Perception, 2000, volume 29, pages 1101 – 1116

DOI:10.1068/p2904

The effect of contrast randomisation on the discrimination of changes in the slopes of the amplitude spectra of natural scenes

C Alejandro Párraga¶, David J Tolhurst
Department of Physiology, University of Cambridge, Downing Street, Cambridge CB2 3EG, UK
Received 26 February 1999, in revised form 25 May 2000

Abstract. It has been suggested (Tadmor and Tolhurst, 1994 *Vision Research* **34** 541 – 554) that the psychophysical task of discriminating changes in the slope of the amplitude spectrum of a complex image may be similar to detecting differences in the degree of blur. It has also been suggested that human observers may perform this discrimination by detecting changes in the effective contrast within single narrow spatial-frequency bands, rather than by detecting changes in the slope per se which would involve the use of contrast information across many different frequency bands. To distinguish between these two possibilities, we have developed an experiment where observers were asked to discriminate changes in the spectral slope while different amounts of random contrast variation were introduced, with the purpose of disrupting their performance. This disruptive effect was designed to be particularly manifest if the observer really was performing a single-frequency-band contrast discrimination but to be unnoticeable if the observer was discriminating the change of slope per se. Our results imply that the observers do not usually detect changes in contrast in just one narrow spatial-frequency band when they discriminate changes in the slope of the amplitude spectrum. Rather, they must compare contrast between bands or, at least, they use contrast information from more than one band. However, for edge-enhanced (whitened) pictures, there is some evidence to suggest that observers rely on contrast changes in only a limited low-spatial-frequency band.

1 Introduction

There has been a resurgence of interest in Kretzmer's (1952) discovery that much of the variety of images which the visual system may naturally encounter are actually described by a simple statistical relationship (Carlson 1978; Burton and Moorhead 1987; Field 1987). The statistical regularity of natural images is shown by a simple relationship describing the amplitude of the frequency coefficients in their Fourier spectra:

$$\text{amplitude}(f) \propto f^{-\alpha}, \tag{1}$$

where amplitude is averaged across all orientations in the spectrum, f is spatial frequency, and α (also called the slope parameter since the relationship is a straight line on log – log coordinates) is greater than zero. The value of α, although it varies from image to image, has been found to be within a fairly narrow range (0.7 – 1.5) for achromatic images (Burton and Moorhead 1987; Tolhurst et al 1992) and across the visible spectrum (Párraga et al 1998). This means that the degree of correlation in luminance between nearby points as a function of distance between the points is remarkably similar from one natural scene to the next. It has been suggested that the overall organisation of the visual system, including the response properties of individual neurons, might exploit any redundancies or correlations in natural scenes (Barlow 1961; Laughlin 1981; Srinivasan et al 1982; Field 1987, 1989; Attick 1992; Attick and Redlich 1992; Brelstaff and Trościanko 1992; Olshausen and Field 1997; van Hateren and van der Schaaf 1998; Tadmor and Tolhurst 2000).

Knill et al (1990) hypothesised that the human visual system should be tuned to best discriminate those scenes with the same α-value as the real world and, therefore,

---

¶ Present address and address for correspondence: Department of Experimental Psychology, University of Bristol, 8 Woodland Road, Bristol BS8 1TN, UK; e-mail: alej.parraga@bris.ac.uk

that visual discriminations should be favoured when images have natural statistics (cf Párraga et al 2000). Knill et al (1990) and Tadmor and Tolhurst (1994) measured the threshold of human observers for discriminating changes in the slope, α, of stimuli derived either from patterns of dots of random luminance or from digitised photographs of natural scenes. Tadmor and Tolhurst argued that discriminating changes in the slopes of amplitude spectra is analogous to discriminating changes in the degree of blur (Hamerly and Dvorjak 1981; Watt and Morgan 1983; Walsh and Charman 1988; Hess et al 1989). A greater value of α is roughly comparable to a greater degree of blur (Goodman 1968). There had been speculation as to whether blur discrimination relied upon detecting changes in the contrast in some limited part of the frequency spectrum of the image (eg Hess et al 1989), and Tadmor and Tolhurst (1994) tried to model the α discrimination task in terms of contrast discrimination too.

The contrast of a complex image such as a natural scene is not simple to evaluate, but must be calculated in several band-limited spatial scales (Peli 1990). Contrast is evaluated in several spatial-frequency bands of about 1 octave bandwidth, and so local band-limited contrast, evaluated at a particular spatial scale, may be considered as representative of the activity of one spatial-frequency channel (Campbell and Robson 1968).

Tadmor and Tolhurst proposed that the discrimination of changes in α could be modelled as a discrimination of changes in local contrast within one specific frequency band. They argued that changing the slope α of the amplitude spectrum also causes this local contrast of the image to change. The magnitude of this change (at one or more spatial scales) agrees reasonably well with that required to discriminate changes in Michelson contrast of a simple sinusoidal grating of the appropriate frequency. The apparent success of this local band-limited contrast model was further shown by Tolhurst et al (1996) and Tolhurst and Tadmor (1997a, 1997b). They demonstrated that the task of discriminating $\Delta\alpha$ did seem to rely upon contrast changes within a limited spatial-frequency band by using stimuli in which the value of α had been changed only within restricted spatial-frequency bands. However, it should be noted that these restricted bands were actually 2 octaves wide.

1.1 *Alternative explanations to the local band-limited contrast model*

Despite the suggestive evidence, the local band-limited contrast model might not be the only possible explanation for discriminating changes in the slope of the amplitude spectra. Indeed, Rohaly et al (1997) have shown that many different models or variant models of discrimination sometimes produce very similar predictions. It is still arguable that the observer may somehow estimate the spectral slope for both the reference and the test stimuli, perhaps by comparing the energy or contrast at two or more different spatial-frequency bands within each image. Figures 1a and 1b illustrate the two alternative explanations. In figure 1a, it is presumed that the observer makes use of cues only in a narrow frequency band (as proposed by Tadmor and Tolhurst), ignoring the rest of the spectrum. In figure 1b, it is presumed that observer makes use of the whole spectral range to estimate slope per se. A slope comparison might, for instance, result if the observer did concentrate on the contrast in one band, but after contrast had been normalised by division by the contrast in some other frequency bands (Heeger 1992; Foley 1994; Rohaly et al 1997).

The purpose of this paper is to investigate the plausibility of the alternative model where observers might be estimating the changes in the amplitude slope by utilising Fourier energy at two or more spatial-frequency bands. To do so, we have designed an experiment where the observer's performance for a simple single-frequency-band contrast discrimination task should be severely disrupted. This is done by introducing different amounts of random variation into the overall contrasts of the reference and test images. Figures 1c and 1d show the effect of this disruption on both models.

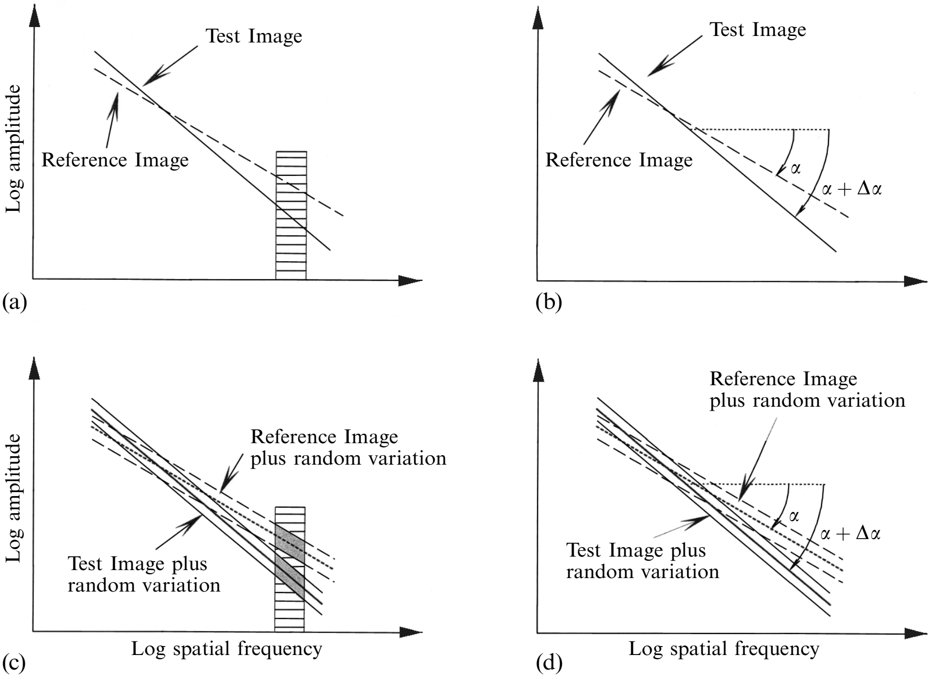

Figure 1. The local band-limited contrast model and its alternative explanation. (a) and (b) The amplitude spectrum of a reference picture and a test picture are shown schematically on double logarithmic axes. Note that the two spectral lines cross at an intermediate frequency; power normalisation was carried out to ensure that the observer could not analyse changes in power instead of changes in amplitude slope, α. In the model of Tadmor and Tolhurst, (a), the observer is presumed to evaluate the Fourier energy or contrast within a single spatial-frequency band (hatched area). The alternative model, (b), consists in estimating the slope (by comparing the energy at two or more frequency bands). (c) and (d) Effect of random perturbations in the overall contrast in both models. The effect of this variation should be greater on the observer's performance if he/she is evaluating changes between the stimuli by discriminating changes in only one frequency band (c). If the observer really does detect slope, then overall contrast perturbations will have no effect on performance, since the perturbation does not change the spectral slopes, only their positions on the y-axis (d).

As seen in the figures, the random shifting (upwards or downwards) of the amplitude spectra ought to vary the amounts of Fourier energy in a given spatial-frequency band while leaving the amplitude slope the same. The disruptive effect of this contrast variation should be much greater in the case where the observer is estimating the spectral slope by discriminating changes in only one spatial-frequency band.

2 Methods

2.1 *Apparatus*

The stimuli were displayed on a custom-made Joyce Electronics raster display (P4 phosphor) at a 100 Hz frame rate with a mean luminance of 45 cd m^{-2}. The screen measured 30.3 cm × 22.9 cm, which was equivalent to 7.58 deg × 5.73 deg at the eye from the observer's viewing distance of 2.30 m from the display. The overall controlling system was an IBM compatible PC housing a VSG2/3 (Cambridge Research Systems) graphics card. This graphics card has a framestore memory that is coupled to two colour palettes, each with 8-bit digital-to-analogue converters (DACs), and 8-bit look-up tables (LUTs). A simple 8-bit display with only 256 grey levels is not adequate for experiments in which it is necessary to change contrast in small steps while still retaining the integrity of the picture or grating. For most of our experiments we needed subtle control of contrast.

For these we used the Pelli and Zhang (1991) modification which allowed us to add together the two palette outputs to form a higher-resolution output. This mode allows the choice of 256 grey levels on any one frame out of a total choice of 4096 grey levels ('pseudo 12-bit resolution').

All images were viewed binocularly, in a darkened room with constant dim artificial lighting so that the display screen was viewed against a background with luminance of about 1 cd m^{-2}. The observer was seated in a comfortable chair with a high neck support that ensured a fixed viewing distance of 2.3 m. Two image sizes were used, size 1 (4 cm × 4 cm, or 1 deg × 1 deg) and size 2 (8 cm × 8 cm, or 2 deg × 2 deg). Each stimulus was composed of 128×128 pixels whose size was about 0.5 and 1 min of arc on the retina for sizes 1 and 2, respectively. Two observers participated; one was emmetropic, the other was a fully corrected myope.

2.2 *Stimuli*

Figure 2 shows the three photographs of natural scenes that were employed as stimuli, the same as in earlier work (Tolhurst and Tadmor 1997b). A detailed account of how these were photographed and digitised can be found elsewhere (Tolhurst et al 1992; Tadmor and Tolhurst 1994). Once the photographs were in a digital form, nonlinearities of the negative's optical density were corrected. Following this, the images were represented at 128×128 pixels and with 256 grey levels, and were Fourier transformed. The resulting amplitude spectra were filtered so as to be consistent with equation (1) with α values equal to 0.4 (whitened), 1.0 (near-normal), and 1.4 (pseudo-blurred). From each of these original images, a set of stimuli was constructed. Each set consisted of a *reference* stimulus whose amplitude slope was equal to α and a sequence of 20 *test* stimuli whose amplitude slope was gradually increased ($\alpha + \Delta\alpha$) in steps of 0.02. The final images were obtained by performing an inverse Fourier transform on these altered spectra.


Picture 1 Picture 3 Picture 10

Figure 2. The original digitised photographs which served as the basis of the stimuli used in this work: picture 1 'tree', picture 3 'portrait', and picture 10 'street'.

Modifying the slope of the amplitude spectrum in this way causes the overall power spectrum to change as well (Bracewell 1986). To avoid this spurious cue for discriminating between a reference and a test stimulus, all images within a set had their powers normalised to the same value before being inverse-Fourier transformed. This effectively means that, when the negative slope, α, is increased, the line is simultaneously shifted upwards, so that the reduction in power at high spatial frequencies is compensated for by an increase in power at low spatial frequencies. This can be seen schematically in figure 1a. Thus, this normalisation of the amplitude spectra differs from a true blurring operation, where the amplitude is decreased at all frequencies. The normalisation was carried out only within a given set of pictures, allowing different sets to have different powers. Figure 3 illustrates the differences between three different sets derived from the same photograph. Although some of the image quality is lost when printed on paper, one still can see the tangible differences produced by altering the value of α. At $\alpha = 0.4$, high spatial frequencies are emphasised although the image appears rather shadowy and of low contrast. When $\alpha = 1.0$, the image appears closer to our everyday perception and is reasonably in focus. At $\alpha = 1.4$, the image looks softer and blurred, consistent with the lack of high frequencies.

2.3 *Measurements of discrimination thresholds*

2.3.1 *Stimulus presentation.* To measure the thresholds for discriminating changes in α within a given set of images, we used a modified two-alternative forced-choice (2AFC) paradigm. A single trial consisted of presenting three images from the set successively for 500 ms, with 200 ms time intervals between them. Two of these three images were the reference image and were identical. The other was one of the test images from the set. The amplitude spectra for the two reference images was f^{α} while the test image had an amplitude spectra of $f^{(\alpha+\Delta\alpha)}$. The order in which the stimuli were presented was decided at random by the computer, except that the middle interval was always the reference. Since the middle image was always the reference, the observer's experimental task was to decide whether the test image (ie the odd one out) was presented in the first or the third interval. The onset of each image presentation was indicated by an auditory signal, and the observer's response was recorded by pressing the appropriate button (left or right) on a remote-control console. Auditory feedback was given whether the observer's choice was correct. The overall power of all three images was always the same and their mean luminance was the 45 cd m^{-2}, the same as the immediately surrounding display. The stimulus was centred on the screen and the observer was allowed to fixate freely. All other parts of the screen (not occupied by the stimulus) had a fixed luminance of 45 cd m^{-2}.

2.3.2 *Random contrast variation.* As described in section 1, the purpose of this work was to determine the plausibility of the local band-limited contrast model by attempting to disrupt the observer's performance for single-frequency-band contrast discrimination. To do this, random contrast variation was imposed on the stimuli during presentation; a different amount of variation was imposed on the three stimulus presentations within a single trial. The amount of random contrast was specified as a percentage of the original contrast of the image (ε%). In practice, this means that the overall power (or rms contrast) of any image was reduced by an equiprobable random value between 0% and ε%. Consequently, all three images (including the two reference ones) presented on a single trial were now likely to differ from each other in their overall power.

2.3.3 *Calculation of the discrimination threshold.* Once the set of images was chosen (eg one of the series in figure 3), two different staircases were initialised. One was initialised with a test stimulus that was well below the discrimination threshold (ie a very low $\Delta\alpha$) and the other with a stimulus well above the discrimination threshold. Both staircases had the same reference stimulus. During a single experimental 'run' each of the test stimuli was presented 5 times. If, for example, 5 out of 5 responses were correct, the value of $\Delta\alpha$ was decreased, making the task harder in the next trial. If the number of correct responses was 3 or fewer, then $\Delta\alpha$ was increased, making the task easier. α was increased or decreased in linear steps. In general, the two staircases converged within 10 – 15 individual runs (of 5 presentations for each staircase), and 20 runs were carried out. The discrimination threshold was estimated by fitting the measured psychometric function with the integral of a normal distribution, which was constrained to fall within 50% correct answers (the guess rate for this 2AFC experiment) and 98% correct answers (allowing 2% of 'finger error'). The discrimination threshold was defined as the value of $\Delta\alpha$ eliciting 74% correct responses and was found by interpolation on the fitted curve. The goodness of the fit was estimated by verifying that χ^2 was lower than the number of

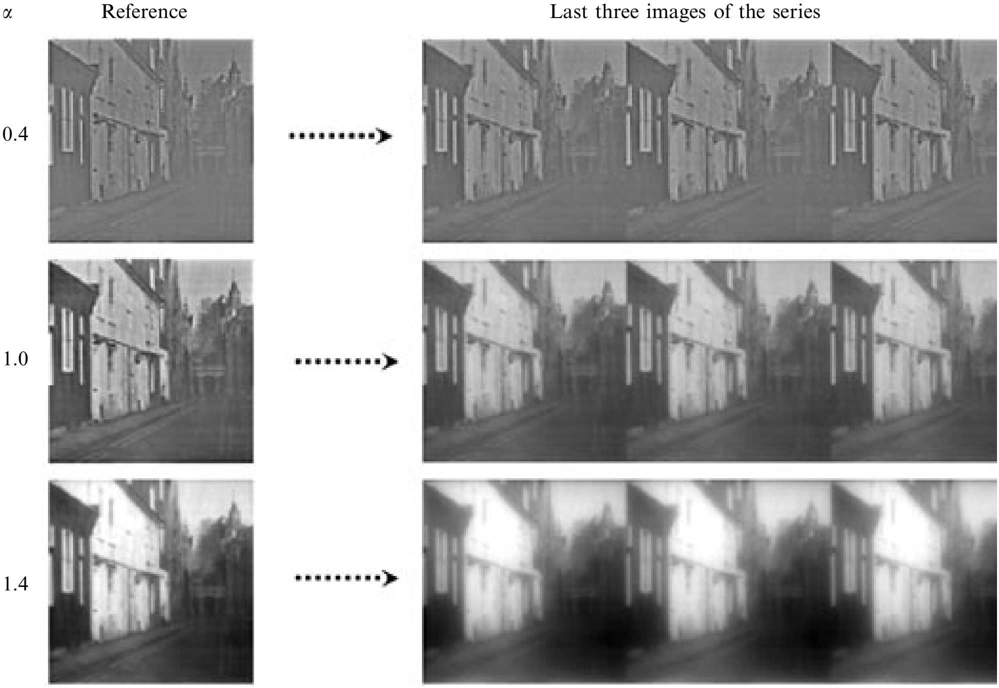

Figure 3. Examples of stimuli made from the three different sets of images based on the same original photograph (picture 10). The value of α for the reference in the three rows is 0.4, 1.0, and 1.4. The figure shows the reference and the last three images of each set.

degrees of freedom in most of the cases. The standard error of the threshold and the value of the slope, β, of the fitted function were also estimated from the inverse of the second-differential of the merit function at the point where the parameters gave the best fit (Edwards 1976). In fact, the thresholds for several experimental conditions were determined concurrently, by randomly interleaving the staircases for the different conditions.

2.4 *Calculation of local contrast*

We model the human observer's ability to discriminate changes in the amplitude spectra of complex scenes in terms of a relatively simple task: the discrimination of changes in the Michelson contrast of simple sinusoidal gratings. To do this, we need to find an appropriate definition of contrast.

A physiologically plausible definition should not only characterise contrast as a measure of local luminance modulation divided by the local mean luminance, but it should take into account how well the neurons or channels (which are optimally responsive to certain spatial frequencies) are activated by the stimulus. Consequently, this definition should be calculated at a series of different spatial scales (or spatial-frequency bands) with operators whose properties broadly resemble those of visual neurons (see Peli 1990; Tolhurst and Tadmor 1997a, 1997b).

We define local contrast at a given position in the picture $[x, y]$ and within a specified spatial-frequency band F as:

$$C_F(x, y) = \frac{a_F(x, y)}{l_{\mathrm{m}}(x, y)}\,, \tag{2}$$

where $a_F(x, y)$ is a band-pass filtered version of the image convolved with a circularly symmetric operator which has the following spatial-frequency characteristic:

$$A_F(f) = \exp\left[-\frac{(f-F)^2}{2\sigma^2}\right]. \tag{3}$$

$l_{\mathrm{m}}(x, y)$ is the result of low-pass filtering the image with a circularly symmetric operator with the following spatial-frequency characteristics:

$$L_{\mathrm{m}}(f) = \exp\left(-\frac{f^2}{2\sigma^2}\right) , \tag{4}$$

f is the spatial frequency, F is the centre frequency of the particular contrast band, and σ is the standard deviation of the Gaussian operator. This parameter is the same in equations (3) and (4), ensuring that both the 'modulating' signal (a_F) and the mean-luminance signal (l_{m}) had the same spatial weighting and extent. This local-band contrast operator roughly represents an 'on' centre ganglion cell. A_F represents the antagonistic centre–surround of a typically band-pass cell. The division by L_{m} represents local light adaptation (Shapley and Enroth Cugell 1984). We set σ to be 0.3 times F, giving the operator a spatial-frequency bandwidth at half-height of about 1 octave, similar to that of human channels (Blakemore and Campbell 1969; Watson and Robson 1981) and neurons in primary visual cortex of a variety of mammalian species (Movshon et al 1978; Tolhurst and Thompson 1981; De Valois et al 1982; Baker et al 1998). The contrast in a frequency band at a single point in a picture is given by equation (2); the absolute value of contrast in the band was then averaged across the picture to give the 'contrast in the band'.

2.5 *Just noticeable differences in real contrast*

2.5.1 *Sine waves.* To evaluate whether the observer's discrimination thresholds for changes in α in complex stimuli can be explained as discrimination of changes in contrast for a given frequency band, F, we must first determine how well the observer is capable of detecting such changes for simple stimuli. This is done by measuring how well the observer can discriminate changes in the 'real' Michelson contrast of simple sinusoidal gratings of frequency F and by evaluating whether the corresponding change in contrast in a similar frequency band (between the test and reference images) is large enough to be detectable.

The observer's discrimination functions for sinusoidal gratings were not measured directly. We used a template corresponding to the familiar 'dipper' function for contrast discrimination (Campbell and Kulikowski 1966; Nachmias and Sansbury 1974; Tolhurst and Barfield 1978; Legge and Foley 1980; Legge 1981; Foley 1994). The position of this template in the plot was determined by measuring the observer's contrast thresholds for detecting sinusoidal gratings of spatial frequency F. To achieve this, we used the same experimental procedure as before, since the detection threshold is equivalent to the discrimination threshold when the reference image has zero contrast.

Figure 4 shows how we obtained the predicted 'dipper' function or contrast discrimination function for a given spatial frequency. Once the observer's contrast thresholds for detecting sinusoidal gratings of frequency F are determined (figure 4a), the predicted dipper function for gratings of the same spatial frequency is aligned. The low-contrast asymptote and the lowest point of the dip are made equal to the detection threshold (figure 4b). All detection thresholds were measured on square patches of sinusoidal grating, 4 cm × 4 cm, presented foveally for 500 ms with a similar 2AFC paradigm to that described above. Michelson contrast of simple sinusoidal gratings was defined as:

$$\text{Michelson contrast} = \frac{L_{\mathrm{max}} - L_{\mathrm{min}}}{2L_{\mathrm{mean}}} , \tag{5}$$

where L represents luminance. Once these measurements have been completed, we are in a position to determine whether the change in equivalent contrast within a certain spatial-frequency band is actually large enough to be detectable by the observer.

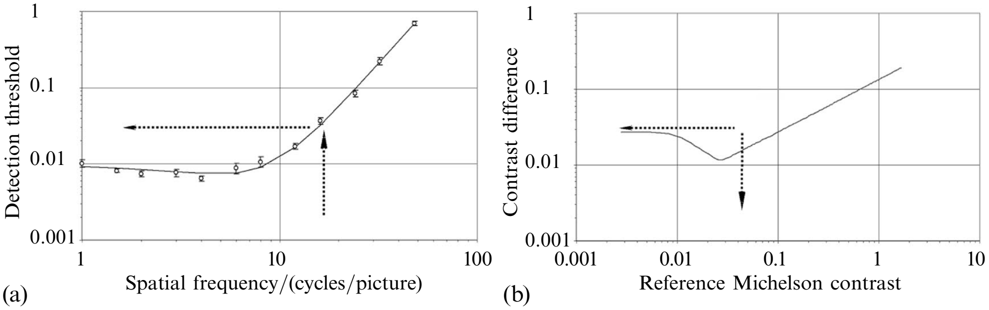

Figure 4. (a) The contrast thresholds for detecting a sinusoidal grating as a function of spatial frequency. The threshold corresponding to 15 cycles/picture has been marked with an arrow. This threshold was used to position the predicted 'dipper' for that spatial frequency on a graph of contrast discrimination threshold plotted against reference contrast (b); the arrows on abscissa and ordinate point at the contrast threshold from (a).

2.5.2 *Band-passed natural scenes.* We tested the validity of this theoretical dipper function directly (see figures 8 and 9). As well as measuring the discrimination of contrast with sinusoidal gratings as the stimuli, we also used band-passed versions of the natural images (figure 2) themselves. The digitised photographs were Fourier-transformed and were filtered with an isotropic filter with a spatial-frequency bandwidth of about 1 octave. The test images were then constructed by inverse Fourier transformation. The contrast of the stimuli was calculated as in section 2.4 as the local contrast in the band centred on the filter used in the band-pass operation. We used the same 2AFC technique to measure the minimum contrast increment needed to detect a change in contrast of the band-passed images at a number of pedestal contrasts.

3 Results

3.1 *The effects of contrast disruption on slope discrimination*

As state above, any change in spectral slope will be accompanied by a change in local contrast within certain spatial-frequency bands. Our experiments were designed to disrupt the performance of the visual system if the observer is detecting contrast changes in only one of these bands, but to interfere very little if the observer is evaluating changes in the amplitude slope per se by comparison between bands.

Figure 5a shows the discrimination thresholds for changes in the amplitude slope, $\Delta\alpha$, when the random contrast disruption is applied. The ordinate represents range of the random contrast disruption (ε%) at each measurement (see section 2). The graphs show the results for two observers. The image numbers correspond to those of the images shown in figure 2. The reference value of the slope α was 1.0 (close to natural statistics) in all cases. Observer JTCL showed no noticeable increment in the discrimination threshold even when a disruption of up to 100% of the contrast was added to the pictures (pictures size 2, left plot). Observer CAP showed a small increment in the discrimination threshold (of about three times the average standard errors) when up to 75% contrast disruption was added (pictures size 1, right plot).

Figure 5b shows similar results for reference α of 1.4. Both observers showed no noticeable (larger than the standard errors) increment in their discrimination thresholds despite massive contrast disruption. Again, observer JTCL (left plot) used pictures size 2 and observer CAP used pictures size 1. Thus, at α values of 10 (near normal images) and 1.4 (seemingly blurred images), the added contrast disruption did not cause a major disruption of the threshold for detecting changes in the amplitude slope, $\Delta\alpha$; this implies that the simple model of Tadmor and Tolhurst (1994) and Tolhurst and Tadmor (1997a, 1997b) is not applicable in these situations.

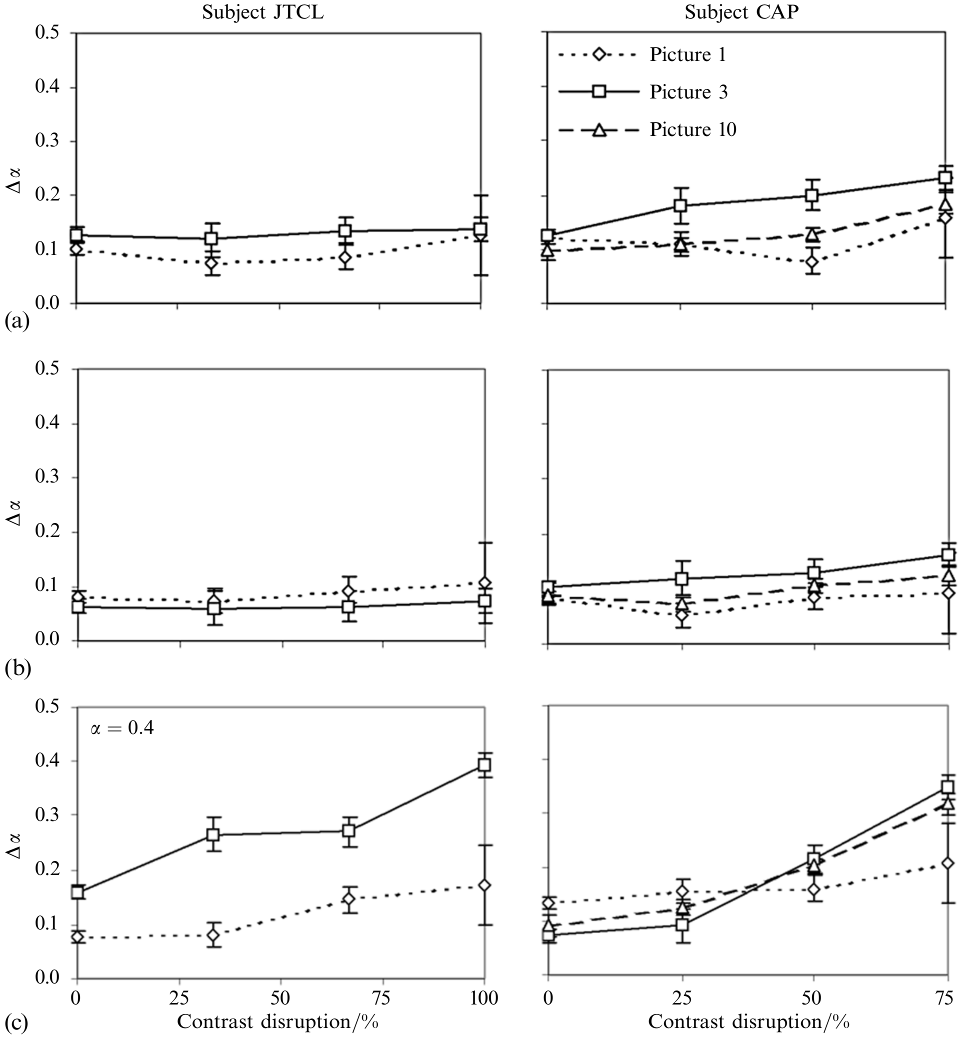

Figure 5. Influence of different degrees of contrast disruption on the discrimination thresholds, $\Delta\alpha$. Two observers participated in the measurements: JTCL and CAP. Image sequences from all three original images were tested on both observers; JTCL was tested with size 2, CAP with size 1. The results here and in the rest of the paper are much the same for the two observers, despite the difference in image size. Reference α values were equal to (a) 1.0, (b) 1.4, and (c) 0.4. Error bars are ±1 SE. Note that the results with whitened images (c) have a different character to those in (a) and (b).

However, figure 5c shows the results for image series where the reference α was equal to 0.4 (edge-enhanced or whitened images). In these series, the contrast disruption produced a remarkable effect in both observers for the two different image sizes. In the most extreme case (subject CAP, image 3, size 1) thresholds were increased by a factor of about 5 (ten times the size of the standard error bar). These results imply that in the case of edge-enhanced images, observers may be detecting changes in contrast in just one narrow spatial-frequency band. It is clear, at least, that a different strategy is employed for whitened images than for the series with steeper spectral slopes.

3.2 *Contrast reduction versus contrast disruption*

These results suggest that the band-limited local contrast model is not applicable in its simplest form and that, instead, some kind of slope analysis is being employed in this discrimination task. However, it is possible that these results are affected by the reduction in average contrast that necessarily accompanies the contrast disruption; the effects of the disruption itself may have been obscured. In order to dismiss this possible artifact, we performed a control experiment that follows the same experimental paradigm. In this case, instead of introducing a random contrast disruption we simply reduced the average contrast of our images by a determined fixed percentage; all three stimuli in a trial would now have the same (but low) contrast. Figure 6 shows the results for this control experiment. Both graphs show little change when contrast was reduced by a fixed amount, confirming that a mere reduction of the overall contrast has no effect on the observer's ability to discriminate changes in α.

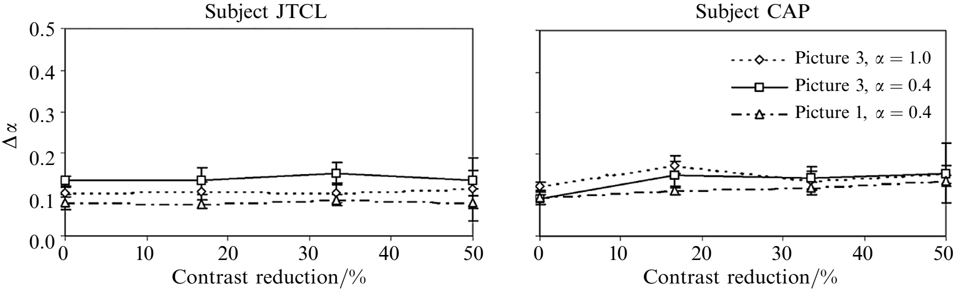

Figure 6. Results for a control experiment showing that reducing the contrast by a fixed amount (but without random disruption) does not produce a significant change in the observer's ability to discriminate variations in α. Subject JTCL used images size 2, and subject CAP used images size 1.

3.3 *Disruption of real contrast discrimination in gratings*

Figure 5 showed that the addition of random contrast disruption disturbed the observer's performance very little at reference α values of 1.0 and 1.4, but was greatly disturbed at an α value of 0.4. However, it is not immediately obvious how much threshold elevation would be expected on the scale of $\Delta\alpha$ in response to a disruption of contrast. For comparison and for 'calibration' of the effect, we examined the effect of the same random kind of contrast disruption on the observer's ability to detect changes in real Michelson contrast of simple sinusoidal gratings. We performed experiments on gratings with spatial frequency and contrast similar to those that we surmised that the observer had been using to make the discriminations in the complex images.

Table 1 lists the calculated contrasts in one limited (1 octave) spatial-frequency band for the reference stimuli used in all of the above experiments. Three image sets (images 1, 3, and 10) and three reference values of α (0.4, 1.0, and 1.4) are considered. The frequency bands chosen as the most likely to have been involved in the discrimination task were 2.5 and 10 cycles/picture for observer CAP at different α values; and 4, 15, and 12.5 cycles/picture for observer JTCL. These spatial-frequency bands were selected because the change in contrast in that band between the reference and threshold test images within a given image set lay closest to the observer's contrast discrimination function (as described in section 2.5).

Figure 7 shows two examples of discrimination experiments in which sinusoidal gratings of the spatial frequencies and pedestals shown in table 1 have been used. The discrimination experiments were carried out with gratings of the same size as the images previously employed (size 2 for observer JTCL and size 1 for observer CAP). Both experiments show a very large increment in the contrast discrimination threshold for sinusoidal gratings when the disruption is applied. This shows the extent of the disruption expected when a visual discrimination is restricted to only one spatial-frequency band.

To be able to compare the results with sinusoidal gratings (figure 7) directly with those of the previous experiments on the discrimination of changes in the amplitude

Table 1. The band-limited contrasts, *C*, of the reference stimuli (three image sets at three reference αs) are shown at the 1 octave spatial-frequency band most likely to have been involved in the discrimination, for each of the observers.

Reference α	Observer CAP			Observer JTCL		
	frequency band cycles/picture	picture	*C*	frequency band cycles/picture	picture	*C*
0.4	2.5	1	0.0214	4	1	0.0233
		3	0.0108		3	0.0143
		10	0.0132			
1.0	10	1	0.1016	15	1	0.1014
		3	0.0564		3	0.0537
		10	0.1242			
1.4	10	1	0.0880	12.5	1	0.0829
		3	0.0849		3	0.0708
		10	0.1199			

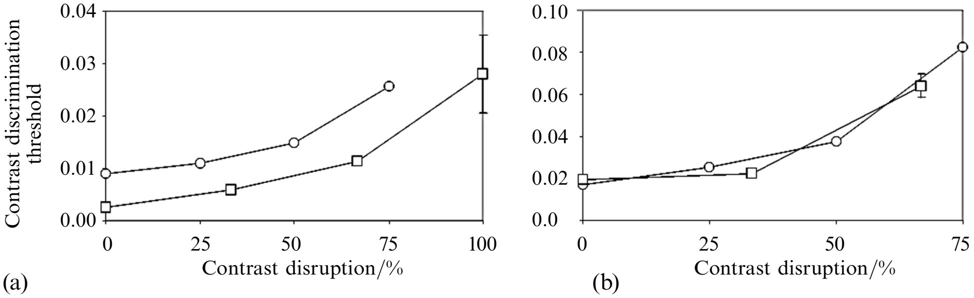

Figure 7. The effect of contrast disruption on contrast discrimination thresholds for sinusoidal gratings. (a) The results for gratings of 2.5 cycles/picture (observer CAP, circles) and a pedestal contrast of 0.032; and 4 cycles/picture (observer JTCL, squares) and a pedestal contrast of 0.018. (b) The results for gratings of 10 cycles/picture (observer CAP, circles) and a pedestal contrast of 0.056; and 15 cycles/picture (observer JTCL, squares) and a pedestal contrast of 0.1.

spectra of natural scenes, we must have a common threshold metric for the two kinds of experiment. It is conceivable that, say, a small change in $\Delta\alpha$ for a reference α of 1.4 corresponds to a large change in contrast. Table 2 shows how the discrimination thresholds, $\Delta\alpha$, for each image at each reference α value were increased when a contrast disruption of 75% was applied. It also shows how the calculated band-limited contrast difference, ΔC, between the (threshold) test and the reference stimulus varies. The spatial frequency chosen for analysis was the same as in table 1. The observer here is CAP.

Table 2 shows that the increments of calculated band-limited contrast, ΔC, are actually bigger than the corresponding changes in $\Delta\alpha$. We can compare these calculated contrast increments with the elevation of the contrast threshold actually produced in the sine-wave grating experiments. Table 3 shows the increments in contrast, ΔC, necessary to discriminate between real reference and test gratings in the presence of up to 75% contrast disruption (eg figure 7). As we can see, the contrast disruption causes a much bigger percentage change in threshold for simple gratings than it does for complex images with α values of 1.4 and 1.0 (table 2). The elevation threshold for detecting changes in spectral slope from α of 0.4 is as great as that found for simple gratings when both thresholds are expressed as change in contrast.

Table 2. The discrimination thresholds, $\Delta\alpha$, for each of three image sets at each of three values of α, with and without an amplitude disruption of 75%. The band-limited contrast differences between test and reference stimulus at threshold, ΔC, are also shown, calculated for the spatial-frequency band stated in column 2. The percentage elevation of threshold, $\mathit{\Delta}$, is expressed in terms both of slope, α, and of band-limited contrast C. Observer CAP.

α	Frequency band	Amplitude disruption	Image 1		Image 3		Image 10	
	cycles/picture		$\Delta\alpha$	ΔC	$\Delta\alpha$	ΔC	$\Delta\alpha$	ΔC
0.4	2.5	0%	0.134	0.0081	0.074	0.0022	0.091	0.0032
		75%	0.206	0.0136	0.348	0.0135	0.317	0.0147
		$\mathit{\Delta}$	53.7%	67.9%	370%	514%	248%	359%
1.0	10	0%	0.121	0.0083	0.123	0.0037	0.101	0.0080
		75%	0.158	0.0111	0.232	0.0092	0.187	0.0178
		$\mathit{\Delta}$	30.6%	33.7%	88.6%	149%	85.1%	123%
1.4	10	0%	0.086	0.0118	0.102	0.0124	0.090	0.0160
		75%	0.093	0.0131	0.160	0.0195	0.125	0.0219
		$\mathit{\Delta}$	8.14%	11.0%	56.9%	57.3%	38.9%	36.9%

Table 3. Elevation of the thresholds, $\mathit{\Delta}$, for discriminating differences in contrast between sinusoidal gratings of the spatial frequency and pedestal contrast shown. Observer CAP.

Contrast	Frequency band	Amplitude disruption	$\Delta\alpha$	ΔC
	cycles/picture			
0.031	2.5	0%	–	0.0089
		75%	–	0.0257
		$\mathit{\Delta}$		189%
0.056	10	0%	–	0.0169
		75%	–	0.0825
		$\mathit{\Delta}$		388%

4 Discussion

Previous studies (Tadmor and Tolhurst 1994; Tolhurst et al 1996; Tolhurst and Tadmor 1997a, 1997b) seemed to provide evidence that an observer's ability to discriminate changes in the amplitude spectra of natural images could be modelled as a simpler task—that of discriminating changes in the Michelson contrast of simple sinusoidal gratings. For each observer and for each reference α [equation (1)], it seemed that there was one spatial-frequency band where the changes in contrast were almost large enough alone to explain the observer's performance. This conclusion was supported by experiments with stimuli in which the spectral slope was changed only within restricted frequency bands: the ability to perform the task was much more reliant on some frequency bands than on others (Tolhurst et al 1996; Tolhurst and Tadmor 1997a, 1997b). However, it was pointed out that, given the spectral complexity of the stimuli, it would be surprising if the observer's discriminations should rely on only a limited portion of the available amplitude spectrum. In fact, the changes in band-limited contrast in complex stimuli were sometimes not large enough on their own to account for the magnitude of the discrimination threshold at a single spatial scale. Thus, there is scope for an alternative hypothesis. There might be some interaction or comparison between two or more spatial-frequency bands.

In this paper, we report an experiment designed to determine whether this alternative view is correct or not. Our results suggest that the observer must somehow be detecting change in the amplitude slope by using information from more than a single spatial-frequency band. We applied random amounts of contrast or power variation to the natural-image stimuli and found that, for reference α values of 1.0 and 1.4, this disruption had little effect on discrimination. This surely means that the observer could not have been relying on contrast information in single frequency bands, since the same contrast disruption did cause the expected large disruption of contrast discriminations with simple sinusoidal gratings.

The logic of the experiment, however, neglects contrast normalisation (Heeger 1992; Foley 1994; Rohaly et al 1997). It is increasingly clear that single frequency channels or frequency bands cannot be considered as independent coding entities, especially when the stimuli contain energy in many different frequency bands (as here). Rohaly et al (1997) considered a number of single-channel and multiple-channel models for predicting the discriminability of objects in natural scenes, and they concluded that there was surprisingly little difference in the predictions of the different models, except that inclusion of contrast normalisation in any model made it much better. If contrast normalisation were absolute, it would completely negate the logic of our experiment, even though we might change the physical contrast of the image, the contrast normalisation would ensure that the 'internal representation' of the stimulus did not actually change. However, the contrast normalisation cannot be so perfect: discrimination of whitened images (α of 0.4) and of grating contrast was disrupted by contrast randomisation. Furthermore, reduction in the physical contrast of all our stimuli was accompanied by a reduction in perceived contrast. It remains possible that contrast normalisation could be a means by which an observer might 'compare' the contrast energy in two bands: normalisation simply means that contrast in one band is represented relative to contrast in more than one band.

The arguments of Tadmor and Tolhurst (1994) and Tolhurst and Tadmor (1997a, 1997b) were based on the finding that the changes in calculated contrast within the test stimuli were similar to those needed to discriminate contrast for real sinusoidal gratings. We should consider whether the magnitudes of the discrimination thresholds in the present experiments are also compatible with those for discriminating changes in the contrast of simple sinusoidal gratings. We follow the procedures of the previous papers by investigating whether the change in equivalent contrast between the reference and test stimuli (calculated within a certain spatial-frequency band) is actually large enough to be detectable by the observer. If this change is not large enough, it is unlikely that a simple contrast discrimination forms the basis of any discrimination task. Contrast discrimination with simple sinusoidal gratings is described by the characteristic 'dipper' template (figure 8); the template is positioned on a graph where the ordinate represents just-detectable contrast difference and the abscissa represents reference or pedestal contrast, with the use of the results of a sinusoidal-grating detection experiment (as described in section 2.5.1). In figure 8, we show the theoretical 'dipper' for discriminating differences in contrast for gratings with a spatial frequency of 10 cycles/picture. The filled symbols show that the template does describe the contrast discrimination thresholds for gratings and also natural images band-pass filtered on a centre frequency of 10 cycles/picture. The open symbols show the changes in calculated contrast in an experiment with the natural image stimuli, ie the effects of contrast reduction on the slope discrimination thresholds with a reference α of 1.0. As the figure shows, the results lie close to the dipper but they do not fit the dipper as closely as might be expected if the observer was performing a simple contrast discrimination and if the observer was relying solely on information within the frequency band that we have illustrated. We have performed this analysis at other frequency bands and the match of the calculated contrast differences to the dipper was generally worse than that illustrated in figure 8. Thus, no single frequency band is adequate alone to account for the observer's ability to detect changes in spectral slope from a reference α of 1.0. The observer must have used information from more than one frequency band, and the nature of the comparison must have made the observer's ability resistant to our imposition of random contrast perturbations.

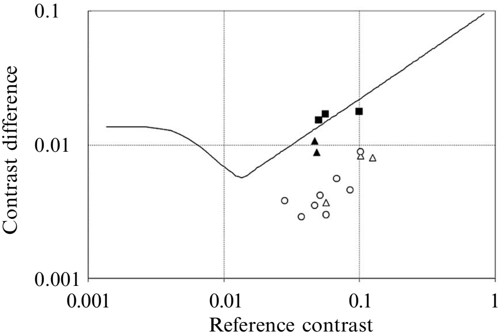

Figure 8. Modelling slope discrimination at reference $\alpha = 1.0$ by contrast discrimination at 10 cycles/picture. The solid line represents the theoretical 'dipper' function for contrast discrimination for simple sinusoidal gratings with spatial frequency of 10 cycles/picture, given the observer's measured contrast detection threshold for a grating of this frequency. The filled squares show actual measurements of the just-noticeable difference (jnd) in contrast for real gratings of 10 cycles/picture. The filled triangles represent jnds for contrast discrimination in band-passed images (pictures 1 and 10); the pass band was about 1 octave wide, centred on 10 cycles/picture. The open circles and open triangles show the calculated contrasts in the reference images and the calculated contrast differences at threshold in the experiment where the observer was obliged to detect differences in spectral amplitude slope from a reference α of 1.0. Contrast is calculated in the band centred on 10 cycles/picture, and the results modelled are from the contrast reduction experiment (pictures 1, 3, and 10 have not been distinguished in the plots to avoid unnecessary complexity). Observer CAP.

However, this is not the case for a reference α of 0.4 (whitened images). Figure 9 shows a similar plot for the contrast reduction experiment when the reference α is 0.4 and the modelled spatial frequency is 2.5 cycles/picture. As we can see from the open symbols in the figure, most of the transformed data for the slope discrimination experiment with natural images do now lie very close to the dipper template. This is just the behaviour expected from an observer who is making the slope discrimination in the complex images as if discriminating the Michelson contrast of a sinusoidal grating of only the given spatial frequency. Thus, the observer might have been using

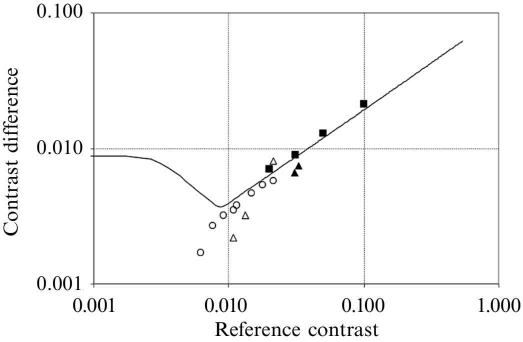

Figure 9. As figure 8, but modelling slope discrimination at reference $\alpha = 0.4$ by contrast discrimination at 2.5 cycles/picture. The solid line represents the theoretical dipper function for contrast discrimination of simple sinusoidal gratings with spatial frequency of 2.5 cycles/picture. The filled squares are the jnd results for real gratings of 2.5 cycles/picture, and the filled triangles represent jnds for pictures 1 and 10 band-passed on a centre frequency of 2.5 cycles/picture. The open circles and open triangles show the calculated contrasts in the reference images and the calculated contrast differences at threshold in the experiment where the observer was obliged to detect difference in spectral amplitude slope from a reference α of 0.4. Contrast is calculated in the band centred on 2.5 cycles/picture, and the results modelled are from the contrast-reduction experiment (pictures 1, 3, and 10 again are not differentiated in the plots). Observer CAP.

just this single spatial-frequency band to perform the task, and this may explain why the contrast disruption described in section 3.1 produced such a large effect in the observer's performance when the reference stimuli had a spectral slope of 0.4.

5 Conclusions

The results shown here challenge the validity of the Tadmor and Tolhurst (1994) band-limited contrast model for discrimination of natural images which differ only in their amplitude spectrum slope α. We found that the observers' thresholds for this task generally remain unaltered in the presence of a random contrast component designed to severely impair an observer's ability to make contrast judgments in just one independent frequency band. Thus we conclude that a band-limited local contrast model can only describe the processes underlying the slope discrimination task if there is some form of slope analysis involved. This might involve a comparison of local contrast between two or more spatial-frequency bands. However, one set of results (whitened images) suggests that the observers may, indeed, be relying on only one spatial-frequency band to perform the slope discrimination task.

Acknowledgements. We are grateful to Fight for Sight, who funded this research; CAP was a Fight for Sight Fellow at Cambridge University. We are grateful to Jason T C Lee for acting as an observer, and to Peter Joyce and Mick Swann for maintaining our equipment.

References

Attick J J, 1992 "Could information theory provide an ecological theory for sensory processing?" *Network* **3** 231 – 251

Attick J J, Redlich A N, 1992 "What does the retina know about natural scenes?" *Neural Computation* **4** 196 – 210

Baker G E, Thompson I D, Krug K, Smyth D, Tolhurst D J, 1998 "Spatial frequency tuning and geniculocortical projections in the visual cortex (areas 17 and 18) of the pigmented ferret" *European Journal of Neuroscience* **10** 2657 – 2668

Barlow H B, 1961 "Possible principles underlying the transformation of sensory messages", in *Sensory Communication* Ed. W A Rosenblith (Cambridge, MA: MIT Press) pp 217 – 274

Blakemore C, Campbell F W, 1969 "On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images" *Journal of Physiology (London)* **203** 237 – 260

Bracewell R N, 1986 *The Fourier Transform and Its Applications* (New York: McGraw-Hill)

Brelstaff G J, Trościanko T, 1992 "Information content of natural scenes: implications for neural coding of color and luminance" *SPIE—Human Vision, Visual Processing and Digital Display* **1666** 302 – 309

Burton G J, Moorhead I R, 1987 "Color and spatial structure in natural scenes" *Applied Optics* **26** 157 – 170

Campbell F W, Kulikowski J J, 1966 "Orientation selectivity of the human visual system" *Journal of Physiology (London)* **187** 437 – 445

Campbell F W, Robson J G, 1968 "Application of Fourier analysis to the visibility of gratings" *Journal of Physiology (London)* **197** 551 – 566

Carlson C R, 1978 "Thresholds for perceived image sharpness" *Photographic Science and Engineering* **22** 69 – 71

De Valois R L, Albrecht D G, Thorell L G, 1982 "Spatial frequency selectivity of cells in macaque visual cortex" *Vision Research* **22** 545 – 559

Edwards A W F, 1976 *Likelihood* (Cambridge, UK: Cambridge University Press)

Field D J, 1987 "Relations between the statistics of natural scenes and the response properties of cortical cells" *Journal of the Optical Society of America A* **4** 2379 – 2394

Field D J, 1989 "What the statistics of natural images tell us about visual coding" *SPIE—Human Vision, Visual Processing and Digital Display* **1077** 269 – 276

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

Goodman J W, 1968 *Introduction to Fourier Optics* (New York: McGraw-Hill)

Hamerly J R, Dvorjak C A, 1981 "Detection and discrimination of blur in edges and lines" *Journal of the Optical Society of America* **71** 448 – 452

Hateren J H van, Schaaf A van der, 1998 "Independent component filters of natural images compared with simple cells in primary visual cortex" *Proceedings of the Royal Society, Series B* **265** 359 – 366

Heeger D J, 1992 "Normalization of cell responses in cat striate cortex" *Visual Neuroscience* **9** 181 – 197

Hess R F, Pointer J S, Watt R J, 1989 "How are spatial filters used in the fovea and parafovea?" *Journal of the Optical Society of America A* **6** 329 – 339

Knill D C, Field D J, Kersten D, 1990 "Human discrimination of fractal images" *Journal of the Optical Society of America A* **7** 1113 – 1123

Kretzmer E R, 1952 "Statistics of television signals" *Bell Systems Technical Journal* **31** 751 – 763

Laughlin S B, 1981 "A simple coding procedure enhances a neuron's information capacity" *Zeitschrift für Naturforschung, Section C: Biosciences* **36** 910 – 912

Legge G E, 1981 "A power law for contrast discrimination" *Vision Research* **21** 457 – 467

Legge G E, Foley J M, 1980 "Contrast masking in human vision" *Journal of the Optical Society of America* **70** 1458 – 1471

Movshon J A, Thompson I D, Tolhurst D J, 1978 "Spatial and temporal contrast sensitivity of neurones in areas 17 and 18 of the cat's visual cortex" *Journal of Physiology (London)* **283** 101 – 120

Nachmias J, Sansbury R V, 1974 "Grating contrast: discrimination may be better than detection" *Vision Research* **14** 1039 – 1042

Olshausen B, Field D J, 1997 "Sparse coding with an overcomplete basis set: a strategy employed by V1?" *Vision Research* **37** 3311 – 3325

Párraga C A, Brelstaff G J, Trościanko T, Moorhead I, 1998 "Color and luminance information in natural scenes" *Journal of the Optical Society of America A* **15** 563 – 569

Párraga C A, Tolhurst D J, Trościanko T, 2000 "The human visual system is optimised for processing the spatial information in natural visual images" *Current Biology* **10** 35 – 38

Peli E, 1990 "Contrast in complex images" *Journal of the Optical Society of America A* **7** 2032 – 2040

Pelli D G, Zhang L, 1991 "Accurate control of contrast on microcomputer displays" *Vision Research* **31** 1337 – 1350

Rohaly A M, Ahumada A J, Watson A B, 1997 "Object detection in natural backgrounds predicted by discrimination performance and models" *Vision Research* **37** 3225 – 3235

Shapley R M, Enroth-Cugell C, 1984 "Visual adaptation and retinal gain controls" *Progress in Retinal Research* **3** 263 – 346

Srinivasan M V, Laughlin S B, Dubs A, 1982 "Predictive coding: a fresh view of inhibition in the retina" *Proceedings of the Royal Society of London, Series B* **216** 427 – 459

Tadmor Y, Tolhurst D J, 1994 "Discrimination of changes in the second-order statistics of natural and synthetic images" *Vision Research* **34** 541 – 554

Tadmor Y, Tolhurst D J, 2000 "The Difference-of-Gaussians receptive-field model and the contrasts in natural scenes" *Vision Research* **40** 3145 – 3157

Tolhurst D J, Barfield L P, 1978 "Interactions between spatial frequency channels" *Vision Research* **18** 951 – 958

Tolhurst D J, Tadmor Y, 1997a "Discrimination of changes in the slopes of the amplitude spectra of natural images: band limited contrast and psychometric functions" *Perception* **26** 1011 – 1025

Tolhurst D J, Tadmor Y, 1997b "Band-limited contrast in natural images explains the detectability of changes in the amplitude spectra" *Vision Research* **37** 3203 – 3215

Tolhurst D J, Tadmor Y, Arthurs G, 1996 "The detection of changes in the amplitude spectra of natural images is explained by a band-limited local contrast model" *SPIE—Human Vision and Electronic Imaging* **2657** 154 – 165

Tolhurst D J, Tadmor Y, Tang Chao, 1992 "Amplitude spectra of natural images" *Ophthalmic and Physiological Optics* **12** 229 – 232

Tolhurst D J, Thompson I D, 1981 "On the variety of spatial frequency selectivities shown by neurons in area 17 of the cat" *Proceedings of the Royal Society of London, Series B* **213** 183 – 199

Walsh G, Charman W N, 1988 "Visual sensitivity to temporal change in focus and its relevance to the accommodation response" *Vision Research* **28** 1207 – 1221

Watson A B, Robson J G, 1981 "Discrimination at threshold—labelled detectors in human vision" *Vision Research* **21** 1115 – 1122

Watt R J, Morgan M J, 1983 "The recognition and representation of edge blur: Evidence for spatial primitives in human vision" *Vision Research* **23** 1465 – 1477