Is color patchy?

Ali Yoonessi,* Frederick A. A. Kingdom, and Samih Alqawlaq

Department of Ophthalmology, McGill Vision Research Unit, 687 Pine Avenue W. Room H4-14, Montréal, Québec, H3A 1A1, Canada
**Corresponding author: ali.yoonessi@mcgill.ca*

Received January 17, 2008; revised March 25, 2008; accepted March 28, 2008;
posted April 7, 2008 (Doc. ID 91859); published May 16, 2008

In many natural scenes, shadows and shading, which are primarily luminance-defined features, proliferate. Hence one might expect that the chromatic variations of natural scenes, which more faithfully represent the layout of object surfaces, will contain relatively fewer and larger uniform regions than the luminance variations, i.e., will be more "patchy." This idea was tested using images of natural scenes that were decomposed into chromatic and luminance layers modeled as the responses of the red–green, blue–yellow, and luminance channels of the human visual system. Patchiness was defined as the portion of pixels falling within a ± threshold in the bandpass-filtered image, averaged across multiple filter scales. The red–green layers were found to be the most patchy, followed by the blue–yellow layers, with the luminance layers the least patchy. The correlation between image-layer patchiness and the slope of the Fourier amplitude spectrum was small and negative for all layers, the maximum value being for red–green (−0.48). We conclude that the chromatic layers of natural scenes contain more uniform areas than the luminance layers and that this is unpredicted by the slope of the Fourier amplitude spectrum.

OCIS codes: 330.0330, 330.1720.

1. INTRODUCTION

Knowledge of the physical structure of the natural visual environment guides our study of human vision [1]. It has been known for some time that the spatial structure of the chromatic and luminance content of natural scenes is different [2] and more recently that the differences have manifest physiological [3] and behavioral consequences [4]. In this communication we consider whether the chromatic and luminance layers of natural scenes differ in the degree to which they form regions with relatively uniform chromaticities or luminances. In other words, we consider whether the color and luminance layers of natural scenes differ in the degree to which they are structurally sparse or "patchy." We use the term "patchiness" to describe the proportion of the image that contains uniform areas—the higher the patchiness, the larger and/or more numerous the uniform areas. An alternative term that emphasizes the proportion of pixels in the image comprising edges rather than uniform areas is "structural sparseness" [5]. Either term would suffice here, but we prefer patchiness, as it best captures the idea of the amount of uniform regions.

A much-vaunted characteristic of natural scenes is the shape of their Fourier amplitude spectra. In natural scenes, Fourier amplitude is inversely proportional to spatial frequency f, or, more precisely, proportional to $1/f^n$, where the exponent n, which defines the (negative) slope of the spectrum, varies slightly between scenes but is on average close to unity [6,7]. Important for the present discussion is that the exponent n has been shown to be similar for the chromatic and luminance layers of natural scenes [8,9]. The $1/f$ characteristic embodies the fact that energy is relatively strong at low spatial frequencies, consistent with the presence of relatively large aggregates of pixels with similar luminances/chromaticities. The slightly steeper slopes in the amplitude spectra of some scenes ($n>1$) has been explained by the presence of particularly large aggregates of similar pixel values, for example, from a large evenly colored object [9]. The finding that the spectral slopes of the luminance and chromatic content of natural scenes are similar might lead one to suppose that the luminance and chromatic layers are equally patchy. However, the presence of uniform regions bounded by edges, a defining feature of natural scenes, is primarily a higher-order structural property. The simplest evidence for this is that phase scrambling a natural scene, which has no affect on the amplitude spectrum, destroys any uniform regions. Thus a similarity in spectral slope is not necessarily an indicator of a similarity in patchiness.

Indeed, other considerations lead one to suppose that the chromatic and luminance layers of natural scenes might differ in patchiness. Chromatic differences tend to arise from differences in material composition, such as from differently colored objects and surfaces. Luminance differences, on the other hand, arise not only from material differences, i.e., from different object/surface reflectances, but from inhomogenous illumination, such as shadows and shading (sometimes termed "attached" shadows). Because shadows and shading are primarily luminance-defined features, and because they tend to proliferate in natural scenes [2,10,11], it might be expected that the chromatic layers of natural scenes would be more patchy than their luminance layers. The aim of this communication is to test this prediction.

In daylight vision the image is first transduced by three photoreceptors: the long- (L-), medium- (M-), and short- (S-) wavelength-sensitive cones. The cone responses are

then combined into three channels, one luminance sensitive and two chromatically sensitive. The luminance-sensitive channel sums the outputs of the L and M cones and is known as the L+M channel. One chromatically sensitive channel takes the difference between the outputs of the L and M cones and is known as the L−M channel, while the other chromatically sensitive channel takes the difference between the sum of the L- and M-cone responses and the S-cone response and is known as the S −(L+M) channel. Because these channels are very often (though strictly speaking incorrectly) referred to as the luminance, red–green, and blue–yellow channels, we will adopt these terms from now on.

In this communication the patchiness indices of the modeled responses of the three postreceptoral channels to natural scenes are compared. The reason for basing the comparison on the modeled postreceptoral channel responses is that some previous studies have noted that the red–green and blue–yellow layers of natural scenes, while in the main insensitive to shadows and shading, are nevertheless differentially sensitive to them, with the blue–yellow channel the more sensitive [9,12]. Scenes with green foliage are full of shadows, and these seem to shift the wavelength spectrum slightly toward the yellow end of the blue–yellow channel continuum [12]. On the other hand, the deep shadows seen in open-lit sunny days tend to be bluish, because whereas the areas outside the shadows are bathed in a mixture of yellow sunlight and blue skylight, the shadows tend to be bathed mainly in blue skylight [13]. These considerations lead to the testable prediction that the blue–yellow channel response to natural scenes will be less patchy than the red–green channel response but more patchy than the luminance channel response.

2. MEASUREMENT OF PATCHINESS

A higher-order statistic that has been suggested to capture natural scene patchiness is the kurtosis of the pixel histogram [14,15]. Kurtosis is a measure of the peakedness of a distribution and is defined as μ^4/σ^4, where μ is the mean pixel value and σ the standard deviation of pixel values. Hence a possible measure of patchiness would be the kurtoses of the raw pixel histograms of the red–green, blue–yellow, and luminance channel responses. However, there are problems with kurtosis as a measure of patchiness. Suppose one has a scene that contains, among other things, a large uniform red region and a large uniform green region, i.e., is relatively patchy along the red–green dimension. This would produce peaks in the red–green pixel histogram on either side of the mean and a low value of kurtosis. Figure 1 shows three images, their red–green layers (rendered as gray-level images), and the histograms and kurtoses of the layers. Two of the three red–green layers have very low kurtosis, despite having large uniform areas of flowers and leaves. The intuitive solution to this problem is to measure kurtosis not in the raw but in the bandpass-filtered version of the image [15], as this would shift any peaks in the histogram arising from uniform areas onto the mean, hence revealing any high patchiness via high kurtosis. In our sample images we found that the kurtosis of the filtered image is a reasonable measure of patchiness, but still not the best. Why?

Kurtosis is very sensitive to noise, in that the presence of low-amplitude noise around the mean reduces kurtosis. Our images are contaminated by noise from the photographic process. Photographic noise is Gaussian, with an amplitude that increases with the brightness of the scene [16]. Although the signal and noise levels may be similar for the R (red), G (green), and B (blue) camera responses, this may not be the case for the three postreceptoral channel images derived from them. Remember that the luminance channel sums cone responses, the red–green channel subtracts cone responses, and the blue–yellow channel both sums and subtracts cone responses. While the cone signals will add or subtract, the noise amplitude levels will not. The amplitude of the noise in the postreceptoral layers will depend not only on whether the cone images are added or subtracted but also on the degree to which the noise is correlated between cone images. For example, suppose the noise in the L and M images was 100% correlated. Subtracting the M- from the L-cone image to produce the red–green channel image would eliminate the noise, whereas adding the cone images to produce the luminance channel image would double the noise amplitude. On the other hand, if the noise was completely uncorrelated, noise amplitude would increase by a factor of $\sqrt{2}$, irrespective of whether the L- and M-cone images were added or subtracted. Using our method, the problem is even more complicated. The color space we have employed, for reasons given later, is based on a log (logarithmic) transform of cone signals. The transform subtracts the average log value of cone signals across the whole image from each log pixel cone value in order to center the cone signals around a midpoint of zero [17]. In an image with a wide signal variation, a log transform can decrease the signal-to-noise ratio.

In our set of natural images the ratios of the average absolute magnitudes of the luminance, red–green, and blue–yellow signals were 1.0, 0.11, and 0.88 respectively. (These values are measured in the logarithmic color space described later and are unitless.) Taken together with the argument above that the signal-to-noise ratios resulting from photographic noise are dependent on both the type of layer and the scene content, it is likely that the signal-to-noise ratios will vary considerably among the three postreceptoral layers and render kurtosis a bad estimator of patchiness. To examine this possibility, the following simulation was performed. A test image consisting of two halves separated by an edge was allocated red–green, blue–yellow, and luminance contrasts with relative magnitudes equal to those given above for the natural image set. The image was then converted into RGB (see Section 3 for details), and uncorrelated Gaussian noise, with σ (standard deviation) multiples of $1\,(x=1,2,4,8)$ was added to each plane $[N(0,\sigma^2)$ where N is the normal distribution, 0 is the mean, and $\sigma=1,2,4,8]$. The RGB images were then converted back to the three postreceptoral layers. The images were then filtered with a log-Gabor filter (see below) with a spatial frequency of 12 cycles per degree (cpd) and kurtosis was measured for each layer. With no noise added, the kurtosis was the same for all three layers: 40.34. The noise-added results are shown in

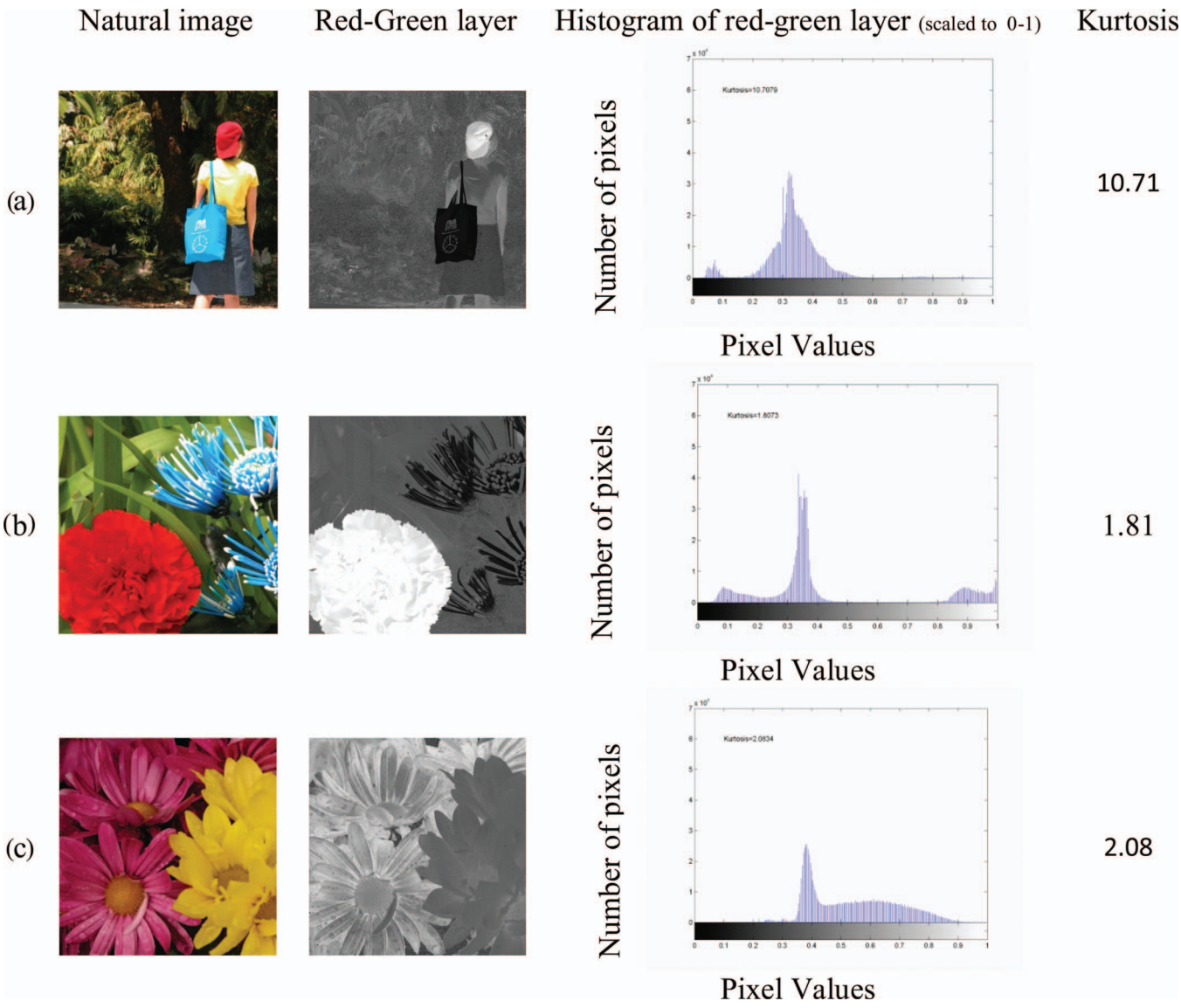


Fig. 1. Measurement of kurtosis in the red–green layer of three natural scenes with different perceived patchiness (see text for details).

Fig. 2. Each point represents the average of five random noise samples. As the figure shows, kurtosis falls with noise σ as expected, but more important, at a different rate for the three layers (blue–yellow >red–green >luminance). Therefore, assuming the simulation reflects what happens with natural scene images, patchiness as measured by kurtosis, even in bandpass-filtered images, will differ among the three postreceptoral channel layers solely because of the differential effects of photographic noise. This suggests that kurtosis of the filtered image is not a valid measure of patchiness when using different postreceptoral channel images of natural scenes.

To minimize the impact of photographic noise on the estimate of patchiness, we employed a modified version of the structural sparseness index developed by Hansen and Hess [5]. In their method, the images were first filtered using log-Gabor filters in 3 spatial frequencies and 12 orientations. The portion of pixels that fell within $\pm 1.5\,\sigma$ from the mean of each filtered image was calculated, and the average value across filter scales was the index of sparseness. The important difference in our method is that we used a fixed threshold rather than one defined in terms of the standard deviation (we also did not filter across orientations and used four rather than three filter scales—see below). A fixed threshold was employed for two reasons. First, a fixed threshold is more robust to noise. To demonstrate this, a similar simulation was performed to the one described above. This time 30 bipartite test stimuli were generated, each with red–green, blue–yellow, and luminance contrasts randomly selected from zero to the average values determined from the natural image set. As in the previous simulation, the images were converted into RGB, and normally distributed noise added to each plane. This time the noise σ was set to 10% of the signal strength of each RGB plane. The images were then converted back into postreceptoral channel images and filtered as described above. Patchiness was measured according to three indices; threshold based on $1.5\,\sigma$ (as in Hansen and Hess [5]), a fixed threshold (±0.07; the range of values in the filtered images varies between ±0.26 and ±1.12), and kurtosis. The formal statement of the threshold based on the standard deviation is

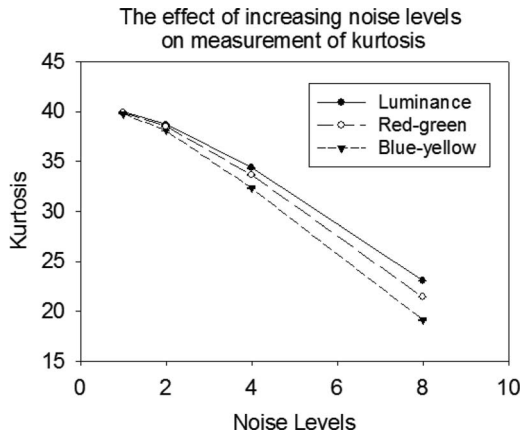


Fig. 2. Effect of noise in the RGB images on the measurement of kurtosis in the three postreceptoral channel responses to a bipartite field. Each point represents the average kurtosis for five sample images with normally distributed noise.

$$F_T(x_i,y_i) = \begin{cases} 1 & F_R(x_i,y_i) < +1.5^*\sigma | F_R(x_i,y_i) > -1.5^*\sigma \\ 0 & \text{elsewhere} \end{cases},$$

where $F_T(x,y)$ is the thresholded filter response, $F_R(x,y)$ is the filtered response, and σ is the standard deviation of the pixel values in $F_R(x,y)$. For the fixed threshold, it is

$$F_T(x_i,y_i) = \begin{cases} 1 & F_R(x_i,y_i) < +0.07 | F_R(x_i,y_i) > -0.07 \\ 0 & \text{elsewhere} \end{cases},$$

where the parameters are the same as above except that the threshold is fixed. Kurtosis has been described earlier.

Because we used a filter that was range independent, the noise-free images produced the same value of patchiness for each of the 30×3 image layers. These values were 0.95, 0.94, and 40.34, respectively, for the three indices of the fixed threshold, standard-deviation threshold, and kurtosis. For the noise-added images, patchiness varied both as a function of noise level and postreceptoral layer, as shown in Table 1. A valid index of patchiness is evidenced by a measure that is similar for all three layers, and Table 1 shows this is the case for both the standard-deviation-based and fixed-threshold indices, but not for kurtosis. A robust measure of patchiness is evidenced by a relatively low variance of the measure across noise levels. We measured the standard deviation of each patchiness index across 30 images. The relative standard deviations of patchiness were 300, 3, and 1, for the kurtosis, standard-deviation-based threshold, and fixed-threshold indices, respectively. Hence the fixed-threshold index appears to be the most robust to noise.

A potential problem with the use of a fixed threshold rather than one based on the standard deviation is the different absolute magnitudes of the responses in the red–green, blue–yellow, and luminance layers (see above), which would normally result in the lowest estimate of patchiness for the red–green layer. However, the log-Gabor filters used here are range independent; that is, the range of the input has no effect on the range of the filter response. Instead, it is the relations among pixels that determine the filter response. Put another way, if we were to multiply an image by any number, the log-Gabor filter response would be unchanged. This is why the measures of patchiness for the three layers in both simulations described above are nearly identical. Log-Gabor rather than Gabor filters were also employed because they reset the dc level of the image to zero. Finally, it has been suggested that the log-Gabor filter better represents the responses of cortical cells than do Gabor filters [18,19].

To estimate patchiness we combined individual estimates of patchiness across four spatial scales. Why? Fine-scale filters will signal the presence of both small and large uniform areas, while coarse-scale filters will signal only relatively large uniform areas. So one could argue from this that only the finest-scale filters should be used. However, the coarse filters are able to detect coarse-scale edges, such as from shading and shadows, which would be encoded as uniform areas by the finest-scale filter. Therefore averaging across scales gives a better indication of the overall patchiness of an image.

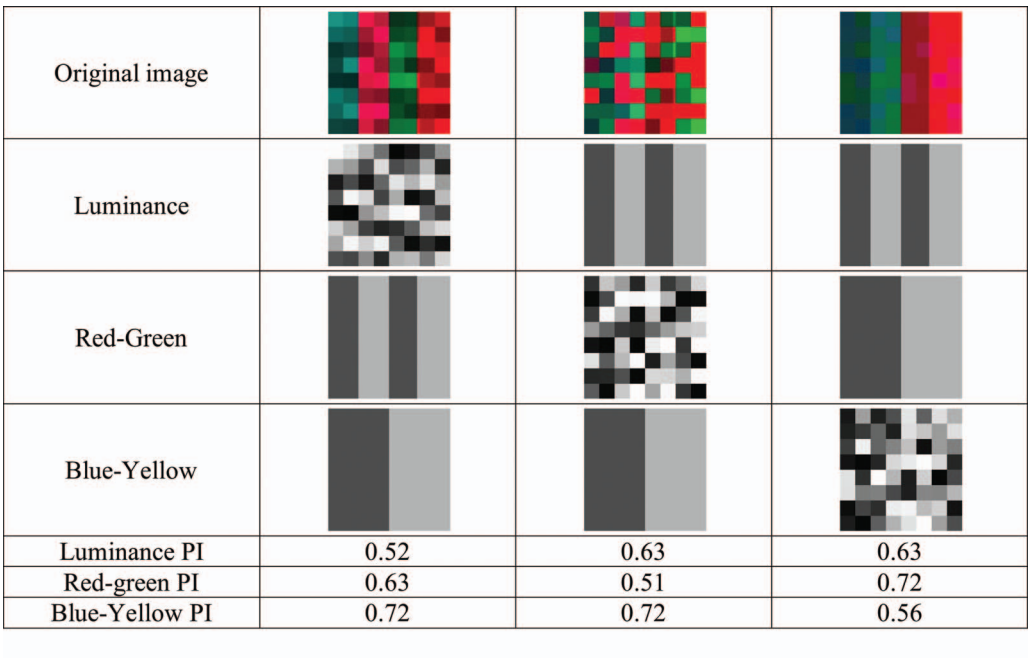
Figure 3 shows the resulting patchiness indices for synthetic stimuli. The three postreceptoral layers have different patch sizes, and this is reflected in their different measures of patchiness, verifying that the patchiness index is directly correlated with the portion of image-layer pixels that form uniform areas.

Figure 4 illustrates the sequence of stages involved in our estimation of natural scene patchiness. The images were first decomposed into luminance, red–green, and blue–yellow images. Each layer was then filtered with isotropic log-Gabor filters at four spatial scales. The filtered images were then converted to binary images using an arbitrary fixed threshold set to ±0.07 from the mean. The portion of pixels falling within the upper and lower

Table 1. Comparison of Three Methods, Kurtosis, Standard-Deviation Threshold and Fixed Threshold for Estimating Patchiness, Using 30 Artificial Images with Added Noise[a]

	Kurtosis			Standard-Deviation Threshold			Fixed Threshold		
Measure	Luminance	Red–Green	Blue–Yellow	Luminance	Red–Green	Blue–Yellow	Luminance	Red–Green	Blue–Yellow
Mean	6.01	4.66	4.46	0.89	0.88	0.88	0.97	0.98	0.98
SD	3.44	2.24	1.72	0.02	0.02	0.02	0.006	0.006	0.007
Mean across layers		4.98			0.87			0.97	
SD across layers		2.64			0.020			0.009	

[a]SD, standard deviation.



Luminance PI	0.52	0.63	0.63
Red-green PI	0.63	0.51	0.72
Blue-Yellow PI	0.72	0.72	0.56

Fig. 3. Synthetic images and resulting measures of patchiness. PI, patchiness index.

threshold bounds was measured, and the index of patchiness was calculated as the average of these values across the four filter scales. Finally, as mentioned earlier, since previous research has suggested that the slope of the Fourier amplitude spectrum may be correlated with the amount of uniform areas in the image [9], we correlated the estimates of patchiness with the measured slopes of the images' Fourier amplitude spectra.

3. METHOD

A. Equipment and Calibration

The scenes were photographed with two Nikon CoolPix-7500 digital cameras. Matlab version 7 was used for all image processing tasks. The cameras were calibrated to remove gamma nonlinearity as follows. Each one of a set of gray Munsell papers was illuminated by an incandescent light with a constant-dc power and photographed. Additionally, the luminance of the light reflected from each paper was measured with a Topcon SR-1 spectroradiometer. The average R, G, and B pixel values were plotted against the corresponding measured luminance, and fitted with the following function: $L=a(b^s+1)$, where L is luminance, s is the pixel level value obtained for each of the camera sensors (R, G, and B), and b is a constant that determines the slope of the curve.

To measure the spectral sensitivity functions for the three R, G, and B camera sensors, a white target was photographed through a series of narrowband optical interference filters from 400 to 700 nm at 10 nm intervals. Each R, G, and B value was recorded, gamma corrected, and used to construct a spectral sensitivity function for each sensor, which was then normalized to produce equal responses of a flat-spectrum light.

B. Images

From the McGill Calibrated Color Image Database [20], we took 758 natural scenes, representing a range of natural environments (forests, mountains, flowers, fruits, snow, and textures) and urban scenes (buildings, traffic signs, and manmade objects), photographed under a variety of different illumination conditions (sunny and cloudy) and at a variety of distances (0.5 m–1000 m). The images were photographed by the camera (see above) and stored in uncompressed Tagged Image File Format (TIFF) files with a resolution of 1920×2560 pixels and a color depth of 24 bits (256 levels for each R, G, and B image). The camera's smallest aperture setting (f7.4) was chosen to capture the images with minimum within-image differences in focus. Then the images were resized and cropped to 960×960 pixels wide using the nearest-neighbor interpolation algorithm.

C. Conversion of Stimuli from RGB to LMS Color Space

Using the spectral sensitivities of the camera sensors and the sensitivities of the L, M, and S cones from Smith and Pokorny [21], a conventional 3×3 linear matrix was used to convert the RGB camera values to LMS cone excitations [20]. The matrices for conversion from RGB to LMS and vice versa were taken from the database website [20].

D. Color Space and Postreceptoral Channels

A modified version of the Ruderman color space was used to model the three postreceptoral channels of human vision [17]. Cone contrasts for each pixel were defined as

$$L_C = \log L - \overline{\log L}, \quad M_C = \log M - \overline{\log M},$$

$$S_C = \log S - \overline{\log S}$$

where $\log L$, $\log M$, and $\log S$ are log pixel cone excitations and the same notations with overbars are log pixel cone excitations averaged across the image. The three postreceptoral responses to each pixel were defined as

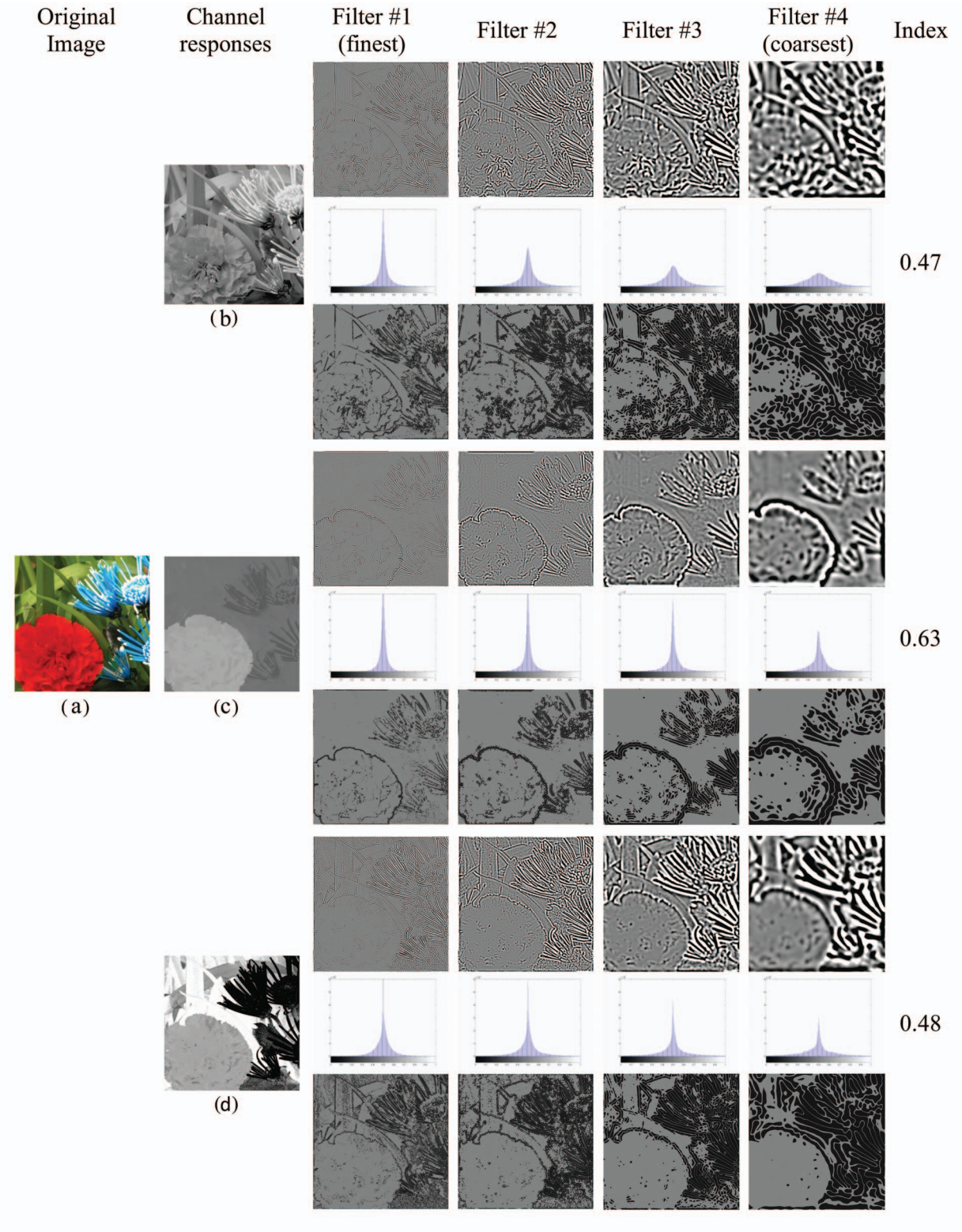


Fig. 4. Method for measuring patchiness. A sample image (a) with its luminance (b), red–green (c), and blue–yellow (d) channel responses. Below each filtered image is shown the pixel histogram and below each histogram is shown the binarized image after applying the fixed threshold. The index of patchiness on the right gives the average portion of light gray pixels in the four binarized images.

$$\hat{l} = (\hat{L}_C + \hat{M}_C), \quad \hat{\alpha} = (\hat{L}_C + \hat{M}_C - 2\hat{S}_C), \quad \hat{\beta} = (\hat{L}_C - \hat{c}_C),$$

where $\hat{l}$, $\hat{\alpha}$, and $\hat{\beta}$ are the luminance, blue–yellow, and red–green axes, respectively.

E. Image Processing

The log-Gabor filter in the Fourier domain is defined as

$$LG(f,\theta) = \exp\left\{-\left[\frac{\log\left(\frac{R(f_i,\theta_j)}{F_{peak}}\right)^2}{2\log\left(\frac{\sigma_1}{F_{peak}}\right)^2}\right]\right\},$$

where f and θ are the polar coordinates of any point, R is the radius vector, F_{peak} is the central spatial frequency of the log-Gaussian function, and σ_1 is the spatial frequency bandwidth of the log-Gaussian function.

Four spatial frequencies (3, 6, 12, and 24 cpd at the viewing distance of 1 m) were used (Fig. 4). The threshold of ±0.07 (the range of the images varied from ±0.26 to ±1.12), although somewhat arbitrary, was chosen so that patchiness ranged between 10% and 90% for all images. This ensured that any differences in patchiness between images and layers would not be masked by floor and ceiling effects. The portion of pixels inside the ± threshold was measured. These portions were averaged over the four spatial scales to give the estimate of patchiness, as shown in Fig. 4. The spectral slope was measured using a two-dimensional fast Fourier transform. After removing the dc component, the slope was measured using linear regression from the lowest spatial frequency (1 cpd) to the Nyquist frequency (wavelength of 2 pixels) on a log–log scale.

4. RESULTS

The mean values of patchiness for the luminance, red–green, and blue–yellow layers across all images were 0.49±0.13, 0.71±0.08, and 0.65±0.13, respectively. The averages at each filter scale are shown in Table 2, which shows that the patchiness index declines with filter size (see also Fig. 4). Two-tailed paired t-tests tested for significant differences among the three layers. After Bonferroni correcting the p values to allow for multiple comparisons (by increasing the calculated p values threefold), the difference between all pairs of layers was highly significant: luminance versus red–green, $t(1514)=12.07$; luminance versus blue–yellow, $t(1514)=8.98$; red–green versus blue–yellow, $t(1514)=4.64$; all p values <0.0001.

The images were then categorized into animals, flowers, foliage, fruits, landscape, manmade objects, shadows, snow, and textures. Table 3 and Fig. 5 show that in all categories patchiness was significantly different between layers (again after Bonferroni correction). The most significant differences among the three layers were found in the foliage category, and the least significant differences in the fruit category. The smallest difference in mean patchiness between the luminance and the red–green layers was found for landscape scenes, and the highest difference for textures and foliage. In manmade and shadow scenes, the red–green and blue–yellow patchiness indices were similar.

Table 2. Patchiness Index at Different Filter Scales

Layer	Filter #1 (Finest)	Filter #2	Filter #3	Filter #4 (Coarsest)	Average
Luminance	0.58	0.51	0.46	0.41	0.489
Red–green	0.73	0.72	0.72	0.70	0.717
Blue–yellow	0.66	0.66	0.65	0.64	0.653

Table 3. Patchiness Index in Different Categories Sorted by the Patchiness Index in Luminance[a]

	Luminance		Red–Green		Blue–yellow	
Category (No. of Images)	PI	Slope	PI	Slope	PI	Slope
Textures (41)	0.41	−0.76	0.75	−0.97	0.71	−1.03
Foliage (238)	0.44	−0.71	0.70	−0.82	0.58	−0.87
Animals (29)	0.49	−0.78	0.77	−0.89	0.68	−0.97
Shadows (49)	0.49	−0.74	0.72	−0.86	0.66	−0.94
Manmade (177)	0.50	−0.73	0.73	−0.85	0.72	−0.94
Flowers (61)	0.52	−0.72	0.70	−0.81	0.59	−0.87
Snow (40)	0.52	−0.72	0.76	−0.92	0.76	−0.97
Fruits (40)	0.54	−0.69	0.66	−0.73	0.61	−0.81
Landscape (81)	0.61	−0.81	0.72	−0.91	0.71	−0.98

[a]PI, patchiness index.

Correlations between layer patchiness across all scenes were low: 0.10, 0.38, and 0.32 for luminance with red–green, luminance with blue–yellow, and red–green with blue–yellow, respectively.

The average slopes of the amplitude spectra across all scenes for the red–green, blue–yellow, and luminance layers were −0.73, −0.85, and −0.91, respectively. After Bonferroni correction, the differences were all significant at the $p<0.001$ level: luminance versus red–green, $t(1514)=8.08$; luminance versus blue–yellow, $t(1514)=12.72$; red–green versus blue–yellow, $t(1514)=5.97$; all p values <0.0001. Correlations between spectral slope and patchiness were low: −0.14, −0.48, and −0.22 for luminance, red–green, and blue–yellow, respectively.

Relatively strong correlations were found among the spectral slopes of the three layers: 0.82 for red–green with blue–yellow, 0.76 for luminance with blue–yellow, and 0.6 for red–green with luminance.

5. DISCUSSION

We established a measure of patchiness that was both valid and robust to noise and used it to compare the patchiness of the red–green, blue–yellow, and luminance layers of images of natural scenes. Across a large number of scenes ($n=758$), the red–green layer was the most patchy, followed by the blue–yellow layer, with the luminance layer the least patchy. Scenes with foliage, manmade objects, trees, and nonplant scenes appeared to produce the biggest differences in patchiness between the red–green and luminance layers, while the smallest differences were found in scenes with landscapes and fruit.

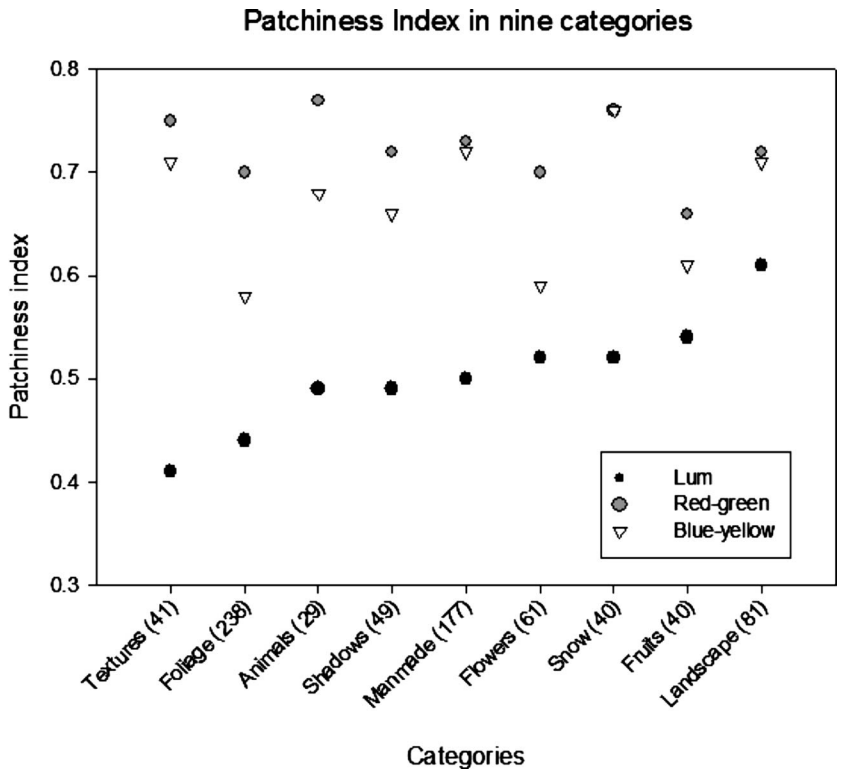


Fig. 5. Comparison of patchiness indices of three layers in different categories. Lum, luminance.

In the Introduction we said how previous studies had shown that the three postreceptoral layers were differentially sensitive to shadows and shading in natural scenes, with the order of sensitivity being luminance > blue–yellow > red–green [8,9,22]. Since this ordering of sensitivity is the reverse of the ordering of patchiness found here, a reasonable hypothesis is that the differences in patchiness between scenes, but more important among the three postreceptoral layer representations of those scenes, is determined by the amount of shadows/shading. Figure 6 shows two groups of scenes, the upper row with higher patchiness in the red–green than in the luminance layers, and the lower row the opposite. The images in the upper row proliferate with shadows and shading, whereas in the lower row shadows and shading are much less evident.

What might be the significance of these findings for the organization of the primate color vision system? Psychophysical studies have established that the red–green and blue–yellow channels have poorer spatial resolution than the luminance channel, but higher contrast sensitivity at low spatial frequencies [23,24]. The low-pass characteris-

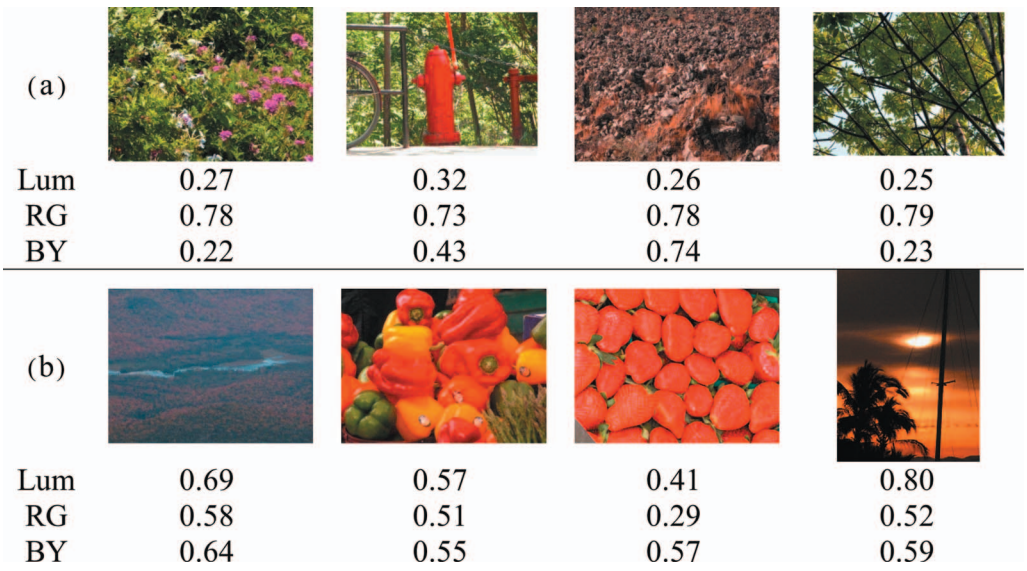


Fig. 6. (Top row) Four images whose red–green patchiness is greater than their luminance patchiness. (Bottom row) Four images whose luminance patchiness is greater than their red–green patchiness. Values of patchiness are shown underneath each image. Lum, luminance; RG, red–green; BY, blue–yellow.

tic of the red–green channel may have evolved to detect large patches of either red or green and/or to reduce noise levels [3], while the high-resolution luminance channel may have evolved to capture the fine details of shadows and shading. Many natural textures, for example, grass, are defined by their pattern of shadows and shading, a pattern contained primarily in the high spatial frequencies. However, since the spatial resolution of the blue–yellow channel is lower than that of the red–green channel [23,24], the idea that resolution evolved to reflect patchiness would predict that the blue–yellow layers would be more patchy than the red–green layers, whereas the opposite was found. It is more likely that the differences in resolution between the two chromatic postreceptoral channels reflect anatomical constraints imposed by the layout of the retinal cone mosaic [25,26] and/or by constraints imposed by chromatic aberration [27,28].

In the Introduction we mentioned the suggestion of Parraga *et al.* [9] that when pixels are grouped into large uniform areas, the slope of the Fourier amplitude spectrum tends to be slightly steeper than average. Consistent with this idea is our finding that the red–green layers, which were the most patchy, had significantly steeper spectral slopes than the luminance layers and that spectral slope and patchiness were in general negatively correlated. However, the correlations were low, indicating that spectral slope is a poor predictor of patchiness in natural scenes.

Finally, the ordering of kurtosis values for the three layers (red−green > luminance > blue−yellow) was different from that of our measure of patchiness (red−green > blue−yellow > luminance). We attribute this difference to the fact that the blue–yellow channel responses to the photographs of natural scenes were particularly susceptible to photographic noise, as shown in our simulations.

REFERENCES

1. H. B. Barlow, "Sensory mechanisms, the reduction of redundancy, and intelligence," in the *NPL Symposium on the Mechanisation of Thought Processes* (National Physical Laboratory, 1959), pp. 371–394.
2. J. M. Rubin and W. A. Richards, "Color vision and image intensities: when are changes material?" Biol. Cybern. **45**, 215–226 (1982).
3. J. J. Atick, Z. Li, and A. N. Redlich, "Understanding retinal color coding from first principles," Neural Comput. **4**, 559–572 (1992).
4. S. K. Shevell and F. A. A. Kingdom, "Color in complex scenes," Annu. Rev. Psychol. **59**, 143–166 (2008).
5. B. C. Hansen and R. F. Hess, "Structural sparseness and spatial phase alignment in natural scenes," J. Opt. Soc. Am. A **24**, 1873–1885 (2007).
6. D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," J. Opt. Soc. Am. A **4**, 2379–2394 (1987).
7. D. J. Tolhurst, Y. Tadmor, and T. Chao, "Amplitude spectra of natural images," Ophthalmic Physiol. Opt. **12**, 229–232 (1992).
8. C. A. Parraga, G. Brelstaff, T. Troscianko, and I. R. Moorehead, "Color and luminance information in natural scenes," J. Opt. Soc. Am. A **15**, 563–569 (1998).
9. C. A. Párraga, T. Troscianko, and D. J. Tolhurst, "Spatiochromatic properties of natural images and human vision," Curr. Biol. **12**, 483–487 (2002).
10. M. F. Tappen, W. T. Freeman, and E. H. Adelson, "Recovering intrinsic images from a single image," Adv. Neural Inf. Process. Syst. **15**, 1459–1472 (2002).
11. A. Olmos and F. A. A. Kingdom, "A biologically inspired algorithm for the recovery of shading and reflectance images," Perception **33**, 1463–1473 (2004).
12. A. P. Johnson, F. Kingdom, and C. J. Baker, "Spatiochromatic statistics of natural scenes: first- and second-order information and their correlational structure," J. Opt. Soc. Am. A **22**, 2050–2059 (2005).
13. M. E. Churma, "Blue shadows: physical, physiological, and psychological causes," Appl. Opt. **33**, 4719–4722 (1994).
14. D. J. Field, "What is the goal of sensory coding?" Neural Comput. **6**, 559–601 (1994).
15. F., A. A. Kingdom, A. Hayes, and D. J. Field, "Sensitivity to contrast histogram differences in synthetic wavelet-textures," Vision Res. **41**, 585–598 (2001).
16. R. A. Boie and I. J. Cox, "An analysis of camera noise," IEEE Trans. Pattern Anal. Mach. Intell. **14**, 671–674 (1992).
17. D. L. Ruderman, T. W. Cronin, and C. C. Chiao, "Statistics of cone responses to natural images: implications for visual coding," J. Opt. Soc. Am. A **15** 2036–2045 (1998).
18. D. J. Field, "Scale-invariance and self-similar 'wavelet' transforms: an analysis of natural scenes and mammalian visual systems," in *Wavelets, Fractals, and Fourier Transforms*, M. Farge, J. C. R. Hunt, and J. C. Vassilicos, eds. (Clarendon, 1993), pp. 151–194.
19. D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," J. Opt. Soc. Am. A **4**, 2379–2394 (1987).
20. A. Olmos and F. Kingdom, "McGill Calibrated Colour Image Database" (McGill Vision Research, 2004), http://tabby.vision.mcgill.ca.
21. V. C. Smith and J. Pokorny, "Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm," Vision Res. **15**, 161–171 (1975).
22. F. A. Kingdom, C. Beauce, and L. Hunter, "Colour vision brings clarity to shadows," J. Vision **3**, 637–637 (2003).
23. K. T. Mullen, "The contrast sensitivity of human colour vision to red–green and blue–yellow chromatic gratings," J. Physiol. (London) **359**, 381–400 (1985).
24. G. J. C. Van Der Horst and M. A. Bouman, "Spatio-temporal chromaticity discrimination," J. Opt. Soc. Am. **59**, 1482–1488 (1969).
25. P. K. Ahnelt, H. Kolb, and R. Pflug, "Identification of a subtype of cone photoreceptor, likely to be blue sensitive, in the human retina," J. Comp. Neurol. **255**, 18–34 (1987).
26. C. A. Curcio, K. A. Allen, K. R. Sloan, C. L. Lerea, J. B. Hurley, I. B. Klock, and A. H. Milam, "Distribution and morphology of human cone photoreceptors stained with anti-blue opsin," J. Comp. Neurol. **312**, 610–624 (1991).
27. A. van Meeteren, "Calculations on the optical modulation transfer function of the human eye for white light," Opt. Acta **21**, 395–412 (1974).
28. A. Bradley, X. Zhang, and L. Thibos, "Failures of isoluminance caused by ocular chromatic aberrations," Appl. Opt. **31**, 3657–3667 (1992).