PERGAMON

Vision Research 40 (2000) 1695–1709

Vision Research

www.elsevier.com/locate/visres

Texture luminance judgments are approximately veridical

Jong-Ho Nam [a], Charles Chubb [b,*]

[a] Department of Psychology, Sungshim Campus, Catholic University of Korea, Puchon, South Korea
[b] Department of Cognitive Sciences, Institute for Mathematical Behavioral Sciences, University of California at Irvine, Irvine, CA 92697-5100, USA

Received 25 September 1998; received in revised form 27 September 1999

Abstract

This paper investigates intensity coding in human vision. Specifically, we address the following question: how do different luminances influence the perceived total luminance of a composite image? We investigate this question using a paradigm in which the observer attempts to judge, with feedback, which of two texture patches has higher total luminance. All patches are composed of nine luminances, ranging linearly from 0 (black) to a maximum luminance (white: 160 cd/m^2 in one condition; 20.2 cd/m^2 in another condition). Luminance histograms of the patches being compared are experimentally varied to derive, for each luminance v, the impact exerted by texture elements (texels) of luminance v on texture luminance judgments. We find that impact is approximately proportional to texel luminance; That is, a texture element exerts, on average, an impact on texture brightness (i.e. perceived texture luminance) that is proportional to its (the texel's) luminance. The only exception occurs for texels of maximal luminance, which surprisingly exert an impact that is slightly, but significantly, less than that exerted by texels of the next lower luminance. We conclude that visual intensity coding for purposes of assessing overall luminance of inhomogeneous patches is approximately veridical. In particular, texture luminance judgments are not mediated by a significant, compressive nonlinearity.
© 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Histogram contrast analysis; Texture; Brightness

1. Introduction

1.1. Background

A prevailing supposition is that adaptive gain control mechanisms play a critical role in photopic visual processing (e.g. Barlow, 1965; Barlow & Levick, 1969; Shapley & Enroth-Cugell, 1984; MacLeod & He, 1993; He & MacLeod, 1996, 1998). By the time the signal reaches the cortex, information about the mean luminance of the display being viewed has been largely adapted out of the signal transmitted. Accordingly, most models of cortical processing (e.g. of pattern discrimination, spatial localization, motion sensing, or texture segregation) currently assume that, given a stimulus S (i.e. a pattern of luminances imaged on the retina), the signal that reaches the cortex is approximated by a contrast modulation function; i.e. a function I_S of the following sort:

$$I_S(x, y, t) \approx \frac{S(x, y, t) - \text{local\_average}_S(x, y, t)}{\text{local\_average}_S(x, y, t)} \quad (1)$$

where local_average$_S$ (x, y, t) is the average luminance of S taken over all points (x', y', t') in some neighborhood of (x, y) during a brief time interval prior to t.

It should not be assumed, however, that the transformation of Eq. (1) is accomplished solely by the retina. On the contrary, Makous (1997) reviews the neurophysiological literature pertaining to photopic, retinal luminance coding and comes to the following conclusion: "Such neurophysiological evidence as is applicable to primate cone vision, then, shows no clear evidence of retinal multiplicative adaptation in the cone pathway of the parvo-ganglion cells that form some 80% of the output of the retina and carry the signals on which, some argue (e.g. Lennie, 1993), nearly all visual performance is based. At least such multiplicative adaptation is not noticeable until luminance levels exceed almost

* Corresponding author. Tel.: + 1-714-8241481; fax: + 1-714-8242517.
E-mail address: cfchubb@uci.edu (C. Chubb)

0042-6989/00/$ - see front matter © 2000 Elsevier Science Ltd. All rights reserved.
PII: S0042-6989(00)00006-7

unphysiological levels, where the cones themselves adapt." Thus, although subtractive gain control (such as that carried out by the numerator of Eq. (1) occurs in the retina, current evidence argues that the retina itself may not actually perform the divisive normalization embodied by the denominator in Eq. (1).

Of course, the precortical coding of intensity need not be perfectly linear. More generally, we imagine that intensities are coded for cortical processing by a system

$$I_S(x, y, t) \approx f\left(\frac{S(x, y, t) - \text{local\_average}_S(x, y, t)}{\text{local\_average}_S(x, y, t)}\right) \quad (2)$$

for f a strictly increasing (e.g. sigmoidal) function mapping $[-1, \infty)$ onto [0, 1).

Currently, most psychophysical researchers in the field of luminance-defined pattern perception, brightness perception and luminance-defined motion perception are willing to assume that their various target processes are preceded by an up-front transformation of the general form captured by Eq. (2).

It is well-known (Weber's law for luminance) that the threshold luminance increment required to detect a patch of light on a uniform background of luminance b is approximately proportional to b. (It should be noted, however, that Weber's law is found only after long term adaptation to luminance b, and holds well only for low frequency test stimuli (Yang & Makous, 1995).) The brighter the background, the greater the increment required to support threshold detection performance. Specifically, for α, the 'Weber fraction,' performance at detecting a patch of luminance $b + \delta$ on a background of luminance b is found to be at threshold when δ is approximately αb. Contemporaneously, Fechner (1860) and Maxwell (1860) both inferred from this result that the transformation mapping luminance onto perceived intensity involved a logarithmic compression: after all, if $\text{perceived\_intensity}(b) = \log(b)$, then consonant with Weber's law, $\text{perceived\_intensity}(b + \alpha b) - \text{perceived\_intensity}(b) = k$, for $k = \text{perceived\_intensity}(1 + \alpha)$. Maxwell proposed that this compression results from cone saturation. By contrast, Fechner thought that the cones were veridical (i.e. linear) luminance transducers, but that the mapping of retinal responses onto subjectively experienced intensity was compressively nonlinear.

However, Weber's law need not imply that luminance is transformed by an instantaneous nonlinearity, as Fechner and Maxwell both assumed (e.g. Luce & Edwards, 1958; Falmagne, 1971). Indeed, Weber's law is an obvious consequence of Eq. (2), provided the function f is approximately linear in the neighborhood of 0. Eq. (2), though, implies a fundamentally different coding scheme than an instantaneous, compressive nonlinearity. Note, in particular, that for a fixed value μ of $\text{local\_average}_S(x, y, t)$ (i.e. a fixed state of adaptation)

$$I_S(x, y, t) \approx f\left(\frac{S(x, y, t)}{\mu} - 1\right) \quad (3)$$

That is, for luminances $S(x, y, t)$ that are not too large in comparison to μ (and hence in the linear domain of f), the model of Eq. (2) implies that the subcortical representation of intensity is approximately a linear function of $S(x, y, t)$ (not a logarithmic compression).

1.2. The goal of the current experiments

The experiments reported here attempt to shed light on visual intensity coding by investigating performance in a novel luminance discrimination task. In this task, given a single, brief display, the observer is required to judge which of two texture patches has greater total luminance. The patches are composed of many small square texture elements (texels), each of which is painted with one of nine, linearly increasing luminances, spanning a broad range. Thus, the judgment requires the observer to additively combine the subjective intensity representations produced by the component texels. We use the methods of histogram contrast analysis (Chubb, Econopouly & Landy, 1994; Chubb, 1999) to measure, for each luminance, v, the average impact $m(v)$ exerted by v on texture brightness (i.e. perceived texture luminance).

What is the concrete meaning of the texture brightness function m? It should be noted first that our measurement methods allow us to determine m only up to an arbitrary positive scale factor and an arbitrary additive constant. In other words, we cannot determine the mean value of m; nor can we determine the absolute amplitude of m's deviation from its mean. What we can determine are the relative deviations of all of m's values from m's mean value.

In order to explain the convention we use to scale m, it will be useful to consider the following hypothetical texture brightness judgment: imagine a texture patch $Patch_1$ containing equal numbers of all nine luminances, randomly permuted within a rectangular region. Such a texture patch is said to have a uniform histogram. Now imagine producing another texture patch $Patch_2$, also with a uniform histogram, and then replacing a randomly chosen texel in $Patch_2$ with a texel of luminance v to produce a new patch $Patch_{2,\text{Altered}}$. Then $m(v)$ reflects the expectation of the difference in the brightness of $Patch_{2,\text{Altered}}$ compared to $Patch_1$.

The magnitude of $m(v)$ can be understood in terms of a hypothetical experiment in which the observer is repeatedly asked to judge which is higher in luminance, $Patch_1$ versus $Patch_{2,\text{Altered}}$ (where the construction of each patch is carried out independently on each trial). Suppose that $m(v) = 0.01$. This indicates that the alteration produces, on average, an increase in patch brightness equal to 0.01 standard deviations of the total noise by which the observer's comparisons of patch brightness are degraded. Thus, if $m(v)=0.01$, then the observer will judge $Patch_{2,\mathrm{Altered}}$ more luminant than $Patch_1$ with probability slightly greater than 0.5. Specifically, for Φ the standard normal cdf, the observer will judge $Patch_{2,\mathrm{Altered}}$ more luminant than $Patch_1$ with probability $\Phi(0.01)=0.504$. On the other hand, if $m(v)=-0.023$, then the observer will judge $Patch_{2,\mathrm{Altered}}$ more luminant than $Patch_1$ with probability less than 0.5, specifically, with probability $\Phi(-0.023)=0.491$.

1.3. What is the relationship between the texture brightness impact function m and the function f (of Eq. (2))?

The relationship between subcortical intensity coding and the texture brightness impact function m may well be complex. It would be ill advised simply to assume that m must be one and the same as the function f of Eq. (2). Indeed, it is not hard to imagine scenarios in which m differs in important ways from f. For example, suppose that the observer attempts to judge texture luminance by summing the activations produced by a certain array of neurons in the cortex whose input is given by Eq. (2). Unless the responses of the neurons in this cortical array depend linearly on their inputs, m will surely differ in form from f. However, there is a wealth of evidence (e.g. simultaneous contrast and related effects) to suggest that the brightness of a homogeneous region (such as a texel) depends in complex, nonlinear ways on the context (i.e. the surrounding constellation of luminances) in which that region occurs (e.g. Grossberg & Todorovic, 1988). Thus, we must be alert to the possibility that the form of m may well depend not merely on f but also on nonlinear interactions between neurons in the hypothetical array mediating texture brightness judgments. We shall take up this issue in greater detail in Section 4.

1.4. List of symbols

For the reader's convenience, we include the following listing of the symbols we use in this paper. Most of the terms in this list have yet to be defined. A reasonable strategy is to skip past this list on initial reading and refer back to it as necessary.

G	the set of luminances of which texture patches will be composed.
v	a luminance in G.
τ	a texel (i.e. a small, rectangular region in a texture patch to be painted with some luminance of G).
q, r	texel distributions (i.e. probability distributions on G).
P_q	an IID texture patch with texel distribution q. Thus, the luminances of the texels of P_q are jointly independent random variables all with distribution q.
$P_q(\tau)$	the luminance (from G) assigned by patch P_q to texel τ. Thus, $P_q(\tau)$ is a random variable with distribution q.
$\mathrm{dom}(P_q)$	the set of all texels of which P_q is composed.
$L(P_q)$	The subjective estimate of the luminance of IID texture patch P_q (Eq. (4)). ($L(P_q)$ is a random variable.)
Y	A normal random variable with mean 0 and standard deviation σ (Eq. (4)).
$\Delta(q, r)$	The random variable that determines the observer's decision as to which is more luminant, P_q vs. P_r (Eq. (4)).
$X_q(\tau)$	A random variable produced by texel $\tau \in dom(P_q)$ that will additively influence $L(P_q)$.
F_v	The cdf of random variable $X_q(\tau)$, given that $X_q(\tau)=v$ (Eq. (5)).
ξ_v	A random variable with cdf F_v.
$m(v)$	The expectation of the impact exerted on texture brightness by an occurrence of a texel of luminance v (Eq. (6)).
m	The function $m: G \to \mathbb{R}$ is called the (texture brightness) impact function.
$s(v)$	The standard deviation of the impact exerted on texture brightness by an occurrence of a texel of luminance v (Eq. (6)).
s	The function $s: G \to \mathbb{R}$ is called the (texture brightness) noise injection function.
s^2	The square of the noise injection function s. That is, $s^2(v)=(s(v))^2$ for all $v \in G$.
$f \cdot g$	for any functions $f: G \to \mathbb{R}$ and $g: G \to \mathbb{R}$, $f \cdot g = \Sigma_{v \in G} f(v)\, g(v)$
W_{P_q}	A function mapping $dom(P_q)$ into $\mathbb{R}$ reflecting the different weights (e.g. due to nonhomogeneous allocation of attention) with which the random variables $X_q(\tau)$ are combined to produce the patch luminance estimate $L(P_q)$.
K_1	Sum of $W_{P_q}(\tau)$ over all texels τ in $dom(P_q)$ (Eq. (8)).
K_2	Sum of $W^2_{P_q}(\tau)$ over all texels τ in $dom(P_q)$ (Eq. (9)).
U	The uniform texel distribution on G (Eq. (11)).
θ (or θ_i)	A reversible U-modulator (Sec. 1.7.1).
A	The component of $var(\Delta(U+\theta, U-\theta))$ that does not depend on m (Eq. (13)).
B	Model parameter governing the degree to which $var(\Delta(U+\theta, U-\theta))$ depends on m (Eq. (13)).

1.5. IID textures

The methods used here are a variant of those used by Chubb et al. (1994). In the experiments to be described, the observer is asked on each trial to judge which of two abutting, rectangular texture patches (the right patch or the left patch) has higher total luminance.

All of the textures we use will be generated from a set G comprising nine luminances, ranging linearly from black to white. We use exclusively IID texture patches, i.e. patches in which the intensities of all texture elements are jointly independent, identically distributed, random variables. The probability distribution from which texel intensities are drawn to generate a given patch is called the texel distribution of the patch.

It will be convenient to write P_q for an IID texture patch with texel distribution q. For any texel τ in the patch, we shall write $P_q(\tau)$ for the luminance assigned texel τ in P_q. Thus, to generate P_q one might load tokens of different intensities into an urn, such that, for any possible luminance v, the proportion of tokens in the urn of luminance v is $q(v)$. Then intensities could be assigned to the texels of P_q by successive draws (with replacement) from the urn. Note that the luminance histogram of P_q will be likely to resemble q in form, but will almost certainly differ from q in random detail.

Fig. 1 shows examples of IID textures composed from G. Within each of the IID texture patches of Fig. 1A–F is a bar graph that shows the texel distribution used to generate the patch. The luminances of G are arranged in increasing order along the bottom of the bar graph. The height of each bar above a given luminance v shows the probability that a given texel will be assigned luminance v. Note that the texel distribution of Fig. 1A has a higher mean luminance than that of Fig. 1B, yielding an obvious difference in brightness between the two patches. Evidently texture brightness is determined largely by texture mean luminance; however, other, higher order moments of the texel distribution may also influence patch brightness. The current experiment measures these secondary influences. The IID patches in Fig. 1C–D are equal in expected texel luminance, but differ in variance, leading to a difference in the apparent contrast of the two textures. The patches in Fig. 1E–F are equal in both expected texel mean and variance; however, they differ in their higher central moments. The fact that there is no very obvious difference in texture brightness between any of the patches 1C,D,E and F adumbrates the main result of this paper: texture luminance judgments depend approximately linearly on texel luminance.

1.6. Model assumptions

In attempting to judge whether patch P_q is more luminant than patch P_r, the observer is assumed to construct (noisy) subjective patch luminance assessments $L(P_q)$ and $L(P_r)$. Then, for Y a normal random variable with mean 0 and standard deviation σ, the observer is assumed to judge P_q more luminant than P_r iff

$$\Delta(q, r) = L(P_q) - L(P_r) + Y > 0. \tag{4}$$

The random variable Y occurring in Eq. (4) is intended to capture trial-to-trial variability that is independent of the texel distributions q and r. The standard deviation σ of Y might depend on a broad range of factors, including the size and shape of a patch, the specific luminances in G, the viewing distance of the observer, etc.

The strongest model assumptions concern the production of random variables $L(P_q)$ and $L(P_r)$. For any texel τ of P_q, the micropattern $P_q(\tau)$ is assumed to generate a random variable $X_q(\tau)$. (These texel-generated random variables will be linearly combined to generate $L(P_q)$.) The random variables $X_q(\tau)$, for all texels τ of P_q, are assumed to be jointly independent and identically distributed.

In addition, for any given luminance v, the distribution of the random variable $X_q(\tau)$ generated by painting texel τ with luminance v is assumed invariant with respect to q. To frame this assumption more precisely, we need the following notion.

1.6.1. Definition (of the conditional cdf F_v)

For any distribution q on G, any texel τ, and any luminance $v \in G$, the conditional cdf of $X_q(\tau)$ given that $P_q(\tau) = v$ is defined by

$$F_v(x) = \mathbb{P}[X_q(\tau) \leq x \mid P_q(\tau) = v]. \tag{5}$$

As suggested by the notation used in Eq. (5), it is assumed that F_v does not depend upon q. In particular, this means that the functions

$$m(v) = E[X_q(\tau) | P_q(\tau) = v] \text{ and}$$
$$s(v) = \text{std\_dev}[X_q(\tau) | P_q(\tau) = v] \tag{6}$$

are assumed invariant with respect to texel distribution q. For any $v \in G$, $m(v)$ determines the expected impact on $L(P_q)$ exerted by an occurrence in patch P_q of a texel of luminance v. Accordingly, m is called the (texture brightness) impact function. On the other hand, $s(v)$ determines the amount of variability injected, on average, into $L(P_q)$ by a texel of luminance v. Thus, s is called the (texture brightness) noise injection function.

It is important to note that, although $L(P_q)$ is assumed to be a linear function of the random variables $X_q(\tau)$, the expectations of these random variables are not necessarily assumed to be linearly related to the corresponding luminances $P_q(\tau)$. In other words, the impact function m is not required to be linear.

As discussed elsewhere (Chubb, 1999), it is not plausible to assume that the conditional cdf's $F_v, v \in G$, are invariant with respect to texel distribution across the set of all IID textures; however, this assumption is likely to hold approximately across a family of textures whose texel distributions all differ only slightly from the uniform distribution U (See Eq. (11)), as is the case in the current experiment.

We assume now that $L(P_q)$ is produced by additively

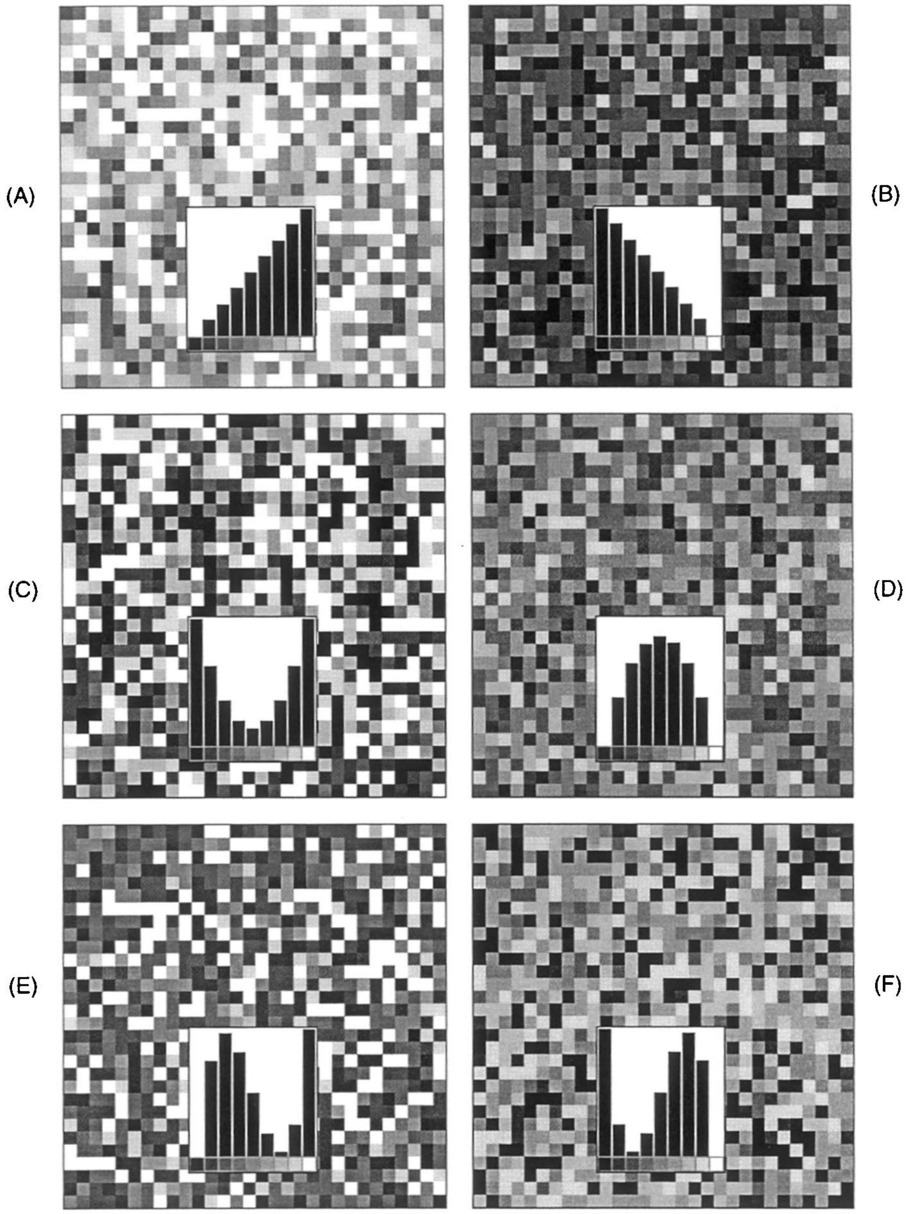

Fig. 1. Examples of visual IID texture patches. The set G used to generate the IID texture patches shown here contains nine small squares of (approximately) linearly increasing luminance. (A) A patch of IID texture containing $N_{\mathrm{Tex}} = 30 \times 30 = 900$ texels. The bar graph enclosed in the patch shows the texel distribution used to generate the patch. The luminances of G are arrayed in increasing order along the bottom of the bar graph. The height of the black bar above each luminance v is proportional to the probability that a given texel in the patch will be assigned luminance v. Thus, the texel distribution of Fig. 1A increases linearly with luminance. (B) A patch of IID texture whose texel distribution decreases linearly with luminance. (C) A patch of IID texture whose texel distribution varies quadratically with texel luminance. (D) A patch of IID texture whose texel distribution varies quadratically, but inversely to that of (C); (E) and (F) are patches of IID texture whose texel distributions have equal mean and variance, but which differ in their third and higher order, odd moments. The fact that there is no obvious difference in overall brightness between any of the patches C, D, E and F adumbrates the main result of this paper: texture luminance judgments depend approximately linearly on texel luminance.

combining the random variables $X_q(\tau)$. However, we admit the possibility that the random variables produced by texels in different locations may be combined with different weights (e.g. due to inhomogeneous allocation of attention across patch P_q). Specifically, for some function W:texels of $P_q \to \mathbb{R}$, it is assumed that

$$L(P_q) = \sum_{\text{all texels } \tau \text{ of } P_q} W(\tau) X_q(\tau). \tag{7}$$

The function W is called the texel weighting function for patch P_q. Although it is natural to suppose that W is everywhere nonnegative, the theory does not require this assumption. It is, however, required that the sum of W over all texels of P_q be positive.

Two additional assumptions (captured in Eqs. (8) and (9)) are required to effectively insure that the texel weighting functions for patches P_q and P_r are comparable in their spatial sampling properties. (These assumptions are very reasonable in the current experiment because the locations (left versus right) of patches P_q and P_r are randomly assigned from trial to trial; in essence, the patches P_q and P_r probabilistically share the same spatial locations.) For W_{P_q} and W_{P_r} the texel weighting functions for patches P_q and P_r, it is assumed that

$$K_1 = \sum_{\text{all texels } \tau \text{ in dom}(P_q)} W_{P_q}(\tau) = \sum_{\text{all texels } \tau \text{ in dom}(P_r)} W_{P_r}(\tau) \tag{8}$$

and also that

$$K_2 = \sum_{\text{all texels } \tau \text{ in dom}(P_q)} W^2_{P_q}(\tau) = \sum_{\text{all texels } \tau \text{ in dom}(P_r)} W^2_{P_r}(\tau). \tag{9}$$

1.7. Measuring the texture brightness impact function

For any functions f:$G \to \mathbb{R}$, g:$G \to \mathbb{R}$, we write $f \cdot g$ for the cross-correlation of f with g:

$$f \cdot g = \sum_{v \in G} f(v) g(v). \tag{10}$$

In the experiments to be described the texel distributions of the textures compared on a given trial are manipulated to infer the form of the texture brightness impact function m. As shall be shown, these methods yield results that are invariant with respect to all unmeasured model parameters (including the standard deviations $s(v)$, $v \in G$). The following notion is required to describe the experimental manipulations used to measure the impact function m.

1.7.1. Definition of a reversible modulator

Let h be a texel distribution. Then any function θ:$G \to \mathbb{R}$ is called an h-modulator if $h + \theta$ is a texel distribution. If $h - \theta$ is also a texel distribution, then θ is called reversible.

In the experiments to be described, the texel distributions of patches being compared are modulated away from the uniform distribution U:

$$U(v) = \frac{1}{9}, \quad v \in G. \tag{11}$$

In particular, our methods exploit the following result, derived in the Appendix. Let Φ denote the standard normal cdf. Then under the model assumptions described above, for any U-modulator θ, the probability $p(\theta)$ that the observer judges $P_{U+\theta}$ more luminant than $P_{U-\theta}$ is given by

$$p(\theta) = \Phi\left(\frac{m \cdot \theta}{\{A + B(m^2 \cdot U - (m \cdot \theta)^2)\}^{1/2}}\right) \tag{12}$$

for

$$B = \frac{K_2}{2K_1^2} \quad \text{and} \quad A = B\left(s^2 \cdot U + \frac{\sigma^2}{2K_2}\right). \tag{13}$$

The argument to Φ in Eq. (12) is a signal-to-noise ratio; the numerator is proportional to the mean, and the denominator to the standard deviation of the random decision variable $\Delta(U + \theta, U - \theta)$ (See Eq. (4)). Note that the term A in the denominator captures the portion of the variance that does not depend on m. For purposes of comparing impact functions m across observers, it will be convenient to express m in multiples of $\sqrt{A}$. For m expressed in this way, Eq. (12) becomes

$$p(\theta) = \Phi\left(\frac{\sqrt{A} m \cdot \theta}{\{A + B((\sqrt{A} m)^2 \cdot U - ((\sqrt{A} m) \cdot \theta)^2)\}^{1/2}}\right) = \Phi\left(\frac{m \cdot \theta}{\{1 + B(m^2 \cdot U - (m \cdot \theta)^2)\}^{1/2}}\right). \tag{14}$$

2. Experimental procedure

Our strategy for measuring m is simple. We select a number of U-modulators $\theta_1, \theta_2, \ldots, \theta_{\text{Nconds}}$, defining our different experimental conditions. On a given experimental trial, for some condition $i = 1, 2, \ldots$, Nconds, the observer is required to judge (with audible feedback on each trial) which of patches $P_{U+\theta_i}$ vs. $P_{U-\theta_i}$ is more luminant. In all cases θ_i is chosen so that $U - \theta_i$ is slightly more luminant than $U + \theta_i$. However, all conditions are constructed so that the observer is inconsistent to some degree in performing this task (neither always correct, nor always incorrect).

For each condition i, some number of trials are conducted, yielding k_i correct and n_i incorrect responses. The total number of trials in condition i is thus $n_i + k_i$. This number differs across conditions and from observer to observer, with more informative conditions receiving more trials.

The resulting data are processed to obtain a maximum likelihood estimate of the texture brightness impact function m. Specifically, the value of B as well as those of $m(v)$, $v = 0, 1, \ldots, 8$ are found so as to maximize the likelihood function

$$\Lambda(m, B) = \prod_{i=1}^{\text{Nconds}} p(\theta_i)^{k_i} (1 - p(\theta_i))^{n_i}, \tag{15}$$

where $p(\theta_i)$ is defined by Eq. (14). Matlab code to accomplish the model fitting is available on the web at http://texel.ss.uci.edu/ChubbAddenda.html.

2.1. Apparatus

The experiment was conducted for two observers using a mean luminance of 80 cd/m^2, and for two observers using a mean luminance of 10 cd/m^2. The monitors differed for these two luminance conditions. For the high mean luminance condition, an Ikegami DM516A monochrome monitor was used. For the low mean luminance condition, a TVM high-resolution monochrome monitor, model MG-11P U03 was used. In both cases, the stimuli were presented using an IBM-compatible computer equipped with an ATVista graphics system.

2.2. Observers

Two of the authors (JN and CC) and one additional experienced psychophysical observer (JM) were used in this experiment. Two observers (JN and CC) had corrected-to-normal vision. The other (JM) had normal vision. One observer, JN, participated in both the high luminance and low luminance experiments. CC participated only in the high luminance experiment; JM participated only in the low luminance experiment.

2.3. Stimuli

All stimuli were composed of small square texels, painted with one of 9, linearly increasing intensities. For the high mean luminance condition, these intensities were $20i$ cd/m^2, $i = 0, 1, \ldots, 8$. For the low mean luminance condition these intensities were $2.5i$ cd/m^2, $i = 0, 1, \ldots, 8$. A stimulus in a given trial comprised two abutting texture patches (one on the left, and one on the right), each composed of 30 rows by 15 columns of texels. Each texel (5×5 pixels) subtended 6.68 min^2. The viewing distance in the high (low) luminance condition was 136 (75) cm.

2.4. Conditions

On a given trial, the observer was asked to judge which of the component IID texture patches $P_{U+\theta_i}$ vs. $P_{U-\theta_i}$ had greater total luminance. The specific modulators used (as well as the rest of the raw data) are available on the world wide web at texel.ss.uci.edu/ChubbAddenda.html. There were 120 conditions (modulators) in all.

These modulators are selected to span the space of all candidate impact functions. Additionally, each modulator θ_i, is constructed so that, in practice, the total physical luminance of $P_{U+\theta_i}$ is practically certain to be greater than the total luminance of $P_{U-\theta_i}$. Thus, on a trial in which the subject is required to judge which of $P_{U+\theta_i}$ vs. $P_{U-\theta_i}$ has greater luminance, the correct response is $P_{U+\theta_i}$. Finally, all conditions are chosen so that performance is a priori likely to be neither perfect, nor at chance.

Different observers collected different amounts of data in different conditions. In the low luminance experiment, observers JN and JM each performed a minimum of 100 trials in each condition. In the high luminance experiment, JN performed a minimum of 130 trials in each condition, while CC performed a minimum of 82 trials in each condition. Thereafter, additional trials were run only in the most informative conditions (conditions yielding performance in the steep region of the psychometric curve). The experiment was conducted in blocks of 240 trials. In those blocks in which all 120 conditions were mixed, each condition occurred twice, once with patch $P_{U+\theta_i}$ on the right and once on the left. Trials were randomly sequenced. In those blocks focusing on informative conditions, a minimum of 30 conditions (each receiving 8 trials) were randomly mixed. In all, JM (low luminance) ran 18,866 trials; JN (low luminance) ran 16,800 trials; CC (high luminance) ran 19,312 trials; JN (high luminance) ran 27,340 trials.

2.5. The structure of a trial

A trial consisted of the following steps: the observer fixated a small central cue spot, then initiated a trial with a button press. The cue spot disappeared, and the pair of patches was presented for 33 ms, and then replaced by the cue spot again. The observer then was asked to report (with a button press) which of the two patches had the higher total luminance. Following the observer's response, an auditory signal was given if it was wrong.

2.6. Linearization

For the low-luminance experiment, linearization was achieved using a by-eye procedure in which a regular grid of texture elements containing three intensities lum_{lo}, lum_{hi} and lum_{mid} (half with luminance lum_{mid}, 1/4 with lum_{lo} and 1/4 with lum_{hi}) was made to alternate in a coarse vertical square-wave with texture comprising a checkerboard of texture elements alternating between intensities lum_{lo} and lum_{hi}. The screen was then viewed from sufficiently far away that the fine granularity of the texture was barely visible. At this distance, the square-wave modulating between the two types of texture had a spatial frequency of approximately 4 c/deg. Since the texture itself could not be resolved, the square-wave is visible only if the mean luminance of alternating texture bars is different. Thus, the luminance lum_{mid} that makes the square-wave vanish is equal to the average of the intensities lum_{lo} and lum_{hi}. We generated a lookup table by reiterating this procedure with different luminances lum_{lo} and lum_{hi} to determine, in each case, the lum_{mid} midway between lum_{lo} and lum_{hi}.

One might wonder why we use a three-luminance checkerboard rather than a uniform field of luminance lum_{mid} in our linearization display. On most monitors, two horizontally adjacent bright pixels on a black background emit more than twice the light emitted by a single bright pixel. More generally, the amount of light emitted by a block of pixels cannot, typically, be derived from the individual amounts of light emitted by pixels turned on in isolation. In other words, monitors often manifest spatial display nonlinearities. What is important in the current experiment is that the amounts of light emitted by square pixel-blocks of a particular fixed size be linearized. The three-luminance checkerboard was used rather than a uniform field of luminance lum_{mid} in order to achieve this goal. In the three-luminance checkerboard, each check of luminance lum_{mid} was (i) equal in size to the checks used in our experimental textures, and (ii) surrounded on all four adjacent sides by squares of luminance either lum_{lo} or lum_{hi}.

A similar by-eye procedure, with a somewhat different test stimulus, was used to linearize the high-luminance display. In this case, we were able to ascertain that the monitor (an Ikegami DM516A monochrome monitor) was remarkably free from spatial nonlinearities of the sort discussed above. Accordingly, we used a linearization stimulus in which uniform bars of luminance lum_{mid} were made to alternate in a coarse vertical square-wave with texture comprising a checkerboard of texture elements alternating between intensities lum_{lo} and lum_{hi}.

3. Results

Raw data are available on the world wide web at http://texel.ss.uci.edu/ChubbAddenda.html.

Maximum likelihood estimates of parameter B (Eq. (14)), as well as parameters $m(v)$, $v \in G$, were obtained using the Matlab program Fit_m_IID.m also available at http://texel.ss.uci.edu/ChubbAddenda.html. Estimated values of B were near 0 in all cases: for CC, $\hat{B} = 0.076$; for JM, $\hat{B} = 0.034$; for JN in both the high and low luminance conditions, $\hat{B} = 0.000$. These small values of $\hat{B}$ indicate that the variance of the decision variable $\Delta(U+\theta, U-\theta)$ depended very little on m. Indeed, for CC (for whom $\hat{B}$ was the largest), only 0.0007% of $var[\Delta(U+\theta, U-\theta)]$ depended on m, on average. Thus, for all observers, this term could have been omitted from the model without significantly altering the estimated impact functions, m.

Estimated (maximum likelihood) texture brightness impact functions are plotted in Fig. 2. The top (bottom) two boxes of Fig. 2 show the impact functions for low (high) mean luminance textures. (Error bars give bootstrapped 95% confidence intervals for the estimated impact function values.)

Thus, for example, the fact that the texture brightness impact function for observer JM assigns the value -0.012 to luminance 0 indicates that an occurrence of a texel of luminance 0 in a patch has (on average) the effect of decreasing the brightness of the patch by 0.012 standard deviations of the total noise compromising JM's judgments.

The texture brightness impact functions are similar for all observers, irrespective of the mean luminance of the display. Note in particular that there are no dramatic differences between the estimates obtained in the low versus high luminance conditions. In all cases, the impact exerted on patch brightness increases as an approximately linear function of texel luminance. The only significant departure from this pattern occurs at the highest luminance: for all observers in all conditions, the impact on patch brightness exerted by texels of the highest luminance is less than or equal to the impact exerted by texels of the next-to-highest luminance.

This pattern emerges very clearly in Fig. 3, in which we have taken the liberty of fitting our model to the pooled data from our four observers. Although we use this curve to summarize the main patterns in our data, it should be noted that there are indeed statistically significant differences across conditions and between observers. In particular, the impact function produced by JN in the low luminance condition differs markedly from the other three curves at the lowest luminance. Indeed, an infinitesimal p-value is derived from a likelihood ratio test comparing the fit provided by the ten parameters used to generate Fig. 3 (the nine values of m shown in the figure, plus the estimate of B, which was 0.000) to the 40 parameter fit obtained by modelling the data of each observer separately. This shows that at least some of our four observers differed significantly in

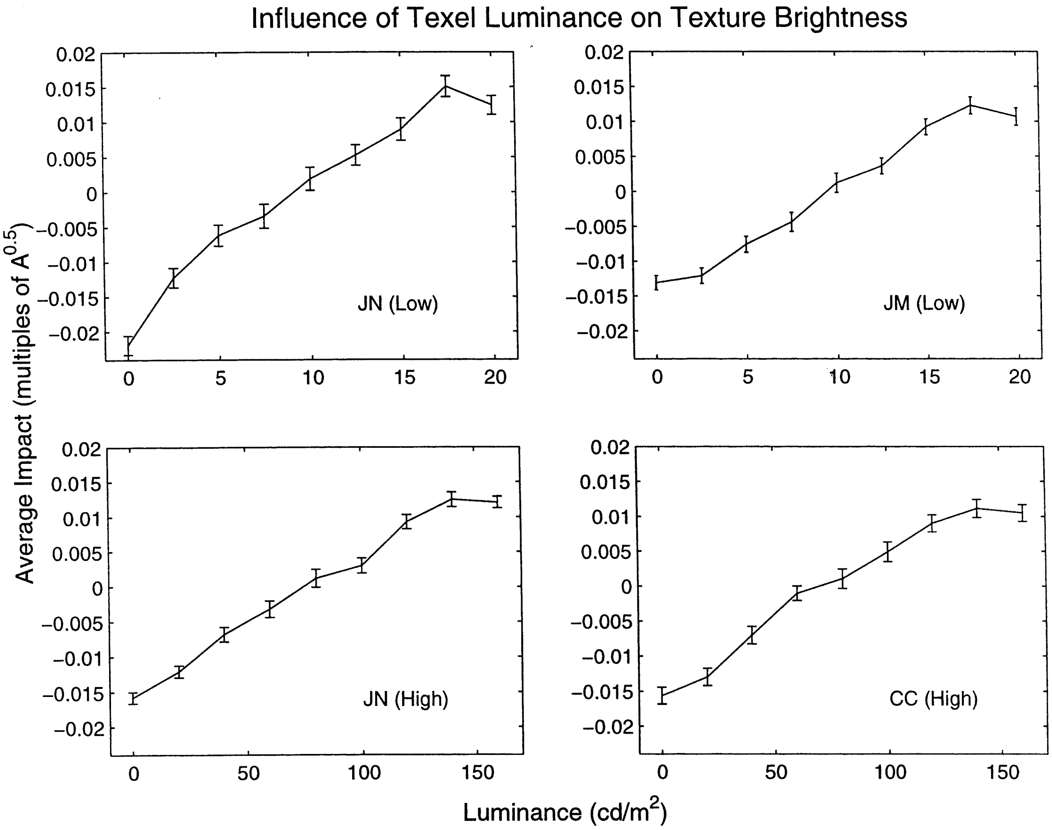

Fig. 2. Texture brightness impact functions. The top two boxes of Fig. 2 show the impact functions for low (high) mean luminance textures. Error bars give bootstrapped 95% confidence intervals for the estimated impact function values. Thus, for example, the fact that the texture brightness impact function for observer JM assigns the value -0.012 to luminance 0 indicates that the effect of substituting a randomly chosen texel in a patch with a uniform histogram by a texel with luminance 0 is to lower patch brightness by $0.012A^{0.5}$, for $A^{0.5}$ defined by Eq. (13); $A^{0.5}$ is the standard deviation of that component of the noise degrading the observer's judgments that does not depend on the texture brightness impact function. Model fits show that the component of the noise that does depend on the impact function is negligible. Thus $A^{0.5}$ is effectively the standard deviation of the total noise degrading judgments.

their response patterns. In particular, although the curves obtained in the high versus the low luminance conditions appear quite similar, our data reject the hypothesis that the curves were generated by the same process; thus, we cannot conclude that the texture brightness impact function is independent of stimulus mean luminance.

The curve of Fig. 3 rises almost perfectly linearly over the lower eight luminances. Interestingly, the curve seems to show a slight nonmonotonicity for the highest luminance. Texels of the highest luminance exert slightly but significantly less impact on patch brightness than do texels of the next lower luminance. This same effect is to be seen in the data of both JN (low luminance condition) and JM (low luminance condition); in each case, the estimated impact exerted on patch brightness by texels of the highest luminance falls below the 95% confidence interval for the estimated impact exerted by texels of the next lower luminance.

4. Discussion

4.1. The possible influence of context effects

Our model assumes that individual texels in a patch generate jointly independent random variables that are combined in a weighted sum to estimate patch luminance. However, the assumption of independence is almost certainly wrong. Consider a texel τ of luminance v: phenomena such as simultaneous contrast suggest that the context in which τ occurs (i.e. the constellation of luminances surrounding τ) may play an important role in determining τ's brightness. Plausibly, in the randomly scrambled textures used in the current experiments, only τ's very local context differentially influences τ's brightness. One must remember, however, that within a texture patch there are many texels of luminance v occurring in many different local contexts. Due to their different contexts, these different texels of

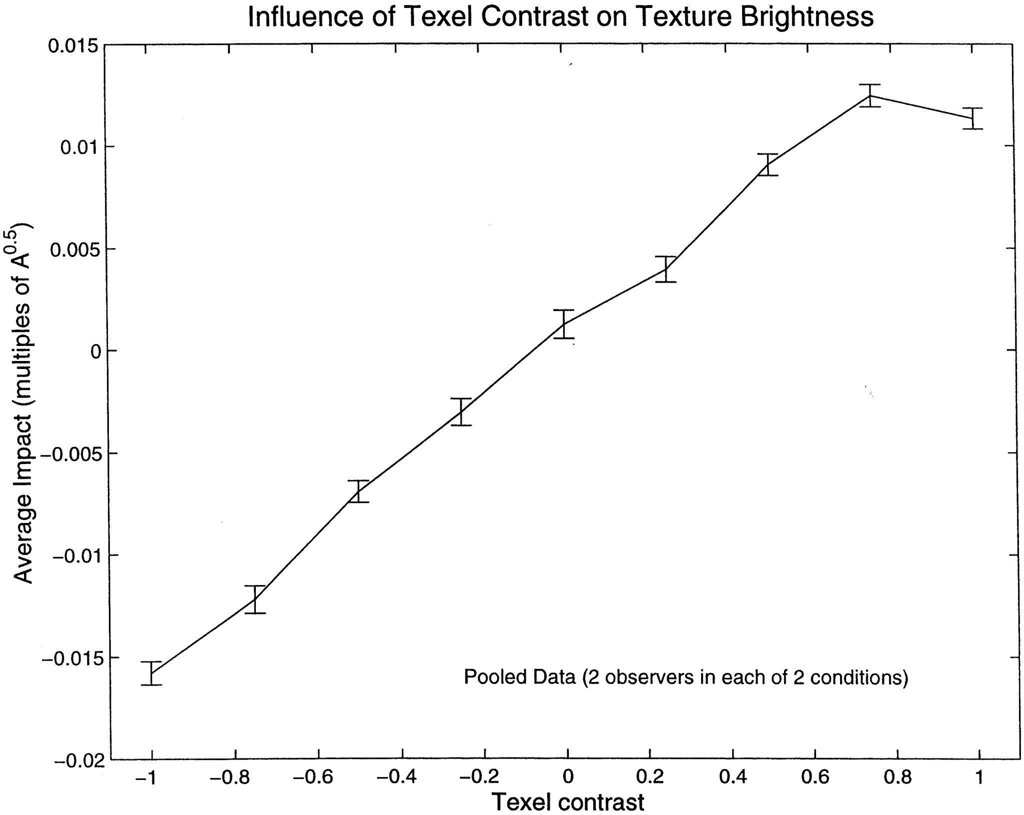

Fig. 3. Average texture brightness impact function. A maximum likelihood estimate of the texture brightness impact function derived from the data pooled across all observers and conditions. Error bars give bootstrapped 95% confidence intervals for the estimated impact function values. Note (i) the linearity of the function across the lower eight texel luminances, and (ii) the slight but significant failure of monotonicity for the highest texel luminance.

luminance v may well have different brightnesses. Suppose it is these different brightnesses that are added together to estimate the luminance of a texture patch. (Note, however, that this need not be the case. Individual texel brightnesses, such as might be gauged by asking the observer to make judgments about single texels, may have nothing to do with assessments of texture brightness.) These different contexts, then, introduce variability in texel brightness, and thus (under our supposition that texture brightness is assessed by adding up texel brightnesses) influence the noise injection function s.

What interpretation should we give to the texture brightness impact function m in light of such context effects? Our answer rests on the following observations. First, because the patches used in our experiments are fairly large, the luminance of a texel is approximately independent of the local context in which that texel is embedded (i.e. one's ability to predict the luminance of a given texel is improved very little by knowing the local context of that texel.) Second, in all experimental conditions we use histograms that differ only slightly from the uniform histogram (only up to the modest levels sufficient to support less than perfect performance at judging patch luminances). Thus, the distribution of available contexts differs little across different textures actually used; in each case the distribution of local contexts is approximately uniform (rendering all local contexts approximately equally probable). Summarizing: in any given experimental condition, the distribution of local contexts is approximately identical for texels of all luminances; moreover, this distribution of contexts is approximately uniform (all local contexts equally probable) across the different conditions used in these experiments. Thus, for any texel luminance v, $m(v)$ approximates the average impact (across all different local contexts) exerted on the brightness of a uniformly distributed texture patch by a texel of luminance v.

4.2. The deviation of the texture brightness impact function from the power law for brightness

As Stevens (1961, 1962, 1967) and Stevens and Stevens (1963) emphasized, direct judgments of the brightness are related, over many log units, to luminance by a power function. The exponent of the power function depends on various aspects of the stimulus such as the size of the target, the state of adaptation of the observer, the context in which the target appears, etc. For a small spot (5 min) viewed by a dark adapted observer, the exponent of the power function is around 0.33 (Stevens, 1967). Adaptation to higher light levels tends to elevate the exponent to higher values less than 1. Of course, Stevens' paradigm (magnitude estimation of the brightness of an isolated patch of light) is fundamentally different from the methods used here. Most importantly, Stevens' observers were basing their judgments on a single target (rather than a texture) and assigning a number (without feedback) to the percept elicited by that target (rather than making a binary forced choice, with feedback). Nonetheless, it is instructive to see how well power functions fare in fitting the current data.

Note that the obtained texture brightness impact functions are approximately linear (yielding veridical intensity representation) over our middle seven texel luminances, and (over these luminances) the curves agree well across conditions and observers. However, for luminance 0, there is substantial variability between observers in the value of $m(0)$ obtained. Most notably, for JN the estimate of $m(0)$ obtained in the low luminance condition is much lower than the estimates obtained for the other observers in other conditions. Moreover, for both observers in the low (high) luminance condition, texels of the highest luminance exerted an impact on texture brightness that was slightly but significantly lower than (no greater than) the impact exerted by texels of the next lower luminance.

Thus, it is really only over the middle seven texel luminances that luminance is veridically coded for purposes of making texture brightness judgments. However, this range (from 2.5 to 17.5 cd/m^2 in our low luminance condition) spans only 0.85 log units. Plausibly, a power function might reasonably fit the data over such a restricted region. To test this possibility, power functions were individually fit (using a least-squares method) to the seven middle data points of each observer. The results are shown in Fig. 4.

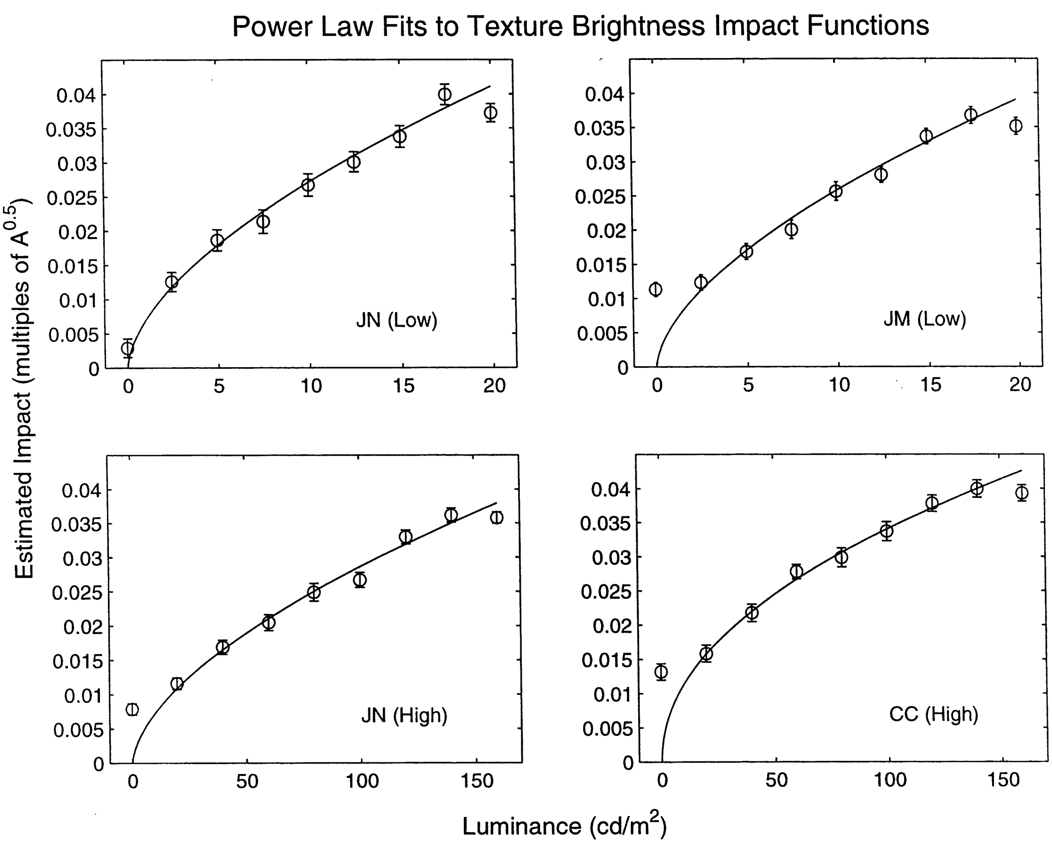

Fig. 4. Power function estimates of texture brightness impact functions. Least-squares estimates of power functions fit to the middle seven data points of the estimated texture brightness impact functions. Impact functions have been shifted upward to optimize fits to these middle seven values. Power function exponents are 0.60 for JM (low luminance condition) and JN (in both high and low luminance conditions), and 0.48 for CC (high luminance condition). Note (i) power functions provide a reasonable fit to the obtained impact functions over the middle seven texel luminances, but (ii) tend to undershoot impact function values for texel luminance 0, and (iii) overshoot impact function values for the maximum texel luminance.

Recall that the texture brightness impact function is only defined up to an arbitrary additive constant and positive scale factor. Thus, when we attempt to approximate m by a power function (of the form $av^b + c$), the parameters a and c are unconstrained by the data; only the exponent b is significant. For each of JN and JM in the low luminance condition, as well as JN in the high luminance condition, the least-squares estimate of the power function exponent was 0.60. For CC (high luminance condition), the exponent was 0.48.

In each case, the power function provides a reasonable fit to the middle seven data points. Indeed, for JN in the low luminance condition, the power function does a moderately good job at approximating the impact function for all except the highest texel luminance. For all other observers, the power function both overshoots the impact function for the highest luminance and undershoots it for texel luminance 0.

We conclude that (i) the apparent linearity of the texture brightness impact function over the middle seven texel luminances is also consistent with a power function coding scheme. However, (ii) the behavior of the texture brightness impact function at the extreme values tested deviates significantly from a power function.

4.3. *Implications to be drawn from the form of the texture brightness impact function*

What can we infer about subcortical intensity coding from the current results? Perhaps not very much. Because all of our texel luminances (except possibly the top two for some observers) exert different impacts on texture luminance judgments, it follows that they must all receive distinct subcortical codes; but this much was obvious from the outset because all our texel luminances are easily discriminable. It is difficult, however, to draw stronger conclusions with much confidence. After all, as long as distinct texel luminances receive distinct subcortical codes, it is logically possible for the cortical system responsible for judging texture luminance to map those distinct subcortical codes onto the distinct impacts that they exert on texture luminance judgments. Only by making assumptions about the computation used by the cortex to estimate texture luminance can we draw any stronger inferences about subcortical intensity coding. For example, only if we are willing to assume that the cortical system used to make texture luminance judgments maps subcortical codes linearly onto the impacts they produce, can we infer that the texture brightness impact function m exactly reflects the form of the function f of Eq. (2). Granted, it would be strange if the mapping from subcortical intensity codes to impact exerted on texture brightness were highly nonlinear.

However, the form of m itself suggests that this mapping is not perfectly linear. We assume with high confidence that our texel luminances produce strictly increasing subcortical intensity codes. If the mapping from subcortical intensity codes to texture brightness impacts were linear, m should be strictly increasing. On the contrary, though, for all observers in all experimental conditions, texels of maximal luminance exerted either equal or lower impact on texture brightness than did texels of the next lower luminance. This forces us to conclude that the cortical module used to make texture luminance judgments must be nonlinear in its mapping of subcortical intensity codes onto texture brightness impacts.

A possible explanation of the nonmonotonicity of the texture brightness impact function runs as follows. Suppose that texture luminance judgments are made by summing the responses of neurons in some array. In addition, imagine that the responses of retinal receptors can (1) excite some neurons in this array, and (2) inhibit others. (We leave open the mechanisms by which these two modes of influence are exerted.) Under one scenario, a direct chain of excitatory connections transmits the receptor signal to an excitatory connection with one or more neurons in the hypothesized array, thus enabling excitatory influence. Inhibitory influence, however, might be achieved indirectly via lateral inhibitory connections between neurons within the array. Under this scheme a bright texel could increase the firing rates of those neurons it directly stimulates (via an excitatory chain beginning with the retinal receptors); however, it could also decrease the firing rates of other neurons via lateral inhibitory connections from neurons it directly stimulates.

Let $m_{\text{excitatory}}$ ($m_{\text{inhibitory}}$) be the function that relates texel luminance to the average, combined output of those neurons receiving excitation (inhibition) from that texel. We assume that $m_{\text{excitatory}}$ ($m_{\text{inhibitory}}$) is monotonically increasing (decreasing). Let v_{max} and $v_{\text{max but one}}$ be the highest and second highest texel intensities. We can explain the nonmonotonicity of the texture brightness impact function by assuming that $m_{\text{excitatory}}$ saturates more rapidly than $m_{\text{inhibitory}}$. In this case, we can imagine a scenario in which the net decrease of $m_{\text{inhibitory}}$ from $v_{\text{max but one}}$ to v_{max} is greater than the net increase of $m_{\text{excitatory}}$ from $v_{\text{max but one}}$ to v_{max}. Hence,

$$\begin{aligned} m(v_{\text{max}}) &= m_{\text{excitatory}}(v_{\text{max}}) + m_{\text{inhibitory}}(v_{\text{max}}) \\ &< m_{\text{excitatory}}(v_{\text{max but one}}) + m_{\text{inhibitory}}(v_{\text{max but one}}) \\ &= m(v_{\text{max but one}}). \end{aligned} \tag{16}$$

It is via the hypothesized functions $m_{\text{excitatory}}$ and $m_{\text{inhibitory}}$ that our conjecture introduces nonlinearity into the mapping from subcortical intensity codes to texture brightness impacts. One naturally assumes that each of these functions is sigmoidal. Moreover, it seems plausible to think that $m_{\text{excitatory}}$ might be steeper than $m_{\text{inhibitory}}$, in which case $m_{\text{excitatory}}$ might well saturate earlier than $m_{\text{inhibitory}}$.

However, having concluded that texture brightness judgments are nonlinearly related to subcortical intensity codes for our two maximal texel luminances, we must acknowledge that the same may be true for other of our texel luminances. Although our middle seven (lower eight in some cases) luminances seem to be veridically represented for purposes of assessing texture luminance, this does not imply that the subcortical codes for these lower eight luminances must, themeselves, be linear. All that is logically entailed by our results is that the subcortical codes for these eight luminances must be distinct.

On the other hand, it should also be noted that if subcortical intensity codes embody a significant compressive nonlinearity, this compression induces little distortion in texture luminance judgments. We conclude that, in making judgments of texture luminance, human vision effectively compensates for any such nonlinearities in precortical intensity coding. This suggests more generally that precortical nonlinearities in the coding of intensity may have little effect on cortical processes.

Acknowledgements

Much of the research reported here was supported by National Science Foundation Human Cognition and Perception Program grant DBS-92032 91. Additional support was provided by AFOSR Life Sciences Visual Information Processing Program grant F49620-94-1-0345. Thanks to Adam Reeves and Walt Makous for helpful comments, and to Geoff Iverson for crucial observations about an earlier version of this paper.

Appendix

Here we derive Eq. (12). It will be convenient to write ξ_v for a random variable with cdf F_v, as defined by Eq. (5).

The theorem of total probabilities implies that for any texel τ of patch P_q the distribution function F_q characterizing the random variable $X_q(\tau)$ is defined as follows: for any $x \in \mathbb{R}$,

$$\begin{aligned}F_q(x) &= \mathbb{P}[X_q(\tau) \le x] \\ &= \sum_{v \in G} \mathbb{P}[X_q(\tau) \le x \mid P_q(\tau) = v]\, \mathbb{P}[P_q(\tau) = v] \\ &= \sum_{v \in G} F_v(x) q(v). \end{aligned} \tag{17}$$

from which it follows that for any function $f: \mathbb{R} \to \mathbb{R}$,

$$\begin{aligned}E[f(X_q(\tau))] &= \int_{-\infty}^{\infty} f(x) dF_q(x) \\ &= \int_{-\infty}^{\infty} f(x) \sum_{v \in G} q(v) dF_v(x) \\ &= \sum_{v \in G} q(v) E[f(\xi_v)]. \end{aligned} \tag{18}$$

Thus, in particular, we have

$$E[X_q(\tau)] = q \cdot m. \tag{19}$$

In addition,

$$\begin{aligned}var[X_q(\tau)] &= E[X_q^2(\tau)] - E^2[X_q(\tau)] \\ &= \sum_{v \in G} q(v)\, E[\xi_v^2] - (q \cdot m)^2 \\ &= \sum_{v \in G} q(v)\, (s^2(v) + m^2(v)) - (q \cdot m)^2 \\ &= q \cdot (s^2 + m^2) - (q \cdot m)^2. \end{aligned} \tag{20}$$

In the following lines,

$$\sum_{\tau} \quad \text{indicates} \quad \sum_{\text{all texels } \tau \text{ in dom}(P_q)}. \tag{21}$$

From Eqs. (7) and (19) it follows that

$$E[L(P_q)] = \sum_{\tau} E[X_q(\tau)] W(\tau) = (m \cdot q) \sum_{\tau} W(\tau) = K_1 m \cdot q \tag{22}$$

for K_1 defined by Eq. (8). Also (because the random variables $X_q(\tau)$ are independent, for all texels τ in dom(P_q))

$$\begin{aligned}var[L(P_q)] &= \sum_{\tau} W^2(\tau) var[X_q(\tau)] \\ &= K_2(q \cdot (s^2 + m^2) - (q \cdot m)^2) \end{aligned} \tag{23}$$

for K_2 defined by Eq. (9).

Under the proposed model, the observer judges P_q more luminant than P_r iff

$$\Delta(q, r) = L(P_q) - L(P_r) + Y > 0. \tag{4}$$

The central limit theorem implies that $\Delta(q, r)$ will be approximately normally distributed, provided (a) the additive pooling performed by each of $L(P_q)$ and $L(P_r)$ is over a moderately large number of texels, and (b) the variance of $\Delta(q, r)$ is not dominated by the contributions of only a few texels.

In addition, from Eqs. (4) and (22) it follows that

$$K_1 m \cdot (q - r) \tag{24}$$

On the other hand, under the natural assumption that $L(P_q)$ and $L(P_r)$ are independent, Eqs. (4) and (23) imply that

$$var[\Delta(q,r)] = K_2((q+r)\cdot(s^2+m^2) - (q\cdot m)^2 - (r\cdot m)^2) + \sigma^2 \quad (25)$$

Under the assumption that $\Delta(q,r)$ is normally distributed, we thus observe that

$$\mathbb{P}[\text{Observer judges } P_q \text{ more luminant than } P_r] = \mathbb{P}[\Delta(q,r) > 0] = \Phi\left(\frac{K_1 m\cdot(q-r)}{\{K_2((q+r)\cdot(s^2+m^2)-(q\cdot m)^2-(r\cdot m)^2)+\sigma^2\}^{1/2}}\right) \quad (26)$$

where Φ denotes the standard normal cdf.

To follow the next strand of argument, one must observe several simple facts concerning histogram modulators. Note first that for any texel distribution h, any h-modulator θ must sum to 0 because

$$\sum_{v\in G}\theta(v) = \sum_{v\in G}(h+\theta)(v) - \sum_{v\in G}h(v) = 0. \quad (27)$$

The uniform distribution U is an example of a constant function of G. More generally, any $k\in\mathbb{R}$ can be construed as the function mapping each $v\in G$ onto k. For any such constant function of G,

$$\theta\cdot k = \sum_{v\in G}\theta(v)k = k\sum_{v\in G}\theta(v) = 0 \quad (28)$$

for any U-modulator θ.

Suppose the texel distribution q in Eq. (26) is set to $U+\theta$, and r is set to $U-\theta$ for some reversible U-modulator θ. Then Eq. (26) implies that the observer judges $P_{U+\theta}$ more luminant than $P_{U-\theta}$ with probability $p(\theta)$ given by

$$p(\theta) = \mathbb{P}[\Delta(U+\theta, U-\theta) > 0] = \Phi\left(\frac{2K_1 m\cdot\theta}{\{2K_2(s^2\cdot U+m^2\cdot U-(m\cdot U)^2-(m\cdot\theta)^2)+\alpha^2\}^{1/2}}\right). \quad (29)$$

Writing $\bar{m}$ for $m\cdot U$ (the mean value of m), note that

$$m^2\cdot U-(m\cdot U)^2 = m^2\cdot U-\bar{m}^2 = (m^2-2\bar{m}m+\bar{m}^2)\cdot U = (m-\bar{m})^2\cdot U \quad (30)$$

enabling us to rewrite Eq. (29) as

$$p(\theta) = \Phi\left(\frac{2K_1 m\cdot\theta}{\{2K_2(s^2\cdot U+(m-\bar{m})^2\cdot U-(m\cdot\theta)^2)+\sigma^2\}^{1/2}}\right). \quad (31)$$

Note also (with reference to Eq. (28)) that

$$m\cdot\theta = m\cdot\theta-\bar{m}\cdot\theta = (m-\bar{m})\cdot\theta \quad (32)$$

Thus the value of Eq. (31) remains unchanged if we substitute $(m-\bar{m})\cdot\theta$ for each occurrence of $m\cdot\theta$ to derive

$$p(\theta) = \Phi\left(\frac{2K_1(m-\bar{m})\cdot\theta}{\{2K_2(s^2\cdot U+(m-\bar{m})^2\cdot U-((m-\bar{m})\cdot\theta)^2)+\sigma^2\}^{1/2}}\right). \quad (33)$$

Inspection of Eq. (33) shows that only the deviation of m from its mean value (and not $\bar{m}$ itself) influences performance; accordingly, we assume with no loss of generality that $\bar{m}=0$, yielding the desired result:

$$p(\theta) = \Phi\left(\frac{m\cdot\theta}{\{A+B(m^2\cdot U-(m\cdot\theta)^2)\}^{1/2}}\right) \quad (12)$$

for

$$B = \frac{K_2}{2K_1^2} \quad \text{and} \quad A = B\left(s^2\cdot U+\frac{\sigma^2}{2K_2}\right) \quad |||||| \quad (34)$$

References

Barlow, H. B. (1965). Optic nerve impulses and Weber's law. *Cold Spring Harbor Symposium on Quantum Biology*, *30*, 539–546.

Barlow, H. B., & Levick, W. R. (1969). Three factors limiting the reliable detection of light by retinal ganglion cells of the cat. *Journal of Physiology*, *200*, 1–24.

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

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

Falmagne, J-C. (1971). The generalized Fechner problem and discrimination. *Journal of Mathematical Psychology*, *8*, 22–43.

Fechner, G.T. (1860). *Elements of Psychophysics*, In D.H. Howes and E.C. Boring (eds.), H.E. Adler, trans., Holt, Reinhardt & Winston, New York, 1966 (originally published, 1860).

Grossberg, S., & Todorovic, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: a unified model of classical and recent phenomena. *Perception & Psychophysics*, *43*, 241–277.

He, S., & MacLeod, D. I. A. (1996). Local luminance nonlinearity and receptor aliasing in the detection of high frequency gratings. *Journal of the Optical Society of America A*, *13*, 1139–1151.

He, S., & MacLeod, D. I. A. (1998). Contrast modulation flicker: dynamics and spatial resolution of the light adaptation process. *Vision Res.*, *38*(7), 985–1000.

Lennie, P. (1993). *Roles of M and P pathways in contrast sensitivity*. In R. Shapley and D. M.-K. Lam (Eds.), MIT Press, Cambridge, MA, pp. 201–213.

Luce, R. D., & Edwards, W. (1958). The derivation of subjective scales from just noticeable differences. *Psychological Review*, *65*, 222–237.

MacLeod, D. I. A., & He, S. (1993). Visible flicker from invisible patterns. *Nature*, *361*, 256–258.

Makous, W. L. (1997). Fourier models and the loci of adaptation. *Journal of the Optical Society of America A*, *14*, 2323–2345.

Maxwell, J. C. (1860). On the theory of compound colors and the relation of the colors of the spectrum. *Philosophical Transactions of the Royal Society*, *150*, 57–84.

Shapley, R., & Enroth-Cugell, C. (1984). Visual adaptation and retinal gain controls. *Prog. Retinal Res.*, *3*, 263–346.

Stevens, J. C., & Stevens, S. S. (1963). Brightness function: effects of adaptation. *Journal of the Optical Society of America A*, *53*, 375–385.

Stevens, S. S. (1961). To honor Fechner and repeal his law. *Science*, *133*, 80–86.

Stevens, S. S. (1962). The surprising simplicity of sensory metrics. *American Psychologist*, *17*, 29–39.

Stevens, S. S. (1967). Intensity functions in sensory systems. *International Journal of Neurology*, *6*, 202–209.

Yang, J., & Makous, W. L. (1995). Modeling pedestal experiments with amplitude instead of contrast. *Vision Research*, *35*, 1979–1989.