Stability of the color-opponent signals under changes of illuminant in natural scenes

P. G. Lovell

Department of Experimental Psychology, University of Bristol, 8 Woodland Road, Bristol BS8 1TN, UK

D. J. Tolhurst

Department of Physiology, University of Cambridge, Downing Street, Cambridge CB2 3EG, UK

C. A. Párraga, R. Baddeley, and U. Leonards

Department of Experimental Psychology, University of Bristol, 8 Woodland Road, Bristol BS8 1TN, UK

J. Troscianko

Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

T. Troscianko

Department of Experimental Psychology, University of Bristol, 8 Woodland Road, Bristol BS8 1TN, UK

Received January 31, 2005; revised manuscript received April 7, 2005; accepted April 28, 2005

Illumination varies greatly both across parts of a natural scene and as a function of time, whereas the spectral reflectance function of surfaces remains more stable and is of much greater relevance when searching for specific targets. This study investigates the functional properties of postreceptoral opponent-channel responses, in particular regarding their stability against spatial and temporal variation in illumination. We studied images of natural scenes obtained in UK and Uganda with digital cameras calibrated to produce estimated L-, M-, and S-cone responses of trichromatic primates (human) and birds (starling). For both primates and birds we calculated luminance and red–green opponent (RG) responses. We also calculated a primate blue–yellow-opponent (BY) response. The BY response varies with changes in illumination, both across time and across the image, rendering this factor less invariant. The RG response is much more stable than the BY response across such changes in illumination for primates, less so for birds. These differences between species are due to the greater separation of bird L and M cones in wavelength and the narrower bandwidth of the cone action spectra. This greater separation also produces a larger chromatic signal for a given change in spectral reflectance. Thus bird vision seems to suffer a greater degree of spatiotemporal "clutter" than primate vision, but also enhances differences between targets and background. Therefore, there may be a trade-off between the degree of chromatic clutter in a visual system versus the degree of chromatic difference between a target and its background. Primate and bird visual systems have found different solutions to this trade-off. © 2005 Optical Society of America

1. INTRODUCTION

The spectrum of light reaching an observer's eye from an object is determined not just by the reflectance function of the object's material but also by the spectral properties of the illuminant. The illuminant typically varies markedly in intensity over time and space and this is especially true of objects of high three-dimensional spatial complexity, such as the foliage of a tree. An important task of vision is therefore to be able to detect the *invariant* material properties of the surface (e.g., the reflectance function) while being invariant to the highly variable illumination. It has been suggested that opponent-color vision might fulfill this task.[1,2]

The work of DeValois and his colleagues[3–5] has provided ample evidence that the postreceptoral channels in monkeys consist of three opponent channels, which can be thought of as encoding the red–green (RG), blue–yellow (BY), and light–dark (Lum) aspects of a scene. There is also ample evidence that human perception of color uses RG and BY opponent channels, as originally proposed by Hering.[6] The BY (or blue–green) opponent system is found in many mammalian orders,[7] but the paradoxical RG system has much more limited mammalian distribution, mostly in old-world monkeys and apes. The paradox is that the amount of RG outflow from the monkey retina is immensely greater than the BY,[8,9] and yet RG changes within the natural world are relatively rare.[10] In the present paper, we explore the manner in which such opponent channels might sample natural images containing real noise such as shadows and specularities. We also explore how the opponent channels respond in an environment in which light changes naturally in intensity and

"color" over time as a result of changes in solar position and atmospheric properties.

Color vision is proposed as a means of removing the camouflaging effects of shadows in the belief that directly illuminated and shadowed parts of an object will differ only in the intensity of illuminant.[11] However, shadowed areas do not differ solely in the intensity of illumination. The "color" of the natural illuminant is not constant. It varies over the time course of a day, primarily in the BY direction of color space,[12] although the light filtering though a tree canopy also varies in its greenness.[13] The light from a cloudy sky is bluer than direct sunlight. This means that on a sunny day, any shadowed area receives scattered light from the blue sky, which has a higher proportion of short wavelengths than direct sunlight. Shadows are therefore blue in comparison with areas of the same material lit by direct sunshine. Even on a cloudy day, it has been shown that the area of sky corresponding to the position of the sun is more yellow than sky more distant from the sun's location,[14] so similar, though smaller, effects may be expected on cloudy days as well. Furthermore, the shadowed areas may receive illumination locally reflected from other objects and, in natural scenes, those other objects are likely to be green foliage. The BY system may therefore not be invariant across shadow boundaries.[15,16] If the visual system is attempting to extract reflectance, and therefore changes due to illumination are viewed as noise, the YB opponent system will show substantial illumination noise both as a function of space and time.

The purpose of this paper is to explore these issues more formally. Specifically, does the RG opponent system provide a means of encoding the spectral properties of objects such as edible fruit in an invariant manner, over the course of a day, during which time illumination will change markedly? And how badly affected is the BY opponent system by shadows and daily changes in the illuminant? This proposal for a RG system as adapted to minimize illumination noise complements proposals that the primate RG opponent system is optimized for detecting and differentiating potentially edible objects.[17–21]

We will also consider how some of the scenes in this paper will be encoded by birds—specifically by starlings, *Sturnus vulgaris*, whose receptor sensitivity is described by Hart *et al.*[22] We do this because birds have a different set of spectral sensitivities for their L and M cones compared with primates, and are thought to have a RG opponent-color mechanism.[23] A comparison of the neural encoding of trichromatic primates and birds is expected to give an interesting insight into the design of both systems.

Part of this work has been described briefly elsewhere.[24–26]

2. METHODS

A. Cameras and Calibration

We took photographs of natural scenes with Nikon digital cameras and used their uncompressed outputs to calculate how human long-, medium-, and short-wavelength (L, M, S) cones and how starling L and M cones would have responded at every point in a scene. This required thorough characterization of each camera's nonlinear gamma function and the spectral activation functions of its three (R,G,B) sensors.

The cameras used in this work were a Nikon Coolpix 950 (camera 1) and a Nikon Coolpix 5700 (camera 2). All automatic settings were turned off; these included image sharpening, selection of white balance, selection of exposure aperture, and integration time. White balance was set to "cloudy," as was the case during the calibration process, and central-spot metering was used. The lens aperture value was manually fixed to allow the maximum available depth of focus (*f*11.4 in camera 1 and *f*7.4 in camera 2) and the cameras' built-in software was free to find the optimal integration time (shutter speed); the flash was disabled. The picture outputs were uncompressed .tif images. Image sizes were 1600 X 1200 pixels for camera 1 and 2560 X 1920 pixels for camera 2. However, photographs taken with camera 2 were subsampled by averaging odd and even rows and columns in order to reduce computer-processing time; this gave an effective resolution of 1280 X 960 pixels. The cameras were mounted on tripods, and a remote shutter release was used to avoid small camera movements and registration problems between successive pictures in time-lapse sequences, in which the cameras were programmed to take pictures at intervals of 4 min of the same scene. Figure 1 shows a montage of some of the colored photographs that we use as the basis for analysis in this paper.

To obtain the LMS cone activations for every point of the visual scene from the RGB pixel values in the .tif files, we characterized the cameras, using the methodology described below, in terms of the LMS cone representations of humans[27,28] (see Fig. 2(*A*)) and starlings (see Fig. 2(*B*)).[22] It will be noted that, by comparison with the human, the starling (like many birds) has four cones whose action spectra are narrower than in the human and are more evenly spaced across the visible spectrum. One might surmise that birds would have "better" color vision than a human. In any case, the L and M cone characteristics are markedly different and a comparison with human vision should be revealing. We concentrate upon those cones of the starling that correspond best to the LMS sensitivities of primates, and so we have excluded those cones sensitive to UV and double cones. We verified the extent of the error in our polynomial RGB-to-LMS mapping as follows:

Characterization of a camera in terms of LMS responses would yield exact results only if the spectral sensitivities of the imaging system are exact transformations of the LMS cone representations (e.g., one could transform exactly from camera RGB space to LMS space using a 3 X 3 transformation matrix). However, this condition is not normally satisfied by commercial camera manufacturers, and therefore our color space transformations will be approximations and subject to error. There is also the issue of device metamerism, where theoretically, two different surfaces under the same illumination may produce the same camera response and be modeled with the same LMS values. To overcome these limitations of the imaging device, we took advantage of the fact that the spectral reflectance of most natural surfaces (as well as natural illumination) are relatively smooth functions of

Fig. 1. (Color online) This montage shows examples of the digitized color photographs that we analyze in this paper. Photograph A features ripe and unripe tomatoes; the opponent responses to tomatoes x, y, and z are examined later. Note that x is shadowed while y and z are illuminated by direct sunlight. Photographs B and C are examples of the 113 photographs of ripe fruit taken from the Kibale dataset. Photograph D is one of a time-lapse sequence taken at 4 min intervals from dawn till dusk; note the gray card in upper left, which was used to take radiometric measurements of the illuminant. These images are the regular output of the cameras; they have not been linearized or gamma corrected for display. In D, thin blue lines outline one or more fruits whose pixels were compared with the surrounding area defined by the thin red lines.

wavelength[29–32] and can be represented by a small number of basis functions[33] (this issue has been a subject of discussion in the literature).[34,35]

The cameras' RGB pixel values were characterized in two stages. First, the cameras were pointed at a Macbeth ColorChecker card illuminated by a tungsten-halogen light source (Osram HLX 64657FGX, 250 W). The light source was supplied with constant current from a stabilized DC power supply (custom-made accurate to 30 parts per million in current). We illuminated a white card with this lamp and measured the CIE-Y value with a spectroradiometer (TopCon, Model SR1, calibrated by the National Physical Laboratory, UK). Once both pieces of equipment were warm, the CIE-Y value varied with a SD of less than 0.25% of the mean. Several pictures (at different exposure durations) were taken of the card's lower row of gray squares, and their RGB values were computed by averaging the central part (under- and over-exposed values were discarded) and scaled by dividing by the corresponding integration time. The spectral radiance of the same regions was measured in the range 380 – 760 nm in 10 nm steps using the TopCon radiometer, allowing us to relate each sensor's gray-level output to a physical measure of the total spectral energy stimulating them. The linearity of the camera's responses as a function of integration time was corroborated by a set of three neutral density filters (0.5, 1.0, and 2.0 log units) that forced the camera to adopt different shutter speeds when photo-

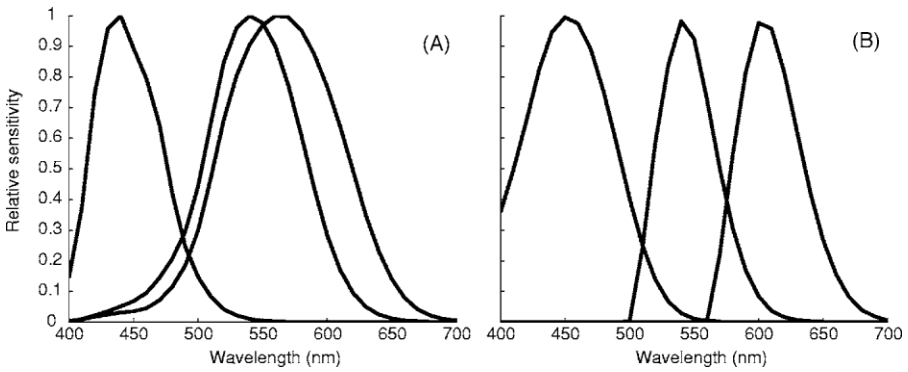

Fig. 2. (A) Cone sensitivities for humans, scaled to unity. (B) Cone sensitivities for the selected starling cones. Note that the L and M cones are much closer to one another for humans than they are in the case of the starling. The bandwidths of the L and M cones also differ; they are narrower for the starling.

graphing the white target. The data allowed us to characterize the nonlinear relation between scene radiance and the magnitude of the pixel values inherent in the camera (we did this for the "cloudy" setting) and to produce inverting functions for each of the RGB sensors that would linearize the camera output.

The second stage consisted of measuring the camera sensor's spectral sensitivities by pointing the cameras at a white target (Kodak-Eastman "standard white" cyanoacrylate powder of approximately 99% reflectance, constant through the visible spectrum) illuminated by the same light source. Images of the target were taken through a set of 31 narrowband color filters (10 nm bandwidth, Ealing Electronics, Watford, UK) spanning the range 400 – 700 nm. Spectral radiance was measured through the same filters by the TopCon radiometer, making it possible to determine the camera's RGB sensor's spectral sensitivities. We used a second-order polynomial model to map linearized RGB value triplets into LMS triplets.[36–41] This characterization technique differs from others in that it allows us to find the optimal colorimetric mapping for a given set of reflectances and illuminations. For example, given a typical set of Northern European vegetation and soil reflectances[42] and natural illumination,[43] it is possible to estimate both the camera response values and the LMS values for the natural scenes that we investigate. For the Kibale forest photographs we used another database of illuminants and reflectance.[44]

We computed the LMS output by calculating the product of the cone sensitivities with half of the samples in our training database using

$$L = \sum_{\lambda} l(\lambda)^{*} Q(\lambda)^{*} I(\lambda), \qquad M = \sum_{\lambda} m(\lambda)^{*} Q(\lambda)^{*} I(\lambda),$$

$$S = \sum_{\lambda} s(\lambda)^{*} Q(\lambda)^{*} I(\lambda), \quad (1,3)$$

where l, m, s are the Smith and Pokorny[27,28] (or starling) cone sensitivities, λ is the wavelength, Q is the spectral reflectance of the samples, and I is the spectral radiance of the illuminant.

At the same time we calculated the camera's response to the same reflectances and illuminations using

$$R = \sum_{\lambda} r(\lambda)^{*} Q(\lambda)^{*} I(\lambda), \qquad G = \sum_{\lambda} g(\lambda)^{*} Q(\lambda)^{*} I(\lambda),$$

$$B = \sum_{\lambda} b(\lambda)^{*} Q(\lambda)^{*} I(\lambda), \qquad (4\text{-}6)$$

where r, g, b are the camera's spectral sensitivities.

The predicted camera RGB responses where then mapped to LMS activities using our polynomial transform, created using the other half of the training database.

The relative error of the polynomial transformation was calculated for each of the samples of the test dataset according to

$$\mathrm{Err} = \frac{\sqrt{(L-\hat{L})^2 + (M-\hat{M})^2 + (S-\hat{S})^2}}{\min(\sqrt{L^2+M^2+S^2}, \sqrt{\hat{L}^2+\hat{M}^2+\hat{S}^2})}, \quad (7)$$

where $\hat{L}, \hat{M}, \hat{S}$ represent the mapped cone activities. The mean error (Err) was 0.034 (SD = 0.034 , n = 1095) for the Northern European dataset and 0.016 (SD = 0.0159 , n = 783) for the Ugandan dataset for human (unity) cones. For starlings the mean errors were 0.056 (SD = 0.05 , n = 783) and 0.01 (SD = 0.01 , n = 1095) for the Northern Europe and Ugandan datasets, respectively.

B. Datasets

We use two main datasets in our analyses, one of which consists of 113 images of fruit taken in Kibale Forest, Uganda (Fig. 1(*B*) and 1(*C*)). This area features in some important studies of optimization of primate vision to frugivory and folivory[18,19] because of its large population of foraging primates. The images within this dataset were taken with camera 1. Figure 1(*A*), taken with the same camera, is part of the British garden dataset used previously.[21]

The second main dataset is a time-lapse sequence (4 min intervals) taken from dawn till dusk in a British garden, in the village of Garndiffaith (South Wales), of a scene in which edible fruits have been situated amongst foliage. These were taken with camera 2. Photographs near the beginning or end of the sequence were excluded where the exposure duration exceeded 2 s since the conditions were then too dark for the camera to provide a noise-free image. At the same time, the TopCon radiom-

eter was used to make measurements (1–2 min intervals) of the total radiance and of the spectrum (380–720 nm, 10 nm intervals) of a 2° patch gray card at the top left of the composition (Fig. 1(*D*)). A time-lapse animation of the opponent responses and spectral measurements can be downloaded from the following website along with the raw images: http:// llpsy223.psy.bris.ac.uk/george/ timelapse/.

C. Modeling and Calculation of Opponent Signals

Following calculation of the LMS activation at each pixel, we established the opponent activity using the MacLeod–Boynton[45] formulas. Equations (8)–(10) define the calculations for the Lum, RG, and BY, opponent channels, respectively, as

$$\text{Lum} = \text{L} + \text{M}, \tag{7'}$$

$$\text{RG} = \text{L}/\text{Lum}, \tag{8}$$

$$\text{BY} = \text{S}/\text{Lum}. \tag{9}$$

It will be noted that our present formulation of the RG signal [Eq. (9)] is directly proportional to the one we have used before[46,47]:

$$\text{RG} = (\text{L} - \text{M})/\text{Lum}. \tag{8a}$$

Noise within the L, M, and S cones was simulated[48] following estimates derived from reported thresholds[43]:

$$n = s^{*} \exp(y), \tag{10}$$

where s is the L, M, or S signal estimated from our calibrated cameras; n is the noisy signal; and y is a random variable with a standard deviation matching the Weber fractions for the S, M, and L channels (8.7%, 1.9%, and 1.8%, respectively, with 1°, 0.2 s stimulus at the level of 10X absolute threshold).

D. Distinguishing Fruit from Leaves

To illustrate how well an opponent signal might help separate a fruit from surrounding leaves, we calculate d' values [Eq. (11)], where the x and y values represent the activation levels for the signal (fruit) and noise (leafy surround), respectively, in a particular opponent channel. The metric is intended to show how much the signals from fruit and background overlap rather than to suggest whether one could tell that a fruit was different from background, given enough measurements

$$d' = \frac{\sqrt{2|\langle x\rangle - \langle y\rangle|}}{\sqrt{\sigma^2(x) + \sigma^2(y)}}. \tag{11}$$

3. RESULTS

A. Examining Opponent Channel Activation for Directly Illuminated and Shadowed Fruit

The potential role for RG opponency in helping primates to find ripe fruit is shown in Fig. 3. Figure 1(*A*) shows a photograph of tomatoes (some ripe, some not) against a background of leaves; the scene was illuminated by direct sunlight, so it contains many shadows [one of the ripe tomatoes (x) is shadowed]. Figures 3(*A*) and 3(*B*) show gray-level representations of the signals generated in the opponent channels. In both the RG and BY channels, the individual tomatoes are obvious as separate entities, but the ripe tomatoes are distinguished only from all else in the RG channel. All ripe tomatoes (whether in shadow or not) have the same RG signal magnitude (solid black curves Fig. 3(*C*)). For the BY signals the shadowed tomato has a signature different from those under direct illumination (Fig. 3(*D*), solid black curves). The plots in Figs. 3(*E*) and 3(*F*) show the d' value for the ripe fruit regions (x, y, and z) compared to their immediate surround (excluding other ripe fruit). Clearly, the d' values are greater for the RG channel than they are for the BY channel, illustrating the benefit of the RG channel for frugivory. But why has the RG channel performed so well? Clearly the primary reason for the success of the RG channel is that the system contrasts red and green activity. Another reason for the strong performance of the RG system is the proximity of the L and M cones; by having these cones close to one another the RG opponent channel becomes insensitive to variations in illumination, both over time and across space. This issue is explored in Subsection 3.B.

The relatively poor performance of the BY channel compared with the RG channel is perhaps unsurprising. The color of the illuminant is not constant, either during the day (see below) or even within a scene.[14,15] The light that falls directly on an object in bright sunlight will be "white" or yellow, but the light falling on shadowed objects will have come from the blue part of the sky or will be reflected from local objects (probably green ones). Consequently the BY system will respond very differently to the shadowed fruit compared with the directly illuminated fruit. Conversely, the A_{max} values for the L and M cones are so close that their responses to a varying illuminant will correlate to a very high degree.

B. Changes in the Illuminant During a Day.

Subsection 3.A has shown that at an instant in time, the BY opponent system seems disrupted by *spatial* discontinuities in illumination. In this section we look at temporal changes in the illuminant and how the BY and RG systems are affected. Figure 4 shows the results of an experiment where we measured the radiance spectrum (see Section 2) of a gray card placed in an open garden; measurements were made every 1–2 min from dawn to dusk on an autumn day when the weather was mostly cloudy.

Figure 4(*A*) shows examples of normalized spectra, i.e., the radiance in each wavelength is divided by the total radiance at all wavelengths in a particular recording. T1 represents the spectrum at dawn; here the illuminant has a reddish tint. T2 shows how the illuminant becomes bluer as clouds became thicker. Over the course of the day the spectrum of the illuminant has changed substantially, depending on weather conditions and the elevation of the sun in the sky. Figure 4(*B*) shows calculations of how the primate RG chromatic opponent signal from the gray card (solid curve) and the BY opponent signal (dotted curve) would have changed. Despite the spectral changes in illumination, the RG signal is fairly stable, as we would hope from an opponent chromatic mechanism. However, the BY signal is affected more substantially, which is bound

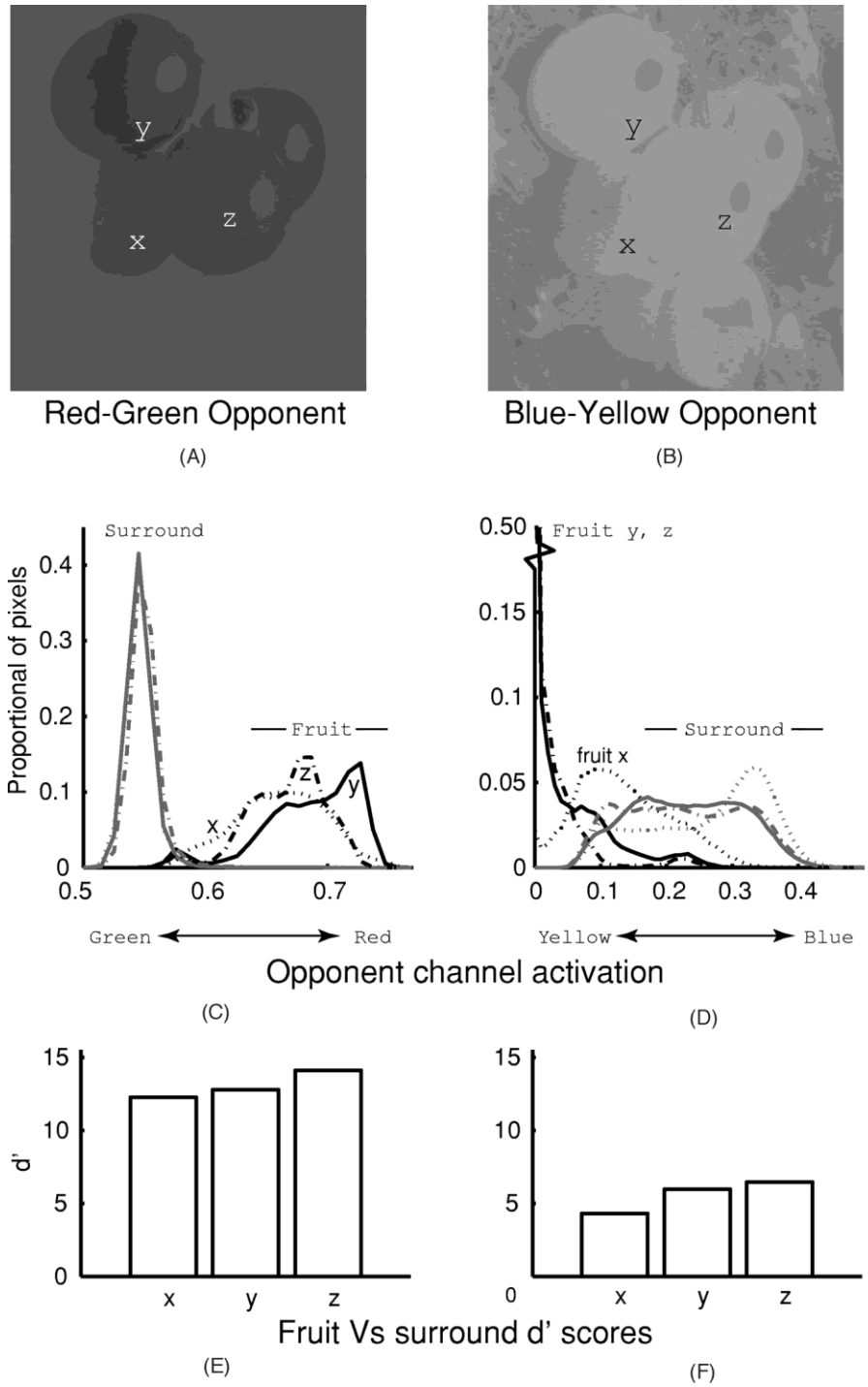

Fig. 3. Top row, gray-level representations of the activation in the primate (A) RG- and (B) BY opponent channels calculated for the image of tomatoes in Fig. 1(*A*). Middle row, histograms of pixel activity levels for the ripe fruit (x, y, z) and the area surrounding each fruit; note the largely separate distributions for fruit (black curves) and the surround (gray curves) in the RG channel [plot (C)]. In the BY channel the distribution for the shadowed fruit (x) is similar to that of its surround [plot (D)]. Bottom row, d' values for each ripe-fruit region compared with its surrounding (nonripe fruit and leaves) area in the RG channel [(plot E)] and the BY channel [(plot F)].

to reduce the likelihood of the BY system producing invariant responses from differently colored objects at different times of day. In contrast to the primate RG signal the equivalent activation for the bird (dotted-dashed curve) varies to a greater extent because of the greater separation of the bird L and M cones (standard deviations for these RG signals were 0.004 for primate and 0.009 for bird, while the primate BY signal had a standard deviation of 0.034.

C. Opponent Encoding of Fruit and Foliage at Different Times of Day

At the same time these radiance measurements were made, we took time-lapse photographs with camera 2 of

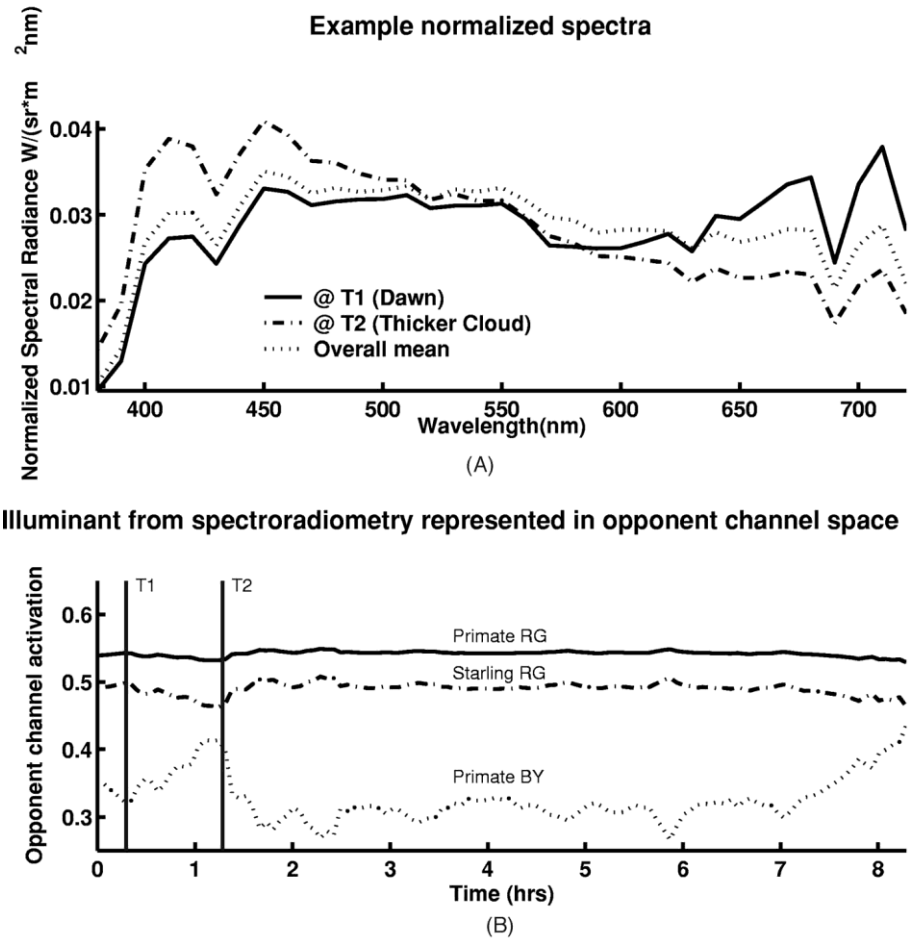

Fig. 4. Summary of spectral radiance measures of sunlight during the day November 23, 2004, in a British garden (Fig. 1(*D*)); time zero was 07:50 GMT. (A) Plots of the normalized spectra at T1 and T2; these times are indicated as vertical lines on (B). The average of the normalized spectra is also shown. (B), The primate RG (solid curve) and By (dotted curve) chromatic opponent signals [Eqs. (8) and (9)] of the light reflected from the gray card are plotted against time. The starling RG signal is also plotted (dotted–dashed) curve.

the scene depicted in Fig. 1(*D*). The illuminant measurements of Fig. 4 are thus a description of the illuminant at one location within this scene. From the sequence of photographs, we determined how the luminance signal and the primate RG and BY chromatic opponent signals generated by each of the fruits would change during the day; we also examined how the bird RG signal from each fruit would change. We also examined the light reflected from a region of interest around each fruit—this was a region the same shape as the fruit but with five times the area (see Fig. 1(*D*)). The fruit itself was obviously excluded from the surround region of interest, and, where there were two fruits close together (regions 3 and 6), the neighboring fruit was also excluded from the analysis of the surrounding area.

Figure 5(*A*) shows how the d' scores for the primate (black curves) and starling (gray curves) RG-opponent signals vary as the illuminant changes during the day. The d' scores are shown only for two of the fruit in Fig. 1(*D*); the results for the others were similar. The d' score for the primate is consistently higher than it is for the starling. However, note that plums actually reflect UV light so that a starling (with UV-sensitive cones) may well use this as an added aid to detection. As one would expect from the stability of the illuminant in RG opponent space (Fig. 4), the d' values for both fruit remain fairly constant all day. By contrast, the d' scores for the tomato in the primate BY-opponent system change markedly during the day, as one might expect given the instability of the illuminant in BY opponent space (Fig. 4). It is at first surprising that the d' scores for the plum are so constant. Figure 5(*C*) shows how this can be; it plots the actual BY signals from plum, tomato, and their leafy surrounds during the day. All signals, especially from the plum, do vary as expected. However, the plum signal and the leafy background signal covary so that d' remains roughly constant. In this case, the BY opponent system might allow a constant level of detection for the plum, but the changing BY signal would mean that its "color" and identity might be confused as the illuminant changes during the day.

Figure 6 presents the mean d' score for each fruit against its surround over the course of the day. A high average d' score could be considered evidence that the opponent channel was successful at contrasting the fruit from the surround areas under many variations of the illumination. The primate RG signals provide the best discrimination, consistently surpassing the starling RG signals. The primate BY signals vary more during the day as a consequence of the varying illuminant (see Fig. 3) and as fruits become illuminated directly or have shadows

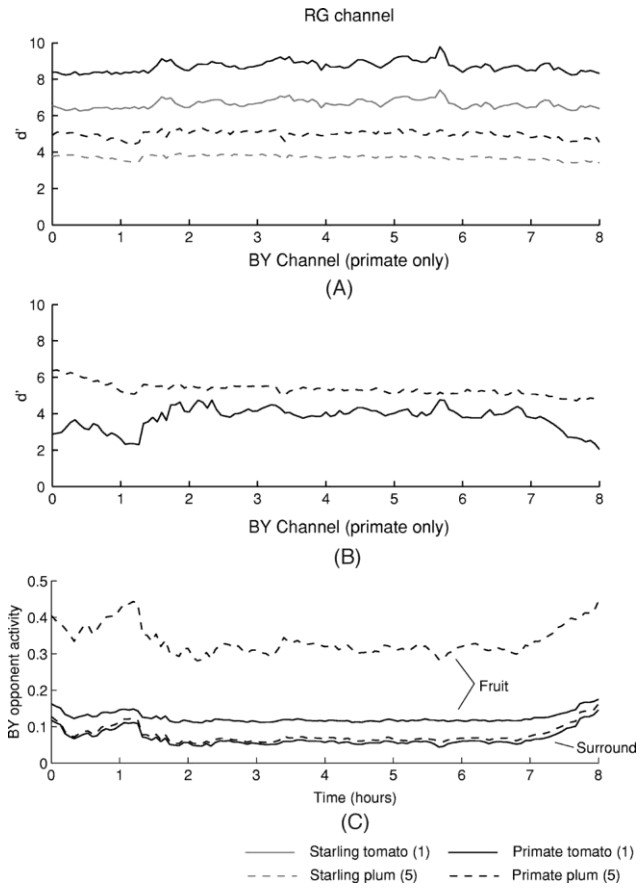

Fig. 5. Plots (A) and (B) show the *d'* values for the RG and BY signals, respectively, for a tomato in Fig. 1(*D*) and for a plum as a function of the time of day. The *d'* values for both fruit are very stable in the RG opponent system despite the changes in illuminant (see Fig. 4) during the day. The *d'* values for the plum in the BY system are also surprisingly stable. Plot C shows the actual BY signals generated from the plum, the tomato, and their leafy surrounds during the day. The BY signals *do* vary considerably (consistent with Fig. 4), but the signals from plum and leafy surround co-vary, so that the *d'* value stays fairly constant.

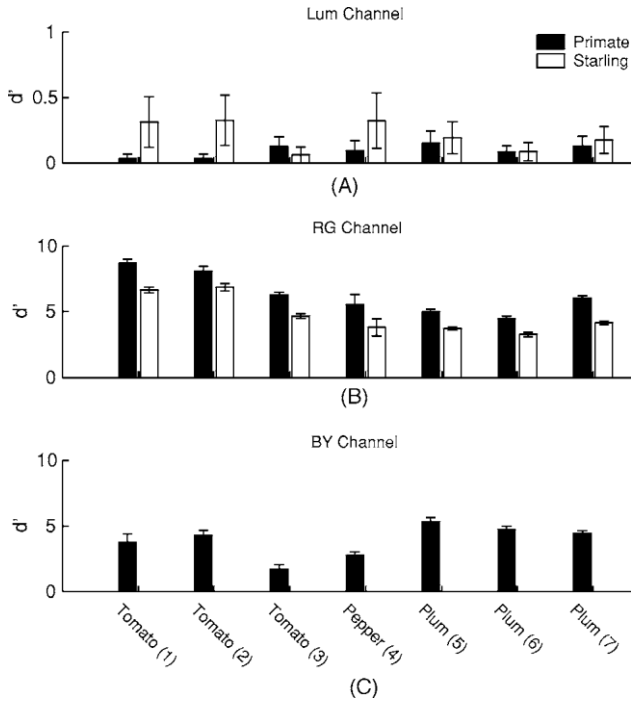

Fig. 6. (A) Averaged *d'* scores for each fruit versus the surrounding area in the Lum channel. (B) shows the mean *d'* scores for each fruit and surround region, i.e., the leftmost black bar shows the average of the topmost trace in Fig. 5. (C) *d'* scores for the primate BY channel. Note the *y* scale for the Lum channel is an order of magnitude smaller than for the RG and BY channels. The error bars represent the standard deviations of the *d'* scores. Solid bars are for primate channels; open bars for starling channels.

cast on them. The variation during the day is shown in Fig. 6 by the sizes of the standard deviation bars which, in most cases are 2–3 times larger for the (smaller) BY *d'* values than for the (larger) RG *d'* values.

D. Opponent Encoding of Fruit and Foliage Photographed in Kibale Forest

We performed similar analyses on each of 113 photographs of fruit taken in Kibale Forest in Uganda.[25] Two of these photographs are shown as Figs. 1(*D*) and 1(*E*). In each image we traced a region of interest (ROI) around a single isolated fruit or cluster of fruit; isolated fruits were chosen so comparisons between fruit and the surround area would not be confounded by the presence of fruit within the surround. All pixels within the ROI constituted "the fruit." To give the surround pixels, the fruit ROI was iteratively rescaled and recentered until the number of pixels in the surround region was five times greater than that of the fruit region—this operation excluded those pixels that constituted the fruit. The distribution of signal values (Lum, RG, or BY) within the fruit region was compared with the distribution of corresponding signal values from the surrounding background area (generally of green leaves).

For each of the 113 photographs, we calculated *d'* scores for the primate Lum, RG, and BY channels and Lum and RG for the starling. Figure 7 summarizes these analyses. Clearly the values given by the RG system (Fig. 7(*B*)) are much higher than those given by either the Lum (Fig. 7(*A*)) or the BY (Fig. 7(*C*)) systems, implying that the RG system would be substantially better at allowing identification of food than the other two systems. The white blocks in Figs. 7(*A*) and 7(*B*) show that putative starling Lum and RG systems would behave similarly to human ones. So regardless of species, the RG channel is most successful at achieving a strong separation between fruit and leafy background. The *d'* scores for the RG channel are once again slightly better for primate than for starling and have a magnitude similar to that reported in Subsection 3.D.

E. What Factors Led to the Discrepancy in the *d'* Scores for the Primate and Starling RG Channel?

In all three datasets reported above, the *d'* scores for the human RG channel were consistently greater than they were for the starling RG channel. The starling and primate L and M cones differ in two respects (see Fig. 2): First the primate cones are more closely spaced, second the bandwidths of the action spectra of the cones also differ (for the human cones the bandwidth is approximately

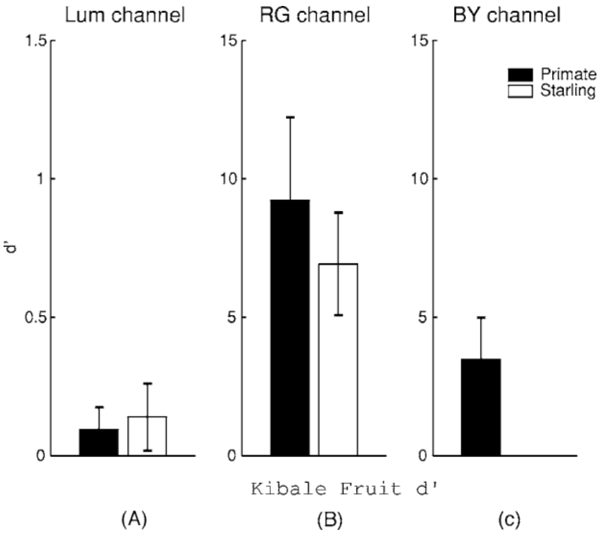

Fig. 7. 113 photographs of fruits in the Kibale Forest (Uganda) were analyzed as in Fig. 3. The values of the three opponent channels for both human and starling were measured for all the pixels within a ROI comprising the outline of a ripe fruit, and were compared with the values in a region of the leafy background surrounding the fruit. The bar charts show the mean d' scores for (A) the luminance signal, (B) the RG opponent signal, and (C) the BY opponent signal. Solid bars are for primate opponent systems; open bars in A and B are for putative starling systems. Note that the y scale for the Lum channel bars is an order of magnitude smaller than for the BY and RG bars.

40 nm while for the starling this is 20 nm). In order to understand the significance of these differences, we have calculated the RG d' scores for all fruit versus surround regions in the time-lapse photographs while varying the characteristics of simulated L and M cones whose bandwidths and peaks were systematically varied. The action spectra of the L and M cones were modeled with normal probability-density functions using a least-squares fit. The A_{max} of the M cone was fixed at 543 nm while the A_{max} of the L cone was varied from 553 – 633 nm in steps of 10 nm. The bandwidths of the action spectra of the L and M cones were varied together from 10 – 80 nm in steps of 10 nm. d' scores were computed for each combination of L cone position and cone bandwidth.

The d' scores are presented in Fig. 8, along with symbols indicating the actual cone action spectra of humans and starlings. The Fig. 8 surface plot reveals that the relative improvement in d' scores for the human RG channel over the starling RG channel is due to both the proximity of the L and M cones and to the increased bandwidth of the human cones. This simulation was repeated for the Kibale photo dataset reported in Subsection 3.D and the results correlated closely. We estimated the degree of error likely in the polynomial transformation of camera RGB values to the L value of our moving red cone. This error was never greater than 12% for the red cone in the positions tested [see Eqs. (1–7) for details of the error calculation].

4. DISCUSSION

The basic question underlying this paper concerns the functional roles of the opponent-color systems, both in primates and in a bird species (starling) that has rather different color vision. When De Valois and his co-workers[3,4] obtained evidence for the existence of postreceptoral opponent channels, little was known about the relationship between these and the structure of the visual environment and the visual tasks that need to be performed on it.

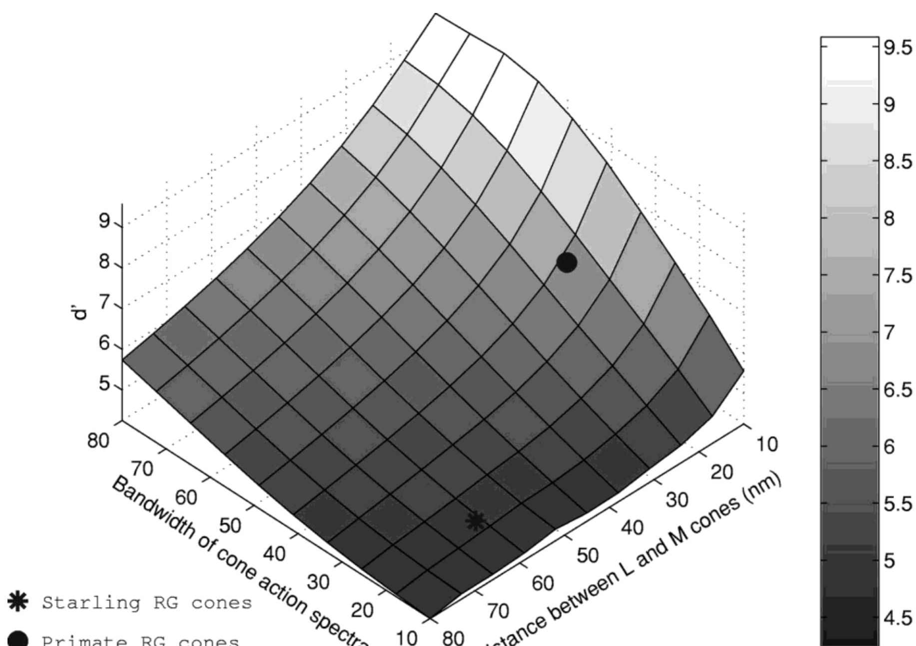

Fig. 8. Time-lapse photographs were reanalyzed as in Fig. 3 while varying the position of the L cone and the bandwidth of the L and M cones. The d' scores represent the mean d' score for all fruit and region pairings over all time intervals. The symbols show how the difference in d' scores reported for starlings and primates is due to the distance between the cones and the bandwidth of their action spectra.

The results presented here suggest a major role for color channels in informing the system about material (reflectance) properties of objects. This in itself is not a new story; many others have discussed color vision along these lines.[11,16,17] Our results are in keeping with a low-level "cleaning up" of the information incident in the retinal image: removing shadows, compensating for changes in illumination, and thus allowing a cleaner object-based segmentation to be carried out.[16] It is worth comparing this with the shadow problem in lightness perception (monochrome vision). Here the work of Gilchrist and others[49,50] suggests that a strategy for removing illumination changes needs to be high level, since it needs to know about scene 3D geometry: the "coplanar ratio principle" of Gilchrist makes this explicit. On the other hand, color vision can remove shadows much more simply; see for example the work of Olmos and Kingdom.[16] This in itself suggests a major role for color information in early vision.

But how is this role distributed between the two color-opponent systems (RG and BY) in primate vision, and how does this vary compared with a different species? The first point of note here is that the BY system does a surprisingly imperfect job of removing shadows, since these often have a blue tint due to the inhomogeneity of solar/sky illumination. The BY system is best thought of as carrying out an approximate segmentation of the scene into areas of different spectral reflectance, but this segmentation is confounded with shadow/illumination effects. It therefore follows that a dichromatic mammal (as most orders of mammal are likely to be) will be able to function well in situations in which light is relatively "flat" and/or the chromatic target is large/strongly colored. Foraging will work well in simple viewing conditions, but a visual search will become less efficient as shadows increase, or target salience decreases. Also since the output of the BY system is larger than that of the RG, it will operate better when the individual receptors are operating at lower signal-to-noise ratios (say at lower luminance levels). This ability to work at lower light levels would present advantages to an animal foraging primarily at dawn and dusk.

The RG system in primates seems to be much less confounded by capricious changes in illumination and will remain efficient at such search tasks for much longer, since (in primates at least) it has two important properties that the BY system does not have. First, it is optimally set up to distinguish edible fruit/leaves from inedible leaves. Second, it ignores shadows and changes in illumination except just after dawn and just before sunset. It is this combination of properties that gives trichromatic mammals a foraging advantage. The work presented here predicts that differences in foraging success in primate dichromats and trichromats should depend on the type of illumination prevailing at the time. Indeed, a recent experimental study has shown that trichromatic primates are more efficient at selecting ripe fruit than dichromats.[51]

However, we know that primates are not the only creatures with a RG-opponent system. Birds, which have cone spectral sensitivities that are different from primate L and M cones, also have RG opponency.[23] Indeed, we might infer that birds would have better color vision than we have, since they have *four* cone types with narrower activation spectra that are more evenly spaced across the spectrum. However, our conclusions are somewhat surprising: In some sense, the even spacing of the cone peaks is not always an advantage, nor is the narrower bandwidth of the starling cones.

How does a bird RG system compare with the primate one? Our results suggest that a bird's RG-opponent signals from a single surface would be more confounded by shadows and changes in natural lighting conditions than would a primate's. The bird RG is confounded by such changes almost as much as the primate's BY system. This difference arises because primate L and M cones have peak absorptions at very close wavelengths, so that gradual changes over the width of the spectrum will have less differential effect on the cones. The greater spectral separation of L and M cones in the starling retina (like the large spacing of the S cones from the L or M cones in the primate) has the consequence that invariance to shadows and changes in illumination is compromised. Areas of the image such as dappled foliage, which appear uniform to a primate RG system, may appear less uniform to a starling RG system. However, we predict that, if one were to equate the cone capture ratios of a target against a surround for humans and starlings, then humans would be less impaired by the incursion of shadows into the surround in a search task.

Many questions remain to be settled. Are there tasks for which the BY system is particularly good and the RG system pretty useless? Some data, not presented here, suggest that the discrimination of Rayleigh-scattered light from direct sunlight (i.e., sky versus solar disk) is well detected by mammalian BY vision. Thus, the detection of distance over hundreds of meters or more might be well encoded by BY opponent channels. Such work awaits experimental verification.

In general, we have provided more questions than answers. It is relatively easy to run mathematical models of detection across images of foliage and fruit; much harder to gather relevant behavioral data. However, we suggest that the results of this paper provide some indications about the functional role of the peculiar color vision systems in primates and one other species.

ACKNOWLEDGMENTS

This research was supported by BBSRC (S11501) and EPSRC/Dstl (RG/S56405/01) grants to D.J. Tolhurst and # T. Troscianko by a BBSRC grant (S18903) to T. Troscianko and I. Cuthill. P.G. Lovell was supported by the EPSRC/Dstl grant; C.A. Párraga was supported by the BBSRC grants. U. Leonards was supported by the Swiss National Science Foundation (83R-069377). We are grateful to I. Cuthill and M. Vorobyev for helpful discussions. We are also grateful to two referees for their challenging suggestions. We also thank the Ugandan Wildlife Authority and all the personnel at Makerere University Biological Field Station in Uganda, especially John Kasenene, its director. Karen Spencer is thanked for assisting in the UK time-lapse recording. Finally, we thank George Hayward for the use of his garden in the United Kingdom time-lapse sequences.

e-mail, p.g.lovell@bristol.ac.uk; phone, 44-117-928-8581; and fax, 44-117-928-8588.

REFERENCES

1. K. T. Mullen and F. A. A. Kingdom, "Colour contrast in form perception," in *The Perception of Colour,* P. Gouras, ed. (Macmillan, 1991), pp. 198–217.
2. V. V. Maximov, "Environmental factors which may have led to the appearance of colour vision," Philos. Trans. R. Soc. London, Ser. B **355**, 1239–1242 (2000).
3. R. L. De Valois, "Analysis and coding of color vision in the primate visual system," Cold Spring Harbor Symp. Quant. Biol. **30**, 567–580 (1965).
4. R. L. De Valois, I. Abramov, and G. H. Jacobs, "Analysis of response patterns of LGN cells," J. Opt. Soc. Am. **56**, 966–977 (1966).
5. R. L. De Valois and K. K. De Valois, "A multistage color model," Vision Res. **33**, 1053–1065 (1993).
6. L. M. Hurvich and D. Jameson, "An opponent-process theory of colour vision," Psychol. Rev. **64**, 384–404 (1957).
7. G. H. Jacobs, "Primate photopigments and primate color vision," Proc. Natl. Acad. Sci. U.S.A. **93**, 577–581 (1996).
8. T. N. Wiesel and D. H. Hubel, "Spatial and chromatic interactions in the lateral geniculate nucleous of the rhesus monkey," J. Neurophysiol. **29**, 1115–1156 (1966).
9. F. M. De Monasterio and P. Gouras, "Functional properties of ganglion cells of the rhesus monkey retina," J. Physiol. (London) **251**, 167–195 (1975).
10. 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).
11. J. D. Mollon, ""Tho she kneeld in that place where they grew." The uses and origins of primate colour vision," J. Exp. Biol. **146**, 21–38 (1989).
12. G. D. Finlayson and S. D. Hordley, "Color constancy at a pixel," J. Opt. Soc. Am. A **18**, 253–264 (2001).
13. C. C. Chiao, D. Osorio, M. Vorobyev, and T. W. Cronin, "Characterization of natural illuminants in forests and the use of digital video data to reconstruct illuminant spectra," J. Opt. Soc. Am. A **17**, 1713–1721 (2000).
14. T. Troscianko and J. P. Harris, "Phase discrimination in chromatic gratings," Perception **15**, A18 (1986).
15. D. Steverding and T. Troscianko, "On the role of blue shadows in the visual behaviour of tsetse flies," Proc. R. Soc. London, Ser. B **271**, S16–S17 (2003).
16. A. Olmos and F. A. A. Kingdom, "A biologically inspired algorithm for the recovery of shading and reflectance images," Perception **33**, 1463–1473 (2004).
17. M. G. Nagle and D. Osorio, "The tuning of human photopigments may minimize red-green chromatic signals in natural conditions," Proc. R. Soc. London, Ser. B **252**, 209–213 (1993).
18. P. Sumner and J. D. Mollon, "Catarrhine photopigments are optimized for detecting targets against a foliage background," J. Exp. Biol. **203**, 1963–1986 (2000).
19. N. J. Dominy and P. W. Lucas, "Ecological importance of trichromatic vision to primates," Nature (London) **410**, 363–365 (2001).
20. B. C. Regan, C. Julliot, B. Simmen, F. Vienot, P. Charles-Dominique, and J. D. Mollon, "Fruits, foliage and the evolution of primate colour vision," Philos. Trans. R. Soc. London, Ser. B **XXX**, 229–284 (2001).
21. C. A. Párraga, T. Troscianko, and D. J. Tolhurst, "Spatiochromatic properties of natural images and human vision," Curr. Biol. **12**, 483–487 (2002).
22. N. S. Hart, J. C. Partridge, and I. C. Cuthill, "Visual pigments, oil droplets and cone photoreceptor distribution in the European starling (*Sturnus vulgaris*)," J. Exp. Biol. **201**, 1433–1446 (1998).
23. D. Osorio, M. Vorobyev, and C. D. Jones, "Colour vision of domestic chicks," J. Exp. Biol. **202**, 2951–2959 (1999).
24. C. A. Párraga, T. Troscianko, and D. J. Tolhurst, structure of colour in natural scenes is changed," Perception **32**, 168b (2003).
25. T. Troscianko, C. A. Párraga, U. Leonards, R. J. Baddeley, J. Troscianko, and D. J. Tolhurst, "Leaves, fruit, shadows, and lighting in Kibale Forest, Uganda," Perception **32**, 51 (2003).
26. T. Troscianko, C. A. Párraga, P. G. Lovell, D. J. Tolhurst, R. J. Baddeley, and U. Leonards, "Natural illumination, shadows and primate colour vision," Perception **33**, 45A (2004).
27. V. C. Smith and J. Pokorny, "Spectral sensitivity of color-blind observers and the cone photopigments," Vision Res. **12**, 2059–2071 (1972).
28. V. C. Smith and J. Pokorny, "Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm," Vision Res. **15**, 161–171 (1975).
29. J. A. Endler, "The color of light in forests and its implications," Ecol. Monogr. **63**, 1–27 (1993).
30. L. T. Maloney and B. A. Wandell, "Color constancy: a method for recovering surface spectral reflectance," J. Opt. Soc. Am. A **3**, 29–33 (1986).
31. W. S. Stiles, G. Wyszecki, and N. Ohta, "Counting metameric object-colour stimuli using frequency-limited spectral reflectance functions," J. Opt. Soc. Am. **67**, 779–784 (1977).
32. L. T. Maloney, "Evaluation of linear models of surface spectral reflectance with small numbers of parameters," J. Opt. Soc. Am. A **3**, 1673–1683 (1986).
33. J. L. Dannemiller, "Spectral reflectance of natural objects: how many basis functions are necessary?" J. Opt. Soc. Am. A **9**, 507–515 (1992).
34. D. H. Foster, K. Amano, and S. M. C. Nascimento, "Color anisotropy for detecting violations of color constancy in natural scenes under daylight changes," Invest. Ophthalmol. Visual Sci. **42**, 3870 (2001).
35. E. K. Oxtoby, D. H. Foster, K. Amano, and S. M. C. Nascimento, "How many basis functions are needed to reproduce coloured patterns under illuminant changes?" Perception **31**, 66–66 (2002).
36. V. Cheung, S. Westland, D. Connah, and C. Ripamonti, "A comparative study of the characterisation of colour cameras by means of neural networks and polynomial transforms," Coloration Technol. **120**, 19–25 (2004).
37. D. Connah, S. Westland, and M. G. A. Thomson, "Recovering spectral information using digital camera systems," Coloration Technol. **117**, 309–311 (2001).
38. G. Hong, M. R. Luo, and P. A. Rhodes, "A study of digital camera colorimetric characterization based on polynomial modeling," Color Res. Appl. **26**, 76–84 (2000).
39. T. Johnson, "Methods for characterising colour scanners and digital cameras," Displays **16**, 183–191 (1996).
40. M. Shi and G. Healey, "Using reflectance models for color scanner calibration," J. Opt. Soc. Am. A **19**, 645–656 (2002).
41. S. Westland and C. Ripamonti, *Computational Color Science Using Matlab* (Wiley, 2004).
42. J. Parkkinen, T. Jaaskelainen, and M. Kuittinen, "Spectral representation of color images," presented at the IEEE 9th International Conference on Pattern Recognition, Rome, Italy, November 14–17, 1988.
43. G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulas* (Wiley, 1967), pp. xiv, 628.
44. P. Sumner, B. C. Regan, and J. D. Mollon, "Cambridge database of natural spectra" (2004).
45. D. I. A. MacLeod and R. M. Boynton, "Chromaticity diagram showing cone excitation by stimuli of equal luminance," J. Opt. Soc. Am. **68**, 1183–1187 (1979).
46. C. A. Párraga, G. Brelstaff, T. Troscianko, and I. R. Moorhead, "Color and luminance information in natural scenes," J. Opt. Soc. Am. A **15**, 563–569 (1998).
47. C. A. Párraga, T. Troscianko, and D. J. Tolhurst, (2000). "The human visual system is optimised for processing the spatial information in natural visual images," Curr. Biol. **10**, 35–38 (2001).

48. M. Vorobyev, personal communication,
49. A. Gilchrist, "Perceived lightness depends on perceived spatial arrangement," Science **195**, 185 (1977).
50. A. Gilchrist and V. Annan, "Articulation effects in lightness: historical background and theoretical implications," Perception **31**, 141–150 (2002).
51. A. C. Smith, H. M. Buchanan-Smith, A. K. Surridge, D. Osorio, and N. I. Mundy, "The effect of colour vision status on the detection and selection of fruits by tamarins (Saguinus spp.)," J. Exp. Biol. **206**, 3159–3165 (2003).

C