PERGAMON

Vision Research 42 (2002) 311–330

Vision Research

www.elsevier.com/locate/visres

Attentional control of texture orientation judgments

Charles Chubb *, Jennifer Talevich

Department of Cognitive Sciences, University of California at Irvine, Irvine, CA 92697-5100, USA

Received 21 February 2001; received in revised form 20 August 2001

Abstract

Recent models of texture processing use low level, spatially parallel computations to extract texture properties. The rapid, preattentive nature of texture segregation suggests that these computations are bottom–up in nature. However, the immunity of texture judgments to top–down influences remains to be tested. Here we investigate the degree to which judgments of texture orientation are susceptible to top–down attentional control. Observers view a brief display composed of variously luminant texture elements (line segments) alternately (in checkerboard arrangement) oriented up/right (at 71.5°) or up/left (at 108.5°), and are asked to make various judgments. In a given task, the observer attempts on each trial to judge which oriented population of line segments has an intensity histogram that best matches a given target histogram. Performance demonstrates adaptive flexibility across different tasks, suggesting that observers are able to exercise significant top–down control over texture orientation computations. Specifically, observers can attend selectively to positive contrast texture elements, to negative contrast texture elements, or to high (positive and negative) contrast texture elements. More generally, observers perform well if the target histogram can be approximated by a weighted average of positive and negative half-wave rectifiers. Performance is poor for histograms that cannot be captured in this way. These results suggest that attentional control in these tasks is limited to adjusting the relative gain of the on- and off-center systems. © 2002 Published by Elsevier Science Ltd.

Keywords: Attention; Spatial vision; Histogram contrast analysis; On-center system; Off-center system

1. Introduction

The nearly 40 years of effort devoted to texture processing has yielded a class of models termed "back pocket models" (Chubb & Landy, 1991), because texture perception researchers routinely pull such models from their back pockets to account for new instances of texture segregation. These models all share three basic processing stages, although they vary in the details of each stage (Beck, 1982; Beck, Graham, & Sutter, 1991; Beck, Prazdny, & Rosenfeld, 1983; Bergen & Adelson, 1988; Bergen & Landy, 1991; Bovik, Clark, & Geisler, 1990; Caelli, 1985; Clark & Bovik, 1989; Fogel & Sagi, 1989; Graham, 1991; Graham, 1994; Graham, Beck, & Sutter, 1992; Graham, Sutter, & Venkatesan, 1993; Grossberg & Mingolla, 1985; Julesz, 1981; Julesz & Bergen, 1983; Knutsson & Granlund, 1983; Landy & Bergen, 1991; Malik & Perona, 1990; Nothdurft, 1994; Rubenstein & Sagi, 1990; Sutter, Beck, & Graham, 1989; Sutter & Graham, 1995).

In the first processing stage, a battery of spatially local, linear filters is applied to the visual input. The receptive fields of these filters are usually hypothesized to resemble simple cell receptive fields of various spatial frequencies and orientations. It is these linear filters that give the hypothetical texture-sensing transformations their sensitivity to local image structure.

In the second stage, the outputs of these linear filters (usually thought of as "neural images" (Robson, 1980)) are passed through some pointwise nonlinearity (e.g., a rectifier) whose purpose is to transform highly variable regions of the stage-1 output into regions of high average value. This stage may involve nonlinear, lateral interactions between units of a given neural array, or even between units of different arrays. For example, cross-orientation normalization (Heeger, 1991, 1992), which has been implicated in threshold masking experiments (Foley, 1994) is a plausible component stage of transformations mediating texture processing. Also evidence is accumulating to suggest that divisive inter-unit normalization to achieve statistical independence of unit

* Corresponding author. Tel.: +1-714-949-1481/+1-949-824-1481; fax: +1-714-949-2517/+1-949-824-2517.

E-mail address: cfchubb@uci.edu (C. Chubb).

0042-6989/02/$ - see front matter © 2002 Published by Elsevier Science Ltd.
PII: S0042-6989(01)00282-6

responses to natural images (Wainwright & Simoncelli, 1999) is likely to be involved.

Taken together, stages 1 and 2 of the back pocket model achieve a battery of spatially local transformations, each of which can be thought of as gauging the level at each location in the input image of some elementary type of "visual stuff". This is in accordance with the proposal that the purpose of the initial stages of visual processing is to measure "the amounts of various kinds of visual 'substances' present in the image" (Adelson & Bergen, 1991). For this reason, we refer to the transformations achieved by the first two stages of the back pocket model as "stuff-sensing" transformations.

In the third stage of the back pocket model of texture segregation, distinctions (e.g., boundaries, modulations, gradations) are drawn throughout the visual field based on the responses of the stuff-sensing transformations achieved in stages 1 and 2. In particular, two regions A and B are assumed to emerge as visually distinct if (i) the responses of all of the stuff-sensing transformations are relatively homogeneous within each of A and B, while (ii) the average response of one or more of the stuff-sensing transformations is significantly different between A and B (e.g., Malik, Belongie, Shi, & Leung, 1997). Recent work (Rosenholtz, 1999) underscores the importance of taking account of the internal variability of A and B in predicting the degree to which they segregate.

It is important to note that in all versions of the back pocket model, processing of the visual input is implicitly assumed to be bottom–up. The activation of any given neuron in one of the hypothesized neural arrays mediating texture segregation is assumed to depend exclusively on the pattern of light stimulating the retina. In particular, the computations performed under standard back pocket models are supposed to be invariant with respect to the attentional state of the observer; emergent boundaries are explicitly held to be *preattentive*.

However, there are reasons to suppose that attentional control plays an important role in many common texture judgments. Consider the problem faced by a hiker of assessing whether or not the surface of a rock will support the sole of her boot without slipping. The input to this assessment is the image of the rock surface, and the output (which must be quickly derived) is the decision either to trust the rock with her boot or to avoid the rock. To arrive at this decision, one needs to compute some well-chosen statistic, with positive values leading to trusting the rock, and negative values leading to avoidance. The type of statistic that naturally suggests itself for this task requires two stages of computation: first, a carefully selected, spatially local transformation must be applied to the rock surface image. This transformation should be as well-correlated with local boot-gripping potential as the hiker can make it. The application of this transformation yields an output image whose value at each point (x,y) reflects the boot-gripping potential of the rock surface in the neighborhood of (x,y). (For the moment, we leave aside the question of how this image might be coded in the brain.) If indeed, the texture of the rock surface is more or less homogeneous, then a natural way to convert this output image into a decision statistic is to compute the spatial average. If this space-average estimate of boot-gripping potential exceeds the hiker's threshold of safety, then the rock gets trusted; otherwise avoided.

This is one example of a real-world texture judgment. It is easy to think of others. Will the fabric in the shop window be soft enough for my baby? Will I be able to bushwhack through the scrub across the valley? Is the cereal in my bowl fresh or stale? Will I get dirty if I lean against the wall beside me? Is the grass in front of me dry enough to walk in? Which of the towels in the rag bin will serve best to soak up the milk I have just spilled?

In each case, a visual texture judgment is made to assess the suitability of some sort of stuff for some particular purpose, and in each case, we base our judgment on some statistic extracted from the retinal image of the texture. However, the statistic we compute must be specially tailored to each situation. Different purposes require different computations. It is difficult to imagine how such computational flexibility might be achieved without significant, top–down attentional control.

It seems natural to suppose that we recruit for such, everyday texture judgments the same machinery that produces preattentive boundaries between different sorts of texture. How might this work?

We propose the following working model: human vision comprises (as proposed by the back pocket model) a fixed set of hard-wired, stuff-sensing transformations implemented as neural arrays (possibly including intra- and/or inter-array interactions). These arrays perform spatially parallel, bottom–up processing of the visual input continuously in real time, producing dynamic neural images reflecting variations of different sorts of visual stuff throughout the scene. Suppose there are N such stuff-sensing arrays: $Q_1, Q_2, \ldots, Q_N$, and for $i = 1, 2, \ldots, N$, write $Q_i(J)$ for the neural image that results from applying Q_i to a given input image J. In addition, write $Q_i(J)[x,y]$ for the level of activation produced at point (x,y) in array Q_i in response to input image J. Thus, $Q_i(J)[x,y]$ reflects the amount of Q_i-stuff in image J in the neighborhood of (x,y).

We suppose that the computations performed by the stuff-sensing arrays Q_i, $i = 1, 2, \ldots, N$, are strictly bottom–up, as typically assumed under the back pocket model. However, we suppose that for purposes of making everyday texture judgments, attention can be used to control judgments in the following way.

Let R be the restriction of the input image (e.g. the view of the trail ahead as you hike) to the spatial region about which the judgment is to be made (e.g. the rock's surface), and for any function $f : R \to \mathscr{R}$, write $\langle f[x,y] \rangle_R$

for the average value of f taken over all points in R. Then we assume observers can compute statistics of the following sort:

$$G_\alpha(R) = \langle g_\alpha(R)[x,y] \rangle_R, \tag{1}$$

where

$$g_\alpha(R)[x,y] = \sum_{i=1}^{N} \alpha_i Q_i(R)[x,y], \tag{2}$$

for $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_N)$ a vector of "stuff weights" that is at least partially under the attentional control of the observer.

Thus, e.g., in trying to assess whether the surface of a rock will hold her boot without slipping, a hiker needs to select stuff weights α_i, $i = 1, 2, \ldots, N$, so that the quantity $g_\alpha(R)[x,y]$ reflects as well as possible the boot-gripping potential afforded by the rock surface in the neighborhood of point (x,y). Then, by selecting R to be the region of the input image I corresponding to the candidate rock face, the hiker can base her decision about whether to trust the rock on the statistic $G_\alpha(R)$.

Given the working model summarized by Eqs. (1) and (2), several questions emerge. First, what are the elementary visual substances mediating texture judgments? In other words, what local image properties do the transformations Q_i, $i = 1, 2, \ldots, N$, sense? Second, what constraints (if any) limit the weights α_i, $i = 1, 2, \ldots, N$, that can be used to synthesize transformations g_α appropriate for specific texture judgments?

The experiments reported here aim to provide insight into both of these issues. In particular, these experiments investigate (i) whether or not those stuff-sensing arrays Q_i of Eq. (2) sensitive to local texture orientation are selective for contrast polarity, and (ii) how the stuff weights α_i of Eq. (2) can be used to adjust sensitivity to different intensities occurring in ambiguously oriented texture.

2. The current project

A simple but influential variant of the back pocket model (Bergen & Adelson, 1988; Julesz, 1962, 1975; Julesz, Gilbert, Shepp, & Frisch, 1973; Knutsson & Granlund, 1983) proposes that the stuff-sensing arrays used by human vision are tuned to local stimulus energy in various spatial frequency bands. Although there exist well-known counterexamples to the texture energy model (Julesz, Gilbert, & Victor, 1978; Victor & Brodie, 1978; Victor & Conte, 1991; Victor, Conte, Purpura, & Katz, 1994), this simple model seems to account well for most instances of preattentive texture segregation. Under the texture energy model, the impulse responses of the stage 1 linear filters are assumed to resemble simple cell receptive fields (gabor-like functions of various spatial frequencies and orientations). Crucially, the pointwise nonlinearities used in stage two of the back pocket model are all assumed to be full-wave rectifiers. Specifically, the output of each linear filter is assumed to be squared prior to further processing.

Thus, the hard-wired stuff-sensing arrays used by the texture energy model discard all information about contrast polarity. Under such a model, it would be impossible for any statistic $G_\alpha(R)$ defined by Eq. (1) to be at all sensitive to variations in the local contrast polarity.

On the other hand, there are several reasons to suppose that texture judgments can indeed be attentionally tuned to differences in local contrast polarity. First, the visual system comprises separate, parallel pathways for information about positive contrasts (the "on-center" pathway) versus negative contrasts (the "off-center" pathway) in the scene (e.g., Kuffler, 1953; Hubel & Wiesel, 1961; Enroth-Cugell & Robson, 1966). Indeed sensitivity to contrast polarity is observed throughout the visual system (e.g., Ito, Fujita, Tamura, & Tanaka, 1994). Thus, information about local stimulus polarity might well be available for texture judgments in the separate on-center and off-center systems. Moreover, examples exist of preattentively discriminable texture pairs with equal mean luminance that differ only in being photographic negatives of each other (Chubb, Econopouly, & Landy, 1994; Malik & Perona, 1990). Any such textures should be preattentively indiscriminable under the energy model.

The current study investigates the nature of the stuff-sensing transformations Q_i, $i = 1, 2, \ldots, N$ occurring in Eq. (2). If all of the transformations Q_i use full-wave rectification (as would be the case under the energy model), then observers should find it difficult (or impossible) to tune their texture judgments to a single contrast polarity.

Our expectation, is that some of the transformations Q_i use either positive or negative half-wave rectification after their linear filters (as should be the case if some of these transformations are on- or off-center channels). Two predictions follow in this case. First, observers should be able to selectively tune their texture judgments to either positive or negative contrast polarities in the texture. Observers should also be able (by equalizing weights in Eq. (2) of positive and negative half-wave rectifying stuff-sensing arrays) to synthesize full-wave rectifying statistics $G_\alpha(R)$. However, suppose that all of the transformations Q_i occurring in Eq. (2) use either full-wave or positive or negative half-wave rectification. In this case, observers should have difficulty making texture judgments requiring pointwise nonlinearities that cannot be achieved as a linear combination of positive and negative half-wave rectifiers.

Indeed, as we shall show, for the restricted range of texture orientation judgments considered here,

1. observers display significant flexibility in tuning their texture judgments to specific contrast polarities. However,
2. observers seem unable to compute texture judgment statistics $G_\alpha(R)$ that use intensive nonlinearities other than linear combinations of positive and negative half-wave rectifiers.

These two results are consistent with the attentional model of Eq. (2), and imply moreover that the pointwise nonlinearities used by the stuff-sensing transformations Q_i are restricted to positive and negative half-wave rectifiers (or perhaps linear combinations of such rectifiers).

3. Experimental set-up

3.1. Observers

The two authors (JT and CC) were used in this experiment. JT had normal vision, CC had corrected-to-normal vision.

3.2. Apparatus

An IBM-compatible computer was used with an ATVista graphics system attached to a TVM monochrome monitor.

3.3. Stimuli

The stimulus field comprised a rectangular array of small, diagonally oriented bars on a mean-luminant gray background. Half the bars were oriented up/right, and half were oriented up/left. Thus the stimuli appeared to observers as a mesh of small, oppositely oriented diagonal lines of various intensities. More specifically, the stimulus field was partitioned into 8 rows by 16 columns of rectangular regions, each region taking up 11 rows by 5 columns of pixels, subtending $0.49° \times 0.25°$. The entire display subtended 3.93° in height by 3.76° in width. Into the center of each region was painted a bar oriented either up/right (at 71.5°) or up/left (at 108.5°). The structure of a bar of contrast 0.0 (1.0) oriented up/right (up/left) is schematized in Fig. 1a (Fig. 1b).

As this figure indicates, each bar comprised 9 pixels. Individual pixels subtended 0.045° (2.7 min) of visual angle. Each bar thus diagonally spanned a rectangular region 24 min in height by 8 min in width.

An example stimulus is illustrated in Fig. 1c. The background upon which each bar was painted remained mean gray. Bar orientation alternated columnwise across the stimulus field (the orientation of the bar in a given rectangle differed from that of the bars to its left and right, but was identical to that of the bars above and below it). Display mean luminance (equal to background luminance) was 71 cd/m^2. The luminance of each bar was drawn from a set $\Lambda = \{\lambda_i | i = 0, \ldots, 7\}$ of eight

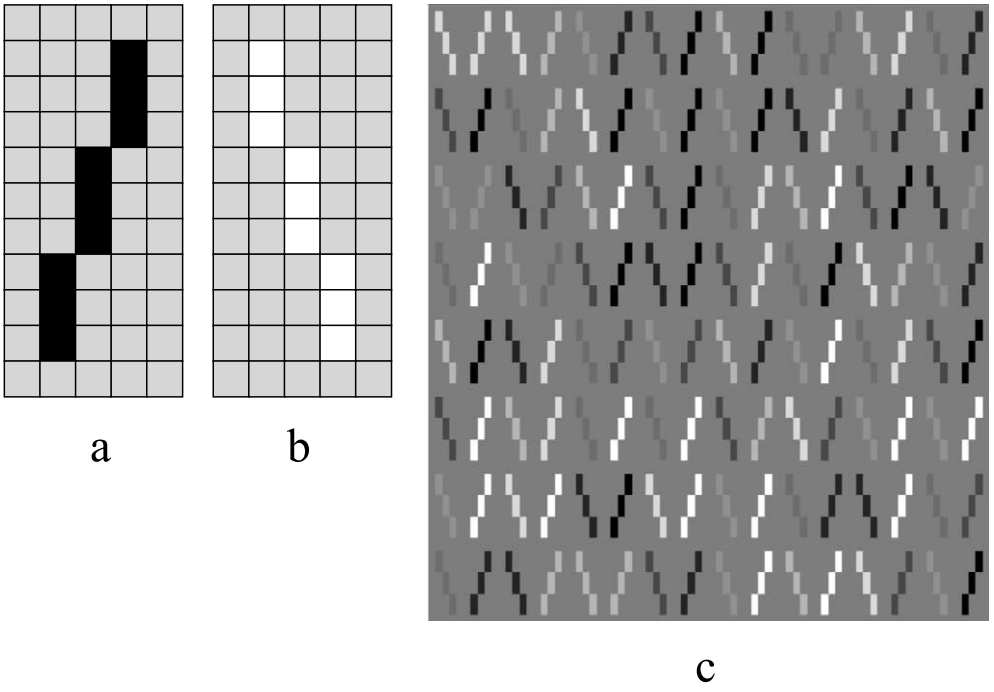

Fig. 1. Stimuli. Texture element regions comprise 11 rows by 5 columns of pixels, each pixel subtending ≈2.7 min of visual angle. Each texture element region is mean gray in background, with a bar painted into its center. Each bar diagonally spanned a rectangular region 24 min in height by 8 min in width: (a) a schematized example of a texture element of contrast 0.0 oriented up/right, (b) a schematized example of a texture element of contrast 1.0 oriented up/left, (c) an example stimulus display.

linearly increasing values. Specifically, $\lambda_i \approx 20i$ cd/m^2, for $i = 0, 1, \ldots, 7$. To be precise, λ_0 was measured at 1.5 cd/m^2, mean luminance was measured at 71 cd/m^2, and white was set so that an alternating black/white checkerboard pattern of granularity comparable to our stimuli disappeared into a mean luminant background, when viewed from far away.

Relative to the background luminance, the luminances $\lambda_i, i = 0, 1, \ldots, 7$, produce contrasts -0.98, -0.70, -0.42, -0.14, 0.14, 0.42, 0.70, 0.98. Although $\lambda_0, \lambda_1, \ldots, \lambda_7$ are luminances, we shall usually refer to them as "contrasts" because (as will soon emerge) the judgments we are investigating depend crucially on contrast polarity. Since mean luminance is fixed throughout these experiments, no ambiguity arises from this usage.

There have been many indications that negative contrasts tend to exert stronger influence over various perceptual judgments than do positive contrasts (e.g., Benton & Johnston, 1999; Sperling & Lu, 1999). However, no effort was made to equate corresponding positive and negative contrasts for perceptual salience. The purpose of the current study was not to compare the relative efficacy of opposite contrast polarities in determining judgments of texture orientation. Rather, our intent was to understand the ways in which texture orientation judgments can be attentionally modified to meet changing task demands.

4. Tasks

In the experiments reported here, various types of judgment are required of the observer. In each case, the judgment is defined by a real-valued *target function* f_{target} of Λ. In the task defined by a given such function f_{target}, the following occurs on each trial: the observer fixates a small cue spot in the middle of a mean luminant display field, then initiates a trial with a button-press. Immediately a stimulus P, comprising 64 bars of various intensities oriented up-right and 64 bars oriented up-left, is displayed for 17 ms, and is then replaced by the fixation field. (Such brief displays are used in order to prevent strategies involving attentional shifts, thereby inducing the observer to rely predominantly on spatially parallel mechanisms in making her/his judgments.) Following stimulus presentation, the observer attempts to judge whether

$$G_{\text{target}}(P) = \sum_{\text{All bars R oriented up-right}} f_{\text{target}}[\text{Contrast[R]}] - \sum_{\text{All bars L oriented up-left}} f_{\text{target}}[\text{Contrast[L]}] > 0, \tag{3}$$

where Contrast[B] is the value from Λ assigned to a given bar B. Finally, the observer registers his/her response with a button press, and receives audible correctness feedback.

In each of the experiments reported here, the target function f_{target} is fixed across all trials. For j ranging from 1 to the number of conditions in the given experiment, the jth condition is specified by a pair of histograms (p_j, q_j). Each of p_j and q_j is a nonnegative, integer valued function of Λ, with

$$\sum_{i=0}^{7} p_j[\lambda_i] = \sum_{i=0}^{7} q_j[\lambda_i] = 64. \tag{4}$$

Moreover, p_j and q_j are always chosen so that

$$\sum_{i=0}^{7} f_{\text{target}}[\lambda_i] p_j[\lambda_i] > \sum_{i=0}^{7} f_{\text{target}}[\lambda_i] q_j[\lambda_i]. \tag{5}$$

On a given experimental trial in this condition, one (randomly chosen) oriented population of bars, either up/right or up/left, is randomly assigned intensities conforming to histogram p_j, and the oppositely oriented bars are assigned intensities conforming to histogram q_j. (Thus, the numbers of different intensities assigned each orientation are precisely stipulated by p_j and q_j; only the locations to which these intensities are assigned are random.) Eq. (5) mandates that the correct response is whichever orientation conforms to histogram p_j.

4.1. Impact functions

The assumptions of our model are spelled out in detail in Appendix A. In this section we provide an informal description to enable the reader to appreciate the experimental manipulations and results.

Consider the task defined by a given target function f_{target}. In this task, given a particular stimulus P, comprising 64 bars of various intensities oriented up-right and 64 bars oriented up-left, the observer would like to compute $G_{\text{target}}(P)$ (Eq. (3)). We can think of the observer's attempt to compute $G_{\text{target}}(P)$ as involving three logical stages. In the first stage, the observer must (rapidly, in parallel over space) estimate $f_{\text{target}} \times$ [Contrast[B]] for each bar B in the display. Second, the observer must sum the resulting values for each of the two differently oriented sets of bars (the up-right and up-left oriented sets of bars). Finally, the observer must compare the sum total of the values estimated for the up-right oriented bars with the sum total of those estimated for the up-left oriented bars, and select whichever orientation has yielded the greater total.

There is room for error in any of these three stages. The trial-by-trial error likely to be introduced either in summing contributions of individual bars (the second stage in processing) within an orientation or in comparing totals across orientations (the final stage in processing) might naturally be modeled by additive normal random variables. Of greater interest is the error that enters in the first (the rapid, spatially parallel) stage of processing, because much of this error will

systematically reflect the processing constraints imposed by the human visual system. Let us then focus more carefully on this stage.

4.2. Rapid, spatially parallel approximation of f_{target}

During the stimulus presentation, the observer must coax each bar B in the image to produce a value (to be somehow coded in the visual system) approximating $f_{\text{target}}[\text{Contrast}[\text{B}]]$ as closely as possible. Since the stimulus is only presented for 17 ms, this computation must be accomplished rapidly in parallel across the stimulus field. On a given trial, something very complicated will undoubtedly occur in the brain of the observer during this initial stage of processing. Vast numbers of neurons will be activated. The nature of the task (defined by Eq. (3)) leads us to expect that those neurons relevant to the required judgment will be retinoptically arrayed, each neuron responsive to stimulus properties in a restricted region of the visual field. Indeed, we anticipate that multiple such neural arrays may well be recruited by the observer for the task. We imagine that the observer can selectively combine (in accordance with Eq. (2)) the activations of these retinotopic neural arrays to fashion the neural device that is used to approximate f_{target}.

A given bar is likely to contribute to the activation of many neurons in any one of these hypothetical neural arrays, and conversely, the activation of a given neuron in any one of these arrays may well be influenced by more than one bar. The sum total of activations produced within this neural device by a bar B on a given trial is called the *impact* exerted by B on the judgment performed on that trial. We expect that the impact exerted by a bar of a given contrast will vary from trial to trial due to both internal noise and also to such systematic factors as the context of the bar (i.e., the contrasts of surrounding bars) and its location in the stimulus.

However, the nature of the task naturally motivates the observer to minimize variability in the impact exerted by a bar B of a given contrast. Ideally, the impact exerted by B should depend only on B's contrast (and, of course, on B's orientation). For any given bar B, the observer wants the impact of B to approximate $f_{\text{target}}[\text{Contrast}[\text{B}]]$ as closely as possible. However, limitations imposed by the hardware resident within the visual system may make it difficult or impossible to achieve a good approximation of f_{target}. In this case, we anticipate that, on average, the impact exerted by a given bar B on the required judgment may well deviate significantly from $f_{\text{target}}[\text{Contrast}[\text{B}]]$. Accordingly, we let

$$m_{\text{target}}[C] = \text{the average impact exerted on judgments of orientation defined by } f_{\text{target}} \text{ by a bar of contrast } C; \quad (6)$$

m_{target} is called the *impact function* achieved by the observer for judgments of orientation defined by f_{target}.

4.3. Plots of impact functions

Many of our graphs plot impact functions estimated for our two observers in various tasks. In the absence of further explanation, the units in which we express these impact functions are likely to be somewhat mysterious. We assume that in attempting to judge whether $G_{\text{target}}(P) > 0$, the observer computes a statistic $\widetilde{G}_{\text{target}}(P)$, which approximates $G_{\text{target}}(P)$ to the extent that this is within the observer's power. We imagine that $\widetilde{G}_{\text{target}}(P)$ results from (i) summing the impacts exerted individually by the bars in P oriented up-left to obtain total_up_left_impact (P), (ii) summing the impacts exerted individually by the bars in P oriented up-right to obtain total_up_right_impact (P), and (iii) taking the difference

$$\widetilde{G}_{\text{target}}(P) = \text{total\_up\_right\_impact} - \text{total\_up\_left\_impact}. \quad (7)$$

If $\widetilde{G}_{\text{target}}(P)$ is positive, the observer responds "up-right"; otherwise the observer responds "up-left". As discussed above, we suppose that the impact exerted by a given bar of contrast C may vary randomly from trial to trial. For any contrast C, $m_{\text{target}}[C]$ gives the average impact (across trials) exerted by a bar of contrast C on the statistic $\widetilde{G}_{\text{target}}(P)$. Under the assumptions detailed in Appendix A, $\widetilde{G}_{\text{target}}(P)$ is approximately normal in distribution. We write $\text{SD}_{\text{target}}$ for the standard deviation of $\widetilde{G}_{\text{target}}(P)$, and as a matter of convention we express m_{target} in multiples of $\text{SD}_{\text{target}}$. We think of $\text{SD}_{\text{target}}$ as the standard deviation of the total noise by which the observer's performance is compromised.

Let us take a hypothetical example. Suppose $m_{\text{target}}[0] = 0.11$. This means that an occurrence of a bar of contrast 0 in, say, the contingent of bars oriented up-right slightly increases the probability that the observer will judge the f_{target}-defined orientation to be up-right. Specifically, on average, an occurrence of a bar of contrast 0 in the contingent of bars oriented up-right induces an increase in the statistic $\widetilde{G}_{\text{target}}(P)$ equal to 0.11 of $\widetilde{G}_{\text{target}}(P)$'s standard deviations.

4.4. Measuring an impact function

For any real-valued functions f and g of Λ, we write $f \cdot g$ for the *dot product* of f with g:

$$f \cdot g = \sum_{i=0}^{7} f[\lambda_i] g[\lambda_i]. \quad (8)$$

Consider a particular experiment in which the task is defined by a fixed target function f_{target}. Histograms of bar intensities for the competing orientations are experimentally varied according to the methods described in Chubb (1999). Central to these methods is the concept of a *histogram modulator*. For current purposes, it suf-

fices to call a given, integer-valued function ϕ of Λ a histogram modulator if

$$\sum_{i=0}^{7} \phi[\lambda_i] = 0, \quad \text{and} \quad |\phi[\lambda_i]| \leqslant 8, \qquad i = 0, 1, \ldots, 7. \quad (9)$$

There are 64 bars oriented up-right and 64 oriented up-left; thus, the right-hand condition in Eq. (9) insures that one obtains a histogram of bar contrasts for one orientation either by adding or by subtracting ϕ from the uniform histogram, which allots to a given orientation eight bars of each contrast in Λ.

Each of the experiments we describe is designed to investigate performance in the task defined by a given target function, f_{target}. Each such experiment subsumes a number of different experimental conditions. Suppose, e.g., that the experiment investigating f_{target} subsumes N_{target} conditions (this number will vary across the experiments we report). Each of these conditions j is defined by a given histogram modulator ϕ_j, $j = 1, 2, \ldots, N_{\text{target}}$. Moreover, each of these modulators correlates positively with f_{target}. That is,

$$f_{\text{target}} \cdot \phi_j > 0, \quad j = 1, 2, \ldots, N_{\text{target}}. \quad (10)$$

On a given trial in condition j, the modulator ϕ_j is used to generate the stimulus as follows:

1. The correct response (up/right or up/left) is randomly selected.
2. The histogram determining the bar intensities in the correct (incorrect) orientation is now set to $p_j = 8 + \phi_j$ ($q_j = 8 - \phi_j$). Thus, for $i = 0, 1, \ldots, 7$, the number of bars oriented in the correct (incorrect) direction to be painted intensity λ_i is set to $p_j[\lambda_i] = 8 + \phi_j[\lambda_i]$ ($q_j[\lambda_i] = 8 - \phi_j[\lambda_i]$). [1] (Note that Eq. (4) guarantees first that the total number of bars assigned intensities in each orientation must be 64, and second, that for $i = 0, 1, \ldots, 7$, a nonnegative number of bars of intensity λ_i is assigned to each orientation.)
3. The stimulus P is now generated by assigning in random order the 64 bar intensities allotted to each orientation.

[1] The reader might wonder why, in generating stimuli, we choose always to modulate oppositely the histograms of bar intensities assigned to opposite orientations. The answer (discussed in more detail in Chubb (1999)) is that the validity of the model rests on weaker assumptions in this case. To gain an intuition about why this might be the case, note that when a stimulus P is produced by oppositely modulating the histograms of opposite orientations, the total histogram across all bars of P (irrespective of bar orientation) remains constant across trials (for $i = 0, 1, \ldots, 7$, P comprises exactly 16 bars of contrast λ_i). By keeping this histogram constant, we insure (under weak assumptions) that the variance of the decision statistic, $\widetilde{G}_{\text{target}}(P)$ is constant across different experimental conditions. Formally, the constraint that each stimulus P be generated by oppositely modulating the histograms of opposite orientations insures that the right side of Eq. (A.8) of the appendix does not depend on the particular histograms used in P.

Note the following: on a trial in condition j, the observer judges correctly if either the correct orientation is up/left and $\widetilde{G}_{\text{target}}(P) > 0$, or if the correct orientation is up/right and $\widetilde{G}_{\text{target}}(P) < 0$. These two events are completely symmetric. For concreteness, suppose the correct orientation is up/left. In this case, as we show in Appendix A, under the assumptions of the model,

$$\text{Prob}[\widetilde{G}_{\text{target}}(P) > 0] = \Phi(m_{\text{target}} \cdot \phi_j), \quad (11)$$

where Φ is the standard normal cdf, and m_{target} is expressed in multiples of $\text{SD}_{\text{target}}$.

The maximum likelihood estimator of m_{target}. Let N_{target} be the number of conditions used in the experiment whose task is defined by f_{target}, and for $j = 1, 2, \ldots, N_{\text{target}}$, let k_j and n_j be the number of correct and incorrect responses given by the observer in condition j. Then for any guess m_{guess} at the $\widetilde{G}_{\text{target}}$ impact function, let

$$\text{Likelihood}(m_{\text{guess}}) = \prod_{j=1}^{N_{\text{target}}} \Phi^{k_j}(m_{\text{guess}} \cdot \phi_j) \times (1 - \Phi(m_{\text{guess}} \cdot \phi_j))^{n_j}. \quad (12)$$

Thus $\text{Likelihood}(m_{\text{guess}})$ gives the joint probability of the obtained data under the assumption that $m_{\text{target}} = m_{\text{guess}}$. The maximum likelihood estimate of m_{target} is the value $\hat{m}_{\text{target}} = m_{\text{guess}}$ maximizing Eq. (12). All the impact function estimates we report here are maximum likelihood estimates.

4.5. Training

Each of the experiments described below yields, for each observer, an estimate $\hat{m}_{\text{target}}$ of the impact function used to make texture orientation judgments based on a given target function f_{target}. In order to draw strong conclusions about the processing capabilities of our observers, it is crucial that each estimated impact function reflect nearly optimal performance in the given task. Several considerations make us confident that this aim was achieved.

Prior to data collection in the experiment to estimate a given impact function $\hat{m}_{\text{target}}$, the corresponding target function f_{target} was presented explicitly to the observer both verbally and graphically. (As the observers were the two authors, there is no doubt that each observer in each experiment was fully aware of the form of the target function defining the task.) Next, prior to collecting data, the observer performed (with trial-by-trial feedback) a practice block (comprising 130 trials), identical in content to each of the 10 data-collection blocks used to estimate $\hat{m}_{\text{target}}$.

In addition, prior to the current study, each observer participated in a preliminary study providing extensive practice in each of the experimental tasks employed in

the current study. Like the current study, the preliminary study comprised five experiments, each experiment requiring an orientation judgment based on a different target function f_{target}. Although stimuli for the current study were modified slightly from those of the preliminary study (the number of bars in the stimulus was substantially reduced in the current study, and the number of intensities in Λ was reduced from 9 to 8), corresponding target functions were very similar in form across the preliminary and current experiments. [2] Thus the five texture judgments required in the preliminary study were essentially identical to the corresponding judgments required in the current study. As is the case in the current study, at the start of a given experiment in the preliminary study, f_{target} was presented explicitly to the observer both verbally and graphically. The observer then performed 11 blocks (the first being a practice block), each consisting of 150 trials (10 trials each of 15 conditions, randomly sequenced). Thus, in the preliminary study, each observer performed 1650 trials in tasks nearly identical to the corresponding tasks used in the current study.

5. Interpreting an impact function

Note that it is not possible to measure the mean value of an impact function. (This is because our paradigm allows us to measure only the projection of the impact function into the space of functions spanned by the modulators we use; however, any modulator must have a mean value equal to 0. It follows that our estimated impact function must also have a mean of 0.) In other words, any impact function is only defined up to an arbitrary additive constant (its unmeasurable mean value). Accordingly, as a matter of convention, all the impact functions we show are plotted with means set to 0.

Let f and g be real-valued functions of Λ, each of which has mean 0. Write $|f|$ for the *norm* of f: $|f| = \sqrt{f \cdot f}$. Then the *correlation coefficient* between f and g is given by

$$r(f,g) = \frac{f \cdot g}{|f||g|}. \tag{13}$$

We shall generally use correlation coefficients to gauge the degree to which two functions (each with mean 0) match in form, irrespective of their norms.

[2] The results of the preliminary study were essentially the same as those of the current study. We were, however, concerned that the stimulus used in the first experiment subtended too large a region in the visual field. Moreover, the texture patch in the preliminary experiment was vertically elongated in a way which both observers felt made judgments unnaturally difficult. For this reason we decided to rerun the experiment using the current stimulus configuration.

Let m_{target} be the impact function achieved by an observer in striving to judge orientation based on a given target function f_{target}. One might suppose that the observer can optimize performance by synthesizing m_{target} so as to maximize $m_{\text{target}} \cdot f_{\text{target}}$. However, this may not be true in a given experiment for the following reason. If $r(m_{\text{target}}, f_{\text{target}}) < 1$, then there exists a nonzero function ρ orthogonal to f_{target} that correlates positively with m_{target}. Specifically, as can be easily checked, this is true for

$$\rho = m_{\text{target}} - \left(\frac{f_{\text{target}} \cdot m_{\text{target}}}{f_{\text{target}} \cdot f_{\text{target}}}\right) f_{\text{target}}. \tag{14}$$

Suppose that half of the modulators used in the experiment to measure m_{target} are made to correlate strongly with ρ, and the other half with $-\rho$. If $|m_{\text{target}} \cdot \rho|$ is large in comparison to $|m_{\text{target}} \cdot f_{\text{target}}|$, then judgments are likely to be dominated by the spurious correlations $\pm m_{\text{target}} \cdot \rho$ occurring from trial to trial, leading to depressed performance. In such circumstances, we say that the observer "has incentive to cleanse his/her impact function of its dependency on ρ". In this case, optimal performance is likely to be achieved using an impact function m^*_{target} that strikes a compromise between maximizing $m^*_{\text{target}} \cdot f_{\text{target}}$ and minimizing $|m^*_{\text{target}} \cdot \rho|$.

As these considerations suggest, the impact function that optimizes performance in a given task may depend both on f_{target} as well as the specific set of modulators used. Nonetheless, for purposes of gauging the overall sensitivity of the observer in judging orientation based on f_{target}, it will be convenient to ignore the possible influence of the modulators used, and to compute a statistic that simply reflects the overall correlation of the impact function m_{target} with f_{target}. Specifically, we compute the dot product of m_{target} with the normalized target function:

$$\begin{aligned}\text{Eff}(m_{\text{target}}, f_{\text{target}}) &= \frac{m_{\text{target}} \cdot f_{\text{target}}}{|f_{\text{target}}|} \\ &= |m_{\text{target}}| r(m_{\text{target}}, f_{\text{target}}).\end{aligned} \tag{15}$$

$\text{Eff}(m_{\text{target}}, f_{\text{target}})$ is called the observer's *efficiency* at the task defined by f_{target}. As the rightmost component of Eq. (15) makes clear, the observer's efficiency will be low if his/her impact function either has a small norm or has a low correlation coefficient with f_{target}.

6. Experiment 1

We begin by measuring the impact function that results if observers are asked to judge the predominant orientation of texture energy in the stimulus patch. The target function f_{energy} used to define the task (and to provide trial-by-trial feedback) is the quadratic function,

$$f_{\text{energy}}(\lambda_i) = (i - 3.5)^2, \quad i = 0, 1, \ldots, 7. \tag{16}$$

For observer JT, the modulators ϕ_j, $j = 1, 2, \ldots, 13$, used in this experiment can be found by going to http://texel.ss.uci.edu/Links.html, clicking first on the link to the current paper, and then on the link to Table 1. The first 8 columns in the jth row of Table 1 define ϕ_j, with the integer in the ith column of row j giving the value of $\phi_j[\lambda_i]$. The modulator ϕ_1 was chosen to correlate as well as possible with f_{energy}, under the constraints (a) that it assume only integer values, (b) its amplitude be sufficiently small as to support less than perfect performance. The other 12 modulators were constructed so that modulators have approximately equal dot products with f_{energy}. They were generated by first obtaining a set of 6, random, orthonormal functions, ρ_k, $k = 1, 2, \ldots, 6$, each of which was also orthogonal to ϕ_1. Then for $j = 2, 3, \ldots, 7$, ϕ_j was obtained by summing ϕ_1 and ρ_{j-1}, rounding the resulting values, and making minimal adjustments to make sure the resulting function satisfied Eq. (9). For $j = 8, 9, \ldots, 13$, ϕ_j was obtained by taking the difference $\phi_1 - \rho_{j-1}$, rounding the resulting values, and making minimal adjustments to make sure the resulting function satisfied Eq. (9).

Observer JT performed 10 blocks, each comprising 130 trials, 10 from each condition, randomly mixed. This yielded 100 trials per condition.

A glance at JT's results (column 9 of row j of Table 1 gives number k_j correct in condition j; column 10 gives number incorrect (n_j)) shows that she was performing with a high rate of success (her mean percent correct was 88%). We were concerned that the conditions used for her might be too easy, and that the resulting impact function might be artificially depressed in overall amplitude compared to the function she would have achieved with more challenging conditions. Accordingly, we used different modulators for observer CC, each of which correlated less strongly with the target function f_{energy}. The process used to generate the thirteen modulators used for CC was precisely analogous to that used for JT. CC also performed 10 blocks, each comprising 130 trials, 10 each from each condition, randomly mixed.

6.1. Results

The raw data are shown in Table 1 for JT and in Table 2 for CC (also accessible via http://texel.ss.uci.edu/Links.html). The estimated impact functions for the two observers are shown in Fig. 2. The correlation coefficient between the impact function of CC (JT) and f_{energy} is 0.99 (0.93), and the efficiency of CC (JT) in the current task is 0.20 (0.17).

6.2. Discussion

As reflected by the high correlation coefficients for both observers, the impact functions of CC and JT match quite well the form of the target function f_{energy} (Fig. 2, bottom). CC's performance was closer to threshold (70%) than JT's; nonetheless, his impact function was similar both in form and in overall modulation depth (norm) to that of JT. This suggests the conditions used for JT were sufficiently challenging to reveal her actual impact function.

The similarity between the quadratic target function and the impact functions used by observers in the current task suggests that the architecture of the visual system is well suited to the task. The current results are thus consistent with a system in which $\widetilde{G}_{\text{energy}}(P)$ is the difference between opposite orientation strengths, each of which was individually extracted using orientation energy analyzers (Bergen & Adelson, 1988). Thus, the results of this experiment lend prima facie support to energy-based models of texture processing.

On the other hand, it is also possible that observers are basing their judgments not simply on the responses of orientation energy analyzers, but on the combined responses of separate, orientation-tuned channels selective for positive and negative contrast polarities (Malik & Perona, 1990; Morgan & Watt, 1997; Watt & Morgan, 1985).

If indeed observers have separate access to distinct positive and negative half-wave rectified channels for making judgments of texture orientation, then we might expect observers to be able to exercise some control over the impact functions they can achieve in making such judgments. Specifically, in an experiment in which only luminances greater (lower) than the mean are informative about the correct orientation, we might expect observers to use impact functions sensitive exclusively to positive (negative) contrast polarities. This possibility is investigated in Experiment 2.

7. Experiment 2

We next asked whether it is possible for the observer to exert any attentional control over the weights with which positive contrasts (produced by luminances greater than the mean) and negative contrasts (produced by luminances less than the mean) are combined in making orientation judgments. To investigate this issue, we used target functions that rendered only a single contrast polarity informative.

The negative-polarity selective target function we used, f_{neg}, assigned the values 3, 1, −1, −3, 0, 0, 0, 0, to corresponding intensities, $\lambda_0, \lambda_1, \ldots, \lambda_7$.

JT was tested at this task using the experimental conditions shown in Table 3 (accessible via http://texel.ss.uci.edu/Links.html). As in Tables 1 and 2, the first 8 columns in a given row define the modulator used in that condition, and the last two columns give the number correct and number incorrect. The top 13 modulators

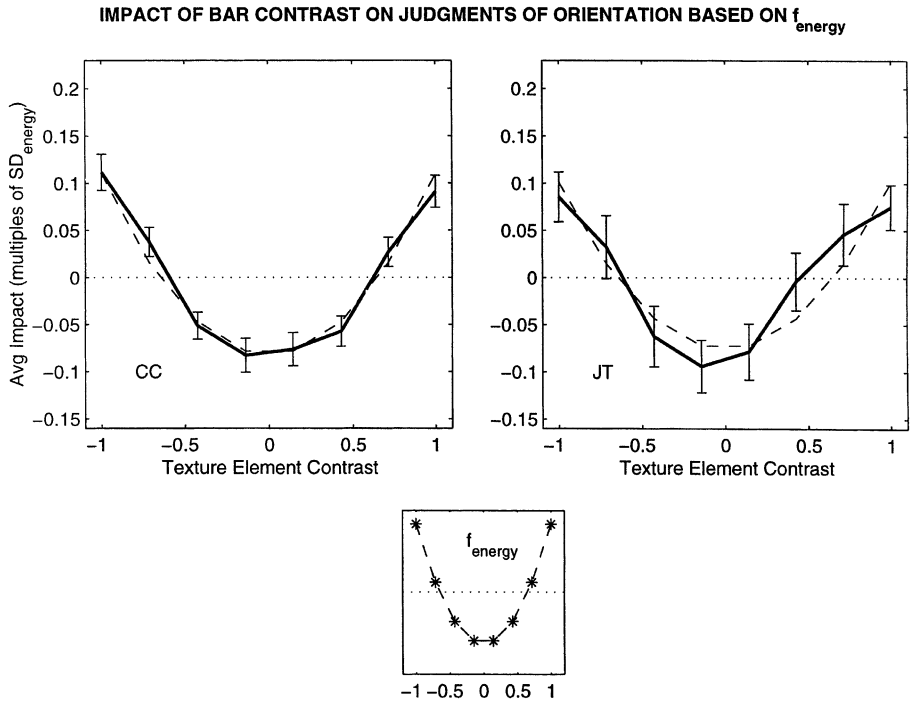

Fig. 2. The estimated impact functions for the two observers for judgments of orientation based on f_{energy}. Error bars indicate (bootstrapped) 95% confidence intervals of estimated impact function values. (Left) The impact function for CC (JT) is shown on the left (right). The dashed line in each figure shows f_{energy} with norm set equal to that of the corresponding impact function. The target function f_{energy} is shown by itself in the inset (center/bottom). The close matches of impact functions to f_{energy} suggest that human vision is well-suited to the task of extracting orientation based on overall orientation energy.

ϕ_j, $j = 1, 2, \ldots, 13$, used for JT were constructed as follows. The modulator ϕ_1 was chosen to correlate strongly with f_{neg}. The other 12 modulators were constructed so that modulators have approximately equal dot products with f_{neg}. Specifically, they were generated by obtaining a set of 6, random, orthonormal functions, ρ_k, $k = 1, 2, \ldots, 6$, each of which was also orthogonal to ϕ_1. Then for $j = 2, 3, \ldots, 7$, ϕ_j was obtained by summing ϕ_1 and ρ_{j-1}, rounding the resulting values, and making minimal adjustments to make sure the resulting function satisfied Eq. (9). For $j = 8, 9, \ldots, 13$, ϕ_j was obtained by taking the difference $\phi_1 - \rho_{j-1}$, rounding the resulting values, and making minimal adjustments to make sure the resulting function satisfied Eq. (9). This method of construction, in which the same ρ_js are positively and negatively added to ϕ_1 to generate different experimental conditions, insures that overall success rate is uniquely governed by the correlation of the observer's impact function with ϕ_1. (Any function orthogonal to ϕ_1 will have mean correlation 0 to all modulators used.) Modulators ϕ_j, $j = 14, 15, \ldots, 26$ for JT were constructed in a manner exactly the same as conditions using ϕ_{14} in the role ϕ_1 played in generating the first 13 modulators. ϕ_{14} has a slightly lower correlation with f_{neg} than does ϕ_1; this makes the last 13 conditions slightly harder than the first 13.

Observer JT performed 20 blocks, each comprising 130 trials. Each of the first 10 blocks contained 10 trials (randomly mixed) of each of conditions 1–13; each of the last 10 blocks contained 10 trials (randomly mixed) of each of conditions 14–26. This yielded 100 trials per condition.

The conditions used in testing CC in the f_{neg} task were constructed somewhat differently (Table 4, accessible via http://texel.ss.uci.edu/Links.html). The method of construction was prompted by the following reflections. Bars of positive contrast are irrelevant to the current task. Any sensitivity of the observer's impact function to bars of positive contrast is thus likely to depress performance (as such sensitivity will detrimentally bias performance in at least some of conditions 2–13 and 15–26). However, the degree of this depression of performance depends on the strength with which the observer's impact function correlates with the experimental modulators across intensities λ_4, λ_5, λ_6, λ_7. If strength of correlation is typically low, then the observer has little motivation to eliminate sensitivity to bars of positive contrast from his/her impact function. Both observers had previously performed judgments using quadratic impact functions that showed high sensitivity to both black and white bars. We attempted to discourage CC from perseverating in the use of this pre-

viously learned impact function by designing experimental conditions that would provide CC with strong incentive to cleanse his impact function of differential sensitivity to intensities λ_4, λ_5, λ_6, λ_7. Toward this end, we made sure that each condition used for CC correlated strongly with the portion of f_{energy} defined on intensities λ_4, λ_5, λ_6, λ_7. However, we arranged things so that on half the trials this correlation would be positive, and on half the trials, it would be negative.

The specific construction method was as follows. ϕ_1 (which assigns values 4, 0, −1, −3, 0, 0, 0, 0 to intensities $\lambda_0, \lambda_1, \ldots, \lambda_7$) was chosen to correlate strongly with f_{neg}. Let θ be the function assigning values 0, 0, 0, 0, −3, −1, 0, 4 to intensities $\lambda_0, \lambda_1, \ldots, \lambda_7$. (Thus, θ correlates strongly with the portion of f_{energy} defined on intensities λ_4, λ_5, λ_6, λ_7.) We then set $\phi_2 = \phi_1 + \theta$, and $\phi_3 = \phi_1 - \theta$. Next we obtain a set of 5, random, orthonormal functions, ρ_k, $k = 1, 2, \ldots, 5$, each of which was also orthogonal to both of ϕ_1 and θ. Then for $j = 4, 5, \ldots, 8$, ϕ_j was obtained by adding ρ_{j-3} to ϕ_2, rounding the resulting values, and making minimal adjustments to make sure the resulting function satisfied Eq. (9). For $j = 9, 10, \ldots, 13$, ϕ_j was obtained by taking the difference $\phi_2 - \rho_{j-8}$, rounding the resulting values, and making minimal adjustments to make sure the resulting function satisfied Eq. (9). An analogous procedure was used to generate ϕ_j, $j = 14, 15, \ldots, 23$, with ϕ_3 playing the role formerly played by ϕ_2. This procedure yielded 23 modulators, all of which correlate strongly with ϕ_1 (and hence also with f_{neg}), 11 of which correlate strongly and positively with θ, 11 of which correlate strongly and negatively with θ, and one of which (ϕ_1) correlates 0 with θ.

Observer CC performed five blocks, each comprising 230 trials, 10 trials in each condition, randomly ordered. This yielded 50 trials per condition.

7.1. Results

Raw data for JT and CC are shown in Tables 3 and 4 (accessible via http://texel.ss.uci.edu/Links.html). The resulting impact functions are shown in Fig. 3. The correlation coefficient between f_{neg} and the impact function of CC (JT) is 0.83 (0.83). The efficiency of CC (JT) at the current task was 0.20 (0.18).

7.2. Discussion

Both observers show relatively heightened sensitivity to negative (vs. positive) contrasts in the current task. We conclude that observers are not strictly constrained to using full-wave rectification in making

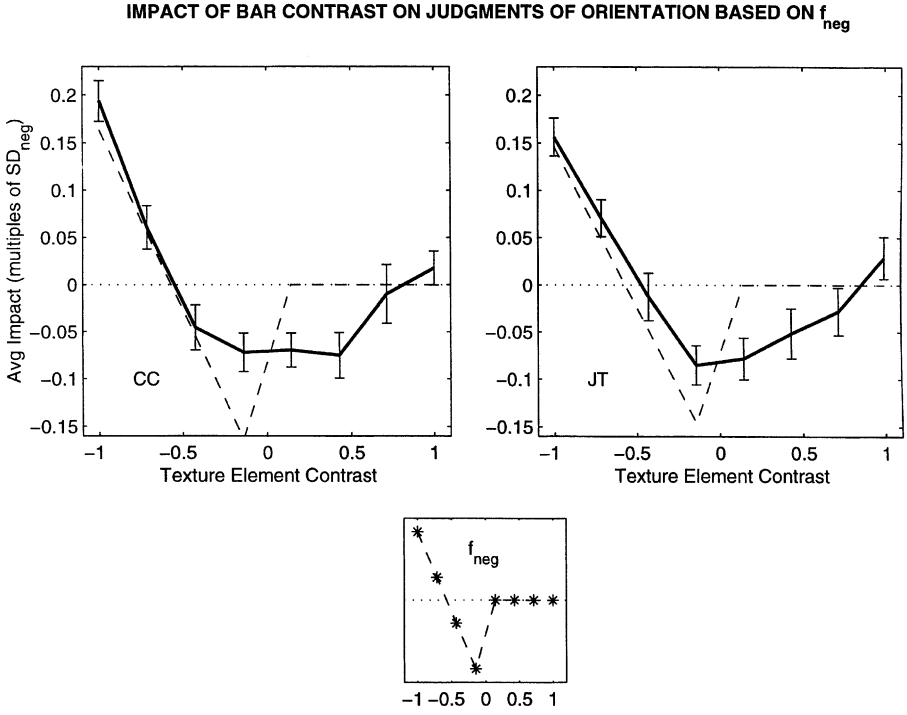

Fig. 3. The estimated impact functions for the two observers for judgments of orientation based on f_{neg}. Error bars indicate (bootstrapped) 95% confidence intervals of estimated impact function values. The impact function for CC (JT) is shown on the left (right). The dashed line in each figure shows f_{neg} with norm set equal to that of the corresponding impact function. The target function f_{neg} is shown by itself in the inset (center/bottom). Impact functions for both observers show increased (decreased) sensitivity for negative (positive) contrasts compared to the impact functions obtained in Experiment 1. This suggests that judgments of predominant texture orientation are subject to top–down control.

judgments of texture orientation. Evidently they can synthesize statistics for making orientation judgments that are selectively sensitized to negative contrasts.

It should be noted, however, that neither observer is able to completely cleanse his/her impact function of sensitivity to positive contrasts; neither observer's impact function approximates f_{neg} within measurement error. This is hardly surprising, however. For an impact function to perfectly mimic target function f_{neg}, texture elements of contrast -0.14 would need to exert significantly different average impact on judgments from texture elements of contrast 0.14. As one might expect, however, for both observers, the impact exerted by low contrast texture elements of both polarities is approximately equal. Under the assumption that impact functions must vary smoothly, both observers have achieved reasonable approximations of f_{neg}, by allowing impact to increase gradually from low to high positive contrasts.

Comparison of the impact functions obtained in Experiments 1 and 2 (Figs. 2 and 3) indicates that observers are able to successfully adapt their judgment statistics to the demands of the new task. Experiment 3 investigates whether observers can similarly tune their judgments to positive contrast polarities.

8. Experiment 3

The procedure was exactly the same as in Experiment 2 with the exception that modulators used for JT (CC) were mirror reflections of the modulators used in Experiment 2. The target function in Experiment 3 was also mirror symmetric to the target function f_{neg} used in Experiment 2. Specifically, f_{pos} assigned the values 0, 0, 0, 0, -3, -1, 1, 3 to corresponding intensities, $\lambda_0, \lambda_1, \ldots, \lambda_7$.

8.1. Results

The raw data are given in Tables 5 (JT) and 6 (CC) (accessible via http://texel.ss.uci.edu/Links.html). The corresponding impact functions are shown in Fig. 4. The correlation coefficients between f_{pos} and the impact function of CC (JT) was 0.75 (0.79). The efficiency of CC (JT) in the current task was 0.18 (0.19).

8.2. Discussion

The orientation judgments of both observers now show heightened sensitivity for texture elements of positive contrast. However, for both observers, but especially for JT, the difference in impact exerted by

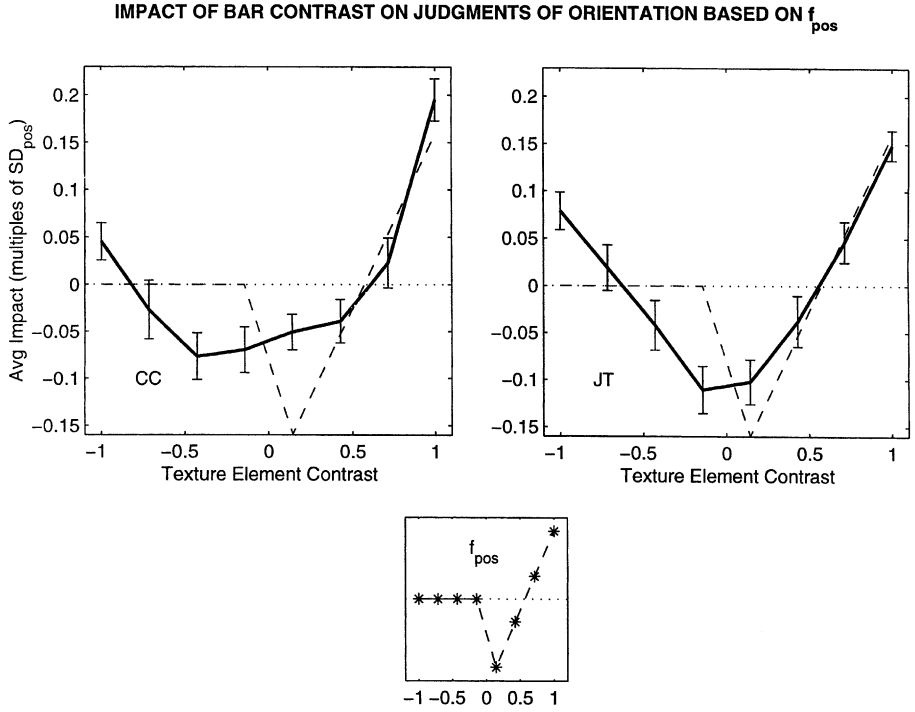

Fig. 4. The estimated impact functions for the two observers for judgments of orientation based on f_{pos}. Error bars indicate (bootstrapped) 95% confidence intervals of estimated impact function values. The impact function for CC (JT) is shown on the left (right). The dashed line in each figure shows f_{pos} with norm set equal to that of the corresponding impact function. The target function f_{pos} is shown by itself in the inset (center/bottom). Impact functions for both observers show increased (decreased) sensitivity for positive (negative) contrasts compared to the impact functions obtained in Experiment 1, again suggesting that texture orientation judgments are subject to top–down control.

positive versus negative contrast texture elements is less pronounced than the corresponding difference obtained in Experiment 2. Even though black texture elements are irrelevant to the required judgment, they nonetheless continue to exert significant impact on JT's judgments. It should be noted, however, that JT had less incentive than did CC to cleanse her impact function of sensitivity to negative contrasts. CC's modulators were specifically constructed so that performance would be adversely affected by using an impact function with a strong linear trend across the positive contrasts.

The important point is that the impact functions of Fig. 4 differ significantly from those obtained in either of the first two experiments. Evidently, observers are able to selectively tune their judgments of texture orientation to either positive or negative polarities.

A possible account of these findings proposes that at least some of the stuff-sensing arrays Q_i combined in Eq. (2) use positive half-wave rectification, and others use negative half-wave rectification. This would be the case, e.g., if some of the Q_i were on-center channels, and others were off-center channels.

In Experiments 4 and 5, we investigate whether any of the stuff-sensing transformations Q_i use nonlinearities other than either full-wave, or positive or negative half-wave rectification.

9. Experiment 4

The goal of Experiment 4 was to construct as simple as possible a target function that would be difficult to synthesize as a linear combination of positive and negative half-wave rectifiers. The function f_{order3} tested in this experiment assigns the values -1.0, 0.714, 1.0, 0.429, -0.429, -1.0, -0.714, 1.0 to intensities λ_0, $\lambda_1, \ldots, \lambda_7$ (shown in Fig. 5). f_{order3} is the (discrete analog of the) third order Legendre polynomial. It is the lowest order Legendre polynomial that is orthogonal to all polynomials of order two or less. As all of the impact functions achieved by our observers in Experiments 1–3 can be well-approximated by second order polynomials, f_{order3} posed a novel challenge.

JT was tested at this task using the experimental conditions shown in Table 7 (accessible via http://texel.ss.uci.edu/Links.html). The modulators ϕ_j, $j = 1, 2, \ldots, 13$ used for JT were constructed as follows. ϕ_1 was chosen to correlate positively with f_{order3}. The other 12 modulators were generated by first obtaining a set of 6, random, orthonormal functions, ρ_k, $k = 1, 2, \ldots, 6$, each of which was also orthogonal to ϕ_1. Then for $j = 2, 3, \ldots, 7$, ϕ_j was obtained by summing ϕ_1 and ρ_{j-1}, rounding the resulting values, and making minimal adjustments to make sure the resulting function satisfied

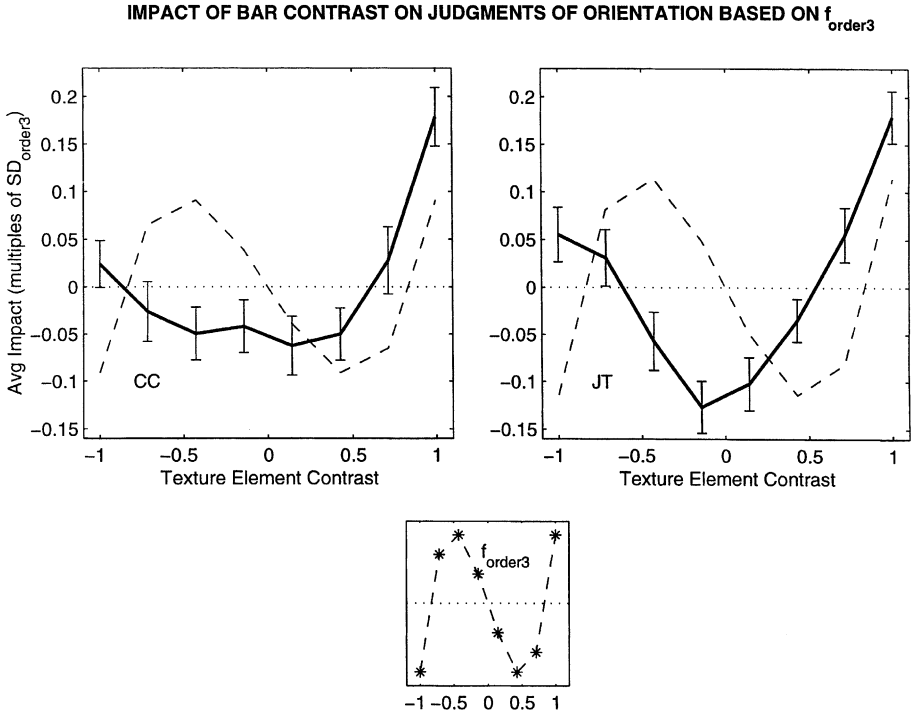

Fig. 5. The estimated impact functions for the two observers for judgments of orientation based on f_{order3}. Error bars indicate (bootstrapped) 95% confidence intervals of estimated impact function values. The impact function for CC (JT) is shown on the left (right). The dashed line in each figure shows f_{order3} with norm set equal to that of the corresponding impact function. The target function f_{order3} is shown by itself in the inset (center/bottom). Impact functions for both observers closely resemble those obtained in Experiment 3 (Fig. 4), suggesting that cannot effectively adapt to the task posed by f_{order3}.

Eq. (9). For $j = 8, 9, \ldots, 13$, ϕ_j was obtained by taking the difference $\phi_1 - \rho_{j-1}$, rounding the resulting values, and making minimal adjustments to make sure the resulting function satisfied Eq. (9).

Observer JT performed 10 blocks, each comprising 130 trials, 10 from each condition, randomly mixed. This yielded 100 trials per condition.

The conditions used for observer CC are shown in Table 8 (accessible via http://texel.ss.uci.edu/Links.html). These 45 conditions were designed to discourage the use of strategies leading to impact functions resembling any of the impact functions achieved by CC in any of Experiments 1, 2 or 3. Each modulator correlated strongly and positively with the target function, f_{order3}. In addition, however, modulators were designed to correlate strongly but randomly with previously obtained impact functions. Thus reliance on any previously used impact function would yield near-chance performance.

Observer CC performed six blocks, each comprising 225 trials, five from each condition, yielding a total of 30 trials per condition.

9.1. Results

The raw data are given in Tables 7 (JT) and 8 (CC) (accessible via http://texel.ss.uci.edu/Links.html). The corresponding impact functions are shown in Fig. 5. The correlation coefficient between the impact function of JT (CC) and f_{order3} is 0.11 (0.25). The efficiency of JT (CC) in the current task was 0.03 (0.05).

9.2. Discussion

The low efficiencies achieved by both observers reflect the difficulty they had in synthesizing a judgment strategy tailored to the task. Indeed, the impact functions produced by both observers resemble very closely those that they produced in Experiment 3. On the one hand, this is perhaps not so surprising, given that the correlation coefficient between f_{order3} and f_{pos} is 0.45. On the other hand, it is revealing that no options seem to be available to the observers to improve performance.

We infer that f_{order3} is largely orthogonal to the space of impact functions achievable by our observers.

It should be noted that both observers were tested with target function f_{order3} after having been tested with target functions f_{energy}, f_{neg}, and f_{pos}. It is possible that our observers' experience in these prior tasks led to learning that blocked effective learning in the current task. This seems unlikely, however. Both observers knew the precise form of the target function; moreover, both observers had a clear, informal grasp of the task. Specifically, both observers were aware that the correct orientation comprised high numbers of white bars mixed with high numbers of moderately dark (but nonblack) bars, whereas the incorrect orientation comprised high numbers of black bars mixed with high numbers of moderately bright (but nonwhite) bars. In short, observers had a clear, conscious understanding of the task goal. Moreover, as previously discussed, both observers had extensive practice (in a preliminary study) in each of the five texture orientation judgments prior to the start of the current study. Similar observations apply to Experiment 5.

10. Experiment 5

Experiment 5 was similar to Experiment 4. As in Experiment 4, the goal in the current experiment was to construct a simple target function that would be difficult to synthesize as a linear combination of positive and negative half-wave rectifiers. The function f_{order4} tested in this experiment assigns the values 0.539, -1.0, -0.231, 0.692, 0.692, -0.231, -1.0, 0.539 to intensities $\lambda_0, \lambda_1, \ldots, \lambda_7$ (shown in Fig. 6). f_{order4} is the (discrete analog of the) fourth order Legendre polynomial. It is the lowest order Legendre polynomial that is orthogonal to all polynomials of order three or less (it is thus orthogonal to f_{order3}), and the lowest order even-symmetric Legendre polynomial orthogonal to all polynomials of order two or less.

JT was tested at this task using the experimental conditions shown in Table 9 (accessible via http://texel.ss.uci.edu/Links.html). The modulators ϕ_j, $j = 1, 2, \ldots, 13$ used for JT were constructed as follows. ϕ_1 was chosen to correlate positively with f_{order3}. The other 12 modulators were generated by first obtaining a set of 6, random, orthonormal functions, ρ_k, $k = 1, 2, \ldots, 6$, each of which was also orthogonal to ϕ_1. Then for $j = 2, 3, \ldots, 7$, ϕ_j was obtained by summing ϕ_1 and ρ_{j-1}, rounding the resulting values, and making minimal adjustments to make sure the resulting function satisfied Eq. (9). For $j = 8, 9, \ldots, 13$, ϕ_j was obtained by taking the difference $\phi_1 - \rho_{j-1}$, rounding the resulting values, and making minimal adjustments to make sure the resulting function satisfied Eq. (9).

Observer JT performed 10 blocks, each comprising 130 trials, 10 from each condition, randomly mixed. This yielded 100 trials per condition.

The conditions used for observer CC are shown in Table 10 (accessible via http://texel.ss.uci.edu/Links.html). These 45 conditions were designed (in the same way as were the conditions for CC used in Experiment 4) to discourage the use of strategies leading to impact functions resembling any of the impact functions achieved by CC in any of Experiments 1, 2 or 3. Each modulator correlated strongly and positively with the target function, f_{order4}. In addition, however, modulators were designed to correlate strongly but randomly in sign

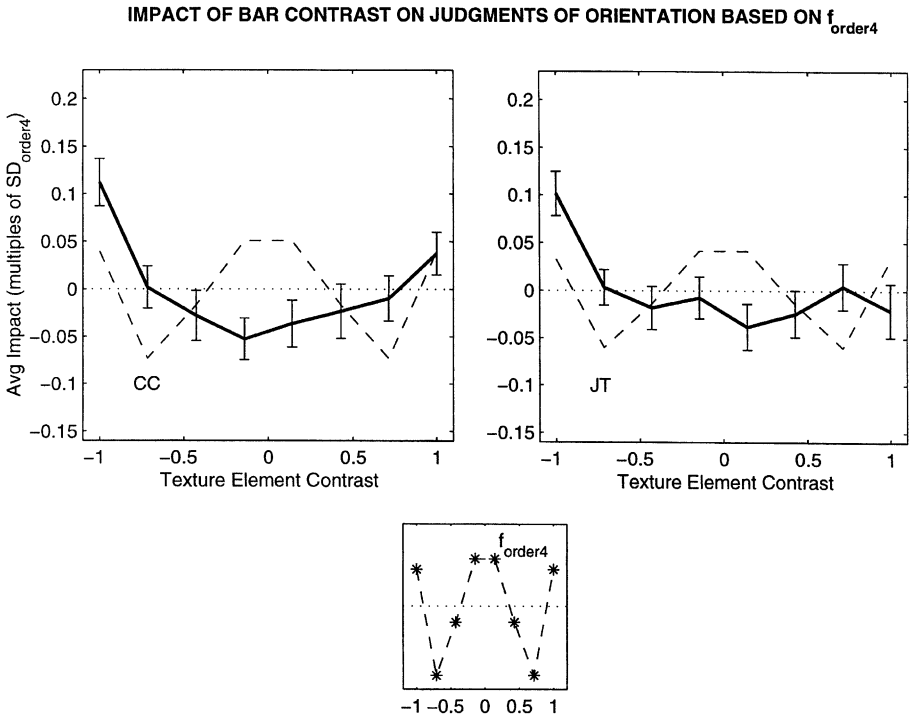

Fig. 6. The estimated impact functions for the two observers for judgments of orientation based on f_{order4}. Error bars indicate (bootstrapped) 95% confidence intervals of estimated impact function values. The impact function for CC (JT) is shown on the left (right). The dashed line in each figure shows f_{order4} with norm set equal to that of the corresponding impact function. The target function f_{order4} is shown by itself in the inset (center/bottom). Impact functions for both observers closely resemble those obtained in Experiment 2 (Fig. 3), suggesting that observers cannot effectively adapt to the task posed by f_{order4}.

with previously obtained impact functions. Thus reliance on any previously used impact function would yield poor performance.

Observer CC performed 6 blocks, each comprising 225 trials, five from each condition, yielding a total of 30 trials per condition.

10.1. Results

The raw data are given in Tables 9 (JT) and 10 (CC) (accessible via http://texel.ss.uci.edu/Links.html). The corresponding impact functions are shown in Fig. 6. The correlation coefficient between the impact function of JT (CC) and f_{order3} is 0.07 (0.14). The efficiency of JT (CC) in the current task was 0.01 (0.02).

10.2. Discussion

Efficiencies of both observers are quite low compared to those obtained in Experiments 1–3. Observers were unable to synthesize effective judgment strategies for this task. Moreover, the impact functions produced by both observers resemble very closely those that they produced in Experiment 2, even though these impact functions were ill-suited to the current task.

We conclude that f_{order4} is nearly orthogonal to the space of impact functions achievable by our observers.

11. General discussion

Experiments 1–3 indicate that observers can exercise some top–down control over the strategies they use in judging texture orientation. The optimal strategy for the task posed in Experiment 1 involved the use of a parabolic impact function, and the impact functions of both observers closely matched this ideal. However, Experiments 2 and 3 gave clear evidence that observers were not strictly constrained to using the parabolic impact functions they achieved in Experiment 1.

In Experiment 2, the task was designed to spur observers to tune their judgments to texture elements of negative contrast; in Experiment 3, observers were prompted to tune their judgments to texture elements of positive contrast. In each case, observers were able to meet the challenge effectively. The tasks were inherently more difficult than that posed in Experiment 1 because the target functions used in Experiments 2 and 3 contained more abrupt jumps in value. Nonetheless, observers were able to achieve impact functions that correlated strongly with the target functions, and efficiencies were similar to those obtained in Experiment 1.

Most importantly, the impact functions obtained in Experiments 1–3 showed substantial variability in form. The impact functions obtained in Experiment 2 were highly sensitized to negative contrasts, with suppressed

sensitivity to positive contrasts, whereas the opposite pattern held for the impact functions obtained in Experiment 3.

Under the model of Eqs. (1) and (2), the results of Experiments 1–3 suggest that at least some of the stuff-sensing arrays Q_i use positive half-wave rectification, and others use negative half-wave rectification. (Possibly, there may also be some channels that use full-wave rectification.)

The straightforward account runs as follows. We shall assume (for the sake of parsimony) that all channels are either positively or negatively half-wave rectifying. (This would be the case, e.g., if all of the stuff-sensing transformations Q_i were either on-center or off-center channels.) Then observers perform the task of Experiment 1 by assigning equal attentional weights to positive and negative half-wave rectifying, stuff-sensing channels with appropriate orientation tuning. By contrast, in Experiment 2, observers assign high attentional weights to negative half-wave rectifying channels with appropriate orientation tuning, and low attentional weights to positive half-wave rectifying channels.

Experiments 4 and 5 investigate whether observers' attentional flexibility is limited to what was demonstrated in Experiments 1–3. In Experiments 4 and 5, we used target functions that were as simple as we could devise under the constraint that they be approximately orthogonal to the impact functions achieved by observers in Experiments 1–3. Observers showed no ability to selectively tune their impact functions to either of these additional tasks. Indeed, the impact functions produced in each case were very similar in form to one or another of the impact functions observed in Experiments 1, 2, or 3—even though these impact functions were poorly suited to the current task.

In light of these results, it seems unlikely that observers will demonstrate proficiency at judging orientation based on f_{target} for any function f_{target} that cannot be reasonably well approximated by a linear combination of half-wave rectifiers.

This places an important constraint on those of the stuff-sensing transformations Q_i (occurring in Eq. (2)) that are tuned to the orientations in our stimuli. Specifically, we conclude with fair confidence that each of these transformations uses a pointwise transformation that can be well approximated as a linear combination of positive and negative half-wave rectifiers—or equivalently as a quadratic function of contrast.

Acknowledgements

Much of the research reported here was supported by National Science Foundation Human Cognition and Perception Program grant DBS-9203291. The authors also thank Adam Reeves for helpful comments on an earlier version.

Appendix A

Here we give a formal presentation of the model underlying the methods used to estimate impact functions for our observers in our various tasks. Consider the task (defined by Eq. (3)) of attempting to judge orientation defined by f_{target}.

In striving to judge whether $G_{\text{target}}(P) > 0$, the observer is assumed to compute a statistic $\widetilde{G}_{\text{target}}(P)$, which approximates $G_{\text{target}}(P)$ to the extent that this is within the subject's power. It is moreover assumed that $\widetilde{G}_{\text{target}}(P)$ results from summing impacts X_k, $k = 1, 2, \ldots, 128$, exerted individually by all of the oriented bars in P. The impact X_k exerted by the kth bar in texture patch P is assumed to be a random variable whose distribution depends on both the orientation and intensity of bar k. Moreover, the random variables, X_k, $k = 1, 2, \ldots, 128$ are assumed to be jointly independent (more on this assumption below).

For our purposes, only the means and standard deviations of these random variables matter. In this connection, we assume that there exist functions $m_{\text{target}} : \Lambda \to \Re$ and $s_{\text{target}} : \Lambda \to \Re$, such that

$$E[X_k] = \begin{cases} m_{\text{target}}[\lambda_i] & \text{if } P[k] \text{ has intensity } \lambda_i \text{ and is oriented up/right,} \\ -m_{\text{target}}[\lambda_i] & \text{if } P[k] \text{ has intensity } \lambda_i \text{ and is oriented up/left,} \end{cases} \tag{A.1}$$

and

$$\text{Std\_Dev}_{\text{target}}[X_k] = s_{\text{target}}[\lambda_i] \quad \text{if } P[k] \text{ has intensity } \lambda_i. \tag{A.2}$$

Finally, for Y a normal random variable with mean 0 and standard deviation σ, the observer is assumed to respond "up/right" if

$$\widetilde{G}_{\text{target}}(P) = \sum_{k=1}^{128} X_k + Y > 0. \tag{A.3}$$

The random variable Y is meant to capture trial-by-trial variability that does not depend on experimental condition. The standard deviation of Y might depend on various factors, e.g., size and shape of the texture patch, the specific intensities in Λ, viewing distance of the observer, etc., all of which we keep fixed throughout all the reported experiments. The function m_{target} reflects the average impact exerted on $\widetilde{G}_{\text{target}}(P)$ by bars of varying intensity, with impact assumed to be opposite in sign for opposite orientations. For this reason, we call m_{target} the *impact function* mediating the observer's judgments in the task defined by function f_{target}. The function s_{target}

reflects the amount of variability injected into the statistic $\widetilde{G}_{\text{target}}(P)$ by bars of varying intensity. Accordingly we call s_{target} the *noise injection function* compromising performance in the task defined by function f_{target}.

In any particular experiment, we fix the target function f_{target}, and apply the techniques of histogram contrast analysis (Chubb, 1999; Chubb et al., 1994; Chubb & Landy, 1991) in order to estimate the corresponding impact function m_{target}. These methods yield estimates of m_{target} that are invariant with respect to the values of all unmeasured model parameters (i.e., to σ, and to $s[\lambda_i]$, $i = 0, 1, \ldots, 7$).

Under our assumptions, $\widetilde{G}_{\text{target}}(P)$ should be approximately normal in distribution (this follows from the central limit theorem and the assumption that Y is normal). Thus, we can completely specify the distribution of $\widetilde{G}_{\text{target}}(P)$ if we can determine its mean and standard deviation. We can infer these statistics as follows.

On a trial in condition j, suppose that r and l are the histogram of up/right and up/left bar intensities. Then $r[\lambda_i]$ ($l[\lambda_i]$) bars in P are both oriented up/right (up/left) and painted with intensity λ_i. Write $R_{i,k}$ ($L_{i,k}$) for the impact exerted on $\widetilde{G}_{\text{target}}(P)$ by the kth such bar. Then we can express $\widetilde{G}_{\text{target}}(P)$ as follows:

$$\widetilde{G}_{\text{target}}(P) = \sum_{i=0}^{7} \sum_{k=1}^{r[\lambda_i]} R_{i,k} + \sum_{i=0}^{7} \sum_{k=1}^{l[\lambda_i]} L_{i,k} + Y. \tag{A.4}$$

Under the assumptions of Eqs. (A.1) and (A.2), we thus find that

$$\begin{aligned} E[\widetilde{G}_{\text{target}}(P)] &= \sum_{i=0}^{7} m_{\text{target}}[\lambda_i] r[\lambda_i] - \sum_{i=0}^{7} m_{\text{target}}[\lambda_i] l[\lambda_i] \\ &= m_{\text{target}} \cdot (r - l), \end{aligned} \tag{A.5}$$

and

$$\begin{aligned} \text{Var}[\widetilde{G}_{\text{target}}(P)] &= \sum_{i=0}^{7} s^2_{\text{target}}[\lambda_i] r[\lambda_i] + \sum_{i=0}^{7} s^2_{\text{target}}[\lambda_i] l[\lambda_i] + \sigma^2 \\ &= s^2_{\text{target}} \cdot (r + l) + \sigma^2, \end{aligned} \tag{A.6}$$

where $s^2_{\text{target}}[\lambda_i] = (s_{\text{target}}[\lambda_i])^2$.

Let ϕ_j be one of the N_{target} modulators used in an experiment to estimate m_{target}. (Thus, ϕ_j satisfies Eq. (9).) Then consider a stimulus P constructed so that the contingent of up-right oriented bars has histogram $r = 8 + \phi_j$, and the contingent of up-left oriented bars has histogram $l = 8 - \phi_j$. Then, Eqs. (A.5) and (A.6) imply

$$\begin{aligned} E[\widetilde{G}_{\text{target}}(P)] &= m_{\text{target}} \cdot ((8 + \phi_j) - (8 - \phi_j)) \\ &= 2 m_{\text{target}} \cdot \phi_j, \end{aligned} \tag{A.7}$$

and

$$\begin{aligned} \text{Var}[\widetilde{G}_{\text{target}}(P)] &= s^2_{\text{target}} \cdot ((8 + \phi_j) + (8 - \phi_j)) + \sigma^2 \\ &= 16 |s_{\text{target}}|^2 + \sigma^2. \end{aligned} \tag{A.8}$$

Thus (assuming the normality of $\widetilde{G}_{\text{target}}(P)$), for Φ the standard normal cdf,

$$\text{Prob}[\widetilde{G}_{\text{target}}(P) > 0] = \Phi\left(\frac{m_{\text{target}} \cdot \phi_j}{\text{SD}_{\text{target}}}\right), \tag{A.9}$$

for

$$\text{SD}_{\text{target}} = \left(4 |s_{\text{target}}|^2 + \frac{\sigma^2}{4}\right)^{1/2}. \tag{A.10}$$

Note first that $\text{SD}_{\text{target}}$, the standard deviation of $\widetilde{G}_{\text{target}}(P)$, is constant across different experimental conditions (i.e., $\text{SD}_{\text{target}}$ does not depend on ϕ_j). The units of m_{target} are arbitrary. Accordingly, as a matter of convention, we express m_{target} in multiples of $\text{SD}_{\text{target}}$. This lets us write Eq. (A.9) more simply:

$$\text{Prob}[\widetilde{G}_{\text{target}}(P) > 0] = \Phi(m_{\text{target}} \cdot \phi_j). \tag{A.11}$$

Concerning model plausibility. The model involves two rather strong assumptions.

1. We assume that the impacts $R_{i,k}$ and $L_{i,k}$ (in Eq. (A.4) exerted on the statistic $\widetilde{G}_{\text{target}}(P)$ by separate bars in texture patch P are jointly independent.
2. In assuming that the impact exerted by a bar depends only on its orientation and its intensity, we implicitly assume that impact does not depend on the location of the bar within the patch.

Neither of these assumptions (independence, nor spatial homogeneity of influence) is likely to be strictly true. However, as we now discuss, neither assumption is likely to be grossly false, and model fits are robust with respect to modest failures of both assumptions.

A.1. Assumption 1

We anticipate that Assumption 1 (independence) will generally fail to hold for the following reason: for any given bar B in P, we expect that the impact exerted on $\widetilde{G}_{\text{target}}(P)$ by B will depend not only on the orientation and intensity of B, but also on the intensities of the bars near to B in P. For example, suppose that bar B is assigned the maximal intensity λ_7. Plausibly, in this case, the impact exerted on $\widetilde{G}_{\text{target}}(P)$ by B may be greater if all the bars around B have minimal intensity λ_0, than if they are all equal in intensity to B. Under this scenario, the impact exerted by B is a systematic function not only of B's orientation and intensity, but also of the (randomly generated) context in which B is embedded.

We thus acknowledge that some of the variability in the impact exerted by a bar B is likely to be due at least in part to systematic influences of the random contexts

in which B occurs from trial to trial. However, one may ask what does this matter? In any case, one can still assess the mean impact $m_{\text{target}}[\lambda_i]$ exerted by an up/right oriented bar of intensity λ_i. Why should it matter whether the variability in impact exerted by such a bar is due to random noise or to random context variations?

The problem emerges when one notes that the probability of obtaining different contexts in P depends on the histograms r and l. To take an extreme case, suppose that r_1 assigns intensity λ_3 to all up/right bars, while l_1 assigns intensity λ_3 to all but one (randomly chosen) up/left bar which receives intensity λ_7. In this case, almost all bars in the patch P are assigned the same intensity, λ_3 (which is slightly below the background luminance of the display). Only a single bar in the entire display receives a different intensity. This bar is randomly chosen from among those oriented up/left, and it receives the maximal intensity λ_7. Note (ignoring differences in context that result if the bar occurs at the edge of the patch) that with probability 1, the single bar of intensity λ_7 is surrounded on all sides by bars of intensity λ_3. Thus, for the given histograms r_1 and l_1, there is no variability in the context in which the bar of intensity λ_7 occurs. By contrast, suppose that each of r_2 and l_2 is the uniform histogram (assigning 8 bars to each of the intensities in Λ). Then for a bar of intensity λ_7, any context is possible; indeed, most contexts occur with roughly equal probability.

Consider, then, the impact $L_{7,1}$ that is exerted on $\widetilde{G}_{\text{target}}(P)$ by the 1st up/left bar in P of intensity λ_7. If variability of impact depends on random variations in bar contexts from trial to trial, then the variance of $L_{7,1}$ will be much higher for P generated using up/right and up/left histograms r_2 and l_2 than for P generated using histograms r_1 and l_1 (since context is highly variable if P is generated using l_2 and r_2, but perfectly constant if P is generated using l_1 and r_1).

The model assumes that the random variable $R_{i,k}$ ($L_{i,k}$) generated by the kth up/right (up/left) bar of intensity λ_i has mean and variance that do not depend on the two histograms characterizing the patch in which this bar occurs. The previous remarks make it clear that context effects may well falsify this assumption.

Nonetheless, further considerations suggest that histogram-dependent changes in the distributions of random variables $R_{i,k}$ ($L_{i,k}$) are likely to be negligible. First, note that in all of the experiments reported here the observer strives to make a judgment of the sort given by Eq. (3). The target statistic $G_{\text{target}}(P)$ is, itself, devoid of contextual dependency in the following sense: the impact exerted on $G_{\text{target}}(P)$ by a given bar in patch P depends only on the intensity and orientation of the bar; the context in which a bar occurs has no influence on the impact exerted by that bar on $G_{\text{target}}(P)$. Thus, the observer has reason to try to eradicate contextual dependencies from the statistic $\widetilde{G}_{\text{target}}(P)$ as far as possible. Although these considerations do not insure that $\widetilde{G}_{\text{target}}(P)$ will be free of contextual dependencies, they do suggest that such dependencies are unlikely to predominate in determining $\widetilde{G}_{\text{target}}(P)$. Second, all of the histograms used in the current experiments embody substantial variability in intensity. Indeed, in every patch P presented as a stimulus in these studies, each intensity λ_i, $i = 0, 1, \ldots, 7$, is assigned to exactly 16 bars, some oriented up/left, and some up/right. Thus, on any given trial, for each intensity λ_i, we always have $l[\lambda_i] = 16 - r[\lambda_i]$. This means that the net ensemble of intensities in a stimulus patch is constant from trial to trial. Perhaps most importantly, each bar in the texture stimuli we use is embedded in a surrounding patch of uniform mean luminance. Thus even neighboring bars are buffered from each other by a sizeable expanse of blank field. Thus the immediate context of every bar in every stimulus is identical in all of the current experiments. Together these observations suggest that the means and variances of the random variables $R_{i,k}$ and $L_{i,k}$ should be approximately constant across different experimental conditions.

A.2. Assumption 2

Plausibly, the impact exerted on $\widetilde{G}_{\text{target}}(P)$ by a given bar B in P might depend not only on B's intensity and orientation, but also on its spatial location within P. For example, an observer might distribute her attention unevenly in making judgments, perhaps giving greater weight to bars near the center of the patch. These considerations lead to the following generalization of the model of Eq. (A.3). For some nonnegative, real-valued function W of bar location $k = 1, 2, \ldots, 128$,

$$\widetilde{G}_{\text{target}}(P) = \sum_{k=1}^{128} W[k]X_k + Y > 0, \tag{A.12}$$

where $W[k]$ reflects the spatial weight with which the random variable X_k contributes to $\widetilde{G}_{\text{target}}(P)$. Note first that the observer is intrinsically motivated to distribute attention evenly across the stimulus patch P, as all bars in P are equally informative about the value of $G_{\text{target}}(P)$. Of course, this does not guarantee that the observer will succeed in equalizing $W[k]$ across all bar locations k. However, as discussed in the appendix to Chubb (1999), impact function estimates are quite robust to moderate variation in W.

References

Adelson, E. H., & Bergen, J. R. (1991). The plenoptic function and the elements of early vision. In J. A. M. Michael, & S. Landy (Eds.), *Computational Models of Visual Processing* (vol. xii) (p. 394). Cambridge, MA: MIT Press.

Beck, J. (1982). Textural segmentation. In J. Beck (Ed.), *Organization and Representation in Perception*. Hillsdale, NJ: Erlbaum.

Beck, J., Graham, N., & Sutter, A. (1991). Lightness differences and the perceived segregation of regions and populations. *Perception and Psychophysics*, *49*, 257–269.

Beck, J., Prazdny, K., & Rosenfeld, A. (1983). A theory of textural segmentation. In J. Beck, B. Hope, & A. Rosenfeld (Eds.), *Human and Machine Vision* (pp. 1–38). New York: Academic Press.

Benton, C. P., & Johnston, A. (1999). Contrast inconstancy across changes in polarity. *Vision Research*, *39*(24), 4076–4084.

Bergen, J. R., & Adelson, E. H. (1988). Visual texture segmentation based on energy measures. *Journal of the Optical Society of America A*, *3*(13), 98–101.

Bergen, J. R., & Landy, M. S. (1991). The computational modeling of visual texture segregation. In M. S. Landy, & J. A. Movshon (Eds.), *Computational Models of Visual Processing* (vol. 1) (pp. 253–271). Cambridge, MA: MIT Press.

Bovik, A. C., Clark, M., & Geisler, W. S. (1990). Multichannel texture analysis using localized spatial filters. *IEE Transactions on Pattern Analysis and Machine Intelligence*, *12*, 55–73.

Caelli, T. (1985). Three processing characteristics of visual texture segmentation. *Spatial Vision*, *1*, 19–30.

Chubb, C. (1999). Texture based methods for analyzing elementary visual substances. *Journal of Mathematical Psychology*, *43*, 539–567.

Chubb, C., Econopouly, J., & Landy, M. S. (1994). Histogram contrast analysis and the visual segregation of IID textures. *Journal of the Optical Society of America A*, *11*(9), 2350–2374.

Chubb, C., & Landy, M. S. (1991). Orthogonal distribution analysis: a new approach to the study of texture perception. In J. A. M. Michael, & S. Landy (Eds.), *Computational Models of Visual Processing* (vol. xii) (p. 394). Cambridge, MA: MIT Press.

Clark, M., & Bovik, A. C. (1989). Experiments in segmenting texton patterns using localized spatial filters. *Pattern Recognition*, *22*, 707–717.

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

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

Foley, J. M. (1994). Human luminance pattern vision mechanisms: masking experiments require a new model. *Journal of the Optical Society of America A*, *11*, 1710–1719.

Graham, N. (1991). Complex channels, early local nonlinearities, and normalization in texture segregation. In J. A. M. Michael, & S. Landy (Eds.), *Computational Models of Visual Processing* (vol. xii) (p. 394). Cambridge, MA: MIT Press.

Graham, N. (1994). Non-linearities in texture segregation. In J. A. Goode, & G. R. Bock (Eds.), *Ciba Foundation Symposium on Higher-order Processing in the Visual System, 184* (pp. 309–329). Chichester, England: Wiley.

Graham, N., Beck, J., & Sutter, A. (1992). Nonlinear processes in spatial-frequency channel models of perceived texture segregation: effects of sign and amount of contrast. *Vision Research*, *32*(4), 719–743.

Graham, N., Sutter, A., & Venkatesan, C. (1993). Spatial-frequency and orientation-selectivity of simple and complex channels in region segregation. *Vision Research*, *33*(14), 1893–1911.

Grossberg, S., & Mingolla, E. (1985). Neural dynamics of perceptual grouping: textures, boundaries, and emergent segmentations. *Perception and Psychophysics*, *38*, 141–171.

Heeger, D. J. (1991). Nonlinear model of neural responses in cat visual cortex. In M. S. Landy, & J. A. Movshon (Eds.), *Computational Models of Visual Processing* (pp. 119–133). Cambridge, MA: MIT Press.

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience*, *9*, 181–198.

Hubel, H. H., & Wiesel, T. N. (1961). Integrative action in the cat's lateral geniculate body. *Journal of Physiology*, *155*, 385–398.

Ito, M., Fujita, I., Tamura, H., & Tanaka, K. (1994). Processing of contrast polarity of visual images in inferotemporal cortex of the macaque monkey. *Cerebral Cortex*, *4*(5), 499–508.

Julesz, B. (1962). Visual pattern discrimination. *IRE Transactions of Information Theory*, *IT-8*, 84–92.

Julesz, B. (1975). Experiments in the visual perception of texture. *Scientific American*, *(April)*, 34–43.

Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature*, *290*(12), 91–97.

Julesz, B., Bergen, J. R. (1983). Textons, the fundamental elements in preattentive vision and perception of textures. Orginally published in *Bell Systems Technical Journal 62*(6) (Part II), 1619–1645; Reprinted In M. A. Fischler, & O. Firschein (Eds.) (1986) *Readings in computer vision* (pp. 243–256). Los Altos, CA: Morgan Kaufmann Publishers.

Julesz, B., Gilbert, E. N., Shepp, L. A., & Frisch, H. L. (1973). Inability of humans to discriminate between visual textures that agree in second order statistics. *Perception*, *2*, 391–405.

Julesz, B., Gilbert, E. N., & Victor, J. D. (1978). Visual discrimination of textures with identical third order statistics. *Biological Cybernetics*, *31*, 137–149.

Knutsson, H., & Granlund, G. H. (1983). Texture analysis using two-dimensional quadrature filters. In *IEEE Computer Society Workshop on Computer Architecture for Pattern Analysis and Image Database Management*.

Kuffler, S. W. (1953). Discharge patterns and functional organization of the mammalian retina. *Journal Neurophysiology*, *5*, 18–26.

Landy, M. S., & Bergen, J. R. (1991). Texture segregation and orientation gradient. *Vision Research*, *31*(4), 679–691.

Malik, J., Belongie, S., Shi, J., & Leung, T. (1997). Textons, contours and regions: cue combination in image segmentation. In *International Conference on Computer Vision, September 1999, Corfu, Greece*.

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

Morgan, M. J., & Watt, R. J. (1997). The combination of filters in early spatial vision: a retrospective analysis of the MIRAGE model. *Perception*, *26*(9), 1073–1088.

Nothdurft, H.-C. (1994). Common properties of visual segmentation. In G. R. Bock, J. A. Goode, et al. (Eds.), *Ciba Foundation symposium on Higher-order Processing in the Visual System* (vol. viii) (pp. 245–268, 347 pp.). Chichester, England: Wiley.

Robson, J. G. (1980). Neural images: the physiological basis of spatial vision. In C. S Harris (Ed.), *Visual Coding and Adaptability* (pp. 177–214). Hillsdale, NJ: Erlbaum.

Rosenholtz, R. (1999). Significantly different textures: a texture segmentation algorithm. *Investigative Ophthalmology and Visual Science*, *40*(4), S200.

Rubenstein, B. S., & Sagi, D. (1990). Spatial variability as a limiting factor in texture discrimination tasks: implications for performance asymmetries. *Journal of the Optical Society of America A*, *7*, 1632–1643.

Sperling, G., & Lu, J.-L. (1999). Unequal representation of black and white in human vision. *Investigative Ophthalmology and Visual Science*, *40*(4), S200.

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*(4), 312–332.

Sutter, A., & Graham, N. (1995). Investigating simple and complex mechanisms in texture segregation using the speed-accuracy tradeoff method. *Vision Research*, *35*(20), 2825–2843.

Victor, J. D., & Brodie, S. (1978). Discriminable textures with identical Buffon needle statistics. *Biological Cybernetics*, *31*, 231–234.

Victor, J. D., & Conte, M. M. (1991). Spatial organization of nonlinear interactions in form perception. *Vision Research*, *31*, 1457–1488.

Victor, J. D., Conte, M. M., Purpura, K., & Katz, E. (1994). Isodipole textures: a window on cortical mechansisms of form processing. In T. V. Papathomas, C. Chubb, A. Gorea, & E. Kowler (Eds.), *Early Vision and Beyond*. Cambridge, MA: MIT Press (Chapter 10).

Wainwright, M., & Simoncelli, E. P. (1999). Characterization of adaptation in V1 neurons with a statistically-derived normalization model, in preparation.

Watt, R. J., & Morgan, M. J. (1985). A theory of the primitive spatial code in human vision. *Vision Research*, *25*, 1661–1674.