

A Horizontal Bias in Human Visual Processing of Orientation and its Correspondence to the Structural Components of Natural Scenes

Bruce C. Hansen

Department of Psychological and Brain Sciences,
University of Louisville, Louisville, KY, USA



Edward A. Essock

Department of Psychological and Brain Sciences,
Department of Ophthalmology and Visual Science,
University of Louisville, Louisville, KY, USA



Many encoding mechanisms and processing strategies in the visual system appear to have evolved to better process the prevalent content in the visual world. Here we examine the relationship between the prevalence of natural scene content at different orientations and visual ability for detecting oriented natural scene content. Whereas testing with isolated gratings shows best performance at horizontal and vertical (the “oblique effect”), we report that when tested with natural scene content, performance is best at obliques and worst at horizontal (the “horizontal effect”). The present analysis of typical natural scenes shows that the prevalence of natural scene content matches the inverse of this horizontal effect pattern with most scene content at horizontal, next most at vertical, and least at obliques. We suggest that encoding of orientation may have evolved to accommodate the anisotropy in natural scene content by perceptually discounting the most prevalent oriented content in a scene, thereby increasing the relative salience of objects and other content in a scene when viewed against a typical natural background.

Keywords: horizontal effect, oblique effect, contrast gain control, normalization, natural scene perception

Introduction

There is a clear relationship between the encoding mechanisms of the visual system and the prevalence of content in the natural world. For example, researchers have suggested correspondences between the prevalence of content at particular spatial scales in natural scenes and the scale and shape of human spatial filters (Field, 1987; Brady & Field, 1995; Olshausen & Field, 2000). In addition, the characteristics of low-level filters that encode color and contrast have also been reported to match the prevalence of natural scene content on those dimensions (Webster & Mollon, 1997; Simoncelli & Olshausen, 2001; Brady and Field, 2000; Taylor, Finkel, & Buchsbaum, 2000). Similarly, the relative amount of scene content at different spatial scales (i.e., the slope of the amplitude spectrum) also relates to special perceptual properties at slopes typical of natural scenes. Specifically, it has been shown that the amplitude spectrum of typical scenes can be generally characterized as a linear decline of amplitude with increasing spatial frequency, f , with a slope of ≈ -1.0 , on a logarithmic plot (Kretzmer, 1952; Deriugin, 1956; Field, 1987; Tolhurst, Tadmor, & Chao, 1992; Hansen & Essock, 2004). Thus the amplitude spectrum is commonly said to have a slope of -1.0 . As alluded to above, it has been suggested (Barlow, 1959) that the visual system might exploit this regularity of the statistical structure of natural scenes in such a way that optimizes the transfer of information to higher levels of visual processing. Specifi-

cally, Atick and Redlich (1992) have suggested that the retina acts as a whitening filter, whereby the image signal is decorrelated, increasing efficiency by effectively transmitting only the deviations from the slope typically encountered in natural scenes (i.e., approximately -1.0).

Several psychophysical studies have set out to obtain behavioral data to support the hypothesis that the human visual system is “optimized” to process scenes with content matching that typical of natural scenes. In some experiments visual sensitivity to manipulations of the slope of the amplitude spectrum of broadband visual noise or natural images was measured (Knill, Field, & Kersten, 1990; Tolhurst & Tadmor, 1997; Párraga & Tolhurst, 2000). These studies found a greater range of perceptual ‘tolerance’ to alterations of slope, as might be needed in order to effectively process visual scenes that often have an amplitude spectrum slope within a range centered around -1.0 . Higher-level visual/cognitive processing has also been examined by requiring participants to respond to a change in the semantic content within various natural scenes as a function of slope of the amplitude spectrum and found best performance at slopes near -1.0 (Párraga, Troscianko, & Tolhurst, 2000; Tolhurst & Tadmor, 2000). Although there does exist much room for debate in this area, there does appear to be a general consensus that the visual system is optimized to process natural scene imagery having amplitude spectra that fall off with a slope at or near -1.0 .

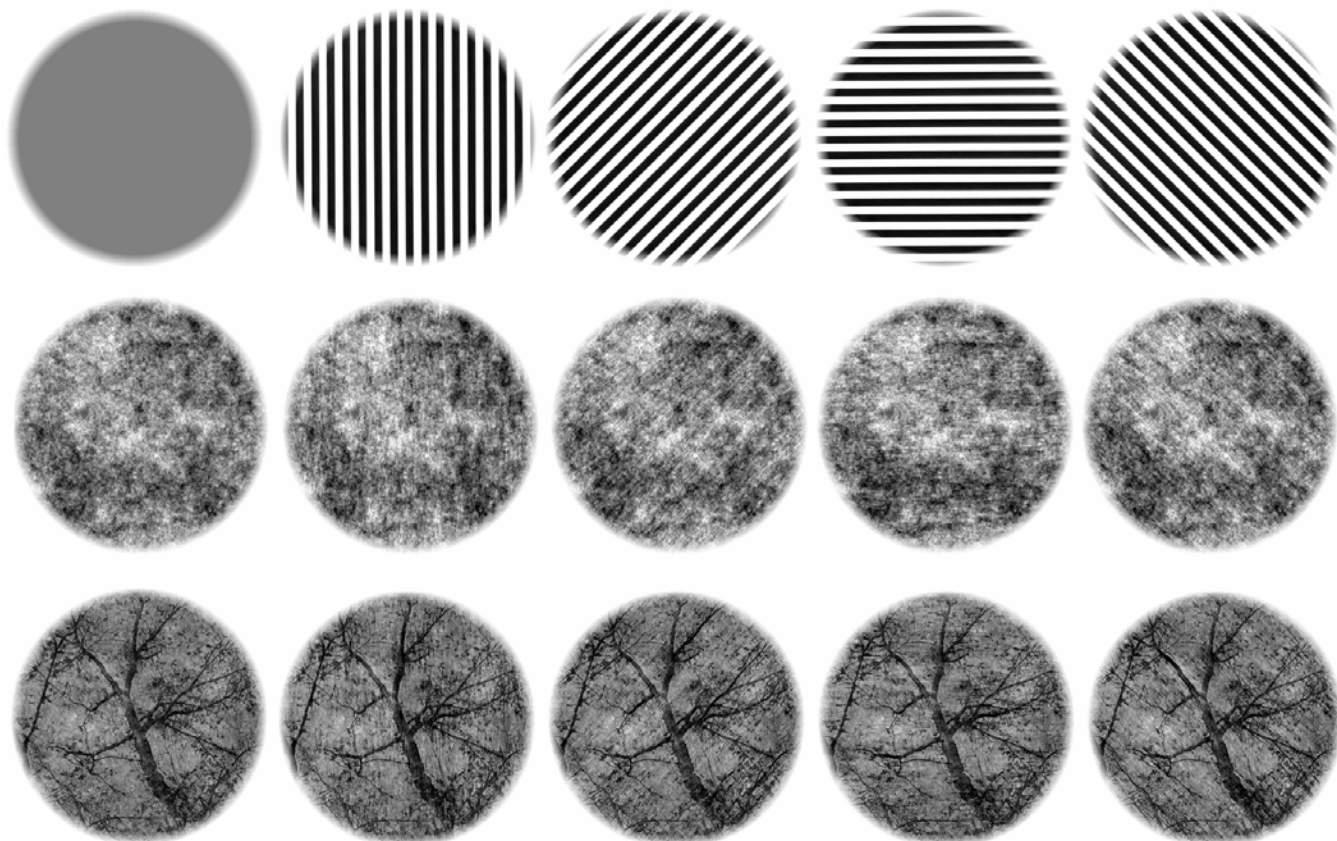


Figure 1. Illustration of test orientation (left to right), showing a non-oriented comparison pattern, followed by vertical, 45° , horizontal, and 135° test orientations. Top row: Un-patterned homogeneous background followed by four square-wave gratings. Most observers (if no astigmatism is present) can demonstrate the Class 1 oblique effect by increasing viewing distance until the oblique stimuli can not be resolved but the horizontal and vertical stimuli can be. Middle row: $1/f$ visual noise pattern without an oriented increment followed by patterns with an oriented increment of amplitude of the same noise pattern at each of the four orientations. Most observers report that the otherwise-identical oblique stimuli are of a *higher* salience than the horizontal stimulus, with the vertical intermediate – the horizontal effect. Bottom row: Natural scene that has been made isotropic and contains no oriented increment is followed by the same isotropic natural image with amplitude increments at each of the four test orientations. As with the middle row, observers report greater salience for the obliques and least for horizontal with these stimuli made of natural scene content.

In the present study we examine the relation between the content of typical natural scenes and behavioral performance with respect to the dimension of orientation. As with the ecological relationships concerning amplitude spectrum slope discussed above, orientation also has often been discussed from an ecological standpoint. Such a standpoint links three lines of research. First, it has been shown that human visual sensitivity and acuity is typically better at horizontal and vertical orientations than at oblique orientations (Campbell, Kulikowski, & Levinson 1966; Mitchell, Freeman, & Westheimer, 1967), which has been termed the “oblique effect” (Appelle, 1972) or specifically the “class 1” oblique effect (Essock, 1980; Essock, Krebs, & Prather, 1992). Secondly, a neurophysiological oblique effect has been documented in humans with visual evoked potentials (Maffei & Campbell, 1970; Zemon, Gutowski & Horton, 1983) as well as with functional magnetic resonance imaging (Furmanski & Engel, 2000) where horizontal and vertical orientations are

favored over oblique orientations in humans. Such a bias has also been demonstrated in other animals with respect to numbers of cortical cells devoted to the different orientations via single unit recordings (e.g., Mansfield, 1974; Mansfield and Ronner, 1978; Kennedy, Martin, Orban, & Whitteridge, 1985; De Valois, Yund, & Hepler, 1982; Li, Peterson, & Freeman, 2003), with more cells sampled that prefer horizontal and vertical orientations relative to the oblique orientations. Consistent with a greater number of neurons at these orientations, larger cortical regions are observed at horizontal and vertical orientations relative to oblique orientations (Chapman, Stryker, & Bonhoeffer, 1996; Chapman, & Bonhoeffer, 1998; Coppola, White, Fitzpatrick, & Purves, 1998; Yu & Shou, 2000). The third related line of research findings, the ecological component, arises when one considers the multiple reports where the content of natural scenes (and scenes with ‘carpentered’ content) is reported to be least prevalent at oblique orientations and most prevalent at horizontal and vertical

orientations (Switkes et al., 1978; Baddeley & Hancock, 1991; Hancock, Baddeley, & Smith, 1992; Coppola, Purves, McCoy, & Purves, 1998; Keil & Cristóbal, 2000). Specifically, these three types of anisotropy have led to the oft-cited dogma that it would be advantageous for animals to have best vision at those orientations that are most prevalent in the environment, and that, through either ontologic (Annis and Frost, 1973) or phylogenetic (Timney and Muir, 1976) means, this anisotropy develops to match the most prevalent content in the anisotropic world.

Where the conventional reasoning goes awry is the presumption that the stimulus orientations that are seen best when tested with isolated stimuli (e.g., a grating) would also be seen best in a natural scene with its broadband content and the potential for contrast normalization or other interactions shown to occur with more complex stimuli (Bonds, 1989; 1991; Albrecht, & Geisler, 1991; Albrecht, Geisler, Frazor, & Crane, 2002; Heeger, 1992; Wilson & Humanski, 1993; Carandini, Heeger, & Movshon, 1997; Wainwright, Schwartz, & Simoncelli, 2001). That is, while sensitivity for simple stimuli is widely reported to be superior at horizontal and vertical orientations and worst at oblique orientations, it has recently been reported that when tested with broad-band oriented stimuli, visual sensitivity is best at oblique orientations, worst at horizontal, and intermediate at vertical (the “horizontal effect”; Essock, DeFord, Hansen, & Sinai, 2003; see Figure 1, middle row). Thus, with customary clinic or laboratory testing, humans typically display an oblique effect, but when viewing the “regular” every-day visual world, the visual anisotropy is a horizontal effect.

A second problem with the traditional anisotropy dogma is that measurements at horizontal and vertical orientations may not have been compared with each other that carefully, and there may actually be a bias of horizontal over vertical in not only visual performance, but also in the neurophysiologic anisotropy and the anisotropy of natural scenes content. A recent extensive survey of single-unit orientation preferences indicates that neurons tuned to horizontal are more prevalent than vertical (Li, Peterson, & Freeman, 2003) and larger cortical areas can be seen at horizontal compared to vertical in some imaging studies (Chapman, Stryker, & Bonhoeffer, 1996; Chapman, & Bonhoeffer, 1998; Coppola, White, Fitzpatrick, & Purves, 1998). With respect to scene content, the prior examinations of the anisotropy of natural scene content (Switkes et al., 1978; Baddeley & Hancock, 1991; Hancock, Baddeley, & Smith, 1992; Huang & Mumford, 1999; Keil and Cristóbal, 2000), while reporting more horizontal and vertical content than oblique content, are not clear on whether the amount of horizontal and vertical content is equal or not. Some of the reports are conflicting and certain methodological and image sampling issues likely confounded the results (refer to the Methods section). Of the reports that specifically state differences in overall magnitude between horizontal and vertical orientations, Switkes et al (1978) reported a bias in favor of ver-

tical relative to horizontal for imagery possessing naturalistic content, however this conclusion is limited due to the very small set of scenes sampled. Keil and Cristóbal (2000) conducted a systematic comparison between the content biases at horizontal and vertical as a function of spatial frequency and found greater horizontal bias at certain spatial frequencies with a preponderance of vertical content at other spatial frequencies. However, since very narrow measurement sectors were utilized, their data likely suffer from differential discrete sampling error at the different orientations (refer to the Methods section). On the other hand, some studies (e.g., Baddeley & Hancock, 1991; Hancock, Baddeley, & Smith, 1992) report a greater bias for horizontal relative to vertical content. Of course, no definitive answer for all natural scenes can be determined as natural scene composition varies and the extent to which horizontal and vertical content differs within any given sample will depend on the specific environments in which the imagery is gathered. However, the findings summarized above argue for a larger bias of horizontal content relative to vertical content for typical or modal outdoor scenes. In the present study we address this issue by (1) using a very large sample of images, (2) considering different types of natural scene content, (3) using a method of analysis that avoids potentially confounding biases, and (4) specifically comparing horizontal and vertical content across spatial frequency.

To provide comparison psychophysical data, and to extend our prior reports (Essock et al., 2003; Hansen, Essock, Zheng, & DeFord, 2003), we assessed human performance for detecting oriented content in natural scenes and in broadband noise as a function of the orientation of the scene’s predominant content. We then compared this to the prevalence of natural scene content (i.e., amplitude) measured at different orientations in a large “random” sample of natural images from two independent image databases. We found a clear correspondence between amount of oriented content in natural scenes at various orientations and perceptual performance at corresponding orientations. Specifically, at orientations where visual performance with natural scenes is worst (horizontal), typical natural scenes contain the most content, and where visual performance is best (obliques), scenes contain the least content. The results from the natural scene analyses and the psychophysical experiments are then considered in the context of a divisive normalization model adapted from Wainwright, Schwartz, and Simoncelli (2001). Portions of this research were presented at the 2nd Annual Vision Sciences Society Meeting.

The organization of this report is as follows. First two sets of psychophysical experiments examining orientation performance with either broadband visual noise patterns or images taken of actual natural scenes are reported. Next analyses of the content bias for three different image sets with imagery obtained by our lab as well as from a different lab are reported. Lastly, we propose a divisive normalization model that builds on a previously proposed

model, but takes into account the results of the current psychophysical experiments and natural scene content bias analyses.

Methods

Psychophysical Experiments: Natural-Scene Selection and Stimulus Image Generation

One thousand seventeen natural scene images were obtained under a variety of lighting, weather, and seasonal conditions from various parts of Kentucky and Michigan in areas free of any man-made structures. A Sony DSC-F505 digital camera set at 1600 x 1200 resolution, with an F8.0 aperture and a fixed field of view of 40.0° x 31.5° was used. Photos were taken with the camera level to the ground plane. Although an effort was made to select a wide variety of scene content, no attempt was made to define "random" image sampling. The camera itself was assessed for any orientation biases that would potentially influence the amplitude spectra of the imagery. This was done by gathering control images where the camera was either aligned with the ground plane or rotated 45° and comparing the amplitude contained in corresponding 45° sectors centered at one of four orientations (i.e., vertical, 45°, horizontal, and 135° in the environment) obtained from the aligned and 45° rotated images. The camera orientation was found to have no significant effect upon the amplitude spectrum at any scene orientation (i.e., ground-plane aligned or 45° rotated), indicating for example that the CCD pixelation did not significantly affect the amplitude spectrum as a function of orientation within the range of spatial frequencies assessed.

Only well-focused images were included in the final sample pool. The images were loaded into Photoshop 5.0 at their original 1600 x 1200 size, transformed into grayscale and then resized to match the resolution setting of the display on which they were to be presented (800 x 600). The conversion into grayscale was achieved by converting the original RGB images into HSB images, then eliminating the hue and saturation image planes while maintaining the luminance plane. The process of down-sampling and conversion to grayscale had a very minimal effect on the luminance distribution of the images (e.g., ~0.9 shift of the standard deviation) which was represented in integer values in the range of 0 to 255. Lastly, a 512 x 512 section was cropped from the larger 800 x 600 image at a position selected at random, and served as the image to be considered for inclusion in the stimulus set.

Next, the orientation content bias conveyed in these images was assessed in order to evaluate scenes with different types of orientation bias. Accordingly, images were placed into vertical-, 45°-, horizontal-, or 135°-dominated categories, and an additional, "non-biased," category based on the relative amount of the image's total amplitude contained in each of four 45° orientation bands (at all

spatial frequency bands – see below). Orientation was defined clockwise from vertical (i.e., vertical = 0°). The oriented content-bias of an image for a particular orientation was defined as the percentage of an image's total amplitude that was contained in a 45° band (centered at 0°, 45°, 90° or 135°) (i.e., the ratio between the summed amplitude contained within a 45° sector and the value obtained from summing across the entire amplitude spectrum). Images were defined as non-biased if the percentage of oriented amplitude at the four orientations did not exceed that of the other orientations by 10%, and were judged as biased if the percentage at one orientation exceeded that at each of the others by 10% in each one of all three spatial frequency bands (high, 4-16 cpd; medium, 1-4 cpd, low, .25-1 cpd). For example, an image with a bias ~33% at a given orientation and spatial frequency band possessed amplitude biases at the other three orientation bands within the same 2-octave range ~22%. The average bias in all three bands was ~33%. Refer to Hansen, Essock, Zheng, and DeFord (2003) for a more detailed account of oriented content-bias calculation. Five image sets were assembled, one "naturally isotropic" (i.e., less than 10% bias at each orientation within each of the three spatial frequency bands) and four sets containing an oriented bias in amplitude at one of the four nominal orientations. Each set contained eight images, resulting in a total of 40 images.

Apparatus

Stimuli were presented by a SGI 540 Visual Workstation on a SGI 420C monitor with maximum luminance of 80 cd/m². To eliminate edge contours from the room and monitor bezel, a circular mask (visual angle: 59°) with a 17° circular stimulus aperture was fit to the monitor. For the natural scene stimuli, the monitor's gamma was set so that output was linear with the camera values (i.e., 2.5); for the visual noise stimuli, this value was set to 1.0. The frame rate was 120 Hz, with a resolution of 800 x 600 pixels. Single pixels subtended .021° visual angle (i.e., 1.25 arc min.) as viewed from 1.325 m.

Participants

Four people (naïve to the purpose of this study) participated in the current study. All participants had normal vision and were further screened by a series of vision tests to assure that they had no residual astigmatism. The age of the participants ranged between 18 and 21. IRB-approved informed consent was obtained.

Stimuli and Procedure

Natural Scene Images All images (described above) were fast Fourier transformed (FFT) using MATLAB (version 6.5) and corresponding Signal Processing Toolbox (version 6.1) and then filtered to be isotropic by aver-

aging each spatial frequency coefficient across orientation and replacing the coefficient at that spatial frequency for all orientations with the average value, thus ensuring equal amplitude at each orientation while maintaining the form (slope) of the amplitude spectrum for each image. To create the test stimuli containing the oriented increment to be detected, the spectra were then multiplied by a filter constituting a broadband increment within a 45° orientation band and all spatial frequencies (i.e., within the filter consisting of a "wedge" or "bow-tie" shaped region in the frequency domain) with a triangular weighting function (the peak of the triangle increment was 30% of the corresponding amplitude coefficients) centered on one of the four test orientations (refer to Figure 2, bottom for further details). Each resultant spectrum was then inverse Fourier transformed to create the spatial images (Figure 1 and 2, bottom), with pixel integer values ranging from 0 to 255. The operations described above resulted in only very minimal 'clipping' of pixel values in the resultant spatial stimuli ($\sim 0.5\%$ of the total viewable pixels), and was assumed negligible (cf. Knill, Field, & Kersten, 1990; Webster & Miyahara, 1997). All stimulus images were windowed with an edge-'blurred' circle and subtended 10.5° visual angle at a viewing distance of 1.325m.

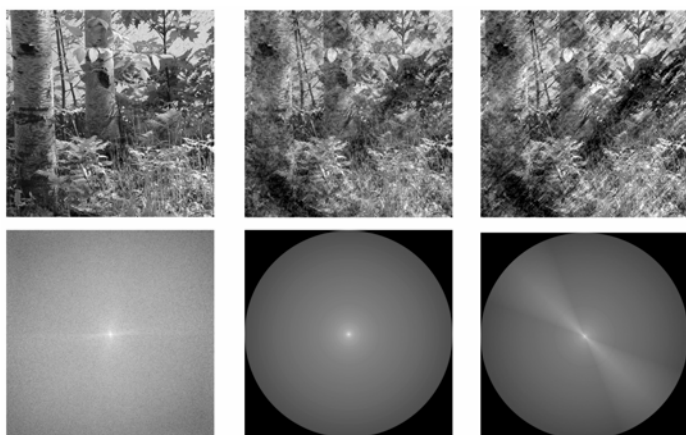


Figure 2. Example of the method by which an oriented increment in amplitude was applied to each stimulus image. Top row: (left to right) the original unaltered spatial image, the isotropic image, and the test image containing an increment of oriented content at one orientation (i.e., a 45° band weighted in the orientation dimension by a triangle filter). Bottom row: the frequency domain spectrum is shown immediately below the spatial image to which it corresponds.

A Yes/No single-interval task was employed in which observers responded by key-press to indicate whether or not the presented stimulus contained the oriented increment. In order to familiarize the participants with the stimuli, all participants were shown examples of stimuli with and without oriented increments (with a 50% increment of amplitude at the peak of the triangle weighting function) prior to testing. An experimental session was formed by splitting the set of 40 images into two random

groups of 20 images (i.e., selecting four of the eight images from each of the five content types -- the naturally isotropic images and the content biases at the four different orientations employed). This resulted in two subsets which were run in separate experiment sessions. For a given subset, a single session consisted of four blocks, one for each orientation, with each block consisting of replications of each stimulus in the subset (e.g., 20 with increment and 20 without increment trials). Orientation order and trials within blocks were presented in random order. All subjects repeated each subset session four times (6,400 trials total). A single trial consisted of a fixation pattern (500msec), followed by the stimulus image (400msec), followed by a white-noise mask (500msec). An unbiased measure of sensitivity, d' , was used as the measure of sensitivity (Creelman & Macmillan, 1991) and calculated on the basis of 320 trials (per orientation of the increment for each content-biased image set). All participants were allowed four practice sessions (one for each of the nominal test orientations) with auditory feedback. Auditory feedback was not given during the actual experimental sessions.

Visual Noise Images A second set of stimuli consisted of visual noise patterns that were generated in the Fourier domain by analogous means (see Essock et al. 2003). In this case however a $1/f$ amplitude spectrum, characteristic of natural scenes (van der Schaaf & van Hateren, 1996), was generated and combined with each of 50 different random phase spectra (each with the random phase values assigned with respect to conjugate symmetry) which then were inverse Fourier transformed to form the space-domain stimuli; with r.m.s contrast set at 70% (see Essock et al, 2003). Each of the 50 resultant random noise patterns were then altered to contain an oriented increment of amplitude in the same manner described above for the natural scene stimuli. These spatial visual noise stimuli were also windowed with an edge-'blurred' circle subtending 10.5° visual angle. The experimental paradigm was identical to that described in the natural scene procedures, with the exception of a somewhat briefer stimulus interval duration and that each noise pattern was presented twice within a block of trials, once with an increment, and once without (i.e. 100 trials per block, 400 trials per session, and 1,600 total trials).

Natural Scene Analysis Procedure

To assess the anisotropy of natural scene content we gathered three different image sets (two from our lab and one from a different lab). Details about how the images from these image sets were processed for this analysis are provided below. Note that these images were not modified (i.e., no increments of amplitude were applied to these 'analysis only' images as for the psychophysical test stimuli). The first set (Image Set 1) consisted of two-hundred

thirty-one images (cropped to 1024 x 1024 pixels) that were selected at random from the full set of 1017 images described earlier with the one stipulation that an equal number of scenes were selected from each annual season. The camera output was linearized by applying the appropriate correction exponent to the image values (i.e., the inverse of the exponent applied by the camera -- i.e., 2.5). Since our spatial measurements of the stimuli were made in the frequency domain, we were careful to assess the potential impact of the edge of the image on the amplitude spectrum (i.e., "edge-effects"). That our edge-'blurred' spatial window effectively eliminated this concern was verified by windowing all imagery with a Gaussian function (a common approach to avoiding significant edge-effects) and obtaining essentially identical results.

Ratios indicating the magnitude of the orientation biases in the imagery were obtained for each image at each orientation in the manner described above. Three categories of images (15 images each) were defined and selected from the 231-image random sample; the three categories were defined on the basis of either containing a dominant horizon line, only the ground plane (i.e., various textures varying with season), or neither (e.g., images of general foliage, shrubbery, etc.). In addition, a fourth category of 186 images was created by removing all images from Image Set 1 that contained a predominant horizon line or receding ground plane (i.e., creating a set without any of the obvious spatial content presumed to create a horizontal bias). These four categories were then analyzed in terms of oriented content as described earlier.

The second set of imagery was obtained from a widely used and well calibrated image database compiled by a different lab (<http://hlab.phys.rug.nl/archive.html>, see van Hateren and van der Schaaf (1998) for detailed information about this imagery) for the purpose of an independent confirmation of the results from Image Set 1. Two-hundred images were randomly selected from this database to form the second image set (Image Set 2) which resulted in a variety of scene types. The random sampling process was conditioned so that only imagery devoid of man-made content would be selected. Given that this imagery is currently made available in 21 sets of 200 images, random sampling was also conditioned on sampling ten images per set (excluding set 1401-1600 as it consisted only of images of man-made content). These images were then cropped to 1024 x 1024 pixels.

The third image set (Image Set 3) was gathered in order to provide a highly detailed analysis of the distribution of the amplitude across orientation and spatial frequency. As mentioned in the Introduction, some previous reports have attempted to provide a detailed measurement of amplitude within very narrow orientation bands (e.g., 5° sectors) as a function of spatial frequency. However, a fundamental problem with such approaches is that, due to the discrete sampling of the digital Fourier transform, very narrow orientation band sectors centered at orientations other than 0°, 45°, 90°, and 135° will not sample from the

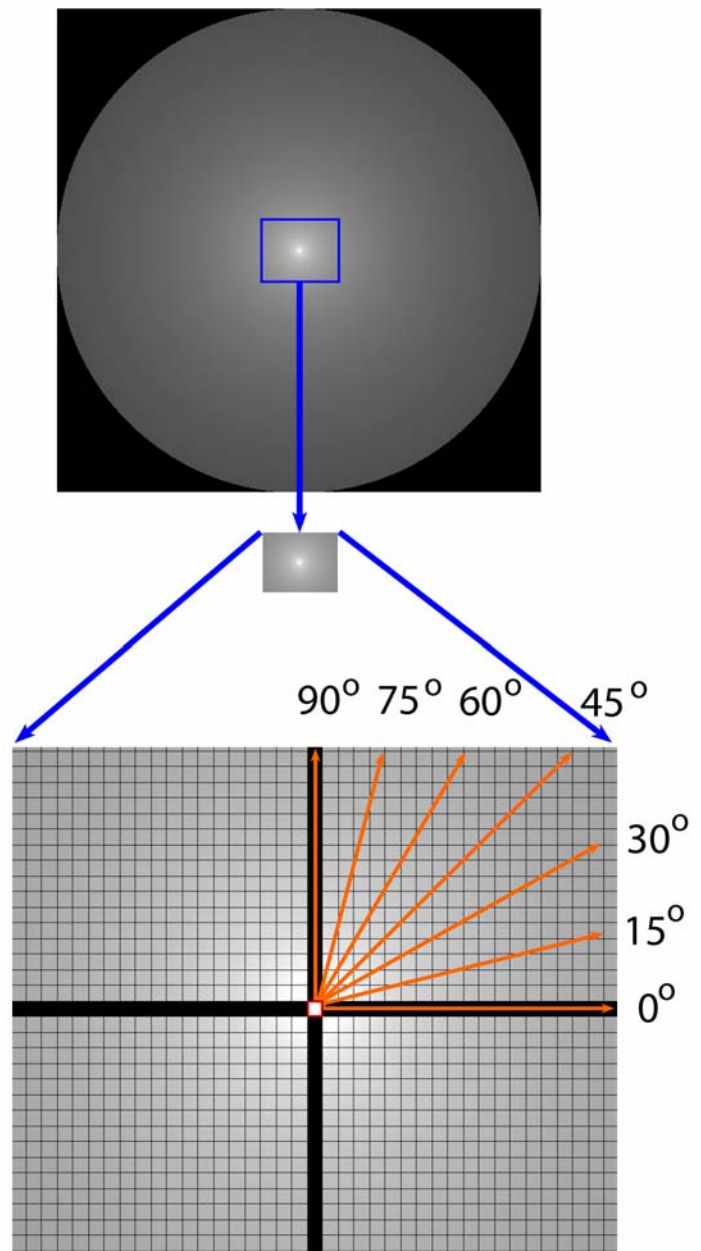


Figure 3. Schematic representing the lack of a set of continuous orientation vectors in the Fourier domain (i.e., the discrete Fourier domain amplitude spectrum). Top: Example isotropic amplitude spectrum with a region of interest highlighted in blue. Middle: The same region of interest shown in isolation. Bottom: The same isolated region of interest enlarged with a matrix grid laid on top. This grid represents the nature of the discrete sampling incurred by the discrete Fourier transform. All of the 0° and 90° coordinates have been shaded in black for clarity. The center square (white, highlighted in red) represents the DC component. The orange arrows represent orientation vectors drawn at arbitrary orientations in steps of 15°. Notice how the 0°, 45°, and 90° vectors contain sampling points (i.e., centers of the squares) along their entire lengths, whereas the other vectors do not. Refer to the text for further details.

lower range of spatial frequencies. Specifically, a continuous Fourier transform will produce an amplitude spectrum, that, when plotted in polar coordinates, will yield a vector for each possible orientation, with each point on a given vector representing amplitude at a specific spatial frequency at that given orientation. However, as shown in [Figure 3](#), due to the discrete representation of the amplitude spectrum (produced via the discrete Fourier transform), not all of the discrete steps of spatial frequency can be contained in most of the oblique orientation vectors (with lower spatial frequencies being most underrepresented). Only the vectors at the nominal orientations mentioned above possess amplitude coefficients across the full range of spatial frequencies produced by the digital Fourier spectrum, with many orientations not having any of the lower spatial frequencies represented. Note that at the 45° and 135° diagonals, the same number of spatial frequency samples are present as at the cardinal orientations, but that the values are slightly higher. A procedure that sums along orientation vectors within a specific segment (spatial frequency range) is the best that can be achieved but will strongly bias amplitude measurements thereby yielding underestimates at orientations other than the four mentioned. The problem with such an approach is that one is left with only a measurement of amplitude for different spatial frequencies at four orientations. Thus if one required samples at orientations other than those four orientations, those vectors would have to be 'aligned' with that content. For example, if one wished to measure the distribution of amplitude across a full range of spatial frequencies at, for example, 3° in an image, the spatial content of that image would have to be sampled in a way such that it would be plotted along one of the four nominal vectors in the Fourier amplitude spectrum. One way to achieve this would be to rotate the image 3° via some interpolation algorithm (e.g., bicubic interpolation). However, such algorithms involve considerable amounts of error in the interpolation process which could have deleterious effects in the frequency domain. An alternative method that avoids this problem, although time intensive, is to physically rotate the imaging device such that the spatial content at 3° would be depicted along one of the ideal vectors mentioned above. The latter approach was employed in the current analysis (i.e., Image Set 3).

Image Set 3 consisted of 60 natural scene images that were obtained from various parts of Kentucky and Michigan in areas free of any carpentered structures. A Minolta Dimage 7_{Hi} digital camera with 2560 x 1920 resolution, an F8.0 aperture and a fixed field of view of 51.6° x 42° was used. For each of the 60 scenes, the camera was rotated in 3° steps, which resulted in 31 images per scene. The rotation of the camera was achieved by a precision rotatable mount (a standard rear double filter box -- Lindahl Specialties Inc.) that was fixed to a professional quality tripod (Star-D Mfg. co. Inc.). The camera was attached to the filter box by fitting it with a 48 mm filter box adapter ring which fit tightly into the filter box. The filter box was

then scribed with calibration marks in 3° steps, starting at 0° lateral (camera level to the filter box – which was aligned with the ground plane during the image sampling process) to 90° vertical (camera perpendicular to the ground plane). With the camera level to the filter box, the adapter ring was marked with a single point (alignment point) aligned with the 0° point on the filter box, this allowed for the rotation of the camera by the specified amount by aligning this point to any one of the 3° points engraved onto the filter box.

The sampling procedure for a given scene involved the following steps. First the tri-pod was set up and the position of the attached filter box was leveled with the ground plane. Next the camera was mounted to the filter box via the adapter ring with the alignment point lined up with the 0° point on the filter box. Thus, in this position, the camera was aligned such that it was level to the ground plane. Once aligned, an image was sampled. The camera was then rotated 3° counter-clockwise to the second mark on the filter box, at which point another image was sampled. This process was repeated until the alignment point on the camera's adapter ring was aligned with the 90° point on the filter box (camera perpendicular to the ground plane). Therefore, this process results in a sampling of the same scene rotated in 3° counter-clockwise steps. This procedure was carried out under conditions that minimized changes in the scene across time (e.g., winds less than 5mph, full sun, or fully overcast conditions) to assure that the various positions of natural content was the same for each rotation of the camera. As with the other image set, an effort was made to select a wide variety of scene content, however no attempt was made to define "random" image sampling. Thumbnail examples of the sampled natural scenes are provided in the [on-line supplement](#).

The analysis of each scene involved the same procedures mentioned above with respect to cropping (the central 1024 x 1024 pixel region), linearization (gamma exponent 2.5), and spatial windowing with a Gaussian function to reduce any influence of the 'edge effects.' The 30 rotations (not counting the first sample – i.e., the aligned image) allowed us to utilize two of the four optimal vectors mentioned above for a complete sampling of orientation across the range of 0° to 180° in steps of 3° (with 0° and 180° being identical measurements of the same orientation, in this case vertical). That is, across the 31 images for a given scene, the 0° vector (see [Figure 4](#)) in the Fourier domain allowed for the measurement of spatial content in the range of 90° to 180° (again at 3° steps – i.e., 90°, 93°, 96°.....180°), and the 90° vector allowed for the measurement of spatial content in the range of 0° to 90° (refer to [Figure 4](#)). All four vectors could not be used together because of the lack of a one-to-one correspondence between the 0°/90° vectors (or the 'cardinal vectors') and the 45°/135° vectors (or the 'oblique vectors') with respect to spatial frequency. For example, a given position on either of the oblique vectors corresponding to the same position on either of the cardinal vectors will contain the

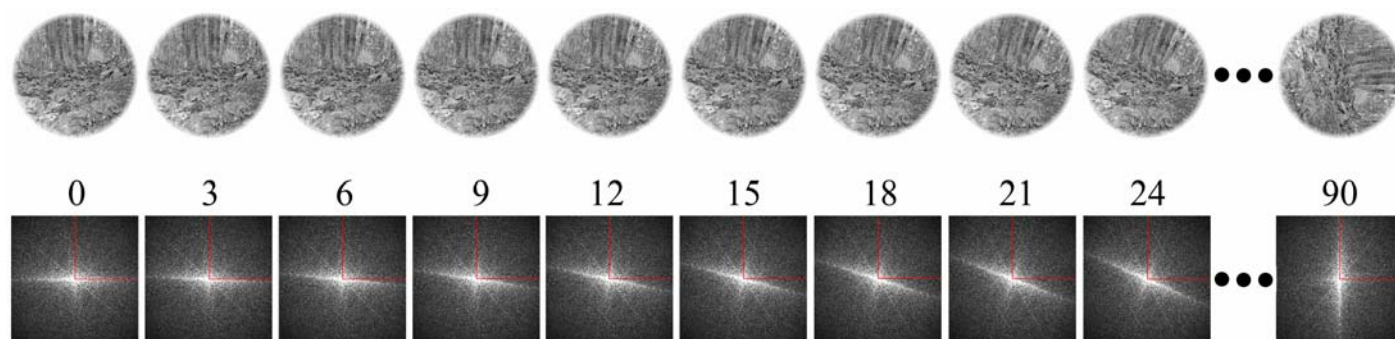


Figure 4. Schematic depicting the process involved in calculating the amplitude biases at 61 orientations (i.e., 0° through 180°). The top row shows a series of rotations for an exemplar scene starting at 0° (camera aligned) to 90°. The bottom row shows each amplitude spectrum corresponding to its respective spatial image. The two red lines drawn on each spectrum indicate the 0° and 90° vectors that were taken from each image rotation. Note that for the un-rotated image sample and the full 90° rotated image sample, camera-aligned vertical (0°) and horizontal (90°) content will be sampled twice. Here, vertical (0 on the theta axis) was taken from the un-rotated sample and horizontal (90 on the theta axis) was taken from the full 90° rotated image. Since 0°/180° correspond to the same spatial content, the same vector sample (from the un-rotated image) was used for both (refer to the text for further details).

amplitude for a spatial frequency equal to the corresponding cardinal spatial frequency multiplied by $\sqrt{2}$. For the current analysis, we chose to use the 0° and 90° vectors. The measurements obtained were analyzed in two ways. First, overall magnitude, collapsed across spatial frequency, obtained by averaging all of the amplitude coefficients along the respective vector for each of the 61 orientations sampled up to the Nyquist limit of the imagery (i.e., 512 cycles per picture). Second, in order to examine orientation biases as a function of spatial frequency, each orientation's respective vector was parsed into 20 bins (the maximum allowed by the Nyquist limit of this imagery), with each bin's amplitude coefficients being summed in order to provide a measure of amplitude contained at each of the 61 sampled orientations for each cycle per degree, ranging from 1cpd to 20cpd.

Results and Discussion

Visual Performance

Performance for detecting oriented content in natural-like noise stimuli and in natural scenes is shown in Figure 5a and 5b, respectively. With the natural-like noise stimuli (Figure 1 middle row), a horizontal effect (worst performance for horizontal, best for obliques) was obtained (expanding on the test conditions reported in Essock et al., 2003 and Hansen et al., 2003).

When tested with stimuli containing natural scene content, a horizontal effect was also obtained. Specifically, when viewing natural scene stimuli with no predominate orientation bias (Figure 5b, non-biased group), oriented content contained in the scenes is hardest to see at horizontal, and easiest to see at oblique orientations (one-way repeated measures ANOVA: $F_{(3, 9)} = 10.48$, $p = .013$). In-

deed, the ability to detect oriented content in the (non-biased) natural scene images showed a horizontal effect quite comparable to that obtained in the noise comparison condition (one-way repeated measures ANOVA: $F_{(3, 9)} = 21.82$, $p = .01$). The non-biased group of natural scene images had been pre-selected on the basis of being naturally "un-oriented," that is, their subject matter contained a mixture of approximately equal amounts of content at all orientations, and furthermore, were filtered to make them exactly isotropic (see Methods section). Thus, the anisotropy obtained with these natural scene stimuli can not be attributed to a bias of global orientation or content. Of course, since the horizontal effect was obtained with the noise control condition, natural scene structure per se (e.g., semantic meaning or phase relations creating local edges) can not account for the horizontal effect.

In addition to testing "un-oriented" images, sensitivity for detecting oriented content in scenes that did have an initial bias of oriented content at one of the four orientations (before isotropic filtering) was also tested. An analogous performance anisotropy was observed with these scenes in that detectability of horizontal content was still poor regardless of the orientation of the predominant content in the original scene (Figure 5b, "biased" images). However, even though these scenes were all filtered to remove any orientation difference in amplitude prior to testing, the scenes that formerly had more amplitude at a particular orientation produced poor performance when the test orientation was at the orientation of the former content bias (i.e., in addition to poor horizontal performance).

Natural Scene Content

We suggest that the consequence of decreased perceptual salience of some contours in a natural scene would be to make contours of other orientations, and thus objects

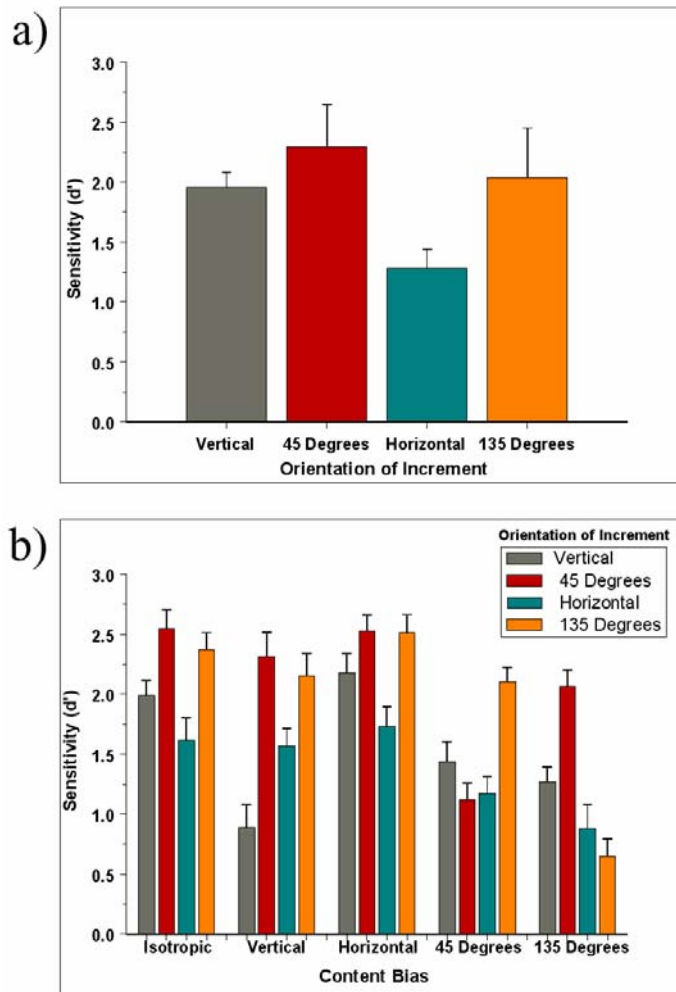


Figure 5. Results show a horizontal effect for both the noise (random phase) and natural scene stimuli. (a) Data from the $1/f$ visual noise condition. Each bar is labeled in terms of the orientation of the increment of amplitude. (b) Data from the natural scene condition. The abscissa is grouped into sections based on the content type (i.e., predominant orientation) of the images used (eight images for each section). Each of the four bars plotted in each section represent the orientation of the increment of amplitude. All error bars are S.E.M).

containing significant power at a range of orientations, relatively more salient. Specifically, this would increase the salience of a typical object when viewed against a background of a natural scene to the extent that a typical natural scene contains relatively more horizontal content, intermediate vertical content, and least oblique content. We suggest that this is indeed the typical make-up of natural scenes in that: (1) due to the frequent presence of the horizon and the preponderance of contours at or near horizontal in addition to the nature of texture gradients associated with foreshortening, natural scenes, on average, seem likely to contain predominantly horizontal content, and (2) from the nature of vegetation (i.e., the phototropic and gravity-directed growth) there is likely to be a secondary preponderance of vertical structure.

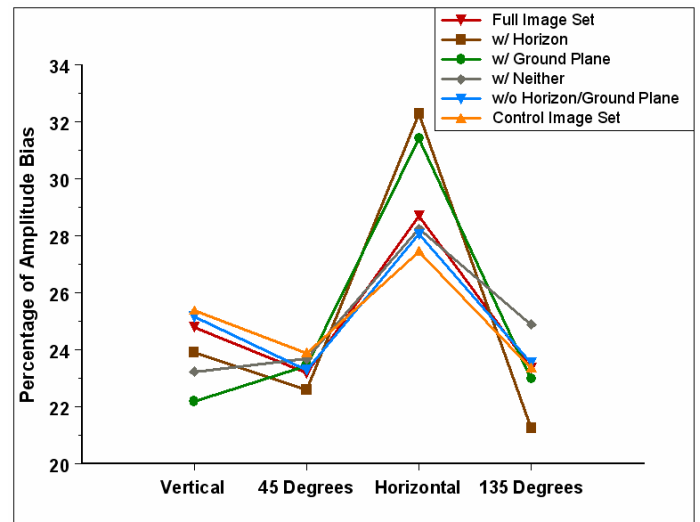


Figure 6. Orientation analysis of natural scene content. Plotted is the average ratio of amplitude at a given orientation relative to the other orientations. The three categories: “w/Horizon”, “w/Ground-plane”, or “w/Neither” plot the averages from subsets of images containing a clear horizon; containing ground-planes consisting of various textures; or containing general foliage and shrubbery, thus containing neither a horizon or a ground plan. Also plotted are the measurements made over the entire 231 natural scene image set from which the subsets were drawn (“All Images”), and the set of all images remaining after those containing a horizon or a ground plane were removed (“w/o Horizon/Ground Plane”). Finally, measurements made on a control set of images obtained by a different lab (see text). Note that all six conditions show a strong *horizontal* bias. Second most prevalent is vertical content in typical scenes (although the vertical bias is not present in scenes of uniform ground planes (where horizontal dominates) or of general shrubbery).

Frequency analyses were performed on two independent, large random samples of natural scene images consisting of 231 of our images (Image Set 1) and 200 images from another lab (Image Set 2, van Hateren & van der Schaaf, 1998), as well as a third natural scene image set obtained with rotation of the camera to allow for a more precise measurement of amplitude across orientation as a function of spatial frequency (Image Set 3, see Methods). It was the intent of these analyses to quantitatively evaluate the above conjecture that in addition to the cardinal bias, content at or near horizontal typically dominates that at or near vertical. The results of the analyses carried out on the three images sets show that the most content (i.e., the most amplitude in the frequency domain) was indeed at the horizontal orientation and second most was at the vertical orientation (Figure 6). The exact distribution of content obtained, of course, depends upon the image sample analyzed, and one might suspect that the horizontal bias obtained is due to horizon-containing images. However, we find this predominance of horizontal content in a high percentage of images that do not contain a horizon as well as those images that do (Figure 6).

What are the correct scenes to select in order to consider the possible evolutionary significance of the natural environment? Unfortunately, in terms of selecting a set of scenes for analysis that mimics the natural visual input of an animal during typical activity, there can be no “correct” or “unbiased” sample of images, especially if one considers that this should be done with respect to an evolutionary time scale. That is, even if one obtained images at random time intervals while moving through a modern-day natural environment, the orientation content will depend upon the type of “natural” environment selected (e.g., forest versus sand dunes), the locations within the environment one chose to walk, and how one oriented the camera (i.e., “framing” of the picture, including the rotation and the tilt of the camera) to mimic the visual orienting of the “evolutionary” animal. Here, we simply note that the results of analyses of image sets 1 and 2, as shown in Figure 6, demonstrated that: (1) horizontal physical content indeed predominates in horizon-containing images, (2) horizontal content predominates even in non-horizon-containing scenes composed of ground surfaces, hillsides, or other regions consisting of similar vegetation or structure, (3) horizontal content predominates in a sample of scenes that contain neither a horizon or ground plane (such as close-ups of bushes, brush or general foliage), (4) horizontal structure persisted in dominating the analysis even when all imagery containing a receding ground plane or predominant horizon line were removed from the image sample, (5) a horizontal content bias was also found in an alternative set of “standardized” calibrated imagery frequently used in natural scene analysis (van Hateren & van der Schaaf, 1998), and (6) there is a suggestion of a predominance of horizontal content evident in certain prior published reports (Baddeley & Hancock, 1991; Hancock, Baddeley, & Smith, 1992; Keil & Cristóbal, 2000). Secondly, we note that the results of this analysis show that the vertical orientation (i.e., the 45°-wide bin centered on vertical) is second most prominent in typical scenes in our sample as well as in the large random sample taken from the imagery of van Hateren & van der Schaaf (1998).

Since the analyses described above involved ratios of the summed amplitude coefficients in a 45° wedge centered at each of the four primary orientations to the summed amplitude of the entire spectrum, it cannot be determined from the data just how the distribution of amplitude at orientations at or near the nominal orientations contributes to their respective biases at specific orientations or across spatial frequency. However, for reasons discussed earlier, the analysis carried out on the camera rotation imagery (i.e., Image Set 3) allowed for a more continuous measurement of amplitude at numerous orientations as a function of spatial frequency. These data (Figure 7a) clearly show that there is a bias in summed amplitude at and near 90° (horizontal content) that indicates more horizontal content relative to the other orientations. Figure 7b plots the averaged amplitude for each

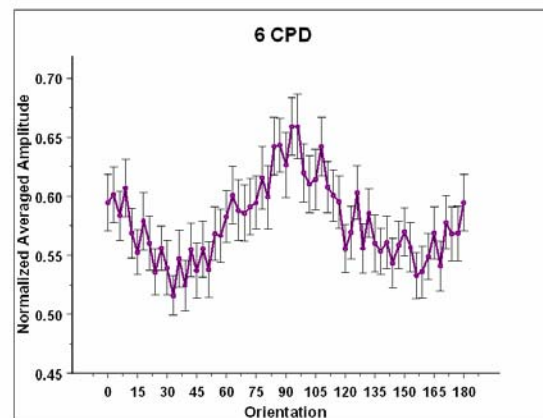
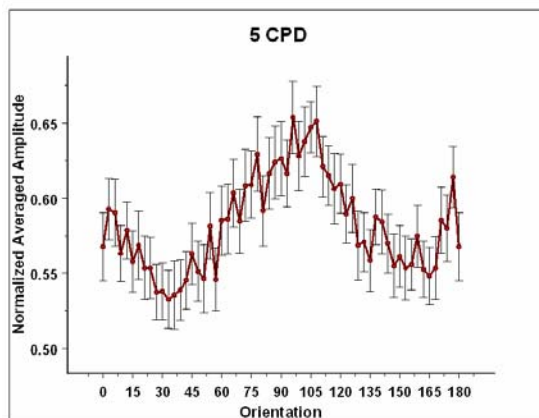
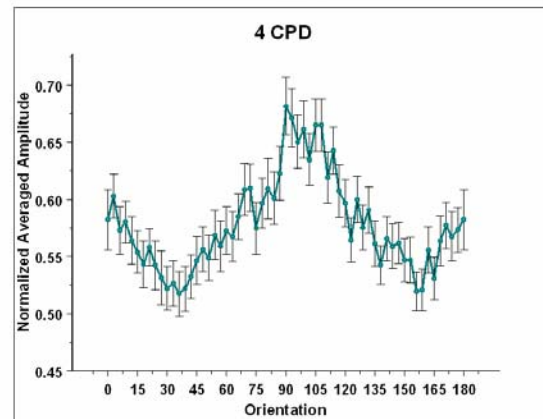
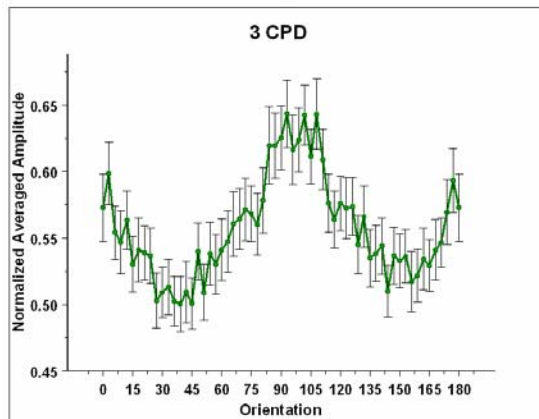
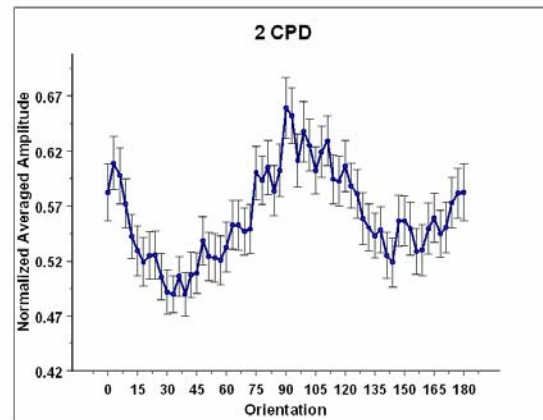
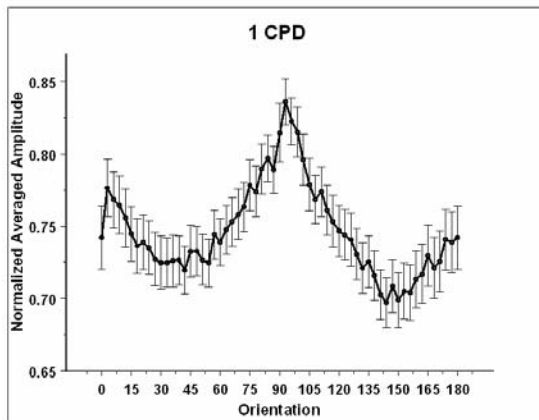
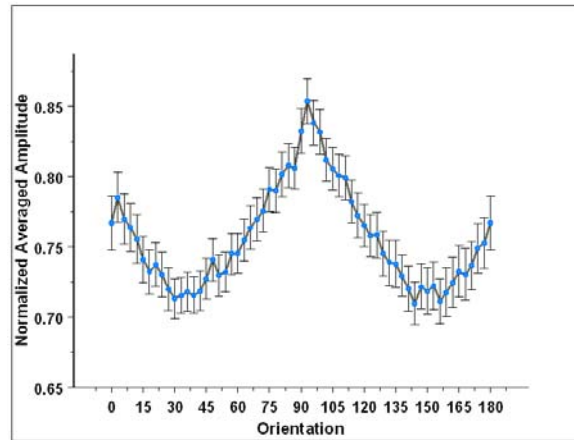
cycle per degree in the range of spatial frequencies allowed by the Nyquist limit of this imagery. There is a clear bias in amplitude at the cardinal orientations at each spatial frequency. Horizontal content is the most prevalent at all spatial frequencies. The second most prevalent content is always at vertical, although its prominence diminishes at the highest spatial frequencies.

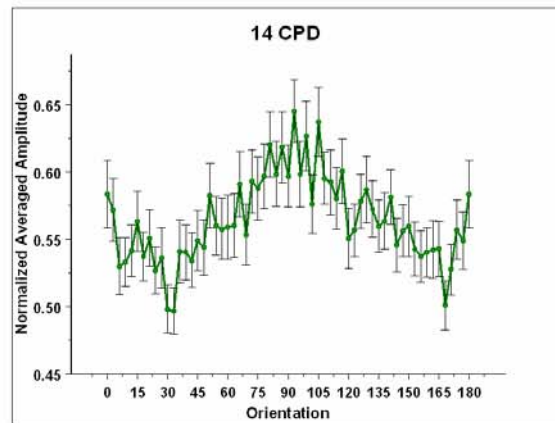
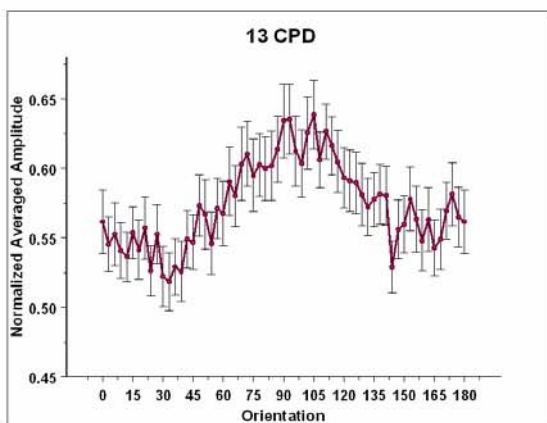
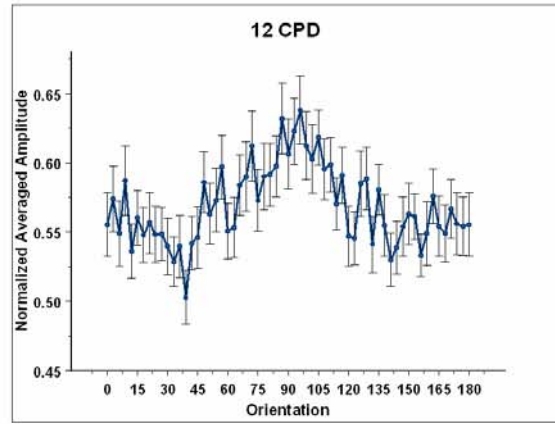
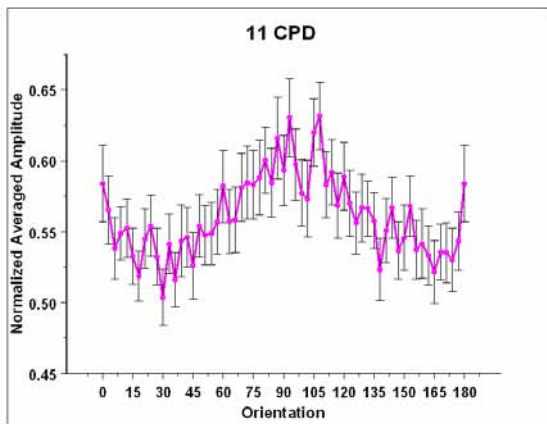
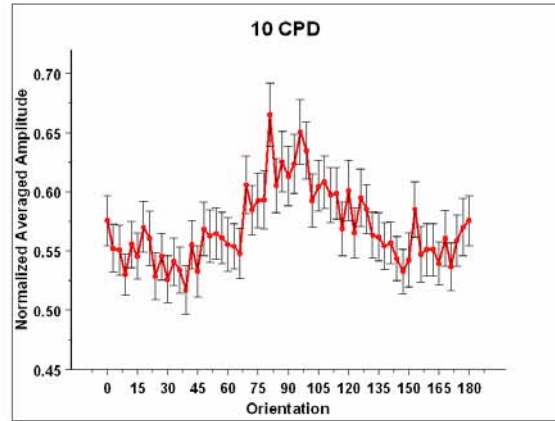
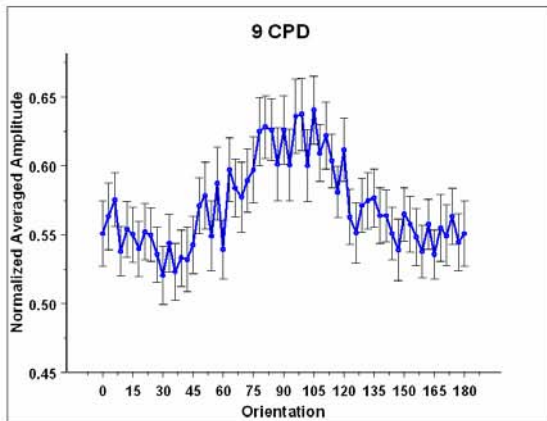
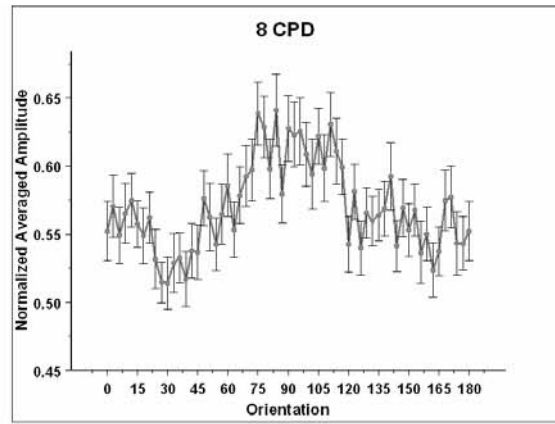
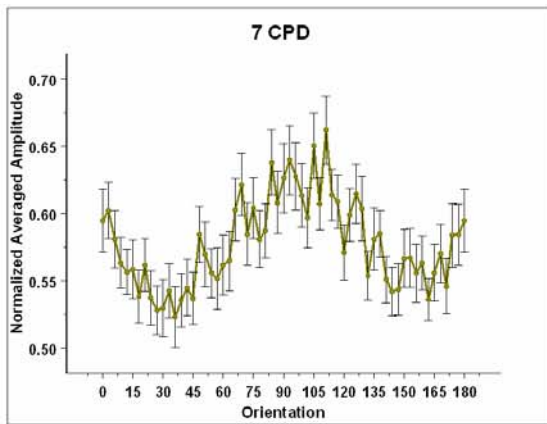
Toward A Model of Orientation Processing in Broadband Scenes

Two main findings from the current paper constrain a model of orientation processing of broadband scenes. First, the results from the psychophysical experiments indicated that instead of broadband targets at oblique orientations being seen most poorly, horizontal stimuli were seen most poorly and oblique stimuli were seen best, with vertical performance intermediate. Thus, when compared to the perception of an isolated grating or line stimulus, the presence of additional orientation components in a visual stimulus results in interactions that strongly alter the relative visibility of oriented content at various orientations. Due to these interactions, the oblique effect obtained with simple stimuli does not extend to naturalistic viewing situations as many have presumed; specifically, a horizontal effect is obtained instead (Essock et al., 2003). The second finding is that although sensitivity to oriented content was examined in the context of natural scenes that contained a natural bias in oriented content (containing predominant content at either 0°, 45°, 90°, or 135°), the horizontal effect could still be observed, when the orientation of the broadband increment matched the orientation of the content bias of the imagery, performance for detecting those orientation increments was dramatically reduced, thus suggesting the presence of an additional ‘content-dependent’ effect that changes with changes in viewed content. Thus the data support previous findings (Essock et al., 2003; Hansen et al., 2003) of the existence of two types of visual processing anisotropies. The first, the horizontal effect, can be considered a ‘static’ or ‘inherent’ effect; the second, the content-dependent effect, can be considered to be more ‘dynamic’ as it depended on the type of content bias present in the natural scene imagery. The following two subsections address each of these effects with respect to a model of orientation processing in visual cortex for naturalistic broadband imagery.

The Inherent Orientation Processing Anisotropy: A Horizontal Effect

The association between the behavioral performance horizontal effect observed with broadband naturalistic stimuli (i.e., broadband visual noise and natural scene imagery) and the prevalence of content at the nominal orientations in natural scenes was re-examined in the current





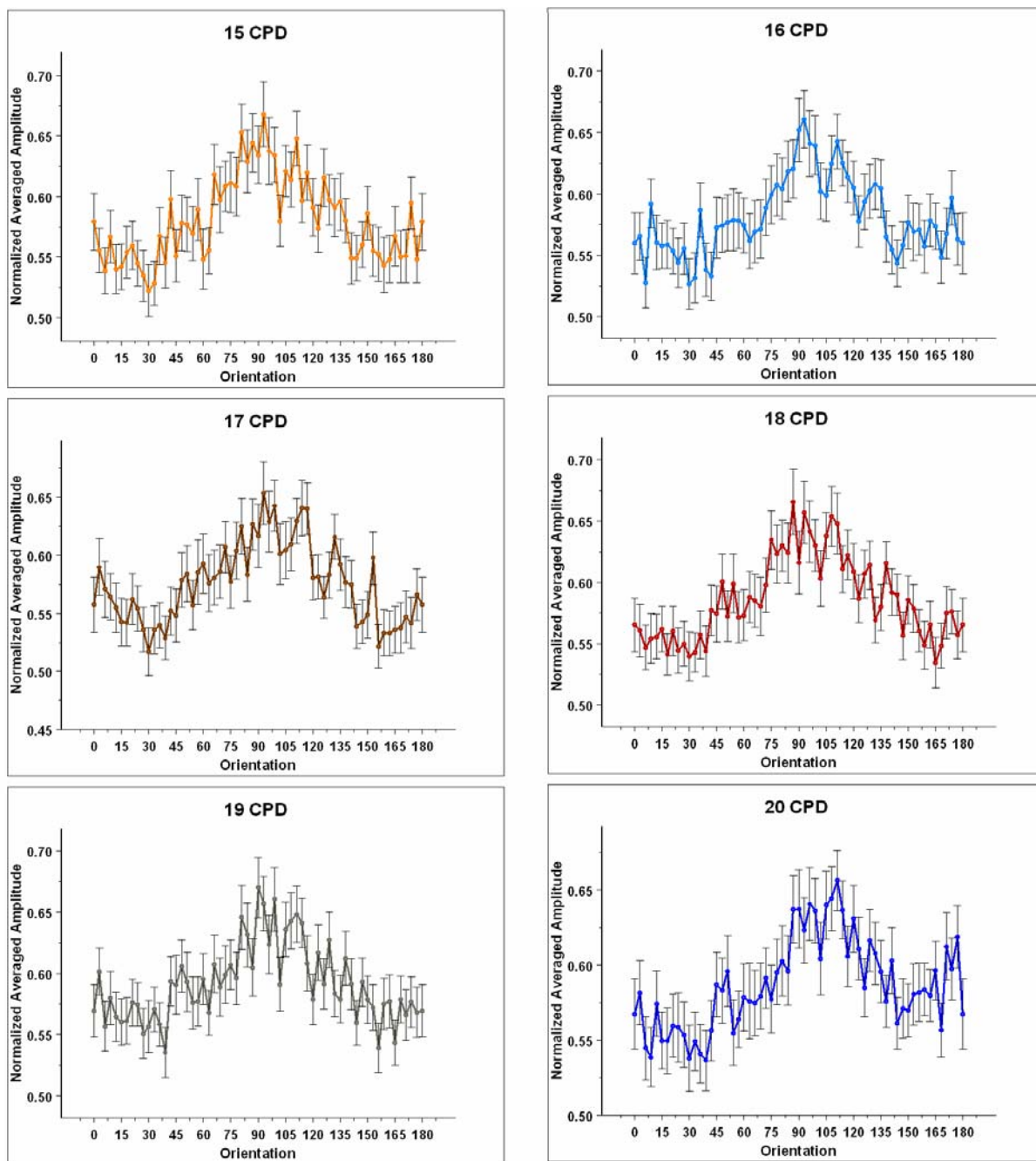


Figure 7. Plots of normalized amplitude at each of the sampled orientations. **a**—single graph: top) Average normalized amplitude for each of the sampled orientations. Each point on this plot represents the average over the 60 images of the vector for that orientation (abscissa) summed across spatial frequency. Note that the orientations with the most amplitude are located at or near horizontal (here 90° corresponds to horizontal spatial content). **b**—20 graphs: bottom) Average normalized amplitude bias for each of the sampled orientations plotted with respect to spatial frequency (cycles per degree). Each point on these plots represents the average summed vector segment for that orientation (abscissa) and respective spatial frequency averaged across the 60 images (refer to text for further details). Note that the bias in amplitude at or near horizontal orientations is clearly present at all spatial frequencies, and that the second peak at or near vertical orientations is largest at the lower spatial frequencies and diminishes towards higher spatial frequencies.

study. That the content contained in typical scenes was found to exhibit a horizontal bias is an important finding as we have previously proposed that the behavioral horizontal effect would have evolutionary utility in such environments (Essock et al., 2003; Hansen et al., 2003). Spe-

cifically, such a hypothesis predicts the existence of a cortical mechanism (presumably at the level of striate cortex) that acts to reduce the perceptual salience of the most prevalent content (i.e., horizontally oriented structures) in a scene, thereby relatively enhancing the less often occur-

ring content of natural scenes. That is, a mechanism that turns down sensitivity for the 'expected' content in a typical scene would serve to relatively enhance the salience of 'unexpected', or novel, content at off-horizontal orientations. This pattern of sensitivity adjustment could most likely be accounted for by a type of specialized cortical gain control mechanism. However, the change in the orientation sensitivity obtained with broad-spectrum stimuli cannot be expected from standard models of contrast gain control (e.g., Bonds, 1989; Heeger, 1992; Geisler & Albrecht 1992; Wilson & Humanski, 1993, Carandini & Heeger, 1994; Carandini, Heeger, & Movshon, 1997). Typical models propose that the output of V1 cortical units is modulated by division of their response (or their input) by the summed activity of other units pooled equally across all orientations and some (if not all) spatial frequencies; thus these models assume equal amounts of modulation upon these cortical units ('channels') of different orientations and spatial frequencies. However, the results from the psychophysical experiments in the current study indicate that: (1) the weights for various orientations contributing to the normalization pool are not equal, and (2) the divisive effect acts more selectively with respect to orientation (i.e., only similarly tuned units adjust the other's output). Specifically, when a given broadband test pattern in the current experiments was oriented obliquely, it apparently caused the gain to be turned down less than when it is oriented horizontally (and to a lesser extent, vertically). Consistent with this proposal, numerous studies have indicated that among striate cortical neurons mediating central vision, horizontal and vertical preferred orientations are somewhat more prevalent than oblique orientations (Mansfield, 1974; Tiao & Blakemore, 1976; Mansfield & Ronner, 1978; Orban & Kennedy, 1980; De Valois et al., 1982; Chapman et al., 1996; Coppola et al., 1998; Li et al., 2003). Thus, when the output of the different units is pooled in restricted orientation (and presumably spatial frequency) ranges, the divisive signal would be weaker at oblique orientations, resulting in the observed stronger response at oblique orientations when viewing broadband patterns. In other words, when the horizontal orientations of the amplitude spectrum of any given broadband pattern are incremented, this might cause more total pooled activity at the horizontal (and to a lesser extent vertical) test orientation than when oblique orientations are incremented, thereby turning down the output of the units detecting the test pattern at horizontal more than when the pattern is at oblique orientations. Accordingly, such an adjustment would produce a relatively smaller perceptual response for horizontally oriented content compared to obliquely oriented content in a broadband pattern.

Dynamic Orientation Processing Anisotropy: The Content-Dependent Effect

Considerable research has shown that when differently oriented simple stimuli (e.g., sine waves) are presented simultaneously (e.g., in cross-orientation inhibition or surround suppression experiments), a contrast normalization mechanism alters the sensitivity to a test stimulus. A reasonable assumption is that when viewing natural scenes with their broad spatial scale and broad orientation content also evokes contrast gain adjustments. Most of these contrast normalization models suggest that the current image (or "recent" image assuming a processing delay in the mechanism) is filtered by the array of orientation and spatial frequency filters (or a subset in certain models) and that their responses to the current/recent stimulus is pooled in a normalization pool that alters the gain of the output unit under consideration (Bonds, 1989; Heeger, 1992; Geisler & Albrecht 1992; Wilson & Humanski, 1993, Carandini & Heeger, 1994; Carandini, Heeger, & Movshon, 1997). That is, the activity level of this pool varies as the overall image content changes, but units tuned to different orientations are affected equally by the normalization pool. If the input is anisotropic, such isotropic adjustment of the output units might be too much for the less-stimulated orientation channels and less than desired for the more-stimulated orientation channels. A more ideal normalization mechanism would take into account the relative content in a scene to which a given neuron is sensitive and have only that content weight the output of the unit dynamically based on the strength of that content present at a particular instant or recent past. A cortical model recently proposed by Wainwright, Schwartz, and Simoncelli (2001) is quite similar to this ideal, but makes the dynamic weights of the filters a function of the likelihood of the presence of the natural scene's content that stimulates one filter given the presence of content that stimulates another filter. Specifically, the modeled neural responses are weighted by the conditional probability of image features as specified by the joint conditional histograms constructed from different filter responses to sets of natural scene imagery. That is, instead of the response of units tuned to a given orientation and spatial scale being weighted (i.e., turned down) by the relative activity of units tuned to all orientations and spatial scales, units selective for a particular orientation and spatial scale are weighted more by units tuned to similar orientations and spatial scales. Specifically, the model of Wainwright and colleagues posits that the output response is determined by each linear filter's output being half-wave rectified, squared, and then divided by a normalization signal consisting of the sum of the weighted squared responses from neighboring filters and an additive constant. The weights represent the extent to which the response of one filter is predictive of the response of the other when viewing a typical natural scene. The actual weights in their model are based on observations of conditional probabilities of

simulated neural responses obtained from the statistical properties of natural signals (i.e. natural scene imagery) processed with linear filters resembling the response profile of receptive fields obtained in early visual processing areas (Simoncelli, 1999; Wainwright et al., 2001; Schwartz & Simoncelli, 2001; see also Hansen, et al. 2003 for a full review). The primary implication of their model was that for a given natural scene, at any location containing 'salient' image features (i.e., prominent edges or lines), differently tuned filters will concurrently signal the presence of the same content. Thus, by using the amount of response overlap as a weighting factor to adjust the responses, Wainwright et al. (2001)'s model reduces the transmission of redundant information to successive visual processing areas. Further, this dependency is 'dynamic' in that it is completely driven by the unique structural components (i.e., image statistics) of a given natural image¹. However, and most importantly, the more structural content at or near a particular orientation, the more the neural responses selective for this content will be reduced thereby effectively increasing the response thresholds at that orientation. Such a model speaks directly to the content-dependent effect reported in the current report. Specifically, a bias in natural scene content at a given orientation and scale would drive the sensitivity of cortical units tuned to that orientation and scale down more, thus accounting the psychophysical results described in that section.

A Striate Model of Orientation Processing in Broadband Stimuli

As just described, the results from the current experiments argue that the dynamic normalization of neural responses emphasizes activity in units with neighboring orientations, rather than all orientations contributing to the normalization pool equally. However, such a model does not take into account the inherent horizontal effect bias found to occur in all of the psychophysical experiments carried out in the current study. Since the general horizontal effect was demonstrated to occur with stimuli consisting of natural scenes as well as with broadband visual noise stimuli, it appears to be due to a static anisotropy inherent in the divisive signal. Such a static component could most likely arise directly from the neurons with a horizontal preferred orientation contributing more heavily to the pooled response due to their greater numbers. This numerical bias (a horizontal effect of orientation preferences) was most clearly documented by Li, Peterson, & Freeman, (2003) recently in a survey of about 4,400 neurons, but is also apparent in the data of several other re-

ports (Tiao & Blakemore, 1976; Chapman, Stryker, & Bonhoeffer, 1996 (easily seen in their Figures 1 and 2); Chapman & Bonhoeffer, 1998 (easily seen in their Figures 1 and 2); Coppola, White, Fitzpatrick, & Purves, 1998; Mansfield, 1974; Mansfield & Ronner, 1978). Thus a static weighting factor needs to be added to normalization models such that the divisive pool is influenced by both the dynamic weighting factors described earlier as well as a static anisotropic weighting factor:

$$R_i = \frac{[L_i]^2}{\sum_j ([L_j]^2 o_{ij}) w_{ij} + \sigma_i^2}$$

This present model accounts for the horizontal effect and the content-dependent effect. In this model, adapted from the Wainwright et al. (2001) model, where the response of linear filter i , L_i , is half-wave rectified and squared and then divided by a weighted sum of the squared responses of the other linear filters, L_j , in its respective 'neural neighborhood' weighted by the probability of these responses occurring (w_{ij}) plus an error term (σ_i^2). The proposed static weight (o_{ij}) serves to scale the nearest neighbor responses at various orientations (i.e. L_j) according to the numerical bias of neurons tuned to different orientations.

While the content-dependent effect can likely be explained in terms of the divisive normalization model posited by Wainwright et al (2001), the horizontal effect cannot. Further, it is important to note that the general horizontal effect observed argues for a 'static' weighting factor to be implemented in the proposed normalization model whether the dynamic portion is based on response probabilities as with the Wainwright et al. (2001) model, or on simple filter output to a present/recent image as formulated here (e.g., Heeger, 1992).

The proposed model provides a general account for the two types of effects summarized in the preceding sections (with respect to the functions of striate simple cells – that is, the functions of the striate complex cells are not considered in such a model). However, it only shows how the different weights would be applied to the responses of striate neurons tuned to different ranges of spatial frequencies and orientations. That is, it does not show how the different weights will change as a function of the type of content bias inherent in the different types of natural scene imagery one may encounter on an everyday basis, or how the relative magnitude of inherent (or static) bias in horizontally tuned units will contribute to the general reduction of horizontal sensitivity observed in the results of the current psychophysical experiments. Specifically, if it is indeed the bias in the number of horizontally, and to a lesser extent vertically, tuned striate cells that causes the reduction of horizontal sensitivity to a broad spatial frequency/orientation amplitude increment, then one would expect that if the extent of the increment (in terms of total

¹ This dependency was eliminated when the neural response simulation was carried out of images consisting of white noise (Schwartz & Simoncelli, 2001; Wainwright et al., 2001).

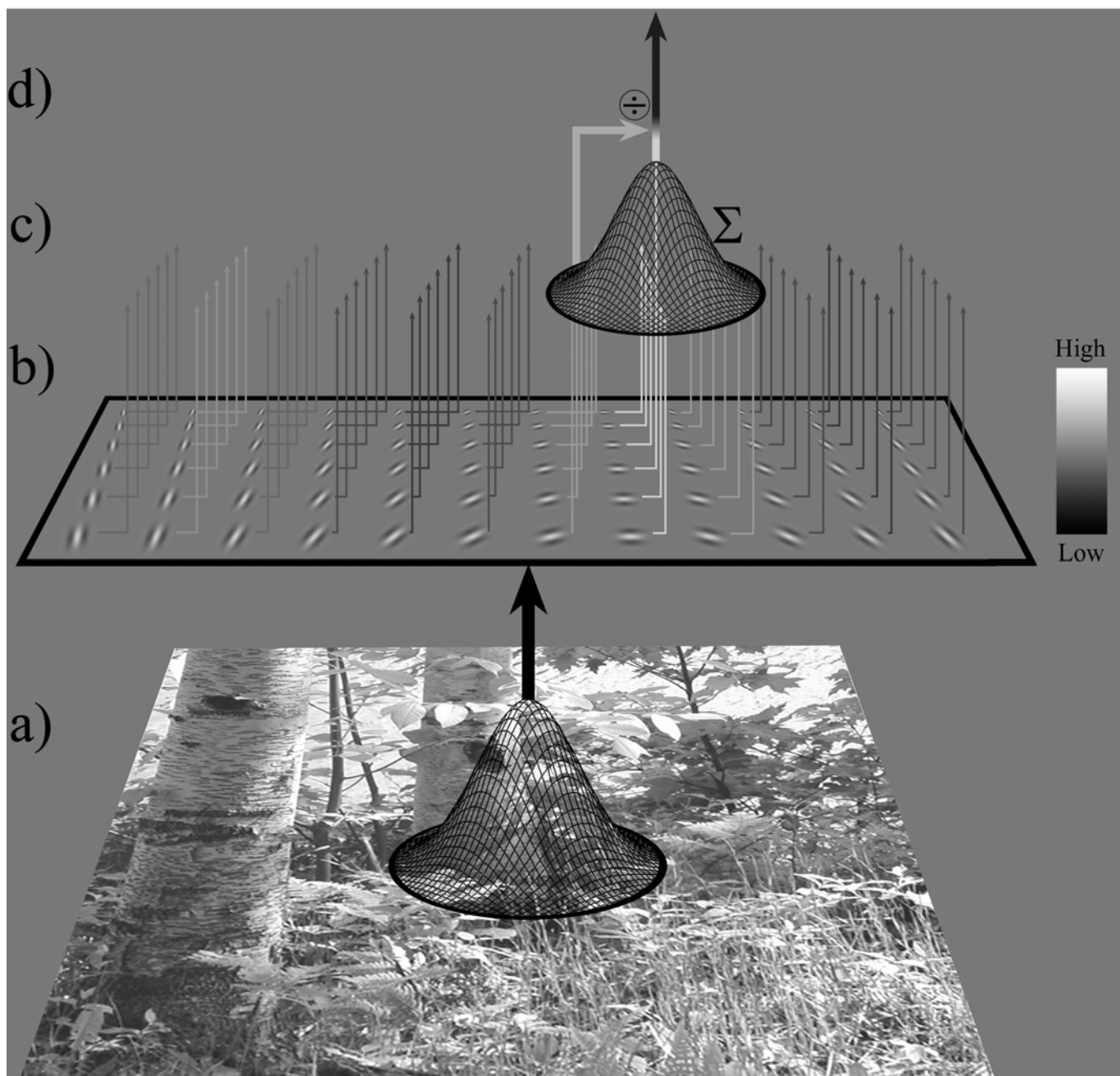


Figure 8. Contrast gain control showing the pooling of filter activity and the divisive signal that adjusts the output of the exemplar perceptual channel shown in this illustration (i.e., a horizontal, middle spatial frequency perceptual channel). Activity elicited by the stimulus image is pooled across nearby spatial locations (shown as a Gaussian-weighted envelope of distance in 'a'), and across neighboring spatial scales and neighboring orientations (shown as summation within a Gaussian-weighted envelope across spatial and orientation difference in 'c'). Note an intrinsic anisotropy reflecting the numerical bias of preferred orientations of neurons (H>V>Oblique) is shown in 'b' (weights indicated by the brightness of lines; higher weights brighter, see scale bar to right) that leads to the horizontal effect (stimuli of otherwise equal content will contribute to the normalization pool more at orientations near horizontal orientations, thereby turning down the gain of an output channel near horizontal more (and secondarily at orientations near vertical). A second orientation effect (see text), is implicit in this model, determined by the predominance of various orientations in the input image ('a'). Magnitude of perceptual channel output of an exemplar channel ('d') is thus diminished based on the amount of activity elicited by the scene at orientations, scales, and locations that are similar to the "label" of the perceptual channel.

number of spatial frequencies and/or orientations incremented) is reduced, the presence of the horizontal effect should also diminish. Thus, it would be useful to know exactly how these weights change in the normalization pool of the proposed model as a function of stimulus content-bias as well as as a function of the extent of the increment in the Fourier domain. In order to show how the inherent weights (i.e., o_{ij}) might change as a function of the extent of the broadband pattern (that is, extent in the Fourier domain), a series of psychophysical experiments need to be carried out. In addition, in order to show how the dynamic weights (i.e., w_{ij}) change as a function of content bias contained in natural scene stimuli, a series of out simulated neural response experiments with the natural scene stimuli used in the current will also need to be carried out. We are currently in the process of carrying out the above mentioned psychophysical and neural response simulation experiments and the results will be reported in a subsequent article.

Currently, we can propose only a general static/dynamic divisive normalization model (shown in Figure 8) where a numeric predominance of horizontally-tuned units, somewhat fewer vertically-tuned units, and fewest obliquely-tuned units (Li, Peterson, & Freeman, 2003), would have the effect of causing anisotropic orientation weights in the normalization pool (horizontal most, obliques least), and consequently creating intrinsically anisotropic gain control. Figure 8 illustrates this model by showing the magnitude of the output of an exemplar orientation/spatial-frequency perceptual channel (d) being reduced divisively (c) by the activity of similar units (b) summed in the normalization pool (c) at nearby orientations and spatial frequencies, all within units localized at nearby spatial locations (a) with the intrinsic anisotropy represented by weighting the output of the linear filter units (b) by their numerical prevalence (weights are indicated by the grayscale shown). Thus, when viewing physically equivalent broad-band stimuli at different orientations, a larger pooled signal will occur at horizontal, and secondarily vertical, than at oblique orientations, and gain for channels reporting stimuli of those orientations will be turned down more, resulting in a lesser ability to see horizontal and vertical stimuli in the presence of adequately broad spatial content.

Furthermore, when viewing typical natural scenes, due to their inherent horizontal-effect anisotropy of image content that we show in Figure 5, a second perceptual horizontal effect may be caused to the extent that the normalization pool of on-going/recent filter activity reflects the content present in the region of the scene. This second effect is apparent in the orientation-biased conditions tested in the present study (Figure 5b) and reflects dynamic suppression of the perceptual output channels (d). We suggest that this suppression is an effect occurring in the orientation dimension comparable to what has been reported previously for the suppressive effect in the spatial frequency dimension of viewing natural scene content

(Webster & Miyahara, 1997). That is, the bias of natural scene content towards more horizontal (and secondarily, vertical) content, as well as the bias of greater amplitude at low spatial frequencies together serve to dynamically alter sensitivity when viewing natural scenes.

Conclusions

We conclude that visual coding of orientation corresponds closely, and inversely, to the content typical of natural scenes. The inverse nature of the effect is most likely due to the divisive nature of contrast gain control. We assume that this anisotropy of contrast gain control stems directly from an apparent numerical bias in preferred orientations of striate cortex neurons. The salience of horizontal content, which is the most prevalent content in typical natural scenes, is turned down, thus serving to discount the perceptual salience of the horizon and other predominant horizontal content. Vertical content is secondarily de-emphasized while oblique content is comparatively enhanced. Thus, objects with broad orientation content would be made relatively more salient when viewed in a typical natural scene. This is an efficient coding strategy, serving to “whiten” the typical natural scene image. Put another way, this mechanism serves to “discount” the anisotropy in natural scenes’ orientation content. In addition to this inherent static anisotropic factor, a dynamic factor adjusts perceptual magnitude based on current/recent scene content that decreases perceptual magnitude of the output channels that correspond to the most prevalent content in a scene (Wainwright et al., 2001). This dynamic gain adjustment will typically reinforce this horizontal effect that is due to the static anisotropy since the content of typical natural scenes is predominated by horizontal, and secondarily, vertical, content.

Acknowledgement

This work was supported by grant # N00014-03-1-0224 from the Office of Naval Research and the Kentucky Space Grant Consortium (KSGC-NASA).

References

- Albrect, D.G. & Geisler, W.S. (1991). Motion selectivity and the contrast-response function of simple cells in the visual cortex. *Visual Neuroscience*, 7, 531-546. [PubMed]
- Albrecht, D.G., Geisler, W.S., Frazor, R.A., & Crane, A.M. (2002). Visual cortex neurons of monkeys and cats: Temporal dynamics of the contrast response function. *Journal of Neurophysiology*, 88, 888-913. [PubMed]

- Annis, R.C. & Frost, B. (1973). Human visual ecology and orientation anisotropies in acuity. *Science*, 182, 729-731. [PubMed]
- Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: The oblique effect in man and animals. *Psychological Bulletin*, 78, 266-278. [PubMed]
- Atick, J.J. & Redlich, A.N. (1992). What does the retina know about natural scenes? *Neural Computation*, 4, 196-210. [PubMed]
- Baddeley, R.J. & Hancock, P.J.B. (1991). A statistical analysis of natural images matches psychophysically derived orientation tuning curves. *Proceedings of the Royal Society of London B*, 246, 219-223. [PubMed]
- Barlow, H.B. (1959). Sensory mechanisms, the reduction of redundancy, and intelligence. *Proceedings of the National Physical Laboratory Symposium*, (pp. 537-559). London: H.M. Stationary Office.
- Bonds, A.B. (1989). Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. *Visual Neuroscience*, 2, 41-55. [PubMed]
- Bonds, A.B. (1991). Temporal dynamics of contrast gain in single cells of the cat striate cortex. *Visual Neuroscience*, 6, 239-255. [PubMed]
- Brady, N. & Field, D.J. (1995). What's constant in contrast constancy? The effects of scaling on the perceived contrast of bandpass patterns. *Vision Research*, 35, 739-756. [PubMed]
- Brady, N. & Field, D.J. (2000). Local contrast in natural images: Normalization and coding efficiency. *Perception*, 29, 1041-1055. [PubMed]
- Campbell, F.W., Kulikowski, J.J., & Levinson, J. (1966). The effect of orientation on the visual resolution of gratings. *Journal of Physiology*, 187, 437-45. [PubMed]
- Carandini, M. & Heeger, D.J. (1994). Summation and division in primate visual cortex. *Science*, 264, 1333-1336. [PubMed]
- Carandini, M., Heeger, D.J., & Movshon, J.A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *Journal of Neuroscience*, 17, 8621-8644. [PubMed]
- Chapman, B. and Bonhoeffer, T. (1998). Overrepresentation of horizontal and vertical orientation preferences in developing ferret area 17. *Neurobiology*, 95, 2609-2614. [PubMed]
- Chapman, B., Stryker, M.P., & Bonhoeffer, T. (1996). Development of orientation preference maps in ferret primary visual cortex. *The Journal of Neuroscience*, 16, 6443-6453. [PubMed]
- Creelman, C.D. & Macmillan, N.A. (1991). Detection theory: A user's guide, Cambridge University Press, Cambridge.
- Coppola, D.M., Purves, H.R., McCoy, A.N., & Purves, D. (1998). The distribution of oriented contours in the real world. *Proceedings of the National Academy of Sciences USA*, 95, 4002-4006. [PubMed]
- Coppola, D.M., White, L.E., Fitzpatrick, D., & Purves, D. (1998). Unequal representation of cardinal and oblique contours in ferret visual cortex. *Proceedings of the National Academy of Sciences USA*, 95, 2621-2623. [PubMed]
- Deriugin, N.G. (1956). The power spectrum and the correlation function of the television signal. *Telecommunications*, 1, 1-12.
- De Valois, R.L., Yund, E.W., & Hepler, N. (1982). The orientation and direction selectivity of cells in macaque visual cortex. *Vision Research*, 22, 531-544. [PubMed]
- Essock, E.A. (1980). The oblique effect of stimulus identification considered with respect to two classes of oblique effects. *Perception*, 9, 37-46. [PubMed]
- Essock, EA, Krebs, WK and Prather, JR (1992). An Anisotropy of human tactile sensitivity and its relation to the visual oblique effect. *Experimental Brain Research*, 91, 520-524. [PubMed]
- Essock, E.A., DeFord, J.K., Hansen, B.C., & Sinai, M.J. (2003). Oblique stimuli are seen best (not worst!) in naturalistic broad-band stimuli: A horizontal effect. *Vision Research*, 43, 1329-1335. [PubMed]
- Field, D.J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, 4, 2379-2394. [PubMed]
- Furmanski, C.S. & Engel, S.A. (2000). An oblique effect in human primary visual cortex. *Nature Neuroscience*, 3, 535-536. [PubMed]
- Geisler, W.S. & Albrecht, D.G. (1992). Cortical-neuron isolation of contrast gain control. *Vision Research*, 32, 1409-1410. [PubMed]
- Hancock, P.J.B., Baddeley, R.J., & Smith, L.S. (1992). The principle components of natural images. *Network: Computation in Neural Systems*, 3, 61-70.
- Hansen, B.C., Essock, E.A., Zheng, Y., & DeFord, J.K. (2003). Perceptual anisotropies in visual processing and their relation to natural image statistics. *Network: Computation in Neural Systems*, 14, 501-526. [PubMed]
- Hansen, B.C. & Essock, E.A. (2004). Influence of scale and orientation on the visual perception of natural scenes. *Visual Cognition* (In Press).

- Haug, J. & Mumford, D. (1999). Statistics of natural images and models. *Proceedings of the ICCV*, 1, 541-547.
- Heeger, D.J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience*, 9, 181-197. [[PubMed](#)]
- Keil, M.S. & Cristóbal, G. (2000). Separating the chaff from the wheat: Possible origins of the oblique effect. *Journal of the Optical Society of America A*, 17, 697-710. [[PubMed](#)]
- Kennedy, H., Martin, K.A.C., Orban, G.A., & Whitteridge, D. (1985). Receptive field properties of neurons in visual area 1 and visual area 2 in the baboon. *Neuroscience*, 14, 405-415. [[PubMed](#)]
- Knill, D.C., Field, D., & Kersten, D. (1990). Human discrimination of fractal images. *Journal of the Optical Society of America A*, 7, 1113-1123. [[PubMed](#)]
- Kretzmer, E.R. (1952). Statistics of television signals. *Bell System Technologies Journal*, 31, 751-763.
- Li, B., Peterson, M.R., & Freeman, R.D. (2003). Oblique effect: A neural bias in the visual cortex. *Journal of Neurophysiology*, 90, 204-217. [[PubMed](#)]
- Maffei, L; Campbell, F. W. (1970). Neurophysiological localization of the vertical and horizontal visual coordinates in man. *Science*, 167, 386-387. [[PubMed](#)]
- Mansfield, R.J.W. (1974). Neural basis of orientation perception in primate vision. *Science*, 186, 1133-1135. [[PubMed](#)]
- Mansfield, R.J.W. & Ronner, S.P. (1978). Orientation anisotropy in monkey visual cortex. *Brain Research*, 149, 229-234. [[PubMed](#)]
- Mitchell, D.E., Freeman, R.D., & Westheimer, G. (1967). Effect of orientation on the modulation sensitivity for interference fringes on the retina. *Journal of the Optical Society of America* 57, 246-249. [[PubMed](#)]
- Olshausen, B.A. & Field, D.J. (2000). Vision and the coding of natural images. *American Scientist*, 88, 238-245.
- Orban, G.A. & Kennedy, H. (1980). Evidence for meridional anisotropies in orientation selectivity of visual cortical neurons. *Archives Internationales de Physiologie et de Biochimie*, 88, 13-14. [[PubMed](#)]
- Párraga, C.A. & Tolhurst, D.J. (2000). The effect of contrast randomization on the discrimination of changes in the slopes of the amplitude spectra of natural scenes. *Perception*, 29, 1101-1116. [[PubMed](#)]
- Párraga, C.A., Troscianko, T., & Tolhurst, D.J. (2002). Spatiochromatic properties of natural images and human vision. *Current Biology*, 12, 483-487. [[PubMed](#)]
- Yu, H-B. & Shou, T-D. (2000). The oblique effect revealed by optical imaging in primary visual cortex of cats. *Acta Physiologica Sinica*, 52, 431-434. [[PubMed](#)]
- Schwartz, O. & Simoncelli, E.P. (2001). Natural signal statistics and sensory gain control *Nature Neuroscience*, 4, 819-825. [[PubMed](#)]
- Simoncelli, E.P., Freeman, W.T., Adelson, E.H., & Heeger, D.J. (1992) Shiftable multi-scale transforms. *IEEE Transactions in Information Theory*, 38, 587-607.
- Simoncelli, E.P. (1999). Modeling the joint statistics of images in the wavelet domain. *Proceedings of the SPIE*, 3813, 188-195.
- Simoncelli, E.P. & Olshausen, B.A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, 24, 1193-1216. [[PubMed](#)]
- Switkes, E., Mayer, M.J., & Sloan, J.A. (1978). Spatial frequency analysis of the visual environment: Anisotropy and the carpentered environment hypothesis. *Vision Research*, 18, 1393-1399. [[PubMed](#)]
- Taylor, D.R., Finkel, L.H., & Buchsbaum, G. (2000). Color-opponent receptive fields derived from independent component analysis of natural scenes. *Vision Research*, 40, 2671-2676. [[PubMed](#)]
- Tiao, Y-C. & Blakemore, C. (1976). Functional organization in the visual cortex of the golden hamster. *Journal of Comparative Neurology*, 168, 459-482. [[PubMed](#)]
- Timney, B.N. & Muir, D.W. (1976). Orientation anisotropy: Incidence and magnitude in caucasian and chinese subjects. *Science*, 193, 699-701. [[PubMed](#)]
- Tolhurst, D.J. & Tadmor, Y. (1997). Band-limited contrast in natural images explains the detectability of changes in the amplitude spectra. *Vision Research*, 37, 3203-3215. [[PubMed](#)]
- Tolhurst, D.J., Tadmor, Y., & Chao, T. (1992). Amplitude spectra of natural images. *Ophthalmic and Physiological Optics*, 12, 229-232. [[PubMed](#)]
- Tolhurst, D.J. & Tadmor, Y. (2000). Discrimination of spectrally blended natural images: Optimization of the human visual system for encoding natural images. *Perception*, 29, 1087-1100. [[PubMed](#)]
- van der Schaaf, A. & van Hateren, J.H. (1996). Modeling the power spectra of natural images: Statistics and Information. *Vision Research*, 36, 2759-2770. [[PubMed](#)]
- van Hateren, J.H. & van der Schaaf, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal society of London:B*, 265, 359-366. [[PubMed](#)]

- Wainwright, M.J., Schwartz, O. & Simoncelli, E.P. (2001). Natural image statistics and divisive normalization: modeling nonlinearities and adaptation in cortical neurons. In R. Rao, B. Olshausen and M. Lewicki (Eds.) *Probabilistic Models of the Brain: Perception and Neural Function*, Cambridge, MA: MIT Press.
- Webster, M.A. & Miyahara, E. (1997). Contrast adaptation on the spatial structure of natural images. *Journal of the Optical Society of America, A*, 14, 2355-2366. [[PubMed](#)]
- Webster, M.A. & Mollon, J.D. (1997). Adaptation and the color statistics of natural images. *Vision Research*, 37, 3283-3298. [[PubMed](#)]
- Wilson, H.R. & Humanski, R. (1993). Spatial frequency adaptation and contrast gain control. *Vision Research*, 33, 1133-1149. [[PubMed](#)]
- Zemon, V., Gutowski, W., & Horton, T. (1983). Orientational anisotropy in the human visual system: An evoked potential and psychophysical study. *International Journal of Neuroscience*, 19, 259-286. [[PubMed](#)]