

Natural image statistics mediate brightness ‘filling in’

Steven C. Dakin^{1*} and Peter J. Bex²

¹Department of Visual Science, and ²Department of Visual Rehabilitation Research, Institute of Ophthalmology, University College London, 11–43 Bath Street, London EC1V 9EL, UK

Although the human visual system can accurately estimate the reflectance (or lightness) of surfaces under enormous variations in illumination, two equiluminant grey regions can be induced to appear quite different simply by placing a light–dark luminance transition between them. This illusion, the Craik–Cornsweet–O’Brien (CCOB) effect, has been taken as evidence for a low-level ‘filling-in’ mechanism subserving lightness perception. Here, we present evidence that the mechanism responsible for the CCOB effect operates not via propagation of a neural signal across space but by amplification of the low spatial frequency (SF) structure of the image. We develop a simple computational model that relies on the statistics of natural scenes actively to reconstruct the image that is most likely to have caused an observed series of responses across SF channels. This principle is tested psychophysically by deriving classification images (CIs) for subjects’ discrimination of the contrast polarity of CCOB stimuli masked with noise. CIs resemble ‘filled-in’ stimuli; i.e. observers rely on portions of the stimuli that contain no information *per se* but that correspond closely to the reported perceptual completion. As predicted by the model, the filling-in process is contingent on the presence of appropriate low SF structure.

Keywords: vision; lightness and brightness; natural images

1. INTRODUCTION

The problem of determining the *lightness* of a surface from an image is under-constrained: the amount of light falling on the retina is affected both by the reflectance of objects and the intensity of their illumination. To operate effectively with a changing illuminant, the visual system must disentangle the contribution of these two factors by making *assumptions* about the world, and in particular by relying on the *contrast* between regions to reveal their relative lightness (Land & McCann 1971). This has the side-effect of making the system poor at estimating absolute luminance. It is this aspect of lightness perception that would appear to be the basis of many illusions including the well-known Craik–Cornsweet–O’Brien effect (CCOB) (O’Brien 1958; Craik 1966; Cornsweet 1970), where the perceived lightness of a region of uniform luminance can be profoundly altered by the presence of a luminance gradient along all or part of the region’s enclosing boundary. Figure 1c shows a variant on this effect constructed by manipulating a high-contrast natural image (figure 1a). Within this image gross changes in surface lightness, for example between the hair and forehead, produce large changes in luminance (the plot below figure 1a shows the luminance along a horizontal slice through the image, at the location indicated by the dashed line). These gradual changes are captured by the *low spatial frequencies* (SFs) of the image (figure 1b). Figure 1c demonstrates that filtering the original image with a centre-surround, Laplacian-of-Gaussian (LoG) filter, to minimize coarse changes in luminance, produces a series of light–dark transitions at the locations of the edges in the original image, interspersed by large uniform-grey areas. The regions corresponding to the forehead and hair are now of identical luminance (the

inset verifies this) but continue to induce a strong but demonstrably *illusory* sense of relative lightness and darkness.

Thus, the contrast polarity of visual borders seems critical in determining the relative perceived lightness of surfaces, and theoretical efforts to understand this phenomenon have focused on ‘low-level’ mechanisms driven by local contour information. However, the *context* within which these light–dark contours are presented can modify the magnitude of the CCOB effect. In particular, the illusion can be reduced by having other features conspire to make it more likely that the two surfaces in question have similar reflectance (Knill & Kersten 1991; Purves *et al.* 1999). Knill & Kersten (1991), for example, showed that when cues to surface curvature suggest that the luminance gradient is due to shading, rather than to a reflectance change, the magnitude of the effect is reduced. Thus, whatever neural mechanism is being driven by local border information, its output can be modified by the high-level interpretation of a scene. However, the illusion can be driven by pure border information in the absence of coherent high-level information; figure 1d shows a black–white noise pattern filtered in the same way as figure 1c. This pattern also produces a strong illusory percept of light and dark but unlike figure 1c, where prior knowledge about the scene (e.g. ‘hair tends to be darker than skin’) could be guiding our percept, this illusion must be being driven only by contour information. Thus although prior knowledge of the world is critical, there is a low-level mechanism involved in this phenomenon, which produces a sense of relative lightness in the absence of high-level knowledge about the scene. In this paper we focus on this component of the CCOB effect. Several candidates for this ‘low-level’ neural mechanism have been proposed based on local energy (Burr & Morrone 1994) and neural networks (Gerrits & Vendrik 1970; Davidson & Whiteside 1971; Cohen & Grossberg 1984; Grossberg & Todovorac

*Author for correspondence (s.dakin@ucl.ac.uk).

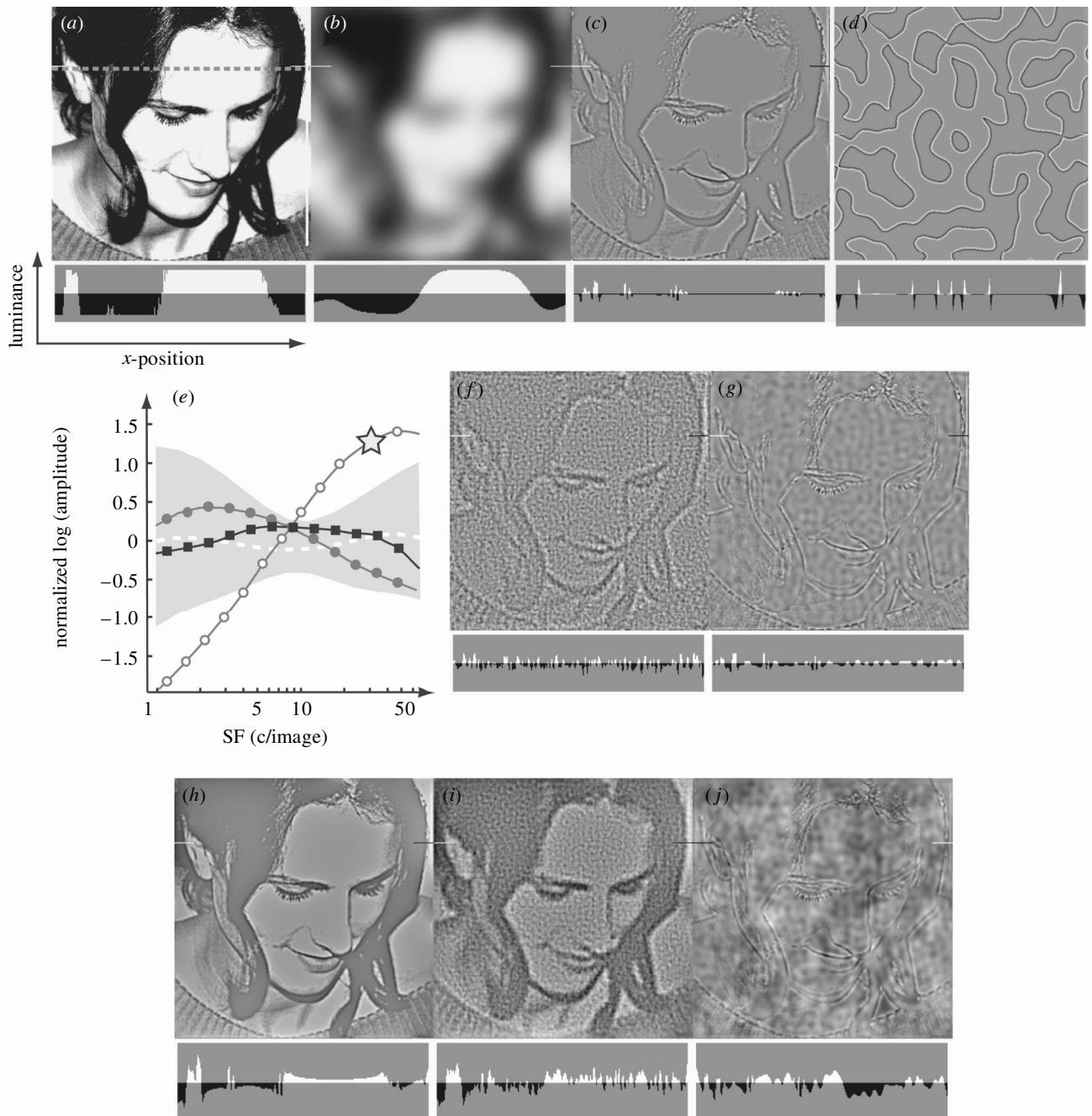


Figure 1. (a) An image and, inset below, the luminance along a horizontal slice through it (indicated by the dashed line). Coarse-scale luminance change is carried by (b) the low SF structure, but when this is attenuated by a centre-surround filter (LoG) as in (c), a strong sense of light and dark is retained. This is an illusion—the inset plot below shows that hair and face are now of identical luminance—and is a variant of the CCOB effect. This effect suggests that light–dark borders initiate ‘filling in’ of brightness across space, which can arise in the absence of ‘high-level’ image structure (e.g. as in (d)). (e) Plots of amplitude against SF for various images. Dashed white line, mean response to 1000 natural images is near constant for a bank of log Gabor bandpass SF filters with similar widths in octaves (Field 1987); the shaded region shows 95% confidence intervals on this estimate. The response to (a) (filled circles) is attenuated at low SFs by LoG filtering (open circles) but low SF structure is not completely removed. This residual low SF structure drives the illusion because ‘scrambling’ information either in the (f) high or (g) low SF range (cut-off point is indicated by the star in (e)) preserves the effect in the former but not the latter case. The luminance noise between the original contours (see insets below (f) and (g)) confounds spatial ‘filling-in’ mechanisms. We propose that ‘filling in’ is the result of normalizing the responses of a filter bank to conform to the expected response to a natural image (shading in (e)); see Appendix A for details. (h,i) Reconstructions of (c) and (f) based on this principle restore physical differences between the face and hair regions (see insets in (h) and (i)) by effectively *boosting* low frequencies (the line of filled squares in (e) is the amplitude plot for (h)). (j) This process generates random low SF structure for (g), mirroring our noisy perception.

1988; Paradiso & Nakayama 1991; Rudd & Arrington 2001). These models share a conceptual similarity, which is that they all use significant image features to infer a brightness value that is then propagated (at a cortical level) to 'fill-in' regions of homogenous luminance. The exception is Campbell *et al.* (1971, 1978) who observed that the appearance of a square-wave grating is not significantly altered by the removal of its low SF (fundamental) component, especially at low contrasts (Campbell *et al.* 1971, 1978; Sullivan & Georgeson 1977; Burr 1987), and concluded that the visual system *assumes* that images have square-wave structure in the absence of evidence to the contrary.

These approaches are confounded by the observation that the CCOB effect can persist when noise is added to the stimulus (see Burr 1987; figure 1*f*). This manipulation introduces numerous features and luminance variation in the formerly homogenous regions (see inset in figure 1*f*) and it is difficult to see how a spatial filling-in process could resist such luminance variation, or why it should in figure 1*f* but not in figure 1*g*. In the case of Campbell's proposition one can see neither how the visual system would spot the 'square-wave' SF profile in such a noisy image, nor, if it did, how a 'simple default' law could account for the fact that the average brightness of noise and other stimuli can be influenced by the Craik–O'Brien illusion (Burr 1987). We now present an alternative model based on the known properties of the visual system and of natural scenes.

2. THE MODEL

An important source of redundancy in natural images is that their amplitude spectra scales in inverse proportion to SF such that $\text{amplitude}(f) = c \cdot 1/f^\alpha$, where amplitude is averaged across all orientations, c is a constant, f is SF and α represents the negative slope on log–log coordinates. The value of α varies from image to image, but lies within a fairly narrow range (0.7–1.5) and is commonly referred to as the $1/f$ statistic. The cause of this property of natural images is contentious, but is thought to originate from correlations across space of both illuminant and material properties—which in turn produce a correlation in luminances (Ruderman 1997)—and the predominance of *edges* in natural scenes (typically arising from occlusion), which have a $1/f$ spectrum and whose phase structure is correlated across SFs. When images with this property are processed by a bank of scaled filters (i.e. that have constant bandwidth on log axes) that resemble those found in the primate visual system (Field & Tolhurst 1986) the amplitude of the responses across SF channels is roughly constant (Field 1987). This property is demonstrated by the dashed white line in figure 1*e* which shows the mean response amplitude of a bank of such filters (log Gabors) to over 1000 calibrated natural images (van Hateren & van der Schaaf 1998). The 95% confidence intervals on that mean profile are shown by the shaded region. The response of such a system to the image in figure 1*a* (filled circles in figure 1*e*) is typical of this distribution.

The open circles in figure 1*e* show the response of the filter bank to the LoG-filtered image shown in figure 1*c*. This illustrates an important property of LoGs: although they are frequently referred to as 'bandpass' filters (i.e.

completely attenuating SFs outside their pass band), in reality they are 'broad-band' filters passing all SFs and merely reducing the amplitude of components above or below their peak SF. This distinction is highlighted with LoG filtering of natural images, in which low SFs are relatively over-represented (recall that they have a $1/f$ amplitude spectrum) making it more likely that the low SFs notionally removed by LoG filtering are not only still present but are also *visible*. In fact, the slope of the LoG-filtered image does not differ greatly from the steepest slopes observed in a large sample of natural images (compare the open circles with the steepest positive slope accommodated by the shaded 95% confidence intervals). Thus, although LoG filtering has removed the slow-changing luminance structure from the image to produce a result that is dominated by large areas of uniform luminance, the low SF structure is still present and its phase is encoded in the shape of the luminance waveform at the location of edges. We argue that it is this low frequency structure that drives the illusion. To demonstrate this more explicitly we generated two versions of figure 1*c*, where we randomized the phase of high (figure 1*f*) or low (figure 1*g*) SFs (above or below 30 cycles per image, indicated by a star in figure 1*e*). The resultant images have identical power spectra to figure 1*c*. Figure 1*f* shows that scrambling high SFs destroys the fine-scale contour structure but preserves the illusion; figure 1*g* shows that scrambling low SFs maintains fine contours but destroys the illusion.

We have formulated a simple model of lightness perception that can explain these findings. As in many models of early visual processing, the model represents the structure of any image with a bank of scaled filters or *channels*. We propose that the response gains of each channel are assigned strictly in proportion to $1/f^\alpha$ to produce a set of channel responses that is nearly constant (see Appendix A for details). This approach embodies the principle that the inference of lightness requires that the visual system use filter responses to *reconstruct* the source image (Blakeslee & McCourt 1999, 2001); in our model, we seek to reconstruct the image that is most likely to have generated the observed energy profile across SF. This is motivated by findings that the visual system is relatively insensitive to departures in the amplitude spectrum of a stimulus from $1/f$, when these do not produce changes in either the blur or the contrast of the stimulus (Tadmor & Tolhurst 1994). The initial response of the filter bank to the CCOB image in figure 1*c* is shown by the open circles in figure 1*e* and demonstrates that the LoG filter has attenuated low SFs. The filled squares in figure 1*e* show the channel responses of the model after re-weighting: this response profile is the closest to a flat function that the filter bank can produce from this input. The model has relatively amplified low SF structure. Figure 1*h–j* shows reconstructions of the image represented by this flattened function for three input images, respectively: the CCOB stimulus (figure 1*c*) and the two variants with scrambled high and low frequencies (figure 1*f,g*). Note that phase randomization does not change the amplitude spectrum and is therefore identical for figure 1*c,f,g*. For the CCOB image in figure 1*h*, the low SF amplification produces luminance differences between regions of the image that are perceived as light and dark. In figure 1*i*, where high

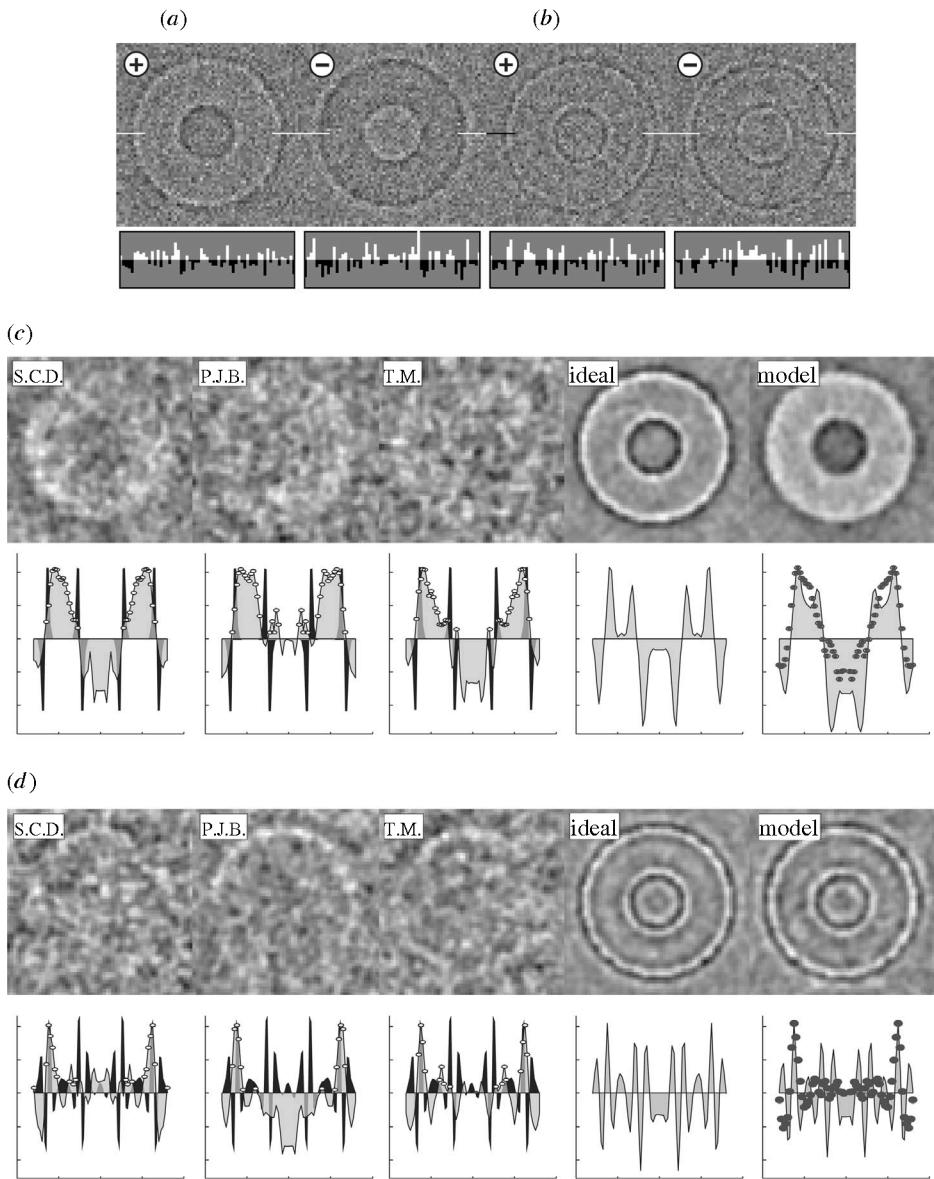


Figure 2. CI experiment. Subjects discriminated between positive- and negative-contrast polarity annuli, filtered with (a) LoG or (b) log exponential filters, and embedded in random luminance noise. (a) The LoG stimuli produce a strong 'filling in' although the mean luminance within the central region of the annulus is actually zero. (b) Removing low SFs destroys 'filling in', making polarity discrimination difficult. (c) CIs derived from three subjects' performance with LoG stimuli. In agreement with subjective experience, the CIs are 'completed' suggesting that observers rely on pixels within the 'filled-in' region *even though these points within the (unmasked) stimulus have background luminance and are therefore uninformative about contrast polarity*. Observers' perception of the stimulus stops them behaving like an ideal observer (which essentially reproduces the signal) but the predictions of the reconstruction model (rightmost image) also show completion. Radial averaging of the stimuli highlights these findings: black regions are the stimulus profile; grey-shaded regions are the average CI; and open circles are points that are statistically different from zero ($p < 0.05$). Could these results arise because luminance information in the centre of the annulus *always* affects the judgement of contrast polarity? Figure part (d) shows that this is not the case. CIs derived using true high-pass stimuli do not show 'filling in', and subjects rely on a limited portion of the stimulus at the edge of the annulus. 1D predictions from the reconstruction model (rightmost plots in (c) and (d)) capture these trends in data averaged across the three subjects (black circles).

SFs have been phase randomized, the unaltered phases of low SF components generate the coarse-scale lightness pattern of the original image. However, in figure 1*j*, the random phases of low SF components generate random lightness changes across the image that correspond to our perception of irregular lightness.

The essential difference between our model and established filling-in models is its dependence on the SF con-

tent of an image. We sought to test this aspect of the model with a psychophysical experiment. We used a classification image (CI) paradigm (Ahumada & Beard 1998), which requires that the subject perform a task (typically detection or 2AFC identification) using a target embedded in a white noise mask. In a typical CI experiment, the contrast of the target is set to be close to the observer's detection threshold. All masks that lead to suc-

cessful detection of the target are then averaged and the average mask that fails to produce detection is subtracted. Over the course of many trials, this technique estimates the correlation between the value of the mask at many locations, and the observers' response. In this way one can construct a 'behavioural receptive field' revealing which parts of the stimulus the subjects are using to perform the task. The CI paradigm has the advantage over more subjective tasks that the observers' task is fully specified (since there is a correct answer from trial to trial). Our use of CIs to investigate the CCOB illusion was inspired by Gold *et al.* (2000) who used the technique to show perceptual completion of illusory contours. Using CI methodology, one measures the extent of the illusion not by asking the subject what they see, but by formulating a task that would be difficult to perform if the subject did not see the illusion. Figure 2a,b shows examples of the stimuli used in our experiment. Observers reported if the annular stimuli were of positive (left image) or negative (right image) contrast polarity. Stimuli were filtered with (i) LoG filters, or (ii) log exponential bandpass SF filters. The latter have a sharper SF cut-off and completely remove low SF structure (see Appendix B for details). The targets were presented in levels of noise that maintained 75% correct performance on this judgement. Note the weakness of the illusion in the absence of low SFs (figure 2b), making the contrast polarity of the target much harder to discriminate. Stimuli were presented for 400 ms and were followed by 150 ms of a fresh noise mask to limit the effects of visual persistence. The task was also performed at a series of shorter exposure durations and the results confirmed earlier findings that the CCOB as reported takes *ca.* 100 ms to 'fill in' (Davey *et al.* 1998). Auditory feedback was provided after an incorrect response. The results are shown for LoG-filtered annuli in figure 2c and for log exponential-filtered annuli in figure 2d. Two-dimensional (2D) plots are CIs, one-dimensional (1D) plots are radial averages from CIs (grey-shaded regions), and the 1D waveform of the target (black regions). Open symbols show points on the radially averaged CIs that are statistically different from zero ($p < 0.05$), assessed with a *t*-test. Note that the noisy white annuli in the 2D CI plots, and the shaded grey regions in the 1D plots, reveal clear 'filling-in' of the LoG-filtered annulus.¹ This finding is counter-intuitive; the centres of the annuli are of uniform luminance and are therefore uninformative about the contrast polarity of the annulus. Nevertheless the behavioural receptive fields show that subjects are influenced by the noise that falls there and the result is a map that accords with one's perceptual completion of the figure.

A sceptic might argue that these results are to be expected: noise structure that is, by chance, darker within the annulus region will encourage an observer to report that the annulus is of negative polarity and vice versa, regardless of the target. However, a similar argument applies to the log exponential filtered annuli, the results of which are shown in figure 2d. Here, again, the central regions of the annuli are uninformative about the contrast polarity of the annulus, but in this case observers use information only at the edges of the annuli and the CI shows no evidence of filling in. It is the combination of low SFs in the image, and the presence of noise that is both suitably located and of consistent contrast polarity,

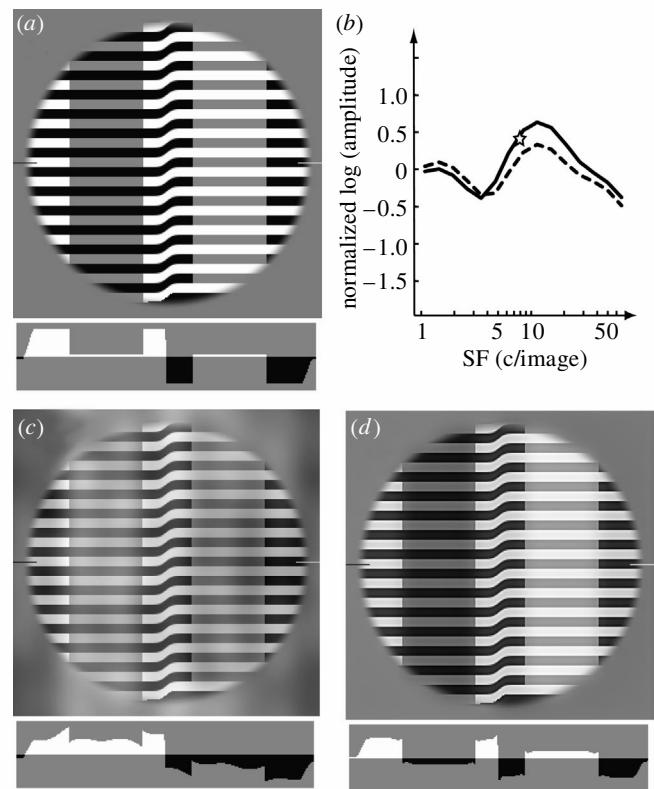


Figure 3. (a) White's effect (White 1981). The grey bars are identical on left- and right-hand sides of the plot but appear quite different. (b) The amplitude plot (solid line) reveals that replacement of light bars with grey on the left of the image, and dark with grey on the right, has introduced a second peak at low SFs in addition to the main peak due to the periodic black and white horizontal stripes. Like the CCOB effect, it is this structure that drives the illusion because removing frequencies below the point indicated by the star produces (c), which does not elicit a strong illusion. (d) The reconstruction model only subtly changes the amplitude information (dashed line in (b)) but introduces physical differences in the grey patch luminance that accord with our percept, while maintaining the integrity of the image.

which lead to the outcome that we observe. These findings extend one's subjective impression of the illusion and demonstrate that 'filling in' is contingent on the presence of low SF structure in the image.

3. DISCUSSION

Our approach shares much in common with, and indeed was in part inspired by, the model developed by Blakeslee & McCourt (1999, 2001). Both systems employ linear filtering and reconstruction by re-weighting of the filter responses. They differ in a number of respects but primarily in that our system uses purely isotropic mechanisms (compared with the broadly oriented mechanisms used by Blakeslee & McCourt) and in the re-weighting scheme employed, which is driven by image statistics in our case and is a fixed function of frequency in the case of the Blakeslee & McCourt model. Although the Blakeslee & McCourt (1999) model is, to our knowledge, untested on Craik-Cornsweet stimuli, it does predict White's effect (White 1981), a filling-in illusion that is not

based on bandpass stimuli (like the CCOB effect). We wondered whether such effects might not also be explicable in terms of a re-weighting of under-represented low SFs. Figure 3 shows White's illusion, a stimulus that is broad-band in terms of SF. All grey bars within the stimulus are similar (see luminance plot in figure 3a) but appear quite different on the left and right side of the image. This is a particularly interesting phenomenon because the induced change in lightness is in the opposite direction to the well-known *simultaneous brightness contrast* effect (where a grey patch appears lighter on a dark background than on a light background). Bars on the left of the White's stimulus are induced to appear darker, even though their surrounds are predominantly dark, and bars on the right are induced to appear lighter by their predominantly light surrounds. Figure 3b (solid line) shows two peaks in the amplitude information for this image (again derived using a bank of log Gabor filters with octave bandwidth). The rightmost peak (at *ca.* 16 cycles per image) is due to the periodic black-white stripes; the leftmost peak (at *ca.* 1.5 cycles per image) is due to the placement of the grey bars, which differentially affect the mean luminance of the left and right halves of the image. We can confirm that low SFs (below 4 cycles per image, indicated by the star in figure 3b) drive White's illusion by removing them, and observing that the effect is now largely absent (figure 3c). The model's re-weighted SF components (dashed line in figure 3b) only very slightly differ from the original (solid line in figure 3b) but this modest change in channel output produces a substantial change in the predicted lightness of the grey bars in an image reconstructed from this representation (figure 3d). Note also that the results appear neither unduly blurred nor discontinuous across SF. Thus, we have identified a commonality between two well-established illusions that has allowed us to explain them using a single simple computational model. Many other lightness illusions are elaborate variations on simultaneous brightness contrast and preliminary findings suggest that an adapted version of the model, which incorporates a local contrast gain control mechanism, can also accommodate this class of illusion (Dakin & Bex 2002).

The research was funded by a BBSRC grant no. 31/S17766 to the authors. We thank Fred Kingdom, Mark Georgeson and Joshua Solomon for useful discussions of this project.

ENDNOTES

¹The negative 'filling in' of the central circular region observed for S.C.D. and T.M. is less reliable because, as a consequence of having performed polar averaging on a Cartesian sampling grid, it is based on a much smaller sample of noise values. That is why points on the 1D shaded plots have typically failed to reach statistical significance within the central circular region.

²The reader may be wondering why we simply do not directly calculate the weights required to force the image to have a near zero slope. This would be possible if all channels were truly independent (i.e. for infinitely narrow SF tuned channels) and did not interact. Channels in the visual system have neither of these properties.

APPENDIX A: NATURAL SCENE ANALYSIS AND MODELLING

Estimates of SF amplitude were made using a conventional bank of log Gabor filters (Field 1987), which are defined in the Fourier domain as:

$$G(f) = \exp\left(-\frac{[\ln(f/f_{\text{peak}})]^2}{2[\ln(\sigma/f_{\text{peak}})]}\right), \quad (\text{A } 1)$$

where f_{peak} always varied from 1 cycles per image to the Nyquist frequency of the image, in half-octave steps. σ was set to 1.0 octaves, in broad agreement with recent electrophysiological studies of SF bandwidth in macaque V1 (Bredfeldt & Ringach 2002; Sceniak *et al.* 2002). We ensured accurate estimation of low SF information by padding images (to twice their size) prior to convolution. For the natural image analysis, we applied this filtering process to the first 1000 images of a calibrated image set (van Hateren & van der Schaaf 1998), each cropped to 256 pixels square. For each image, we estimated energy across SF (i.e. the variance of each filter response). We then plotted $\log(\text{energy})$ against $\log(\text{SF})$ and performed a least-squares fit of this function to a straight line. The mean slope of these fits was -0.04 (i.e. on average, slightly low-pass), and although this value is, as expected, close to zero (Field 1987), we used it rather than zero for subsequent modelling, on the basis that it captured minor departures from independence of each channel, as might be introduced by, for example, the increased size of low-frequency filters.

Computationally, we pose lightness perception as a reconstruction problem, where the visual system must infer the source image that is most likely to have caused the particular observed response of the filter bank. The filter bank yields two sources of information: filtered images at various SFs and their amplitudes. The sum of the filtered images, each weighted by their original amplitude, will return the original stimulus (provided that the filter bank adequately tiles the SF domain). Similarly, the model reconstructs the source image from the sum of the filtered images, but after having actively inferred each filtered image's contribution to the image (i.e. the relative amplitude of each band). If the source image is a natural scene, this imposes two constraints on this process.

- (i) The amplitude of each octave-bandwidth component should contribute as $1/f_{\text{peak}}^{\alpha}$.
- (ii) The sum of these re-weighted channels—call this a *candidate reconstruction*—should have near-constant amplitude across SF.

The model works iteratively, making a series of candidate reconstructions from the original filter responses with amplitudes set using different values of α , then filtering each candidate reconstruction with the same log Gabor filter bank to determine its amplitude slope, until the model has found a value of α that has produced a reconstruction with a slope near to -0.04 (the mean slope derived from our natural image set).²

In practice, we do not envisage an active reconstruction process occurring in human visual processing; rather this is simply a way of evaluating the way in which images are represented when encoded via the responses of a bank of SF filters. Lightness perception and lightness illusions are based wholly on this response; departures from veridical image representation occur when the weights of the individual channels differ significantly from those original values. We conjecture that interactions among SF channels may serve to normalize the contrast response range for natural images, at the expense of veridical representation of luminance.

APPENDIX B: CLASSIFICATION IMAGE EXPERIMENT

(a) Apparatus

The experiment was written in the MATLAB programming environment (MathWorks Ltd), incorporating elements of the PSYCHTOOLBOX (Brainard 1997) and VIDEOTOOBOX (Pelli 1997). Stimuli were displayed on a LaCie colour monitor (screen resolution: 640×480 pixels running at 75 Hz) driven by the built-in graphics card of a Macintosh computer. We achieved pseudo-12 bit contrast resolution in grey-scale by attenuating and combining the RGB outputs from the graphics card (Pelli & Zhang 1991) and then amplifying and copying the resulting signal to all three guns of the monitor. The display was calibrated using a photometer and linearized using look-up tables in software.

(b) Stimuli

Stimuli were 64×64 pixel images (viewed at 115 cm; 1 pixel subtended 1 arc min) composed of a Gaussian/white noise mask added to an annulus. The annulus was defined as:

$$A(r) = \begin{cases} 1 & (r \geq r_{\text{in}}) \& (r \leq r_{\text{out}}) \\ 0 & \text{otherwise} \end{cases}, \quad (\text{B } 1)$$

where r_{in} , the inner radius, was 9.6 arc min and r_{out} , the outer radius, was 25.6 arc min. This image was filtered in one of two ways. The first stimulus (figure 2a,b) was generated by convolving the annulus with an isotropic LoG filter:

$$\nabla^2 G(r) = \frac{1}{\pi \sigma^4} \left[\left(1 - \frac{r^2}{2\sigma^2} \right) e^{-r^2/2\sigma^2} \right], \quad (\text{B } 2)$$

where $\sigma = 1.0$ arc min. The second stimulus (figure 2c,d) was constructed using a bandpass filter that has a sharper cut-off: a log-exponential filter, defined in the frequency domain as:

$$L(f) = \exp \left(-\frac{\ln 2 |\ln(f/f_{\text{peak}})|^3}{[\sigma \ln 2]^3} \right). \quad (\text{B } 3)$$

We used a filter with a peak SF (f_{peak}) of 14.1 cycles per degree (to match approximately the SF of the LoG stimulus) and set the bandwidth parameter, σ , to 0.6 octaves. This produced a filter with a bandwidth (half-height at half-width, estimated directly from the power spectrum) of 0.45 octaves, and the bandwidth resulting from convolving this filter with the annulus was 0.6 octaves.

Stimuli were composed of a target stimulus (of variable contrast) added to a white noise mask (fixed root mean square contrast of 8%). All stimuli were normalized to have a mean luminance of 50 cd m^{-2} . Fresh masking noise was generated on each trial and stored (with 8 bit accuracy) according to the stimulus polarity and the subjects' response.

(c) Procedure

Observers were first familiarized with noise-free versions of the stimuli. On each trial of the experimental phase, they were then presented with a target masked by white noise and were required to judge whether the target was

of positive or negative contrast polarity. Stimuli were presented for 400 ms and were followed by a 150 ms mask (a fresh sample of white noise) to reduce visual persistence of the target. No systematic relationship was found between the subject's response and the sample of noise used in the post-mask. The experiment continued when the subject made a key-press, with errors being indicated by an audible beep. QUEST, an adaptive staircase method (Watson & Pelli 1983), was used to adjust the contrast of the target until subjects were performing at 75% correct identification. For subject S.C.D. these contrast levels for the two conditions were 28.0% and 47%, respectively; for P.J.B. they were 30% and 51% and for T.M. 27.5% and 58%. Observers each completed 3000 trials per condition, run in blocks of between 200 and 1000 trials. Conditions were not interleaved.

(d) Classification images

Subjects performing the psychophysical task could make one of two possible responses (R^+ , R^-) given either a positive or negative contrast polarity stimulus (S^+ and S^-). This gives four stimulus-response combinations: S^+R^+ , S^-R^- , S^+R^- , S^-R^+ . If one computes the average noise image for each stimulus-response combination and denotes them $\mu_{S^+R^+}$, $\mu_{S^-R^-}$ etc., then one can compute the CI (the correlation between the luminance of each noise pixel and the observers' response) simply by taking the difference between the sum of all noise fields from trials when the subject reported a positive polarity and the sum of all the noise fields when they reported a negative polarity stimulus:

$$C = (\mu_{S^+R^+} + \mu_{S^-R^-}) - (\mu_{S^-R^-} + \mu_{S^+R^+}). \quad (\text{B } 4)$$

2D CIs were smoothed using a Gaussian filter ($\sigma = 1$ pixel) and because our stimuli were radially symmetric, we reduced the degrees of freedom in our dataset by averaging pixels falling within a fixed range of distances from the centre (Abbey *et al.* 1999) over a number of distances ranging from zero to the half-width of the image. This gives 1D CIs. CIs have *normalized magnitude* (matched to the magnitude of the stimulus in the data plots) and have means of zero. Statistical significance was assessed by converting values in the CIs into *t*-values.

REFERENCES

Abbey, C. K., Eckstein, M. P. & Bochud, F. 1999 Estimation of human-observer templates for 2 alternative forced choice tasks. *Proc. SPIE* **3663**, 284–295.

Ahumada, A. J. & Beard, B. L. 1998 Response classification images in vernier acuity. *Investigative Ophthalmol. Vis. Sci.* **39**(Suppl. 4), S1109.

Blakeslee, B. & McCourt, M. E. 1999 A multiscale spatial filtering account of the White effect, simultaneous brightness contrast and grating induction. *Vision Res.* **39**, 4361–4377.

Blakeslee, B. & McCourt, M. E. 2001 A multiscale spatial filtering account of the Wertheimer–Benary effect and the corrugated Mondrian. *Vision Res.* **41**, 2487–2502.

Brainard, D. H. 1997 The psychophysics toolbox. *Spatial Vision* **10**, 433–436.

Bredfeldt, C. E. & Ringach, D. L. 2002 Dynamics of spatial frequency tuning in macaque V1. *J. Neurosci.* **22**, 1976–1984.

Burr, D. C. 1987 Implications of the Craik–O'Brien illusion for brightness perception. *Vision Res.* **27**, 1903–1913.

Burr, D. C. & Morrone, M. C. 1994 The role of features in structuring visual images. *Ciba Found Symp.* **184**, 129–141; discussion 141–146, 269–271.

Campbell, F. W., Howell, E. R. & Robson, J. G. 1971 The appearance of gratings with and without the fundamental Fourier component. *J. Physiol.* **217**, 17–18.

Campbell, F. W., Howell, E. R. & Johnstone, J. R. 1978 A comparison of threshold and suprathreshold appearance of gratings with components in the low and high spatial frequency range. *J. Physiol.* **284**, 193–201.

Cohen, M. A. & Grossberg, S. 1984 Neural dynamics of brightness perception: features, boundaries, diffusion and resonance. *Perception Psychophys.* **36**, 428–456.

Cornsweet, T. N. 1970 *Visual perception*. New York: Academic.

Craik, K. J. W. 1966 *The nature of psychology: a selection of papers, essays and other writings by the late K. J. W. Craik*. Cambridge University Press.

Dakin, S. C. & Bex, P. J. 2002 1/f channel reweighting predicts many aspects of lightness perception. *Perception* **31**(Suppl.), 149.

Davey, M. P., Maddess, T. & Srinivasan, M. V. 1998 The spatiotemporal properties of the Craik–O’Brien–Cornsweet effect are consistent with ‘filling-in’. *Vision Res.* **38**, 2037–2046.

Davidson, M. & Whiteside, J. 1971 Human brightness perception near sharp contours. *J. Optical Soc. Am.* **61**, 530–536.

Field, D. J. 1987 Relations between the statistics of natural images and the response properties of cortical cells. *J. Optical Soc. Am. A* **4**, 2379–2394.

Field, D. J. & Tolhurst, D. J. 1986 The structure and symmetry of simple-cell receptive-field profiles in the cat’s visual cortex. *Proc. R. Soc. Lond. B* **228**, 379–400.

Gerrits, H. J. & Vendrik, A. J. 1970 Simultaneous contrast, filling-in process and information processing in man’s visual system. *Exp. Brain Res.* **11**, 411–430.

Gold, J. M., Murray, R. F., Bennett, P. J. & Sekuler, A. B. 2000 Deriving behavioural receptive fields for visually completed contours. *Curr. Biol.* **10**, 663–666.

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

Knill, D. C. & Kersten, D. 1991 Apparent surface curvature affects lightness perception. *Nature* **351**, 228–230.

Land, E. H. & McCann, J. J. 1971 Lightness and retinex theory. *J. Optical Soc. Am.* **61**, 1–11.

O’Brien, V. 1958 Contour perception, illusion and reality. *J. Optical Soc. Am.* **48**, 112–119.

Paradiso, M. A. & Nakayama, K. 1991 Brightness perception and filling-in. *Vision Res.* **31**, 1221–1236.

Pelli, D. G. 1997 The VIDEOTOOBOX software for visual psychophysics: transforming number into movies. *Spatial Vision* **10**, 437–442.

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

Purves, D., Shimpi, A. & Lotto, R. B. 1999 An empirical explanation of the Cornsweet effect. *J. Neurosci.* **19**, 8542–8551.

Rudd, M. E. & Arrington, K. F. 2001 Darkness filling-in: a neural model of darkness induction. *Vision Res.* **41**, 3649–3662.

Ruderman, D. L. 1997 Origins of scaling in natural images. *Vision Res.* **37**, 3385–3398.

Sceniak, M. P., Hawken, M. J. & Shapley, R. 2002 Contrast-dependent changes in spatial frequency tuning of macaque V1 neurons: effects of a changing receptive field size. *J. Neurophysiol.* **88**, 1363–1373.

Sullivan, G. D. & Georgeson, M. A. 1977 The missing fundamental illusion: variation of spatio-temporal characteristics with dark adaptation. *Vision Res.* **17**, 977–981.

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

van Hateren, J. H. & van der Schaaf, A. 1998 Independent component filters of natural images compared with simple cells in primary visual cortex. *Proc. R. Soc. Lond. B* **265**, 359–366. (DOI 10.1098/rspb.1998.0303.)

Watson, A. B. & Pelli, D. G. 1983 QUEST: a Bayesian adaptive psychometric method. *Perception Psychophys.* **33**, 113–120.

White, M. 1981 The effect of the nature of the surround on the perceived lightness of grey bars within square-wave test gratings. *Perception* **10**, 215–230.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.