Available online at www.sciencedirect.com
SCIENCE DIRECT®
Vision Research 44 (2004) 669–683
Vision Research
www.elsevier.com/locate/visres

BOLD fMRI and psychophysical measurements of contrast response to broadband images

Cheryl A. Olman [a,*], Kamil Ugurbil [b], Paul Schrater [c], Daniel Kersten [c]

[a] Department of Neuroscience, University of Minnesota, Center for Magnetic Resonance Research, 2021 6th Street SE, Minneapolis, MN 55455, USA
[b] Department of Radiology, University of Minnesota, Center for Magnetic Resonance Research, 2021 6th Street SE, Minneapolis, MN 55455, USA
[c] Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, MN 55455, USA

Received 26 June 2003; received in revised form 25 September 2003

Abstract

We have measured the relationship between image contrast, perceived contrast, and BOLD fMRI activity in human early visual areas, for natural, whitened, pink noise, and white noise images. As root-mean-square contrast increases, BOLD response to natural images is stronger and saturates more rapidly than response to the whitened images. Perceived contrast and BOLD fMRI responses are higher for pink noise than for white noise patterns, by the same ratio as between natural and whitened images. Spatial phase structure has no measurable effect on perceived contrast or BOLD fMRI response. The fMRI and perceived contrast response results can be described by models of spatial frequency response in V1, that match the contrast sensitivity function at low contrasts, and have more uniform spatial frequency response at high contrasts.

Keywords: Functional brain imaging; Psychophysics; Spatial vision; Models and theory; *Natural images*

1. Introduction

Because image contrast strongly modulates neural response in early visual areas, any study of visual response to a particular set of images requires an understanding of the underlying contrast response. Contrast response and perceived contrast are well understood for simple stimuli such as sine gratings, but these results cannot easily be extended to more natural images. As the contrast of a sine wave grating is increased, individual neurons in cat and monkey striate cortex show no increase in firing rate up to a threshold contrast, and then show a rapid increase with increasing contrast beyond the threshold at a rate that slows and saturates at high contrasts (Albrecht & Hamilton, 1982). Thresholds, rates of increase, and saturation points vary across the neural population, but the averaged supra-threshold neurophysiological measurements can be fit by a power law ($\text{response} \propto \text{contrast}^{\gamma}$) and are consistent with contrast response functions inferred from psychophysical measurements of contrast discrimination thresholds (Boynton, Demb, Glover, & Heeger, 1999; Heeger, Huk, Geisler, & Albrecht, 2000; Legge, 1981) [1]. However, physiological measurements show that nonlinearities in the visual system (see, for example, Vinje & Gallant, 2001) preclude the prediction of contrast response by simple summation of responses to isolated sinusoidal and Gabor patterns. Neural responses in early visual areas are affected by the responses of neighboring neurons, conveying a dependence on image regions and properties outside the classical receptive field.

A key issue in understanding contrast response to complicated images is the interaction of contrast response with spatial frequency. In psychophysical measurements, contrast detection thresholds vary systematically with spatial frequency, with peak sensitivity (lowest threshold) near 4 cycles per degree (Campbell & Robson, 1968). This spatial frequency

[1] Interestingly, EEG measurements of visually evoked potentials in response to sine gratings of increasing contrast find a contrast response that is linear with log contrast (DiRusso, Spinelli, & Morrone, 2001), perhaps implying a log relationship between the EEG signal and firing rates or the BOLD signal.

* Corresponding author. Tel.: +1-612-625-5372; fax: +1-612-624-2004.
E-mail address: cheryl@cmrr.umn.edu (C.A. Olman).

0042-6989/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2003.10.022

dependence of threshold is referred to as the contrast sensitivity function (CSF). However, for contrasts well above detection thresholds, perceived contrast does not vary with spatial frequency, a phenomenon referred to as contrast constancy (Cannon & Fullenkamp, 1991; Georgeson & Sullivan, 1975). Contrast constancy would imply that cortical activity in early visual areas is the same over a range of spatial frequencies at high contrast, but one fMRI study (Singh, Smith, & Greenlee, 2000) as shown a dependence of cortical activity on spatial frequency that more resembles the contrast sensitivity function than the predictions from contrast constancy. It is not known how the BOLD response depends on contrast for broadband images.

Contrast in broadband images can be described or quantified in many ways. One common metric is the root-mean-square (RMS) contrast (i.e. the square root of the average of the squared pixel intensity deviation from the mean luminance normalized by the mean), and perception of contrast in broadband images is better predicted by RMS contrast than by other contrast metrics such as Michelson contrast (Bex & Makous, 2002). However, RMS contrast conveys no information about the distribution of contrast power across spatial frequency (or space), and is not consistent with the known spatial frequency dependence of both neural and perceptual visual responses. Models in which response is distributed across multiple spatial frequency channels do a better job in predicting human observers' discrimination performance on natural images and perception of contrast (Graham & Nachmias, 1971; Peli, 1990, 1997; Tolhurst & Tadmor, 1997). In these models, images are processed by an array of spatial frequency channels, each with a bandwidth of one to two octaves. This type of multiple channel model for spatial frequency response is consistent with both electrophysiological measurements (DeValois, Albrecht, & Thorell, 1982) and psychophysical measurements (Wilson, McFarlane, & Phillips, 1983). It has also been demonstrated that a set of channels like this can maximize efficiency in coding natural images by more evenly distributing responses to natural images across the population of neurons (Field, 1987).

To study the dependence of perceived contrast and activity in early visual areas on the distribution of image contrast across spatial frequency, we have combined psychophysics and fMRI to study visual response to natural and whitened images. Whitening images produces a spatial frequency spectrum with a flattened slope, thus reducing the strength of low spatial frequency components relative to natural images, and increasing power at high spatial frequencies. The structure of whitened images is similar to that of their natural counterparts in that edges are left intact, and objects are recognizable (Fig. 1).

In this work, in addition to quantifying the effect of whitening on perceived contrast and contrast response in V1, we also test the power of several models of spatial frequency processing in early visual cortex to predict the observation that, for a given RMS contrast, apparent contrast of whitened images and white noise is lower than apparent contrast of natural images and pink noise. The simplest explanation for the observed difference is the prediction made by the CSF: lower sensitivity to high spatial frequencies, which are more strongly represented in whitened images, predicts lower response to whitened images. Both single- and multiple-channel models, matched to the CSF of V1, predict significant decreases in the response to whitened images, compared to natural images, but the ratio of these responses does not depend on image contrast. Two additional models are tested: one in which multiple spatial frequency channels have different contrast response thresholds and gains; and one based on the divisive normalization model (Carandini, Heeger, & Movshon, 1997; Heeger, Simoncelli, & Movshon, 1996; Wilson & Humanski, 1993), with multiple spatial frequency channels that are mutually inhibitory. Both of these models predict a large difference between responses to natural and whitened images at low contrasts (regardless of whether images are structured or phase-scrambled noise). The predicted response ratio decreases with increasing contrast—an effect very similar to contrast constancy and one that fits the BOLD fMRI and psychophysical data in this study.

The imaging studies were performed at high field (7 T), where the spatial specificity of the BOLD signal is increased over lower fields (Yacoub et al., 2001), and the increased magnitude of BOLD signal changes (increased contrast-to-noise ratio) allows higher resolution with a good signal-to-noise ratio. This high contrast-to-noise ratio also permits quantification of response to even low contrast images in single scans. We compared psychophysical measurements of the perception of contrast in natural and whitened images and noise patterns against both the observed (BOLD fMRI) and the predicted reduction of contrast response to images with flat amplitude spectra.

2. Methods

2.1. Images

Digitized, calibrated natural images were downloaded from the van Hateren database (van Hateren & van der Schaaf, 1998), and 42 images with predominantly unimodal pixel intensity histogram distributions were selected. Pixel intensity values were reassigned so that the intensity distribution was Gaussian. This manipulation did not significantly affect the appearance of the images, but did ensure that comparisons of RMS contrast between images would correlate well with other contrast metrics, such as Michelson contrast.

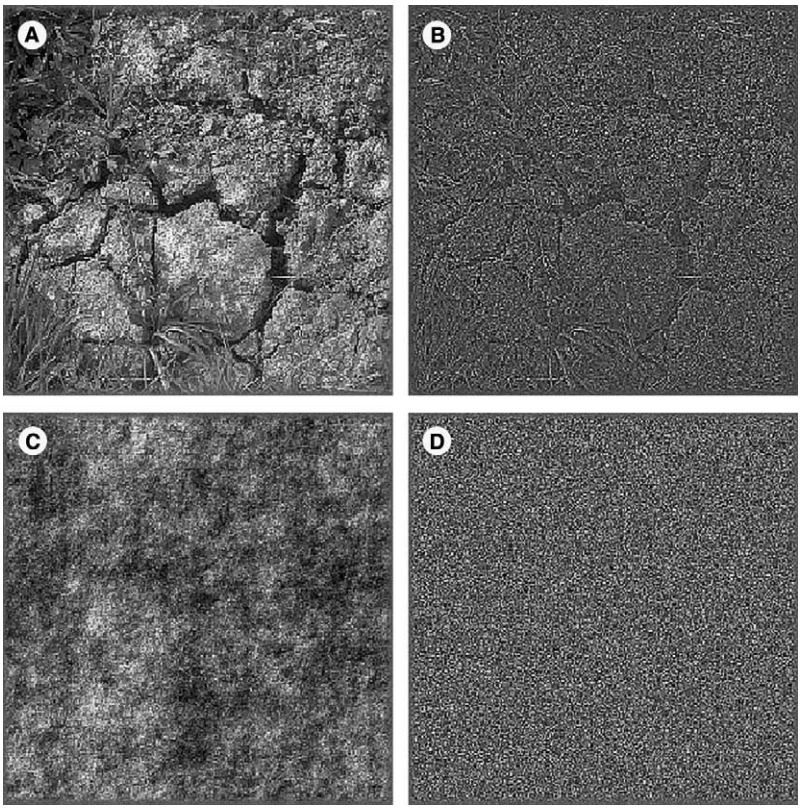

Fig. 1. Examples of stimuli used in experiments. (A) Natural images typically have a spatial frequency amplitude spectrum that drops off as $1/f$, where f is spatial frequency. (B) A whitened image is identical to a natural image in every respect except the relative amplitude of different spatial frequency components: all are represented with equal power. The presence of edges and image structure remain unaffected because the phase portion of the complex spatial frequency spectrum is not manipulated. (C) Pink noise has a $1/f$ spectrum, but none of the original image structure, due to randomization of the Fourier phase spectrum. (D) White noise has a flat amplitude spectrum and no structure.

From each natural image, a whitened image was generated by multiplying the amplitude, A_f, at each spatial frequency (in the 2D Fourier transform of each image) by $1/A_f$, then performing the inverse transform and scaling the pixel intensity values to equate the RMS contrast. Noise images were generated either by scrambling the Fourier phase spectrum of each natural image by adding random perturbations to each component of the phase spectrum (for pink noise), or by drawing pixel intensity values from a normal distribution (white noise). All images (whitened and natural) were saved as zero mean, unit variance matrices, and image contrast was adjusted at the time of display.

2.2. Image presentation

Images were 256×256 pixels, displayed to subtend 14° of visual angle, so the highest spatial frequency present was approximately 9 cycles per degree (cpd). Both whitened and natural images were vignetted by an 8-pixel $\cos^2$ function to eliminate effects from the image edges. Imperfections in the projection path for the fMRI experiments resulted in image blurring. The blurring was quantified by measuring a normal observers' contrast sensitivity function with and without the blur, finding 10% attenuation at 4 cpd, and 50% attenuation at 8 cpd. Therefore, spatial frequency spectra of the images were attenuated at high spatial frequencies before being presented for the psychophysics study, in order to match conditions as closely as possible between the fMRI and psychophysics experiments.

2.3. Psychophysics

Psychophysical measurements of contrast perception were performed on a Macintosh G3 computer, running MATLAB (The Mathworks Inc., Natick, MA, USA) with the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). Images were displayed on a CRT monitor using a video card with 8 bit input resolution, 10 bit output resolution. The PsychToolbox VisualGamma function was used to generate an 8-bit linear contrast look up table. To display images with greater contrast resolution, the range of the table was restricted, and then interpolated to generate an 8-bit look up table with finer resolution that could take advantage of the 10-bit

display range of the video card. For high contrast trials (12% and 24% RMS contrast), the full lookup table was used (finest contrast increment ~0.5%), but the table could be halved and interpolated (up to two times) to increase the contrast resolution for contrast discrimination or contrast matching trials at low pedestal contrasts. Four subjects participated, two of the authors, and two experienced volunteers naïve to the purposes of the experiment.

2.4. Preliminary contrast discrimination measurements

Two of the authors participated in measurements of contrast discrimination thresholds at 3%, 6%, 12%, and 24% pedestal contrasts, which were used to select images for the contrast matching experiment (see below) and to set the contrast increments for the attention-controlling task in the magnet. A two interval forced choice (2IFC) paradigm was used to determine contrast discrimination thresholds for the whitened and natural images. For each trial, the same image was presented in both intervals (at either the pedestal contrast or the pedestal contrast plus an increment), but different images were selected for each trial. At each of four pedestal contrasts (for both image sets), a Quest algorithm was used to find contrast discrimination thresholds (82% correct) in blocks of 50 trials.

2.5. Contrast matching experiments

A subset of eight images was selected for the contrast matching experiments. These representative images were selected from the preliminary contrast discrimination measurements as the images for which measured contrast discrimination thresholds were most consistent. Images were displayed using a temporal 2IFC paradigm, in which subjects responded whether the contrast of the whitened image needed to be increased or decreased to match the natural image. Matches were performed at four pedestal contrasts, and contrast of the whitened image was adjusted either up or down, according to the subject's response, by a fixed ratio after each trial. Images for each trial were drawn randomly; within each trial the natural and whitened image were the same image except for the spatial frequency spectrum manipulation. For each combination, six matches were made at each of four pedestal contrasts; an average of eight trials was required for each match.

An additional set of contrast matching experiments was performed to test the effect of image structure (phase coherence) on perceived contrast. For these experiments, either pink noise or white noise patterns were used as the match images; otherwise experimental conditions were identical.

2.6. fMRI scans

Five healthy subjects (all female, ages 21–30) participated in the first series of experiments (natural and whitened image contrast response, six scans); one subject who participated in the first series, plus three additional subjects (one male) participated in the second series of experiments (white noise and pink noise response measurement). All subjects provided informed consent in accordance with institution guidelines and were paid for their time.

For stimuli, the full set of 42 images was used, generated by a Macintosh G4 computer running MATLAB with PsychToolbox and back-projected by a NEC MT1050 projector (NEC Solutions (America), Itasca, IL, USA) housed outside of the magnet room, using a modified lens (Buhl Optical, Pittsburgh, PA, USA) to focus the output beam through a waveguide into the magnet room, at the appropriate size and focal plane. In the magnet, subjects viewed the image screen behind their head through a mirror. In the first two sessions of the first experiment, subjects were instructed to fixate passively on a fixation mark at the center of the screen as images, randomly selected, were presented at a rate of 10 Hz. In the latter half of the first set of experiments (four sessions), subjects were engaged in a contrast discrimination task during the experiment. Images were presented at a slower rate, in a temporal 2IFC paradigm. Each image in a trial was randomly selected and presented for 0.45 s with 0.05 s blank (mean gray) between images, and then a noise image (with appropriate amplitude spectrum, to match the image set being tested) was displayed for 1 s. Typical experiments of this type would use a blank gray screen between trials instead of a noise image, but preliminary studies had indicated that there could be a difference in the hemodynamic response to the two image sets, so a matched noise image (same spatial frequency spectrum) was displayed during the response period to maintain a consistent level of neural activity throughout the block. It would have been preferable to present an image during the inter-trial blank, but this made the task too difficult. Subjects maintained fixation on the fixation mark and were provided with a button box, instructed to press the left button if the first image was at higher contrast, and the right button if the second image was higher contrast. Performance was monitored, and contrast thresholds were set so that discrimination performance was at approximately 80% during the scan for all image sets.

The basic stimulus paradigm was a block paradigm, with 24 s of image presentation (240 images at 10 Hz for the first two subjects, 12 2 s 2IFC trials for the last four), followed by 24 s of blank. Blocks of each image type (whitened and natural) were presented at five pedestal RMS contrasts: 3.3%, 6.6%, 9.9%, 19.7%, and 33%, in random order. For the first two subjects, the two types

of images were presented in different scans, repeated three times each, for a total of six scans. For the last four experiments, all ten image conditions were combined into one scan, which was repeated three times.

The second series of fMRI experiments presented only four blocks of images: natural images at 20% RMS contrast, pink noise at 20% RMS, whitened images at 33% RMS, and white noise at 33% RMS. For these blocks, noise patterns and images were not mixed, so the contrast discrimination task was much more demanding (inexperienced subjects performed only slightly better than chance), but subjects saw only one image type during each block.

2.7. Data acquisition and pre-processing

All MR images were acquired with gradient echo EPI (TR/TE 2.4 s/20 ms) at 7 T (Magnex scanner, Varian console), using a quadrature surface coil with 14 cm loops. Slices were axial or oblique (parallel to the calcarine sulcus) and 2 mm thick. Field of view was 19.2 cm $\times$ 14.4 cm with matrix size of 128 $\times$ 96 voxels (4 segment acquisition), yielding a nominal resolution of 1.5 mm $\times$ 1.5 mm $\times$ 2 mm. One, three, or five slices were acquired in the first set of experiments. The only preprocessing applied to these data sets was automated correction for phase errors and global B_0 fluctuations due to respiration (Pfeuffer, Van de Moortele, Ugurbil, Hu, & Glover, 2002). Five timepoints were acquired before the onset of the stimulus in each scan (the first two were discarded as the signal reached steady state); 10–15 volumes were acquired after the stimulus offset (24–36 s). Scan durations were therefore 264 s for the short block scans, 504 s for the long block scans. In the second series of experiments, a single slice (either oblique or coronal, through the calcarine sulcus) was acquired, with either 1 mm or 1.5 mm in-plane resolution, and a repetition time of 2 s (170 s for each scan, containing four 20 s blocks of image, separated by 20 s).

2.8. Data analysis

All analysis was performed with code written in MATLAB specifically for this project and with Stimulate analysis software (Strupp, 1996). Visually activated voxels were identified by cross correlation between the BOLD signal and a boxcar representing the blocks of stimulus presentation. For the sessions in which the natural and whitened images were presented in separate, short movies, the first natural image block scan was used to identify active voxels; this scan was not included in further analysis. For the later experiments, in which natural and whitened images were interleaved in a long block paradigm, active voxels were identified by cross-correlation between the first five (first half) of the blocks and a box car with identical height during each epoch. This procedure was adopted because subject motion between scans precluded the independent selection of voxels from different scans. The threshold for the correlation coefficient was set at $r = 0.3$, yielding on average of 125 activated voxels in each slice. Because at most a 1 cm slab of cortex was imaged, functional localization of V1 by field sign mapping was not performed. Voxels clearly located in the calcarine sulcus and directly posterior to it on the occipital pole were selected as V1. All other visually activated voxels were grouped together as extrastriate cortex.

After each group of voxels was delineated (V1 and extrastriate), a quadratic trend was removed from the baseline portions (pre- and post-stimulus) of the data from each scan; this is the only detrending applied to the data. Each voxel intensity was then normalized by the mean baseline value; all numbers are reported as percent change from baseline. Voxels were then averaged across the visual area selected, and the percent increase in BOLD signal (averaged across five to seven timepoints during the stimulus) was used to quantify the response to each block of images.

3. Results

3.1. Psychophysics: apparent contrast

A contrast-matching task was used to test the relationship between the apparent contrast of a natural image and the apparent contrast of the same image with a flat amplitude spectrum. Results are shown in Fig. 2. Subjects consistently increased the RMS contrast of the whitened images to match the apparent contrast of the natural image (Fig. 2A). To test whether the same relationship held for unstructured images, pink noise (Gaussian noise with a frequency spectrum in which the amplitude of the Fourier component decreases as the inverse of spatial frequency) and white noise patterns (flat amplitude spectrum) were then used for the match images. For matching the apparent contrast of pink noise patterns to natural images, RMS contrast was equal when apparent contrast was matched (Fig. 2B). Subjects increased white noise RMS contrast by the same amount as whitened image contrast to match the perceived contrast of natural images (Fig. 2C), indicating that spatial frequency amplitude spectrum plays a significant role in determining apparent contrast, regardless of the presence of edges and image features.

3.2. BOLD fMRI contrast response functions

Having measured the reduction in perceived contrast for whitened images, we tested whether activity in early visual areas corresponds better to RMS image contrast or to perceived contrast. BOLD activity was measured

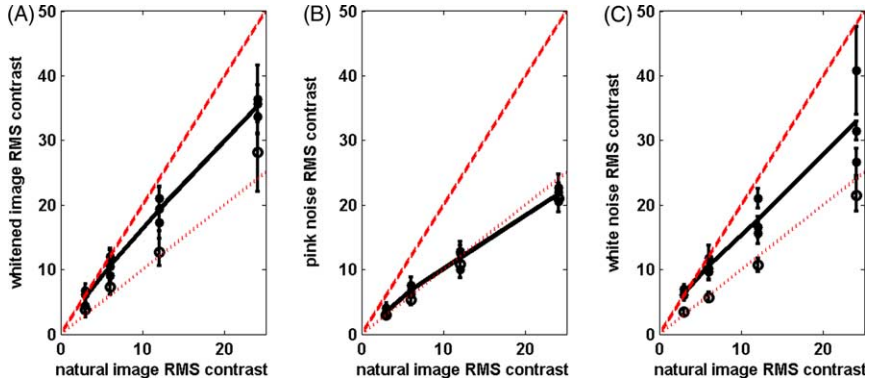

Fig. 2. Results of the contrast matching experiments, showing the effect of spatial frequency amplitude spectrum on perception of global image contrast. Images were displayed in a 2IFC paradigm, 0.5 s each, after which subjects responded whether the apparent contrast of the second image was higher or lower than the target image. Contrast was adjusted accordingly, by a fixed ratio, until subjects indicated a match in perceived contrast. Contrast of test images was adjusted to match the perceived contrast of natural images at 3%, 6%, 12%, and 24% RMS contrast. Four subjects participated: two authors, and two naïve observers. (A) When the RMS contrast of whitened images was adjusted to match the perceived contrast of natural images, subjects increased the contrast of whitened images by as much as a factor of two. Red dashed line indicates a 2:1 ratio between RMS contrast in whitened images and RMS contrast in natural images; red dotted line indicates 1:1 match. One observer's results were consistently different from the rest of the subjects, so this data (shown in open circles) was not included in the calculation of the average response. (B) When contrast of pink noise images was adjusted to match the perceived contrast of the natural images, the relationship between perceived contrast and RMS contrast was the same, in spite of the lack of similarity in the appearance of the images. (C) When white noise is adjusted to match the apparent contrast of natural images, the RMS contrast of the noise is increased by the same amount as the contrast of whitened images.

in primary visual cortex while subjects viewed blocks of either natural or whitened images at several different RMS contrasts (see Section 2). Because the results from scans acquired from subjects engaged in a task designed to control attention matched the results from earlier scans in which no task was employed, all scans were analyzed together (Table 1). Exemplary raw data from one scan, averaged across selected voxels in anatomically delineated striate (red box, Fig. 3A) and extrastriate voxel groups in one subject are shown in Fig. 3C and D.

Measured contrast response functions, calculated individually for each subject and then averaged across six subjects, are shown in Fig. 4. The power law, $R = AC^{\gamma}$, was fit to the data from each subject; the exponent γ from the power law describes the saturation of the contrast response and is shown in Table 1 for each subject. For both striate and extrastriate voxel groups, the amplitude of the contrast response function was greater and was lower (faster saturation) for natural images, which combined to result in significantly larger BOLD signal for natural images than for whitened images at any given RMS contrast. As expected, the amplitude of the response in extrastriate voxel groups was smaller, and response saturation with increasing contrast was more rapid (Avidan et al., 2001). (This result is only qualitative; standard error on fits in extrastriate regions is large, due to the small number of voxels included in the averages, and the fact that averages are across many different visual areas.) The measured reduction in the V1 response to whitened images matched the results from the contrast matching experiments (Fig. 5), indicating that the amplitude of the BOLD fMRI contrast response in V1 was correlated with perceived contrast rather than RMS image contrast.

Table 1
Fitted gamma parameters for each subject, V1 and extrastriate voxel clusters

Subject	Natural images	Whitened images
1	0.38 ± 0.60	0.41 ± 0.44
2	0.40 ± 0.10	0.56 ± 0.26
3	0.23 ± 0.08	0.66 ± 0.30
4	0.42 ± 0.29	0.60 ± 0.50
5	0.50 ± 0.28	0.67 ± 0.91
6	0.38 ± 0.15	0.41 ± 0.14
Average	0.38 ± 0.14	0.52 ± 0.18
Extrastriate		
1	0.31 ± 1.1	0.28 ± 1.2
2	0.25 ± 0.66	0.07 ± 1.0
3	0.10 ± 0.16	0.32 ± 0.18
4	0.22 ± 0.24	0.65 ± 0.75
5	0.46 ± 0.15	0.45 ± 0.72
6	−0.01 ± 0.16	0.09 ± 0.23
Average	0.18 ± 0.15	0.34 ± 0.29

Error estimates are 95% confidence intervals for the fit of the line slope in log–log coordinates.

3.3. BOLD fMRI measurement of response to noise patterns

To test whether V1 activity was more strongly affected by the presence of edges (phase coherence) in the images, or by the spatial frequency spectrum of the

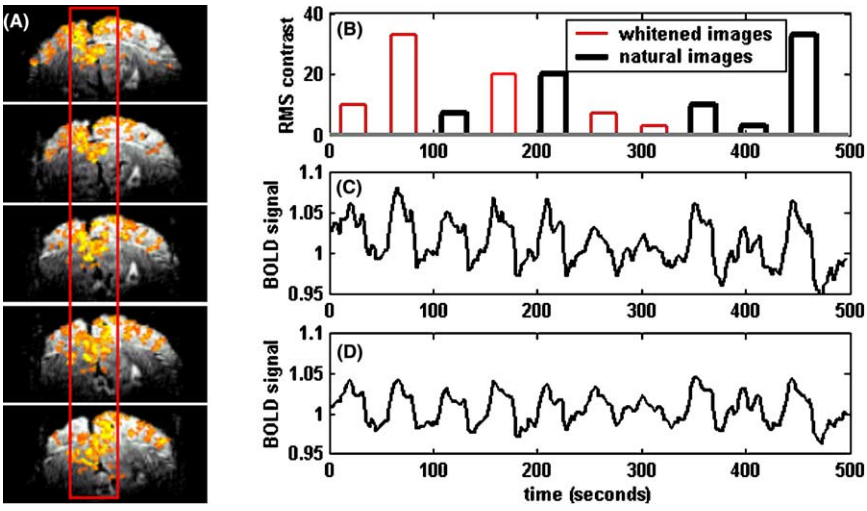

Fig. 3. (A) Representative fMRI images are shown (EPI images, axial slices, covering a 1cm slab of visual cortex centered on the calcarine sulcus), with the calculated activation map overlaid. Red box indicates voxels selected as V1. (B) For the main experiment, stimuli were presented in a block paradigm, interleaving different image types at several different pedestal contrasts. (C) Representative raw data is shown (average activity in 200 voxels in V1, one subject, one scan), plotting relative strength of BOLD signal as a function of time. No baseline detrending or normalization has been performed. (D) Representative raw data from extrastriate regions, same scan.

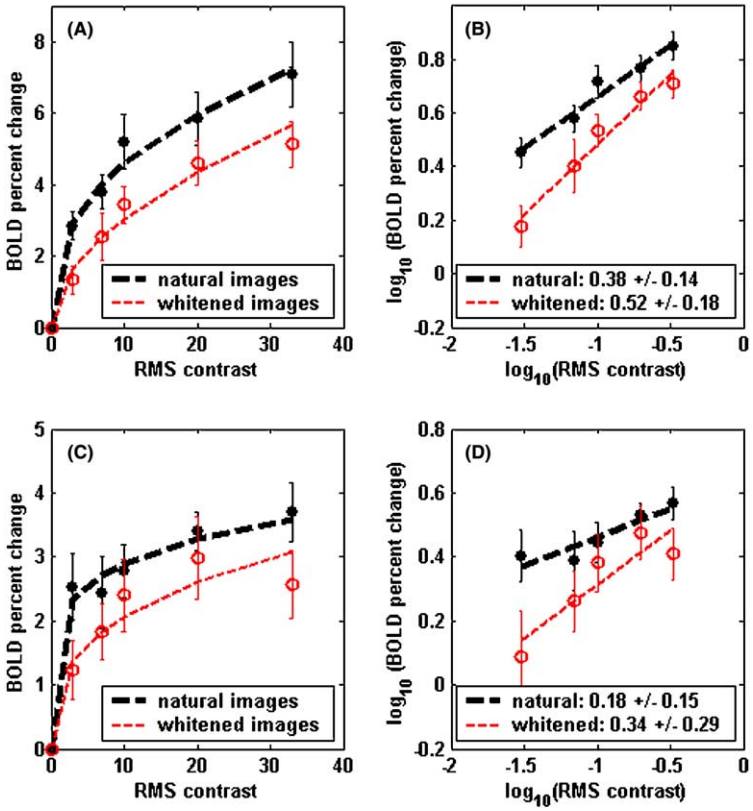

Fig. 4. Contrast response functions measured in striate and extrastriate visual areas. Individual estimates (six experiments) of the contrast response functions for natural and whitened images were combined to generate an average contrast response function. Errorbars show standard error of the mean, ($n = 6$). (A) Contrast response functions in striate cortex for natural (black circles) and whitened (red open circles) images. Dashed lines represent fit to power law: BOLD $\propto C^{\gamma}$. (B) The same contrast response functions as in (A), plotted on log–log coordinates to visualize the difference in contrast response saturation. The calculated value of γ (with 95% confidence intervals) is indicated in the legend for each image set. (C) Extrastriate contrast response, exhibiting the expected rapid saturation and lower amplitude. (D) Log–log plot of extrastriate contrast response. Estimates of contrast response exponent, γ, are again indicated in the legend.

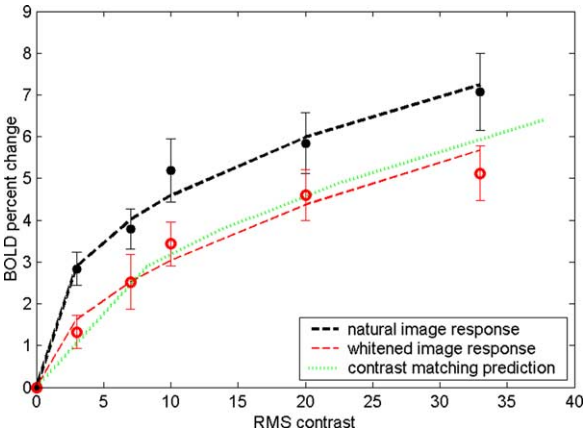

Fig. 5. Perceived contrast for the two image classes matches the relationship between BOLD fMRI measurements of contrast response functions in V1. Scaled against measured contrast response to natural images (black dashed line, showing fit to BOLD data), the lower perceived contrast of whitened images predicts lower contrast response in V1: the green dotted line is generated by scaling the response to natural images by the perceived contrast ratio measured in the contrast matching experiments. The results show good agreement with measured contrast response to whitened images (red dashed line shows fit to BOLD data).

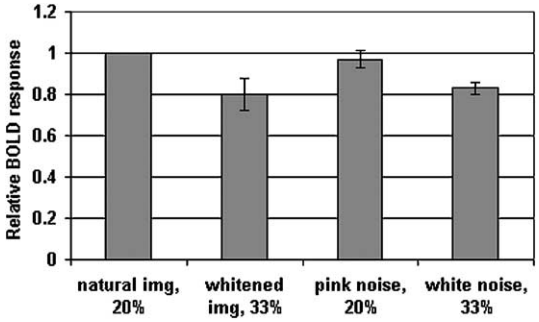

Fig. 6. Response to pink noise and white noise, compared against response to natural and whitened images. For each subject, response was normalized by the response to natural images. Errorbars indicate standard error of the mean ($n = 4$ subjects) for the ratio of whitened image, pink noise, and white noise response to natural image response. Psychophysical measurements of perceived contrast in noise images predict that pink noise should match natural image contrast response (100%), and white noise should match whitened image contrast response (and response to whitened images at 33% RMS contrast should be 80% of the response to natural images at 20% RMS).

images, we also measured the BOLD fMRI response to pink noise and white noise images. Because the psychophysical measurements of perceived contrast indicated that pink noise and natural images had the same perceived contrast for a given RMS contrast, these two images were displayed at 20% RMS contrast, and compared against whitened images and white noise displayed at 33% RMS contrast. If spatial frequency content were a sufficient indicator of contrast response, then the response to pink noise should match the response to the natural images, and the response to whitened images and white noise should be 80–90% of the response to natural images. (Using the results in Fig. 5, a whitened image at 33% RMS contrast produced a 4.6% change in the BOLD signal at 7T, while a natural image at 20% RMS contrast produced a 5.5% change in the BOLD signal.) Results, shown in Fig. 6, indicated that the prediction was successful, and the relationship between RMS contrast and perceived contrast measured for whitened and natural images held for noise patterns. The local energy model of feature detection and visual response would not have predicted this finding, a topic that will be taken up in the discussion.

3.4. Results—models of spatial frequency processing

We compared our results to predictions from three types of models: one model that simply matched the contrast sensitivity function, and two that accounted for more uniform spatial frequency response in V1 at high contrasts. Model details are discussed in Appendix A. The goal was to understand whether our results could be described by what is known of V1 spatial frequency response and neural behavior in response to simple sine wave gratings.

The first type of model predicted the contrast response simply as the variance of an image filtered by the contrast sensitivity function (Campbell & Robson, 1968), raised to an exponent ($\gamma = 0.6$) that modeled neural response saturation. This model was tested with both a single spatial frequency channel and with multiple channels (Brady & Field, 1995; Peli, 1997). For multiple channel models, response was calculated in each channel, and then pooled by an appropriate pooling mechanism (see e.g. Wilson & Bergen, 1979). The highest spatial frequency represented in the images in this study was around 9 cycles per degree, at which point contrast sensitivity is approximately 15% below peak. A single-channel model (matching the contrast sensitivity function for images presented at 2 Hz) predicted that the response to whitened images would be 70% that of the response to natural images at all RMS contrasts; a multiple-channel model using a reasonable set of spatial frequency channels (centered at 1, 2, 4, and 8 cycles per degree, with 1.5 octave bandwidth and Gaussian spatial frequency profiles) predicted a uniform 80% difference. This type of model was therefore able to predict differences in perceived contrast as large as the measured ratios, but both models failed to capture a significant aspect of the measured contrast response: the difference between the two types of images decreased at high contrasts. A summary of the various model predictions is shown in Fig. 7, plotting, as a function of increasing contrast, the ratio of predicted whitened image response over predicted natural image response.

The second type of model tested in this study used sigmoidal functions to provide for both different

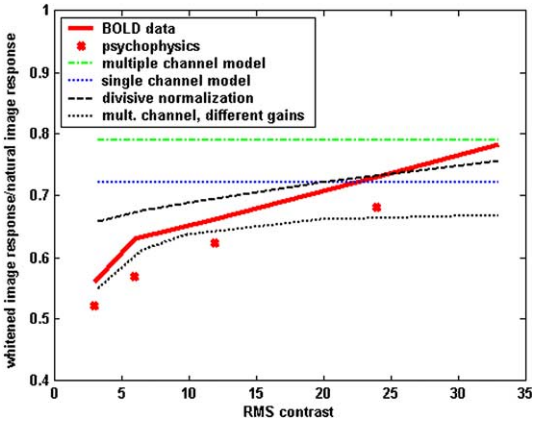

Fig. 7. Predictions from several models of spatial frequency processing. Single- and multiple-channel models with point nonlinearities predict a consistent difference in perceived contrast or contrast response, independent of contrast (green dotted and dashed line and blue dotted line). A second model uses different contrast response thresholds and gains to produce a spatial frequency response that matches the CSF at low contrasts but is more uniform at high contrasts. This produces predicted ratios of whitened image response to natural image response that match the data (black dotted line). The divisive normalization model also captures the measured behavior of the system (black dashed line), with a larger difference in perceived contrast at low contrasts, and more closely matched responses at higher contrasts. FMRI and psychophysical data are shown in red (solid line and X's).

contrast response thresholds for each channel and smooth contrast response functions. Each channel had a different contrast response gain, so that contrast response saturation occurred at a similar point in each channel, even though thresholds were different. The result was a spatial frequency response, summed over all channels, which matched the contrast sensitivity function at low image contrasts, but was more uniform at high contrasts. Because this model was based on the same spatial frequency channels modeled in Model 1, the spatial frequency response was not flat at high contrasts, so the ratio of whitened to natural images reached a maximum value of 0.7.

Another way to build a model that matched the CSF at low contrasts but had a flat spatial frequency response at high contrasts was by incorporating a divisive normalization (contrast gain) stage into the multiple channel model (Heeger et al., 1996; Wilson & Humanski, 1993). Such a model has been successfully used to describe the visual phenomena such as the White effect and grating induction, and the appearance of more complicated patterns (Blakeslee & McCourt, 1999, 2001). The measured contrast response was matched by appropriate selection of a pooling weight, or normalization strength, in a model with filters identical to those used in the simpler multiple-channel model (four filters, Gaussian spatial frequency profile, 1.5 octave bandwidth). This model did not explicitly model the different gains in each channel; response saturation arose from inhibition by other spatial frequency channels. At low contrasts, the impact of divisive normalization was low, and the model response was very similar to that of the multiple channel model in Model 1. At high contrasts, the impact of divisive normalization was to equalize the modeled response in each channel, producing a result similar to contrast constancy.

For each of these models, reasonable modeling parameters were selected, but they were not optimized to fit the data or to show perfect congruence with known physiology (e.g. the low contrast response of Model 2 was a more aggressive filter than the contrast sensitivity function, and the high contrast response was less uniform than it should be, because the model was based on the same four spatial frequency channels as the other models, for the sake of simplicity). However, the fact that Models 2 and 3 came close to matching the measured data confirmed the goal of the modeling effort: to demonstrate that any model that accounts for both the low contrast sensitivity function and high contrast constancy could explain the measured contrast response functions for natural and whitened images.

4. Discussion

The strongest findings in this study are that BOLD fMRI measurements of contrast response in early visual areas correspond to lower perceived contrast for whitened images than for natural images, and that the ratio of the two responses changes with increasing image contrast. Understanding how this neural activity depends on both spatial frequency spectrum and contrast power requires a working model for spatial frequency processing and perceived contrast. Models of spatial frequency response that succeed in describing both contrast sensitivity at low contrasts and contrast constancy at high contrasts will also successfully describe our data.

4.1. Linearity and nonlinearity in the BOLD response

The interpretation of these results relies heavily on the assumption of a linear relationship between underlying neural activity and the measured BOLD fMRI response. This linear relationship has been shown to hold for isolated sine wave gratings presented in blocks of images, as long as the block presentations are longer than 6 s in duration (Boynton, Engel, Glover, & Heeger, 1996). However, it is still possible that a nonlinearity exists between the neural response and the BOLD response, particularly in the context of natural images. Indeed, some of our early data showed the following nonlinearity: whitened images, which elicited a weaker BOLD response when presented in blocks, produced a response modulation as strong as that from natural images when presented in continuous movies (image

contrast was modulated according to a sinusoidal function with the same period as the block paradigm). As noted earlier, the amplitude of the response to block presentation of all stimulus types was not affected by adding an attention-controlling task (and reducing the rate of image presentation from 10 to 2 Hz), but the apparent difference in the hemodynamic response (an enhanced undershoot, explaining the discrepant results with continuous image presentation) was eliminated. Therefore, we conclude that nonlinearities are indeed present in the relationship between the neural response and the BOLD response, but that they can be avoided when stimuli are presented slowly in a block paradigm. The comparison between natural, whitened, and scrambled images was judged fair for this particular paradigm, but care must be taken in interpreting these results, or extending them. For example, these results should do not predict the result of presenting natural and scrambled (or whitened) images back-to-back, without intervening blank gray patch. Under those conditions, the expected difference in response to natural and scrambled images (discussed below in the context of local energy models) may be more apparent.

4.2. The effect of phase coherence on perceived contrast and contrast response

The well-established contrast energy model (Morrone & Owens, 1987) predicts decreased activity in V1 due to the lack of phase coherence in the scrambled images. Many psychophysical measurements of the effects of phase coherence in image viewing match this model and are in keeping with a picture of V1 in which edges have a profound effect on neural activity (Morrone & Burr, 1988), even though sensitivity to phase coherence is not evident in this particular study. One possible reason for this is the low sensitivity of the experiment to response modulation driven by changes in phase coherence, which are small relative to the strong effect of changes in spatial frequency amplitude spectrum. Previous work in anesthetized monkeys has found that activity in V1 is lower for phase scrambled or scrambled images than for natural images (Rainer, Augath, Trinath, & Logothetis, 2001, 2002). In the first of these studies, phase coherence was varied from 0% (pink noise) to 100% (natural image), leaving spatial frequency amplitude spectra unaffected. The measured points on the continuum may be subject to an artifact that affects the statistical properties of the images (Dakin, Hess, Ledgeway, & Achtman, 2002), but the 0% and 100% points should be free of artifact, and the response to scrambled images (noise) is clearly lower than the response to natural images.

The divisive normalization model used to describe the results in this paper is based on the local energy model, in that each channel is represented by a quadrature pair of filters (sine-phase and cosine-phase). Simple cell responses can be approximated by the responses of individual filters; complex cells can be approximated by the sum of the squares of the individual filter responses within a channel. Since the BOLD response will be related to summed activity over simple and complex cells in a given region of cortex, the model used here is built to predict the neural population response as a weighted sum of simple and complex cells in a given region of cortex. At low contrasts, the model generates the prediction of the local contrast model: lower response to scrambled images. However, just as the modeled normalization step equalized responses across spatial frequency channels, the normalization also equalizes the responses to natural and scrambled images as contrast is increased. The BOLD fMRI measurements at high contrast, showing the same activity in V1 for natural and phase-scrambled images, are consistent with this modeling result. However, the psychophysical measurements (pink noise control in the contrast matching task) covered a range of contrasts, and there both models would predict a difference. The sensitivity of this particular psychophysical technique for quantifying contrast response is poor, so an existing difference could be missed, but further work is required to understand this discrepancy, as well as the difference between fMRI measurements in the human and anesthetized monkey.

4.3. Connection with natural image contrast perception literature

The broadband models discussed in this paper do a good job of predicting perceived contrast for these images, but previous studies have suggested that contrast perception can be modeled by a single spatial frequency band, centered near 2 cpd (Bex & Makous, 2002; Tolhurst & Tadmor, 1997). When whitened and natural images are matched for RMS contrast, contrast power is approximately equal in a frequency band centered well above 4 cpd. But when the RMS contrast of a whitened image is increased so perceived contrast matches the perceived contrast of the comparable natural image, the point of equivalence is moved below 4 cpd. Therefore, the reduction of perceived contrast for whitened images could be explained simply by the lower power in a spatial frequency band located around 2 cpd. This is in good agreement with Bex and Makous' (2002) finding that the visual system is most sensitive to perturbations of the spatial frequency spectrum in the 0.5–2 cpd range, a finding similar to that of Tolhurst and Tadmor (1997), although the latter was in the context of discrimination of changes in the slope of the spatial frequency amplitude spectrum. The present study therefore does not definitively distinguish between a perceived contrast judgment based on average activity across all spatial frequency bands in V1, and a single-channel discriminator centered below 4 cpd. But even a single channel

discriminator would need to be based on a model like the divisive normalization model to fit the measured dependence of perceived contrast ratios on image contrast (see Parraga & Tolhurst, 2000 for a similar discussion).

Apparent contrast of natural images has also been successfully modeled using the contrast sensitivity function as a nonlinear threshold function (the inverse of the sensitivity function). This function masks all spatial frequency components that do not reach threshold and passes (without attenuation) spatial frequency components that exceed the threshold (Peli, 2001). This type of model can fit our data, predicting a stronger reduction in response to whitened images at low contrast, since the thresholding function eliminates the high spatial frequency components that are more strongly represented in whitened images. However, the abrupt transition from subthreshold to detectable is not neurophysiologically realistic, and the model used in Model 2 is considered as a physiologically plausible implementation of this model.

4.4. An appropriate contrast metric for structured images

A consistent difficulty in predicting visual response to natural images is the spatial heterogeneity of the images. Local contrast varies significantly across the images (Brady & Field, 2000) and needs to be taken into account when predicting perceived contrast (Peli, 1990). Local features can also have a strong effect on perceived contrast and brightness (Chubb, Sperling, & Solomon, 1989; Morrone & Burr, 1988). In this work, we have chosen to study contrast response averaged across a group of natural images, so that the particular features of any one image are unimportant. In fact, we have seen that both our contrast matching and fMRI results are insensitive to the spatial phase structure of natural images. However, the divisive normalization model used to predict this contrast response does permit study of local variations in contrast in different spatial frequency bands. A potentially fruitful test of the model developed here would be an investigation of localized responses to particular image features and the concomitant effect on contrast perception.

Acknowledgements

The authors would like to thank Essa Yacoub for his assistance in fMRI data collection, and the two reviewers for their suggestions for improving the manuscript. Work on this project was funded by NIH RO1 EY02587, NSF/IGERT DGE 9870633, BTRR P41-RR008079, the Keck Foundation, and the MIND Institute.

Appendix A. Model calculations

In each of these model descriptions, calculations are performed in the spatial frequency domain, where the complex-valued two-dimensional Fourier transform of the image, $I(\vec{k})$ is filtered by a spatial frequency response, $F(\vec{k})$. The image created by the inverse Fourier transform of the filtered image transform can be interpreted as a "neural image", or a map of how activity in a channel with a particular spatial frequency response would be distributed across visual cortex. One metric for the neural population response is the standard deviation of activity across this neural image. Because of the equivalence between variance in the image domain and the integrated power spectrum in the frequency domain, modeled responses are calculated as the integrated amplitude spectrum (square root of the power spectrum, thus equivalent to the RMS of the image). For comparing responses between image sets, response to each image at each contrast was normalized by the response to the comparable natural image at the highest contrast. To generate quantitative predictions with these models, images were modeled as subtending 8°, containing spatial frequencies up to 16 cycles per degree.

A.1. Model 1a: Single-channel model for spatial frequency response

A neural image can be generated simply by attenuating high (and very low) spatial frequencies according to the measured contrast sensitivity function, $\mathrm{CSF}(\vec{k})$, which describes the sensitivity of the visual system at each two-dimensional spatial frequency vector, $\vec{k}$. The response, R_0, to an image with unit RMS contrast is

$$R_0 = \int \mathrm{d}\vec{k} I(\vec{k}) CSF(\vec{k})$$

The power law describing the saturation of the contrast response can be modeled by raising the modeled response to an exponent γ, which is generally close to 0.6 for sine wave gratings. A scalar multiplier, c (taking values between 0 and 1) specifies the contrast of an image; therefore the contrast response function can be described by,

$$R(c) = \left(\int \mathrm{d}\vec{k}\, cI(\vec{k}) CSF(\vec{k}) \right)^{\gamma} = (cR_0)^{\gamma}$$

and the ratio between the response to natural and whitened images, as a function of contrast, will be described by

$$\frac{R_{\mathrm{W}}(c)}{R_{\mathrm{N}}(c)} = \frac{c^{\gamma} R_{0,W}^{\gamma}}{c^{\gamma} R_{0,N}^{\gamma}} = \left(\frac{R_{0,W}}{R_{0,N}} \right)^{\gamma}$$

This ratio is constant and independent of contrast.

A.2. Model 1b: Multiple-channel model for spatial frequency response

Instead of a single spatial frequency response function, a combination of spatial frequency channels can be used. For this particular study, we selected channels with a Gaussian spatial frequency profile, each with a bandwidth of approximately one and a half octaves:

$$F_i(k) = e^{-(k-\mu_i)^2/(2\sigma_i^2)}$$

The functions are defined to be radially symmetric (i.e. no orientation tuning), so $k = |\vec{k}|$; μ_i is the center frequency (peak sensitivity) of a channel; σ_i describes the bandwidth of a channel. (To maintain constant bandwidth in octaves, σ_i/μ_i is a constant.) The model was tested with both Gaussian and lognormal filters, with little difference in results.

Individual filter responses are first calculated, and then the total response is calculated as the sum of the activity in the channels:

$$R(c) = \sum_i R_i = \sum_i \left(\int d\vec{k}\, cI(\vec{k}) F_i(\vec{k}) \right)^{\gamma} = c^{\gamma} \sum_i (R_{0,i})^{\gamma}$$

The point nonlinearity to approximate the contrast response saturation is applied in each channel, before the average response is calculated. In this form, it is again clear that the ratio of the response to whitened images and natural images will be constant and independent of contrast.

The single-channel model is illustrated in Fig. 8, Panels A and C. So this model could be compared against the other models, the filter used was the envelope of the multiple channels. Therefore, the shape is somewhat different from the classical CSF, but if a CSF appropriate for the extent of the images and the rate of presentation (2 Hz) is used, the result is not substantially different from the plot shown in Fig. 8. Because so much of the power of natural images is found at low spatial frequencies, model predictions using the single channel are sensitive to assumptions about the shape of the CSF at very low spatial frequencies. The abrupt stimulus presentation used in this study should further increase sensitivity to low spatial frequencies, increase the modeled difference in responses to natural over whitened images. Ideally, the CSF should be measured under the same experimental conditions (image size, luminance, and temporal characteristics of image presentation) to

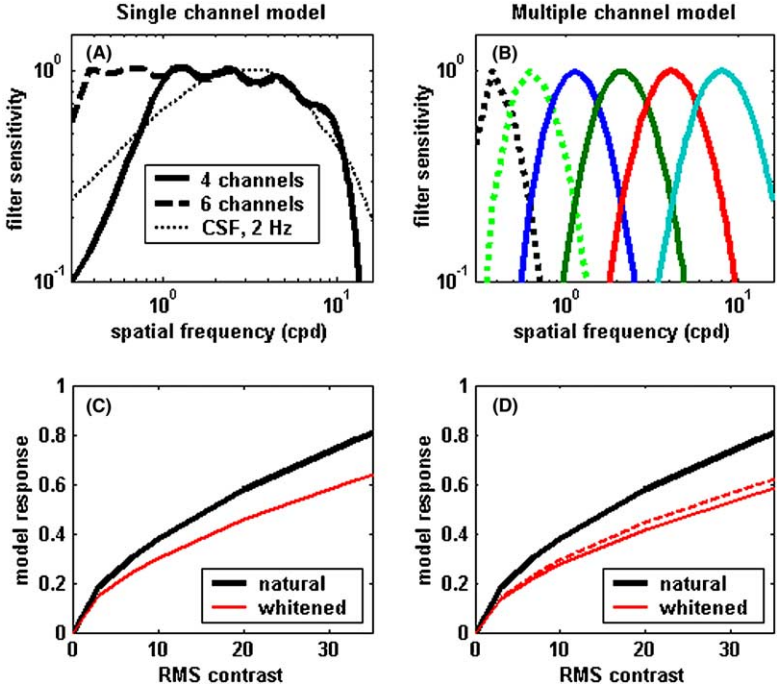

Fig. 8. Single- and multiple-channel linear filter models. (A) Spatial frequency response in the single channel, which for the sake of comparison is defined as the envelope of the spatial frequency channels in the multiple-channel model. To match the contrast sensitivity function for images presented at 2 Hz, only four of the channels were used in the calculation; the two lowest spatial frequency channels were dropped. (B) Spatial frequency profiles of the filters in the multiple-channel model: $\mu = 0.25$, 0.5, 1, 2, 4, 8 cycles per degree, Gaussian profile, with a constant bandwidth in octaves (1.5 octaves, measured as full width at half-maximum). Channels indicated by dashed lines were not used. (C) Response to natural (bold black line) and whitened images (red line) calculated by the single channel model. The contrast response saturation is added as a point nonlinearity on the output of the model (standard deviation of the filtered image raised to an exponent). (D) Response to natural and whitened images calculated by the multiple-channel version of Model 1. Contrast response saturation is added as a point nonlinearity in each channel before the total response is calculated as the average across the channels. The dashed red line indicates the difference if the model had been built with lognormal filter profiles instead of Gaussian.

maximize model accuracy. Yet the relatively low sensitivity of model predictions to the two different CSF's tested (from the literature and as the envelope of the multiple channel model) indicates that resulting changes should be relatively small. The most important feature of this model is that the ratio of whitened image response to natural image response is independent of image contrast.

For the multiple channel model (Fig. 8, Panels B and D), four filters were selected with μ equal to 1, 2, 4, and 8 cycles per degree, and σ_i/μ_i equal to 0.3, which produces a constant bandwidth of ~1.5 octaves. (Two additional channels with lower values of μ could be added to match the human CSF for higher temporal frequencies, as shown in Panels A and B.) The appropriate method for combining the channel responses is still an open question; a simple mean has been chosen for this calculation, since neither the method of combination nor the exact values of μ_i and σ_i changes the basic finding that the ratio between the responses to natural images and whitened images is independent of contrast.

A.3. Model 2: Multiple channels, with different contrast response gain and threshold

To model the dependence of contrast response threshold on spatial frequency, the Naka–Ruston formula,

$$R = \frac{c^{p+q}}{c^q + \sigma_{NR}} \tag{1}$$

can be used to generate sigmoidal functions that describe observed neural contrast response functions (Boynton et al., 1999): no measurable response up to a threshold contrast, then rapidly increasing response with increasing contrast, and finally saturation at high contrasts. By changing the p, q and σ_{NR} parameters for each channel, the threshold and gain can be balanced to produce channels with different thresholds but similar saturation points. For the purposes of this study, a set of filters was built to have an envelope function roughly matched to the contrast sensitivity function at low image contrasts, but with a more uniform spatial frequency response at high contrasts (Fig. 9).

To calculate the spatial frequency response at each contrast, the Fourier transform of the image is multiplied by each of the spatial frequency filters:

$$F_i(k) = \frac{c^{p_i+q_i}}{c^{q_i} + \sigma_{NR,i}} \mathrm{e}^{-(k-\mu_i)^2/(2\sigma_i^2)}$$

The response of the model is then:

$$R(c) = \int \mathrm{d}\vec{k}\, cI(\vec{k}) \left(\sum_i F_i(\vec{k}) \right)$$

The envelope function is dependent on contrast, as shown in Fig. 9, becoming more uniform at higher contrast. An analytical expression for the dependence of the whitened image response to natural image response is not simply expressed, but the results (shown in Fig. 7) demonstrate the contrast dependence of the whitened:natural response ratio. As in the measured data, the ratio increases as contrast increases.

A.4. Model 3: Multiple channel model with divisive normalization mechanism

Rather than selecting different contrast response gains and thresholds for each spatial frequency channel,

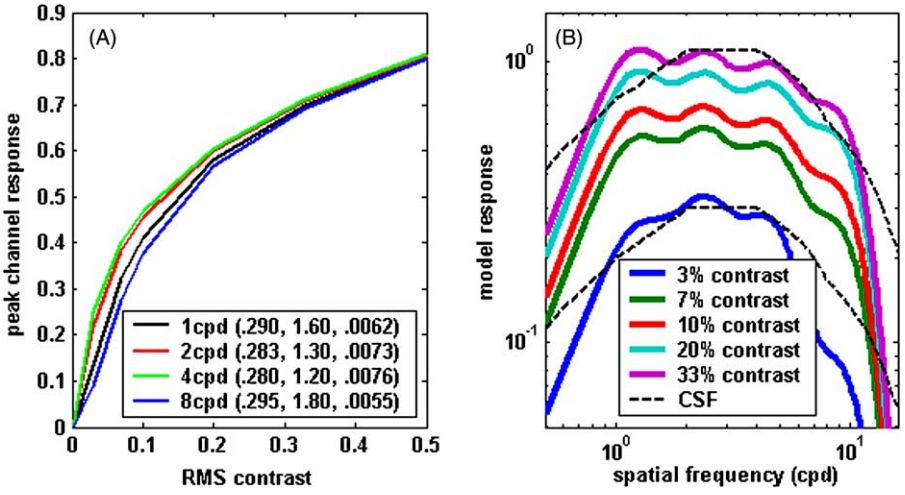

Fig. 9. Model of spatial frequency response with different contrast response thresholds and gains in each channel. (A) Using the Naka–Rushton formula (given in text), contrast sensitivity functions were generated for each channel to allow for higher thresholds in the 1 and 8 cpd channels, and lowest thresholds in the 2 and 4 cpd channels. Parameters were selected so all functions would saturate near 50% RMS contrast; the values for each channel are indicated in the legend in the order (p, q, sigma). (B) Spatial frequency response (the envelope of the four channels) matches the contrast sensitivity function at low contrasts, but is more uniform at higher contrasts.

the divisive normalization model uses inhibition between different neural populations to provide the necessary contrast gain control. The divisive normalization model begins with filters identical to the multiple-channel model. Response is calculated in each channel, and then a total raw response is summed over these responses:

$$R_{\text{raw}}(c) = \sum_i \left(\int \mathrm{d}\vec{k}\, cI(\vec{k})F_i(\vec{k}) \right) = cK$$

with

$$K = \sum_i \left(\int \mathrm{d}\vec{k} I(\vec{k})F_i(\vec{k}) \right)$$

A normalization factor is defined by

$$n_{\text{c}} = 1 + PR_{\text{raw}}(c)$$

where P is a pooling weight that controls the strength of the modeled inhibitory feedback. This normalization factor, n_{c}, is used to inhibit each channel simply by dividing the activity in each channel. At a given contrast, the final model output is then the summed activity across the normalized channels:

$$R_{\text{norm}} = \frac{R_{\text{raw}}(c)}{1 + PR_{\text{raw}}(c)} = \frac{cK}{1 + cKP}$$

No point nonlinearity is built into this model, because the feedback inhibition accomplishes the contrast response saturation. The pooling weight, P, controls the rate of saturation; in this instance, P was chosen to generate a response saturation for the natural images that matched the measured data, and then this pooling weight was used to calculate the relative contrast responses for natural and whitened image sets (Fig. 10).

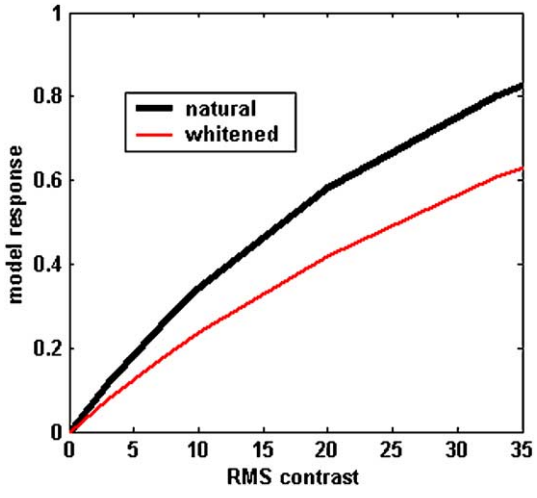

Fig. 10. Output of the divisive normalization model. In this case, the response saturation is not the result of a point nonlinearity, but the result of the increase in divisive inhibition strength as contrast increases. The pooling weight for generating this normalization factor has been selected to generate modeled saturation rates that match measured fMRI data for natural images; the same pooling weight was then used to predict contrast response for whitened images.

References

Albrecht, D. G., & Hamilton, D. R. (1982). Striate cortex of monkey and cat: contrast response function. *Journal of Neurophysiology, 48*(1), 217–237.

Avidan, G., Harel, M., Hendler, T., Ben-Bashat, D., Zohary, E., & Malach, R. (2001). Contrast sensitivity in human visual areas and its relationship to object recognition. *Journal of Neurophysiology, 87*, 3112–3116.

Bex, P. J., & Makous, W. (2002). Spatial frequency, phase, and the contrast of natural images. *Journal of Optical Society of America A, 19*(6), 1096–1106.

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

Blakeslee, B., & McCourt, M. E. (2001). A multiscale spatial filtering account of the Wertheimer–Benary effect and the corrugated Mondrian. *Vision Research, 41*, 2487–2503.

Boynton, G. M., Demb, J. B., Glover, G. H., & Heeger, D. J. (1999). Neuronal basis of contrast discrimination. *Vision Research, 39*, 257–269.

Boynton, G. M., Engel, S. A., Glover, G. H., & Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *Journal of Neuroscience, 16*(13), 4207–4221.

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*(6), 739–756.

Brady, N., & Field, D. J. (2000). Local contrast in natural images: Normalisation and coding efficiency. *Perception, 29*, 1041–1055.

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

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

Cannon, M. W., & Fullenkamp, S. C. (1991). A transducer model for contrast perception. *Vision Research, 31*(6), 983–998.

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.

Chubb, C., Sperling, G., & Solomon, J. A. (1989). Texture interactions determine perceived contrast. *Proceedings of the National Academy of Sciences, USA, 86*, 9631–9635.

Dakin, S. C., Hess, R. F., Ledgeway, T., & Achtman, R. L. (2002). What causes non-monotonic tunin of fMRI response to noisy images? *Current Biology, 12*(14), R476–R477.

DeValois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research, 22*(5), 545–559.

DiRusso, F., Spinelli, D., & Morrone, M. C. (2001). Automatic gain control contrast mechanisms are modulated by attention in humans: Evidence from visual evoked potentials. *Vision Research, 41*, 2435–2447.

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of Optical Society of America A, 4*(12), 2379–2394.

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

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

Heeger, D. J., Huk, A. C., Geisler, W. S., & Albrecht, D. G. (2000). Spikes versus BOLD: What does neuroimaging tell us about neuronal activity? *Nature Neuroscience, 3*(7), 631–633.

Heeger, D. J., Simoncelli, E. P., & Movshon, J. A. (1996). Computational models of cortical visual processing. *Proceedings of the National Academy of Sciences, USA, 93*, 623–627.

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

Morrone, M. C., & Burr, D. C. (1988). Feature detection in human vision: A phase-dependent energy model. *Proceedings of the Royal Society of London, B, 1280*, 221–245.

Morrone, M. C., & Owens, R. A. (1987). Feature detection from local energy. *Pattern Recognition Letters, 6*, 303–313.

Parraga, C. A., & Tolhurst, D. J. (2000). The effect of contrast randomisation on the discrimination of changes in the slopes of the amplitude spectra. *Perception, 29*, 1101–1116.

Peli, E. (1990). Contrast in complex images. *Journal of Optical Society of America A, 7*(10), 2032–2040.

Peli, E. (1997). In search of a contrast metric: Matching the perceived contrast of Gabor patches at different phases and bandwidths. *Vision Research, 37*(23), 3217–3224.

Peli, E. (2001). Contrast sensitivity function and image discrimination. *Journal of Optical Society of America A, 18*(2), 283–293.

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

Pfeuffer, J., Van de Moortele, P. F., Ugurbil, K., Hu, X., & Glover, G. H. (2002). Correction of physiologically induced global off-resonance effects in dynamic echo-planar and spiral functional imaging. *Magnetic Resonance in Medicine, 47*(2), 344–353.

Rainer, G., Augath, M., Trinath, T., & Logothetis, N. K. (2001). Nonmonotonic noise tuning of BOLD fMRI signal to natural images in the visual cortex of anesthetized monkeys. *Current Biology, 11*(11), 846–854.

Rainer, G., Augath, M., Trinath, T., & Logothetis, N. K. (2002). The effect of image scrambling on visual cortical BOLD activity in the anesthetized monkey. *NeuroImage, 16*, 607–616.

Singh, K. D., Smith, A. T., & Greenlee, M. W. (2000). Spatiotemporal frequency and direction sensitivities of human visual areas measured using fMRI. *NeuroImage, 12*, 550–564.

Strupp, J. P. (1996). Stimulate: A GUI based fMRI analysis software package. *NeuroImage, 3*(3), S607.

Tolhurst, D. J., & Tadmor, Y. (1997). Band-limited contrast in natural images explains the detectability of changes in the amplitude spectra. *Vision Research, 37*(23), 3203–3215.

van Hateren, J. H., & van der Schaaf, A. (1998). Independent component filers of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society of London B, 265*, 359–366.

Vinje, W. E., & Gallant, J. L. (2001). Natural stimulation of the nonclassical receptive field increases information transmission efficiency in V1. *Journal of Neuroscience, 22*(7), 2904–2915.

Wilson, H. R., & Bergen, J. R. (1979). A four mechanism model for threshld spatial vision. *Vision Research, 19*, 19–32.

Wilson, H. R., & Humanski, R. (1993). Spatial frequency adaptation and contrast gain control. *Vision Research, 33*(8), 1133–1149.

Wilson, H. R., McFarlane, D. K., & Phillips, G. C. (1983). Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision Research, 23*(9), 873–882.

Yacoub, E., Shmuel, A., Pfeuffer, J., VandeMoortele, P.-F., Adriany, G., Andersen, P., Vaughan, J. T., Merkle, H., Ugurbil, K., & Hu, X. (2001). Imaging brain function in humans at 7 Tesla. *Magnetic Resonance in Medicine, 45*, 588–594.