Contrast gain control in natural scenes

Peter J. Bex
Schepens Eye Research Institute, Harvard Medical School, Boston, MA, USA, &
UCL Institute of Ophthalmology, University College London, London, UK

Isabelle Mareschal
UCL Institute of Ophthalmology, University College London, London, UK

Steven C. Dakin
UCL Institute of Ophthalmology, University College London, London, UK

Behavioral and electrophysiological studies of visual processing routinely employ sine wave grating stimuli, an approach that has led to the development of models in which the first stage of cortical visual processing acts as a bank of narrowband local filters whose responses vary with the contrast of preferred structure falling within their receptive fields. The relevance of this approach to natural vision is currently being challenged. We examine the contrast response of the human visual system to natural scenes. The results support a narrowband approach to visual processing but require its elaboration. Unlike grating patterns, the contrast response to natural scenes depends on the phase structure at remote spatial scales, but over a limited spatial region. The results suggest that contrast gain control acts within, but not across, cortical hypercolumns and serves to reduce the difference between the responses of detectors in regions of high and low contrast. This process tends to normalize the response of the visual system across natural scenes, which contain uneven contrast distributions.

Keywords: contrast, gain control, natural images

Citation: Bex, P. J., Mareschal, I., & Dakin, S. C. (2007). Contrast gain control in natural scenes. *Journal of Vision,* 7(11):12, 1–12, http://journalofvision.org/7/11/12/, doi:10.1167/7.11.12.

Introduction

The goal of human visual processing is to generate functional information from the natural environment based on retinal images that have characteristic and broad distributions of spatial frequency (SF; Bex & Makous, 2002; Bex, Dakin, & Mareschal, 2005; Billock, 1996; Burton & Moorhead, 1987; Field, 1987; Hancock, Baddeley, & Smith, 1992; Tolhurst, Tadmor, & Chao, 1992; Ruderman, 1994; van der Schaaf & van Hateren, 1996; van Hateren & van der Schaaf, 1998) and orientation (Betsch, Einhäuser, Körding, & König, 2004; Coppola, Purves, McCoy, & Purves 1998; Hancock et al., 1992; Hansen, Essock, Zheng, & Deford, 2003; Keil & Cristóbal, 2000; Switkes, Mayer, & Sloan, 1978; van der Schaaf & van Hateren, 1996) are dynamic (Bex et al., 2005; Billock, de Guzman, & Kelso, 2001; Dong & Atick, 1995; van Hateren, 1997) and contain variable luminances and contrasts (Balboa & Grzywacz, 2000, 2003; Frazor & Geisler, 2006; Mante, Frazor, Bonin, Geisler, & Carandini, 2005; Ruderman & Bialek, 1994). Much of our understanding of visual processing is based on the results of experiments employing sine wave grating stimuli that are narrow in SF and orientation content and are presented at uniform, barely visible contrasts (Campbell & Green, 1965; Campbell & Robson, 1968). These data have been used to develop widely accepted models of early vision in which the sensitivity of a set of log-scaled SF-selective neurons (Field & Tolhurst, 1986; Lennie & Movshon, 2005; Ringach, Hawken, & Shapley, 2002) or channels (Blakemore & Campbell, 1969; Campbell & Robson, 1968; Graham & Nachmias, 1971) follows the inverted-U shape of the contrast sensitivity function (CSF). At suprathreshold contrasts, inhibitory connections attenuate neural responses through *contrast gain control* (Bonds, 1989; Carandini, Heeger, & Movshon, 1997; Geisler & Albrecht, 1992; Heeger, 1992; Morrone, Burr, & Maffei, 1982).

In behavioral research, such contrast gain control has been advanced to account for phenomena such as *contrast constancy* and *contextual modulation* of contrast perception. Contrast constancy describes the observation that while the CSF peaks at around 2–4 c/deg, suprathreshold apparent contrast is relatively invariant of SF (Blakemore, Muncey, & Ridley, 1973; Bowker, 1983; Brady & Field, 1995; Bryngdahl, 1966; Cannon, 1979; Georgeson & Sullivan, 1975; Kulikowski, 1976; St. John, Timney, Armstrong, & Szpak, 1987; Watanabe, Mori, & Hiwatashi, 1968). Contextual modulation measures how the threshold (Polat, 1999; Polat & Sagi, 1993,

doi: 10.1167/7.11.12 Received September 15, 2006; published August 31, 2007 ISSN 1534-7362

1994; Sagi & Hochstein, 1985; Snowden & Hammett, 1998; Solomon & Morgan, 2000) and suprathreshold apparent contrast (Cannon, 1993; Cannon & Fullenkamp, 1991, 1993, 1996; Chubb, Sperling, & Solomon, 1989; Snowden & Hammett, 1998; Xing & Heeger, 2000, 2001) of an image can be suppressed (i.e., increased thresholds) or enhanced (i.e., decreased thresholds) when it is presented within a surrounding image. In general, suppression occurs when the surround contrast is higher than the central contrast and enhancement occurs when it is lower (Cannon & Fullenkamp, 1993; Ejima & Takahashi, 1985; Xing & Heeger, 2001), depending on their relative sizes (Cannon & Fullenkamp, 1993), orientations (Wilson, McFarlane, & Phillips, 1983), spatial frequencies (Chubb et al., 1989; Solomon, Sperling, & Chubb, 1993), and phases (Polat, 1999; Polat & Sagi, 1993; Solomon, Watson, & Morgan, 1999). To account for center-surround effects, contrast normalization models (Carandini, Heeger, & Movshon, 1996; Foley, 1994; Heeger, 1992; Snowden & Hammett, 1998) divide the responses of neurons by a quantity proportional to the pooled activity of a large number of other neurons from a nearby cortical neighborhood. This can account for contrast inhibition, but none of these models can account for contrast enhancement, which requires accelerating (Zenger & Sagi, 1996) or multiplicative (Xing & Heeger, 2001) contrast response functions (CRFs) or collinear (selective) lateral excitation and nonselective lateral inhibition (Holmes & Meese, 2004; Meese, 2004; Polat, 1999; Solomon & Morgan, 2000).

The relevance of grating-based approaches to our understanding of visual processing in natural environments have recently been challenged (Olshausen & Field, 2005) and are currently the subject of axiomatic debate (Felsen & Dan, 2005; Rust & Movshon, 2005). Unlike gratings, the broad orientation and the SF content of natural images mean that neurons with different tuning properties respond to the same location of an image simultaneously, but not always in a manner that is predictable from their responses to individual bars or gratings (David, Vinje, & Gallant, 2004; Gallant, Connor, & Van Essen, 1998; Ringach, Hawken, & Shapley, 2002). Although many researchers have examined threshold and suprathreshold contrast perception with grating stimuli, relatively little is known about how the CRF is affected by the broad spectral content of natural images. In this paper, we measure the CRF of human observers presented with natural images in an attempt to relate our understanding of the perception of grating contrast to the perception of the contrast of natural scenes. We adapt a widely used pedestal-plus-test paradigm (Legge & Foley, 1980) to measure threshold versus contrast (TvC or "dipper") functions, and from these we are able to infer the underlying CRF of the visual system to structure at the component spatial scales of natural scenes.

Methods

Stimuli

Stimuli were generated on a PC computer using MatLab™ software and employed routines from the PsychToolbox (Brainard, 1997; Pelli, 1997). Stimuli were displayed with a GeForce4 MX440 graphics card driving a LaCieElectron22 monitor with a mean luminance of 50 cd/m^2 and a frame rate of 75 Hz. The display measured 36° horizontally (1,152 pixels), 27° vertically (864 pixels), and was positioned 57 cm from the observer in a dark room. The luminance gamma functions for each RGB color were measured separately with a Minolta CS100 photometer and were corrected directly in the graphics card's control panel to produce linear 8-bit resolution per color. The monitor settings were adjusted so that the luminance of green was twice that of red, which in turn was twice that of blue. This shifted the white point of the monitor to 0.31, 0.28 (x, y) at 50 cd/m^2. A "bit-stealing" algorithm (Tyler, 1997) was used to obtain 10.8 bits (1,785 unique levels) of luminance resolution under the constraint that no RGB value could differ from the others by more than one look up table step. This allowed us simultaneously to present accurately images with high and low contrast regions.

Natural images (4,165) were downloaded from a database of calibrated, 16-bit grayscale natural scenes (http://hlab.phys.rug.nl/archive.html) whose characteristics have been described in detail elsewhere (van Hateren & van der Schaaf, 1998). On each trial, a 1,536 × 1,024 pixel source image was selected at random from all 4,165 images and a 256 × 256 pixel area was then sampled at random from the source image and became the experimental image. We made no attempt to select images for particular content. A 1-octave band of target SFs was separated in the Fourier domain with log exponential filters:

$$A(\omega) \propto \exp\left(\frac{-\left|\ln\left(\omega/\omega_{\text{peak}}\right)\right|^3 \ln 2}{\left(b_{0.5}\ln 2\right)^3}\right), \quad (1)$$

where ω is the SF, ω_{peak} specifies the peak frequency, and $b_{0.5}$ is the half bandwidth of the filter in octaves, which was fixed at 0.5 octaves (full width = 1 octave). There were three image conditions, illustrated in the first three rows of Figure 1:

1. Band-pass condition. The standard and the test stimuli were 1 octave-wide spatially band-pass filtered natural images; our goal being to compare our results to those derived using gratings (Legge & Foley, 1980). The peak SF was 1, 2, or 4 c/deg. The RMS contrast of the standard image was fixed at one

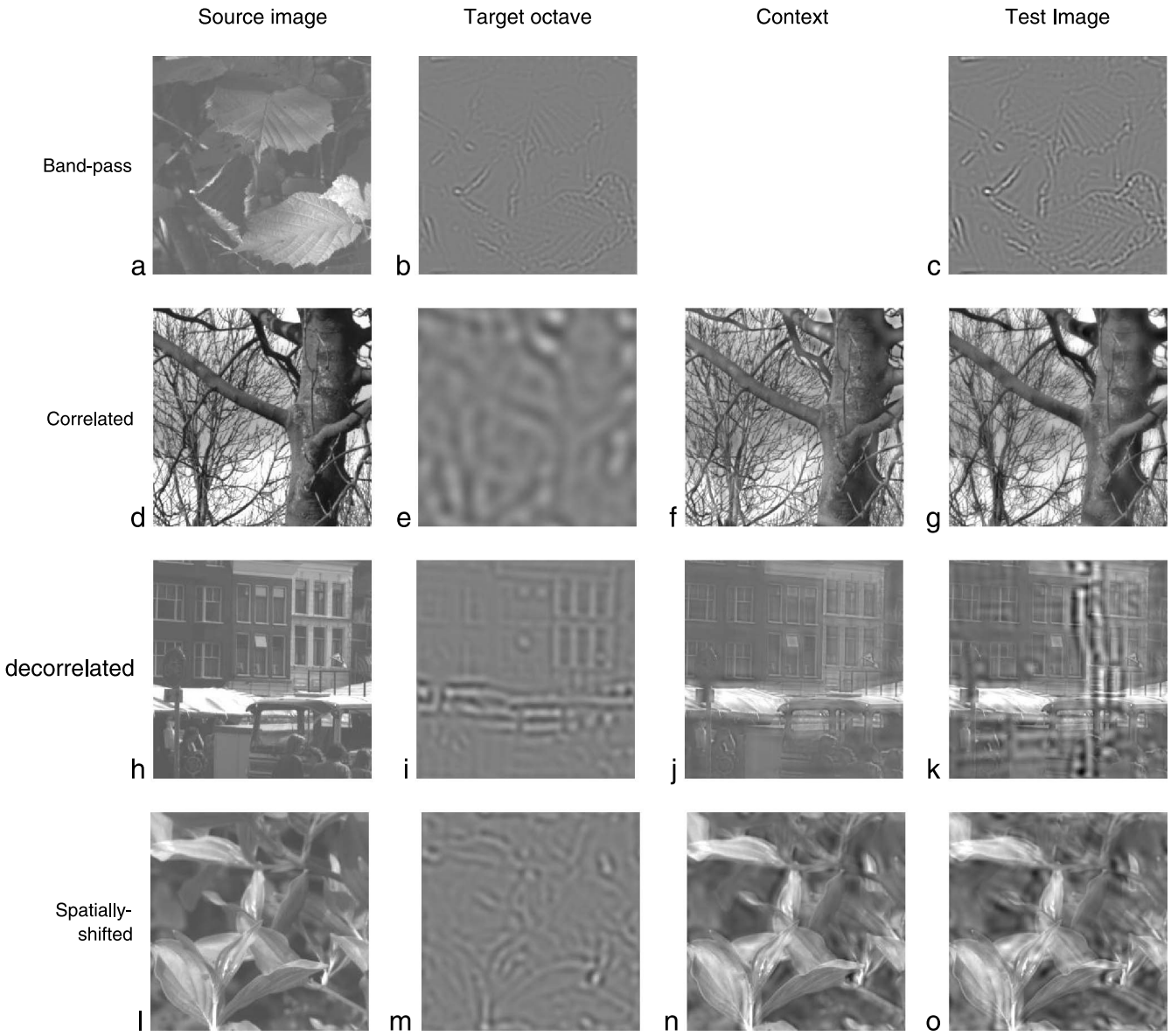

Figure 1. Illustrations of the stimuli. A source image was selected at random each trial from a database of calibrated images (left column—a, d, h, and l). A 1-octave target band was separated from the source image with a digital filter (second column—b, e, i, and m). In the band-pass condition (top row), the RMS contrast of the standard image was fixed (b) and a contrast increment was applied to the test image. This procedure was run with different contrast values (of the standard image and pedestal) to estimate the threshold-versus-contrast (TvC) function. (c). In the correlated condition (second row), the contrast of the full standard image (f) was scaled to fix the RMS contrast of the target band (e) within it. A contrast increment was applied exclusively to the target band (g) and was the only difference between the standard and the test images. The decorrelated condition (third row) was identical to the correlated condition, except that the target octave was randomly rotated and/or mirror inverted in the standard (i) and the test (k) images. In the spatially shifted condition (fourth row), the target band was decorrelated from the remainder of the image by displacing the target octave by a fixed distance in a random direction; shifts of 0.25 wavelength (n) and 1 wavelength (o) are illustrated.

of 10 levels between 0 and 0.1 (one level at 0 and 9 logarithmically spaced steps between 0.00035 and 0.1); this is the pedestal contrast. A contrast increment was added to the test image, which was otherwise identical to the standard image.

2. Correlated condition. Similar to the band-pass condition, except that the SF components outside the target octave were not discarded. The contrast of the entire image was scaled so that the RMS contrast of the target octave was set at a value between 0 and

0.1 (as in the band-pass condition) to produce the standard image, and the contrast increment was applied exclusively to the target octave. The standard and the test images were identical except for the contrast increment at the target octave. We elected to manipulate contrast of the whole image in this manner rather than, say, fixing the nontarget structure at full contrast because our method retains the slope of the original amplitude spectrum and preserves the relationships across spatial scales of the original image, which more closely emulates the effect of natural (global) variation in image lighting and contrast.

3. Decorrelated condition. Similar to the correlated condition above, the contrast of the whole image was scaled to set the contrast of the target octave at the required RMS. In this condition, the target octave in the standard and the test images was decorrelated from the rest of the image by rotation (by 90°, 180°, or 270° at random) and/or mirror reversal at random. The standard and the test images were identical (both contained the same rotation and mirror reversal of the target SF). A contrast increment was added to the target octave in the test image. On average therefore, although the source and the decorrelated images appear quite different, the amplitude spectra of the stimuli in the correlated and the uncorrelated conditions are identical, although the local distribution of amplitude across orientation may differ.
4. Spatially shifted condition. Similar to the decorrelated condition above, except that structure within the target octave was decorrelated from the context image not by rotation but by shifting it in a random direction by a variable multiple of the band's peak wavelength.

Procedure

The standard and the test images were presented for 253 ms in an 8° circular window, centered 4° to the left or right of fixation at random across trials. A raised cosine spatiotemporal envelope was used to smooth the edges of the stimuli over 0.5° and the onset and offset of the stimuli over 40 ms. The observer's 2AFC task was to identify whether the image on the left or right of fixation contained the contrast increment. Visual feedback was provided at the fixation mark, which was white following a correct response or black following an incorrect response. The contrast increment was under the control of a three-down-one-up staircase (Wetherill & Levitt, 1965) designed to converge at a contrast increment producing 79.4% correct response. The step size of the staircase was initially 2 dB and was reduced to 1 dB after three reversals. The raw data from a minimum of four runs for each condition (at least 200 trials per psychometric function) were combined and were fit with a cumulative normal function by minimization of chi-square (in which the data at each point were weighted by the binomial standard deviation). Contrast increment thresholds were estimated from the 75% correct point of the psychometric function. Ninety-five percent confidence intervals on this point were calculated with a bootstrap procedure, based on 1,000 data sets simulated from the number of experimental trials at each level tested (Foster & Bischof, 1991).

Results

Experiment 1

Figure 2 plots TvC as a function of the pedestal contrast for two observers and three SFs (shown in the caption). For all figures, green squares show the data for band-pass-only images (Condition 1, top row Figure 1), red triangles show data for the correlated broadband images (Condition 2, second row, Figure 1), and blue circles show data for decorrelated broadband images (condition 3, third row, Figure 1). In all cases, TvC functions are characteristically "dipper" shaped—contrast increment sensitivity first increases at pedestal contrasts near detection threshold, then decreases at higher pedestal levels. The magnitude of peak contrast facilitation (f) at the dip varies from 2 to 14 dB (where 1 dB = 1/20 log unit) as shown in the figure caption. This pattern of results is in good agreement with many previous studies employing grating stimuli (Chen & Tyler, 2001; Foley, 1994; Foley & Chen, 1999; Holmes & Meese, 2004; Kontsevich & Tyler, 1999; Legge & Foley, 1980; Meese, 2004; Nachmias & Sansbury, 1974; Phillips & Wilson, 1984; Ross & Speed, 1991; Solomon et al., 1999; Wilson, 1980; Wilson et al., 1983;). At low pedestal contrasts, there is little difference across conditions. However, at higher pedestal contrasts, the data separate in a similar way for all subjects and conditions. Contrast increment thresholds are consistently lower for the band-pass-only condition than for the decorrelated broadband condition, which in turn are consistently lower than for the correlated broadband conditions.

Part of the difference between band-pass and broadband images may be attributed to spectral overlap between the nontarget components in the broadband patterns and the tuning of the target detection mechanism. These components may drive up the overall response of the target detection mechanism, resulting in larger contrast increment thresholds (and shallower inferred CRFs, see below). However, note that this cannot explain the difference between the correlated and the decorrelated broadband images in which the nontarget components were, on average, identical.

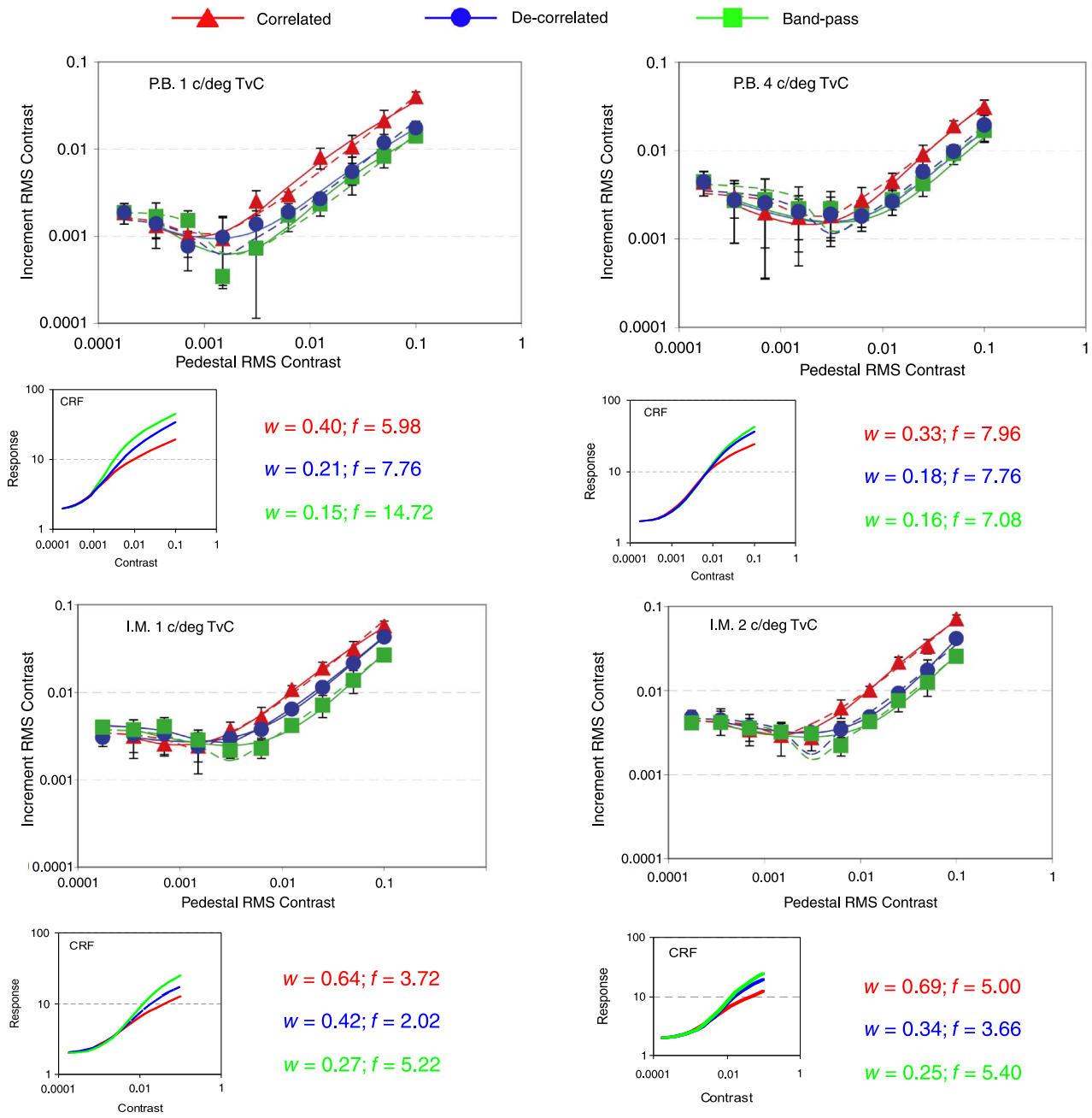

Figure 2. Threshold versus contrast (TvC) functions and contrast response functions (CRF) for two observers (P.B. and I.M.) at the three spatial frequencies indicated by the captions. The main graphs show RMS contrast increment thresholds as a function of the pedestal contrast of the target spatial frequency (SF). Green squares show data for band-pass images (Figure 1, first row), red triangles show data for correlated natural images (Figure 1, second row), and blue circles show data for hybrid images in which the target octave was decorrelated from the rest of the image by rotation and/or image reversal (Figure 1, third row). Error bars indicate 95% confidence intervals. Solid curves are the best-fitting dipper function (the first derivative of the Naka–Rushton function shown in the inset graphs). Dashed curves show the best-fitting hard threshold functions and the Weber fraction for each fit is shown in the caption (*w,* color-coded for each condition). Also shown in each caption is the magnitude of maximum contrast threshold facilitation (*f,* in dB where 1 dB = 1/20 log unit, color-coded for each condition).

The curve fits on Figure 2 show TvC functions derived from the best-fitting Naka–Rushton function:

$$r = \frac{c^P}{c^q + z}, \quad (2)$$

where r is the response, c is the RMS contrast of the pedestal, z is a constant, and p and q control the slope of the accelerating and decelerating parts of the function. Curves were fit using the fmin minimization function in MatLab that minimized the rms difference between the data and the fit, weighted by the 95% confidence intervals on each data point. To derive the TvC functions from the CRFs, we assumed that the standard and the test stimuli could be discriminated when they produced a constant difference in the underlying CRF. On this assumption and assuming additive internal noise, the TvC is the first derivative of the CRF. The first derivative is concerned with changes at in far smaller scale than contrast discrimination thresholds, which means that errors are greatest where the CRF is changing rapidly. The important differences in the present results occur at high contrasts where inferred CRFs and TvC functions are relatively stable, so these errors are likely to be small. With this in mind, we solved for the CRF whose first derivative produced the best-fitting TvC to the data by varying parameters p, q, and z. The TvC function was fit by minimizing the logarithm of least squares difference between the data and the fit, weighted by the 95% confidence intervals on the data. The inset graph in each figure shows the CRF for each TvC function. The scaling of the CRF is arbitrary, so the height of the functions was adjusted to superimpose the threshold contrast data points, where the TvC data are the same across conditions.

The CRF fits indicate that the contrast response is consistently attenuated at high contrasts for the correlated broadband condition. This contrast gain control must be mediated by structure at remote spatial scales in the broadband conditions. Furthermore, the contrast gain control is greatest when the spatial structure is correlated across spatial frequencies. We also fit the data with modified Weber functions with a term for "hard" threshold contrast:

$$\Delta c = \theta \times w + \begin{cases} \theta - c & \theta > c \\ w \times c & \theta \le c \end{cases}, \quad (3)$$

where c is the RMS contrast of the pedestal, θ is the threshold contrast, and w is the Weber fraction. The fits are shown as the dashed lines in Figure 2, and the Weber fraction for each condition is given in the legend for each figure. The Weber fits confirm the TvC fits and show that Weber fractions for the decorrelated broadband condition are on average (across all observers and conditions) 1.79 times greater than that for the band-pass condition. The correlated broadband conditions are on average 2.48 times greater.

Experiment 2

In many cases, contrast increment thresholds for the band-pass and decorrelated conditions in Experiment 1 were very similar, suggesting that observers acted as if the structure at remote spatial frequencies in the decorrelated condition was entirely absent. In Experiment 2, we manipulated the level of decorrelation in a more systematic manner than the gross rotation and mirror reversal applied in Experiment 1. The procedure was the same as for the decorrelated condition (Condition 3) in Experiment 1, except that the target octave was now spatially shifted relative to the nontarget spatial frequencies. The size of the spatial displacement varied between 0 and 64 pixels, and the direction of spatial shift was randomized across trials. The size of the horizontal and the vertical shifts for the random direction was rounded to the nearest whole number of pixels to avoid spatial blur from linear interpolation between pixels. As in Experiment 1, the target and nontarget structures were isolated from the source image with a 1-octave log exponential band-pass filter (Equation 1) with a peak SF at 1 or 4 c/deg. To avoid wrapping at the edges of the image, we took the pixels at the edges of the spatially shifted image from the adjacent pixels in the original source image so that the structure in the target band was spatially contiguous (recall that the stimulus image was a 256 pixel region cropped from the original 1,536 × 1,024 pixel image). Contrast increment detection thresholds were only collected with a pedestal contrast of 0.0125 for the target octave. This pedestal produced good separation between data across all conditions in Experiment 1.

Figure 3 shows contrast increment thresholds for two observers (P.B. and I.M.) for a pedestal contrast of 0.0125. The X-axis shows the size of the spatial shift expressed as the number of wavelengths of the target SF band. A spatial shift of zero is exactly equivalent to the correlated condition in Experiment 1; thresholds are highest at this point, confirming the observation in Experiment 1 that contrast increment sensitivity is lower for correlated images than decorrelated images, although the amplitude spectra of the images are identical.

Contrast increment thresholds fall steadily as the size of the spatial shift increases and asymptote for spatial shifts between ~0.5 and 1 cycle of the target SF. This shows that the modulation of contrast sensitivity by contrast gain control scales with SF rather than with physical spatial offset, although the spatial structure that modulates contrast sensitivity is at SFs that are remote from the target SF. In all figures, there is some evidence of a second peak in contrast increment sensitivity for spatial shifts around 0.25 wavelengths. We speculate that this effect is related to destructive summation of luminance

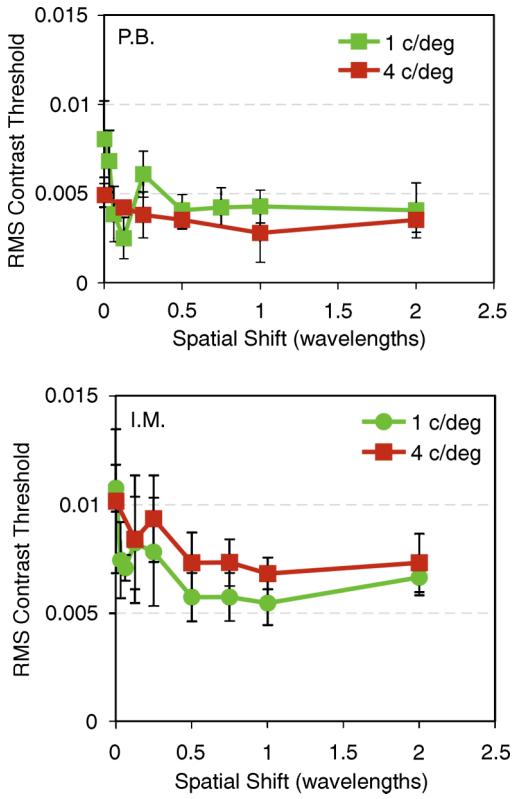

Figure 3. Contrast increment sensitivity as a function of spatial decorrelation for two observers (P.B. and I.M.) at 1 c/deg (green circles) or 4 c/deg (red squares). The *Y*-axis shows contrast increment thresholds (where the pedestal contrast of the target octave set to 0.0125) as a function of the spatial shift of the target octave, shown on the *X*-axis. Error bars indicate 95% confidence intervals.

between the target and the higher SFs. It is unlikely that lower SFs would sum destructively over such short distances.

Discussion

Contrast gain control across spatial scales depends on relative alignment

The results of Experiment 1 demonstrate that psychophysically derived CRFs for narrowband structure embedded within natural scenes are qualitatively similar to those measured with grating patterns, both having a classic "dipper" shape. The results further indicate that when the contrast of the pedestal is greater than detection threshold, contrast increment sensitivity is affected by the presence of structure at spatial scales remote from the target SF. The magnitude of sensitivity loss depends on the relative correlation of structure at these remote scales. There is a relatively small loss in sensitivity when nontarget structure is present, but uncorrelated with that of the target structure (blue circles, Figure 2). Under these conditions, sensitivity is almost as high as when the remote structure is completely removed from the image (green squares, Figure 2). The loss in sensitivity is much greater when the structure at remote scales is spatially correlated with that of the target (red triangles, Figure 2).

For grating stimuli, the relative phases of gratings of differing SF do not affect contrast sensitivity (Graham & Nachmias, 1971), suggesting that SF channels can appear to operate independently. However, for patterns composed of a small number of gratings, a relative phase difference has relatively little effect on the image, only shifting the position of beats. For natural stimuli, phase defines the location of features (Oppenheim & Lim, 1981), and for these patterns, we have now shown that contrast gain control is exerted across spatial scales and depends on the relative spatial alignment of structure.

The contrast discriminations in our stimuli must be based on the output of a population of detectors. The broad distribution of contrast in natural scenes (Balboa & Grzywacz, 2000) means that for any given natural image, individual detectors may be at different levels along their contrast response curves. The presence of contrast facilitation (a dipper) indicates that sensitivity is not based only on the response of the most sensitive detector because at low image contrasts, there would always be a detector at the most sensitive point along its response curve. This applies to the standard and the reference images in all our experimental conditions. The presence of structure at remote spatial frequencies may drive up the response of single detectors by introducing energy within the pass-band of the target SF. This could explain the difference between band-pass images and correlated broadband images. However, we would expect the same results with decorrelated images, and although there is a trend in this direction, the effect is much weaker. Decorrelating the nontarget structure may change the distribution of contrast responses—the responses of some may be driven up whereas others may be driven down by the phase change at remote spatial frequencies—but the mean is the same in correlated and decorrelated conditions.

Previous studies of contrast increment sensitivity following adaptation or masking have reported broad SF tuning of contrast gain control (Foley, 1994; Meese, 2004; Meese & Hess, 2004; Meese & Holmes, 2002; Olzak & Thomas, 1999; Ross & Speed, 1991; Ross, Speed, Morgan, 1993). Our results with correlated natural scenes confirm that contrast gain control acts across spatial scales although, owing to the broad SF content of natural scenes, our results do not specify the bandwidth of this SF tuning. However, comparison between our results with correlated and decorrelated natural scenes shows that contrast gain control depends largely on the relative alignment of structure at remote SFs, and not only their presence.

Can observers access structure within broadband images?

In our stimuli, the contrast increment was applied only to the target SF in the test stimulus; the standard and the target stimuli were otherwise identical. In addition to changing the contrast of the target structure, this manipulation also changes the overall contrast of the image. It is therefore possible that observers based their judgment on a difference in the overall contrast of the standard and the test stimuli, and not on the contrast of the target band. This does not affect our observation of the dependence of contrast increment sensitivity on the relative phase structure of natural scenes, but it does not specify whether the effect acts at narrow or global spatial scales. To address this question, on each trial we recorded the RMS contrast of the standard and the test images, as well as the RMS contrast of the target bands and the observer's responses. The data plotted in Figures 2 and 3 are based on the psychometric function fits to the narrowband contrast only.

We also plotted observers' responses as a function of the difference in contrast of the whole image. However, we were unable to fit any set of such data with a standard psychometric function. Figure 4 shows typical results from one condition (observer I.M., decorrelated condition, 4 c/deg). The *X*-axis shows the difference in RMS contrast each trial; circles show the corresponding proportion of correct trials. These data are orderly and are well fit by the cumulative normal psychometric function (solid curve).

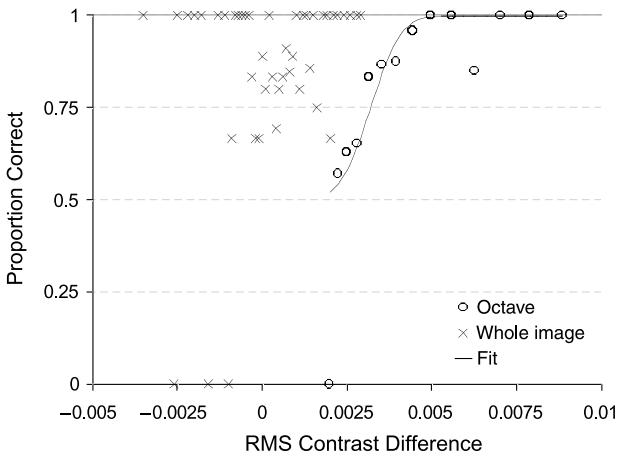

Figure 4. Typical psychometric functions for one condition (observer I.M., decorrelated condition, 4 c/deg). The *Y*-axis represents percent correct, the *X*-axis represents the RMS contrast difference between the standard and the test images, either in the target octave alone (circles) or in the RMS contrast of the whole image (crosses). The curve shows the best-fitting cumulative normal function to the single octave contrast data. It was not possible to obtain a satisfactory fit to any set of the whole image data.

Crosses show the same data plotted against the contrast difference of the overall image; these data are clearly not well behaved. In this case, the target image often had lower overall contrast than the standard image, owing to summation of the target and the nontarget pixel luminances, which often produced lower values when the target and the nontarget were misaligned compared to when they were aligned. Because observers were required to indicate the image with higher contrast, such destructive combination would have induced them to behave below chance on these conditions, but the data clearly indicate that this was not the case. These results demonstrate that observers' performance was consistent with them making their judgments on the contrast difference in the target channel before the visual system recombined the image across channels. This suggests that the visual system has access to structure embedded within natural scenes and, if required, can make discriminations based on this information. This runs contrary to the predictions of schemes such as MIRAGE (Watt & Morgan, 1985) that posit rigid combination of early filter outputs across SF.

Contrast gain control within but not across hypercolumns

Figure 3 shows the effects of gradually decorrelating structure across spatial scales by introducing systematic shifts in relative position between SFs. As in Experiment 1, thresholds were highest for correlated structure (spatial shift zero) and steadily decreased until the decorrelation reached between one half and one cycle of the target component. Thus, contrast gain control is effective across spatial scales in natural scenes, but only over a region limited to one cycle of the target SF, regardless of the target SF. The scale of these effects suggests that contrast gain control operates between neurons tuned for different spatial frequencies. We therefore speculate that contrast gain control occurs within hypercolumns. However, the observation that contrast gain control is spatially localized suggests that inhibition does not transfer across hypercolumns.

Summary

We used a contrast increment detection paradigm to examine the contrast response of the human visual system to natural scenes. The results showed that visual responses to narrowband structure are attenuated at high contrasts. This contrast gain control is mediated by structure at remote spatial frequencies but only over a limited spatial area. The results suggest that contrast gain control acts within, but not across, hypercolumns and serves to reduce the difference between the responses of detectors in

regions of high and low contrast. This process tends to normalize the response of the visual system across images with the uneven contrast distributions of natural scenes.

Acknowledgments

This research was supported by The Wellcome Trust. We thank Tim Meese and an anonymous reviewer for their helpful comments.

Commercial relationships: none.
Corresponding author: Peter Bex.
Email: peter.bex@schepens.harvard.edu.
Address: Schepens Eye Research Institute, Department of Ophthalmology, Harvard Medical School, 20 Staniford Street, Boston MA 02114, USA.

References

Balboa, R. M., & Grzywacz, N. M. (2000). Occlusions and their relationship with the distribution of contrasts in natural images. *Vision Research, 40,* 2661–2669. [PubMed]

Balboa, R. M., & Grzywacz, N. M. (2003). Power spectra and distribution of contrasts of natural images from different habitats. *Vision Research, 43,* 2527–2537. [PubMed]

Betsch, B. Y., Einhäuser, W., Körding, K. P., & König, P. (2004). The world from a cat's perspective-statistics of natural videos. *Biological Cybernetics, 90,* 41–50. [PubMed]

Bex, P. J., Dakin, S. C., & Mareschal I. (2005). Critical band masking in optic flow. *Network, 16,* 261–284. [PubMed]

Bex, P. J., & Makous, W. (2002). Spatial frequency, phase, and the contrast of natural images. *Journal of the Optical Society of America A, Optics, image science, and vision, 19,* 1096–1106. [PubMed]

Billock, V. A. (1996). Fractal properties of natural images and spatial vision. *Investigative Ophthalmology & Visual Science, 37,* 4209–4209.

Billock, V. A., de Guzman, G. C., & Kelso, S. (2001). Fractal time and 1/*f* spectra in dynamic images and human vision. *Physica D. Nonlinear Phenomena, 148,* 136–146.

Blakemore, C., & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *The Journal of Physiology, 203,* 237–260. [PubMed] [Article]

Blakemore, C., Muncey, J. P., & Ridley, R. M. (1973). Stimulus specificity in the human visual system. *Vision Research, 13,* 1915–1931. [PubMed]

Bonds, A. B. (1989). Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. *Visual Neuroscience, 2,* 41–55. [PubMed]

Bowker, D. O. (1983). Suprathreshold spatiotemporal response characteristics of the human visual system. *Journal of the Optical Society of America, 73,* 436–440. [PubMed]

Brady, N., & Field, D. J. (1995). What's constant in contrast constancy? The effects of scaling on the perceived contrast of bandpass patterns. *Vision Research, 35,* 739–756. [PubMed]

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436. [PubMed]

Bryngdahl, O. (1966). Characteristics of the visual system: Psychophysical measurements of the response to spatial sine-wave stimuli in the photopic region. *Journal of the Optical Society of America, 56,* 811–821. [PubMed]

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

Campbell, F. W., & Green, D. G. (1965). Optical and retinal factors affecting visual resolution. *The Journal of Physiology, 181,* 576–593. [PubMed] [Article]

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *The Journal of Physiology, 197,* 551–566. [PubMed] [Article]

Cannon, M. W., Jr. (1979). Contrast sensation: A linear function of stimulus contrast. *Vision Research, 19,* 1045–1052. [PubMed]

Cannon, M. W. (1993). *The influence of long-range spatial interactions on human contrast perception. Sensory research: Multimodal perspectives* (pp. 75–90). Hillsdale, NJ, Lawrence Erlbaum Associates, Inc.

Cannon, M. W., & Fullenkamp, S. C. (1991). Spatial interactions in apparent contrast: Inhibitory effects among grating patterns of different spatial frequencies, spatial positions and orientations. *Vision Research, 31,* 1985–1998. [PubMed]

Cannon, M. W., & Fullenkamp, S. C. (1993). Spatial interactions in apparent contrast: Individual differences in enhancement and suppression effects. *Vision Research, 33,* 1685–1695. [PubMed]

Cannon, M. W., & Fullencamp, S. C. (1996). A model for inhibitory lateral interactions effects in perceived contrast. *Vision Research, 36,* 1115–1125. [PubMed]

Carandini, M., Heeger, D. J., & Movshon, J. A. (1996). Linearity and gain control in V1 simple cells. In E. G. Jones & P. S. Ulinski (Eds.), *Cerebral cortex: Vol. X. Cortical models*. New York: Plenum Press.

Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *Journal of Neuroscience, 17,* 8621–8644. [PubMed] [Article]

Chen, C. C., & Tyler, C. W. (2001). Lateral sensitivity modulation explains the flanker effect in contrast discrimination. *Proceedings of the Royal Society B: Biological Sciences, 268,* 509–516. [PubMed] [Article]

Chubb, C., Sperling, G., & Solomon, J. A. (1989). Texture interactions determine perceived contrast. *Proceedings of the National Academy of Sciences of the United States of America, 86,* 9631–9635. [PubMed] [Article]

Coppola, D. M., Purves, H. R., McCoy, A. N., & Purves D. (1998). The distribution of oriented contours in the real word. *Proceedings of the National Academy of Sciences of the United States of America, 95,* 4002–4006. [PubMed] [Article]

David, S. V., Vinje, W. E., & Gallant, J. L. (2004). Natural stimulus statistics alter the receptive field structure of v1 neurons. *Journal of Neuroscience, 24,* 6991–7006. [PubMed] [Article]

Dong, D. W., & Atick, J. J. (1995). Statistics of natural time-varying images. *Network, 6,* 345–358.

Ejima, Y., & Takahashi, S. (1985). Apparent contrast of a sinusoidal grating in the simultaneous presence of peripheral gratings. *Vision Research, 25,* 1223–1232. [PubMed]

Felsen, G., & Dan, Y. (2005). A natural approach to studying vision. *Nature Neuroscience, 8,* 1643–1646. [PubMed]

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A, Optics and image science, 4,* 2379–2394. [PubMed]

Field, D. J., & Tolhurst, D. J. (1986). The structure and symmetry of simple-cell receptive-field profiles in the cat's visual cortex. *Proceedings of the Royal Society of London Series B: Biological Sciences, 228,* 379–400. [PubMed]

Foley, J. M. (1994). Human luminance pattern-vision mechanisms: Masking experiments require a new model. *Journal of the Optical Society of America A, Optics, image science, and vision, 11,* 1710–1719. [PubMed]

Foley, J. M., & Chen, C. C. (1999). Pattern detection in the presence of maskers that differ in spatial phase and temporal offset: Threshold measurements and a model. *Vision Research, 39,* 3855–3872. [PubMed]

Foster, D. H., & Bischof, W. F. (1991). Thresholds from psychometric functions: Superiority of bootstrap to incremental and probit variance estimators. *Psychological Bulletin, 109,* 152–159.

Frazor, R. A., & Geisler, W. S. (2006). Local luminance and contrast in natural images. *Vision Research, 46,* 1585–1598. [PubMed]

Gallant, J. L., Connor, C. E., & Van Essen, D. C. (1998). Neural activity in areas V1, V2 and V4 during free viewing of natural scenes compared to controlled viewing. *Neuroreport, 9,* 2153–2158. [PubMed]

Geisler, W. S., & Albrecht, D. G. (1992). Cortical neurons: Isolation of contrast gain control. *Vision Research, 32,* 1409–1410. [PubMed]

Georgeson, M. A., & Sullivan, G. D. (1975). Contrast constancy: Deblurring in human vision by spatial frequency channels. *The Journal of Physiology, 252,* 627–656. [PubMed] [Article]

Graham, N., & Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: A comparison of single-channel and multiple-channels models. *Vision Research, 11,* 251–259. [PubMed]

Hancock, P. J. B., Baddeley, R. J., & Smith, L. S. (1992). The principal components of natural images. *Network, 3,* 61–70.

Hansen, B. C., Essock, E. A., Zheng, Y., & DeFord, J. K. (2003). Perceptual anisotropies in visual processing and their relation to natural image statistics. *Network, 14,* 501–526. [PubMed]

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience, 9,* 181–197. [PubMed]

Holmes, D. J., & Meese, T. S. (2004). Grating and plaid masks indicate linear summation in a contrast gain pool. *Journal of Vision, 4*(12):7, 1080–1089, http://journalofvision.org/4/12/7/, doi:10.1167/4.12.7. [PubMed] [Article]

Keil, M. S., & Cristóbal, G. (2000). Separating the chaff from the wheat: Possible origins of the oblique effect. *Journal of the Optical Society of America A, Optics image science and vision, 17,* 697–710. [PubMed]

Kontsevich, L. L., & Tyler, C. W. (1999). Nonlinearities of near-threshold contrast transduction. *Vision Research, 39,* 1869–1880. [PubMed]

Kulikowski, J. J. (1976). Effective contrast constancy and linearity of contrast sensation. *Vision Research, 16,* 1419–1431. [PubMed]

Legge, G. E., & Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America, 70,* 1458–1471. [PubMed]

Lennie, P., & Movshon, J. A. (2005). Coding of color and form in the geniculostriate visual pathway (invited review). *Journal of the Optical Society of America A,*

Optics, image science, and vision, 22, 2013–2033. [PubMed]

Mante, V., Frazor, R. A., Bonin, V., Geisler, W. S., & Carandini, M. (2005). Independence of luminance and contrast in natural scenes and in the early visual system. *Nature Neuroscience, 8,* 1690–1697. [PubMed]

Meese, T. S. (2004). Area summation and masking. *Journal of Vision, 4*(10):8, 930–943, http://journalofvision.org/4/10/8/, doi:10.1167/4.10.8. [PubMed] [Article]

Meese, T. S., & Hess, R. F. (2004). Low spatial frequencies are suppressively masked across spatial scale, orientation, field position, and eye of origin. *Journal of Vision, 4*(10):2, 843–859, http://journalofvision.org/4/10/2/, doi:10.1167/4.10.2. [PubMed] [Article]

Meese, T. S., & Holmes, D. J. (2002). Adaptation and gain pool summation: Alternative models and masking data. *Vision Research, 42,* 1113–1125. [PubMed]

Morrone, M. C., Burr, D. C., & Maffei, L. (1982). Functional implications of cross-orientation inhibition of cortical visual cells: I. Neurophysiological evidence. *Proceedings of the Royal Society of London Series B: Biological Sciences, 216,* 335–354. [PubMed]

Nachmias, J., & Sansbury, R. V. (1974). Grating contrast: Discrimination may be better than detection. *Vision Research, 14,* 1039–1042. [PubMed]

Olshausen, B. A., & Field, D. J. (2005). How close are we to understanding v1? *Neural Computation, 17,* 1665–1699. [PubMed]

Olzak, L. A., & Thomas, J. P. (1999). Neural recoding in human pattern vision: Model and mechanisms. *Vision Research, 39,* 231–256. [PubMed]

Oppenheim, A. V., & Lim, J. S. (1981). The importance of phase in signals. *Proceedings of the IEEE, 69,* 529–541.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442. [PubMed]

Phillips, G. C., & Wilson, H. R. (1984). Orientation bandwidths of spatial mechanisms measured by masking. *Journal of the Optical Society of America A, Optics and image science, 1,* 226–232. [PubMed]

Polat, U. (1999). Functional architecture of long-range perceptual interactions. *Spatial Vision, 12,* 143–162. [PubMed]

Polat, U., & Sagi, D. (1993). Lateral Interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments. *Vision Research, 33,* 993–999. [PubMed]

Polat, U., & Sagi, D. (1994). The architecture of perceptual spatial interactions. *Vision Research, 34,* 73–78. [PubMed]

Ringach, D. L., Hawken, M. J., & Shapley, R. (2002). Receptive field structure of neurons in monkey primary visual cortex revealed by stimulation with natural image sequences. *Journal of Vision, 2*(1):2, 12–24, http://journalofvision.org/2/1/2/, doi:10.1167/2.1.2. [PubMed] [Article]

Ross, J., & Speed, H. D. (1991). Contrast adaptation and contrast masking in human vision. *Proceedings of the Royal Society B: Biological Sciences, 246,* 61–69. [PubMed]

Ross, J., Speed, H. D., Morgan, M. J. (1993). The effects of adaptation and masking on incremental thresholds for contrast. *Vision Research, 33,* 2051–2056. [PubMed]

Ruderman, D. L. (1994). The statistics of natural images. *Network, 5,* 517–548.

Ruderman, D. L., & Bialek, W. (1994). Statistics of natural images: Scaling in the woods. *Physical Review Letters, 73,* 814–817. [PubMed]

Rust, N. C., & Movshon, J. A. (2005). In praise of artifice. *Nature Neuroscience, 8,* 1647–1650. [PubMed]

Sagi, D., & Hochstein, S. (1985). Lateral inhibition between spatially adjacent spatial-frequency channels? *& Psychophysics, 37,* 315–322. [PubMed]

Snowden, R. J., & Hammett, S. T. (1998). The effects of surround contrast on contrast thresholds, perceived contrast and contrast discrimination. *Vision Research, 38,* 1935–1945. [PubMed]

Solomon, J. A., & Morgan, M. J. (2000). Facilitation from collinear flanks is cancelled by non-collinear flanks. *Vision Research, 40,* 279–286. [PubMed]

Solomon, J. A., Sperling, G., & Chubb, C. (1993). The lateral inhibition of perceived contrast is indifferent to on-center/off-center segregation, but specific to orientation. *Vision Research, 33,* 2671–2683. [PubMed]

Solomon, J. A., Watson, A. B., & Morgan, M. J. (1999). Transducer model produces facilitation from opposite-sign flanks. *Vision Research, 39,* 987–992. [PubMed]

St. John, R., Timney, B., Armstrong, K. E., & Szpak, A. B. (1987). Changes in perceived contrast of suprathreshold gratings as a function of orientation and spatial frequency. *Spatial Vision, 2,* 223–232. [PubMed]

Switkes, E., Mayer, M. J., & Sloan, J. A. (1978). Spatial frequency analysis of the visual environment: Anisotropy and the carpentered environment hypothesis. *Vision Research, 18,* 1393–1399. [PubMed]

Tolhurst, D. J., Tadmor, Y., & Chao, T. (1992). Amplitude spectra of natural images. *Ophthalmic & Physiological Optics, 12,* 229–232. [PubMed]

Tyler, C. W. (1997). Colour bit-stealing to enhance the luminance resolution of digital displays on a single pixel basis. *Spatial Vision, 10,* 369–377. [PubMed]

van der Schaaf, A., & van Hateren, J. H. (1996). Modelling the power spectra of natural images: Statistics and information. *Vision Research, 36,* 2759–2770. [PubMed]

van Hateren, J. H., & van der Schaaf, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society B: Biological Sciences, 265,* 359–366. [PubMed] [Article]

van Hateren, J. H. (1997). Processing of natural time series of intensities by the visual system of the blowfly. *Vision Research, 37,* 3407–3416. [PubMed]

Watanabe, A., Mori, T., Nagata, S., & Hiwatashi, K. (1968). Spatial sine-wave responses of the human visual system. *Vision Research, 8,* 1245–1263. [PubMed]

Watt, R. J., & Morgan, M. J. (1985). A theory of the primitive spatial code in human vision. *Vision Research, 25,* 1661–1674. [PubMed]

Wetherill, G. B., & Levitt, H. (1965). Sequential estimation of points on a psychometric function. *British Journal of Mathematical and Statistical Psychology, 18,* 1–10. [PubMed]

Wilson, H. R. (1980). A transducer function for threshold and suprathreshold human vision. *Biological Cybernetics, 38,* 171–178. [PubMed]

Wilson, H. R., McFarlane, D. K., & Phillips, G. C. (1983). Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision Research, 23,* 873–882. [PubMed]

Xing, J., & Heeger, D. J. (2000). Center-surround interactions in foveal and peripheral vision. *Vision Research, 40,* 3065–3072. [PubMed]

Xing, J., & Heeger, D. J. (2001). Measurement and modeling of center-surround suppression and enhancement. *Vision Research, 41,* 571–583. [PubMed]

Zenger, B., & Sagi, D. (1996). Isolating excitatory and inhibitory nonlinear spatial interactions involved in contrast detection. *Vision Research, 36,* 2497–2513. [PubMed]