Available online at www.sciencedirect.com
ScienceDirect
Vision Research 47 (2007) 2237–2248
Vision Research
www.elsevier.com/locate/visres

The three dimensions of human visual sensitivity to first-order contrast statistics

Charles Chubb [a,*], Jong-Ho Nam [c], Daniel R. Bindman [b], George Sperling [a]

[a] Department of Cognitive Sciences and Institute for Mathematical Behavioral Sciences, University of California at Irvine, Irvine, CA 92697-5100, USA
[b] ALEKS Corporation, 15641 Red Hill Avenue, Suite 150, Tustin, CA 92780, USA
[c] Department of Psychology, The Catholic University of Korea, 43-1 Yokkok 2-dong Wonmi-gu, Puchon City, Kyonggi-do 420-743, Korea

Received 12 October 2006; received in revised form 18 March 2007

Abstract

This work studies the preattentive discrimination of achromatic textures composed of mixtures of different (Weber) contrasts. These textures differ not at all in local spatial structure, but only in the relative proportions of different contrasts they comprise. It is shown that, like chromatic discrimination, preattentive discrimination of such textures is three-dimensional. The current results do not uniquely determine the characteristics of the three texture filters mediating human discrimination of these textures; they do, however, define the space of textures with 4th-order polynomial histograms to which human vision is sensitive. Three real-valued functions of contrast that collectively capture human sensitivity to the textures in this space are presented.

Keywords: Texture perception; Texture segregation; Blackshot; Scramble; Contrast

1. Introduction

It has often been observed that vision is not just a camera. Vision does not merely capture a 2-d image of the pattern of light playing over the retina; rather it operates with remarkable inferential power to construct a world populated by objects in 3-d space directly revealing their manifold "affordances" (Gibson, 1979).

However, vision does use some camera-like processes. The retina itself operates like a camera. Under photopic conditions, the retina captures three, time-varying images: one with the L-cones, another with the M-cones and a third with the S-cones. These three images provide all the input to photopic vision. And, as dramatized by Julesz (1962, 1975) and Beck (1966, 1982), these are not the only camera-like processes in vision. In experiments using textures composed of short, oriented line segments, they observed that it is easy to find textures X and Y such that the location and shape of a patch of X embedded in a background of Y are immediately apparent despite the fact that on average, both textures produce equal activation in each cone class. In the case of retinal cones, we know that a cone of a particular type captures light of certain wavelengths in its small local receptive field. And the visual field is more-or-less fully tiled with cones of each of the three types. In the case of textures, we suppose that there are more complex detectors, sensitive to lines of particular orientations, and that the visual field is more-or-less fully tiled with each particular type. These early examples and conceptualizations launched the field of research into preattentive texture segregation.

The many "back pocket" models (Chubb & Landy, 1991) proposed to account for preattentive texture segregation (e.g., Bergen & Adelson, 1988; Bergen & Landy, 1991; Bovik, Clark, & Geisler, 1990; Caelli, 1985; Fogel & Sagi, 1989; Graham, 1991; Graham & Sutter, 1998; Graham,

* Corresponding author. Fax: +1 949 824 2517.
E-mail address: cfchubb@uci.edu (C. Chubb).

0042-6989/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2007.03.025

Sutter, & Venkatesan, 1993; Grossberg & Mingolla, 1985; Knutsson & Granlund, 1983; Landy & Bergen, 1991; Malik & Perona, 1990) all posit two primary processing stages:

1. A *filtering stage* in which some fixed battery of *texture filters* is applied to the visual input to produce a corresponding set of output images. Such a filter is analogous to a particular cone class.
2. A subsequent *surveying stage* in which the variations perceived within the visual field are derived from the filtering stage output images. The outcome of a survey is analogous to the assignment of a particular color and brightness to a region.

To be useful, the filtering must be fast; each texture filter must work like a movie camera, capturing a continuous image stream. This leads us to view all texture filters as retinotopically organized arrays of neurons, and the output images they produce as "neural images" (Robson, 1980). Each texture filter thus requires a substantial commitment of resources: it sits at the top of a vast, pyramid of subsidiary neurons, all of which are continuously active in normal vision. Given the cost in resources of a single texture filter, it seems likely that the number of distinct texture filters in human vision is small.[1]

A basic question about any texture filter is: what does it sense? In one of these texture filter arrays, any given filter is assumed to monitor a fixed, local region of the visual field, computing from the input to that region the same statistic as do all the other filters in the array for their regions. It is this *texture filter statistic* that functionally defines the filter. The central questions about preattentive visual processing are how many texture filters are needed to characterize human vision? and what statistics do they sense?

The method we develop here involves searching for "texture metamers," a general approach first explored by Richards and colleagues (Richards, 1979; Richards & Riley, 1977). Two lights with different spectra are called "metamers" if they appear identical in color and brightness. From color theory we know that this will be true only if the two lights produce the same levels of activation in each of the three cone classes. We shall make precisely analogous assumptions about textures. That is, we will assume that human vision embodies some number of "texture filters" (analogous to the different cone classes) that enable preattentive texture discrimination and that two textures will appear identical (at least to preattentive vision) if they produce the same levels of activation in each of these texture filters. We expect that most readers will be familiar with color theory, and we will try to sharpen intuitions about our texture manipulations by drawing the analogy between textures and colors wherever appropriate.

Suppose we create a display that consists of a patch of one color surrounded by a background of a different color. To make the central patch disappear, i.e., to match the background exactly in color and brightness, we need to be able to manipulate at least three different components, paints or lights, that comprise the central patch. For example, in computer monitors three differently colored lights are combined to create the patch and background. There is a single, unique combination that produces a match between the patch and the background. With four or more lights or paints there are infinitely many different combinations that can produce an exact match, and with only two lights, matches are usually not possible. From this we infer that color vision is three-dimensional, i.e., that there are three receptors. However, this experiment does not enable us to say what the receptors are (Krantz, 1975). The present experiment deals with random textures, in which one texture is used as a background, and a patch is composed as a mixture of different textures. As in the case of color, we ask how many different component textures are needed in the mixture to enable an appropriate combination of the components to match any arbitrarily chosen background? And, what are the essential properties of the component textures to enable a match? The texture experiment is completely analogous to the color experiment.

We address these questions in a restricted domain in the current project. Here, we focus on achromatic textures composed of randomly arranged elements of different intensities. The first question is, what's the minimum number of texture filters required to account for human discrimination of these sorts of textures? We find that the number is three. The second question is, what do these texture filters sense? Since our textures vary only in the proportions of different intensities they contain, we can only talk about the sensitivity of these filters to different intensities. We cannot say anything about the sensitivity of the implicated texture filters to orientation, spatial frequency or various other aspects of the local visual input. In fact, the current experiments by themselves do not uniquely determine the sensitivities of individual texture filters to different Weber contrasts. However, they do determine the

[1] This statement slightly oversimplifies the situation. Consider, for example, sensitivity to texture orientation in a particular spatial frequency band. Suppose, as seems likely, that within a given spatial frequency band, striate cortex contains neurons tuned to all orientations. This is quite different from the case of only three types of retinal cones. Infinitely many differently tuned neurons could be taken to mean that there exist infinitely many mechanisms tuned to distinct center orientations. However, suppose the tuning of any particular neuron is no sharper than approximately 60 deg. Suppose the orientation tuning curves (sensitivity as a function of stimulus line-segment orientation) of all of these neurons have similar, Gaussian profiles, differing only in the central orientation. If so, then the tuning functions of all of these neurons span a space of dimension 3 (oriented line segments are described within 180 deg). Dimension 3 in turn implies that, when any small spatial neighborhood is presented with a complex mixture of differently oriented line segments (i.e., a spectrum of orientations), one needs to know only the responses of three neurons (analagous to three color receptors) with linearly independent tuning curves in order to predict the response to the orientations mixture. This example illustrates that even though the population of neurons represents a continuum of center frequencies, it would be nonetheless true that the number of distinct (taken here to mean nonredundant) orientation-selective filters within a given spatial frequency band could be quite modest.

three-dimensional space of textures spanned by the three sensitivity functions. An orthonormal basis of this space is provided later in Fig. 7.

2. The generation and description of textures

We begin with essential definitions (e.g., Chubb, Landy, & Econopouly, 2004).

2.1. Texture patches

2.1.1. Weber contrast

We use the term "Weber contrast" to refer to the normalized deviation of stimulus luminance from the background luminance (to which it is assumed the observer is adapted). Thus, if the background luminance is B, then the Weber contrast of a texture element with luminance L is $(L - B)/B$.

2.1.2. The set of 17 contrasts, C

Our stimuli are arrays of small square texels (texture elements) painted with values from C, the set of 17 (Weber) contrasts -1, $-7/8, \ldots, 7/8, 1$. (Thus the contrasts in C increase linearly from -1 (black) through shades of gray up to *1* (white).)

Under the analogy to color theory, the set C corresponds to the set of quanta of different wavelengths. Just as a light can be viewed as a mixture of quanta of different wavelengths, one of our textures can be viewed as a mixture of the different contrasts in C.

2.1.3. Histograms

The term "histogram" usually refers to a function that gives the number of pixels in an image that take any given value. We deviate from this usage in calling any probability distribution on C a *histogram*.

Under the analogy to color theory, the histogram of a texture corresponds to the spectrum of a light. However, the analogy breaks down in one way that should be noted. Light spectra can vary in amplitude, but texture histograms cannot. Thus, for example, one can double the intensity of a light by doubling its quantal flux at each wavelength; however, there is no way to cram more elements into a patch of texture of some fixed area. In this regard, our textures are like equiluminant lights (i.e., lights that all share the same fixed intensity).

U denotes the uniform histogram: i.e., $U(c) = 1/17$ for all $c \in C$. Thus a texture with histogram U is like a white light—i.e., a light with a flat spectrum.

2.1.4. Scrambles

To generate a texture patch with N texels, we first stipulate a histogram p. Then we load a virtual urn with N contrasts from C in proportions conforming as closely as possible to p and draw N times from the urn *without replacement* to assign contrasts to texels. The resulting patch S_p is called a *scramble* with histogram p. Thus, S_U is a scramble with equal proportions of all 17 contrasts.[2] A texture patch composed of a particular scramble is analogous to a patch of light with a particular spectrum.

2.2. Impact functions

A cone's response to a light is very simple: the cone just adds up all the quanta in the light weighted by its differential sensitivity to their different wavelengths. Take an S-cone, for example. The S-cone's sensitivity to different wavelengths w is captured by its *spectral sensitivity function* $f_S(w)$. Thus the response of an S-cone to a light with spectrum r is given by $f_S \cdot r = \int f_S(w) r(w)\,dw$, the *dot product* of r with f_S. Similarly, for f_M and f_L the M- and L-cone spectral sensitivity functions, $f_M \cdot r = \int f_M(w) r(w)\,dw$ and $f_L \cdot r = \int f_L(w) r(w)\,dw$ give the corresponding responses of M- and L-cones.

In the domain of scrambles, we assume that human vision comprises some set of texture filters $F_1, F_2, \ldots, F_N$ (analogous to the three cone classes) that enable discrimination of scrambles. We summarize the scramble-sensitivity of given one of these filters, F_k, by its *impact function* f_k. Just as, for example, f_S gives the sensitivity of an S-cone to quanta of different wavelengths, f_k gives the sensitivity of F_k to texels of different contrasts, and just as the dot product $f_S \cdot r$ gives the S-cone activation produced by a light of spectrum r, $f_k \cdot p = \sum_{c \in C} f_k(c) p(c)$ gives the average F_k-activation produced by a scramble with histogram p.

It is important to realize that the impact functions f_k give us only very limited information about the computations performed by the texture filters F_k. A given F_k computes a fixed texture filter statistic at each location (x, y) in space. This texture filter statistic is some (perhaps complicated and/or nonlinear) function of the input values in a spatial neighborhood of (x, y) called the *receptive field* of F_k. If we think of F_k as being embodied by an array of neurons, then a given texel, τ, of contrast c in an input scramble will influence the output values of all of those F_k neurons in whose receptive fields it falls. Thus, τ will distribute its influence across a number of F_k output neurons. Moreover, the response of a given one of these neurons will depend not only on τ but also on the contrasts of all of the other texels that fall in its receptive field. Thus it is natural to expect the influence of a texel of contrast c to depend to some extent on the configuration of texel contrasts surrounding it. All of the details of this process and all context-dependent variations in the impact exerted by a texel of contrast c are hidden by f_k. For any contrast $c \in C$,

[2] Chubb, Econopouly, and Landy (1994) used independent, identically distributed (IID) textures, whose texel values are jointly independent random contrasts identically distributed as p. Unlike a scramble, an IID texture patch may have different contrasts in proportions that deviate randomly from p. We use scrambles to eliminate uncontrolled histogram variability. Note, however, that as a consequence scrambles contain spatial dependencies. Thus, for example, if the top-half of S_p happens by chance to have a mean that is higher than the mean of p, then the bottom-half must have a mean that is lower.

$f_k(c)$ tells us only the net, average impact exerted by a texel of contrast c on the space-average output of F_k. However, the texture filters most useful for discriminating scrambles will be the ones whose space-average responses are strongly and systematically influenced in different ways by texels of different contrasts. It is precisely these systematic influences that are captured by the impact functions f_k.

2.3. Modulators and modulation spaces

Different scrambles are generated by perturbing their histograms away from uniformity. The functions we use to achieve these perturbations are called *modulators*. Specifically, a function $m : C \to \Re$ is called a *modulator* if each of $U + m$ and $U - m$ is a probability distribution (i.e., is nonnegative for all $c \in C$ and sums to 1). In addition m is called *maximal*[3] if $\max(|m|) = 1/17$. In fact, all of the modulators we use to generate stimuli in this project are maximal.

In our experiments, subjects discriminate between scrambles S_{U+m} and S_{U-m} for various maximal modulators m. In fact, often, as a shorthand for saying that an observer correctly discriminates S_{U+m} from S_{U-m} with probability ρ we will instead say that "The observer discriminates m with success rate ρ" or, if ρ is close to 1.0, "m is easily discriminable".

The set of functions spanned by N linearly independent modulators is called a *modulation space* of dimension N.[4] Ω denotes the 16-dimensional space containing all modulators (Ω is only 16-dimensional because any modulator is orthogonal to the constant function U).

2.4. Projection

For a light with spectrum r, much of the structure of r is likely to be irrelevant to the color produced by the light. In fact, we can split r uniquely into the part that is relevant and the part that is irrelevant. For Θ the (three-dimensional) space of functions spanned by the three cone spectral sensitivity functions f_S, f_M and f_L, the relevant part of r is called the projection of r into Θ; it produces S-, M- and L-cone activations equal to those produced by r. The irrelevant part is orthogonal to each of f_S, f_M and f_L; as a result, it produces S-, M- and L-cone activations equal to zero.

Moving back to scrambles, let $\phi_1, \ldots, \phi_N$ be an orthogonal basis of Φ, a subspace of Ω (An orthogonal basis can always be derived using the Gram–Schmidt procedure.). For any function $f : C \to \Re$, the *projection* of f into Φ is

$$\mathrm{Proj}_{\Phi}(f) = \sum_{i=1}^{N} W_i \phi_i, \tag{1}$$

where,

$$W_i = \frac{\phi_i \cdot f}{\phi_i \cdot \phi_i}, \quad i = 1, \ldots, N. \tag{2}$$

One can think of $\mathrm{Proj}_{\Phi}(f)$ as the portion of f that resides in (is accounted for by) Φ. The following basic fact is easily proven:

2.4.1. Proposition

For any function $f : C \to \Re$ and any subspace Φ of Ω, $f - \mathrm{Proj}_{\Phi}(f)$ is orthogonal to every function in Φ.

2.5. The discrete Legendre modulators

The maximal modulators $\lambda_1, \ldots, \lambda_7$ (Fig. 1) play a central role here. These modulators (discrete domain analogues of the Legendre polynomials) are orthogonal. Moreover, λ_i is an ith-order polynomial, for $i = 1, \ldots, 7$; thus, $\lambda_1, \ldots, \lambda_i$ span the space of all ith-order polynomial modulators, and in particular $\lambda_1, \ldots, \lambda_7$ span the space of all 7th-order modulators. Fig. 2 shows examples of scrambles S_{U+m} and S_{U-m} for $m = \lambda_1, \ldots, \lambda_4$.

Λ subscripted by any concatenation of integers $\in\{1, 2, \ldots, 7\}$ denotes the modulation space spanned by the the λ_is whose subscripts are in the concatenated sequence. Thus, for example, we write Λ_{123} for the modulation space spanned by λ_1, λ_2, λ_3 and Λ_{34567} for the modulation space spanned by λ_3, λ_4, $\ldots$,λ_7.

2.6. Assumptions about discrimination

2.6.1. Differential activation, $\Delta_k(m)$

Our task requires the subject to judge the discriminability of the textures S_{U+m} and S_{U-m} in the stimulus diagrammed in Fig. 3. The differential activation produced in texture filter F_k by this stimulus is assumed to be

$$\Delta_k(m) = f_k \cdot (U + m) - f_k \cdot (U - m) = 2 f_k \cdot m. \tag{3}$$

That is, differential activation of a texture filter by a modulator is simply twice the dot product of the filter impact function and modulator.

2.6.2. Discriminability and null points

We assume that discrimination performance is a strictly increasing function of $|\Delta_k(m)|$, for $k = 1, 2, \ldots, N$. For nonnegative scalers α, $|\Delta_k(\alpha m)| = \alpha|\Delta_k(m)|$. Thus, performance is optimized by choosing α so that αm is a maximal modulator. We also assume that chance discrimination performance occurs only if $|\Delta_k(m)| = 0$ (implying that $f_k \cdot m = 0$) for $k = 1, 2, \ldots, N$. Any maximal modulator m yielding chance discrimination is called a *null point*.

The idea of a null point is illustrated by the inability of dichromats to distinguish numbers hidden in an Ishihara color plate. Take the case of a protanope, i.e., a dichromat who lacks L-cones. For two lights to be indiscriminable to a protanope, their spectra r_1 and r_2 must satisfy $f_S \cdot r_1 = f_S \cdot r_2$ and $f_M \cdot r_1 = f_M \cdot r_2$. Suppose a computer monitor

[3] Note that this last condition implies that for any scaler A greater than one, there must exist some $c \in C$ for which either $U(c) + Am(c) < 0$ or $U(c) - Am(c) < 0$, implying that Am is not a modulator.

[4] Note that any modulation space contains many functions g for which $\max(|g|) > 1/17$ and thus are not modulators.

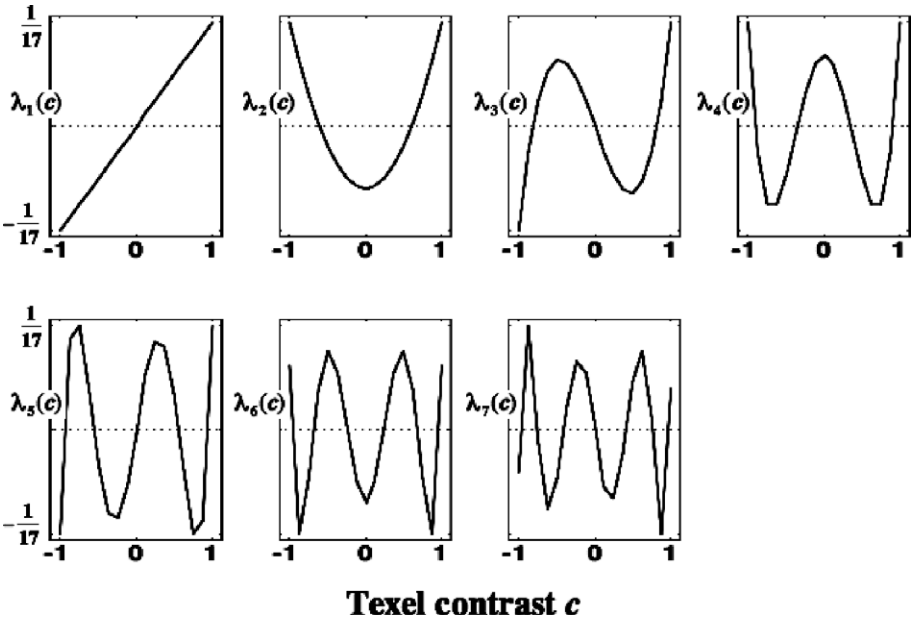

Fig. 1. Modulators $\lambda_1, \lambda_2, \ldots, \lambda_7$. The maximal modulators plotted here (and especially $\lambda_1, \lambda_2, \ldots, \lambda_4$) play a central role in this study. These modulators (discrete domain analogues of the Legendre polynomials) are orthogonal. Moreover, λ_i is an ith-order polynomial, for $i = 1, \ldots, 7$; thus, $\lambda_1, \ldots, \lambda_i$ span the space of all ith-order polynomial modulators.

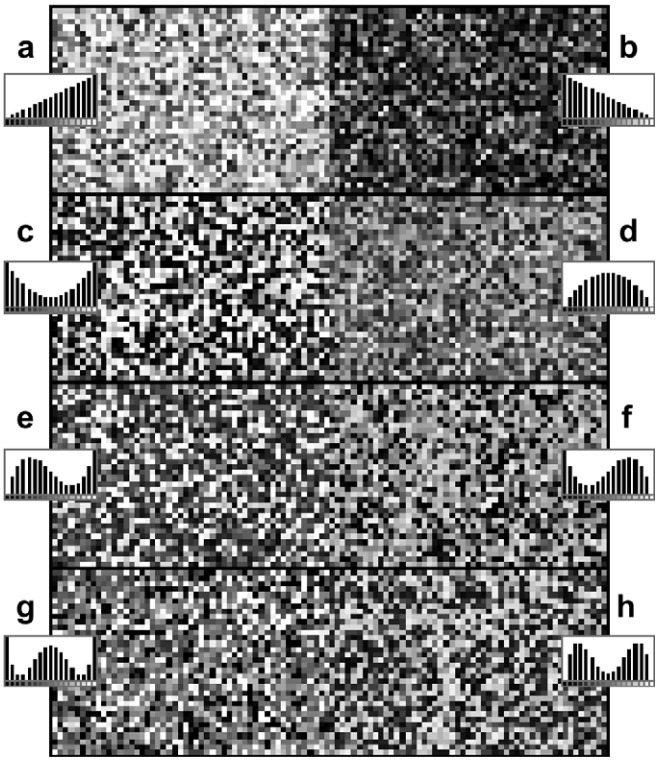

Fig. 2. Some scrambles from Λ_{1234}. The scrambles S_{U+m} and S_{U-m} for $m = \lambda_1, \ldots, \lambda_4$. The inset in each patch of scramble gives the histogram of that scramble.

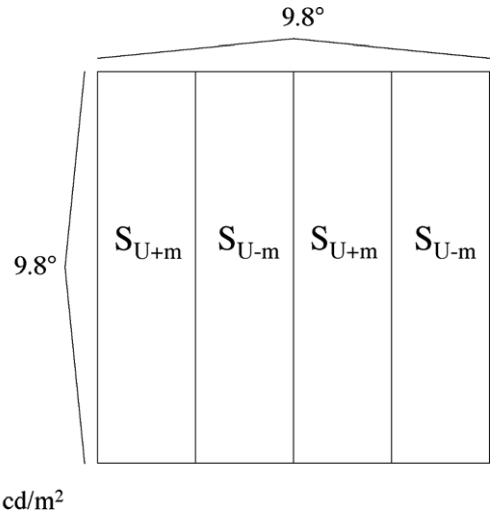

Fig. 3. Diagram of stimulus used to find minimal salience modulators.

creates lights from three sources (e.g., red, green and blue) with spectra s_1, s_2 and s_3 and let Θ be the space of functions spanned by s_1, s_2 and s_3. The projections into Θ of f_S and f_M span a planar subspace, which means there exists a non-zero function q in Θ that is orthogonal to both of f_S and f_M. Thus, for any "base light" with spectrum b, the lights with spectra $b + \alpha q$, $\alpha \in \Re$, all satisfy $f_S \cdot (b + \alpha q) = f_S \cdot b$ and $f_M \cdot (b + \alpha q) = f_M \cdot b$, implying that to the protanope all of these different lights appear identical to the base light with spectrum b. Such a set of lights is called a "protanopic confusion line". By choosing b appropriately so as to be able to take α as large as possible under the constraints of our display system we can create two lights with spectra $r_1 = b + \alpha q$ and $r_2 = b - \alpha q$ that are maximally discriminable to a normal trichromat yet which a protanope will be unable to discriminate. If we paint the digit in the Ishihara plate with one of these lights and the background with the

other, this will give us a maximally sensitive test of protanopia.

In our texture experiments, a null point m in some modulation space Φ is analogous to the function αq (with α taken as large as possible) in the space Θ; also, the scramble S_U is analogous to the "base light" with spectrum b. Just as all of the cones in the retina of the protanope give the same response to all of the lights with spectra $b + \alpha q$, all of the texture filters resident in human vision give the same response to scrambles S_{U+m} and S_{U-m}. As a result, these two scrambles will be preattentively indiscriminable even though m is a maximal modulator.

2.6.3. Guiding principles

Suppose there exist in human vision at most N texture filters with linearly independent impact functions f_k, $k = 1, 2, \ldots, N$.[5] Under our assumptions

1. Any modulation space Φ of dimension greater than N must contain null points. Moreover, if the dimension of Φ is $N + 1$, and if the projections into Φ of the impact functions f_k are linearly independent, then Φ will contain a unique null point.[6]
2. There exist modulation spaces Φ of dimension N devoid of null points.[7] (In particular, this is true of the space spanned by f_k, $k = 1, 2, \ldots, N$.)

See the Appendix for Proofs of 1 and 2. As will become clear below, the logic of the experiments reported here depends crucially on these two simple derivations from linear algebra.

2.7. Empirical background

In light of these observations, a thought experiment shows that human vision has at least two texture filters differentially sensitive to different scrambles. As one may imagine, it is impossible to compensate for a substantial difference in mean luminance of two scrambles by adjusting their relative variances, and similarly impossible to compensate for a substantial difference in variance between two scrambles by adjusting their relative mean luminances. This intuition can be tested by exploring the 1-d curve of all maximal modulators $m = A\lambda_1 + B\lambda_2$. All such m's yield discrimination well above chance, showing that at least two (and possibly more) texture filters discriminate modulators in Λ_{12}. We write F_1 and F_2 for two of the texture filters sensitive to $m \in \Lambda_{12}$ and f_1 and f_2 for their impact functions.

Chubb et al. (1994) demonstrated that a single texture filter, B, strongly predominates in discriminating modulators $m \in \Lambda_{34567}$ (with other texture filters exerting negligible influence on discrimination); they also measured $\mathrm{Proj}_{\Lambda_{34567}}(f_B)$, for f_B the impact function of B. If f_B were

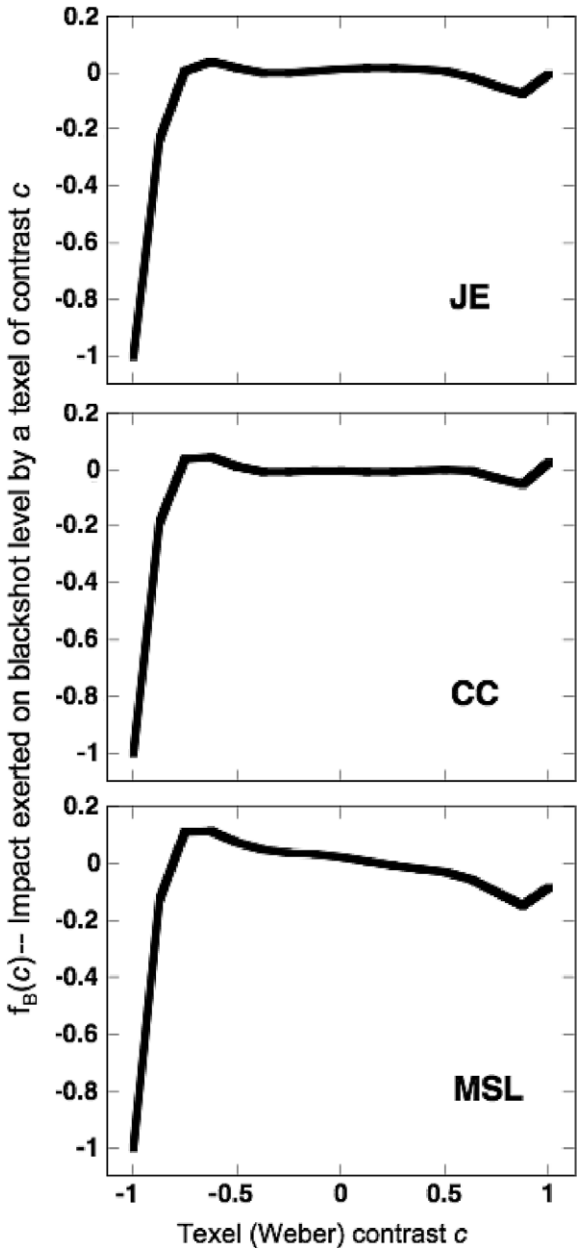

Fig. 4. The blackshot impact function. 7th-order polynomial approximations of the impact function f_B of the blackshot texture filter (for three observers). Note that the blackshot texture filter is sharply tuned to contrasts near −1.0, showing no significant differential sensitivity to contrasts between −0.875 and 1.0.

[5] Under this supposition, human vision may contain more than just these N texture filters; if so, however, the additional texture filters are either uninformative about scrambles (i.e., their impact functions are invariant across C) or redundant (i.e., their impact functions can be expressed as linear combinations of the impact functions f_k, $k = 1,2, \ldots, N$).

[6] As the function q used to generate protanopic confusion lines is unique in the space spanned by the spectra s_1, s_2 and s_3 of a given set of red, green and blue sources.

[7] Consider the case of a normal trichromat viewing an RGB display, and let Θ be the space of functions spanned by the spectra of the R, G and B sources. If the sources have been well-chosen, then the projections into Θ of the cone sensitivity functions f_S, f_M and f_L will be linearly independent, implying that Θ will contain no null points.

equal to $\mathrm{Proj}_{\Lambda_{34567}}(f_B)$, then $\mathrm{Proj}_{\Lambda_{12}}(f_B)$ would be zero. Chubb et al. (2004) showed that this was not the case by measuring $\mathrm{Proj}_{12}(f_B)$. They found that f_B correlates strongly both with λ_1 (positively) and λ_2 (negatively) and is thus quite sensitive to many of the modulators in Λ_{12}.

By combining the results of Chubb et al. (1994) and Chubb et al. (2004) one can derive the full 7th-order polynomial approximation ($\mathrm{Proj}_{\Lambda_{1234567}}(f_B)$) of f_B shown in Fig. 4. As Fig. 4 makes clear, B is sharply tuned to contrasts near −1.0, showing no significant differential sensitivity to contrasts between −0.875 and 1.0. Chubb et al. (2004) called this the "blackshot" texture filter.

There is no question that at least two distinct texture filters differentiate modulators in Λ_{12}. If there are *only* two such texture filters, then because $\mathrm{Proj}_{\Lambda_{12}}(f_B)$ is nonzero, one of them must be B. Alternatively, there may be more than two such texture filters. The main goal of the current experiment is to resolve this question.

3. Methods

3.1. General strategy

Suppose there exist only two texture filters, F_1 and F_2, with impact functions f_1 and f_2 that enable scramble discrimination. Then guiding principle 1 (Section 2.6.3) implies that each of the three-dimensional spaces Λ_{123}, Λ_{124}, Λ_{134}, Λ_{234} must contain at least one null point.

For two modulators m_1 and m_2, if the discriminability of m_1 is greater than that of m_2, then we expect a boundary formed between abutting patches of S_{U+m_1} and S_{U-m_1} to appear more vivid than a boundary between patches of S_{U+m_2} and S_{U-m_2}. The first part of the current study relies on our observers to assess the relative strengths of such apparent differences between pairs of textures. We assume that if one of the spaces Λ_{ijk} contains a null point, then that null point will be identical to the maximal modulator m in that space that yields the weakest apparent difference between S_{U+m} vs. S_{U-m}. We call this modulator the minimal salience modulator in Λ_{ijk} and denote it $\hat{m}_{ijk}$. We shall use an adjustment procedure to find estimates of the minimal salience modulators $\hat{m}_{123}$, $\hat{m}_{124}$, $\hat{m}_{134}$ and $\hat{m}_{234}$.

We then need to test whether the estimates we have found are actually null points. To do this we measure performance in discriminating each of them in a preattentive texture segmentation task. If (and only if) all of these estimated minimum salience modulators prove to be null points (i.e., yield chance discrimination performance), will we be able to conclude that human vision has only two texture filters enabling scramble discrimination.

We shall also use an adjustment procedure to find the minimal salience modulator $\hat{m}_{1234}$ in Λ_{1234}. Previous results strongly suggest that $\hat{m}_{1234}$ must be a null point regardless of the number of texture filters differentially sensitive to modulators in Ω. This is clearly true if the number of such texture filters is three or fewer. It is also true, however, if there are more than three; the reasoning runs as follows: Suppose F is a texture filter other than B that is sensitive to modulators in Ω, and let f_F be its impact function. We assume that none of the modulators λ_k, $k = 8, 9, \ldots, 16$ contribute significantly to f_F. The finding of Chubb et al. (1994) that B uniquely predominates in discriminating modulators in Λ_{34567} implies in addition that none of λ_k, $k = 3, 4, \ldots, 7$, contribute significantly to f_F. It follows that f_F can be well-approximated as a linear combination of λ_1 and λ_2. This means that if there are more than two texture filters other than B that are differentially sensitive to modulators in Ω, then the impact functions of all of them are linear combinations of λ_1 and λ_2, implying that they are linearly dependent. In this case, however, the space spanned by the impact functions of all of the texture filters differentially sensitive to modulators in Ω can never be of dimension greater than three from which it follows that $\hat{m}_{1234}$ should be a null point.

3.2. Procedural details

3.2.1. Stimulus configuration

Each display comprised a vertically oriented square wave, each bar of which was a scramble; the histograms of scrambles filling the bars alternated between $U + m$ and $U - m$ for some maximal modulator m subject to manipulation. The square wave had four bars, each bar consisting of 68 rows by 17 columns of texels. Each texel subtended 8.77 min at the viewing distance of 88 cm. The entire display thus subtended approximately 9.8 deg. The texture square wave was presented in the middle of a monitor screen against a background of luminance 60 cd/m^2. The 17 luminances used were approximately equal to $7.5n$ cd/m^2, for $n = 0, 1, \ldots, 16$. More precisely, the lowest luminance was measured at 1 cd/m^2; the highest at 119 cd/m^2.

3.2.2. Luminance calibration

Linearization was achieved using a by-eye procedure in which a regular grid of texture elements containing three intensities lum_{lo}, lum_{hi} and lum_{mid} (half with luminance lum_{mid}, 1/4 with lum_{lo} and 1/4 with lum_{hi}) was made to alternate in a coarse vertical square-wave with texture comprising a checkerboard of texture elements alternating between intensities lum_{lo} and lum_{hi}. (We used a three-luminance checkerboard rather than a uniform field of luminance lum_{mid} in order to control for possible spatial nonlinearities in the display (Klein, Hu, & Carney, 1996; Mulligan & Stone, 1989).) The screen was then viewed from sufficiently far away that the fine granularity of the texture was invisible. At this distance, the square-wave modulating between the two types of texture had a spatial frequency of approximately 6 c/deg. Since the texture itself could not be resolved, the square-wave was visible only if the mean luminance of alternating texture bars was different. Thus, the pixel value val_{mid} that made the square-wave vanish produced a luminance equal to the average of the luminances lum_{lo} and lum_{hi}. We generated a lookup table by reiterating this procedure with different luminances lum_{lo} and lum_{hi} to determine, in each case, the lum_{mid} midway between lum_{lo} and lum_{hi}. Finally, a smooth function was fit to the resulting data, and the 17 luminances used in our stimuli were extracted.

3.2.3. Adjustment methods, subject CC

In the first phase of the experiment, observer CC used a method of adjustment to estimate the minimal salience point $\hat{m}_{1234}$ in Λ_{1234} and also each of the minimal salience points $\hat{m}_{ijk}$ in Λ_{ijk} for $ijk = 123,124,134,$ and 234. In estimating $\hat{m}_{1234}$, bar histograms alternated between $U + m$ and $U-m$, for $m = A_1\lambda_1 + A_2\lambda_2 + A_3\lambda_3 + A_4\lambda_4$, with the weights A_k, $k = 1,2,3,4$ subject to adjustment. Similarly, in estimating $\hat{m}_{ijk}$ for $ijk = 123,124,134,$ or 234, texture bar histograms alternated between $U + m$ and $U-m$, for $m = A_i\lambda_i + A_j\lambda_j + A_k\lambda_k$ with the weights A_i, A_j, A_k subject to adjustment.

During a given adjustment, the texture-defined squarewave pattern was presented continuously, with fresh, independent stimulus realizations presented at the approximate rate of 2.33/s. For each of $ijk = 123$, 124, 134, and 234, preliminary investigations had revealed that the minimal salience point $\hat{m}_{ijk} = A_i\lambda_i + A_j\lambda_j + A_k\lambda_k$ would have the signs of A_i, A_j, and A_k all equal (Otherwise, the resulting texture-defined square wave was quite obvious.). Accordingly, each of A_1, A_2, A_3 and A_4 was confined to the interval $[0, 1]$ in all adjustments.

In estimating $\hat{m}_{ijk}$ ($ijk = 123,124,134,234$) CC was able to control the display in four ways. Pressing the keyboard "a" slightly lowered A_i while preserving the current ratio of A_j to A_k, yet increasing both so as to keep the modulator $m = A_i\lambda_i + A_j\lambda_j + A_k\lambda_k$ maximal. Pressing "b" and "c" produced analogous changes for A_j and A_k, respectively. Finally, pressing "t" swapped the histograms of square wave bars: if the first and third bars had had histogram $U + m$ and the second and fourth $U - m$, then pressing "t" gave the first and third bars histogram $U - m$ and the second and fourth bars $U + m$. When CC was satisfied that the texture square wave was minimally salient, he entered "O", and the three values A_i, A_j and A_k were recorded.

The procedure for estimating $\hat{m}_{1234}$ was exactly the same, except that in addition to being able to press "a", "b" and "c" to decrease A_1, A_2, and A_3 in $m = A_1\lambda_1 + A_2\lambda_2 + A_3\lambda_3 + A_4\lambda_4$, CC could also press "d" to decrease A_4 while preserving the relative proportions of A_1, A_2, and A_3.

CC obtained 16 estimates of $\hat{m}_{1234}$ and also of each of $\hat{m}_{ijk}$ for $ijk = 123, 124, 134, 234$.

3.2.4. Adjustment methods, subjects DB, JN

A modified procedure was then used to obtain estimates of $\hat{m}_{1234}$ and of $\hat{m}_{ijk}$ ($ijk = 123, 124$, and 134) for observers JN and DB, two of the other authors, both of whom understood the hypotheses under investigation (Estimates of $\hat{m}_{234}$ were not derived for either of these observers because, as described below, CC was found to be less sensitive to $\hat{m}_{234}$ than to any of $\hat{m}_{123}$, $\hat{m}_{124}$, or $\hat{m}_{134}$.).

In this modified procedure, observers viewed a uniform gray screen. If they pressed "s", two texture-defined squarewaves would be presented in succession, each for one second separated by a one second uniform gray ISI. The bars of the first squarewave alternated between histograms $U + m_1$ vs. $U - m_1$ for some maximal modulator m_1, and the second squarewave alternated between histograms $U + m_2$ vs. $U - m_2$ for a different modulator m_2. For each such pair of stimuli, the observer was allowed to review the two squarewaves as many times as desired (by pressing "s" repeatedly) in order to make the following ternary judgment: If the first squarewave was clearly more salient than the second, the observer pressed "1"; if the second was more salient than the first, the observer pressed "2"; if the two squarewaves were roughly equal in salience, then the observer pressed "e". Observers were instructed to use the "e" key only when they could not resolve all doubt as to which of the two squarewaves was more salient. These judgments proved to be quite easy and reliable.

The observer's responses were used to control a search for the minimal salience point in one of the spaces Λ_{1234} or Λ_{ijk} ($ijk = 123, 124, 134$). Specifically, in each adjustment trial, a line search was first conducted (as described below) to find the minimal salience modulator $\hat{m}_{1i}$ in the 1-d space spanned by λ_1 and λ_i, for i the second index in the subscript. Then an additional line search was conducted to find the minimal salience modulator $\hat{m}_{1ij}$ in the 1-d space spanned by $\hat{m}_{1i}$ and λ_j for j the third index in the space subscript. If the target space was Λ_{1ij}, then $\hat{m}_{1ij}$ was taken as an estimate of the minimal salience point. If the target space was Λ_{1234}, a final line search was conducted to find the minimal salience point $\hat{m}_{1234}$ in the space spanned by $\hat{m}_{123}$ and λ_4.

Each line search was controlled as follows. At each stage in the search, the participant is asked to compare the salience of two maximal modulators q and r. At the start of the line search, q and r are initiallized so that the minimal salience point m can be safely assumed to lie between them in the following sense: for some α between *0* and *1*, $m = b(\alpha q + (1 - \alpha)r)$, where the scaler b is chosen to make m maximal. Suppose the participant presses a key indicating that q is clearly more salient than r. We take this to mean that the minimal salience point is closer to r than to q. In this case, we replace q by $q_{\text{new}} = b(q + r)$, where scaler b makes q_{new} a maximal modulator, and we leave r unchanged. Thus, q_{new} is pushed to the point in the search locus midway between r and q. If the minimal salience point is indeed closer to r than q, then it will remain between r and q_{new}. If the participant presses "e" (indicating that q and r are similar in salience) then we assume the miminal salience point is closer to the midpoint between q and r than it is to either q or r. Thus we keep the minimal salience point between q_{new} and r_{new} by setting $q_{\text{new}} = b(3q + r)$ and $r_{\text{new}} = c(3r + q)$, where scalers b and c make q_{new} and r_{new} maximal modulators (this pushes q_{new} (r_{new}) half way from q (r) toward the midpoint between q and r).

For JN (DB) this procedure was used to obtain 22 (15) estimates of $\hat{m}_{1234}$ and 21 (15) estimates each of $\hat{m}_{ijk}$ for $ijk = 123, 124$, and 134.

3.2.5. Null point assessment

Each observer was then tested in a texture discrimination task to assess his sensitivity to $m = \hat{m}_{ijk}$, $ijk = 123, 124, 134$, and to $m = \hat{m}_{1234}$. In addition, CC was tested with $m = \hat{m}_{234}$. All conditions were presented in a mixed design. Each block comprised *20* trials from each condition (*10* trials in which the target patch had histogram $U + m$ and the background had histogram $U - m$, and 10 in which target and patch histograms were reversed). Observers completed five blocks for a total of 100 trials per condition. On a given trial, the observer fixated a cue spot centered in a mean gray field, and initiated immediate stimulus presentation with a button-press. The stimulus was presented for 200 ms, and was followed by a uniform gray field, with the cue spot in the center. As in the adjustment phase, each texel subtended 8.77 min at the viewing distance of 88 cm, and the entire display thus subtended approximately 9.8 deg. On a given trial the stimulus conformed to one of the four patterns shown in Fig. 5, and the observer had to indicate which pattern had been displayed. The subject had to press one of the "right arrow," "left arrow," "up arrow," or "down arrow" keys to indicate the location of the target bar relative to fixation.

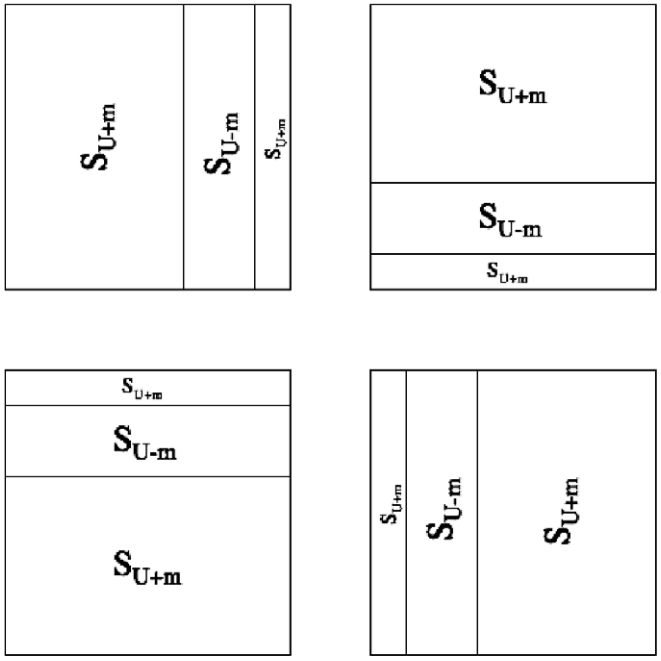

Fig. 5. Stimulus configurations in the 4AFC detection task. On a given trial in the detection experiment, the stimulus conformed to one of the four patterns shown, and the observer had to indicate which pattern had been displayed by pressing one of the "right arrow", "left arrow", "up arrow", or "down arrow" keys to indicate the location of the target bar (the bar containing S_{U-m}) relative to fixation.

4. Results

The results are shown in Fig. 6. The four plots on the left show the estimated minimal salience modulators $\hat{m}_{123}$, $\hat{m}_{124}$, $\hat{m}_{134}$ and $\hat{m}_{1234}$ for each of CC (in blue), DB (in green) and JN (in red). In addition, $\hat{m}_{234}$ is plotted for CC in the lower right of Fig. 6. The curves for the three observers superimpose quite closely for each of $\hat{m}_{123}$, $\hat{m}_{134}$ and $\hat{m}_{1234}$. However, some slight variability is evident in the estimates of $\hat{m}_{124}$. The three numbers shown at the top of a given subplot of Fig. 6 give the number of trials out of 100 in which the three observers (CC, DB and JN, from left to right) succeeded at discriminating the plotted modulator in the 4AFC task. Starred results are significantly higher than expected under the null hypothesis that the actual probability of a correct response is 0.25 (in all cases, $p < 0.002$). p-values for unstarred results are all greater than *0.05*. Specifically, the p-values for 29, 30, 31 and 33 are 0.28, 0.21, 0.15 and 0.07. We conclude that $\hat{m}_{123}$ is not a null point for either of CC or JN and that

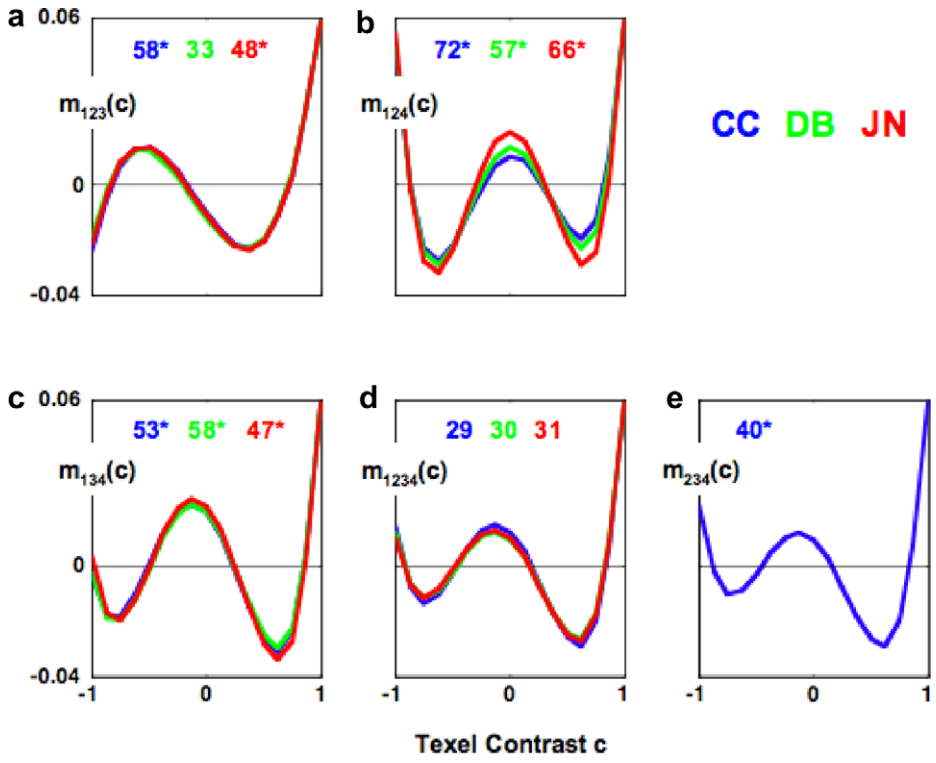

Fig. 6. Minimal salience modulators derived from the 4AFC texture discrimination experiment. (a–d) The estimated minimal salience modulators $\hat{m}_{123}$, $\hat{m}_{124}$, $\hat{m}_{134}$ and $\hat{m}_{1234}$ for each of CC (in blue), DB (in green) and JN (in red). In (e) $\hat{m}_{234}$ is plotted for CC. The three numbers shown at the top of a panel give the number of trials out of 100 in which CC, DB and JN (from left to right) succeeded at discriminating the plotted modulator in the 4AFC detection task. Starred results are significantly higher than expected under the null hypothesis that the actual probability of a correct response is 0.25 (in all cases, $p < 0.002$). p-values for unstarred results are all greater than 0.05: for 29, 30, 31 and 33 the p-values are 0.28, 0.21, 0.15 and 0.07. We conclude that $\hat{m}_{123}$ is not a null point for either of CC or JN and that $\hat{m}_{124}$ and $\hat{m}_{134}$ are not null points for any of our observers. On the other hand discrimination of $\hat{m}_{1234}$ is in the vicinity of chance for all observers, suggesting that $\hat{m}_{1234}$ is at least an approximate null point for all observers. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper).

$\hat{m}_{124}$ and $\hat{m}_{134}$ are not null points for any of our observers. On the other hand discrimination of $\hat{m}_{1234}$ is in the vicinity of chance for all observers, suggesting that $\hat{m}_{1234}$ is at least an approximate null point.

The finding that the minimal salience points $\hat{m}_{124}$ and $\hat{m}_{134}$ are not null points for any of our observers implies that human vision embodies more than two texture filters enabling scramble discrimination. However, the failure to reject $\hat{m}_{1234}$ as a null point supports the three-texture-filter theory.

5. Discussion

The finding that $\hat{m}_{124}$ and $\hat{m}_{134}$ were not null points for any of our three observers supports the claim that human vision embodies at least three scramble-sensitive texture filters, F_1, F_2 and F_3 with linearly independent impact functions f_1, f_2 and f_3. On the other hand, as expected from previous results, $\hat{m}_{1234}$ yielded performance not significantly different from chance for all three observers, consistent with its being a null point and supporting the claim that there exist at most three scramble-sensitive texture filters in human vision with linearly independent impact functions. Note that these results with texture are exactly analogous what we would have observed had we done a color discrimination experiment with three and with four primaries. With four primaries, say red, green, blue, and yellow, there would be a null point signaled by greatly different primary mixtures that are absolutely indistinguishable. By contrast, with only three primaries, R, G and B, there are no null points.

What can we say about these three texture filters? First, we shall assume that all of their impact functions can be well-approximated by 7th-order polynomials. Suppose we identify F_3 with the blackshot texture filter (Chubb et al., 2004). This filter has been isolated and functionally characterized. The claim that discrimination of modulators in Λ_{34567} is approximately univariate is supported by demonstrations that modulators in this space all trade off linearly with each other in controlling discrimination performance (for details, see Chubb et al., 1994). If multiple texture filters were being used to make discriminations in Λ_{34567}, then all would have to have impact functions sharing the same projection into Λ_{34567}. Note however, that the projection of an impact function into Λ_{34567} has five free parameters, and for two impact functions

to share the same projection, their five-dimensional parameter vectors must lie on the same line through the origin. As argued in detail by Chubb et al., 2004, this is highly unlikely unless all but one of these projections is 0—i.e., unless only one texture filter is being used to discriminate scrambles in Λ_{34567}. This finding also implies that each of f_1 and f_2 has a near-zero projection into Λ_{34567}, which in turn implies that each of f_1 and f_2 can be well-approximated by a linear combination of λ_1 and λ_2. The current results show, in addition, that each of f_1, f_2 and f_3 is orthogonal to the null point $\hat{m}_{1234}$.

Consider the three functions, g_1, g_2 and g_3 plotted in Fig. 7. For $k = 1,2,3$, g_k is derived by orthogonalizing λ_k with respect to $\hat{m}_{1234}$ (We used CC's $\hat{m}_{1234}$, but the corresponding plots for JN and DB are very similar, as might be expected given the similarity of their estimates of $\hat{m}_{1234}$). These three functions are mutually orthogonal and span the set of modulators in Λ_{1234} that are orthogonal to $\hat{m}_{1234}$. They thus capture the space of all 4th-order polynomial modulators to which human vision is sensitive. Moreover, g_1 and g_2 are fairly well-approximated by linear combinations of λ_1 and λ_2, whereas g_3 abstracts the sharp tuning of the blackshot texture filter to contrasts near -1.0 from its sensitivity to scramble mean and variance. Indeed, as might be expected from the fact that blackshot strongly predominates in enabling discrimination of scrambles in Λ_{34567}, g_3 matches very closely the projection of the blackshot sensitivity function into Λ_{34} (given by the red line). The reader may wonder why it is that g_3 does not more closely resemble the blackshot impact function $f_3 = f_B$ shown in Fig. 4. The answer is that f_3 contains substantial amounts of λ_1 and $-\lambda_2$ (as shown by Chubb et al., 2004), whereas g_3 does not. This is because, by construction, g_3 is orthogonal to the space spanned by g_1 and g_2, and this space is approximately equal to Λ_{12}.

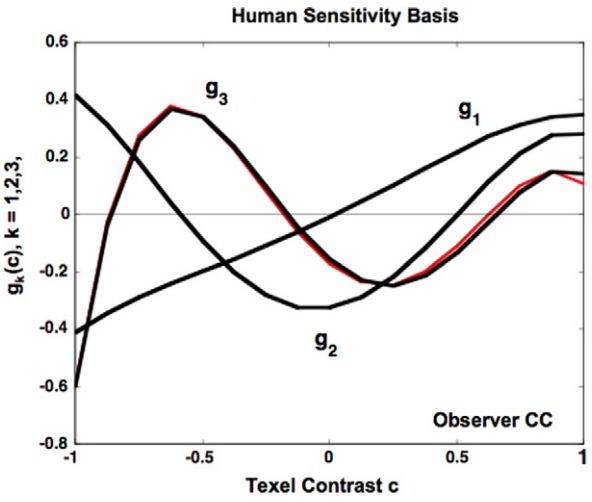

Fig. 7. A basis of the subspace of 4th-order polynomial modulators sensed by human vision. The three functions, g_k, $k = 1, 2, 3$, are derived by orthogonalizing λ_k with respect to $\hat{m}_{1234}$. These three functions are mutually orthogonal and span the set of modulators in Λ_{1234} that are orthogonal to $\hat{m}_{1234}$. They thus capture the space of all 4th-order polynomial modulators to which human vision is sensitive. Moreover, g_1 and g_2 are well-approximated by linear combinations of λ_1 and λ_2, whereas g_3 abstracts the sharp tuning of the blackshot process to contrasts near -1.0 from its sensitivity to scramble mean and variance. Indeed, as might be expected from the fact that blackshot strongly predominates in enabling discrimination of scrambles in Λ_{34567}, g_3 matches very closely the projection of the blackshot sensitivity function into Λ_{34} (given by the red line). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper).

Our current results do not allow us to assert that f_1 and f_2 are identical to the functions g_1 and g_2 in Fig. 7. All we can say with certainty is that f_1 and f_2 span approximately the same space as g_1 and g_2. However, the functions g_1 and g_2 are in accord with theories that posit distinct 1st- and 2nd-order visual processes. First-order processes are typically assumed to be sensitive primarily to the raw Weber contrasts in the visual input. The function g_1 reflects the differential sensitivity to scrambles one would expect such a 1st-order process to provide. Second-order processes are assumed to be sensitive to some rectified transformation (e.g., the square) of Weber contrast. Such a rectifying process might reasonably be expected to provide differential scramble sensitivity in line with the function g_2. Finally, note that since only the blackshot texture filter $f_3 = f_B$ (shown in Fig. 4) is sensitive to modulators outside Λ_{1234}, the three functions, g_1, g_2 and f_3 span the space of all 7th-order polynomial modulators to which human vision is sensitive. Therefore if, as seems likely, human vision has negligible sensitivity to modulations of order higher than 7, then g_1, g_2 and f_3 capture the entire space of human scramble-sensitivity.

5.1. Relation to previous research

We have shown that preattentive scramble discrimination is enabled in human vision by three texture filters, F_1, F_2 and F_3 with linearly independent impact functions. Scramble discrimination based on two or on four filters is clearly excluded. Previous research (Chubb et al., 1994, 2004) shows that F_3, the blackshot texture filter, is very sharply tuned to Weber contrasts near -1.0 and does not discriminate significantly between Weber contrasts between -0.875 and 1.0. The current study implicates two additional texture filters, F_1 and F_2, that collectively provide sensitivity to the 1st- and 2nd-order moments of the scramble distribution.

5.2. Unanswered questions

Note that the current results leave open many questions about the three implicated texture filters. All textures used in the current study are devoid of spatial structure and hence shed no light on the sensitivity of any of these three texture filters to pattern. There is much evidence, both psychophysical and neurophysiological, to suggest that there exist multiple, 1st-order, band-tuned

texture filters (corresponding to classes of simple cells with receptive fields varying in size and orientation). There is also evidence for multiple, band-selective, 2nd-order texture filters (Sutter, Sperling, & Chubb, 1995). These multiple 1st- and 2nd-order texture filters may be the source of the sensitivity conferred by the hypothetical texture filters F_1 and F_2 to the 1st and 2nd moments of the scramble distribution. If so, then in reality each of F_1 and F_2 should be seen not as an individual texture filter, but rather as a class of texture filters with different sensitivities to spatial structure but equivalent sensitivity to spatially random luminance variations. This may also be true of the "blackshot texture filter"; this question awaits investigation.

Appendix A

In this section we give proofs of the following basic facts upon which our methods depend.

If human vision has N texture filters with linearly independent impact functions f_k, $k = 1, 2, \ldots, N$, then (under the assumptions laid out in Section 2.6)

1. Any modulation space Φ of dimension greater than N must contain maximal modulators m orthogonal to the impact function of every texture filter (i.e., null points). (In particular, if the dimension of Φ is $N+1$, and if the projections into Φ of the impact functions f_k are linearly independent, then Φ will contain a unique null point.)
2. There exist modulation spaces of dimension N devoid of null points.

Proof of 1. Let g_k, $k = 1, 2, \ldots, N$, be the projections of the impact functions f_k into Φ. Suppose the g_k's are linearly independent; then the subspace Θ that they span is of dimension N (otherwise Θ is of dimension less than N). The Gram–Schmidt orthogonalization procedure insures that there exists an orthonormal basis b_k, $k = 1, 2, \ldots, N$, of Θ. However, since Φ is of dimension greater than N, it contains a modulator m that is linearly independent with respect to the b_k's. For some nonzero scaler α, let

$$\tilde{m} = \alpha(m - \mathrm{Proj}_{\Theta}(m)) = \alpha\left(m - \sum_{k=1}^{N}(m \cdot b_k)b_i\right). \quad (4)$$

Note that $\tilde{m}$ is (1) nonzero (otherwise, $\tilde{m}$ would be linearly dependent with respect to the b_k's), and (2) orthogonal to all of the b_k's, and hence to all modulators in Θ. Thus, if we choose α to make $\tilde{m}$ maximal, $\tilde{m}$ is a null point. If the dimension of Φ is $N+1$, then the set comprising $\tilde{m}$ along with the b_k's spans Φ, implying that Φ comprises no null points other than $\tilde{m}$.

Proof of 2. Since, as is easily verified, the mean value of any impact function f_k has no influence on $\Delta_k(m)$ for any modulator m, we assume without loss of generality that each of the f_k's has mean *0*; thus, if appropriately scaled, each of the f_k's is a modulator. Hence, the space Φ spanned by the f_k's is itself a modulation space of dimension N that is devoid of null points.

References

Beck, J. (1966). Perceptual grouping produced by changes in orientation and shape. *Science, 154*, 538–540.

Beck, J. (1982). Textural Segmentation. In J. Beck (Ed.), *Organization and representation in perception* (pp. 285–317). Hillsdale, NJ: Erlbaum.

Bergen, J. R., & Adelson, E. H. (1988). Visual texture segmentation based on energy measures. *Journal of the Optical Society of America A, 3*, 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* (pp. 253–271). Cambridge, MA: MIT Press.

Bovik, A. C., Clark, M., & Geisler, W. S. (1990). Multichannel texture analysis using localized spatial filters. *IEEE 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., 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*, 2350–2374.

Chubb, C., & Landy, M. S. (1991). Orthogonal distribution analysis: A new approach to the study of texture perception. In M. S. Landy & J. A. Movshon (Eds.), *Computational models of visual processing* (pp. 291–301). Cambridge, MA: MIT Press.

Chubb, C., Landy, M. S., & Econopouly, J. (2004). A visual mechanism tuned to black. *Vision Research, 44*(27), 3223–3232.

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

Gibson, J.J. (1979), The Ecological approach to visual perception, Erlbaum, London.

Graham, N. (1991). Complex channels, early local nonlinearities, and normalization in texture segregation. In M. S. Landy & J. A. Movshon (Eds.), *Computational models of visual processing* (pp. 273–290). Cambridge, MA: MIT Press.

Graham, N., & Sutter, A. (1998). Spatial summation in simple (Fourier) and complex (non-Fourier) channels in texture segregation. *Vision Research, 38*, 231–257.

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

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

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

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

Klein, S. A., Hu, Q. J., & Carney, T. (1996). The adjacent pixel nonlinearity: Problems and solutions. *Vision Research, 36*, 3167–3181.

Knutsson, H., & Granlund, G. H. (1983). Texture analysis using two-dimensional quadrature filters. In *Proceedings of the IEEE computer society workshop on computer architecture for pattern analysis and image database management* (pp. 206–213). Silver Spring, MD: IEEE Computer Society.

Krantz, D. H. (1975). Color measurement and color theory. I. Representation theorem for Grassmann structures. *Journal of Mathematical Psychology, 12*, 283–303.

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

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

Mulligan, J. B., & Stone, L. S. (1989). Halftoning method for the generation of motion stimuli. *Journal of the Optical Society of America A, 6*, 1217–1227.

Richards, W. (1979). Quantifying sensory channels: Generalizing colorimetry to orientation and texture, touch, and tones. *Sensory Processes, 3*(3), 207–229.

Richards, W., & Riley, M. D. (1977). Texture metamers. *Journal of the Optical Society of America, 67*(10), 1401, abstract.

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.

Sutter, A., Sperling, G., & Chubb, C. (1995). Measuring the spatial frequency selectivity of second order texture mechanisms. *Vision Research, 35*(7), 915–924.