

# Histogram contrast analysis and the visual segregation of IID textures

Charles Chubb

*Department of Psychology, Rutgers University, New Brunswick, New Jersey 08903*

John Econopouly

*Center for Neural Science, New York University, New York, New York 10003*

Michael S. Landy

*Center for Neural Science and Department of Psychology, New York University, New York, New York 10003*

Received November 24, 1992; revised manuscript received February 18, 1994; accepted March 16, 1994

A new psychophysical methodology is introduced, histogram contrast analysis, that allows one to measure stimulus transformations,  $f$ , used by the visual system to draw distinctions between different image regions. The method involves the discrimination of images constructed by selecting texture micropatterns randomly and independently (across locations) on the basis of a given micropattern histogram. Different components of  $f$  are measured by use of different component functions to modulate the micropattern histogram until the resulting textures are discriminable. When no discrimination threshold can be obtained for a given modulating component function, a second titration technique may be used to measure the contribution of that component to  $f$ . The method includes several strong tests of its own assumptions. An example is given of the method applied to visual textures composed of small, uniform squares with randomly chosen gray levels. In particular, for a fixed mean gray level  $\mu$  and a fixed gray-level variance  $\sigma^2$ , histogram contrast analysis is used to establish that the class  $S$  of all textures composed of small squares with jointly independent, identically distributed gray levels with mean  $\mu$  and variance  $\sigma^2$  is perceptually elementary in the following sense: there exists a single, real-valued function  $f_S$  of gray level, such that two textures  $I$  and  $J$  in  $S$  are discriminable only if the average value of  $f_S$  applied to the gray levels in  $I$  is significantly different from the average value of  $f_S$  applied to the gray levels in  $J$ . Finally, histogram contrast analysis is used to obtain a seventh-order polynomial approximation of  $f_S$ .

*Key words:* Texture, spatial vision.

## 1. INTRODUCTION

We are constantly assailed by distinctions drawn for us in perceptual space-time preceding any conscious effort on our part. The spatial, temporal, and spatiotemporal boundaries, gradients, and other more complicated modulations of sensed quality that compose the fabric of reality are all there for us beforehand. They are preliminary to, and taken for granted by, consciousness.

In this paper we describe a new psychophysical methodology for studying the computations that underlie such perceptual distinctions: histogram contrast analysis. The theory underlying the technique is developed in Section 2. Our aim in Section 1 is to motivate that discussion. Toward this end we begin by elaborating the shortcomings of the classical method of increment thresholds as a tool for studying perceptual distinctions. We then introduce histogram contrast analysis as a preferable approach and proceed to describe the method informally as it might be applied to studying the mechanism underlying a certain class of visual texture distinctions. In Section 3, after the formal elaboration (in Section 2) of the theory and methods of histogram contrast analysis, we describe an actual application of the technique to the example introduced at the end of Section 1.

**A. Background: The Method of Increment Thresholds**  
Since the publication of Fechner's book in 1860,<sup>1</sup> the pri-

mary psychophysical tool for investigating how perceptual distinctions are drawn has been the method of increment thresholds. In the simplest sort of application of this method, it is presumed that there exists a single function  $m$  that defines the relative impacts on a given sensory system (e.g., the auditory system or the visual system) as a whole of different values of a given physical parameter  $v$  (e.g., the amplitude of a tone or the intensity of a light). It is always assumed that  $m$  is a monotonic function of  $v$ . The subject is asked to distinguish between regions  $A$  and  $B$  of the visual field (separated in time and/or space) differing in the value of  $v$ . It is assumed that the subject is able to draw the distinction only if

$$m(\text{the greater value of } v) - m(\text{the lesser value of } v) \geq \text{threshold} \quad (1)$$

for some constant threshold (i.e., some threshold independent of the values of  $v$  being compared). If this assumption holds, then we can construct the function  $m$  (up to arbitrary additive and multiplicative constants) by varying the lesser value of  $v$  and in each case determining the greater value of  $v$  for which the subject can just barely draw the distinction between the regions differing in  $v$  value (e.g., she or he makes a correct forced choice 75% of the time). In effect, the subject is providing an estimate that is approximately proportional to the derivative of  $m$ .

## B. Method of Increment Thresholds and the Problem of Adaptation

The basic problem in interpreting increment thresholds derives from the fact that the sensory impact of a region of the visual field painted a given value often depends critically on the context of that region (i.e., on the values surrounding the region in space and time). Consider, for instance, the case of visual lightness perception. It has long been recognized<sup>2</sup> that the lightness of a patch of the visual field depends not so much on the absolute luminance of the patch as on the luminance of the patch relative to the average level of luminance to which the eye has adapted. Indeed, it is now well documented that adaptive gain-control mechanisms play a critical role in visual signal transduction, from the retina on up through higher levels of processing (e.g., Refs. 3–5). Above the level of the retina, information about the mean luminance of the display being viewed has largely been removed from the transmitted signal.

Let  $S$  be a stimulus assigning a luminance value to each point  $(x, y)$  in the visual field. It is now generally thought that the signal that survives these early contrast gain-control mechanisms is reasonably well approximated by a contrast modulation function; that is, a function  $C_S$  of the following sort:

$$C_S(x, y) = \frac{S(x, y) - \text{local average}_S(x, y)}{\text{local average}_S(x, y)}, \quad (2)$$

where  $\text{local average}_S(x, y)$  is the average luminance of  $S$  taken over all points in some neighborhood of  $(x, y)$ .

Note that if  $S$  has local average luminance that is constant over space and is equal to the global mean luminance  $\mu$  of  $S$ , then

$$C_S(x, y) = \frac{S(x, y) - \mu}{\mu}. \quad (3)$$

In light of Eq. (3), consider the impact, for higher-level visual processes, of occurrences in  $S$  of luminance  $v$ . If  $S(x, y) = v$  for some  $(x, y)$  in the visual field, then

$$C_S(x, y) = \frac{v - \mu}{\mu}. \quad (4)$$

That is, given adaptation to a patterned field of mean luminance  $\mu$ , the impact of different luminances  $v$  for higher-level visual processing is a linear function of  $v$  with slope  $1/\mu$ . Indeed, precisely such a linear relationship has been observed in the firing rate of cat retinal ganglion cells in response to suprathreshold increments and decrements over a moderate range from a given level of adaptation.<sup>4</sup>

Suppose, however, that one were to assume (mistakenly) that there exists a single function  $m$ , as given in Eq. (1), that directly registers the impact for higher-level visual processes of any luminance  $v$ , regardless of the context in which  $v$  appears (e.g., regardless of adaptation level). This was, in fact, the assumption underlying Fechner's method.

To measure the function  $m$ , which maps luminance into just-noticeable differences (JND's) of elicited sensation,

one might use the method of increment thresholds as follows: set  $v_0$  equal to a fixed luminance level; then, for  $i = 1, 2, \dots$ , experimentally determine the luminance  $v_i$  for which a subject can just detect a small spot of luminance  $v_i$  on a background of luminance  $v_{i-1}$ .

If reality is reflected by Eq. (4) and adaptation level is set primarily by the background luminance, then for the subject to just perceive a small disk of luminance  $w > v$  appearing on a uniform background of luminance  $v$  we must have

$$\frac{w - v}{v} = \text{threshold}, \quad (5)$$

or Weber's law. From this empirical result and our untested and mistaken assumption of a single function  $m$  governing threshold regardless of context, we conclude that  $m$  is logarithmic.

Here is the main point: we might have thought that the JND was a unit of measurement of "nerve states" (in the sense of Müller<sup>6</sup>); we might have thought that, by increasing stimulus intensity at some point in the visual field so as to double the number of JND's above threshold of that point, we could double the firing rate of some set of neurons monitoring intensity at that point. This was the presumption driving the construction of the mapping  $m$  with use of the method of increment thresholds. However, if the true state of nature is given by Eq. (2), then the JND is a theoretically meaningless unit of measurement because it does not predict how the visual system codes stimulus strength.

It has become clear over the past two decades that it does not make sense to try to summarize with a function such as  $m$  the effect that a point of light of a given intensity has on the visual system as a whole. There are many parallel processes executed by the visual system in which a point of light plays a variety of roles.

However, as we shall argue, it is often possible to determine the effect of a point of light for a visual subsystem that draws a specific sort of distinction in the visual field. Indeed, in Section 2 we describe a method, histogram contrast analysis, that is specifically tailored to this task. In Section 3 we present the results of experiments in which histogram contrast analysis is applied in studying the mechanism that draws distinctions between textures in which the visual field is tiled with small uniform squares of different intensities. We shall show, for a given class of such textures, that texture segregation performance may be modeled with use of a single function that codes intensity. Histogram contrast analysis provides strong empirical tests that justify this claim (Subsections 2.D and 2.E).

## C. One Solution: IID Textures

How can we investigate the way in which the visual system codes light intensity, given that it changes its characteristics as it changes its state of adaptation? One approach (the approach taken here) is to use stimuli that, although they are composed of multiple gray levels, induce a constant state of adaptation in the visual system. In this paper we define and explore a stimulus set that enables this to occur: textures composed of jointly independent, identically distributed, random microcomponents (IID textures).

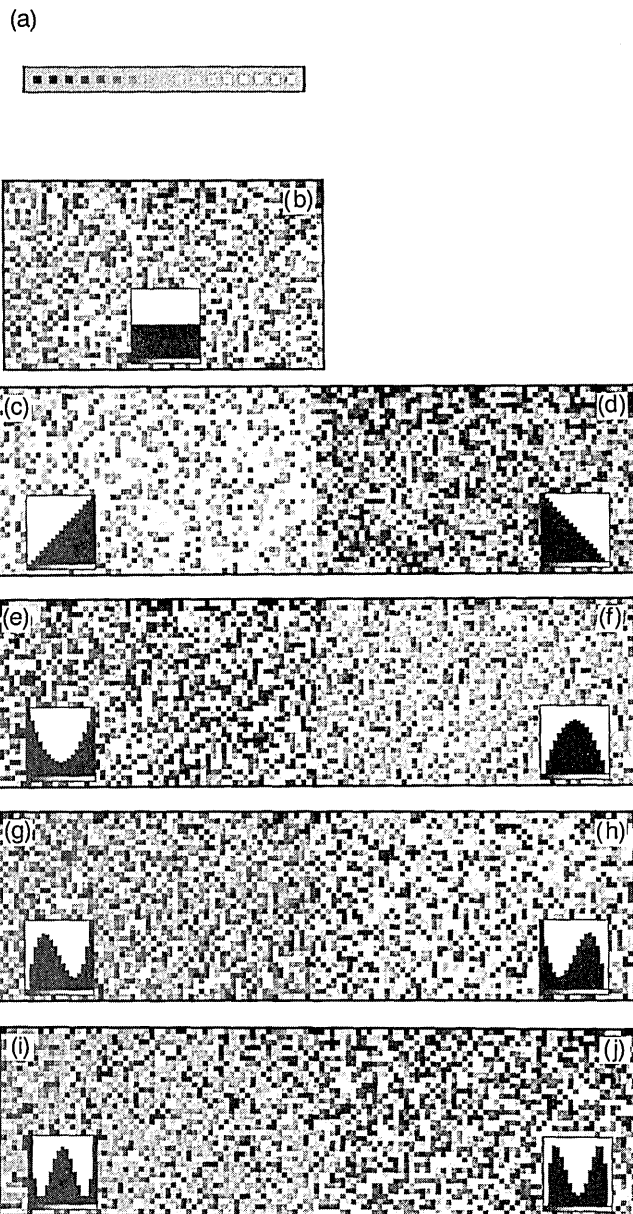


Fig. 1. IID textures for a micropattern set  $\Gamma$  of small gray squares of various luminances. (a) the micropattern set  $\Gamma$ ; (b)  $I_U$  for  $U$ , the uniform distribution (in this and subsequent panels the texel distribution is shown in an inset); (c)  $I_{U+\lambda_1}$ ; (d)  $I_{U-\lambda_1}$ ; (e)  $I_{U+\lambda_2}$ ; (f)  $I_{U-\lambda_2}$ ; (g)  $I_{U+\lambda_3}$ ; (h)  $I_{U-\lambda_3}$ ; (i)  $I_{U+\lambda_4}$ ; (j)  $I_{U-\lambda_4}$ . The set of modulators  $\lambda_i$  is defined in Subsection 3.A. The visible distinctions between (g) and (h) and between (i) and (j) are for texture pairs with equal mean and variance. Hence such distinctions cannot be drawn simply on the basis of average luminance or contrast.

Consider, for example, the textures illustrated in Fig. 1. Each texture consists of a grid of squares of uniform intensity [chosen from among those in Fig. 1(a)]. For a given texture patch the intensity of each square is chosen randomly, and different squares are colored independently. (The square is called a texture element, or texel.) Thus the generation of such a patch of texture is governed entirely by the probability distribution that determines the likelihood of each possible intensity's being chosen. For example, Fig. 1(b) shows a patch of texture in which the uniform distribution was used to generate the texture (so that all gray levels were equally

likely). (In this and subsequent stimulus figures the texel probability distribution is shown in an inset.) In Figs. 1(c) and 1(d) two linear distributions were used: the left-hand texture favors brighter texels, and the right-hand texture favors darker texels. The difference in the mean luminances of these two textures elicits a clear lightness boundary between them. In Figs. 1(e) and 1(f) two parabolic distributions were used: the left-hand texture favors white and black texels, and the right-hand texture favors gray texels. These two textures have the same mean luminance, and hence they have approximately equal average lightness. However, the left-hand texture has a higher variance than the right-hand texture, generating a clear contrast boundary.

In the texture pairs just considered, segregation was possible because of a difference in mean luminance or mean contrast between the two textures. Such differences are precisely the sorts that one might expect to engender a change in state of adaptation by means of a luminance or contrast gain-control mechanism.<sup>5</sup> However, most current theories of adaptation predict that adaptation will remain constant across a set of stimuli that have equal mean (average luminance) and variance (contrast power).

With this in mind, consider the texture pairs in Figs. 1(g) and 1(h). The two texel gray-level distributions here have equal mean and variance, and thus the two textures have approximately equal average luminance and contrast. Thus it is reasonable to expect that the state of adaptation will be constant across these two patterns. The two patterns also have identical spatial structure (a grid of squares). Yet the two patterns segregate reasonably easily. A second example of this sort is given in Figs. 1(i) and 1(j). We shall show how one may use such texture pairs in psychophysical tasks to learn about early nonlinearities in the visual system. Although the techniques introduced in this paper are quite general, for concreteness we shall continue to focus on the segregation of IID textures composed of uniform gray squares (as in Fig. 1) as our canonical example, and in Section 3 we shall demonstrate an application of the method by using such textures.

#### D. Texture Segregation Models

The literature on texture segregation has been concerned primarily with differences in spatial structure on either side of a texture boundary that lead to easy, preattentive texture segregation (see, e.g., Ref. 7). There was recently proposed a class of models of texture segregation that, although different in details, share the same basic form (Refs. 8–16; see also Refs. 17 and 18). We term this the back-pocket model of texture segregation, because texture perception researchers routinely pull it out of their back pockets to make sense of new instances of texture segregation. These models take for granted that visual processing begins by the application to the visual input of a set of linear filters that are spatially localized, bandpass in spatial frequency, and orientation tuned, as has been found both psychophysically and physiologically (Refs. 19 and 20). Back-pocket models differ in how these linear filter outputs are handled, but they all share a common structure. Processing begins with a measurement phase and proceeds to a surveying phase, as follows:

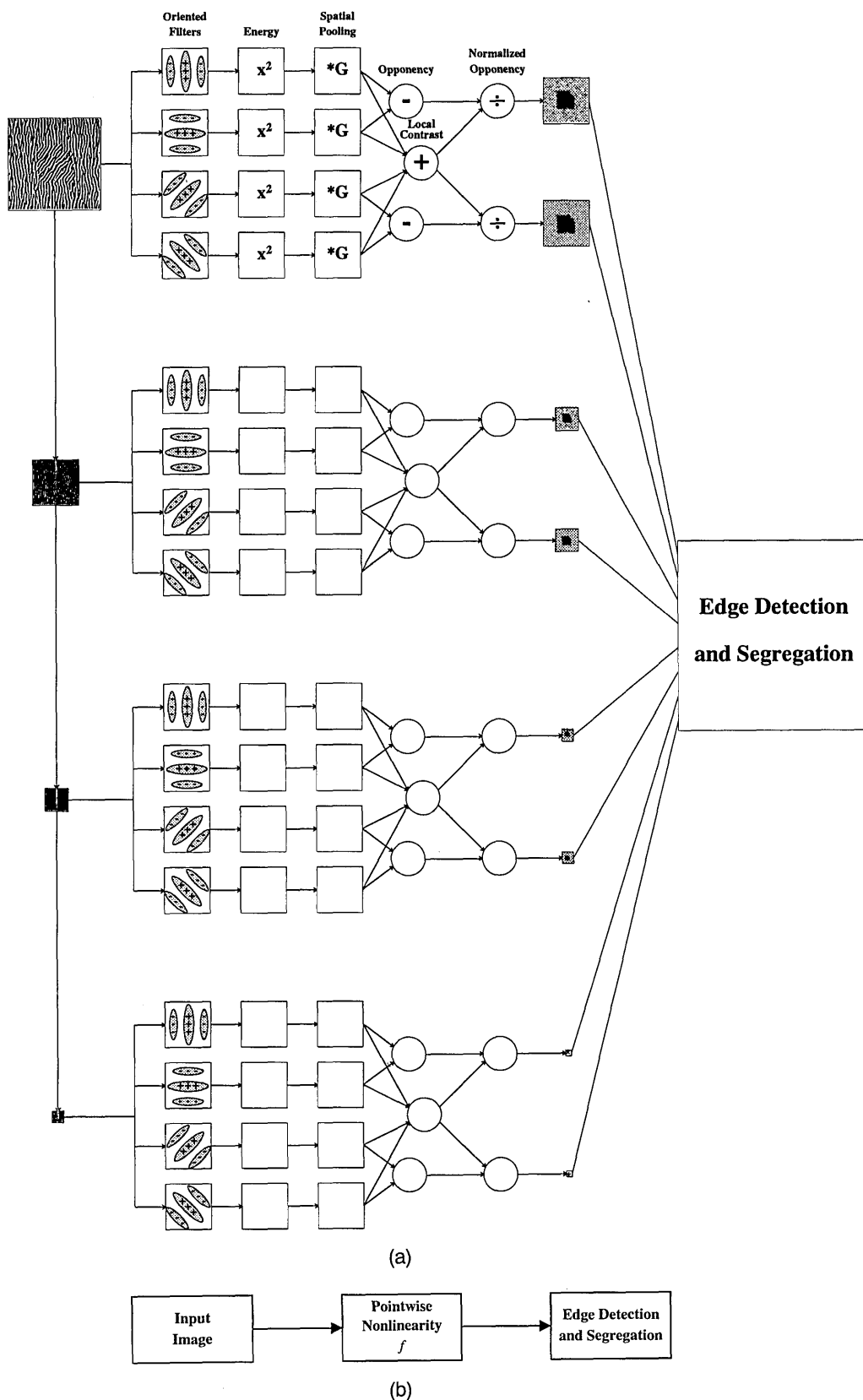


Fig. 2. (a) Landy-Bergen<sup>10,14</sup> model, which is an example of the back-pocket model of texture segregation. The input image is processed by a number of linear filters varying in dominant orientation and spatial frequency. Within a given scale each oriented filter output is squared (computing a texture energy), orthogonal orientations are put in opponency, and these opponent responses are divided by the sum of all four oriented energies as a form of contrast gain control. The intent is to convert an input textural difference into a difference in intensity of response. Finally, an edge detector is applied to localize such intensity edges. (b) For the textures illustrated in Figs. 1(g)–1(j), this much simpler model may suffice. It corresponds to a single channel of the model in (a) with no up-front linear filtering and a single pointwise nonlinearity.

### 1. In the measurement phase

a. Each member of some set of linear filters is applied to the retinal input, producing a set of spatial functions (or neural images), one per filter.

b. The resulting set of spatial functions is collectively subjected to some sort of spatially local nonlinear processing. In the simplest variants of the back-pocket model, this processing consists merely of the application of a function (e.g., a square or absolute value) to the output of each filter point by point in space.

The entire set of resulting spatial output functions is then passed to

2. The surveying phase, in which the visual field is partitioned into regions within which the measurement-phase output functions are all relatively constant in value and between which (at least some of) these output functions diverge in value.

An example is the model proposed by Landy and Bergen<sup>10,14</sup> as illustrated in Fig. 2(a). This model begins by application of a set of linear filters at various spatial scales and orientations. The second, nonlinear stage (stage b of the measurement phase) involves several computations, including a pointwise nonlinearity (a square or texture energy computation), an opponency between orthogonal orientations, and a contrast normalization. In the surveying phase, one looks for changes in the measurement-phase output between different regions by applying an edge-enhancement mechanism.

How would a back-pocket model such as that of Landy and Bergen fare with the IID textures of Fig. 1? Here, there is no difference in spatial structure between the paired texture regions. If a filter has a receptive field that is large relative to the squares that make up the textures, then its output will be a weighted sum of these randomly chosen texels. Thus the central limit theorem implies that the output of this receptive field will be Gaussian and hence a function only of the mean and variance of the texel distribution (see Subsection 3.A below for further discussion). A measurement phase channel of the texture model with use of that filter would be capable of segregating the texture pairs of Figs. 1(c)–1(f) but would be blind to the texture differences in Figs. 1(g)–1(j), because those texel distributions are constrained to have equal mean and variance on either side of the boundary.

The ability to segment IID textures such as those in Figs. 1(g)–1(j) implies that there exists a measurement-phase channel other than those pictured in Fig. 2(a). Figure 2(b) shows a model of a measurement-phase channel that could account for this behavior. This channel differs from standard back-pocket model channels in that it uses no linear filter as the first stage in its measurement transformation; rather, it applies a nonlinear function directly to the retinal input, point by point in space. The technique described in this paper allows one to determine whether the visual system applies a unique channel of this form. Further, once it is verified that only one such measurement-phase channel is being applied, the technique can be used to measure the form of the nonlinear function being used in this channel.

### E. Histogram Contrast Thresholds

To investigate the model shown in Fig. 2(b) we must first determine a psychophysical measure of the strength of texture segregation in texture pairs such as those in Figs. 1(g)–1(j). This paper introduces such a measure: the histogram contrast threshold. The basic method is illustrated in Fig. 3. This figure shows a sequence of texture pairs involving a texel distribution for the left-hand texture and the inverted distribution for the right-hand texture. The functional form of the distributions is identical in all textures; it is the uniform (flat) distribution modulated by a (third-order) polynomial chosen so as to keep the mean and variance of resulting texture gray levels constant and equal to that of the uniform distribution [Fig. 1(b)]. However, the amplitude of this modulating function is varied across texture pairs. As this amplitude decreases, the paired textures increasingly resemble the texture generated by the uniform distribution and thus increasingly resemble one another. Hence, lower amplitude of modulation results in poorer texture segregation.

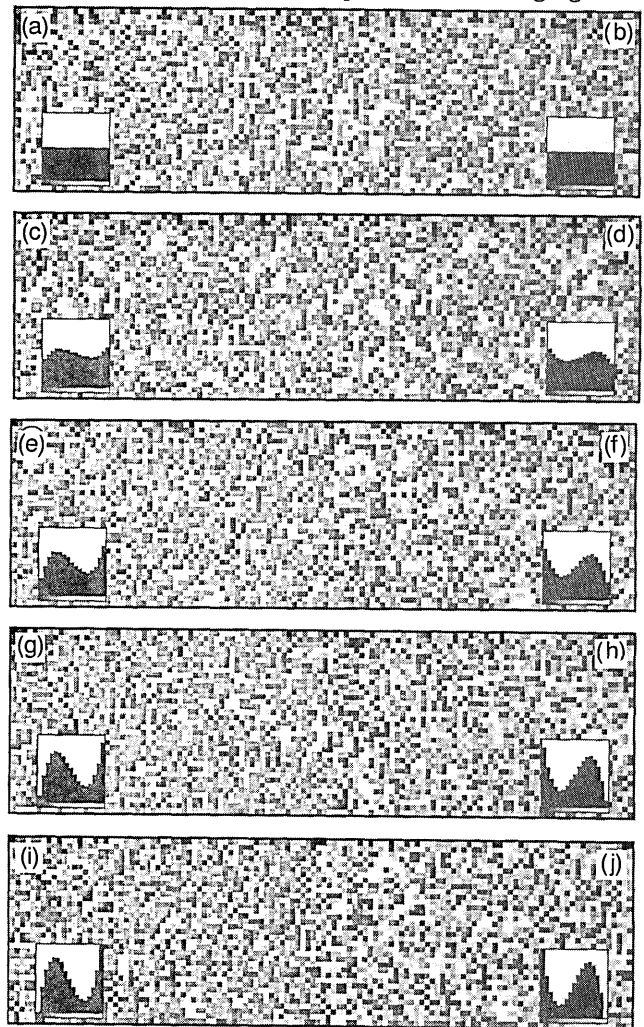


Fig. 3. Histogram contrast threshold. The figure illustrates a series of IID texture pairs with increasing amplitude of modulation of the probability distributions used to generate the textures. Here a third-order polynomial  $\lambda_3$  was used as the modulator of the uniform distribution  $U$  [as in Figs. 1(g) and 1(h)]. The histogram contrast threshold is the amount of modulation that results in threshold segregation performance (e.g., 75% correct on a forced-choice discrimination task).

The psychophysical technique employed is the obvious one: a contrast threshold. However, it is not the stimulus contrast that we vary. It is rather the contrast in the histogram (or probability distribution) used to generate the texture that we vary. By applying typical psychophysical paradigms (forced-choice, staircase procedures), we can determine threshold histogram contrast. It will be shown (Subsections 2.C and 2.D below) that the result of this procedure gives us a means of estimating the form of the unknown nonlinearity shown in Fig. 2(b).

## F. Summary and Overview

The bulk of this paper concerns the mathematical basis of our techniques. Considerable formalism is required for justifying the statements made in the last two sections. However, the broad outline of the technique is as we have just described it: we (1) define a set of functions (to be used to modulate IID texture histograms) sufficient to generate boundaries between textures with equal mean and variance, (2) measure the histogram contrast threshold for each such function, and (3) use these measurements in conjunction with various side measurements and controls to estimate the form of the pointwise nonlinearity in the texture channel used to segregate these patterns. The method includes a number of checks to verify that the simple, single-channel model holds for this class of stimuli.

The actual technique is, in fact, quite general. It can be applied to other sorts of visual stimuli as well as to nonvisual stimuli. It can also be used in situations in which discrimination between textures involves multiple channels. Moreover, we shall introduce additional psychophysical techniques with greater sensitivity than that of the basic histogram contrast threshold. Nonetheless, it will help to keep the above example of IID gray-level textures in mind as you proceed.

Section 2 outlines the technique in broad mathematical generality. Section 3 applies the technique to the case of IID gray-level textures and shows both that the nonlinearity applied to input luminance may be measured precisely by the technique and that the cross checks generally succeed for this class of stimuli. For a first reading you may wish to scan Section 2, concentrate on the application of the method in Section 3, and later return to wade through the notation and proofs.

## 2. DETAILED DEVELOPMENT OF HISTOGRAM CONTRAST ANALYSIS

The purpose of the following prefatory remarks is to lead the reader informally from the view of texture perception expressed by the back-pocket model to the slightly more general perspective that we formalize in this section.

It is important to realize that the back-pocket model works because each measurement phase channel is, in its own characteristic way, differentially sensitive to different textures. By this we mean that the average output value over space of a given channel varies from texture to texture. Consider, for instance, a measurement-phase channel whose linear filter is tuned to high frequency, vertically oriented texture and whose nonlinear stage consists of squaring the filter output point by point in space.

The average response over space of just the linear portion of this channel will be 0 to any texture whatever. Thus, by itself, the linear filter fails to be differentially sensitive to different textures. However, by applying the squaring nonlinearity to the filter output, we derive a channel that is differentially sensitive to high-frequency vertical texture; the average response over space of this nonlinear channel is greater for high-frequency vertical texture than for lower-frequency or nonvertical texture.

This differential sensitivity permits back-pocket-model channels to play a role in texture discrimination analogous to the roles played by different cone classes in light discrimination. Each cone class is differentially sensitive to different lights, and two lights are discriminable only if at least one cone class has a different average response to one light than to the other. In precisely the same way, under the back-pocket model two textures are discriminable only if at least one measurement-phase channel has a different average response to one texture than to the other.

There is nothing sacrosanct about the back-pocket model assumption that measurement-phase channels consist of linear filters followed by pointwise nonlinearities. This is merely a convenient way of constructing an image transformation that is differentially sensitive to different textures. Accordingly, we drop all assumptions about the internal workings of measurement-phase channels and introduce the term (texture) measure to refer to any image transformation whose average response over space is different for different textures.

We suppose that the visual system routinely applies some fixed set of texture measures to the retinal input and that two textures are discriminable only if they elicit different average responses from at least one of these texture measures.

In light of these reflections, consider some set  $\Omega$  of micropatterns, all of the same rectangular shape and size, from which we shall construct various IID textures. In Fig. 1 the micropatterns used are all uniform squares of different intensities. However, the micropatterns contained in  $\Omega$  need not be uniform patches; they can be arbitrarily patterned. A given texture measure is likely to be differentially sensitive to the different elements of  $\Omega$  in a sense akin to that in which a given cone class is differentially sensitive to quanta of different wavelengths. Just as quanta of different wavelengths may elicit different levels of excitation in cones of a given class, different micropatterns of  $\Omega$  may elicit different levels of excitation in a given texture measure.

This effect leads to the notion of the differential sensitivity function  $d_T$  of a given texture measure  $T$ . For any micropattern  $\omega \in \Omega$ ,  $d_T(\omega)$  gives the average level of excitation induced in  $T$  by an occurrence of  $\omega$  (in some IID texture of  $\Omega$ ). Thus the function  $d_T$  is analogous to the absorption spectrum of the photopigment of a given cone class. The absorption spectrum gives the sensitivity of the cone class to quanta of different wavelengths;  $d_T$  gives the sensitivity of  $T$  to different elements of  $\Omega$ .

The methods that we develop will allow us to investigate hypotheses concerning the differential sensitivity functions of texture measures responsible for drawing distinctions between IID textures. For example, for any mi-

croppattern set  $\Omega$  we can test the hypothesis that the visual system has only one texture measure  $T$  that is differentially sensitive to the micropatterns in  $\Omega$ . Moreover, if this hypothesis is supported, our methods will allow us to measure  $d_T$  with arbitrary precision. A full description of the general class of hypotheses that can be addressed by histogram contrast analysis must await the discussion (in Subsection 2.A.8 below) of spaces of IID textures.

The outline of this section is as follows. Subsection 2.A sets out the definitions and notation needed for the basic theory of histogram contrast analysis. The development of the theory begins in Subsection 2.B with the definitions surrounding the notion of a perceptually elementary space of IID textures and concludes in Subsection 2.C with the results from which the usefulness of the histogram contrast threshold derives. There follows an elaboration of the principal methods used in histogram contrast analysis: cancellation tests (Subsection 2.D) and the method of titration (Subsection 2.E).

### A. Preliminaries

This section consists primarily of definitions required for developing the theory. A formal description is given of an IID texture. The notion of a texture measure  $T$  whose response to an IID texture is a linear combination of its responses to the individual texels of that texture leads to a formal definition of the differential sensitivity function of  $T$ .

For definiteness, a simple computation is assumed for the surveying stage of the back-pocket model; specifically, we assume that texture boundaries are detected by the application of linear edge detectors to texture measure output.

Finally, the linear algebraic notion of a space of IID textures is introduced (formalizing the kinds of textures shown in Fig. 1).

#### 1. Notation Used for Stimuli and Stimulus Transformations

We use the term stimulus to refer to any function  $I: \mathbb{R}^2 \rightarrow \mathbb{R}$ , where the plane  $\mathbb{R}^2$  represents visual space and  $I(x, y)$  denotes the value assigned by  $I$  to the point  $(x, y)$  in space. A transformation  $T$  is a function mapping the set of all stimuli into itself. For any stimulus  $I$ , the stimulus that results from applying  $T$  to  $I$  is denoted  $T(I)$ , and the value assumed by  $T(I)$  at the point  $(x, y) \in \mathbb{R}^2$  is denoted  $[T(I)](x, y)$ . We shall refer to some transformations  $T$  as texture measures. Any such  $T$  can be assumed to be both spatially local and shift invariant.

#### 2. Notation Associated with IID Textures

Let  $\Omega$  be a set of spatially local visual stimuli. For instance, we can let  $\Omega$  be a set of small uniform squares of various intensities (as in Fig. 1), or, more generally, we can take  $\Omega$  to be any set of micropatterns. We now partition the visual field into regions (texels) the size of the micropatterns in  $\Omega$ . A texture is generated by a micropattern of  $\Omega$  being painted into each texel. We focus exclusively on visual textures whose texels are painted with jointly independent, identically distributed, random elements of  $\Omega$ . Hence we call any such texture an IID texture of  $\Omega$ . Any IID texture of  $\Omega$  is completely speci-

fied by the probability distribution  $p$  used to draw each texel from  $\Omega$ . Accordingly, we call  $p$  the texel distribution of the texture.

For any probability distribution  $p$  on  $\Omega$ , we write  $I_p$  for an IID texture with texel distribution  $p$ . In addition, we shall sometimes write  $X_p$  for an  $\Omega$ -valued random variable with distribution  $p$ . For a given texel  $\tau$  in the visual field, we write  $I_p(\tau)$  for the micropattern painted into  $\tau$  by  $I_p$ . Thus  $I_p(\tau)$  is an  $\Omega$ -valued random variable distributed as  $X_p$ .

Formally,  $I_p$  is a family of jointly independent random variables (each distributed as  $X_p$ ) assigned to the texels tiling the visual field. Thus, for nontrivial distributions  $p$ ,  $I_p$  typically has many different possible realizations  $J$ . Any such  $J$  is simply an image, a real-valued function of the visual field. As a realization of  $I_p$ , however,  $J$  has a probability. This probability is simply the product, over all texels  $\tau$ , of  $p[J(\tau)]$ .

Note that for large patches of the texture  $I_p$  the proportions of different micropatterns occurring in texels (i.e., the histogram of  $I_p$ ) will be approximately given by  $p$ .

#### 3. Definition (of the Cross Correlation of Two Functions)

Let  $f$  and  $g$  be real-valued functions of a finite set  $\Omega$ . Then the cross correlation of  $f$  and  $g$  is defined by

$$f \cdot g = \sum_{\omega \in \Omega} f(\omega)g(\omega). \quad (6)$$

#### 4. Notation Associated with Random Variables

For any real-valued random variable  $Y$  we write  $E[Y]$  for the expectation of  $Y$  and  $\text{var}[Y]$  for the variance of  $Y$ .

In this connection note the following facts. Let  $p$  be a probability distribution on a set  $\Omega$  of micropatterns, and let  $X_p$  be an  $\Omega$ -valued random variable with distribution  $p$ . Then, for any function  $f: \Omega \rightarrow \mathbb{R}$ ,  $f(X_p)$  is a real-valued random variable. Moreover,

$$E[f(X_p)] = \sum_{\omega \in \Omega} f(\omega)p(\omega) = f \cdot p, \quad (7)$$

$$\text{var}[f(X_p)] = E[f^2(X_p)] - E^2[f(X_p)] = f^2 \cdot p - (f \cdot p)^2, \quad (8)$$

where  $f^2(\omega) = [f(\omega)]^2$  for all  $\omega \in \Omega$ .

#### 5. Notation Used for Constant Functions

For any set  $B$  and any  $\alpha \in \mathbb{R}$  we write  $\alpha$  for the function that assigns the value  $\alpha$  to every  $b \in B$ . In particular, we shall have occasion to refer to constant functions of space (in which case the set  $B = \mathbb{R}^2$ , the image plane) as well as to constant functions of a micropattern set  $\Omega$ . For instance, the function  $\mathbf{1}: \Omega \rightarrow \mathbb{R}$  satisfies  $\mathbf{1}(\omega) = 1$  for all  $\omega \in \Omega$ , and likewise the function  $\mathbf{0}: \mathbb{R}^2 \rightarrow \mathbb{R}$  satisfies  $\mathbf{0}(x, y) = 0$  for all points  $(x, y)$  in space.

#### 6. Surveying Phase: Assumptions and Notation

As discussed at the beginning of Section 2, a texture measure  $T$  that is useful in segregating a given texture will convert a textural difference in its input into an intensity difference in its output. Thus the surveying phase need only compare the output of  $T$  in two abutting image regions and respond if the regions differ sufficiently. Although there are elaborate ways to accomplish this goal,

the straightforward approach is to use a linear edge detector. For the sake of definiteness we assume this model.

When we think of a luminance-edge detector, we think of a neuron whose receptive field has elongated, adjacent excitatory and inhibitory regions that allow the neuron to detect spatial differences in retinal excitation. We suppose that the surveying-phase neurons involved in detecting texture edges on the basis of the output of the texture measure  $T$  work similarly; however, instead of detecting spatial differences in retinal excitation, these mechanisms detect spatial differences in excitation within the array of neurons composing  $T$ . We can conveniently capture the general sort of computation performed by such a linear  $T$ -edge detector in terms of the following notation.

*Definition (of an additive pooler).* For any image  $J$  and any nonnegative, real-valued function  $g$  of the visual field, let

$$\langle J \rangle_g = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) J(x, y) dx dy. \quad (9)$$

Because  $g$  is nonnegative,  $\langle \rangle_g$  can be viewed as aggregating contributions from  $J$  with varying degrees of emphasis from different points in the visual field. Accordingly, we call  $\langle \rangle_g$  an additive pooler.

In this connection we call

$$\langle 1 \rangle_g = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) dx dy \quad (10)$$

the gain of  $\langle \rangle_g$  and

$$\langle g \rangle_g = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y)^2 dx dy \quad (11)$$

the energy of  $\langle \rangle_g$ .

Consider, then, a  $T$ -edge detecting neuron, TED, whose excitatory lobe pools its input according to the spatial weighting function  $g_{\text{excit}}$  and whose inhibitory lobe pools its input according to the spatial weighting function  $g_{\text{inhib}}$ . We can now represent the response of this neuron to an input stimulus  $J$  by writing simply

$$\text{TED}(J) = \langle T(J) \rangle_{g_{\text{excit}}} - \langle T(J) \rangle_{g_{\text{inhib}}}. \quad (12)$$

### 7. Additive Texture Measures and Differential Sensitivity Functions

We call any texture measure  $T$  additive with respect to a given micropattern set  $\Omega$  if  $T$ 's response to any IID texture of  $\Omega$  is equal to the sum of  $T$ 's responses to the individual texels of this texture. For any such  $T$  (additive with respect to  $\Omega$ ) we can predict  $T$ 's response to any IID texture of  $\Omega$  if we know how it responds to the individual micropatterns in  $\Omega$ . Moreover, because (by assumption) the surveying phase pools the output of  $T$  additively over multiple texels, the detailed spatial form of  $T$ 's response to any given micropattern  $\omega \in \Omega$  is irrelevant to the distinction-drawing process. The only thing that matters, the only thing that we need to know, is the sum over space of  $T$ 's response to  $\omega$ . Thus  $T$  might as well be simply replacing each micropattern  $\omega$  in any of these IID textures with a uniformly painted texel of intensity  $d_T(\omega)/\text{texel}$ , for

$$d_T(\omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [T(\omega_{\text{alone}})](x, y) dx dy \quad (\omega \in \Omega), \quad (13)$$

where  $\omega_{\text{alone}}$  is the image containing a single occurrence of  $\omega$  in an otherwise uniform field of mean luminance. We call the function  $d_T: \Omega \rightarrow \mathbb{R}$  the differential sensitivity function of  $T$  with respect to  $\Omega$ .

In light of these observations, consider TED the  $T$ -edge detecting neuron whose response to an arbitrary stimulus  $J$  is given by Eq. (12). If  $J$  is an IID texture of  $\Omega$ , and if  $T$  is additive with respect to  $\Omega$ , then TED's response to  $J$  is dictated by  $d_T$  [Eq. (13)]. Each texel  $\tau$  of  $J$  exerts a total impact in this response equal to  $d_T[J(\tau)]$ . We express this fact by rewriting Eq. (12) as

$$\text{TED}(J) = \langle d_T \circ J \rangle_{g_{\text{excit}}} - \langle d_T \circ J \rangle_{g_{\text{inhib}}}, \quad (14)$$

where we use the notation  $d_T \circ J$  to represent the image that results from painting each texel  $\tau$  uniformly with the intensity  $d_T[J(\tau)]$ .

### 8. Background Linear Algebra and IID Texture Spaces

Histogram contrast analysis makes substantial use of linear algebra. In particular, the modulators that we use (e.g., those shown in Fig. 1) will make up an orthogonal set of functions. They will allow us to determine a linear decomposition of the sensitivity function  $d_T$  of the measure  $T$  used to discriminate the textures for the simple model of Fig. 2(b). In this section we state the necessary facts and definitions.

*a. Definition (of a linear combination of functions and of a set closed under linear combination).* Let  $G$  be a set of real-valued functions of a set  $Y$ . Any function  $f$  is said to be a linear combination of  $G$  if there exist real numbers  $\alpha_g$  such that  $f = \sum_{g \in G} \alpha_g g$ .  $G$  is said to be closed under linear combination if every linear combination of  $G$  is an element of  $G$ . Here we consider linear combinations of real-valued functions of the micropattern set  $\Omega$ .

All applications of histogram contrast analysis focus on spaces of IID textures. This concept depends on the following notion:

*b. Definition (of a reversible modulator).* Let  $p$  be a probability distribution on a set  $\Omega$  of micropatterns. Then any function  $\phi: \Omega \rightarrow \mathbb{R}$  is called a modulator of  $p$  if  $p + \phi$  is a probability distribution. If  $p - \phi$  is also a distribution, then  $\phi$  is called reversible. In addition,  $\phi$  is called maximal if  $A\phi$  fails to be a reversible modulator of  $p$  for all  $A > 1$ .

For example, the texture pairs shown in Fig. 1 all use reversible modulators ( $\lambda_1, \lambda_2, \lambda_3$ , and  $\lambda_4$ ) of the uniform probability distribution  $U$ . For  $i = 1, 2, 3, 4$ ,  $\lambda_i$  is a polynomial of order  $i$ . The texture pairs shown are  $I_{U-\lambda_i}$  next to  $I_{U+\lambda_i}$ .

The definition of a space of IID textures is analogous to the standard, linear algebraic definition of a space of functions, but it is complicated, because linear combinations of modulators of a given distribution  $p$  are not necessarily modulators of  $p$ :

*c. Definition (of spaces of modulators and of IID textures).* Let  $\Omega$  be a set of micropatterns. Call any set  $Q$  of functions  $q: \Omega \rightarrow \mathbb{R}$  a space (of modulators) if there exists a distribution  $p$  on  $\Omega$  such that  $Q$  is maximal with respect to the following two properties:

- (i)  $q$  is a modulator of  $p$ .
- (ii)  $q$  is a linear combination of functions in  $Q$ .



[That is, all  $q \in Q$  must satisfy properties (i) and (ii), whereas any  $q \notin Q$  must fail to satisfy at least one of properties (i) and (ii).] Specifically, we say that  $Q$  is a space (of modulators) relative to  $p$ . In addition, we say that  $Q$  is spanned by a subset  $B \subseteq Q$  if any function in  $Q$  is a linear combination of functions in  $B$ . In this case  $B$  is called a basis of  $Q$  and is said to be independent if none of its elements are linear combinations of the others. The dimension of  $Q$  is  $|B|$  for any such independent basis  $B$ . Any space that is a subset of  $Q$  is called a subspace of  $Q$ .

We call any set  $S$  of IID textures of  $\Omega$  a space if there exists a space  $Q$  of modulators relative to some distribution  $p$  on  $\Omega$  such that  $S = \{I_{p+q} \mid q \in Q\}$ . In this case  $p$  is called the base distribution of  $S$ , and  $Q$  is called the modulator space of  $S$ .

We shall sometimes say “ $Q$  is the modulator space of  $S$ ,” without making explicit reference to the base distribution  $p$ . The reader should not infer from such a statement that there exists only one modulator space of  $S$  (i.e.,  $Q$ ). Rather, such an assertion should be taken as shorthand for “ $Q$  is the modulator space of  $S$  relative to the base distribution  $p$  that is used in the current application of histogram contrast analysis.”

Note a few basic facts:

*Lemma 1.* Let  $\Omega$  be a set of micropatterns:

- a. The set of all IID textures of  $\Omega$  is a space.
- b. The empty set of IID textures of  $\Omega$  is a space.

*Proof:* Let  $p$  be a distribution on  $\Omega$ .

Note that the set  $Q_p$  of all modulators of  $p$  is a space, since any function  $q \in Q_p$  passes both conditions (i) and (ii) of definition 2.A.8.c., whereas any  $q \notin Q_p$  fails condition (i). Thus the corresponding set  $S_p$  of IID textures  $I_{p+q}$  ( $q \in Q_p$ ) is a space. But note that any texel distribution  $r$  can be written as  $p + q_r$ , for the modulator  $q_r = r - p$ . Thus  $S_p$  contains all IID textures of  $\Omega$ , implying condition a above.

On the other hand, the empty set  $e_p$  of modulators of  $p$  vacuously satisfies conditions (i) and (ii) of definition 2.A.8.c [that is, for all  $q \in e_p$  (of which there are none)  $q$  satisfies conditions (i) and (ii)]. However, for any function  $q \notin e_p$ ,  $q$  fails condition (ii), implying condition b above. ■

Lemma 1 gives the largest and smallest spaces of IID textures of a micropattern set  $\Omega$ . Typically, however, we shall be interested in spaces of IID textures falling between these extremes: nonempty, proper subspaces of the space of all IID textures of  $\Omega$ . In applying histogram contrast analysis to spaces of this sort, we shall need the following concept:

*d. Definition (of the projection of a function into a subspace).* Let  $p$  be a distribution on a set  $\Omega$  of micropatterns; let  $Q$  be a space of modulators relative to  $p$ , and let  $Q'$  be a subspace of  $Q$ . A theorem of linear algebra asserts that for any basis  $B'$  of  $Q'$  there exists a basis  $B \supseteq B'$  of  $Q$ . Thus for any  $q \in Q$  there exist real numbers  $\alpha_b$  (for all  $b \in B$ ) such that

$$q = \sum_{b \in B} \alpha_b b. \tag{15}$$

Let  $q'$  be the portion of  $q$  generated by only those basis functions in  $B'$ . That is,

$$q' = \sum_{b \in B'} \alpha_b b. \tag{16}$$

It is proven within the theory of linear algebra that  $q'$  does not depend on the specific choices of  $B'$  or  $B$ . Accordingly, we drop the reference to  $B'$  and  $B$  and call  $q'$  the projection of  $q$  into  $Q'$ .

*e. Orthogonal basis sets.* Let  $Y$  be a set of finite size  $N$ . Then any real-valued functions,  $f$  and  $g$ , of  $Y$  are called orthogonal if  $f \cdot g = 0$ .

Let  $\phi_0, \phi_1, \dots, \phi_{N-1}$  be mutually orthogonal, non-0 functions of  $Y$ . An important theorem of linear algebra states that for any function  $f: Y \rightarrow \mathbb{R}$  there exist weights  $W_0, W_1, \dots, W_{N-1} \in \mathbb{R}$  such that

$$f = \sum_{i=0}^{N-1} W_i \phi_i. \tag{17}$$

Specifically, these weights are given by

$$W_i = \frac{f \cdot \phi_i}{\phi_i \cdot \phi_i} \quad (i = 0, 1, \dots, N - 1). \tag{18}$$

Such functions  $\phi_0, \phi_1, \dots, \phi_{N-1}$  are said to compose an orthogonal basis of the space of all real-valued functions of  $Y$ .

### B. Perceptually Elementary Spaces of IID Textures

In general, histogram contrast analysis is used to test whether a given space  $S$  of IID textures of  $\Omega$  is perceptually elementary. We define this term precisely below. For the moment, the reader may rely on the intuition that a space  $S$  of IID textures is perceptually elementary if discrimination between textures of  $S$  is mediated exclusively by a single texture measure  $T$  that is additive with respect to  $\Omega$ . In this case the model of Fig. 2(b) applies to segregation of the textures in  $S$  with  $d_T$  serving as the nonlinearity of Fig. 2(b). As we shall try to illustrate in this section, if  $S$  is a strict subspace of the space of all IID textures of  $\Omega$ ,  $S$  may be perceptually elementary even if there exist multiple texture measures that are differentially sensitive to the micropatterns in  $\Omega$ .

To see how this can happen, let  $\Omega$  be a set of micropatterns to which an additive texture measure  $T$  is differentially sensitive, and let  $S$  be a space of textures with modulator space  $Q$  relative to base distribution  $p$ . However, suppose that  $d_T$  happens to be orthogonal to all  $q \in Q$  (that is,  $d_T \cdot q = 0$ ). In this case the projection of  $d_T$  into  $Q$  is 0. We learn below (proposition 1) that the observer's ability to segregate texture  $I_{p+q}$  from texture  $I_{p-q}$  is a monotonic function of  $d_T \cdot q$ . In particular, if  $d_T \cdot q = 0$ , then the average response of any  $T$ -edge detector to an  $I_{p+q}/I_{p-q}$  edge will be 0. This means that  $T$  can play no role in the drawing of distinctions within  $S$ .

Conversely, if  $d_T$  has a non-0 projection into  $Q$ , then  $T$  must be relevant to the discrimination of textures of  $S$ . Indeed, if discriminations within  $Q$  are determined exclusively by a unique, additive texture measure  $T$ , then the methods of histogram contrast analysis will allow one (1) to verify that there really is only one such  $T$  and (2) to measure the projection of  $d_T$  into  $Q$ .

### 1. Signal-Detection Theory and Textural Distinctions

We now take a signal-detection approach to texture segregation. We compute the signal-to-noise ratio for a given texture measure  $T$ , additive with respect to micropattern set  $\Omega$ . Let  $T$  be a texture measure, and suppose that  $T$  distinctions are drawn by linear  $T$ -edge detectors of the sort introduced in Subsection 2.A.6. Recall that the response of such a  $T$ -edge detector, TED, to any stimulus  $J$  is given by

$$\text{TED}(J) = \langle T(J) \rangle_{g_{\text{excit}}} - \langle T(J) \rangle_{g_{\text{inhib}}} \quad (19)$$

for some pair of additive poolers  $\langle \rangle_{g_{\text{excit}}}$  and  $\langle \rangle_{g_{\text{inhib}}}$ .

Suppose now that  $J_{p,q}$  is a stimulus that assigns a patch of IID texture  $I_p$  to the region windowed by  $g_{\text{excit}}$  and a patch of  $I_q$  to the region windowed by  $g_{\text{inhib}}$ . In this case we note that

$$\begin{aligned} \text{TED}(J_{p,q}) &= \langle T(I_p) \rangle_{g_{\text{excit}}} - \langle T(I_q) \rangle_{g_{\text{inhib}}} \\ &= \langle d_T \circ I_p \rangle_{g_{\text{excit}}} - \langle d_T \circ I_q \rangle_{g_{\text{inhib}}}. \end{aligned} \quad (20)$$

Note that  $\text{TED}(J_{p,q})$  is a random variable. Moreover, if  $g_{\text{excit}}$  and  $g_{\text{inhib}}$  pool over sufficiently many texels, then (by the central limit theorem)  $\text{TED}(J_{p,q})$  will be approximately normal. In this case signal-detection theory assures us that the probability that  $\text{TED}(J_{p,q})$  signals the presence of the  $I_p/I_q$  edge in  $J_{p,q}$  is a strictly increasing function of the signal-to-noise ratio

$$\begin{aligned} d'(J_{p,q}) &= \frac{|E[\text{TED}(J_{p,q})]|}{\{\text{var}[\text{TED}(J_{p,q})]\}^{1/2}} \\ &= \frac{|E[\langle d_T \circ I_p \rangle_{g_{\text{excit}}} - \langle d_T \circ I_q \rangle_{g_{\text{inhib}}}]|}{\{\text{var}[\langle d_T \circ I_p \rangle_{g_{\text{excit}}} - \langle d_T \circ I_q \rangle_{g_{\text{inhib}}}]\}^{1/2}}. \end{aligned} \quad (21)$$

Of course,  $T$  may be only one of many texture measures that are relevant to the drawing of distinctions within a given space  $S$  of IID textures. However, in the special case in which  $T$  is the only texture measure with a non-0 projection into the modulator space of  $S$ , the probability that the subject will detect the  $I_p/I_q$  edge in  $J_{p,q}$  becomes equal to the probability that  $\text{TED}(J_{p,q})$  will register the  $I_p/I_q$  edge. In this case we call  $S$  perceptually elementary. The definition of this notion follows.

### 2. Definition (of a Perceptually Elementary Space of IID Textures)

Call any space  $S$  of IID textures of a micropattern set  $\Omega$  perceptually elementary if there exists a function  $f: \Omega \rightarrow \mathbb{R}$  such that for any  $I_p, I_q \in S$ , the probability of observing a distinction between textures  $I_p$  and  $I_q$  is a strictly increasing function of

$$d'_S(p, q) = \frac{|E[\langle f \circ I_p \rangle_{g_{\text{excit}}} - \langle f \circ I_q \rangle_{g_{\text{inhib}}}]|}{\{\text{var}[\langle f \circ I_p \rangle_{g_{\text{excit}}} - \langle f \circ I_q \rangle_{g_{\text{inhib}}}]\}^{1/2}} \quad (22)$$

for spatial poolers  $\langle \rangle_{g_{\text{excit}}}$  and  $\langle \rangle_{g_{\text{inhib}}}$  with equal gain [Eq. (10)] and energy [Eq. (11)].

The constraints that  $\langle \rangle_{g_{\text{excit}}}$  and  $\langle \rangle_{g_{\text{inhib}}}$  have equal gain and energy are necessary to support subsequent inferences. Note, however, that these conditions will certainly be satisfied by standard sorts of linear edge detectors in which the excitatory and inhibitory receptive field lobes are symmetrically related.

As we shall show, it is possible to determine empirically whether a given space  $S$  is perceptually elementary. If  $S$  proves to be perceptually elementary, then we presume that distinctions within  $S$  are mediated exclusively by a single texture measure  $T$ , additive with respect to  $\Omega$ , and that Eq. (22) is satisfied by  $f = d_T$ . In this case the methods of histogram contrast analysis allow us to measure the projection of  $d_T$  into the modulator space of  $S$ .

### C. Histogram Contrast Thresholds

The stage is now set to formalize and justify the statements originally made in Section 1. Here we define the histogram contrast threshold, which is a threshold derived from comparisons such as those illustrated in Fig. 3. We have already defined the class of required stimuli: textures  $I_{p \pm A\phi}$  for a reversible modulator  $\phi$  of  $p$ . We have related the observer's performance  $d'$  to the output of the model in response to pairs of textures drawn from a perceptually elementary space. In this section we compute what the results of this task imply about the function  $d_T$ , which governs performance.

Let  $\Omega$  be a set of micropatterns, and suppose that  $S$  is a perceptually elementary space of IID textures of  $\Omega$ . Let  $f: \Omega \rightarrow \mathbb{R}$  satisfy Eq. (22), set  $K_1$  equal to the gain of each of the additive poolers,  $\langle \rangle_{g_{\text{excit}}}$  and  $\langle \rangle_{g_{\text{inhib}}}$ , and set  $K_2$  equal to the energy of each pooler. Throughout this section let  $p$  be a probability distribution on  $\Omega$ , let  $\phi$  be a reversible modulator of  $p$ , and let  $I_{p+\phi}$  and  $I_{p-\phi}$  be elements of  $S$ .

Note the following facts about  $I_{p+\phi}$ . Because  $I_{p+\phi}$  is IID,

$$\begin{aligned} E[\langle f \circ I_{p+\phi} \rangle_{g_{\text{excit}}}] &= K_1 E[f(\text{any texel of } I_{p+\phi})] \\ &= K_1 E[f(X_{p+\phi})], \end{aligned} \quad (23)$$

$$\begin{aligned} \text{var}[\langle f \circ I_{p+\phi} \rangle_{g_{\text{excit}}}] &= K_2 \text{var}[f(\text{any texel of } I_{p+\phi})] \\ &= K_2 \text{var}[f(X_{p+\phi})]. \end{aligned} \quad (24)$$

Similarly, we observe that

$$E[\langle f \circ I_{p-\phi} \rangle_{g_{\text{inhib}}}] = K_1 E[f(X_{p-\phi})], \quad (25)$$

$$\text{var}[\langle f \circ I_{p-\phi} \rangle_{g_{\text{inhib}}}] = K_2 \text{var}[f(X_{p-\phi})]. \quad (26)$$

Consider then a stimulus whose left-hand side is filled with  $I_{p+\phi}$  and whose right-hand side is filled with  $I_{p-\phi}$ . Because  $S$  is perceptually elementary, the probability of perceiving the boundary between the textures  $I_{p+\phi}$  and  $I_{p-\phi}$  is a strictly increasing function of

$$\begin{aligned} d'_S(p + \phi, p - \phi) &= \frac{|E[\langle f \circ I_{p+\phi} \rangle_{g_{\text{excit}}} - \langle f \circ I_{p-\phi} \rangle_{g_{\text{inhib}}}]|}{\{\text{var}[\langle f \circ I_{p+\phi} \rangle_{g_{\text{excit}}} - \langle f \circ I_{p-\phi} \rangle_{g_{\text{inhib}}}]\}^{1/2}} \\ &= \frac{K |E[f(X_{p+\phi}) - f(X_{p-\phi})]|}{\{\text{var}[f(X_{p+\phi}) - f(X_{p-\phi})]\}^{1/2}} \end{aligned} \quad (27)$$

for

$$K = K_1 / \sqrt{K_2}. \quad (28)$$

The preceding observations pave the way for the following proposition, which is the foundation of histogram contrast analysis. It states that the signal-to-noise ra-

tio, and hence observer performance on psychophysical discrimination tasks, is a function of  $f \cdot \phi$  for  $\phi$  an arbitrary modulator of some fixed distribution  $p$ . Therefore it allows one to learn about  $f$ , using a task such as that illustrated in Fig. 3.

*Proposition 1.* Let  $p$  be a probability distribution on a set  $\Omega$ , and let  $\phi$  be a reversible distribution modulator of  $p$ . Let  $X_{p+\phi}$  and  $X_{p-\phi}$  be independent,  $\Omega$ -valued random variables with distributions  $p + \phi$  and  $p - \phi$ , respectively. Then, for any nonnegative constant  $R$  and any real-valued function  $f$  of  $\Omega$ , the following conditions [Eqs. (29) and (30)] are equivalent:

$$\frac{|E[f(X_{p+\phi}) - f(X_{p-\phi})]|}{\{\text{var}[f(X_{p+\phi}) - f(X_{p-\phi})]\}^{1/2}} = R. \tag{29}$$

For  $X_p$ , an  $\Omega$ -valued random variable with distribution  $p$ ,

$$(f \cdot \phi)^2 = \frac{R^2 \text{var}[f(X_p)]}{R^2 + 2}. \tag{30}$$

The proof of proposition 1 uses the following:

*Lemma 2.* Let  $p$  be a probability distribution on  $\Omega$ , and let  $\phi$  be a reversible distribution modulator of  $p$ . Let  $X_{p+\phi}$  and  $X_{p-\phi}$  be independent  $\Omega$ -valued random variables with distributions  $p + \phi$  and  $p - \phi$ , respectively. Then, for any function  $f: \Omega \rightarrow \mathbb{R}$ ,

$$\text{var}[f(X_{p+\phi}) - f(X_{p-\phi})] = 2 \text{var}[f(X_p)] - 2(f \cdot \phi)^2 \tag{31}$$

for  $X_p$ , an  $\Omega$ -valued random variable with distribution  $p$ .  
*Proof:*

$$\begin{aligned} &\text{var}[f(X_{p+\phi}) - f(X_{p-\phi})] \\ &= \text{var}[f(X_{p+\phi})] + \text{var}[f(X_{p-\phi})] \\ &= f^2 \cdot (p + \phi) - [f \cdot (p + \phi)]^2 + f^2 \cdot (p - \phi) \\ &\quad - [f \cdot (p - \phi)]^2 \\ &= f^2 \cdot p + f^2 \cdot \phi - [(f \cdot p)^2 + 2(f \cdot p)(f \cdot \phi) + (f \cdot \phi)^2] \\ &\quad + f^2 \cdot p - f^2 \cdot \phi - [(f \cdot p)^2 - 2(f \cdot p)(f \cdot \phi) \\ &\quad + (f \cdot \phi)^2] \\ &= 2f^2 \cdot p - 2[(f \cdot p)^2 + (f \cdot \phi)^2] \\ &= 2 \text{var}[f(X_p)] - 2(f \cdot \phi)^2. \quad \blacksquare \end{aligned} \tag{32}$$

We now proceed with the proof of proposition 1. Note first that

$$\begin{aligned} |E[f(X_{p+\phi}) - f(X_{p-\phi})]| &= |E[f(X_{p+\phi})] - E[f(X_{p-\phi})]| \\ &= |f \cdot p + f \cdot \phi - (f \cdot p - f \cdot \phi)| = 2|f \cdot \phi|. \end{aligned} \tag{33}$$

From Eqs. (31) and (33) we see that Eq. (29) is equivalent to

$$\frac{2|f \cdot \phi|}{\{2 \text{var}[f(X_p)] - 2(f \cdot \phi)^2\}^{1/2}} = R. \tag{34}$$

Some algebra yields the result. ■

To see the relevance of proposition 1 to the study of visual perception, let us relate the texture-discrimination task illustrated in Fig. 3 to the proposition. Let  $S$  be

a space of IID textures relative to a distribution  $p$  on micropattern set  $\Omega$ . Consider a psychophysical task in which the subject is asked to detect the orientation (e.g., horizontal versus vertical) of a boundary between two IID textures in  $S$ :  $I_{p+\phi}$  and  $I_{p-\phi}$ . More specifically, let us use a staircase to vary the amplitude  $A$  of the modulator  $\phi$  of  $p$  from trial to trial in order to determine that value of  $A$  for which the subject detects the boundary between  $I_{p+A\phi}$  and  $I_{p-A\phi}$  with probability  $\alpha$  (e.g.,  $\alpha$  might be 0.75). (Of course, it is possible that no such value of  $A$  exists. That is, the subject may not be able to detect a boundary between  $I_{p+A\phi}$  and  $I_{p-A\phi}$  for any value of  $A$  satisfying the constraint that both  $p + A\phi$  and  $p - A\phi$  be probability distributions.) If  $S$  is perceptually elementary, then this level of performance corresponds to some fixed value,  $R_\alpha$ , of  $d'_S(p + A\phi, p - A\phi)$  [Eq. (22)]: that is, the subject detects the boundary with probability  $\alpha$  for the distribution modulator amplitude  $A$  such that

$$\begin{aligned} &d'_S(p + A\phi, p - A\phi) \\ &= \frac{|E[\langle f \circ I_{p+A\phi} \rangle_{g_{\text{excit}}} - \langle f \circ I_{p-A\phi} \rangle_{g_{\text{inhib}}}]|}{\{\text{var}[\langle f \circ I_{p+A\phi} \rangle_{g_{\text{excit}}} - \langle f \circ I_{p-A\phi} \rangle_{g_{\text{inhib}}}]\}^{1/2}} = R_\alpha. \end{aligned} \tag{35}$$

Equation (27) implies that Eq. (35) holds if and only if (iff)

$$\frac{|E[f(X_{p+A\phi}) - f(X_{p-A\phi})]|}{\{\text{var}[f(X_{p+A\phi}) - f(X_{p-A\phi})]\}^{1/2}} = \frac{R_\alpha}{K}. \tag{36}$$

But now, from proposition 1, we see that Eq. (36) holds iff

$$(f \cdot A\phi)^2 = \frac{R_\alpha^2 \text{var}[f(X_p)]}{R_\alpha^2 + 2K^2}. \tag{37}$$

And from Eq. (37) we reason that Eq. (35) holds iff

$$|f \cdot \phi| = |A|^{-1} \Psi_\alpha(p) \tag{38}$$

for

$$\Psi_\alpha(p) = \left[ \frac{R_\alpha^2 \text{var}[f(X_p)]}{R_\alpha^2 + 2K^2} \right]^{1/2}. \tag{39}$$

That is,  $|f \cdot \phi|$  is inversely proportional to  $|A|$ . Equation (38) leads us to call  $|A|$  the  $\alpha$  histogram contrast threshold of  $\phi$  relative to  $p$ .

Note: Histogram contrast thresholds are absolute values and hence are never negative. Henceforth any value referred to as a histogram contrast threshold is to be understood as a nonnegative quantity. We shall generally drop the reference to  $\alpha$  when it is irrelevant. In this case the reader may assume an arbitrary, fixed  $\alpha$  (e.g.,  $\alpha = 0.75$ ).

#### D. Cancellation Tests and the Characteristic Function of a Space of IID Textures

We have just seen how a psychophysical discrimination experiment (the experiment of measuring the histogram

contrast threshold of a modulator  $\phi$ ) yields information about the hypothetical visual function  $f$ . In this section we explore the ramifications of these facts when we have mutually orthogonal modulators  $\phi_1, \phi_2, \dots, \phi_N$ . As we show below, it is possible to determine whether the space  $S$  spanned by  $\phi_1, \phi_2, \dots, \phi_N$  relative to some base distribution  $p$  is perceptually elementary; that is, it is possible to determine whether there exists a single function  $f: \Omega \rightarrow \mathbb{R}$  [as in the model of Fig. 2(b)] capable of explaining all discriminations within  $S$  [by satisfying Eq. (22)]. Moreover, if such a function  $f$  exists, it is possible by psychophysical means to measure precisely the projection of  $f$  into  $S$  (i.e., that part of  $f$  responsible for making those texture discriminations).

A strategy for determining the projection of  $f$  into the space of functions spanned by  $\phi_1, \phi_2, \dots, \phi_N$  is suggested by Eq. (38).

Specifically, if we could obtain  $A_i$ , the histogram contrast threshold relative to  $p$ , of each basis element  $\phi_i$ , then we would know that

$$|f \cdot \phi_i| = \frac{\Psi(p)}{A_i}, \quad i = 1, 2, \dots, N, \quad (40)$$

where  $\Psi(p)$ , defined by Eq. (39), is a constant that is independent of  $i$ .

Suppose for the moment that all these thresholds  $A_i$  exist (are empirically obtainable). In this case Eq. (40) implies that

$$f \cdot \phi_i = \pm \frac{\Psi(p)}{A_i} \quad i = 1, 2, \dots, N. \quad (41)$$

If we could untangle the relative signs of the dot products in Eq. (41), then, using Eq. (17), we could synthesize the function

$$f_S = \sum_{i=1}^N \frac{f \cdot \phi_i}{\phi_i \cdot \phi_i} \phi_i. \quad (42)$$

It is a consequence of the theory of linear algebra that the function  $f_S$  is uniquely defined [up to the unknown and perceptually irrelevant scale factor  $\Psi(p)$ ] and is independent of the particular basis used to span the space of modulators that generate  $S$  relative to  $p$ . Accordingly, we call  $f_S$  the characteristic function of  $S$ . Indeed, if we suppose that distinctions within  $S$  are mediated exclusively by a single, additive texture measure  $T$ , then we must suppose that  $f = d_T$ . Hence  $f_S$  is the projection of  $d_T$  into the subspace of modulators spanned by  $\phi_1, \phi_2, \dots, \phi_N$ .

In fact, this untangling of relative signs is not difficult. For this purpose we use a method called the cancellation test. For any pair of basis elements,  $\phi_i$  and  $\phi_j$ , from Eq. (41) we reason that either

$$\begin{aligned} (a) \quad & |f_S \cdot (A_i \phi_i + A_j \phi_j)| = 2\Psi(p), \\ (b) \quad & |f_S \cdot (A_i \phi_i - A_j \phi_j)| = 0, \end{aligned} \quad (43)$$

or else

$$\begin{aligned} (a) \quad & |f_S \cdot (A_i \phi_i + A_j \phi_j)| = 0, \\ (b) \quad & |f_S \cdot (A_i \phi_i - A_j \phi_j)| = 2\Psi(p). \end{aligned} \quad (44)$$

As we shall now explain, Eqs. (43) and (44) allow us to determine the relative signs of  $f_S \cdot \phi_i$  and  $f_S \cdot \phi_j$ . In

the process we also subject our initial assumption that  $S$  is perceptually elementary to a strong empirical test.

If the set  $S$  is perceptually elementary, then for any pair of basis elements,  $\phi_i$  and  $\phi_j$ , one of the condition pairs in Eq. (43) and Eq. (44) must hold.

Condition (a) of Eqs. (43) implies that the histogram contrast of the distribution modulator  $A_i \phi_i + A_j \phi_j$  is at double the threshold level. In other words, Eqs. (43) imply that the histogram contrast threshold of the distribution modulator  $A_i \phi_i + A_j \phi_j$  is  $1/2$ . Thus, if condition (a) of Eqs. (43) holds, the subject should be able to detect the boundary between IID textures  $I_{p+(A_i \phi_i + A_j \phi_j)/2}$  and  $I_{p-(A_i \phi_i + A_j \phi_j)/2}$  with precisely threshold probability. Condition (b) of Eqs. (43) implies that the histogram contrast threshold of  $A_i \phi_i - A_j \phi_j$  relative to  $p$  is empirically unobtainable. Thus, if this condition holds, the subject should be unable to distinguish  $I_{p+B(A_i \phi_i - A_j \phi_j)}$  from  $I_{p-B(A_i \phi_i - A_j \phi_j)}$  for any value of  $B$  whatever.

Equations (44) express the mirror symmetric condition in which the histogram contrast threshold of  $A_i \phi_i + A_j \phi_j$  relative to  $p$  is unobtainable and the histogram contrast threshold of  $A_i \phi_i - A_j \phi_j$  relative to  $p$  is  $1/2$ .

For any given pair of basis functions,  $\phi_i$  and  $\phi_j$ , if the histogram contrast threshold of  $A_i \phi_i + A_j \phi_j$  relative to  $p$  is  $1/2$  and the histogram contrast threshold of  $A_i \phi_i - A_j \phi_j$  relative to  $p$  is unobtainable, then our data are consistent with Eqs. (43), and we presume in this case that the coefficients of  $\phi_i$  and  $\phi_j$  in the synthesis of  $f_S$  have the same sign. Otherwise, if the histogram contrast threshold of  $A_i \phi_i + A_j \phi_j$  relative to  $p$  is unobtainable and the histogram contrast threshold of  $A_i \phi_i - A_j \phi_j$  relative to  $p$  is  $1/2$ , then our data are consistent with Eqs. (44), and we presume that the coefficients of  $\phi_i$  and  $\phi_j$  in the synthesis of  $f_S$  have opposite signs.

The third possibility is that neither of these two empirical outcomes is observed. In this case we must conclude that  $S$  is not perceptually elementary, thus providing a strong test of our theoretical framework.

### E. What to Do When a Histogram Contrast Threshold is Empirically Unobtainable: The Method of Titration

Of course, it may well turn out that many of the histogram contrast thresholds  $A_i$  are not empirically obtainable. That is, no boundary emerges between textures  $I_{p+A_i \phi_i}$  and  $I_{p-A_i \phi_i}$ , even when  $A_i$  is as large as we can make it under the constraint that  $p - A_i \phi_i$  and  $p + A_i \phi_i$  be probability distributions. We cannot necessarily conclude from this fact that  $\phi_i$  contributes nothing to  $f_S$ ; all that we know is that the amplitude of  $\phi_i$  in  $f_S$  is smaller than can be directly measured relative to  $p$ . In this section we describe a method whereby we can determine  $\phi_i \cdot f_S$  without directly measuring the histogram contrast threshold of  $\phi_i$  relative to  $p$ , thus increasing the sensitivity of the psychophysical methodology.

Note first that, whenever we can find two IID textures  $I_q$  and  $I_r$  that can be distinguished (here  $q$  and  $r$  are the texel distributions of the two textures), then there is at least one threshold that we can empirically measure. In particular, we know that we can obtain the histogram contrast threshold of  $\phi$  relative to  $p$  for  $p = (q + r)/2$  and  $\phi = (q - r)/2$ , because

$$I_{p+\phi} = I_q, \quad I_{p-\phi} = I_r. \quad (45)$$

We can, moreover, orthogonalize our other basis elements with respect to  $\phi$  in order to be able to include  $\phi$  in our basis.

Thus, if there exists a pair of IID textures that segregate, then we can construct an orthogonal basis that contains an element  $\phi$  for which the histogram contrast threshold can be obtained relative to some distribution  $p$ . Throughout the rest of this section, all histogram contrast thresholds mentioned should be understood as being relative to a given such distribution  $p$ . As we shall see, for any other element  $\rho$  of this basis (other than the element  $\phi$  whose histogram contrast threshold is assumed to exist), if the histogram contrast threshold of  $\rho$  relative to  $p$  cannot be obtained then we can titrate  $\phi$  with  $\rho$  in order to determine  $\rho \cdot f_S$ . An important implication of this fact is that, whenever any distinctions at all can be drawn between elements of some space of IID textures, the psychophysical tools exist to allow us to determine whether that space of textures is perceptually elementary and, if so, to measure the characteristic function of that space.

1. *Titration Lines*

Let  $\rho$  be another basis element (aside from the element  $\phi$  whose histogram contrast threshold we assume exists relative to  $p$ ) for which the histogram contrast threshold is empirically unobtainable. Suppose, moreover, for real numbers  $x$  and  $y$ , that  $x\rho + y\phi$  is a reversible modulator of  $p$  and that  $A$  is the histogram contrast threshold of  $x\rho + y\phi$  relative to  $p$ . Then from Eq. (38) we derive the fact that

$$|f_S \cdot (x\rho + y\phi)| = A^{-1}\Psi(p) \tag{46}$$

for  $\Psi(p)$  defined by Eq. (39). Thus

$$f_S \cdot (x\rho + y\phi) = \pm A^{-1}\Psi(p), \tag{47}$$

from which we infer that

$$y = \frac{\pm A\Psi(p) - f_S \cdot \rho x}{f_S \cdot \phi}. \tag{48}$$

*Empirical Significance of Eq. (48).* Let  $A_\phi$  be the histogram contrast threshold of  $\phi$  (empirically obtainable by assumption). Although we cannot obtain the threshold of  $\rho$ , we can obtain a nonempty set of real-number pairs  $(x, y)$  such that the histogram contrast threshold of  $x\rho + y\phi$  is 1 [this sets  $A = 1$  in Eq. (48)]. (In particular, we know that this is true for  $x = 0$  and  $y = A_\phi$ .) Equation (48) shows that this set of pairs  $(x, y)$  is the union of two parallel straight lines:

$$y = \frac{\Psi(p) - f_S \cdot \rho x}{f_S \cdot \phi}, \tag{49}$$

$$y = \frac{-\Psi(p) - f_S \cdot \rho x}{f_S \cdot \phi}. \tag{50}$$

One of these lines has a positive  $y$  intercept, and the other has a negative  $y$  intercept (which one is which depends on the sign of  $f_S \cdot \phi$ ). We call the line with the positive

$y$  intercept the titration line relating  $\rho$  to  $\phi$ . Note that these two lines have the same slope and also that, for  $x_0$  the  $x$  intercept of the titration line,  $-x_0$  is the  $x$  intercept of the other line. For either line,

$$|f_S \cdot \rho| = \Psi(p)/|x_0|. \tag{51}$$

Thus

1.  $|f_S \cdot \rho|$  is inversely proportional to  $|x_0|$ . Moreover, as is clear from Eq. (49),
2. The slope of the titration line is negative iff the signs of  $f_S \cdot \rho$  and  $f_S \cdot \phi$  are equal.

We can estimate the titration line by varying  $y$  around  $A_\phi$  and obtaining, in each case, the value of  $x$  for which the probability of discriminating the textures  $I_{p+(x\rho+y\phi)}$  and  $I_{p-(x\rho+y\phi)}$  is at threshold. By fitting a straight line through the data we can thus estimate the sign and magnitude of  $f_S \cdot \rho$  relative to  $f_S \cdot \phi$ .

2. *Three-Point Estimation of the Titration Line*

In practice it is convenient to obtain three points on the titration line relating  $\rho$  to  $\phi$ . One of these points is, of course, the point  $(0, A_\phi)$ .

The other two points require explanation. For notational simplicity we assume that  $\rho$  and  $\phi$  are maximal relative to  $p$  (see definition 2.A.8.2). The histogram contrast threshold of  $\rho$  is not empirically obtainable. Thus the textures  $I_{p+\rho}$  and  $I_{p-\rho}$  segregate (if at all) with subthreshold probability. On the other hand, the textures  $I_{p+\phi}$  and  $I_{p-\phi}$  segregate with suprathreshold probability.

For any proportion  $\pi$ , we can generate the IID texture  $I_{p+\pi\rho+(1-\pi)\phi}$  by simply taking a random mixture of  $I_{p+\phi}$  and  $I_{p+\rho}$ , where, for any texel  $\tau$ ,

$$I_{p+\pi\rho+(1-\pi)\phi}(\tau) = \begin{cases} I_{p+\rho}(\tau) & \text{with probability } \pi \\ I_{p+\phi}(\tau) & \text{otherwise} \end{cases}. \tag{52}$$

A similar procedure produces the IID texture  $I_{p-(\pi\rho+(1-\pi)\phi)}$ . For  $\pi = 0$  the textures  $I_{p+\pi\rho+(1-\pi)\phi}$  and  $I_{p-(\pi\rho+(1-\pi)\phi)}$  segregate with suprathreshold probability. For  $\pi = 1$  these textures segregate with subthreshold probability. We infer that for some proportion  $\pi^+$  (which can be obtained with a staircase procedure)  $I_{p+\pi^+\rho+(1-\pi^+)\phi}$  segregates from  $I_{p-(\pi^+\rho+(1-\pi^+)\phi)}$  with threshold probability. Thus the point  $(\pi^+, 1 - \pi^+)$  is on the titration line relating  $\rho$  to  $\phi$ .

By an analogous procedure we can obtain the proportion  $\pi^-$  for which the IID textures  $I_{p+\pi^-\rho-(1-\pi^-)\phi}$  and  $I_{p-(\pi^-\rho-(1-\pi^-)\phi)}$  segregate with threshold probability. This procedure yields a third point  $(-\pi^-, 1 - \pi^-)$  on the titration line.

Thus we obtain three points (which are well spread in most practical cases) on the titration line relating  $\rho$  to  $\phi$ :  $(-\pi^-, 1 - \pi^-)$ ,  $(0, A_\phi)$ , and  $(\pi^+, 1 - \pi^+)$ .

This three-point method is illustrated graphically in Fig. 4. Here  $\phi$  and  $\rho$  are maximal reversible modulators of a distribution  $p$  defined on  $\Omega$ . Suppose that  $A_\phi$ , the histogram contrast threshold of  $\phi$  relative to  $p$ , is empirically obtainable, whereas the histogram contrast threshold of  $\rho$  relative to  $p$  cannot be measured. Points in the plane of Fig. 4 correspond to IID texture discrimi-

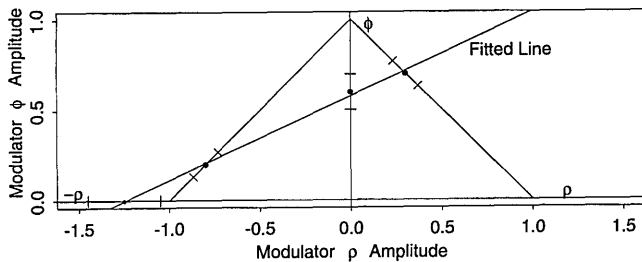


Fig. 4. Method of titration. Points in this plane represent IID texture pairs that are generated by mixing two maximal modulators  $\phi$  and  $\rho$  of the base distribution  $p$ . A point  $(x, y)$  in this plane corresponds to the discrimination of  $I_{p+x\rho+y\phi}$  from  $I_{p-x\rho-y\phi}$ . If the measure used by the subject for this space of textures is one dimensional, then the locus of points corresponding to a given performance level is a straight line. The point  $(0, A_\phi)$  corresponds to a direct measurement of the histogram contrast threshold of  $\phi$ . Here  $A_\rho$  is not directly measurable. Instead, we perform experiments along the two indicated 45-deg lines, which correspond to a mixing of  $\phi$  with  $\pm\rho$  to determine the proportions  $\pi^+$  and  $\pi^-$  of  $\pm\rho$  that must be mixed with  $\phi$  for threshold performance to be achieved. A line fitted through these three empirically measured points is then extrapolated to its  $x$  intercept. The absolute value of the  $x$  intercept is an estimate of  $A_\rho$ . A negative titration-line slope indicates that  $A_\phi$  and  $A_\rho$  have the same sign, and a positive slope indicates opposite signs.

nation stimuli in which the two textures to be discriminated are  $I_{p+x\rho+y\phi}$  and  $I_{p-(x\rho+y\phi)}$ , for  $x$  ranging from  $-1$  to  $1$  along the horizontal axis and for  $y$  ranging from  $0$  to  $1$  along the vertical axis. Thus the point  $(0, A_\phi)$  corresponds to the texture pair  $(I_{p+A_\phi\phi}, I_{p-A_\phi\phi})$ . Of course, this pair of textures segregates with threshold probability, and thus, as we have noted, the point  $(0, A_\phi)$  is on the titration line relating  $\rho$  to  $\phi$ . The point  $(0, 1)$  corresponds to the pair of textures  $(I_{p+\phi}, I_{p-\phi})$ , which segregate with suprathreshold probability. The point  $(1, 0)$  corresponds to the pair of textures  $(I_{p+\rho}, I_{p-\rho})$ , which segregate (if at all) with subthreshold probability. The line between  $(1, 0)$  and  $(0, 1)$  is the locus of points corresponding to texture pairs  $[I_{p+\pi\rho+(1-\pi)\phi}, I_{p-(\pi\rho+(1-\pi)\phi)}]$ , where  $0 \leq \pi \leq 1$ . Thus we can use a staircase procedure to adjust  $\pi$  from trial to trial, thereby moving back and forth along the line from  $(1, 0)$  to  $(0, 1)$  in order to find the point  $[\pi^+, (1 - \pi^+)]$  corresponding to the pair  $[I_{p+\pi^+\rho+(1-\pi^+)\phi}, I_{p-(\pi^+\rho+(1-\pi^+)\phi)}]$  of textures that segregate with threshold probability. This is the second of our three points on the titration line relating  $\rho$  to  $\phi$ . The third point  $(-\pi^-, 1 - \pi^-)$  is obtained in similar fashion by use of a staircase procedure to move back and forth from trial to trial along the line between  $(-1, 0)$  and  $(0, 1)$ , which is the locus of all points  $(-\pi, 1 - \pi)$  corresponding to pairs of stimuli  $[I_{p+\pi\rho-(1-\pi)\phi}, I_{p-(\pi\rho-(1-\pi)\phi)}]$ .

We proceed to obtain the best-fitting straight line through points  $(\pi^+, 1 - \pi^+)$ ,  $(A_\phi, 0)$ , and  $(-\pi^-, 1 - \pi^-)$ . The absolute value of the  $x$  intercept of this line is inversely proportional to  $|f_S \cdot \rho|$ , and the sign of the slope of this line is negative iff the relative signs of  $f_S \cdot \rho$  and  $f_S \cdot \phi$  are equal.

Finally, note that, if the obtained data do not conform (within measurement error) to a straight line, then we must conclude that the space  $S$  of the IID textures that we have been exploring is not perceptually elementary. Thus here again we obtain a powerful test of our model. Indeed, the method of titration frequently provides a

variety of tests of the hypothesis that  $S$  is perceptually elementary. For instance, if the histogram contrast thresholds of multiple basis functions, e.g.,  $\phi_1$  and  $\phi_2$ , can be empirically measured, then for any basis function  $\rho$  whose histogram contrast threshold cannot be measured the two titration lines relating  $\rho$  to  $\phi_1$  and to  $\phi_2$  can be obtained. These two lines provide independent estimates of  $f_S \cdot \rho$ . If these estimates diverge significantly, then we must reject the hypothesis that  $S$  is perceptually elementary.

## F. Summary

Thus far,

1. We have presented an argument to show that the method of increment thresholds is insufficient to allow psychophysicists to determine the internal codes of physical stimuli for perceptual systems.
2. We have described histogram contrast analysis. This is a new psychophysical method that can be used to infer the internal codes assigned to the members of a set  $\Omega$  of micropatterns (e.g., to small square patches of various gray levels) by processes that draw perceptual distinctions between IID textures of  $\Omega$ . Histogram contrast analysis involves a collection of related tools. The tools that we have described are

a. Histogram contrast thresholds (Subsection 2.C). These are the principal tools of histogram contrast analysis. For  $\Omega$  a set of micropatterns, and  $S$  a space of IID textures of  $\Omega$ , suppose that the visual system draws preattentive distinctions between certain textures in  $S$ . It was shown that, if the set  $S$  is perceptually elementary (i.e., if there exists a single function  $f_S$  that characterizes the effective internal code applied to  $\Omega$  by the mechanism that draws these distinctions), then the histogram contrast threshold of a given real-valued function  $\phi$  of  $\Omega$  is inversely proportional to  $|\phi \cdot f_S|$  (the absolute value of the cross correlation of  $\phi$  with  $f_S$ ).

b. Cancellation tests (Subsection 2.D). If one can obtain histogram contrast thresholds  $A_\phi$  and  $A_\rho$  of two different functions  $\phi$  and  $\rho$ , then a method is available for determining the relative signs of  $f_S \cdot \phi$  and  $f_S \cdot \rho$ ; this method simultaneously provides a strong empirical test of the assumption that  $S$  is perceptually elementary. Specifically, for  $\beta^+ = A_\phi\phi + A_\rho\rho$  and  $\beta^- = A_\phi\phi - A_\rho\rho$ , if  $S$  is perceptually elementary then for  $A_{\beta^+}$  and  $A_{\beta^-}$ , the histogram contrast thresholds of  $\beta^+$  and  $\beta^-$ , either (1)  $A_{\beta^+} = 0.5$  and  $A_{\beta^-}$  is empirically unobtainable or (2)  $A_{\beta^-} = 0.5$  and  $A_{\beta^+}$  is empirically unobtainable. If alternative (1) holds, then  $f_S \cdot \phi$  and  $f_S \cdot \rho$  have the same sign. If alternative (2) holds, then  $f_S \cdot \phi$  and  $f_S \cdot \rho$  have opposite signs. If neither alternative holds, we conclude that  $S$  is not perceptually elementary.

c. The method of titration (Subsection 2.E). It may well turn out that the histogram contrast threshold of a given function cannot be obtained directly. However, if any preattentive distinctions at all can be drawn between IID textures in  $S$ , then there always exists a distribution  $p$  relative to which one can measure the histogram contrast threshold of some modulator  $\phi$ , thus obtaining an estimate of  $|f_S \cdot \phi|$ . For any other modulator  $\rho$  of  $p$ , orthogonal to  $\phi$ , one can determine the ratio of  $f_S \cdot \rho$  to  $f_S \cdot \phi$  by seeing how much  $\phi$  is needed to compensate for

a small quantity of  $\rho$  near threshold. The three-point method elaborated in Subsection 2.E.2 is a precise recipe for determining the ratio of  $f_S \cdot \rho$  to  $f_S \cdot \phi$ . In addition, the method of titration also provides a strong test of the hypothesis that  $S$  is perceptually elementary. If  $S$  is perceptually elementary, then the amount of  $\phi$  required for sustaining a fixed level of performance as  $\rho$  is varied by amount  $\delta_\rho$  must be a linear function of  $\delta_\rho$ . Any deviation of titration data from this predicted linear relationship forces the conclusion that  $S$  is not perceptually elementary.

d.  $f_S$  synthesis. If the data support the hypothesis that  $S$  is perceptually elementary, then one can use either histogram contrast thresholds or the method of titration to obtain  $f_S \cdot \phi_i$  for each element  $\phi_i$  of a suitable set of orthogonal functions. One can then compose  $f_S$  as a linear combination of the functions  $\phi_i$ .

### 3. USE OF HISTOGRAM CONTRAST ANALYSIS TO STUDY SPATIALLY LOCAL VISUAL NONLINEARITIES

In this section we describe experiments that apply histogram contrast analysis to IID textures of the set  $\Gamma$  of 0.13-deg squares of intensities  $4.4i \text{ cd/m}^2$ ,  $i = 0, \dots, 16$  (e.g., those illustrated in Figs. 1 and 3). For convenience we identify the micropatterns of  $\Gamma$  with their respective gray levels and write  $i$  for the gray level  $4.4i \text{ cd/m}^2$ . We begin by situating these experiments in the context of current models of texture perception.

#### A. Back-Pocket Model of Texture Segregation Applied to IID Textures of $\Gamma$

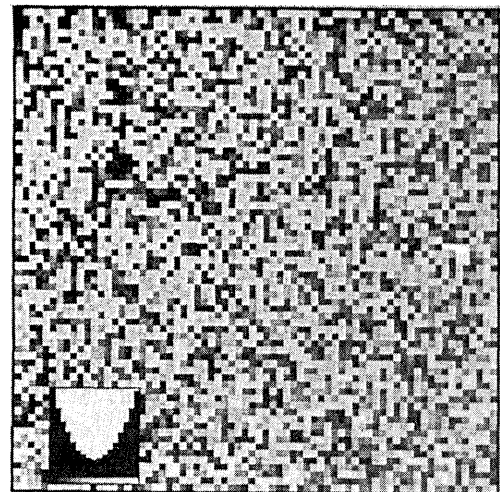
As we mentioned above, most models of texture segregation (e.g., Refs. 8–16; see also Refs. 17 and 18) suppose that the visual system is equipped with a variety of linear filters (realized by arrays of linear neurons with different receptive-field profiles). A given filter will respond differently to different textures; that is, the histogram of values in the neural image resulting from the application of a given filter will vary systematically across different classes of textures. Such histogram differences can be converted to differences in spatial average value by application of one or another pointwise nonlinearity to filter outputs. This is the basis of the back-pocket model of texture perception.

In Subsection 1.D it was mentioned that IID textures with equal mean and variance [such as those of Figs. 1(g)–1(j)] cannot be segmented by a back-pocket model channel that has a sufficiently large receptive field. As a result, in Section 2 we restricted our attention to the far simpler model of Fig. 2(b), which omits the initial spatial filter. Here we elaborate on this observation and point out that the performance of back-pocket model channels with sufficiently small receptive fields is similar to that of the model shown in Fig. 2(b), so that histogram contrast analysis still applies.

Consider what happens when we apply a linear filter  $F^*$  with a large impulse response to an IID texture,  $I_p$ , of  $\Gamma$ . Specifically, suppose that the impulse response  $F$  encompasses a number of different texels in  $I_p$ . Since the gray levels of the texels of  $I_p$  are jointly independent

random variables, and the value produced at any point in space is a linear combination of those gray levels, the central limit theorem implies that the value at any point in the neural array embodying the output image must be (approximately) normally distributed. Thus the histogram of the neural image  $F^* I_p$  will have a Gaussian profile.

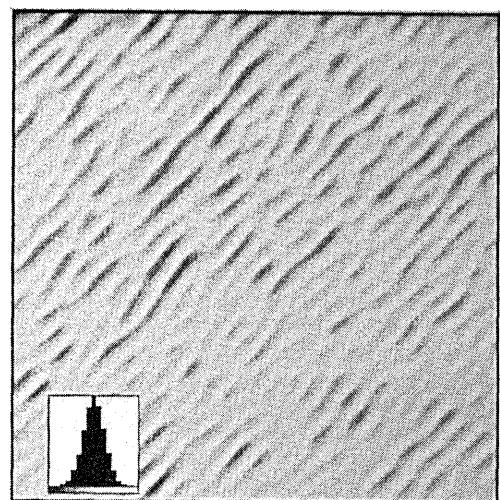
This point is illustrated in Fig. 5. In Fig. 5(a) is shown an IID texture  $I_p$  of  $\Gamma$ . The texel distribution  $p$  characterizing  $I_p$  is shown inset within this patch of  $I_p$ . This



(a)



(b)



(c)

Fig. 5. Back-pocket model and IID textures. (a) An IID texture  $I_p$  and its gray-level histogram, (b) the impulse response of a typical linear filter used in the modeling of early vision, (c) the resulting filtered texture and its intensity histogram. Note that the filtered texture-intensity histogram is approximately Gaussian and thus is determined only by the mean and the variance of  $I_p$  (and by the characteristics of the filter).

distribution  $p$  gives high probabilities to extreme light and dark intensities. Consequently,  $I_p$  appears high in contrast. Figure 5(b) shows the impulse response of a garden-variety, oriented linear filter  $F^*$ , and Fig. 5(c) shows the texture that results from application of  $F^*$  to  $I_p$ . The important point to note here is that the gray-level histogram of  $F^* I_p$  [shown inset within Fig. 5(c)] is Gaussian. It does not matter what sort of texel distribution  $p$  is used to generate the initial IID texture  $I_p$ . The output image will always have a Gaussian gray-level histogram.

Note, however, that (1) the normal distribution is characterized completely by its mean and its variance and (2) the mean and the variance of a linear combination of random variables depend only on the means and variances of the random variables being combined and the weights used in the linear combination. Thus, for any IID textures  $I_p$  and  $I_q$ , if gray-level distributions  $p$  and  $q$  have the same mean and variance, then application of  $F^*$  to  $I_p$  and  $I_q$  will obliterate the difference between them. To be precise, it can be shown that the pixels of the output image  $F^* I_p$  will be jointly normal with a covariance matrix approximately equal to that of the (jointly normal) output image  $F^* I_q$ .

Now, the IID textures  $I_{U+\lambda_3}$  and  $I_{U-\lambda_3}$  of Figs. 1(g) and 1(h) [the same is true for the textures  $I_{U+\lambda_4}$  and  $I_{U-\lambda_4}$ , Figs. 1(i) and 1(j)] do, in fact, have gray-level distributions with equal mean and variance (a statement discussed further below and justified in Appendix A). There is, nonetheless, a perceptual boundary between them.

The above argument rules out the possibility that this boundary is being produced by any back-pocket model channel that uses an up-front linear filter with a receptive field that pools over a sufficient number of texels. Thus the only back-pocket model channels whose output is relevant to the distinction that we see drawn here are those whose up-front linear filters are calibrated to preserve the relative values of individual texels. The best filter for this purpose would be no filter at all: i.e., the identity filter. This is the case illustrated in Fig. 2(b). However, many models assume that all the filters are spatial frequency specific, either low pass or band pass. In this case the optimal filter for preserving the relative values of individual texels is that filter whose output value at a given texel is dominated most thoroughly by the value of that particular texel.

Consider, for instance, a filter whose receptive field has an excitatory center the same size as a single micropattern of  $\Gamma$  and an inhibitory surround that encompasses a region the size of a  $3 \times 3$  block of texels. In this case a cell whose receptive field is centered on a given texel  $\tau$  of an IID texture of  $\Gamma$  has a surround that takes the average of the eight texels forming a ring around  $\tau$ . Figure 6 illustrates the effect of application of such a filter to a stimulus [Fig. 6(a)] composed of a patch of  $I_{U-\lambda_3}$  (at the left) abutting a patch of  $I_{U+\lambda_3}$  (at the right). The histograms enclosed within patches of texture in Figs. 6(a) and 6(c) are all actual image intensity histograms, not texel distributions. (Of course, the texel histograms of the patches of  $I_{U-\lambda_3}$  and  $I_{U+\lambda_3}$  closely resemble the probability distributions  $U - \lambda_3$  and  $U + \lambda_3$ , respectively.) Convolution of this input stimulus with the receptive field shown in Fig. 6(b) yields the (neural) output image shown

in Fig. 6(c). Although the histograms of the filtered textures in Fig. 6(c) at the left and at the right appear similar (both seem to be roughly normal), these histograms are systematically different, as we see in Figs. 6(d)–6(f). Figures 6(d) and 6(e) are the histograms of the left and right patches of filtered texture, respectively, and Fig. 6(f) displays the difference between these two histograms. The important point to note is that the form of the modulator  $\lambda_3$  is preserved quite well in the difference between the histograms of the filtered textures.

Thus we see that the particular isotropic bandpass filter whose receptive field is shown in Fig. 6(b) does a good job of preserving the histogram information in the original IID texture to which it is applied.

The precise characteristics of the up-front filters used to distinguish between  $I_{U+\lambda_3}$  and  $I_{U-\lambda_3}$  are unimportant. What matters is that there evidently exists at least one up-front linear transformation that does, in fact, preserve the information (carried in the histograms) about the relative frequencies of different gray levels in the two textures. If indeed there is such a unique filter (or a set of such filters with highly correlated outputs), then the distinction drawn between  $I_{U+\lambda_3}$  and  $I_{U-\lambda_3}$  (as well as between  $I_{U+\lambda_4}$  and  $I_{U-\lambda_4}$ ) depends only on the pointwise nonlinearity that is presumed to follow this up-front linear filter. And in this case we can use histogram contrast analysis to characterize this nonlinearity partially. Specifically, for  $S$ , the space of all IID textures of  $\Gamma$  with mean and variance equal to  $I_U$ , we can apply histogram contrast analysis to measure the characteristic function  $f_S$ . If, indeed, discrimination between textures in  $S$  depends on a single back-pocket model channel with pointwise nonlinearity  $f$ , then  $f_S$  will be identical to that portion of  $f$  that is relevant to drawing distinctions in  $S$ .

For measurement of  $f_S$ , a set of reversible modulators is needed that form an orthogonal basis set and that have the additional characteristic of having equal mean and variance. To generate a suitable basis set, we started with the discrete-domain monomials  $f_i(v) = v^i$ ,  $i = 0, 1, \dots, 16$ , which form a basis set but are not mutually orthogonal. Then we performed Gram–Schmidt orthogonalization on the  $f_i$  (see, e.g., Ref. 21). We obtained the modulators  $\lambda_0, \lambda_1, \dots, \lambda_{16}$  by scaling the resulting polynomials to be maximal reversible modulators of  $U$ , the uniform distribution on  $\Gamma$ . Each modulator  $\lambda_i$  is a discrete-domain polynomial of degree  $i$ . (The set  $\lambda_i$ ,  $i = 0, 1, \dots, 16$  of polynomials is a discrete-domain analog of the Legendre polynomials.<sup>22</sup>) These are the  $\lambda_i$  shown in Fig. 1.

The functions  $\lambda_3, \lambda_4, \dots, \lambda_{16}$  are orthogonal. Moreover, each of these functions is orthogonal to  $\zeta_0, \zeta_1$ , and  $\zeta_2$ , where, for  $v = 0, 1, \dots, 16$ ,

$$\begin{aligned}\zeta_0(v) &= 1, \\ \zeta_1(v) &= v, \\ \zeta_2(v) &= v^2.\end{aligned}\tag{53}$$

Let  $h$  be any modulator of  $U$  produced by taking a linear combination of the higher-order polynomials,  $\lambda_3, \lambda_4, \dots, \lambda_{16}$ . Because  $\lambda_3, \lambda_4, \dots, \lambda_{16}$  are all orthogonal to each of  $\zeta_1$  and  $\zeta_2$ ,  $h$  must also be orthogo-



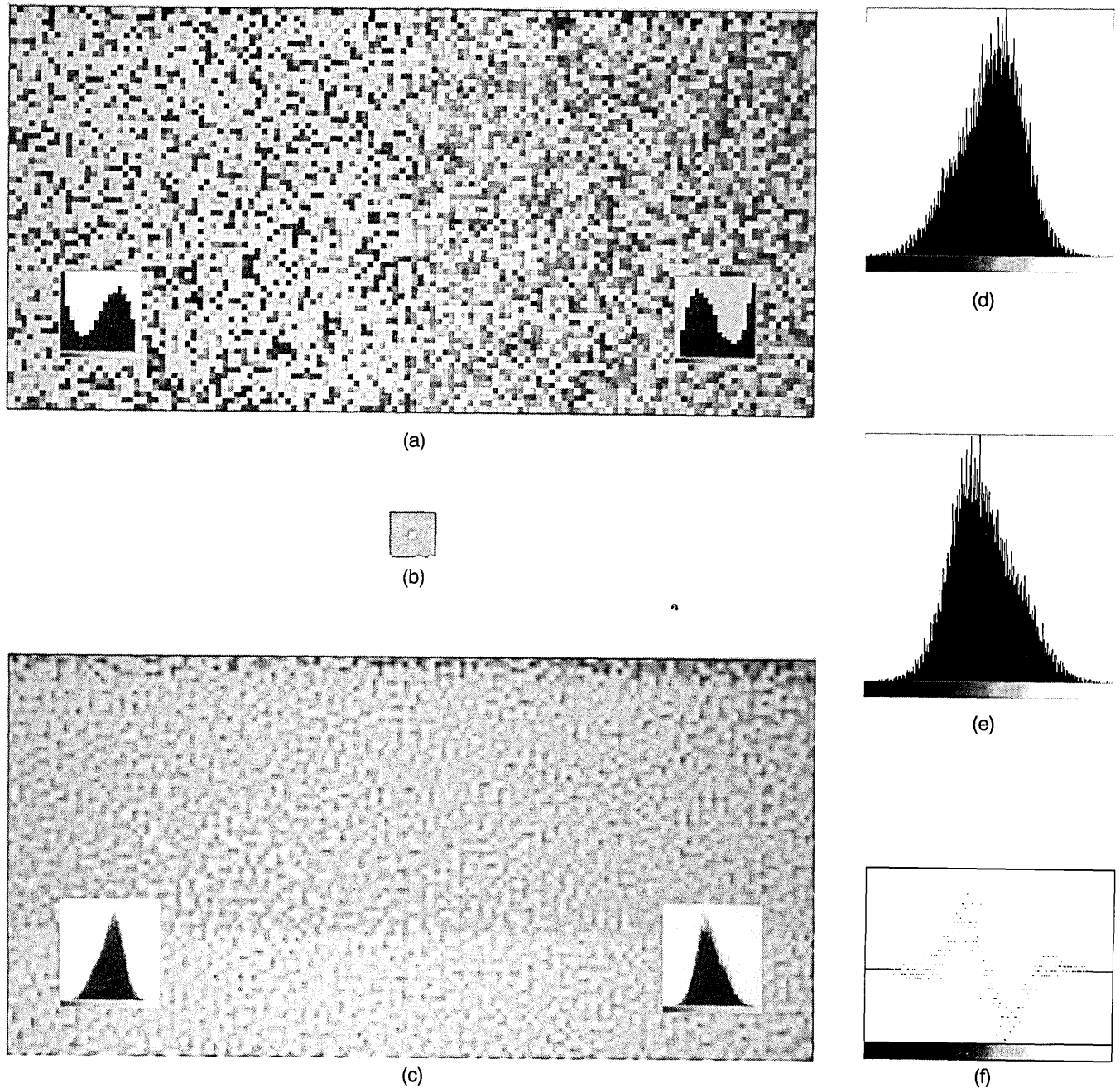


Fig. 6. Spatial filtering that preserves information about the moments of IID texture histograms other than mean and variance. (a) Patches of IID texture  $I_{U-\lambda_3}$  and  $I_{U+\lambda_3}$ . The bar graphs enclosed within the patches of texture are the intensity histograms of the patches. These histograms approximate the texel distributions  $U - \lambda_3$  and  $U + \lambda_3$  (which have equal mean and variance). (b) The receptive field  $F$  of a linear filter capable of preserving information about the higher moments (other than mean and variance) of the histograms of IID textures of  $\Gamma$ . This receptive field has an excitatory center exactly the size of a micropattern of  $\Gamma$  and an annular inhibitory surround covering a region equal in size to the ring of texels surrounding a given texel. If the excitatory center of a neuron with such a receptive field is aligned with a texel  $\tau$ , then the output of that neuron will be dominated by the intensity of  $\tau$ . (c) The result of convolving the IID texture field shown in (a) with receptive field  $F$ . Enclosed bar graphs are intensity histograms of left-hand and right-hand filtered textures,  $F * I_{U-\lambda_3}$  and  $F * I_{U+\lambda_3}$ . (d) The intensity histogram (enlarged) of the left-hand patch of filtered texture,  $F * I_{U-\lambda_3}$ . (e) The intensity histogram (enlarged) of the right-hand patch of filtered texture,  $F * I_{U+\lambda_3}$ . (f) The difference between the two histograms shown in (d) and (e). Note that the form of the modulator  $\lambda_3$  is preserved quite well in the difference between the histograms of the filtered textures. Thus the particular isotropic bandpass filter whose receptive field is shown in (b) does a good job of preserving the histogram information in the original IID texture to which it is applied.

nal to  $\zeta_1$  and  $\zeta_2$ . It follows from proposition A.1 (in Appendix A) that a random variable with distribution  $U + h$  has the same mean and variance as a random variable with distribution  $U$ . This means that, by obtaining histogram contrast thresholds and/or using the

titration method with the modulators  $\lambda_i, i = 3, 4, \dots, 16$ , we can completely characterize that part of the nonlinearity implicated by Figs. 1(g)–1(j) that is relevant to distinguishing between IID textures with equal mean and variance.

## B. Experiments to Measure Histogram Contrast Thresholds

### 1. Procedure

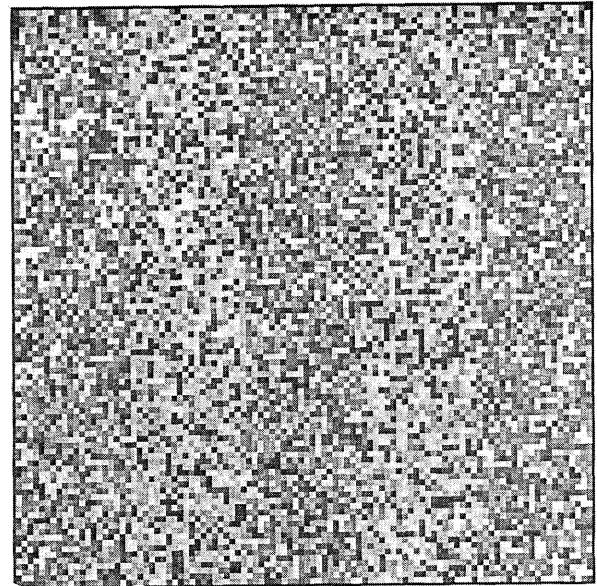
In our experiments the subject was required to judge, in a 250-ms display, the orientation of a square-wave grating (0.26 cycles/deg at a viewing distance of 1 m) that modulated between a texture pair for which we wanted to measure the histogram contrast threshold. The complete display subtended 13 deg, and the texels in these displays ( $4 \times 4$  pixels on our display device) had a width of 0.13 deg. Thus there were approximately 8 texels/deg, each bar of the square wave had a width of 20 texels, and the total display was 100 texels wide. We randomized the phase of the square wave to thwart any attempt by the subject to attend only to a location at which a texture boundary had appeared in the past.

The display device was a BARCO Calibrator monitor driven by a Gateway 2000 486 personal computer equipped with a Truevision ATVista board. The nature of the experiments (manipulation of the gray-level histogram of an image) dictates that extreme care be taken in calibration of the monitor. It is crucial that the function relating the luminance of the different gray levels be linear; otherwise unwanted distortions will be imposed on our histograms. Our procedure was to linearize each gun of our monitor separately, by eye. We created a pattern similar in pixel structure ( $4 \times 4$  pixel squares) to that of the experimental images. The pattern consisted of a central area and a small surround. Half of the pixels in the central patch were gray and of mean luminance. The other half were split evenly between pixels of two other gray levels, and they swapped locations rapidly and randomly. Half of the surround pixels were also gray and of mean luminance. The other half were made up of a third type of pixel, the intensity of which was carefully adjusted so that when these pixels were blurred the surround was indiscriminable from the central region. The central and the surrounding regions combined used only a small area at the center of the monitor; the rest of the monitor was maintained at the same mean luminance. In this manner, adjusting a surround so that after a blur it looked the same as an interior that flickered between two gray levels, we could find the midpoint between any two gray levels. We assigned values to each of the 17 gray levels by finding the midpoint between the widest interval available (e.g., gray level 7 is assigned to the midpoint of gray levels 0 and 14, not 6 and 8). Combining the three guns does not appear to compromise the linearity.<sup>23</sup> One reason that we chose texels to be  $4 \times 4$  pixel squares was to minimize the influence of each texel on the luminance of its neighbors.<sup>24</sup> The resulting mean luminance was  $35.5 \text{ cd/m}^2$  (corresponding to gray level 8) with steps of  $4.4 \text{ cd/m}^2$  between successive gray levels.

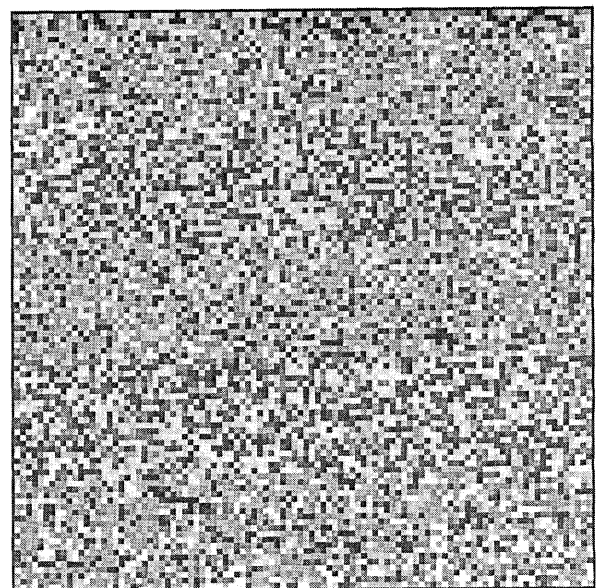
The experiments described in this section measured the histogram contrast threshold  $A_i$  relative to  $U$  for each  $\lambda_i$  in our sequence for which this was possible. That is, we found the amplitude  $A_i$  for which  $I_{U+A_i\lambda_i}$  was just barely discriminable from  $I_{U-A_i\lambda_i}$  for each  $i = 3, 4, \dots, 16$  for which such an  $A_i$  existed. Examples of stimulus frames are shown in Fig. 7. Figure 7(a) shows a vertical square wave modulating between  $I_{U+\lambda_3}$  and  $I_{U-\lambda_3}$ .

Figure 7(b) shows a horizontal square wave modulating between  $I_{U+\lambda_4}$  and  $I_{U-\lambda_4}$ . The black borders of the frames were not present when they were displayed in the experiment; instead, stimulus frames were displayed in the middle of a uniform field of luminance equal to the mean luminance of the textures. Before stimulus onset, the screen was uniform midgray except for a small, low-contrast cue spot.

Subjects viewed several hundred such textures in a single block of trials. In a given block the subject attempted to distinguish the orientations of square waves



(a)



(b)

Fig. 7. Example stimuli: (a) a vertical IID texture square wave modulating between  $I_{U+\lambda_3}$  and  $I_{U-\lambda_3}$ , (b) a horizontal IID texture square wave modulating between  $I_{U+\lambda_4}$  and  $I_{U-\lambda_4}$ . The black border was not present in the experiments; rather, the stimulus was surrounded by a gray field with the same luminance as the mean luminance of the textures.

**Table 1. Histogram Contrast Thresholds for Subject JE**

| Modulator   | Experimental Data |        | Results from Simulations <sup>a</sup> |                    |                     |
|-------------|-------------------|--------|---------------------------------------|--------------------|---------------------|
|             | Threshold A       | Trials | Bias                                  | Standard Deviation | Confidence Interval |
| $\lambda_3$ | 0.214             | 1376   | 0.0000                                | 0.008              | 0.198–0.228         |
| $\lambda_4$ | 0.319             | 1434   | 0.0005                                | 0.012              | 0.293–0.342         |
| $\lambda_5$ | 0.445             | 1494   | –0.0003                               | 0.018              | 0.406–0.478         |
| $\lambda_6$ | <i>b</i>          | 1383   | —                                     | —                  | —                   |
| $\lambda_7$ | <i>b</i>          | 452    | —                                     | —                  | —                   |

<sup>a</sup>Simulation results are taken from 1000 simulated experiments.

<sup>b</sup>Threshold performance was not attainable within the range of histogram contrasts available.

**Table 2. Histogram Contrast Thresholds for Subject CC**

| Modulator   | Experimental Data |        | Results from Simulations <sup>a</sup> |                    |                     |
|-------------|-------------------|--------|---------------------------------------|--------------------|---------------------|
|             | Threshold A       | Trials | Bias                                  | Standard Deviation | Confidence Interval |
| $\lambda_3$ | 0.257             | 1142   | –0.0001                               | 0.010              | 0.236–0.276         |
| $\lambda_4$ | 0.367             | 1084   | –0.0001                               | 0.012              | 0.341–0.390         |
| $\lambda_5$ | 0.538             | 927    | –0.0004                               | 0.020              | 0.497–0.574         |
| $\lambda_6$ | <i>b</i>          | 506    | —                                     | —                  | —                   |
| $\lambda_7$ | <i>b</i>          | 579    | —                                     | —                  | —                   |

<sup>a</sup>Simulation results are taken from 1000 simulated experiments.

<sup>b</sup>Threshold performance was not attainable within the range of histogram contrasts available.

modulating between  $I_{U+AA_i}$  and  $I_{U-AA_i}$  for all the different distribution modulators  $\lambda_i$ . It was critical to mix all conditions in each block of trials because, if the subject were required to distinguish only one sort of IID texture  $I_{U+AA_i}$  from its dual  $I_{U-AA_i}$  in a given block, the subject might be able to adjust his or her texture measure  $T$  to optimize performance on the particular task posed in that block of trials. In this case we would not be learning about the differential sensitivity of a single texture measure that was constant across the set of all distribution modulators  $\lambda_i$ ,  $i = 3, 4, \dots, 16$ .

The paradigm was two-alternative forced choice with a staircase used to place trials. Feedback was provided after each trial. The trials for a given pair of textures  $I_{U+AA_i}$  and  $I_{U-AA_i}$  resulted in a psychometric function (percent correct as a function of the histogram modulation amplitude  $A$ ). These functions were fitted with a cumulative Weibull distribution with use of a maximum-likelihood estimator, and the point at which this function was equal to 75% correct was taken to be the histogram contrast threshold.

The reliability of our psychometric function parameters was determined by use of a variation on Efron's bootstrap method as described by Maloney.<sup>25</sup> This procedure simulates data for 1000 experiments by sampling a binomial distribution for each value of the independent variable for which trials were run. The value of the psychometric function at that point was used as the probability (the parameter  $p$ ) of the binomial distribution, and the number of trials run at that point was used as the other parameter ( $N$ ) of the binomial distribution. 1000 psychometric functions were fitted to the 1000 sets of simulated data, and the 1000 sets of parameters of these psychometric functions were used to determine the variability of the original parameters.

## 2. Results

Using this setup, we ran interleaved staircases to obtain histogram contrast thresholds for each of  $\lambda_3, \lambda_4, \lambda_5, \lambda_6$ , and  $\lambda_7$ . We had previously observed that the higher-order histogram contrast thresholds  $A_i$ ,  $i = 8, 9, \dots, 16$  were not empirically obtainable. For this reason we did not attempt to measure these thresholds in the current study.

There were two subjects, authors John Econopouly (JE) and Charles Chubb (CC). JE was extremely well practiced at the task, and CC was moderately well practiced before data collection began.

Of the five histogram contrast thresholds that were measured, the three low-order thresholds were easily obtainable, but the thresholds for  $\lambda_6$  and  $\lambda_7$  were not. The thresholds obtained are shown in Tables 1 and 2 for subjects JE and CC, respectively. Also shown are estimates of the bias, the standard deviation, and the 95% nonparametric confidence interval for each threshold obtained. These were calculated from 1000 simulated experiments, obtained from the bootstrap method described above.

Parenthetically, we observed that these values were significantly lower if in a given block of trials the subject was required to distinguish only one sort of IID texture  $I_{U+AA_i}$  from its dual  $I_{U-AA_i}$  (i.e.,  $i$  was kept fixed). This suggests that the subject can partially adjust the differential sensitivity of the relevant texture measure  $T$  to the micropatterns of  $\Gamma$  in order to optimize performance on a given discrimination task. These trials were not included in our analysis.

In summary, we were able to measure directly three histogram contrast thresholds for each of two subjects. From Eq. (38) it follows that, if  $S$  is perceptually elementary with characteristic function  $f_S$ , then  $|f_S \cdot \lambda_i|$  is inversely proportional to the  $A_i$  measured, for  $i = 3, 4, 5$ .

However, we still do not know (1) whether  $S$  is perceptually elementary, (2) the relative signs of the dot products  $f_S \cdot \lambda_i$ ,  $i = 3, 4, 5$ , or (3) the contributions to  $f_S$  of basis elements  $\lambda_i$  for  $i > 5$ . We can, however, use the titration method discussed in Subsection 2.E to settle all these issues.

### C. Experiments Involving Titrations

#### 1. Procedure

For texture pairs  $I_{U+A\lambda_i}$  and  $I_{U-A\lambda_i}$  for which we were unable to achieve threshold performance within the range of histogram contrast available (i.e., within the range  $0 \leq A \leq 1$ ), we titrated each of the textures  $I_{U+\lambda_i}$  and  $I_{U-\lambda_i}$  with one of the textures  $I_{U+\lambda_j}$  for which threshold performance was attainable. This produced two pairs of textures, each of which is presented in a square-wave modulated display, in a manner completely analogous to the procedure detailed above for measuring histogram contrast thresholds directly. The only difference is that now instead of the modulation amplitude we are measuring the proportions  $\pi_{-ij}$  and  $\pi_{ij}$  of each of  $I_{U-\lambda_i}$  and  $I_{U+\lambda_i}$  that we have to mix in randomly with  $I_{U+\lambda_j}$  to achieve threshold performance. In other words, we created stimuli that alternated between textures  $I_{U-\pi_{-ij}\lambda_i+(1-\pi_{-ij})\lambda_j}$  and  $I_{U+\pi_{-ij}\lambda_i-(1-\pi_{-ij})\lambda_j}$  to find the proportion  $\pi_{-ij}$  that yielded threshold performance; similarly, we used stimuli that alternated between textures  $I_{U+\pi_{ij}\lambda_i+(1-\pi_{ij})\lambda_j}$  and  $I_{U-\pi_{ij}\lambda_i-(1-\pi_{ij})\lambda_j}$  to find the proportion  $\pi_{ij}$  that yielded threshold performance (Subsection 2.E.2).

The paradigm is again two-alternative forced choice, with a staircase used to place trials. The trials for a given titration result in a psychometric function (percent correct as a function of the proportion  $\pi$ ). These functions were fitted with a cumulative Weibull distribution with use of a maximum-likelihood estimator, and the point at which this function was equal to 75% correct was taken to be the histogram contrast threshold.

Each titration procedure has three data points associated with it: the simple histogram contrast threshold  $A_j$  and the proportions  $\pi_{-ij}$  and  $\pi_{ij}$ . These are plotted at  $(-\pi_{-ij}, 1 - \pi_{-ij})$ ,  $(0, A_j)$  and  $(\pi_{ij}, 1 - \pi_{ij})$ , as shown in Fig. 4. The variabilities of the estimates of  $\pi_{ij}$  and  $\pi_{-ij}$  are determined by use of exactly the same bootstrap procedure as is described in Subsection 3.B.1. A titration line is fitted to these three points by use of a weighted least-squared-error fit, so that we can determine whether the points are well modeled by a straight line and, if so, what the  $x$  intercept of that line is. The error of each data point is taken to be its distance from the titration line along the line where it could occur [i.e., diagonal distance for  $(-\pi_{-ij}, 1 - \pi_{-ij})$  and  $(\pi_{ij}, 1 - \pi_{ij})$ , vertical distance for  $(0, A_j)$ ]. These individual errors are first weighted by the inverse standard deviations of their distributions (since a variable data point should restrict the location of the titration line less than a more definite point), then squared, and finally added together to yield the total error that is to be minimized.

The variability of the  $x$  intercept of the titration line is determined by use of a bootstrap procedure, as well. A simulated experiment is generated for each of the three

psychometric functions involved, and a line is fitted to each of the three simulated thresholds by use of the weighted least-squared-error fit described above. This is repeated 1000 times, giving us 1000 simulated  $x$  intercepts from which to compute the standard deviation.

Because we measured three histogram contrast thresholds directly ( $A_3, A_4$ , and  $A_5$ ), as described in Subsection 3.B, we can measure three independent titration lines for each of  $\lambda_6$  and  $\lambda_7$  by titrating with each of  $\lambda_3, \lambda_4$ , and  $\lambda_5$ . This provides us with three independent estimates of  $f_S \cdot \lambda_6$  and  $f_S \cdot \lambda_7$  (each multiplied by an unknown proportionality constant). These values enter into Eq. (18) to determine the weights for basis set elements  $\lambda_6$  and  $\lambda_7$  in the synthesis of  $f_S$ . Comparison of the three independent estimates of the two weights will provide a measure of consistency. The directly measured histogram contrast thresholds  $A_3, A_4$ , and  $A_5$  yield estimates of  $|f_S \cdot \lambda_3|$ ,  $|f_S \cdot \lambda_4|$ , and  $|f_S \cdot \lambda_5|$  in accordance with Eq. (38) [again we ignore the proportionality constant  $\Psi(U)$ ]. The relative signs of  $f_S \cdot \lambda_3$ ,  $f_S \cdot \lambda_4$  and  $f_S \cdot \lambda_5$  will be determined by the signs of the slopes of the titration lines of  $\lambda_3, \lambda_4$ , and  $\lambda_5$  with  $\lambda_6$  and  $\lambda_7$ .

#### 2. Results

We obtained estimates of  $A_6$  and  $A_7$  for the same two subjects, using the technique of titration. We did not obtain estimates of  $A_i$  for  $i > 7$ . The complete experimental results are tabulated in Tables 3 and 4, along with the results of various estimation procedures that will be explained as this section unfolds.

First, let us examine the results for subject JE, for whom we measured all three possible titration lines:  $\lambda_6$  titrated with each of  $\lambda_3, \lambda_4$ , and  $\lambda_5$ ; and  $\lambda_7$  titrated with each of  $\lambda_3, \lambda_4$ , and  $\lambda_5$ . The results are shown graphically in Fig. 8.

These graphs deserve careful consideration. They are meant to convey two separate points, each important to our conclusions. First, note that each graph contains three actual data points. In Fig. 8(a) these three points are (1) the point obtained by titrating  $I_{U-\lambda_6}$  with  $I_{U+\lambda_3}$ , plotted at  $(-\pi_{-6,3}, 1 - \pi_{-6,3})$ ; (2) the point  $(0, A_3)$ ; and (3) the point obtained by titrating  $I_{U+\lambda_6}$  with  $I_{U+\lambda_3}$ , plotted at  $(\pi_{6,3}, 1 - \pi_{6,3})$ . Each of these points has associated with it a 95% confidence interval, obtained by the bootstrap procedure. A line is fitted to the three points in each graph by a method of weighted least squares. The first point to note is that in none of the six cases displayed did the line fall outside the confidence intervals for the points. In the case of the  $\lambda_6$  titrations very many trials were run (see Table 3), with the express purpose of getting the confidence intervals very tight: the lines still fall within the intervals. This result is consistent with the hypothesis that the space  $S$  is perceptually elementary (Subsection 2.8.2), suggesting that distinctions drawn between IID textures of  $\Gamma$  with equal mean and variance are mediated by a single, additive texture measure  $T$ .

The second point is simply that the absolute values of the three independent estimates of  $A_6$  and  $A_7$  (the  $x$  intercepts of the extrapolated titration lines) match remarkably well. An average of the three estimates gives us the final estimates: 1.02 for  $A_6$  and 1.52 for  $A_7$  for subject JE. The sign of each titration line slope tells us

**Table 3. Titrations for Subject JE**

| Titration <sup>b</sup>       | Modulators              | Results from Simulations <sup>a</sup> |        |         |                    |                     |                            |                                 |                  |
|------------------------------|-------------------------|---------------------------------------|--------|---------|--------------------|---------------------|----------------------------|---------------------------------|------------------|
|                              |                         | Experimental Data <sup>a</sup>        |        | Bias    | Standard Deviation | Confidence Interval | $x$ Intercept <sup>c</sup> | Standard Deviation <sup>c</sup> | $p$ <sup>c</sup> |
|                              |                         | Threshold $\pi$                       | Trials |         |                    |                     |                            |                                 |                  |
| $\lambda_3$ with $\lambda_6$ | $\lambda_3, -\lambda_6$ | 0.008                                 | 1533   | 0.0001  | 0.004              | 0.004–0.015         | –1.032                     | 0.015                           | 0.805            |
|                              | $\lambda_3, \lambda_6$  | 0.346                                 | 1362   | –0.0003 | 0.008              | 0.330–0.360         |                            |                                 |                  |
| $\lambda_4$ with $\lambda_6$ | $\lambda_4, -\lambda_6$ | 0.457                                 | 1372   | –0.0002 | 0.008              | 0.440–0.471         | 1.029                      | 0.012                           | 0.136            |
|                              | $\lambda_4, \lambda_6$  | 0.012                                 | 1566   | 0.0002  | 0.005              | 0.004–0.021         |                            |                                 |                  |
| $\lambda_5$ with $\lambda_6$ | $\lambda_5, -\lambda_6$ | 0.008                                 | 1574   | 0.0016  | 0.008              | 0.000–0.027         | –1.011                     | 0.011                           | 0.407            |
|                              | $\lambda_5, \lambda_6$  | 0.599                                 | 1325   | –0.0004 | 0.011              | 0.577–0.619         |                            |                                 |                  |
| $\lambda_3$ with $\lambda_7$ | $\lambda_3, -\lambda_7$ | 0.323                                 | 291    | –0.0001 | 0.015              | 0.291–0.349         | 1.500                      | 0.117                           | 0.496            |
|                              | $\lambda_3, \lambda_7$  | 0.088                                 | 452    | 0.0001  | 0.013              | 0.061–0.112         |                            |                                 |                  |
| $\lambda_4$ with $\lambda_7$ | $\lambda_4, -\lambda_7$ | 0.118                                 | 310    | –0.0023 | 0.043              | 0.030–0.196         | –1.510                     | 0.212                           | 0.491            |
|                              | $\lambda_4, \lambda_7$  | 0.426                                 | 311    | –0.0010 | 0.013              | 0.394–0.448         |                            |                                 |                  |
| $\lambda_5$ with $\lambda_7$ | $\lambda_5, -\lambda_7$ | 0.525                                 | 336    | 0.0001  | 0.029              | 0.478–0.595         | 1.553                      | 0.231                           | 0.198            |
|                              | $\lambda_5, \lambda_7$  | 0.189                                 | 361    | –0.0036 | 0.045              | 0.088–0.265         |                            |                                 |                  |

<sup>a</sup>Results for each of the component thresholds are shown in columns 3–7.

<sup>b</sup>Each titration has three thresholds associated with it, the two shown here and one each from Tables 1 and 2.

<sup>c</sup>Results of the titration lines fitted to these thresholds are shown in columns 8–10.  $p$  refers to the percentile represented by the true error of the data points to the best-fitting titration line, compared with the errors of simulated points drawn from a linear model to their best-fitting titration lines. 1000 simulated experiments were run, each with the same number of trials as in the original experiments.

**Table 4. Titrations for Subject CC**

| Titration <sup>b</sup>       | Modulators              | Results from Simulations <sup>a</sup> |        |         |                    |                     |                            |                                 |                  |
|------------------------------|-------------------------|---------------------------------------|--------|---------|--------------------|---------------------|----------------------------|---------------------------------|------------------|
|                              |                         | Experimental Data <sup>a</sup>        |        | Bias    | Standard Deviation | Confidence Interval | $x$ Intercept <sup>c</sup> | Standard Deviation <sup>c</sup> | $p$ <sup>c</sup> |
|                              |                         | Threshold $\pi$                       | Trials |         |                    |                     |                            |                                 |                  |
| $\lambda_3$ with $\lambda_6$ | $\lambda_3, -\lambda_6$ | 0.018                                 | 387    | 0.0016  | 0.010              | 0.003–0.042         | –1.064                     | 0.042                           | 0.404            |
|                              | $\lambda_3, \lambda_6$  | 0.387                                 | 316    | 0.0002  | 0.012              | 0.362–0.411         |                            |                                 |                  |
| $\lambda_3$ with $\lambda_7$ | $\lambda_3, -\lambda_7$ | 0.303                                 | 165    | –0.0006 | 0.010              | 0.283–0.324         | 1.381                      | 0.128                           | 0.001            |
|                              | $\lambda_3, \lambda_7$  | 0.054                                 | 318    | 0.0009  | 0.014              | 0.028–0.083         |                            |                                 |                  |

<sup>a</sup>Results for each of the component thresholds are shown in columns 3–7.

<sup>b</sup>Each titration has three thresholds associated with it, the two shown here and one each from Tables 1 and 2.

<sup>c</sup>Results of the titration lines fitted to these thresholds are shown in columns 8–10.  $p$  refers to the percentile represented by the true error of the data points to the best-fitting titration line, compared with the errors of simulated points drawn from a linear model to their best-fitting titration lines. 1000 simulated experiments were run, each with the same number of trials as in the original experiments.

whether the thresholds for the two basis set elements being titrated have the same sign or opposite signs. As it turns out,  $A_3, A_5,$  and  $A_7$  share one sign, and  $A_4$  and  $A_6$  share the opposite sign.

Given the close agreement of the three independent estimates of  $A_6$  and  $A_7$  for subject JE, we believed that it was unnecessary to get three independent estimates for subject CC. Instead, we chose to titrate each of  $\lambda_6$  and  $\lambda_7$  with  $\lambda_3$  alone, with the results shown in Figs. 9(a) and 9(b). These figures are completely analogous to Figs. 8(a) and 8(d), respectively. Again, note that in Fig. 9(a) the titration line falls within the confidence intervals of the three points. However, for the  $\lambda_7$  titrations shown in Fig. 9(b) the titration line is outside two of the confidence intervals. The estimates of  $A_6$  and  $A_7$  from the  $x$  intercepts of the titration lines for CC are 1.06 and 1.38, respectively.

The fact that one of CC's titration lines did not fall within the confidence intervals of all its points motivated us to devise a more comprehensive statistic to determine how likely it is that the data that we measured can be modeled by a straight line. This is crucial because, if the data cannot be modeled by a straight line, then the

space  $S$  of textures cannot be considered perceptually elementary.

The statistic that we devised was again based on a bootstrapped simulation. The idea is to assume that a straight line does model the data, choose the best-fitting straight line available (the one shown on the graphs), and use the data points implied by that line as input to a bootstrap procedure to generate a sequence of simulated experiments. After each simulated experiment (which will produce three simulated titration-line points), we measure the error of these points to the best-fitting straight line, to find out how far off from straight are typical data sets. Finally, we compare this expected error measure with the actual error in the data, that is, with the squared deviations of the measured data points from the line shown in the graphs.

The results of this test are shown in the final column of Tables 3 and 4, and they confirm the impressions drawn from simply looking at the confidence intervals. The first seven titrations do not fail the linearity test: the total squared error measured from the data points to the fit line was within a 95% confidence interval of the typical error displayed by simulated data points to a fit line after

a straight-line model is assumed in each case. However, the final titration,  $\lambda_7$  with  $\lambda_3$  for subject CC, fails this test: it is extremely unlikely ( $p < 0.001$ ) that these data came from a straight line. This means that for CC we must reject the hypothesis that the space  $S$  of textures with mean and variance equal to  $I_U$  is perceptually elementary. This finding suggests that more than a single texture measure is relevant to the distinctions drawn by CC between textures in  $S$ . However, it should be noted that, although the deviation from linearity of CC's data is statistically significant, it is nonetheless a small effect. It is likely that CC's performance on these segregation tasks is determined for the most part by the single texture measure whose characteristic function we shall construct.

**D. Constructing  $f_S$**

Since we have determined that the space  $S$  of textures is perceptually elementary for subject JE and approximately so for CC, we can now synthesize the characteristic function  $f_S$  for each subject by setting

$$f_S = \sum_{i=3}^7 \frac{f_S \cdot \lambda_i}{\lambda_i \cdot \lambda_i} \lambda_i. \quad (54)$$

The sign of  $f_S \cdot \lambda_3$  was arbitrarily chosen to be positive, and the signs of  $f_S \cdot \lambda_i$ ,  $i = 4, 5, 6, 7$  were determined

by the results of the titration experiments. Specifically, assuming (arbitrarily) that  $f_S \cdot \lambda_3$  is positive, the signs of the slopes of the titration lines measured for JE dictate that  $f_S \cdot \lambda_5$  and  $f_S \cdot \lambda_7$  are also positive, whereas  $f_S \cdot \lambda_4$  and  $f_S \cdot \lambda_6$  are negative. For subject CC we had to assume that the relative signs of  $f_S \cdot \lambda_3$ ,  $f_S \cdot \lambda_4$  and  $f_S \cdot \lambda_5$  were the same as for subject JE, since not all the titrations required for determining those signs were run for CC. This assumption is consistent with the signs of the two titrations that were run for CC. Because of the great amount of time required for each titration, and since the higher-order basis elements appear to contribute least to the nonlinearity, we chose not to measure the contributions of the elements higher than  $\lambda_7$ .

Figures 10(a) and 10(b) show the reconstructions of  $f_S$  for subjects JE and CC, respectively.

**E. Discussion**

The results of this experiment demonstrate first that  $S$  is perceptually elementary. This implies that discriminations within  $S$  are mediated exclusively by a single texture measure  $T$ , additive with respect to  $\Gamma$ . Because  $S$  turns out to be perceptually elementary, the methods of histogram contrast analysis yield a characteristic function  $f_S$  (Fig. 10). ( $f_S$  is defined only up to arbitrary ad-

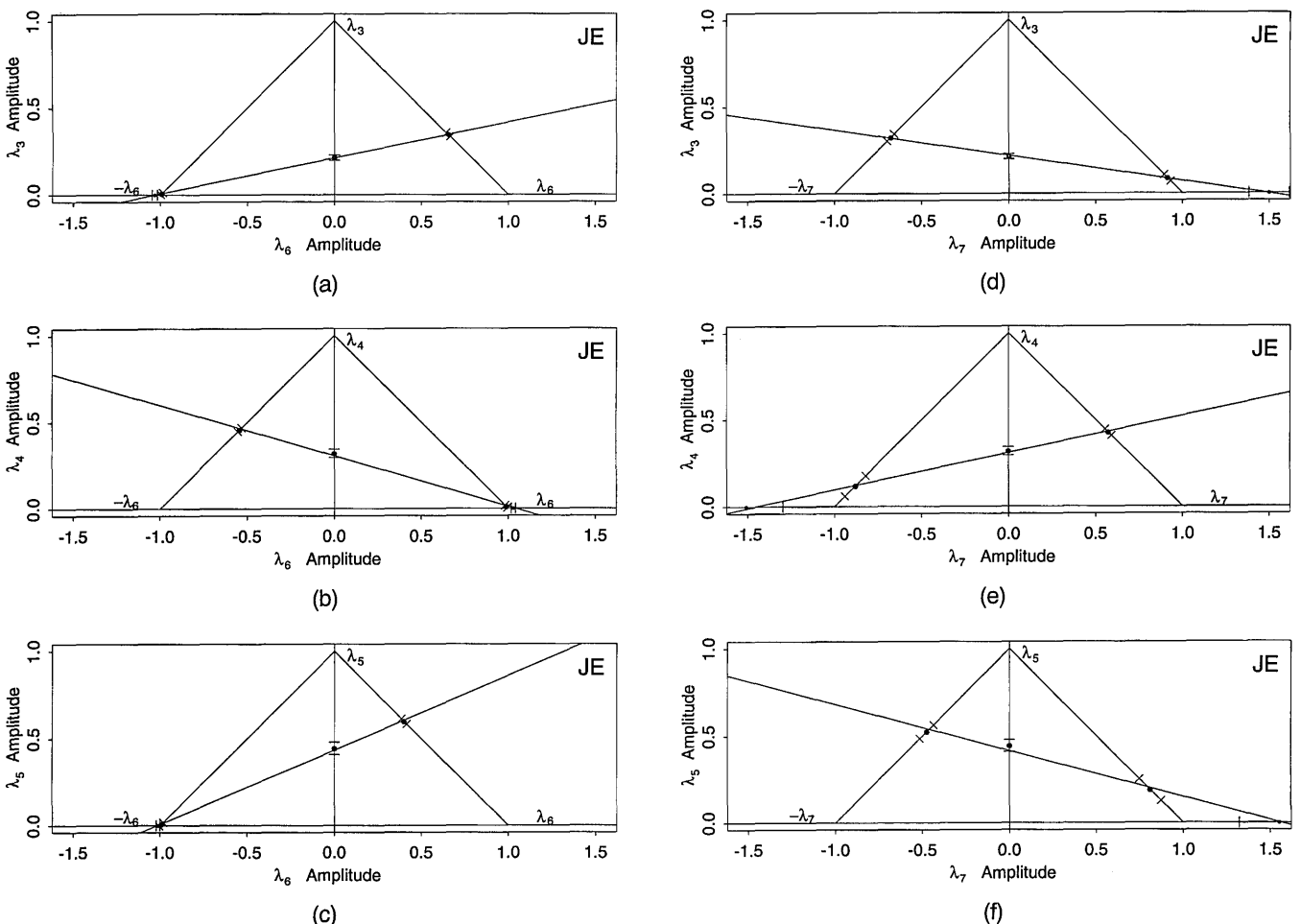


Fig. 8. Titration results for subject JE. Panels correspond to different titrations of each of  $\lambda_3$ ,  $\lambda_4$ , and  $\lambda_5$  with each of  $\lambda_6$  and  $\lambda_7$ . The number of trials per point and other statistics of interest are tabulated in Table 3. Intervals around the data points are 95% nonparametric confidence intervals. A more comprehensive statistic,  $p$ , refers to the ability of a straight line to model the data as a whole. The intervals around the  $x$  intercepts are standard deviations. (See text for details.)

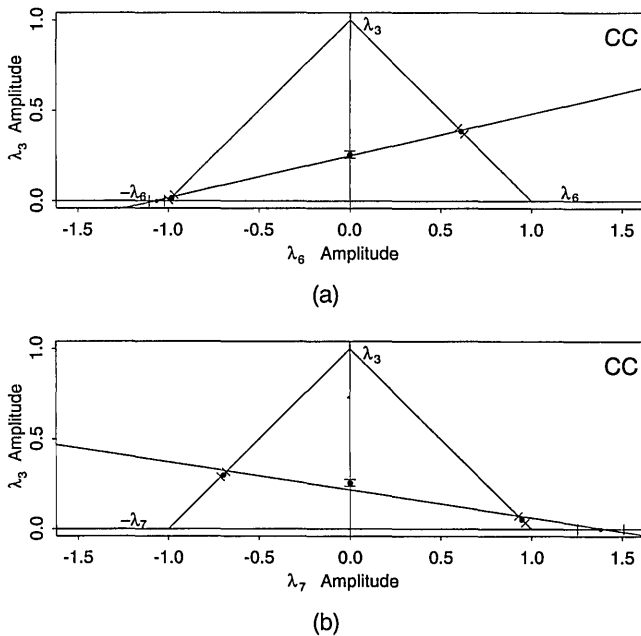


Fig. 9. Titration results for subject CC. Format as in Fig. 8.

ditive and multiplicative constants; see the definition in Subsection 2.D.) What is the relation between  $f_S$ , the peculiar seventh-order polynomial shown in Fig. 10, and  $T$ ? For  $d_T$ , the function reflecting the differential sensitivity of  $T$  to the micropatterns of  $\Gamma$ ,  $f_S$  is the projection of  $d_T$  into the space of functions spanned by orthogonal polynomials  $\lambda_3, \lambda_4, \dots, \lambda_7$ . Thus  $f_S$  does not tell us everything about  $d_T$ . In particular,  $f_S$  does not give us any indication about whether the measure  $T$  can detect differences in gray-level mean between textures. Neither does  $f_S$  give us any indication as to whether  $T$  is sensitive to differences in the gray-level variance between textures.  $f_S$  can tell us nothing about these aspects of  $T$ 's discriminating capabilities because all the textures in the space  $S$  have the same gray-level mean and variance. It may well be that  $T$  is sensitive to differences in gray-level mean and variance between textures. If we knew these sensitivities, we could add appropriate linear and quadratic terms to  $f_S$  to obtain a function that more fully characterized  $d_T$ .

We cannot measure the sensitivities of  $T$  to intertextural differences in gray-level mean and variance because other measures aside from  $T$  are also sensitive to such differences in gray-level mean and variance. This fact was established by pilot studies performed before the investigations reported in this paper. These pilot studies indicate that the space of all IID textures of  $\Gamma$  (not just the subset of this space for which mean and variance equal those of  $I_U$ ) is not perceptually elementary. Specifically, the cancellation tests discussed in Subsection 2.D failed (1) between  $\lambda_1$  and all of  $\lambda_2, \lambda_3$ , and  $\lambda_4$ , and (2) between  $\lambda_2$  and both of  $\lambda_3$  and  $\lambda_4$ .

How does the function  $f_S$  fit into the back-pocket model of texture segregation (Subsection 1.D)? We presume that the unique measure  $T$  relevant to the discrimination of textures of  $S$  is a single back-pocket channel whose linear filter is well suited to preserving the histogram differences between textures of  $S$  (such as that shown in Fig. 6). In this case  $f_S$  emerges as a partial character-

ization of the nonlinear function that is applied to the output of  $T$ 's linear filter.

An alternative possibility is that  $f_S$  is a partial characterization of an early pointwise nonlinearity, a nonlinearity that exists before any filtering of the sort performed by back-pocket model channels. The following argument makes this unlikely, however. Suppose that image intensity at each point were transformed by some function  $g: \mathbb{R} \rightarrow \mathbb{R}$  at a stage of visual processing before the linear filters of the back-pocket model. Suppose, moreover, that  $g = h + f_S$ , for some quadratic function  $h(v) = Av^2 + Bv$ . In this case consider what happens when we submit to the visual system a stimulus  $J$  composed of IID textures  $I_{U+\phi}$  and  $I_{U-\phi}$  with equal gray-level means and variances. The function  $g$  is applied point to  $J$ , yielding a neural output image  $g \circ J$  composed of IID textures  $g \circ I_{U+\phi}$  and  $g \circ I_{U-\phi}$ . Because  $I_{U+\phi}$  and  $I_{U-\phi}$  have equal gray-level means and variances,  $h \cdot \phi = 0$  (this is a consequence of proposition A.1 below). Note then that

$$E[g \circ I_{U+\phi}] = (h + f_S) \cdot (U + \phi) = (h + f_S) \cdot U + f_S \cdot \phi, \tag{55}$$

$$E[g \circ I_{U-\phi}] = (h + f_S) \cdot (U - \phi) = (h + f_S) \cdot U - f_S \cdot \phi. \tag{56}$$

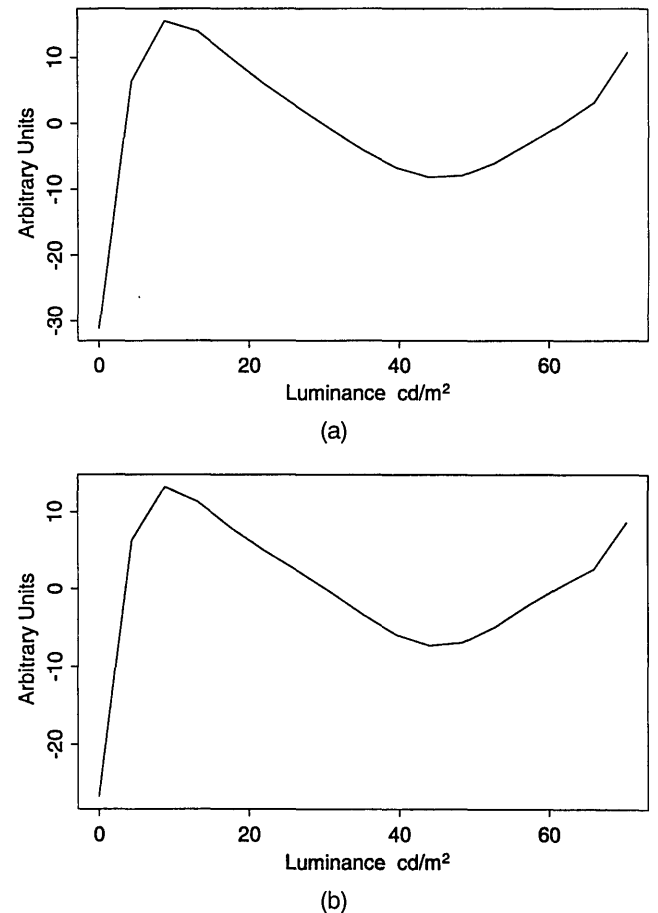


Fig. 10. Reconstructions of  $f_S$  for subjects (a) JE and (b) CC. These reconstructions include neither of components  $\lambda_1$  and  $\lambda_2$ ; these components cannot be measured by use of IID textures with equal mean and variance because measurement for such textures is no longer one dimensional (that is, the space of all IID textures is not perceptually elementary).

Thus, if  $f_S \cdot \phi \neq 0$ , then the neural IID textures  $g \circ I_{U+\phi}$  and  $g \circ I_{U-\phi}$  have different means. If  $g$  is the standard compressive nonlinearity sometimes asserted to modify the coding of light intensity early in the train of visual processing, then the polynomial expansion of  $g$  will certainly have substantial first- and second-order terms; that is,  $g$  will correlate significantly with both the linear modulator  $\lambda_1$  and the quadratic modulator  $\lambda_2$ . Thus, if the difference in the means of  $g \circ I_{U+\phi}$  and  $g \circ I_{U-\phi}$  is the source of our ability to distinguish the textures  $I_{U+\phi}$  and  $I_{U-\phi}$ , then we should be able to cancel the effect of modulation by  $\phi$  with a countermodulation by either  $\lambda_1$  or  $\lambda_2$ . That is, for some amplitude  $C$  we should find that  $I_{U+\phi+C\lambda_1}$  is indiscriminable from  $I_{U-(\phi+C\lambda_1)}$ . And similarly, for some amplitude  $D$  we should find that  $I_{U+\phi+D\lambda_2}$  is indiscriminable from  $I_{U-(\phi+D\lambda_2)}$ . However, the pilot studies discussed above showed that no such cancellation of  $\phi$  by either  $\lambda_1$  or  $\lambda_2$  could be obtained. We therefore conclude that the function  $f_S$  is probably not a component of an early pointwise, visual nonlinearity.

#### 4. SUMMARY

We have introduced a new psychophysical methodology called histogram contrast analysis for analyzing and estimating early stimulus measurements carried out by the visual system in drawing distinctions in the visual field. The basic techniques embraced by histogram contrast analysis are elaborated in Subsections 2.C–2.E and summarized in Subsection 2.F. Histogram contrast analysis can be used first to determine whether discriminations within a space  $S$  of IID textures of a given micropattern set  $\Omega$  (for definitions see Subsection 2.A.8.c) are consonant with the hypothesis that, among all the measures that the visual system uses in drawing visual distinctions, there is only a single such measure  $T$  that plays an effective role in drawing distinctions between textures in  $S$ . If this one-dimensionality hypothesis holds,  $S$  is said to be perceptually elementary. In this case, for  $d_T$ , the function reflecting the differential sensitivity of  $T$  to the micropatterns of  $\Omega$ , histogram contrast analysis can be used to produce a function  $f_S$  that precisely characterizes the portion of  $d_T$  that is relevant to discrimination between textures in  $S$ .

In Section 3 histogram contrast analysis was applied to a particular space of IID textures whose texels were drawn from a small set  $\Gamma$  of gray squares of various luminances. We concentrated on a space  $S$  of such textures for which the mean and the variance of the luminance across the texture were kept constant. It seemed likely (for reasons elaborated in Subsection 3.A) that such a space  $S$  would be perceptually elementary (Subsection 2.B.2). The method was applied successfully. For one subject it turned out that  $S$  was perceptually elementary, and the characteristic function  $f_S$  was estimated (Fig. 10). That  $S$  was perceptually elementary for this subject was confirmed despite a very large number of trials and several strong tests of that assumption. For a second subject the test that  $S$  was perceptually elementary failed in one of two cases, although the failure was relatively small.

For  $d_T$ , the function reflecting  $T$ 's differential sensitivity to the micropatterns in  $\Gamma$ , the characteristic function

$f_S$  is the portion of  $d_T$  that enters actively into the drawing of distinctions within  $S$ . Specifically, for functions  $\zeta_0$ ,  $\zeta_1$ , and  $\zeta_2$  defined by setting  $\zeta_0(v) = 1$ ,  $\zeta_1(v) = v$ , and  $\zeta_2(v) = v^2$ , proposition A.1 below implies that  $f_S$  is the projection of  $d_T$  into the space of all functions orthogonal to all of  $\zeta_0$ ,  $\zeta_1$ , and  $\zeta_2$ .

We argued (Subsection 3.E) that  $T$  is probably a specific back-pocket channel and that the obtained characteristic function  $f_S$  is a partial characterization of the nonlinear, pointwise transformation following the linear filter used by  $T$ .

Finally, it is important to note that the theoretical arguments made in this paper are quite general and may be applied to any sensory task for which analogous random textures may be made. Examples are spatial vision tasks with more complex texels, temporal visual textures (i.e., flicker; see Ref. 26), temporal acoustical textures (i.e., timbres or melodies), and spatial tactile textures.

#### APPENDIX A

This appendix concerns micropattern set  $\Gamma$ , studied experimentally in Section 3. The discussion presented here supports the assertion in Section 3 that all IID textures generated by the space of modulators spanned by  $\lambda_3, \lambda_4, \dots, \lambda_{16}$  relative to the uniform distribution  $U$  have equal mean and variance. Recall that the set  $\Gamma$  consists of small uniform squares of linearly increasing gray levels 0, 1,  $\dots$ , 16 (see Fig. 1). For convenience we identify the elements of  $\Gamma$  with their respective gray levels and write simply  $i$  for the square micropattern having intensity  $i$ .

Let  $\zeta_1$  and  $\zeta_2$  be the polynomial functions of  $\Gamma$  defined by

$$\zeta_1(i) = i, \quad \zeta_2(i) = i^2, \quad (\text{A1})$$

for  $i = 0, 1, \dots, 16$ .

Note that, for any distribution  $p$  of  $\Gamma$ ,

$$E[X_p] = \sum_{i=0}^{16} ip(i) = \zeta_1 \cdot p, \quad (\text{A2})$$

$$\begin{aligned} \text{var}[X_p] &= E[X_p^2] - E^2[X_p] = \sum_{i=0}^{16} i^2 p(i) - \left[ \sum_{i=0}^{16} ip(i) \right]^2 \\ &= \zeta_2 \cdot p - [\zeta_1 \cdot p]^2. \end{aligned} \quad (\text{A3})$$

We now state the main point of this appendix:

*Proposition A.1.* Let  $p$  be a probability distribution on  $\Gamma$ . The set  $A$  of all IID textures  $I_q$  such that  $E[X_q] = E[X_p]$  and  $\text{var}[X_q] = \text{var}[X_p]$  is equal to the set  $B$  of all IID textures  $I_{p+\phi}$  such that  $\phi$  is orthogonal to both  $\zeta_1$  and  $\zeta_2$ .

*Proof.* Suppose that  $I_q \in A$ . Then  $\phi = q - p$  is a modulator of  $p$  (since  $p + \phi = q$ ). Moreover,

$$\zeta_1 \cdot \phi = \zeta_1 \cdot (q - p) = \zeta_1 \cdot q - \zeta_1 \cdot p = E[X_q] - E[X_p]. \quad (\text{A4})$$

But because  $I_q \in A$ , the right-hand side of Eq. (A4) is 0. Thus  $\zeta_1 \cdot \phi = 0$ . Similarly,

$$\zeta_2 \cdot \phi = \zeta_2 \cdot (q - p) = \zeta_2 \cdot q - \zeta_2 \cdot p = E[X_q^2] - E[X_p^2]. \quad (\text{A5})$$



Because  $I_q \in A$ ,  $E^2[X_q] = E^2[X_p]$ . Thus the right-hand side of Eq. (A5) can be rewritten as

$$E[X_q^2] - E^2[X_q] - (E[X_p^2] - E^2[X_p]) \\ = \text{var}[X_p] - \text{var}[X_q] = 0, \quad (\text{A6})$$

implying that  $A \subseteq B$ .

On the other hand, suppose that  $I_{p+\phi} \in B$ . Then

$$E[X_{p+\phi}] = \zeta_1 \cdot (p + \phi) = \zeta_1 \cdot p + \zeta_1 \cdot \phi = \zeta_1 \cdot p \\ = E[X_p], \quad (\text{A7})$$

$$\text{var}[X_{p+\phi}] = \zeta_2 \cdot (p + \phi) - [\zeta_1 \cdot (p + \phi)]^2 \\ = \zeta_2 \cdot p + \zeta_2 \cdot \phi - (\zeta_1 \cdot p + \zeta_1 \cdot \phi)^2 \\ = \zeta_2 \cdot p - (\zeta_1 \cdot p)^2 = \text{var}[X_p]. \quad (\text{A8})$$

Hence  $I_{p+\phi} \in A$ , proving that  $B \subseteq A$ . ■

## ACKNOWLEDGMENTS

The research reported here was supported in part by National Science Foundation Human Cognition and Perception Program grant DBS-9203291 to Charles Chubb and in part by National Institutes of Health grant EY08266 and a National Research Council-NASA Ames Research Associateship (to Michael Landy). John Econopouly is supported by National Institute of Mental Health predoctoral fellowship MH10283. Portions of this work have been presented at conferences and workshops.<sup>27-30</sup>

Address correspondence to Charles Chubb, Department of Psychology, Rutgers University, New Brunswick, New Jersey 08903.

## REFERENCES

- G. T. Fechner, *Elements of Psychophysics*, D. H. Howes and E. C. Boring, eds., H. E. Adler, trans. (Holt, Reinhart & Winston, New York, 1966; originally published, 1860).
- Mach, E. *The Analysis of Sensations and the Relation of the Physical to the Psychical* (Dover, New York, 1959, with use of the fifth German edition revised and supplemented by S. Waterlow; first German edition 1886).
- H. B. Barlow, "Optic nerve impulses and Weber's law," *Cold Spring Harbor Symp. Quant. Biol.* **30**, 539-546 (1965).
- H. B. Barlow and W. R. Levick, "Three factors limiting the reliable detection of light by retinal ganglion cells of the cat," *J. Physiol.* **200**, 1-24 (1969).
- R. Shapley and C. Enroth-Cugell, "Visual adaptation and retinal gain controls," *Prog. Retinal Res.* **3**, 263-346 (1984).
- J. Müller, "Of the senses," in *Visual Perception: the Nineteenth Century*, W. N. Dember, ed. (Wiley, New York, 1964), pp. 35-69; recycled from the original English version published in *Elements of Physiology*, W. Baly, trans. (Taylor & Walton, London, 1942), Vol. 2.
- J. R. Bergen, "Theories of visual texture perception," in *Vision and Visual Dysfunction*, D. Regan, ed. (Macmillan, New York, 1991), Vol. 10, pp. 114-134.
- J. Beck, A. Sutter, and R. Ivry, "Spatial frequency channels and perceptual grouping in texture segregation," *Computer Vision Graphics Image Process.* **37**, 299-325 (1987).
- J. R. Bergen and E. H. Adelson, "Early vision and texture perception," *Nature (London)* **333**, 363-364 (1988).
- J. R. Bergen and M. S. Landy, "Computational modeling of visual texture segregation," in *Computational Models of Visual Processing*, M. S. Landy and J. A. Movshon, eds. (MIT Press, Cambridge, Mass., 1991), pp. 252-271.
- T. Caelli, "Three processing characteristics of visual texture segmentation," *Spatial Vision* **1**, 19-30 (1985).
- I. Fogel and D. Sagi, "Gabor filters as texture discriminator," *Biol. Cybern.* **61**, 103-113 (1989).
- N. Graham, "Complex channels, early local nonlinearities, and normalization in texture segregation," in *Computational Models of Visual Processing*, M. S. Landy and J. A. Movshon, eds. (MIT Press, Cambridge, Mass., 1991), pp. 273-290.
- M. S. Landy and J. R. Bergen, "Texture segregation and orientation gradient," *Vision Res.* **31**, 679-691 (1991).
- J. Malik and P. Perona, "Preattentive texture discrimination with early vision mechanisms," *J. Opt. Soc. Am. A* **7**, 923-932 (1990).
- A. Sutter, J. Beck, and N. Graham, "Contrast and spatial variables in texture segregation: testing a simple spatial-frequency channels model," *Percept. Psychophys.* **46**, 312-332 (1989).
- H. Knutsson and G. H. Granlund, "Texture analysis using two-dimensional quadrature filters," presented at the IEEE Computer Society Workshop on Computer Architecture for Pattern Analysis and Image Database Management, Pasadena, Calif., October 12-14, 1983.
- J. D. Victor, "Models for preattentive texture discrimination: Fourier analysis and local feature processing in a unified framework," *Spatial Vision* **3**, 263-280 (1988).
- R. L. De Valois and K. K. De Valois, *Spatial Vision* (Oxford U. Press, New York, 1988).
- N. Graham, *Visual Pattern Analyzers* (Oxford U. Press, New York, 1989).
- R. C. Fisher, *Linear Algebra* (Dickenson, Encino, Calif., 1970).
- U. W. Hochstrasser, "Orthogonal polynomials," in *Handbook of Mathematical Functions*, Vol. 55 of Applied Mathematics Series, M. Abramowitz and I. A. Stegun, eds. (National Bureau of Standards, Washington, D.C., 1972), Chap. 22, pp. 771-802.
- D. H. Brainard, "Calibration of a computer controlled color monitor," *Color Res. Appl.* **14**, 23-43 (1989).
- J. B. Mulligan and L. S. Stone, "Halftoning method for the generation of motion stimuli," *J. Opt. Soc. Am. A* **6**, 1217-1227 (1989).
- L. T. Maloney, "Confidence intervals for the parameters of psychometric functions," *Percept. Psychophys.* **47**, 127-134 (1990).
- M. Eckstein, J. S. Whiting, J. P. Thomas, and S. Shimozaki, "A novel temporal integration of intensity," in *Annual Meeting*, Vol. 23 of 1992 OSA Technical Digest Series (Optical Society of America, Washington, D.C., 1992), p. 17.
- C. Chubb and M. S. Landy, "Orthogonal distribution analysis: a systematic approach to the study of texture perception," *Invest. Ophthalmol. Vis. Sci. Suppl.* **31**, 561 (1990).
- C. Chubb and M. S. Landy, "Orthogonal distribution analysis: a new approach to the study of texture perception," in *Computational Models of Visual Processing*, M. S. Landy and J. A. Movshon, eds. (MIT Press, Cambridge, Mass., 1991), pp. 291-301.
- J. Econopouly, C. Chubb, and M. S. Landy, "Segregation of 1st-order noise textures," *Invest. Ophthalmol. Vis. Sci. Suppl.* **32**, 714 (1991).
- J. Econopouly, C. Chubb, and M. S. Landy, "Segregation of 1st-order noise textures is perceptually one-dimensional," *Invest. Ophthalmol. Vis. Sci. Suppl.* **33**, 956 (1992).