)^{\beta/2}, \tag{2}$$

where s_{norm} is the depth of modulation in normalized units, s is the depth of modulation in the original units of measurement (i.e., degrees, octaves, or percent), and α and β are the threshold and slope parameters of the best-fitting Weibull curve. In other words, all depth-of-modulation levels were transformed such that the threshold and slope parameters of the best-fitting Weibull curves were 1 and 2, respectively, for all three types of texture modulation after normalization. It should be stressed that this transformation in no way affects the goodness of fit of the curve fits; it merely serves to permit psychometric curves from different types of texture modulation to be combined on a single unitless scale.

The question we set out to answer asks, Is the mechanism that underlies modulation detection labeled such that at detection threshold the texture type can also be identified? If so, the psychometric functions describing the detection and identification performance of the underlying mechanisms should be identical. We cannot determine whether this is so by directly comparing proportions correct of the detection and the identification task. The reason is that whereas the proportion correct for any given stimulus level on the detection task will not be influenced by a response bias, the proportion correct for any given stimulus level on the identification task will be influenced by response bias. This is not because of any fundamental difference in the type of task; the difference is purely one of a statistical nature. There is a similar bias present in the detection task. An observer unable to detect the stimulus might have a propensity toward responding "interval 1." However, this response bias is simply averaged out when we combine trials in which the stimulus was presented in the first interval with those in which the stimulus was presented in the second interval. The resulting function describing proportion correct thus becomes

$$P(\mathrm{correct}|s_{\mathrm{norm}}) = \gamma + (1 - \gamma)\omega_d(s_{\mathrm{norm}}, \alpha_d, \beta_d), \tag{3}$$

where γ corresponds to the guessing parameter (here, 0.5) and $\omega_d(s_{\mathrm{norm}}, \alpha_d, \beta_d)$ corresponds to the Weibull function (with threshold parameter α_d and slope parameter β_d) describing the probability that the underlying mechanism is able to detect the stimulus as a function of normalized stimulus level s_{norm}. Note that α_d and β_d will have values of 1 and 2, respectively, given that stimulus levels were normalized in that manner [Eq. (2)].

In the case of the identification task, however, we cannot simply average out response bias in this manner because responses in trials in which the type of modulation was, say, OM are measured at (normalized) stimulus levels different from those in trials in which the type of modulation was, say, FM. Hence, whereas the proportion correct for different stimulus levels in the detection task follows a single curve, the effect of response bias splits the proportions correct in the identification task into two separate curves. One of the curves describes proportion correct when the stimulus presented was one type of texture modulation, and the other curve describes proportion correct when the stimulus presented was the other type of texture modulation. However, the normalization of the stimulus intensity levels will enable us to model the two split curves with a single curve. This underlying curve describes the proportion of trials on which the mechanisms mediating performance are able to identify the stimulus as a function of normalized stimulus intensity. The underlying function will equal zero when stimulus intensity equals zero and will asymptote toward unity at high levels of stimulus intensity.

In order to model the identification function, we made the following assumptions: If the mechanism is able to identify the stimulus, the observer will make the correct response; but if the mechanism is not able to identify the stimulus, the observer will guess. We assume that the function describing the identification probabilities is a Weibull function $\omega_i(s_{\mathrm{norm}}, \alpha_i, \beta_i)$. The guess is subject

to bias such that observers will respond "stimulus type A" with probability P_A and "stimulus type B" with probability $1 - P_A$. We further assume that P_A is constant and hence independent of $\omega_i(s_{\text{norm}}, \alpha_i, \beta_i)$.

The above assumptions lead to the following function describing the probability of a correct identification given that the stimulus presented was of type A and was presented at stimulus intensity s_{norm}:

$$P(A|\{A, s_{\text{norm}}\}) = \omega_i(s_{\text{norm}}, \alpha_i, \beta_i) + [1 - \omega_i(s_{\text{norm}}, \alpha_i, \beta_i)]P_A,$$

which can be rewritten as

$$P(A|\{A, s_{\text{norm}}\}) = P_A + (1 - P_A)\omega_i(s_{\text{norm}}, \alpha_i, \beta_i). \quad (4a)$$

Similarly, the probability of a correct identification given that the stimulus presented was of type B is given as

$$P(B|\{B, s_{\text{norm}}\}) = (1 - P_A) + P_A\omega_i(s_{\text{norm}}, \alpha_i, \beta_i). \quad (4b)$$

Testing the equivalence of the mechanisms underlying the detection and identification of our textures amounts to testing the equivalence of α_d and α_i and the equivalence of β_d and β_i. To do so, we modeled our detection and discrimination results using two models. Model 1 assumes a common mechanism underlying both detection and identification of the textures, i.e., $\alpha_d = \alpha_i$ and $\beta_d = \beta_i$. We will refer to the common parameters as α_c and β_c. Model 2, on the other hand, assumes separate mechanisms, i.e., $\alpha_d \neq \alpha_i$ and/or $\beta_d \neq \beta_i$. In other words, under Model 1 the proportions correct for the detection and identification task are modeled as follows:

Model 1:

detection

$$P(\text{correct}|s_{\text{norm}}) = 0.5 + (1 - 0.5)\omega_c(s_{\text{norm}}, \alpha_c, \beta_c),$$

identification

$$P(A|\{A, s_{\text{norm}}\}) = P_A + (1 - P_A)\omega_c(s_{\text{norm}}, \alpha_c, \beta_c),$$

$$P(B|\{B, s_{\text{norm}}\}) = (1 - P_A) + P_A\omega_c(s_{\text{norm}}, \alpha_c, \beta_c).$$

Under model 2, the proportions correct for the detection and identification task are modeled as follows:

Model 2:

detection

$$P(\text{correct}|s_{\text{norm}}) = 0.5 + (1 - 0.5)\omega_d(s_{\text{norm}}, \alpha_d, \beta_d),$$

identification

$$P(A|\{A, s_{\text{norm}}\}) = P_A + (1 - P_A)\omega_i(s_{\text{norm}}, \alpha_i, \beta_i),$$

$$P(B|\{B, s_{\text{norm}}\}) = (1 - P_A) + P_A\omega_i(s_{\text{norm}}, \alpha_i, \beta_i).$$

Model 1 requires the estimation of three parameters: α_c, β_c, and P_A. Model 2 requires the estimation of five parameters: α_d, α_i, β_d, β_i, and P_A. Our test statistic is

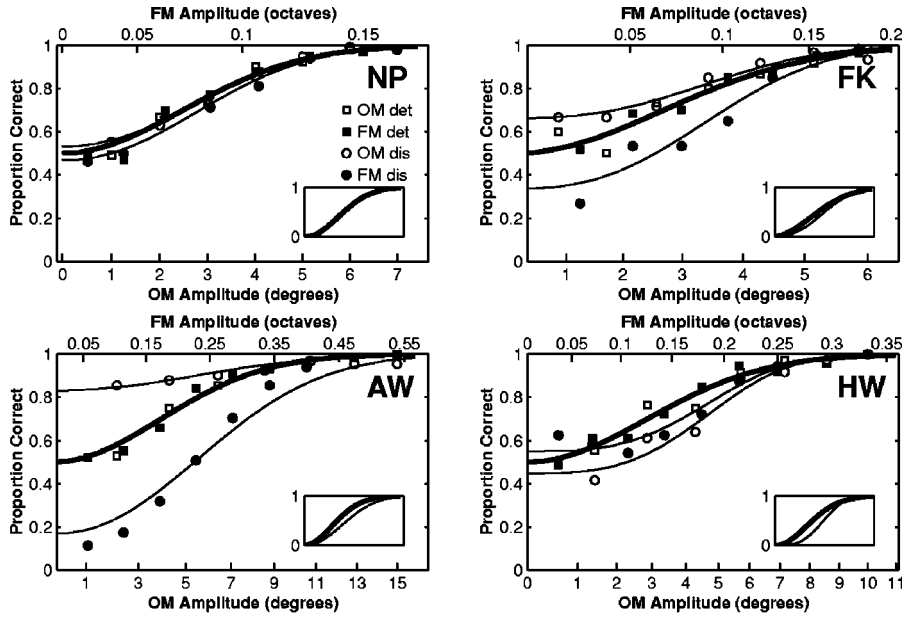
$$\lambda = -2 \log_e[L(\mathbf{X}|\tilde{\alpha}_c, \tilde{\beta}_c, \tilde{P}_A)/L(\mathbf{X}|\tilde{\alpha}_d, \tilde{\beta}_d, \tilde{\alpha}_i, \tilde{\beta}_i, \tilde{P}_A)],$$

where $L(\mathbf{X}|\tilde{\alpha}_c, \tilde{\beta}_c, \tilde{P}_A)$ represents the likelihood of the entire set of responses ($\mathbf{X}$) in a testing condition under Model 1, with maximum-likelihood estimates of the three parameters of Model 1. Similarly, $L(\mathbf{X}|\tilde{\alpha}_d, \tilde{\beta}_d, \tilde{\alpha}_i, \tilde{\beta}_i, \tilde{P}_A)$ represents the likelihood of the same set of responses under Model 2, with maximum-likelihood estimates of the five parameters of Model 2. Since the parameter space under Model 1 is a subset of the parameter space under Model 2, the random variable λ is asymptotically distributed as χ^2 with degrees of freedom equal to 2 (the difference in the number of independent parameters between the two models).[18]

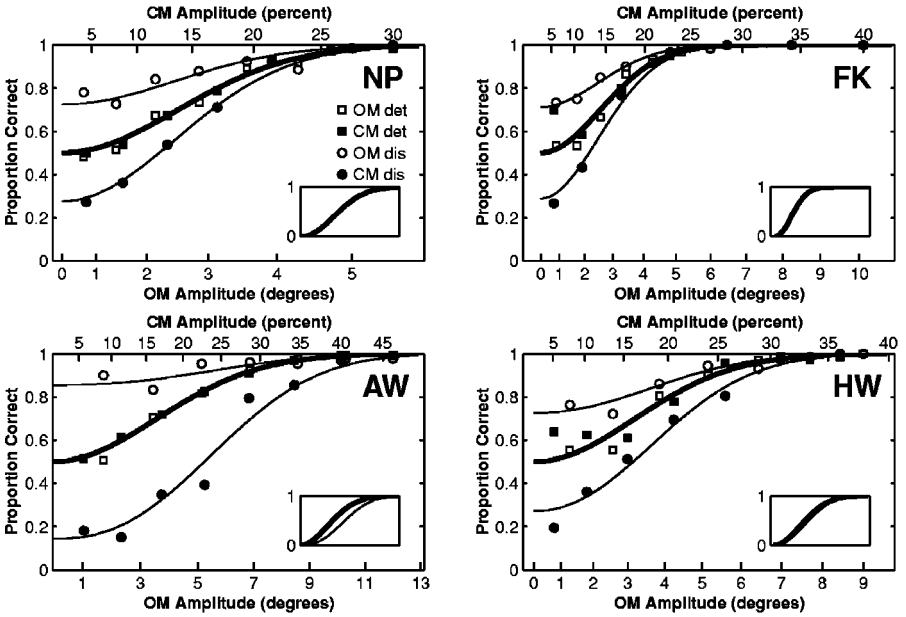
5. RESULTS

Proportions correct as a function of the modulation depth of the texture are plotted separately for all observers and each of the three different combinations of texture type in Fig. 4. The abscissas are linear in terms of the normalized depth-of-modulation values [s_{norm}, formula (2)]. Note that the functions describing proportions correct in the identification task are split into two distinct curves because of response bias, which in some cases was strong [consider, for example, the identification results for AW in the OM–CM combination, Fig. 4(b)]. It is imperative for meaningful comparison of detection and identification performance to remove the bias effect from the data by modeling the probabilities of correct detection and identification by the underlying mechanisms as described in Section 4. Goodness of fit of the individual curves was assessed as suggested by Wichmann and Hill.[19] That is, for each of the fits a "deviance" score was calculated, which was then compared with a distribution ($n = 500{,}000$) of deviance scores obtained by Monte Carlo simulations individually for each of the curves. Reported in Table 1 are the proportions of simulated deviance scores that were higher (indicating a worse fit) than the empirically obtained deviance score. Only four[20] of the twenty-four deviance scores reached statistical significance ($p < 0.05$), indicating that the data were well modeled by the Weibull function even after normalization of depth-of-modulation values and combination of the different texture types. The figure insets show the estimated functions describing the probabilities that the underlying mechanisms were able to detect the stimulus [i.e., $\omega_d(s_{\text{norm}}, \alpha_d, \beta_d)$ in formula (3); heavy curve] and identify the stimulus [i.e., $\omega_i(s_{\text{norm}}, \alpha_i, \beta_i)$ in formula (4); light curve] as a function of the (normalized) modulation depth. The abscissas of the insets are linearly scaled versions of those of the corresponding main figures.

The results of the authors (FK and NP) indicate that for the OM–FM and OM–CM combinations the functions describing the probabilities that the underlying mechanisms will be able to detect the stimulus are virtually identical to those describing the probabilities that the mechanisms will be able to identify the stimulus. When FM and CM textures were paired, however, identification performance was significantly worse than detection performance. These statements can be verified either by visual inspection of the insets in Fig. 4 or by the results of the statistical analysis. Table 1 lists the identification



(a)



(b)

Fig. 4. Continues on next page.

threshold-to-detection threshold ratios, the λ values associated with the statistical comparison between the detection and identification curves (see Section 4), and their corresponding *p* values. The statistical analysis supports both claims: None of the λ values reached magnitudes anywhere near statistical significance in the conditions employing OM–FM or OM–CM pairings. For both authors the λ values did, however, reach statistically significant levels when FM and CM textures were paired.

The results of the naïve observers are not as straightforward. All but one of the six λ values were statistically significant, indicating that in all these conditions identification performance was significantly worse than detection performance for these observers. The one exception lies in the results of HW in the condition in which OM and CM textures were paired. Identification performance here was no worse (at least not significantly) than detection performance.

However, the naïve observers follow the general trend of the authors in that detection and identification performance are very similar (even though significantly different in the statistical sense) in the OM–FM and OM–CM texture combinations but not in the FM–CM texture combination. We hypothesize that the slightly worse performance of the naïve subjects on the identification task reflects secondary factors related to the task (such as memory or response mapping). Remember that each trial required two responses: one indicating in which in-

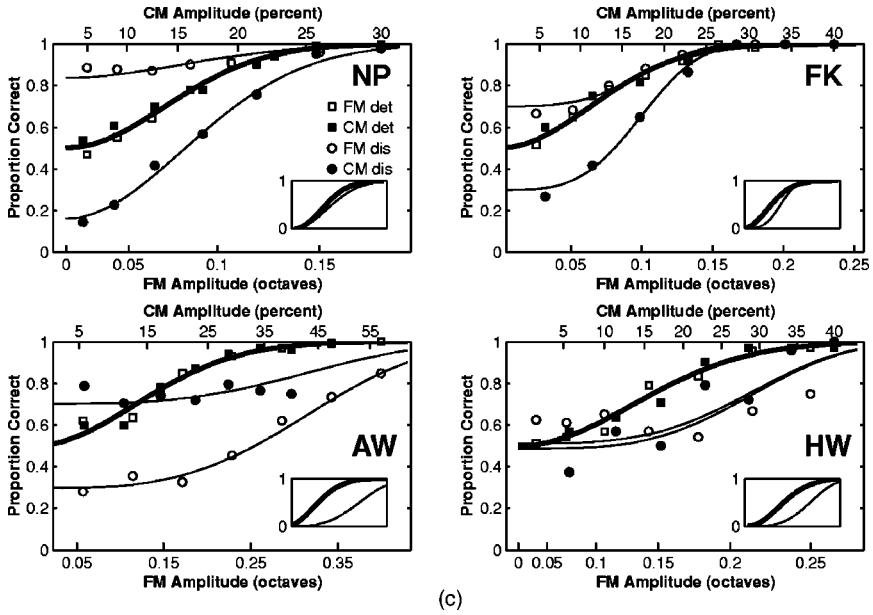


Fig. 4. Detection and identification performance. Plotted are proportion correct detection and identification as a function of depth of modulation for (a) OM–FM pairing of textures, (b) OM–CM pairing of textures, and (c) FM–CM pairing of textures. Abscissas are linear with respect to depth of modulation in normalized units. Insets show modeled proportions of correct detection [i.e., $\omega_d(s_{\text{norm}}, \alpha_d, \beta_d)$ of formula (3); heavy curve] and proportions of correct identification [i.e., $\omega_i(s_{\text{norm}}, \alpha_i, \beta_i)$ of formula (4); light curve] produced by the underlying mechanisms for visual comparison. Abscissas of insets are linearly scaled versions of those of the main figures.

Table 1. Model Fits[a]

Condition	Observer	p_{det}	p_{dis}	α_i/α_d	λ	$p(\lambda)$
OM–FM	NP	0.70	0.85	1.05	1.23	0.54
	FK	0.95	0.43	1.10	3.08	0.21
	AW	0.85	0.00	1.28	35.90	1×10^{-8}
	HW	0.69	0.81	1.22	15.27	5×10^{-4}
OM–CM	NP	0.95	$0.01^{(20)}$	0.98	0.17	0.92
	FK	0.47	0.95	0.98	0.06	0.97
	AW	1.00	0.00	1.33	46.04	2×10^{-10}
	HW	0.35	0.60	1.10	2.10	0.35
FM–CM	NP	0.87	0.42	1.20	17.78	1×10^{-4}
	FK	0.83	0.99	1.23	12.09	2×10^{-3}
	AW	0.67	0.33	2.12	356.02	~ 0
	HW	0.87	0.00	1.60	73.66	1×10^{-10}

[a] Shown are p values for the goodness of fit of the Weibull curves fitted to the detection and discrimination data, obtained by assessing goodness of fit by means of the deviance statistic and generating sampling distributions of this statistic for each individual condition by Monte Carlo simulation, as suggested by Wichmann and Hill.[19] The p values indicate the proportion of simulations in which a deviance score was obtained that was higher (indicating a worse fit) than the empirical deviance score. Also shown are identification-to-detection threshold ratios, observed λ, and corresponding p values. The p values are interpreted as the probability of observing the indicated λ value (or a higher λ value) when detection and identification performance are statistically equal. See text for details.

terval the modulated texture was presented and the other indicating what type of texture was presented. It is interesting to note that one of the naïve observers (HW), perhaps through a practice effect, did produce nearly identical detection and identification curves in the second condition in which he was tested (OM–CM pairing).

6. DISCUSSION

Our results indicate that texture modulations are identified as readily as they are detected when OM textures are to be distinguished from either FM or CM textures. However, when FM textures are paired with CM textures, identification requires a higher depth of modulation compared with that needed for detection. In Section 1 we speculated that texture modulations could potentially be identified by a mechanism that determines the nature of first-order filters that produce maximum differential activation between different texture regions. We showed in Fig. 2 that these filters are, in theory, different for each of the three different types of texture modulation used in this study.

We will discuss three possible reasons why FM textures might be difficult to distinguish from CM textures within the proposed framework. In theory, an FM texture should lead to two groups of active FRFs: one with first-

order filters tuned to spatial frequencies approximately one half octave below the center spatial frequency of the texture, the other with first-order filters tuned to spatial frequencies approximately one half octave above the center spatial frequency of the texture. However, as mentioned in Section 1, FM textures are processed primarily by texture mechanisms with first-order filters tuned to spatial frequencies centered approximately one octave below the texture's center spatial frequency. Prins and Kingdom[14] found only a slight hint that texture mechanisms tuned to spatial frequencies above the center spatial frequency were involved in the processing of FM textures. This is at least partly due to the physics of the stimulus. The difference in energy between the two regions of our FM textures is simply smaller at higher-than-dc frequencies than the difference at lower-than-dc frequencies. This can be verified by studying Fig. 2c where the outer (high-frequency) differential-energy blob is of lower peak amplitude than the inner (low-frequency) blob. Moreover, the human visual system is differentially sensitive to different spatial frequencies. Contrast sensitivity drops off quite rapidly at high spatial frequencies.[21] At threshold amplitudes of FM modulation, then, there will be only one group of active FRFs, namely, those tuned to lower-than-dc spatial frequencies [i.e., the inner (low frequency) blob in Fig. 2c]. As discussed in Section 1, however, a CM texture will also lead to only one group of active FRF mechanisms (i.e., the single blob in Fig. 2a). Moreover, the low-frequency blob in Fig. 2c and the blob in Fig. 2a are both centered at the same orientation (horizontal) and show considerable overlap in spatial frequency. The overlap corresponds to mechanisms that are activated by both FM and CM textures. Especially at low depths of modulation (i.e., low signal-to-noise ratio) it may be that on a proportion of trials detection is triggered by a mechanism that receives its front-end input from filters that have tuning properties that fit both the low-frequency FM blob and the single CM blob. In such a case it is not possible to determine whether the mechanism was activated by an FM texture or a CM texture, and hence identification will fail.

A second explanation for the confusability of CM and FM textures is that contrast normalization does not occur across different spatial frequencies.[22] Some proposed second-order models involve a stage of local contrast normalization (e.g., Ref. 7) that serves to reduce local contrast variations. As a result of the normalization, the overall response strength of second-order mechanisms will depend more reliably on the spatial properties (e.g., orientation, spatial frequency) of the stimulus as random contrast variations are removed. However, when contrast normalization indeed fails between regions of different spatial frequency, random contrast variations are preserved. These contrast variations will be best detected by FRF mechanisms with front-end luminance filters tuned to orientations and spatial frequencies that are close to the dc orientation and spatial frequency of the texture. The failure of contrast normalization may thus lead to occasional activation patterns of FRF mechanisms that are normally indicative of a CM texture. On the other hand, neither of the two activation patterns that can be elicited by an FM texture is similar to the activation pattern elicited by an OM texture. In this case, the modulation can be identified as an OM texture when the pattern of activation fits that of an OM texture or as an FM texture when the pattern of activation fits that of a typical CM or a typical FM texture. As a result, identification performance in the condition in which FM textures are paired with OM textures is essentially identical to detection performance.

A third possible reason for the confusability of FM and CM textures might be that spatial-frequency information is simply not, or not clearly, labeled. Note that both of the conditions in which identification performance was nearly identical to detection performance (OM–CM and OM–FM) involved OM textures. It is possible that the visual system does not place much value on spatial frequency in retinal terms but rather has a tendency to treat different retinal spatial frequencies as having arisen from identical object spatial frequencies (i.e., size constancy).

What then, is the mechanism of identification? Our results are compatible with the idea that the distribution of active FRF mechanism across the orientation and spatial-frequency preferences of their first-order filters is monitored and that this distribution of activity is used to disambiguate different types of texture modulations. The nature of the mechanism that determines the distribution of active FRF mechanism remains, for now, speculative. In its simplest form, the mechanism that identifies a texture as OM might simply consist of an AND gate connecting FRF mechanisms that receive their first-order input from filters tuned to clockwise (relative to the center orientation of the texture) orientations and FRF mechanisms that receive their first-order input from filters tuned to counterclockwise (again relative to the center orientation of the texture) orientations. The two groups of FRF mechanisms connected by the AND gate would have to respond to the modulation in counterphase. For example, an ON-center OFF-surround FRF mechanism tuned to 30° clockwise from horizontal and a counterphase (OFF-center ON-surround) FRF mechanism tuned to 30° counterclockwise from horizontal will respond in phase to our OM textures (see Section 1). If both of these mechanisms are activated by a texture (signaled by the AND gate), it would indicate that the texture is OM modulated. Similarly, the mechanism that identifies a texture as FM might consist of an AND gate connecting FRF mechanisms that receive their first-order input from filters tuned to spatial frequencies that lie on either side of the center spatial frequency of the texture. Again the mechanisms connected by the AND gate would be in counterphase.

Another possibility is that the visual system learns to monitor the FRF mechanisms that receive their first-order input from channels that are relevant for the specific situation. For example, in our OM–CM task the visual system may monitor only three groups of FRF mechanisms: FRF mechanisms receiving first-order input from filters tuned to (1) the dc orientation and dc spatial frequency, (2) the dc spatial frequency and orientations ~30° clockwise from the dc orientation, and (3) the dc spatial frequency and orientations ~30° counterclockwise from the dc orientation. Activation in group (1) of the FRF mechanisms would signal a CM texture, whereas

activation in group (2) or group (3) would signal an OM texture. However, as mentioned, the exact nature of the mechanism identifying texture modulations remains speculative at this point and a possible subject of future investigations.

ACKNOWLEDGMENTS

This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada under grant OGP 0121713.

Nicolaas Prins, the corresponding author, may be reached through e-mail: nicolaas.prins@mcgill.ca.

REFERENCES AND NOTES

1. H. Wilson, "Non-Fourier cortical processes in texture, form, and motion perception," Cereb. Cortex **13**, 445–477 (1999).
2. C. L. Baker, Jr., "Central neural mechanisms for detecting second-order motion," Curr. Opin. Neurobiol. **9**, 461–466 (1999).
3. A. Sutter, G. Sperling, and C. Chubb, "Measuring the spatial frequency selectivity of second-order texture mechanisms," Vision Res. **35**, 915–924 (1995).
4. S. C. Dakin and I. Mareschal, "Sensitivity to contrast modulation depends on carrier spatial frequency and orientation," Vision Res. **40**, 311–329 (2000).
5. J. Malik and P. Perona, "Preattentive texture discrimination with early vision mechanisms," J. Opt. Soc. Am. A **7**, 923–932 (1990).
6. J. R. Bergen and M. S. Landy, "Computational modeling of visual texture segregation," in *Computational Models of Visual Processing*, M. S. Landy and J. A. Movshon, eds. (MIT Press, Cambridge, Mass., 1991), pp. 253–271.
7. M. S. Landy and J. R. Bergen, "Texture segregation and orientation gradient," Vision Res. **31**, 679–691 (1991).
8. F. A. A. Kingdom and D. R. T. Keeble, "A linear systems approach to the detection of both abrupt and smooth spatial variations in orientation-defined textures," Vision Res. **36**, 409–420 (1996).
9. R. Gray and D. Regan, "Spatial frequency discrimination and detection characteristics for gratings defined by orientation texture," Vision Res. **38**, 2601–2617 (1998).
10. N. Prins and A. J. Mussap, "Alignment of orientation-modulated textures," Vision Res. **40**, 3567–3573 (2000).
11. N. Graham, A. Sutter, and C. Venkatesan, "Spatial-frequency- and orientation-selectivity of simple and complex channels in region segregation," Vision Res. **33**, 1893–1911 (1993).
12. S. A. Arsenault, F. Wilkinson, and F. A. A. Kingdom, "Modulation frequency and orientation tuning of second-order texture mechanisms," J. Opt. Soc. Am. A **16**, 427–435 (1999).
13. N. Graham and A. Sutter, "Spatial summation in simple (Fourier) and complex (non-Fourier) texture channels," Vision Res. **38**, 231–257 (1999).
14. N. Prins and F. A. A. Kingdom, "Orientation- and frequency-modulated textures at low depths of modulation are processed by off-orientation and off-frequency texture mechanisms," Vision Res. **42**, 705–713 (2002).
15. We model the spectral amplitude distribution here as $H(f, \theta) = \exp(-0.5[(f - f_0)/(f_0\sigma_f)]^2)\exp(-0.5[(\theta - \theta_0)/\sigma_\theta]^2)$, where f is frequency, f_0 is the dc spatial frequency (5 cpd), σ_f is a constant determining spatial-frequency bandwidth and set at a value of 0.41, θ is orientation, θ_0 is the dc orientation of the texture (0°, horizontal), and σ_θ is a constant determining orientation bandwidth and set at 25.5. This function describes the average spectral content of the textures used here quite well.[14]
16. A. B. Watson and J. C. Robson, "Discrimination at threshold: labeled detectors in human vision," Vision Res. **21**, 1115–1122 (1981).
17. N. Prins and A. J. Mussap, "Adaptation reveals a neural code for the visual location of orientation change," Perception **30**, 669–680 (2001).
18. P. G. Hoel, S. C. Port, and C. J. Stone, *Introduction to Statistical Theory* (Houghton Mifflin, Boston, Mass., 1971).
19. F. A. Wichmann and N. J. Hill, "The psychometric function: I. Fitting, sampling, and goodness of fit," Percept. Psychophys. **63**, 1293–1313 (2001).
20. The statistical test employed to assess goodness of fit indicated a poor fit for the identification curve of NP in the OM–CM condition ($p < 0.01$). However, as the reader may verify by inspection of Fig. 4, the data appear to fit the curve quite well. As it turns out, the poor goodness of fit is due almost entirely to one data point, namely, identification performance for OM textures at the second-highest value of modulation amplitude. This suggests that the poor fit is likely a spurious result. The p value of the deviance score calculated when this data point is omitted is 0.57.
21. F. W. Campbell and J. G. Robson, "Application of Fourier analysis to the visibility of gratings," J. Physiol. **197**, 551–566 (1968).
22. L. A. Olzak and J. P. Thomas, "Neural recoding in human pattern vision: model and mechanisms," Vision Res. **39**, 231–256 (1999).