Published in final edited form as:
J Vis. ; 9(10): 1.1–119. doi:10.1167/9.10.1.

Contrast sensitivity in natural scenes depends on edge as well as spatial frequency structure

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

Samuel G. Solomon, and
Department of Physiology, University of Sydney, Sydney, NSW, Australia

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

Abstract

The contrast sensitivity function is routinely measured in the laboratory with sine-wave gratings presented on homogenous gray backgrounds; natural images are instead composed of a broad range of spatial and temporal structures. In order to extend channel-based models of visual processing to more natural conditions, we examined how contrast sensitivity varies with the context in which it is measured. We report that contrast sensitivity is quite different under laboratory than natural viewing conditions: adaptation or masking with natural scenes attenuates contrast sensitivity at low spatial and temporal frequencies. Expressed another way, viewing stimuli presented on homogenous screens overcomes chronic adaptation to the natural environment and causes a sharp, unnatural increase in sensitivity to low spatial and temporal frequencies. Consequently, the standard contrast sensitivity function is a poor indicator of sensitivity to structure in natural scenes. The magnitude of masking by natural scenes is relatively independent of local contrast but depends strongly on the density of edges even though neither greatly affects the local amplitude spectrum. These results suggest that sensitivity to spatial structure in natural scenes depends on the distribution of local edges as well as the local amplitude spectrum.

Keywords

contrast sensitivity; natural images; adaptation; masking

Introduction

Much of our understanding of visual processing is based on the results of experiments employing sine-wave grating stimuli that are narrow in spatial frequency (SF) and orientation and are presented at uniform, barely visible contrasts (Campbell & Green, 1965; Campbell & Robson, 1968). These data form the foundation of widely accepted models of early vision in which the sensitivity of a set of log-scaled SF selective neurons (Field &

Corresponding author: Peter Bex. peter.bex@schepens.harvard.edu. Address: Schepens Eye Research Institute, Department of Ophthalmology, Harvard Medical School, 20 Staniford Street, Boston, MA 02114, USA, and Institute of Ophthalmology, University College London, 11-43 Bath Street, London EC1V 9EL, UK..

Commercial relationships: none.

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).

The goal of human visual processing is to generate functional information from the natural environment, which casts retinal images that are quite unlike sine-wave grating stimuli. The relevance of the grating-based approach to understanding visual processing has recently been challenged (Olshausen & Field, 2005) and is currently the subject of axiomatic debate (Felsen & Dan,2005; Rust & Movshon, 2005). Natural images have characteristic and broad distributions of spatial frequency (Bex, Dakin, & Mareschal, 2005; Bex & Makous, 2002; Billock, 1996; Burton & Moorhead, 1987; Field, 1987; Hancock, Baddeley, & Smith, 1992; Ruderman, 1994; Tolhurst, Tadmor, & Chao, 1992; van der Schaaf & van Hateren, 1996; van Hateren & van der Schaaf, 1998), 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), and temporal frequency (Bex et al., 2005; Billock, de Guzman, & Kelso, 2001; Dong & Atick, 1995; van Hateren, 1997). A single natural image also contains a wide range of luminances and contrasts (Balboa & Grzywacz, 2000, 2003; Frazor & Geisler, 2006; Mante, Frazor, Bonin, Geisler, & Carandini, 2005; Ruderman & Bialek, 1994).

The broad orientation and SF content of natural images means 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 et al., 2002).

At supra-threshold levels, contrast gain control modulates responses (Bex, Mareschal, & Dakin, 2007; Bonds, 1989; Carandini, Heeger, & Movshon, 1997; Geisler & Albrecht, 1992; Heeger, 1992; Morrone, Burr, & Maffei, 1982), allowing the visual system to achieve *contrast constancy* (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, Nagata, & Hiwatashi, 1968) and *contextual modulation* of threshold (Polat, 1999; Polat & Sagi, 1993, 1994; Sagi & Hochstein, 1985; Snowden & Hammett, 1998; Solomon & Morgan, 2000), and supra-threshold apparent contrast (Cannon, 1993; Cannon & Fullenkamp, 1991, 1993, 1996; Chubb, Sperling, & Solomon, 1989; Snowden & Hammett, 1998; Xing & Heeger, 2000, 2001). The broad range of contrasts in natural images will activate these gain controls in normal visual processing, but their impact is poorly understood.

Although many researchers have examined threshold and supra-threshold contrast perception with grating stimuli, relatively little is known about contrast sensitivity under more natural viewing conditions. Webster and Miyahara (1997) found that adaptation to natural images reduced contrast sensitivity at low spatial frequencies. Here we establish human spatial and temporal contrast sensitivity functions under natural conditions of adaptation and masking compared with (unnatural) blank screen viewing conditions.

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 (1152 pixels),

27° vertically (864 pixels), and was positioned 57 cm from the observer, in an otherwise dark room. The luminance gamma functions for each RGB gun was 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 (1785 unique levels) of luminance resolution for the target stimuli 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 were acquired from commercial DVD movies. We elected to use these images instead of other image databases, which are often dominated by images of foliage (e.g., van Hateren & van der Schaaf, 1998), because they contained a greater variation of image content, including more people, interior images, and man-made objects. We consider that such images are a more representative sample of the everyday environment of urban humans. Elsewhere, we have examined single frames from many different commercial movies (including cartoons) and have shown that with an assumed luminance gamma value of 2, their amplitude spectra are comparable to those of images from calibrated databases (Bex et al., 2005). If there are concerns about the use of this class of image rather than those obtained in rural settings, US individuals 12 years and older spend an average of 4.7 hours per day watching television (not including time spent observing computer images and movies, U.S. Census Bureau, 2008, http://www.census.gov/compendia/statab/tables/09s1089.pdf) and so this class of image—if not entirely natural—represents a significant proportion of retinal input.

Movies were extracted as mp4 files on a chapter-by-chapter basis, using Handbrake software (http://handbrake.fr/) with compression set to zero. This produced movie chapter files varying in size from 25 MB to 643 MB, depending on chapter duration. Next, each frame from each chapter of the movie was extracted at 25 Hz using Matlab software. Each movie frame was converted from RGB to YUV to separate luminance and color information. The luminance plane (Y) was then scaled to the required rms contrast, before converting the YUV image back to RGB and saving the frame as a uint8 720 × 384 pixel matrix. The last chapter of each movie was excluded because it usually contained rolling credits, otherwise we made no attempt to select images for particular content or to avoid scene cuts.

Contrast sensitivity—Targets were 3D white noise patterns of 128 × 128 pixels by 32 frames that were band-pass filtered with log exponential filters:

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

where ω is spatial or temporal frequency, 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). For spatial frequency tuning in Experiment 1, the peak spatial frequency was fixed at 0.5–16 cycles per degree in six log steps (2 to 64 cycles per image) while the peak temporal frequency was set at 2.4 Hz (1 cycle per 32-frame movie at 75 Hz). All spatial frequencies were randomly interleaved in a single run. For temporal frequency tuning in Experiment 1, the peak temporal frequency was fixed at 0 Hz and between 2.4 and 18.75 Hz (1–8 cpi) in 4 log-spaced steps and the peak spatial frequency was fixed at 2 cycles per degree (8 cpi). All temporal frequencies were randomly interleaved in a single run. Noise patterns were presented in envelopes that were spatially circular with a diameter of 4 deg,

smoothed by a raised cosine over 0.25 deg (8 pixels), and temporally smoothed by a raised cosine over 27 ms (3 video frames).

Experiment 1

Figures 2a and 2b show contrast sensitivity as a function of spatial frequency for two observers and three test conditions. Blue squares show contrast sensitivity measured under typical experimental conditions, where the remainder of the screen was at the mean luminance (50 cd/m^2) and the whole screen was held at the mean luminance between trials. Under these conditions, contrast sensitivity peaks at around 2 c/deg, in agreement with many previous studies under similar conditions (Campbell & Green, 1965; Campbell & Robson, 1968; Hoekstra, Van der Goot, Van den Brink, & Bilsen, 1974; Savoy & McCann, 1975).

Red circles show sensitivity when targets were presented on gray screens of the mean luminance, but between trials a natural movie was displayed for a minimum of 5 s. This adapting movie caused a reduction in contrast sensitivity but only at low spatial frequencies; there was little or no loss in sensitivity at spatial frequencies above 4 c/deg, in agreement with previous work (Webster & Miyahara, 1997).

Green triangles show data for conditions in which targets were presented within a natural movie (Figure 1a) and a phase-randomized version of a natural image was presented between trials (Figure 1b). Under these conditions, there was a substantial loss in sensitivity at low spatial frequencies, and the peak in sensitivity shifted to around 8 c/deg.

Figures 2c and 2d show the same data expressed as threshold elevations relative to conventional viewing conditions (blue squares) and illustrate more clearly the low-pass tuning of sensitivity loss under adapted or masked conditions.

Figure 3 shows results for two observers, in the same format as Figure 2, contrast sensitivity as a function of temporal frequency. There is a loss in sensitivity to low temporal frequencies during adaptation to a movie and an even greater loss if the target is presented within a natural image. Masking and adaptation bring about a greater loss in sensitivity at high temporal frequencies than at high spatial frequencies, and therefore cause a rise in threshold elevations at high temporal frequencies. We suspect that this elevation is caused by the update of the adapting and masking movies, which were refreshed at 25 Hz.

Targets were presented in three conditions that were completed in separate runs:

1. Standard Contrast Sensitivity—filtered noise stimuli were presented on a uniform background of mean luminance, with a blank field presented between trials.
2. Natural Image Adaptation—filtered noise stimuli were presented on a uniform background of mean luminance. A natural movie was played at all times, except during the test intervals. The movie was presented for a minimum of 5 s between trials and 10 s before the first trial. All movie frames had the same mean luminance as the background and global rms contrast fixed at 0.5.
3. Natural Image Masking—filtered noise stimuli were presented within a natural movie. A sample frame is shown in Figure 1a, which shows a band-pass-filtered target on the right of a central green fixation point. The global rms contrast of each masking movie image was fixed at 0.5. Noise stimuli were computed at 75 Hz and so each frame of the 25-Hz movie was presented 3 times to synchronize the target with the movie. The movie and the target were presented simultaneously for 427 ms. Between test intervals, a random phase mask pattern was presented to maintain adaptation to images with some of the properties of natural scenes throughout the run. The random phase mask was constructed by adding a random phase offset (a

uniform random deviate in the interval $-\pi$ to $+\pi$) to each component of the phase spectrum of the first frame of the upcoming test movie. In order to preserve the color distribution of the source image, the same offset was added to each RGB component. The effect of this manipulation is shown in Figure 1b, which is a phase-randomized version of Figure 1a.

Procedure—One of the authors served as an observer in all experiments and three naive subjects served as observers in different sections. Each observer had normal or corrected-to-normal vision and was experienced in psychophysical experiments.

The test images were presented for 427 ms in a 4° circular window, centered 2° to the left or right of fixation, chosen at random across trials. The observer's 2AFC task was to maintain steady fixation and to identify whether the target was presented on the left or right of fixation. Visual feedback was provided at the fixation mark, which was 50 cd/m^2 (like the background) green following a correct response, or 50 cd/m^2 red following an incorrect response, and was present at all times throughout the run. The rms contrast of the target was under the control of a 3-down-1-up staircase (Wetherill & Levitt, 1965) designed to converge at a contrast increment producing 79.4% correct responses. The staircases were initialized with random start contrasts within ±4 dB of a threshold estimated in pilot runs; the step size was initially 2 dB and was reduced to 1 dB after 2 reversals. The raw data from a minimum of four runs for each condition (at least 160 trials per psychometric function) were combined and fit with a cumulative normal function by minimization of chi-square (in which the percent correct at each test contrast was weighted by the binomial standard deviation based on the number of trials presented at that contrast). Contrast increment thresholds were estimated from the 75% correct point of the best-fitting psychometric function. Ninety-five percent confidence intervals on this point were calculated with a bootstrap procedure, based on 1000 data sets simulated from the number of experimental trials at each level tested (Foster & Bischof, 1991).

Threshold elevations were estimated by dividing masked and adapted thresholds by unadapted thresholds. Ninety-five percent confidence intervals were estimated with a Monte Carlo simulation, based on 10,000 samples with the variance matched to the unadapted, adapted, and masked thresholds.

Experiment 2

In Experiment 1, the spatial and temporal amplitude spectrum of the adapting and masking movies was inversely proportional to frequency ($1/F$), as is the case in most natural images and movies (Bex et al., 2005; Billock et al., 2001; Dong & Atick, 1995; van Hateren, 1997). It is well known that the magnitude of masking and adaptation effects is proportional to the similarity between adapt and test stimuli (Blakemore & Campbell, 1969). The low-pass tuning of adaptation and masking effects may not therefore be terribly surprising, given the dominance of low-pass structure in natural images. Nevertheless, it has been argued that the log scaling of visual channels means that the total energy within each octave is approximately constant (Brady & Field, 1995; Field, 1994), so one might expect that threshold elevations would be approximately constant across spatial (and possibly temporal) frequency. In order to examine the effect of the distribution of energy in the masking pattern, we repeated the masking conditions of Experiment 1 (which produced the most robust effect) but with movies whose amplitude spectra had been "whitened", so that the global energy at all spatial frequencies was approximately equal. This was achieved by measuring the slope of the amplitude spectrum of over 1,000 movie frames, selected at random. The amplitude spectrum was computed using the fft2() function in Matlab with zero padding to twice the original image size and truncated with a circular Tukey window

(Ramirez, 1985). The amplitude was summed across all orientations within one-octave bands. Log amplitude versus log frequency was fit with linear regression. Each movie frame was filtered with a $1/F^{\alpha}$ filter, where α was set at +1.4, the mean of the sample images. Illustrations of this whitening process are shown in Figure 4.

The chromatic structure of images varies most over low spatial scales; to remove this from our whitened images, the masking images were converted to grayscale and were composed only of the Y (luminance) plane of the YUV image. The rms contrast of the $1/F$ and whitened images was set to 0.25, which was necessary to prevent excessive overflows (where the required LUT value of any pixel falls outside the range of 0–255) of the whitened image. The results were averaged across observers and are shown in Figure 5 as threshold elevations, in the same format as in Figures 2c and 2d. Filled squares show threshold elevations as a function of target spatial frequency in grayscale masks with natural "$1/F$" amplitude spectra.

For the $1/F$ masks, threshold elevations are low-pass but with a lower threshold elevation than in Experiment 1 (Figure 2), which may be attributed to the lower rms contrast of the masking stimuli. Open circles show threshold elevations for the whitened masks. The results show weak band-pass tuning, with a peak at 2 c/deg across observers. The loss of low-pass tuning indicates that masking in natural scenes depends at least in part on the amplitude spectrum of the mask.

Experiment 3

In Experiment 1, we measured contrast sensitivity for targets that were presented within natural movies (green triangles in Figures 2 and 3) and found a pronounced loss in sensitivity compared with conventional experimental conditions. We made no attempt to control the spatial structure at the location in the movie where the target was presented; the 4° diameter targets were simply presented at a location centered 2° to the left or right of fixation.

The contrast of natural scenes is unevenly distributed (Balboa & Grzywacz, 2000, 2003; Frazor & Geisler, 2006; Mante et al., 2005; Ruderman & Bialek, 1994). Previous studies (Legge & Foley, 1980) have shown that contrast discrimination thresholds depend on the contrast of the background—termed the pedestal contrast—on which a target is presented. Contrast sensitivity should therefore depend on the contrast of the image at the target location in our movies. In Experiment 3, we estimate local rms contrast and examine how contrast sensitivity depends on this local contrast. By convention, rms contrast is defined as

$$C_{\mathrm{RMS}} = \frac{\sigma_{\mathrm{Lum}}}{\mu_{\mathrm{Lum}}}, \quad (2)$$

where σ_{Lum} is the standard deviation of pixel luminances and μ_{Lum} is the mean luminance. Standard deviation is computed as

$$\sigma = \sqrt{\frac{\Sigma X^2 - (\Sigma X)^2/N}{N}}, \quad (3)$$

where X represents the value of each sample, and N is the number of samples. We adapted this equation to estimate the local rms contrast of each pixel: The luminance of each pixel was weighted by a Gaussian over space and time, and this value was used to estimate the local rms contrast of any voxel in the movie. For a Gaussian weighted estimate, $N = 1$, corresponding to the area under a Gaussian.

Figure 6b illustrates the local rms contrast for the image shown in Figure 6a. The size of the source image is 720 × 384 pixels and the standard deviation of the Gaussian used for this example was 32 pixels, corresponding to 1° under our experimental conditions. The intensity of each pixel in the rms image indicates the local rms at that location in the image, according to the inset scale. As expected, image locations containing high local contrast produce a high local rms contrast estimate; for example, the high contrast region formed where dark hair occludes the light background wall within the dashed red circle.

The local contrast estimate is obviously affected by the standard deviation of the Gaussian used to estimate it. To examine the influence of this parameter, we estimated the distribution of local contrast for a range of standard deviation values. Figure 6c shows the logarithm of frequency of rms contrast values computed with this technique, for all frames of all movies used in this study. The different symbols show the frequency of local contrast in natural scenes for a range of Gaussian standard deviations (from 5 to 720 pixels in log steps). The area over which the rms contrast is computed has a systematic effect on the distribution of local contrast: Increasing the standard deviation reduces the proportion of low contrast estimates in the image. When the standard deviation is small, the distribution of local contrasts is heavily biased toward low contrasts (note the logarithmic scale on the ordinate), and containing relatively few regions of high contrast, in agreement with previous work (Balboa & Grzywacz, 2000) and consistent with the preponderance of mean luminance pixels in leptokurtic natural scenes (Field, 1994). This in turn arises from the relatively sparse structure present, within a given fixed-size window, at high compared to low spatial frequencies. As the standard deviation of the region of analysis increases, there is a steady increase in the proportion of regions with high contrast. For Gaussian luminance distributions, the maximum rms contrast is approximately 0.5 (so that 95.4% of luminances fall within $\pm 2\sigma$ of the mean). However, the distribution of luminances in natural scenes is not always well described by a Gaussian distribution; regions of very high rms contrast frequently occur at locations where the mean luminance is low, giving a low denominator, but where there are a small number of high luminance pixels, giving a high numerator.

In Experiment 3, we used this estimate of local rms contrast to determine target placement within the natural image. In order to equate the Gaussian estimate of local contrast with the target, we windowed the band-pass-filtered noise targets with a spatio-temporal Gaussian window rather than a raised cosine as in Experiment 1. The standard deviation of the Gaussian window was 1° spatially and 107 ms temporally, which matched the parameters of the Gaussian used to compute local rms contrast. The movie was presented for 427 ms (32 frames) and the spatial coordinates of the movie were translated such that the target always appeared at fixation, which remained at the center of the screen at all times. To ensure that the local contrast was the same in the target and comparison interval, the same movie sequence was presented twice in a temporal 2AFC paradigm. During one interval, chosen randomly, the contrast of the target was ramped on and off with a temporal Gaussian centered midway through the movie. A phase-randomized image of the first frame of the test movie (see Figure 1b) was presented for 250 ms between intervals and between trials to maintain adaptation levels. The observer's task was to fixate the colored fixation mark and to indicate which interval contained the target. The fixation point was removed during stimulus presentation but was present at all other times and was used to provide feedback as in Experiment 1. The contrast of the target noise was under the control of a staircase that converged on a contrast producing 79.4% correct identification of the interval containing the target (Wetherill & Levitt, 1965). The data from a minimum of four repetitions of each run were combined and fit as in Experiment 1 to estimate contrast detection threshold and 95% confidence intervals.

Figure 7a shows contrast detection thresholds as a function of local rms contrast at the target location. Contrast thresholds vary non-monotonically with local contrast. Thresholds rise as local rms contrast increases above zero, then fall again so that thresholds are approximately the same whether the local contrast is near zero or near 0.8. We wondered if this finding could be explained by any change in the power at the target frequency that was correlated with local contrast. We therefore calculated the slope (using the methods described above) of a 32 × 32 pixel region for 125,000 image locations randomly selected from all the movie frames in our database. Any drop in power at 2 c/deg (2 cycles per image in a 32-pixel region) should cause a precipitous increase in slope. This is because relatively more of the local energy should fall at higher spatial frequencies. Figure 7b shows the results. Blue data show the slope as a function of local contrast, red data show the best straight line fit: $y = 0.1319x - 1.589$. The same results were obtained when a 128 × 128 pixel patch was analyzed: here the slope of the best fitting straight line was 0.115. There was a small increase in the slope of the amplitude spectrum with local contrast, but nowhere near large enough to account for the present data, where the image contrast increased by a factor of more than four with little overall effect on contrast detection threshold.

This pattern of results is similar to previous results with sine-wave gratings. As the Micheslon contrast of a pedestal grating increases up to 50%, there is a general increase in contrast discrimination thresholds, with the possibility of a dipper function at very low pedestal contrasts (Legge & Foley, 1980). At higher pedestal contrasts, luminance (Whittle, 1986) and contrast (Kingdom & Whittle, 1996) discrimination thresholds saturate and may decrease again. For contrast discrimination, this reduction at high pedestal contrasts was only present at low spatial frequencies (0.5 c/deg and below), thresholds rose monotonically for 4 c/deg and 8 c/deg gratings. The 2 c/deg stimuli in Experiment 3 fall between these conditions but are in good agreement with Kingdom and Whittle (1996) and extend their finding from gratings to natural scenes.

Experiment 4

While conventional stimuli are typically periodic, natural scenes are generally characterized by aperiodic singularities arising from irregularly spaced edges and specularities (Ruderman, 1997). In Experiment 4, we examine how contrast sensitivity depends on the distribution of edge structure in natural scenes. We used the “canny” edge detector in MatLab to identify the location of edge features in each movie frame. Informal inspection showed that this algorithm tended to produce smoother and more complete contours than the alternatives, and these accorded better with our introspective assessments of the edges in the images. We found that there was very close agreement in the number and location of edges identified by all the common edge-finding algorithms, so this particular choice of edge detector should make little difference to the results.

Figure 8a shows the output of a canny edge detector (using MatLab’s default parameters for threshold and standard deviation) for the image shown in Figure 6a. The perceived and estimated locations of edges identified with this algorithm are in reasonably close correspondence. The edge maps were smoothed over space and time with a Gaussian with the same parameters as the stimulus window ($\sigma_{x,y} = 1°$, $\sigma_t = 107$ ms), and the result is illustrated in Figure 8b. This blurred edge map provided an estimate of local edge density that was used to select the position of the band-pass-filtered noise target, as inExperiment 3. In the same format as Figure 6c, Figure 8c shows distributions of edge densities (where edge density is the number of edge pixels per image pixel) or 75,000 movie frames. The standard deviation of Gaussian smoothing was 720 pixels (1 cycle/image, filled diamonds), 360 (filled triangles), 180 (filled circles), 90 (filled squares), 45 (open diamonds), 22.5 (open triangles), 11.25 (open circles), or 5.125 (open squares) pixels. The data have been fitted with Gaussians, dashed lines for filled symbols and solid lines for open symbols.

Figure 9 shows contrast detection thresholds for two observers as a function of the edge density at the target location. Contrast thresholds rose monotonically with edge density under all conditions tested. This effect is not simply related to any change in contrast with edge density because Experiment 3 showed that local contrast has no systematic relationship with contrast detection thresholds in natural scenes.

Figure 8d shows the slope of the amplitude spectrum as a function of the local edge density in the image. The amplitude spectrum was computed for a 128 × 128 pixel square region centered on the region with the required mean edge density. The amplitude spectrum was computed using the fft2() function in Matlab with zero padding to 256 × 256 deg and truncation with a circular Tukey window (Ramirez, 1985). The amplitude was summed across all orientations within one-octave bands. Log amplitude versus log frequency was fit with linear regression. This process was repeated for 10,000 random patches and the mean and ±1 standard deviation of the slope parameter is shown in Figure 8d. In line with many previous studies, the slope of this function is on average −1.39, characteristic of the "1/*F*" amplitude spectrum of nature scenes. There is a small increase in slope in regions where the edge density is high, which represents a gradual trend toward a flatter amplitude spectrum with a relative increase in energy at high spatial frequencies, but the effect is small and there is significant overlap in the slopes for all edge densities. This indicates that the differences we observed in contrast sensitivity with edge density are unrelated to changes in the average amplitude spectrum.

Experiment 5

Our results so far suggest that in natural scene contexts, contrast sensitivity depends on the distribution of both spatial frequency and edge structure. In order to examine whether our results generalize to more conventional stimuli, we employed one-dimensional grating stimuli. Targets were Gabor patches defined as

$$F_\theta = \exp\left[\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2} \times (\varpi_x s_\theta)\right] \times \exp\left[-\frac{t^2}{2\sigma_t^2}\right], \quad (4)$$

where spatial frequency (ϖ_s) was 1–16 c/deg, in log steps, orientation (θ) was 0°, $s_\theta = x\cos\theta + y\sin\theta$, and the standard deviation of the Gaussian envelope (s_x and s_y) was 2°. The 5 target spatial frequencies were randomly interleaved across trials in the same run. The observer's task was to indicate which of two temporal intervals contained the target Gabor with feedback at the fixation point. The target contrast was under the control of a staircase (Wetherill & Levitt, 1965) and the raw data from a minimum of four runs were combined and fit as described above. Contrast detection thresholds were measured under baseline (no mask) and in the presence of three masking stimuli in separate runs completed in random order. Three masking stimuli were used to separate spatial frequency distribution from edge distribution:

- **i.** a 1.3 c/deg square-wave grating,
- **ii.** a 1.3 c/deg missing fundamental (MF) grating (formed by subtracting a 1.3 c/deg sine grating of amplitude $4/\pi$ from a 1.3 c/deg square wave of unit amplitude) or
- **iii.** a 4 c/deg sine grating, corresponding to the third harmonic of the 1.3 c/deg square wave (i) and the lowest component of the missing fundamental wave (ii).

Each masking pattern was presented within the same spatio-temporal Gaussian envelope as the target Gabor.

Threshold elevations (masked divided by unmasked contrast detection thresholds) are shown for two observers in Figure 10. Blue triangles show threshold elevations for a 3F masking grating and are band-pass in line with many previous studies (Foley & Boynton, 1993; Georgeson & Georgeson, 1987; Wilson, McFarlane, & Phillips, 1983). A square-wave grating also generates a tuning function that is similar to its amplitude spectrum—threshold elevation is greatest near its fundamental frequency, which is its highest contrast component, but is elevated at other frequencies present in the pattern in roughly in proportion to their amplitudes. The broken line shows a slope of the fall off in amplitude spectrum of a square wave.

Red circles shows threshold elevation for an MF grating. The lowest spatial frequency component of the MF grating is 3F; however, this pattern produces substantial threshold elevation at lower frequencies, notably at F, the frequency of edge structure in the MF pattern. This result is in good agreement with previous studies of compound waveforms that produce beats at the difference frequency between component gratings (Derrington & Badcock, 1985; Henning, Hertz, & Broadbent, 1975).

This result and many like it have previously been interpreted as evidence for saturating or rectifying (Chubb & Sperling, 1988) non-linearities in the visual system. It is difficult to interpret the role of such non-linearities in compound stimuli like natural images as the number of possible combinations is enormous. A more parsimonious explanation that is consistent with the natural image and grating data presented here is that the level of masking depends on the correlation between the edge features of the target and mask pattern, in this case edges at the fundamental frequency of the masking patterns.

Discussion

Tuning of contrast sensitivity

This study measured spatio-temporal contrast sensitivity under conventional laboratory conditions, following adaptation to natural movies or within natural movies.Experiment 1 confirmed the well-known inverted U shape of contrast sensitivity as a function of spatial frequency (Campbell & Green, 1965; Campbell & Robson, 1968) and the low-pass tuning of contrast sensitivity as a function of temporal frequency (de Lange, 1958; Kelly, 1975; Roufs, 1972) for narrow-band stimuli presented on uniform backgrounds (Hoekstra et al., 1974; Savoy & McCann, 1975) under comparable viewing conditions.

Adaptation

Following a brief period of adaptation to dynamic natural images, the contrast sensitivity function shows a selective loss in sensitivity at low spatial frequencies, in agreement with a previous study of adaptation to stationary natural images (Webster & Miyahara, 1997). It is well known that the amplitude spectrum of natural scenes decreases at high spatial frequencies (Bex et al., 2005; Bex & Makous, 2002; Billock, 1996; Burton & Moorhead, 1987; Field, 1987; Hancock et al., 1992; Ruderman, 1994; Tolhurst et al., 1992; van der Schaaf & van Hateren, 1996; van Hateren & van der Schaaf, 1998) and temporal frequencies (Bex et al., 2005; Billock et al., 2001; Dong & Atick, 1995; van Hateren, 1997), a property that is commonly referred to as "$1/F$" scaling. It is now widely accepted that the logarithmic scaling of visual sensitivity means that the responses of the visual system to ($\approx 1/F$) natural scenes are approximately the same across spatial scales (Brady & Field, 1995; Field, 1994); this may help the efficient coding of natural scenes (Atick, 1992; Barlow, 1961; Field, 1994; Olshausen & Field, 1996). It has also been shown that adaptation aftereffects are narrowly tuned for spatial frequency (e.g., Blakemore & Campbell, 1969). Thus adapting to a grating of one spatial frequency produces a loss in sensitivity that is greatest for nearby spatial

frequencies. For temporal frequency adaptation, the effects are more complex and may be narrow-band under some conditions (Hammett & Snowden, 1995) or relatively low-pass, implying adaptation at multiple stages of visual processing (Langley & Bex, 2007).

Given the $1/F$ spectrum of natural scenes and the tuning of adaptation aftereffects, it might not be surprising that adaptation to natural movies produces a selective loss in sensitivity at low spatial (Figure 2, red circles) or temporal frequencies (Figure 3, red circles). However, the difference between unadapted and adapted conditions is nevertheless surprising for at least two reasons. Firstly, one might expect that many years of adaptation in the observers' natural environment before starting Experiment 1 would produce a similar effect to 5 s of adaptation to a natural movie during the experiment; but our results did not support this. There was a measurable loss in sensitivity following 5 s of adaptation to a natural movie. Having spent years in the natural environment, this is presumably the observers' normal state of adaptation, which is approximated by our natural movie adaptation condition. Our results suggest a more reasonable interpretation is that viewing a gray, mean luminance screen between trials under traditional experimental conditions *elevates* sensitivity at low spatial or temporal frequencies. This calls into question the generality of the standard contrast sensitivity function since our results suggest that for natural vision, where the observer is chronically adapted to dynamic natural scenes, it is more band-pass tuned than previously supposed.

The second surprising aspect of our results is that adaptation to natural movies had different effects at different spatial and temporal frequencies. The visual system is widely held to be optimally tuned to encode natural scenes (Atick, 1992; Barlow, 1961; Field, 1994), such that the response of early visual channels to high contrast natural images are similar across spatial scales (Brady & Field, 1995; Georgeson & Sullivan, 1975). If adaptation desensitized these channels in proportion to their response to the adaptor, one might expect similar adaptation effects across all spatial and temporal frequencies. The impact of adaptation was instead low-pass over spatial and temporal frequencies, which suggests that low spatio-temporal frequencies might adapt more than high. Adaptation is widely thought to reflect changes in contrast gain control following prolonged stimulation. To achieve contrast constancy, gain control has to attenuate the responses of low to mid spatial frequencies more than high and so threshold elevation should decrease as spatial frequency increases. Our results are therefore generally consistent with such gain control models.

We did observe a modest loss in sensitivity at high temporal frequencies. However, we speculate that this loss in sensitivity to high temporal frequencies arose because our natural movies were updated at 25 Hz (each movie frame was repeated three times at 75 Hz), producing a local amplitude peak at this temporal frequency.

Masking

Like adaptation, numerous masking studies have shown that contrast thresholds are elevated when a target and mask patterns are similar in spatial (Foley & Boynton, 1993; Georgeson & Georgeson, 1987; Wilson et al., 1983; see also Figure 10, blue triangles) or temporal frequency (Cass & Alais, 2006; Hess & Snowden, 1992; Snowden & Hess, 1992). The results of Experiment 1 show that the effects of masking with natural images are low-pass, with little or no effect at high spatial (Figure 2, green triangles) or temporal (Figure 2, green triangles) frequencies. This effect is similar to that for adaptation and suggests that contrast gain control attenuates the responses of low spatial frequencies more than that to high in broad-band natural scenes.

Experiment 2 examined this proposition further with whitened images, in which the amplitude spectrum is approximately equal at all spatial frequencies, as illustrated in Figure

4. The results in Figure 5 show that threshold elevations were weakly band-pass in the presence of whitened masking patterns. This suggests that attenuating the amplitude of low spatial frequencies in natural scenes reduces spatial frequency-dependent contrast gain control.

Local structure in natural scenes

Experiment 3 examined how contrast discrimination thresholds vary with the local contrast in natural scenes. The results showed a non-monotonic relationship between local rms contrast and contrast discrimination thresholds. Figure 5b shows that contrast thresholds are approximately the same whether the local contrast is near zero or near 80%. This finding somewhat surprisingly suggests that the visibility of a target in a natural scene is relatively independent of the contrast in the location the target appears. Nonetheless, the result parallels some of the results observed in studies with sine grating stimuli. As for grating stimuli, at relatively low pedestal contrasts (<50%), increment detection thresholds increase steadily with background contrast, with the possibility of an increase in sensitivity for very low background contrasts, which reflects the slope of a sigmoidal-shaped contrast response function (Legge & Foley, 1980). We recently showed that this characteristic dipper shape of the threshold versus contrast function was qualitatively similar when contrast increments were applied to narrow-band structure within natural scenes (Bex et al., 2007). At higher pedestal contrasts, discrimination thresholds rise monotonically for high spatial frequency (4 c/deg or higher) gratings but decrease for low spatial frequency (0.5 c/deg or lower) pedestals (Kingdom & Whittle, 1996). The present results show similar effects for natural images.

We have previously shown in a threshold versus contrast pedestal paradigm that the magnitude of threshold elevation depends on the correlation among structures at remote spatial frequencies (Bex et al., 2007). We argued that the results were consistent with contrast gain control effects that were local in space, but broad in spatial frequency, indicating that these effects operate within but not across cortical hypercolumns. The narrow-band targets used in the present study were uncorrelated with the spatial structure at the test location and therefore confirm and extend this research.

Taken together, these results suggest that the visibility of targets in natural scenes depends on additional factors to the local contrast. The underlying amplitude spectrum of natural scenes produces a relative loss in sensitivity to low spatial and temporal frequencies, but this only accounts for part of the contrast sensitivity function for targets embedded within natural scenes. Given that the amplitude spectra of natural scenes are remarkably similar, these other factors must account for much of the variability in contrast sensitivity in natural scenes. In Experiment 4, we considered the spatial distribution of features in natural scenes. The results, plotted in Figure 9, show that contrast detection thresholds increased monotonically with the density of edges at the target location. The lack of a systematic relationship between the amplitude spectrum and edge density (Figure 8d) and between local contrast and detection thresholds (Figure 7) indicates that local edge density plays a key role in target detection. Experiment 5 replicated this result with more traditional grating stimuli and thus establishes that this observation is not a unique property of natural images. Others have also considered a role for feature density in target detection. For a bar stimulus presented on a sine-wave grating background, sensitivity decreases as spatial frequency (edge density) increases, an effect that is attributed to local luminance adaptation, pattern masking, and contrast gain control (Freeman & Badcock, 1999).

We offer is a simple explanation for these effects in practice. In many images, high contrast areas occur when there is a large luminance difference between an occluder and its background (e.g., the area ringed in Figure 6a). Under these conditions, local variation in the

target is relatively unmasked in either side of the occluding contour, where the luminance can be relatively uniform. We therefore propose that the number of edges at the test location is an important parameter in attempting to predict contrast sensitivity in natural scenes.

Summary

We examined how contrast sensitivity to narrow-band noise patterns varies under natural and unnatural viewing conditions. Contrast sensitivity was measured as a function of spatial and temporal frequencies for band-pass-filtered noise patterns that were presented within or after adaptation to dynamic natural images. A natural movie context produced a selective loss in sensitivity mostly at low spatial and temporal frequencies. This tuning follows the "1/*F*" scaling of natural scenes and the tuning varies with the slope of the amplitude spectrum of natural scenes.

Contrast sensitivity did not increase monotonically with the image contrast at the target location but was similar for very high or very low *local* contrasts. However, contrast sensitivity did rise monotonically with the density of edge features at the test location. This property was unrelated to changes in the contrast or amplitude spectrum. The results suggest that the visibility of targets within natural contexts can be best predicted by the density of local image features.

Acknowledgments

This research was supported by The Wellcome Trust.

References

Atick JJ. Could information theory provide an ecological theory of sensory processing? Network: Computation in Neural Systems. 1992; 3:213–251.

Balboa RM, Grzywacz NM. Occlusions and their relationship with the distribution of contrasts in natural images. Vision Research. 2000; 40:2661–2669. PubMed. [PubMed: 10958916]

Balboa RM, Grzywacz NM. Power spectra and distribution of contrasts of natural images from different habitats. Vision Research. 2003; 43:2527–2537. PubMed. [PubMed: 13129540]

Barlow, HB. Possible principles underlying the transformation of sensory messages. In: Rosenblith, WA., editor. Sensory communication. MIT Press; Cambridge, MA: 1961. p. 217-234.

Betsch BY, Einhäuser W, Körding KP, König P. The world from a cat's perspective—Statistics of natural videos. Biological Cybernetics. 2004; 90:41–50. PubMed. [PubMed: 14762723]

Bex PJ, Dakin SC, Mareschal I. Critical band masking in optic flow. Network. 2005; 16:261–284. PubMed. [PubMed: 16411499]

Bex PJ, Mareschal I, Dakin SC. Contrast gain control in natural scenes. Journal of Vision. 2007; 7(11): 11–12. 12. http://journalofvision.org/7/11/12/, doi:10.1167/7.11.12. PubMed Article.

Bex PJ, Makous W. Spatial frequency, phase, and the contrast of natural images. Journal of the Optical Society of America A. 2002; 19:1096–1106. PubMed.

Billock VA. Fractal properties of natural images and spatial vision. Investigative Ophthalmology & Visual Science. 1996; 37:4209.

Billock VA, de Guzman GC, Kelso JAS. Fractal time and 1/*f* spectra in dynamic images and human vision. Physica D: Nonlinear Phenomena. 2001; 148:136–146.

Blakemore C, Campbell FW. On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. The Journal of Physiology. 1969; 203:237–260. PubMed Article. [PubMed: 5821879]

Blakemore C, Muncey JPJ, Ridley RM. Stimulus specificity in the human visual system. Vision Research. 1973; 13:1915–1931. PubMed. [PubMed: 4746989]

Bonds AB. Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. Visual Neuroscience. 1989; 2:41–55. PubMed. [PubMed: 2487637]

Bowker DO. Suprathreshold spatiotemporal response characteristics of the human visual system. Journal of the Optical Society of America. 1983; 73:436–440. PubMed. [PubMed: 6864356]

Brady N, Field DJ. What's constant in contrast constancy? The effects of scaling on the perceived contrast of bandpass patterns. Vision Research. 1995; 35:739–756. PubMed. [PubMed: 7740766]

Brainard DH. The Psychophysics Toolbox. Spatial Vision. 1997; 10:433–436. PubMed. [PubMed: 9176952]

Bryngdahl O. 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. 1966; 56:811–821. PubMed. [PubMed: 5963528]

Burton GJ, Moorhead IR. Color and spatial structure in natural scenes. Applied Optics. 1987; 26:157–170. [PubMed: 20454092]

Campbell FW, Green DG. Optical and retinal factors affecting visual resolution. The Journal of Physiology. 1965; 181:576–593. PubMed Article. [PubMed: 5880378]

Campbell FW, Robson JG. Application of Fourier analysis to the visibility of gratings. The Journal of Physiology. 1968; 197:551–566. PubMed Article. [PubMed: 5666169]

Cannon MW. Contrast sensation: A linear function of stimulus contrast. Vision Research. 1979; 19:1045–1052. PubMed. [PubMed: 532118]

Cannon, MW. Sensory research: Multimodal perspectives. Lawrence Erlbaum Associates; Hillsdale, NJ: 1993. The influence of long-range spatial interactions on human contrast perception; p. 75-90.

Cannon MW, Fullenkamp SC. Spatial interactions in apparent contrast: Inhibitory effects among grating patterns of different spatial frequencies, spatial positions and orientations. Vision Research. 1991; 31:1985–1998. PubMed. [PubMed: 1771782]

Cannon MW, Fullenkamp SC. Spatial interactions in apparent contrast: Individual differences in enhancement and suppression effects. Vision Research. 1993; 33:1685–1695. PubMed. [PubMed: 8236856]

Cannon MW, Fullenkamp SC. A model for inhibitory lateral interactions effects in perceived contrast. Vision Research. 1996; 36:1115–1125. PubMed. [PubMed: 8762716]

Carandini M, Heeger DJ, Movshon JA. Linearity and normalization in simple cells of the macaque primary visual cortex. Journal of Neuroscience. 1997; 17:8621–8644. PubMed Article. [PubMed: 9334433]

Cass J, Alais D. Evidence for two interacting temporal channels in human visual processing. Vision Research. 2006; 46:2859–2868. PubMed. [PubMed: 16684555]

Chubb C, Sperling G. Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. Journal of the Optical Society of America A, Optics and Image Science. 1988; 5:1986–2006. PubMed.

Chubb C, Sperling G, Solomon JA. Texture interactions determine perceived contrast. Proceedings of the National Academy of Sciences of the United States of America. 1989; 86:9631–9635. PubMed Article. [PubMed: 2594791]

Coppola DM, Purves HR, McCoy AN, Purves D. The distribution of oriented contours in the real word. Proceedings of the National Academy of Sciences of the United States of America. 1998; 95:4002–4006. PubMed Article. [PubMed: 9520482]

David SV, Vinje WE, Gallant JL. Natural stimulus statistics alter the receptive field structure of v1 neurons. Journal of Neuroscience. 2004; 24:6991–7006. PubMed Article. [PubMed: 15295035]

de Lange H. Research into the dynamic nature of human fovea–cortex systems with intermittent and modulated light. 1. Attenuation characteristics with white and colored light. Journal of the Optical Society of America. 1958; 48:777–784. [PubMed: 13588450]

Derrington AM, Badcock DR. Separate detectors for simple and complex grating patterns? Vision Research. 1985; 25:1869–1878. PubMed. [PubMed: 3832611]

Dong DW, Atick JJ. Statistics of natural time-varying images. Network: Computation in Neural Systems. 1995; 6:345–358.

Felsen G, Dan Y. A natural approach to studying vision. Nature Neuroscience. 2005; 8:1643–1646. PubMed.

Field DJ. 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. 1987; 4:2379–2394. PubMed.

Field DJ. What is the goal of sensory coding? Neural Computation. 1994; 6:559–601.

Field DJ, Tolhurst DJ. The structure and symmetry of simple-cell receptive-field profiles in the cat's visual cortex. Proceedings of the Royal Society of London B: Biological Sciences. 1986; 228:379–400. PubMed.

Foley JM, Boynton GM. Forward pattern masking and adaptation: Effects of duration, interstimulus interval, contrast, and spatial and temporal frequency. Vision Research. 1993; 33:959–980. PubMed. [PubMed: 8506639]

Foster DH, Bischof WF. Thresholds from psychometric functions: Superiority of bootstrap to incremental and probit variance estimators. Psychological Bulletin. 1991; 109:152–159.

Frazor RA, Geisler WS. Local luminance and contrast in natural images. Vision Research. 2006; 46:1585–1598. PubMed. [PubMed: 16403546]

Freeman AW, Badcock DR. Visual sensitivity in the presence of a patterned background. Journal of the Optical Society of America A, Optics, Image Science, and Vision. 1999; 16:979–986. PubMed.

Gallant JL, Connor CE, Van Essen DC. Neural activity in areas V1, V2 and V4 during free viewing of natural scenes compared to controlled viewing. Neuroreport. 1998; 9:2153–2158. PubMed. [PubMed: 9674611]

Geisler WS, Albrecht DG. Cortical neurons: Isolation of contrast gain control. Vision Research. 1992; 32:1409–1410. PubMed. [PubMed: 1455713]

Georgeson MA, Georgeson JM. Facilitation and masking of briefly presented gratings: Time-course and contrast dependence. Vision Research. 1987; 27:369–379. PubMed. [PubMed: 3660598]

Georgeson MA, Sullivan GD. Contrast constancy: Deblurring in human vision by spatial frequency channels. The Journal of Physiology. 1975; 252:627–656. PubMed Article. [PubMed: 1206570]

Graham N, Nachmias J. Detection of grating patterns containing two spatial frequencies: A comparison of single-channel and multiple channel models. Vision Research. 1971; 11:251–259. PubMed. [PubMed: 5579840]

Hammett ST, Snowden RJ. The effect of contrast adaptation on briefly presented stimuli. Vision Research. 1995; 35:1721–1725. PubMed. [PubMed: 7660580]

Hancock PJB, Baddeley RJ, Smith LS. The principal components of natural images. Network: Computation in Neural Systems. 1992; 3:61–70.

Hansen BC, Essock EA, Zheng Y, DeFord JK. Perceptual anisotropies in visual processing and their relation to natural image statistics. Network: Computation in Neural Systems. 2003; 14:501–526. PubMed.

Heeger DJ. Normalization of cell responses in cat striate cortex. Visual Neuroscience. 1992; 9:181–197. PubMed. [PubMed: 1504027]

Henning GB, Hertz BG, Broadbent DE. Some experiments bearing on the hypothesis that the visual system analyses spatial patterns in independent bands of spatial frequency. Vision Research. 1975; 15:887–897. PubMed. [PubMed: 1166622]

Hess RF, Snowden RJ. Temporal properties of human visual filters: Number, shapes and spatial covariation. Vision Research. 1992; 32:47–59. PubMed. [PubMed: 1502811]

Hoekstra J, Van der Goot DPJ, Van den Brink G, Bilsen FA. The influence of the number of cycles upon the visual contrast threshold for spatial sine wave patterns. Vision Research. 1974; 14:365–368. PubMed. [PubMed: 4854900]

Keil MS, Cristóbal G. 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. 2000; 17:697–710. PubMed.

Kelly DH. Spatial frequency selectivity in the retina. Vision Research. 1975; 15:665–672. PubMed. [PubMed: 1138482]

Kingdom FA, Whittle P. Contrast discrimination at high contrasts reveals the influence of local light adaptation on contrast processing. Vision Research. 1996; 36:817–829. PubMed. [PubMed: 8736217]

Kulikowski JJ. Effective contrast constancy and linearity of contrast sensation. Vision Research. 1976; 16:1419–1431. PubMed. [PubMed: 1007022]

Langley K, Bex PJ. Contrast adaptation implies two spatiotemporal channels but three adapting processes. Journal of Experimental Psychology: Human Perception and Performance. 2007; 33:1283–1296. PubMed. [PubMed: 18085943]

Legge GE, Foley JM. Contrast masking in human vision. Journal of the Optical Society of America. 1980; 70:1458–1471. PubMed. [PubMed: 7463185]

Lennie P, Movshon JA. 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. 2005; 22:2013–2033. PubMed.

Mante V, Frazor RA, Bonin V, Geisler WS, Carandini M. Independence of luminance and contrast in natural scenes and in the early visual system. Nature Neuroscience. 2005; 8:1690–1697. PubMed.

Morrone MC, Burr DC, Maffei L. Functional implications of cross-orientation inhibition of cortical visual cells—I. Neurophysiological evidence. Proceedings of the Royal Society of London B: Biological Sciences. 1982; 216:335–354. PubMed.

Olshausen BA, Field DJ. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature. 1996; 381:607–609. PubMed. [PubMed: 8637596]

Olshausen BA, Field DJ. How close are we to understanding v1? Neural Computation. 2005; 17:1665–1699. PubMed. [PubMed: 15969914]

Pelli DG. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spatial Vision. 1997; 10:437–442. PubMed. [PubMed: 9176953]

Polat U. Functional architecture of long-range perceptual interactions. Spatial Vision. 1999; 12:143–162. PubMed. [PubMed: 10221425]

Polat U, Sagi D. Lateral interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments. Vision Research. 1993; 33:993–999. PubMed. [PubMed: 8506641]

Polat U, Sagi D. The architecture of perceptual spatial interactions. Vision Research. 1994; 34:73–78. PubMed. [PubMed: 8116270]

Ramirez, RW. The FFT, fundamentals and concepts. Prentice-Hall; New Jersey: 1985.

Ringach DL, Hawken MJ, Shapley R. Receptive field structure of neurons in monkey primary visual cortex revealed by stimulation with natural image sequences. Journal of Vision. 2002; 2(1):12–24. 2. http://journalofvision.org/2/1/2/, doi:10.1167/ 2.1.2. PubMed Article. [PubMed: 12678594]

Roufs JA. Dynamic properties of vision—1. Experimental relationship between flicker and flash thresholds. Vision Research. 1972; 12:261–278. PubMed. [PubMed: 5033689]

Ruderman DL. Origins of scaling in natural images. Vision Research. 1997; 37:3385–3398. PubMed. [PubMed: 9425551]

Ruderman DL. The statistics of natural images. Network: Computation in Neural Systems. 1994; 5:517–548.

Ruderman DL, Bialek W. Statistics of natural images: Scaling in the woods. Physical Review Letters. 1994; 73:814–817. PubMed. [PubMed: 10057546]

Rust NC, Movshon JA. In praise of artifice. Nature Neuroscience. 2005; 8:1647–1650. PubMed.

Sagi D, Hochstein S. Lateral inhibition between spatially adjacent spatial-frequency channels? Perception & Psychophysics. 1985; 37:315–322. PubMed. [PubMed: 4034349]

Savoy RL, McCann JJ. Visibility of low-spatial-frequency sine-wave targets: Dependence on number of cycles. Journal of the Optical Society of America. 1975; 65:343–350. PubMed. [PubMed: 1123690]

Snowden RJ, Hammett ST. The effects of surround contrast on contrast thresholds, perceived contrast and contrast discrimination. Vision Research. 1998; 38:1935–1945. PubMed. [PubMed: 9797940]

Snowden RJ, Hess RF. Temporal frequency filters in the human peripheral visual field. Vision Research. 1992; 32:61–72. PubMed. [PubMed: 1502812]

Solomon JA, Morgan MJ. Facilitation from collinear flanks is cancelled by non-collinear flanks. Vision Research. 2000; 40:279–286. PubMed. [PubMed: 10793901]

St John R, Timney B, Armstrong KE, Szpak AB. Changes in perceived contrast of suprathreshold gratings as a function of orientation and spatial frequency. Spatial Vision. 1987; 2:223–232. PubMed. [PubMed: 3154948]

Switkes E, Mayer MJ, Sloan JA. Spatial frequency analysis of the visual environment: Anisotropy and the carpentered environment hypothesis. Vision Research. 1978; 18:1393–1399. PubMed. [PubMed: 726283]

Tolhurst DJ, Tadmor Y, Chao T. Amplitude spectra of natural images. Ophthalmic & Physiological Optics. 1992; 12:229–232. PubMed. [PubMed: 1408179]

Tyler CW. Colour bit-stealing to enhance the luminance resolution of digital displays on a single pixel basis. Spatial Vision. 1997; 10:369–377. PubMed. [PubMed: 9176946]

van der Schaaf A, van Hateren JH. Modelling the power spectra of natural images: Statistics and information. Vision Research. 1996; 36:2759–2770. PubMed. [PubMed: 8917763]

van Hateren JH. Processing of natural time series of intensities by the visual system of the blowfly. Vision Research. 1997; 37:3407–3416. PubMed. [PubMed: 9425553]

van Hateren JH, van der Schaaf A. Independent component filters of natural images compared with simple cells in primary visual cortex. Proceedings of the Royal Society of London B: Biological Sciences. 1998; 265:359–366. PubMed Article.

Watanabe A, Mori T, Nagata S, Hiwatashi K. Spatial sine-wave responses of the human visual system. Vision Research. 1968; 8:1245–1263. PubMed. [PubMed: 5682796]

Webster MA, Miyahara E. Contrast adaptation and the spatial structure of natural images. Journal of the Optical Society of America A, Optics, Image Science, and Vision. 1997; 14:2355–2366. PubMed.

Wetherill GB, Levitt H. Sequential estimation of points on a psychometric function. British Journal of Mathematics, Statistics, and Psychology. 1965; 18:1–10.

Whittle P. Increments and decrements: Luminance discrimination. Vision Research. 1986; 26:1677–1691. PubMed. [PubMed: 3617509]

Wilson HR, McFarlane DK, Phillips GC. Spatial frequency tuning of orientation selective units estimated by oblique masking. Vision Research. 1983; 23:873–882. PubMed. [PubMed: 6636547]

Xing J, Heeger DJ. Center-surround interactions in foveal and peripheral vision. Vision Research. 2000; 40:3065–3072. PubMed. [PubMed: 10996610]

Xing J, Heeger DJ. Measurement and modeling of center-surround suppression and enhancement. Vision Research. 2001; 41:571–583. PubMed. [PubMed: 11226503]

Figure 1.
Illustrations of the stimuli. (a) A single frame from a movie. This image contains a band-pass-filtered noise target to the right of fixation that was present in only the masking condition (see text for details). (b) Random phase version of (a) in which the same random phase offset has been applied in each RGB plane to preserve the chromatic properties of the original image. The green and red circles show isoluminant fixation points whose color indicated whether the observer's response was correct or incorrect.

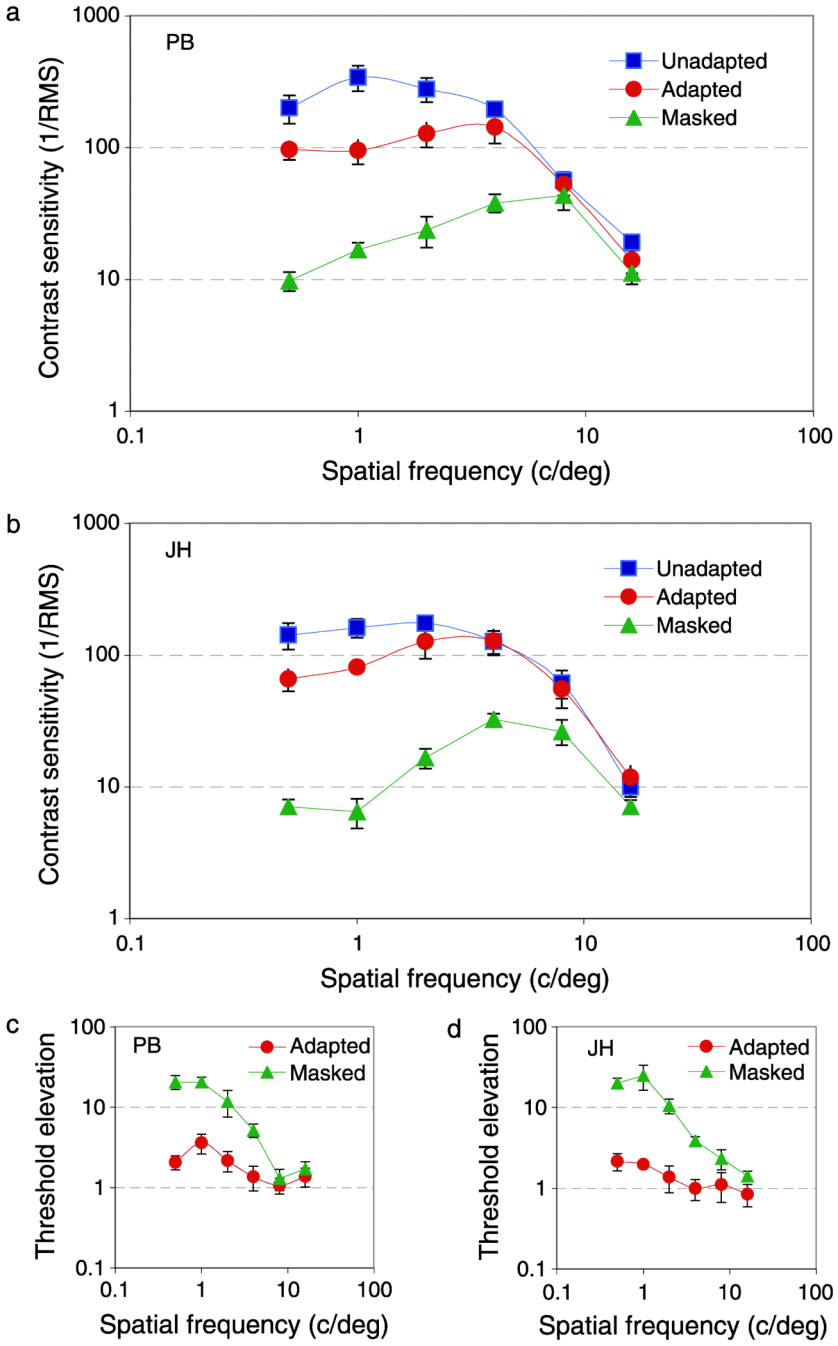

Figure 2.

(a, b) Spatial frequency contrast sensitivity and (c, d) threshold elevations for two observers (PB and JH). Blue squares show standard contrast sensitivity (in which targets were presented on mean luminance backgrounds), red circles show data from adapted conditions (in which a movie was played for at least 5 s between test periods), and green triangles show data for the masking condition (where the target was presented within a real movie). The *x*-axis shows the peak spatial frequency of the band-pass-filtered noise target, the *y*-axis shows contrast sensitivity, defined as the inverse of the rms contrast required for correct detection of the target on 75% trials. Threshold elevations are relative to standard (blue data) conditions. Error bars show ±95% confidence intervals on all figures.

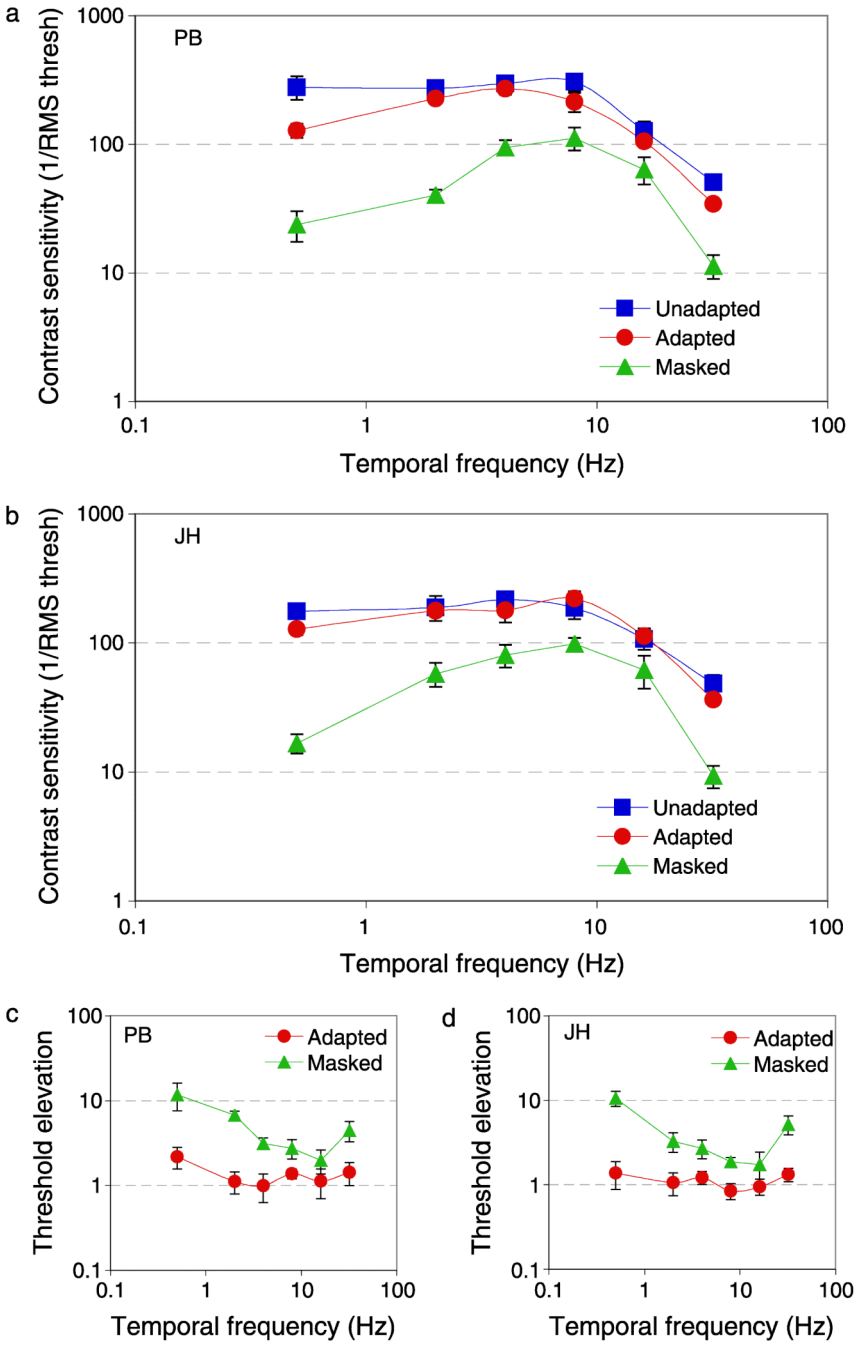

Figure 3.
(a, b) Temporal frequency contrast sensitivity and (c, d) threshold elevations for two observers (PB and JH). As Figure 2 except for temporal frequency.

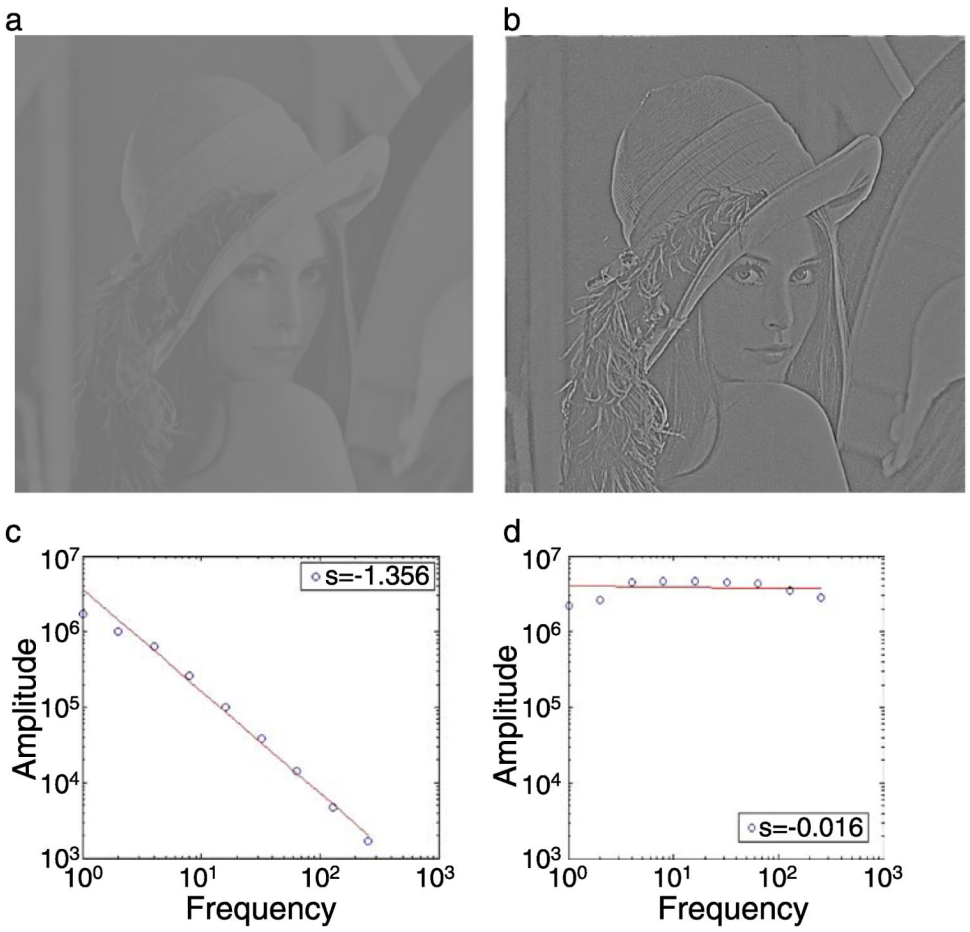

Figure 4.

Illustration of spatial whitening. (a) A natural image whose amplitude spectrum, plotted in (c), falls approximately as "1/*F*" on log–log axes with a slope of −1.4. Whitening the amplitude spectrum produces an image (b) that appears sharpened, but otherwise structurally quite similar. (d) The amplitude spectrum of the whitened image has approximately the same amplitude at all spatial frequencies and a resultant spectral slope close to 0. The rms contrasts of the source and whitened images have been fixed at 0.25.

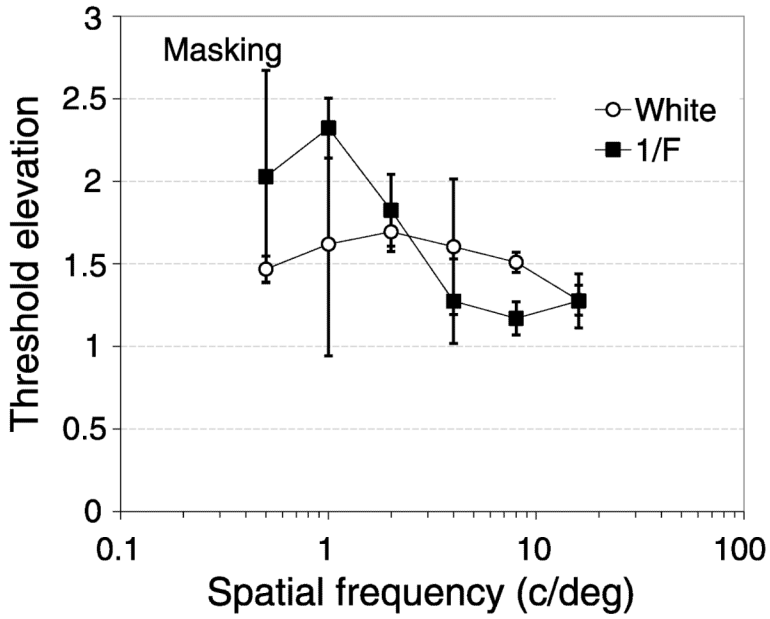

Figure 5.

Threshold elevations (masked divided by unmasked contrast threshold) for band-pass-filtered targets presented in $1/F$ (filled symbols) or spatially whitened (open symbols) movies. Data show the mean results for three subjects for standard grayscale (filled squares) or whitened (open circles) movies. Error bars show ±95% confidence intervals.

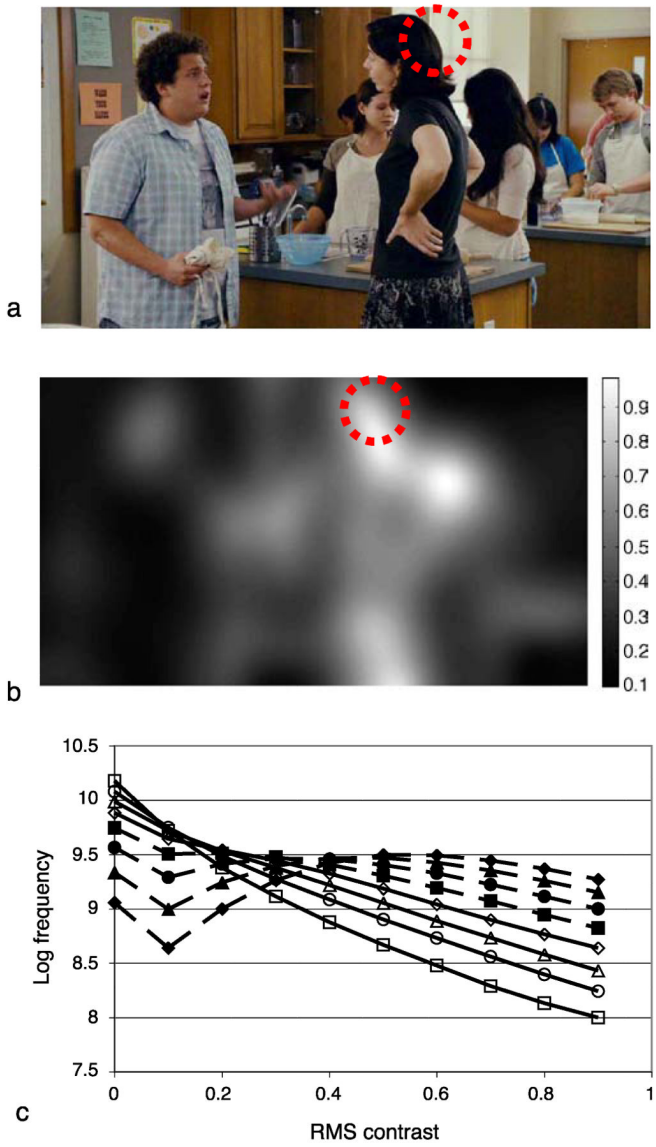

Figure 6.

Illustration of local contrast estimation. (a) A typical frame from one of our movies. A Gaussian weighted function was used to compute the local mean luminance and local standard deviation that were used to estimate local rms contrast. (b) Local rms contrast computed this way for (a). The gray level of each pixel corresponds to the local contrast at that point, see scale bar. (c) The distribution of rms contrast in 75,000 movie frames. The standard deviation of the Gaussian was 720 pixels (1 cycle/image, filled diamonds), 360 (filled triangles), 180 (filled circles), 90 (filled squares), 45 (open diamonds), 22.5 (open triangles), 11.25 (open circles), or 5.125 (open squares).

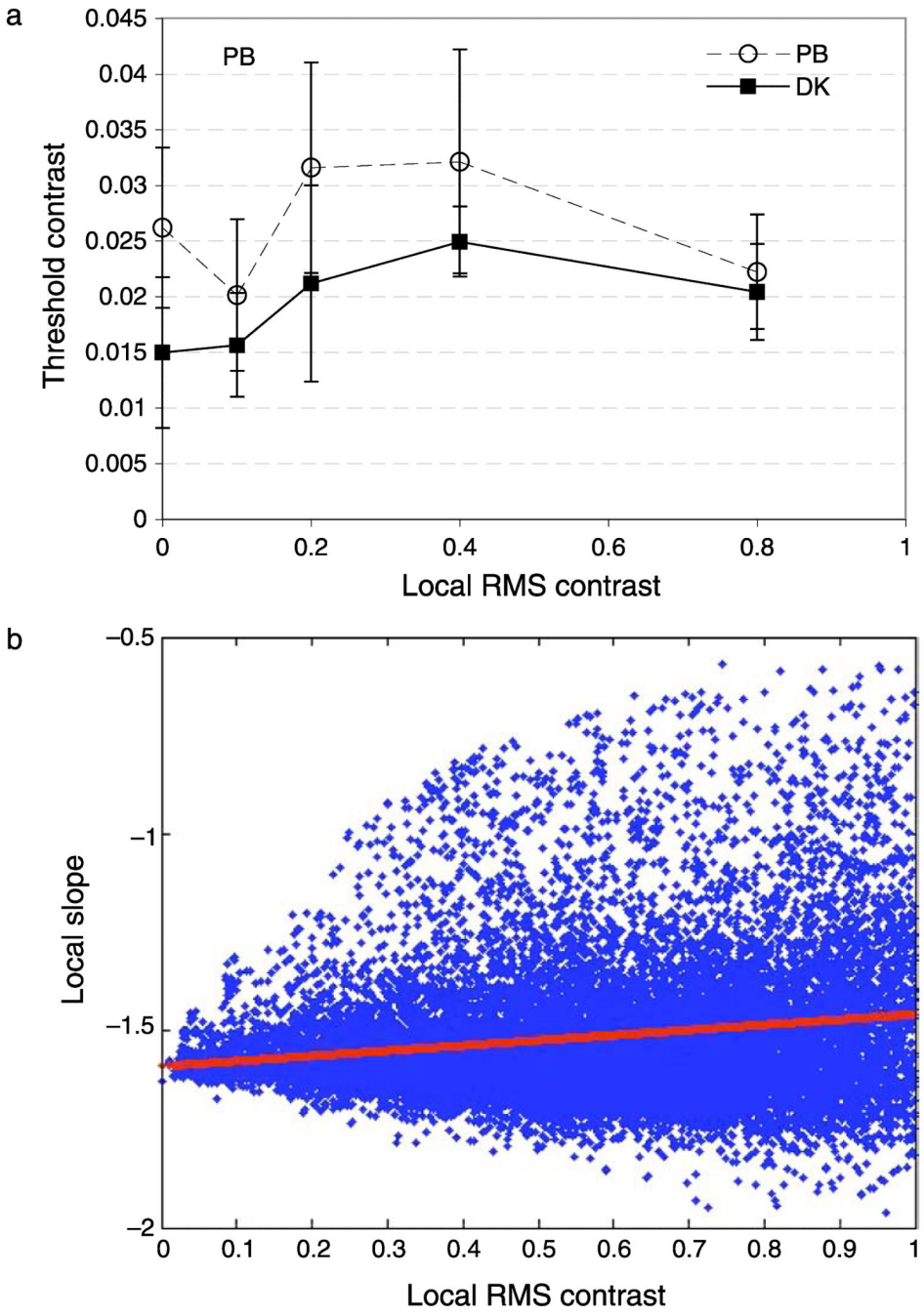

Figure 7.

(a) Contrast detection as a function of local rms contrast for two observers. The target was band-pass-filtered noise with 2 c/deg peak spatial frequency and 2.4-Hz peak temporal frequency, presented within a Gaussian spatio-temporal envelope. The *x*-axis shows the local rms contrast at the test location (see text for explanation). Error bars show 95% confidence intervals. (b) Scatter plot of the slope of the local amplitude spectrum as a function of local rms contrast (blue data). Red data show the best straight line fit.

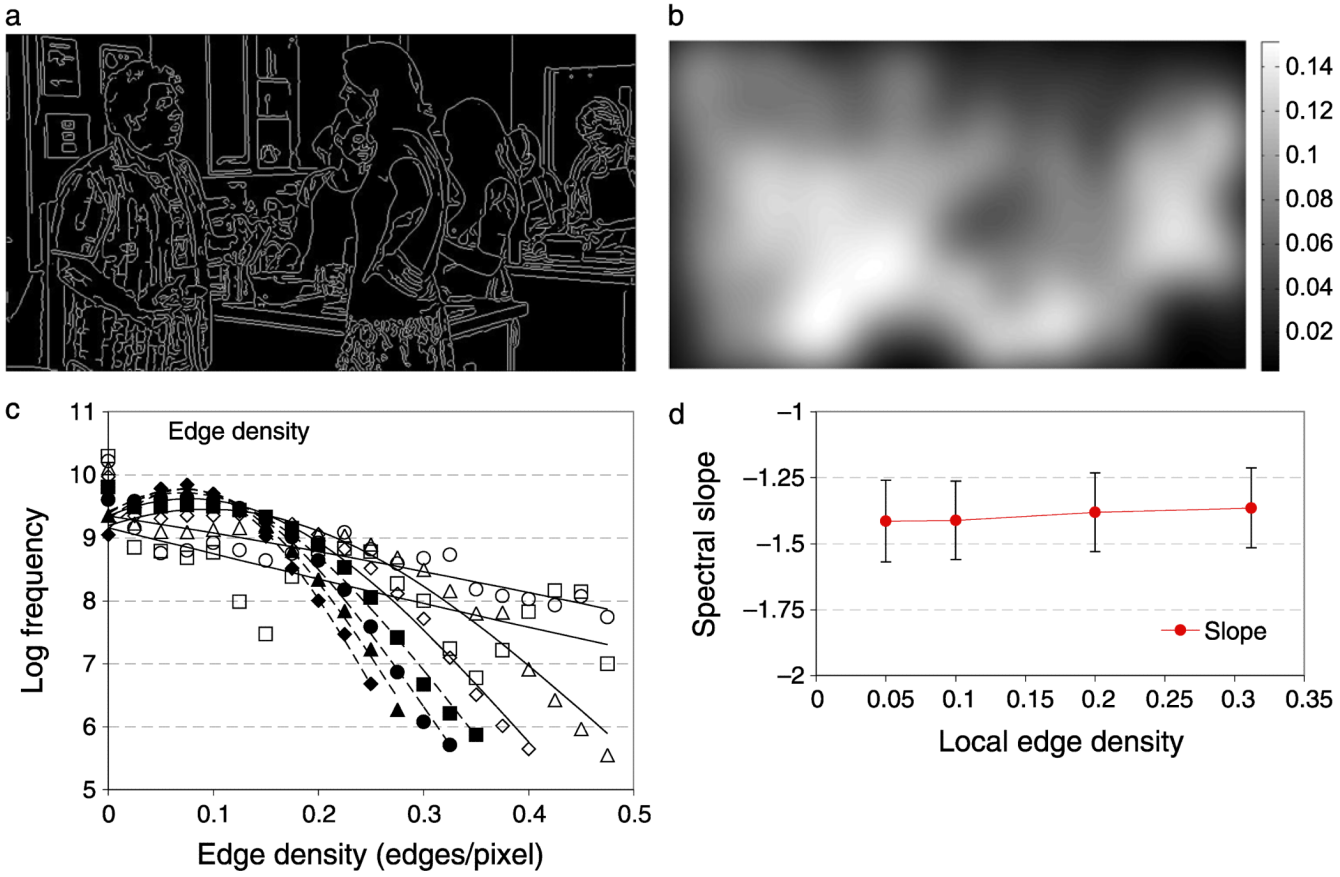

Figure 8.

Illustration of local edge density estimation. The edges of each movie frame were determined by a Canny edge-finding algorithm. (a) The edges identified in Figure 4a. (b) A Gaussian weighted function was used to compute the local edge density (proportion of edge pixels) over the target area, shown by the scale bar. (c) The distribution of edge densities across 75,000 movies. The standard deviation of the Gaussian was 720 pixels (1 cycle/image, filled diamonds), 360 (filled triangles), 180 (filled circles), 90 (filled squares), 45 (open diamonds), 22.5 (open triangles), 11.25 (open circles), or 5.125 (open squares). High edge density regions are not common in natural scenes. Most areas have fewer than 0.2 edges per pixel. The effect of edge density on the slope of the amplitude spectrum (d) is small and there is significant overlap in the slopes at all edge densities tested.

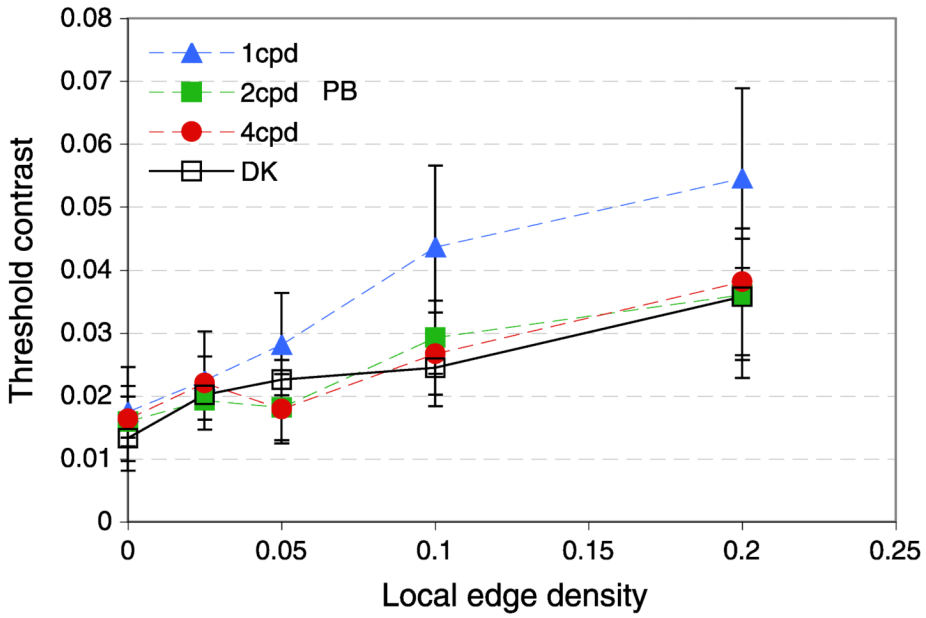

Figure 9.

Edge density in natural scenes. (a) Frequency distribution of local edge density contrast over 2.68×10^{10} pixels. (b) Contrast detection as a function of local edge density for two observers (PB, filled symbols and DK, open squares). The target was band-pass-filtered noise pattern with peak spatial frequency at 1 c/deg (triangles), 2 c/deg (squares), or 4 c/deg (circles) and 2.4-Hz peak temporal frequency, presented within a Gaussian spatio-temporal envelope. The *x*-axis shows the local edge density at the test location (see text for explanation). Error bars show 95% confidence intervals.

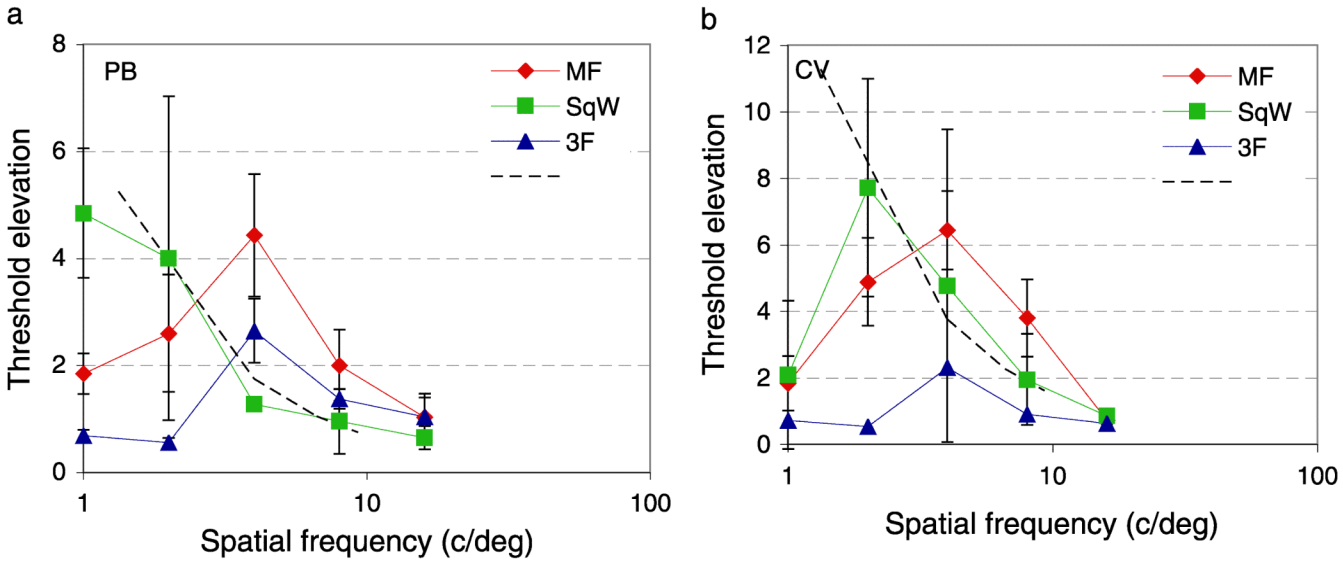

Figure 10.

Contrast threshold elevations for sine grating targets presented with masking waveforms for two observers (PB and CV). The target was a horizontal Gabor patch with fixed spatial (2 deg) and temporal (107 ms) standard deviation and spatial frequency from 1 to 16 c/deg in log steps. Masks were presented within the same Gaussian spatio-temporal envelope and was either a 1.3 c/deg square grating of 25% Michelson contrast (green squares), a 1.3 c/deg missing fundamental grating of 0.25 contrast (red circles), or a 4 c/deg Gabor of 8.3% contrast (blue triangles). Dashed line shows the expected loss in sensitivity for a square-wave grating.