Color vision is inseparable from spatial vision. Chromatic and achromatic aspects of visual experience together subserve our perception of the forms of objects. This view is supported by physiological studies demonstrating that both color and luminance are carried along with form information on the same optic nerve fibers, albeit at different spatial scales. These scale differences can be summarized by contrast sensitivity functions measured with chromatic and achromatic spatial sinusoids, and may be illustrated by digitally filtered images that separate achromatic and chromatic variations. Analyses of the chromatic content of natural images also demonstrate a close link with the chromatic and spatial tuning of neural pathways. While characteristic properties of natural scenes can predict general characteristics of visual coding, color can vary widely across individual images, and thus could not be represented optimally by a fixed visual system. However, color coding is not fixed, but rather adjusts to both the average color and distribution of colors in scenes through processes of adaptation. Such adjustments may support color constancy and coding efficiency, and may also optimize detection and discrimination of colors that are novel in an image. Finally, the spatial properties of color-coding mechanisms are essential to our perception of figure and ground. Chromatic (border) contrast enhances the difference between figure and ground, while homogenization of object surfaces is facilitated by short- and long-range processes of assimilation and color spreading.

Keywords: color and form, color pathways, adaptation, assimilation, contrast, color constancy

1. INTRODUCTION

Color vision refers to the ability to discriminate spectral distributions irrespective of brightness. It is thought that this capacity evolved independently several times so it may not serve the same purposes in all those species that have acquired it. In normal human trichromats, color vision begins with the absorption of quanta by three different classes of cone photoreceptor. These receptor types overlap in their spectral sensitivity, but differ in their peak sensitivity at short (S), middle (M) or long (L) wavelengths. Photoreceptor responses are summed or differenced by postreceptoral pathways so that variations in intensity can be decorrelated with variations in the spectral composition of object reflectance. The summing and differencing pathways are called achromatic (or luminance) and chromatic pathways, respectively.

It is possible to describe the color of an object using terms that are disembodied from that object’s form or pattern. Concepts of redness, greenness, blueness or yellowness are normally understood without reference to specific forms. It might, therefore, be thought that chromatic vision serves functions that are separate from form vision. Indeed, it is common to assume that the primary purpose of color is to carry information about the “substance” of surfaces. Thus it is color that provides the capacity to judge the quality of foods (e.g., whether a fruit is ripe) or of conspecifics (e.g., a healthy complexion). Yet such knowledge is of questionable value if it is not tied to the spatial locations of objects, and our experience of color is far richer than the detection of infrequent sign stimuli. In this paper, we present evidence from physiology, psychophysics and perception that support a different view, that primate color vision is largely about form and object vision. Although the evidence to be presented is relatively recent, the thesis is not new. Hering noted in the last century that colors are always spatial: “Our visual world consists solely of differently formed colors... seen objects, are nothing other than colors of different kinds and forms” (p. 1).
2. CLUES FROM ELECTROPHYSIOLOGY

The same neurons that make chromatic vision possible, also provide spatial constraints for the appearance of colors. Of the three major pathways from the primate retina, two (parvocellular and koniocellular) carry information about color. The parvocellular pathway includes approximately 70% of the optic nerve fibers. Each of these fibers encodes spectral information in a spatially-dependent manner due to the antagonistic center-surround organization of their receptive fields. A single cone (say L-cone) may provide the basis for the receptive field center, while the inhibitory surround is fed by M-cones, or vice versa. These M-L or L-M receptive fields respond differently for patterns varying in luminance and those varying in chromaticity (isoluminance). As can be inferred from Figure 1, when the receptive field of such a cell is uniformly illuminated, the response will be null for broadband or luminance-varying (black-white) stimuli, but will be strong for small spots and increase until the size of the spot matches the receptive field center. The same cell will also respond to chromatic stimuli, but its response will not be attenuated by stimuli filling the center and surround. Indeed, the center and the surround of such a cell will respond synergistically to changes in spectral stimuli. These cells can extract chromatic differences that provide information related to the forms of objects that could not otherwise be detected. The net effect of this organizational scheme is that the majority of cells in primate visual pathways carry information about both chromatic and achromatic properties of images, but at different spatial scales. In general, this scheme results in chromatic processing with relatively low spatial resolution, while high spatial resolution is dependent on luminance differences.

Recent evidence shows that in the first cortical visual area, there is a transformation of signals from cells having receptive fields such as that illustrated in Figure 1 to cortical receptive fields in which the spatial selectivity is similar for chromatic and achromatic patterns. Nevertheless, as shown by the right graph in Figure 1, human contrast sensitivity is different for chromatic and luminance modulation of sinusoidal gratings. The curves show contrast sensitivity as a function of spatial frequency (number of cycles per degree of visual angle) measured with sinusoidal patterns varying only in luminance (black-white) or only in chromaticity (red-green or yellow-blue at equal luminance). The achromatic contrast sensitivity function (CSF) has a band-pass shape; the visual system is most sensitive to a band of spatial frequencies in the middle of the range. The chromatic CSF is low-pass; sensitivity is highest at low frequencies and decreases at middle and high frequencies. The high frequency limit for red-green chromatic modulation is similar to that for achromatic modulation, but that limit is typically much lower for blue-yellow chromatic modulation. This means that resolution of black-white and red-green equiluminant stimuli exceeds that with equiluminant blue-yellow stimuli. The poor resolution for blue-yellow stimuli may be attributed to, among other factors, the sparse distribution of S cones.

Figure 1: Left panel shows a cross-section of a circular receptive field profile of a retinal ganglion cell. The filled curve shows the measured receptive field profile, while the dashed curves show the spatial distribution of excitation and inhibition. In this example, excitatory input is from a long-wave-sensitive (L-) cone and inhibitory input is from middle-wave-sensitive (M-) cones. For a black-white (luminance-varying) stimulus, the cell will be maximally excited by a stimulus filling only the receptive field center; stimuli exceeding this size will be less effective due to the antagonistic surround. When tested with sinusoidal stimuli, the cell will be maximally responsive to a particular spatial frequency and less responsive to lower and higher spatial frequencies (band-pass filtering). Stimuli that change in dominant wavelength, say from middle to long wavelengths, will produce responses that both excite the L-cones through the center and reduce the inhibition from the M-cones in the surround. The net effect is that the response to chromatic variations at low spatial frequencies will not be attenuated (low-pass filtering). Right panel shows human contrast sensitivity as a function of spatial frequency (cycles per degree); solid curve for achromatic gratings and dashed curves for equiluminant red-green and blue-yellow chromatic gratings.
The spatial properties of chromatic and achromatic vision may be illustrated qualitatively with digitally-filtered images that separate the chromatic and achromatic variations, examples of which may be found in Figures 9 and 10 of Werner and Ratliff8 and Figure 1 of Gegenfurtner.9 In these images, we see much more detail (higher spatial frequencies) in the achromatic compared to the chromatic images. Indeed, the chromatic image can be defocused considerably with little deterioration of image quality. As with signals transmitted for television, most of the details in natural images are carried by lightness and darkness variations.

If the main job of color is really to tell us about the spatial properties of scenes, then why is the description carried by color so coarse? The answer is probably that it is as good as the retinal image will allow. Because of chromatic aberrations (a consequence of the fact that the refractive power of the lens varies with wavelength), fine spatial variations in chromatic contrast will be blurred on the retina.9 This problem becomes pronounced for short wavelengths and may explain the paucity of S cones. A second factor is that color necessarily requires comparisons across more than one cone class. For luminance the intermixed L and M cones can be treated as a single sampling array, yet resolution within the L and M submosaics is necessarily lower.10 Still a further factor is that the spectral sensitivities of the L and M cones are similar, with peak sensitivities that differ by only 30 nm. This similarity is probably necessary so that they can be treated equivalently for sampling luminance variations, yet it comes with the cost that the L-M difference signal on which color depends is necessarily weak, and thus may be more susceptible to noise.11 In fact, theoretical derivations of retinal receptive fields based on the same design principles but different noise levels for luminance and color closely predict the bandpass and lowpass CSFs that are actually observed.12 Finally, and perhaps most importantly, it should be emphasized that much of the function of spatial vision — of identifying objects and locations — is little compromised even by large losses in resolution. Thus the coarse spatial information provided by color may nevertheless be highly effective for the main tasks of form perception. Indeed, color appears nearly as efficient as luminance in supporting many spatial discriminations.13-15

Further evidence that color signals are used to define form may be found in a case study of a patient, M.S. He suffered a lesion in his extrastriate cortex that resulted in a loss of phenomenal color vision, a condition called achromatopsia. Achromatopsic patients describe images or scenes as appearing grey, similar to watching a black-and-white television. M.S. was completely blind in one half of his visual field and experienced achromatopsia in the other half. Tests with stimuli placed in the achromatopic region of his visual field demonstrated that he had normal S-, M- and L-cones despite being unable to use differences in cone signals to reliably discriminate colors on standard tests of color vision.16 He could, however, read color plates from the Ishihara test (made up of colored dots of different size and reflectances) if they were placed farther away (2 m), and he was able to detect some isoluminant chromatic borders even though he described the colors as appearing the same shade of gray. M.S. also had normal chromatic contrast sensitivity, mediated presumably by cells in cortical area V1. Thus, signals about chromatic properties were available to extract information about form even though they were not available for the conscious appreciation of hue. These results provide further support for the view that color vision is used for form vision.

3. COLOR STATISTICS OF NATURAL IMAGES

Important clues about the role of color in form perception can be gleaned from the kinds of signals available to the observer in natural images. Consider first that we live in a three-dimensional world filled with reflecting surfaces that are usually illuminated by a directional light source. Consequently, most natural scenes are filled with luminance variations that result from shadows and shading. Spatial variations in luminance are therefore potentially ambiguous, because they can arise either from changes in lighting or changes in surfaces. On the other hand, the color of the illuminant remains relatively (though not completely, see below) unaltered by shadows, so that regions of common color, though they may differ greatly in brightness, are more likely to belong to a common object. Thus one of the main advantages of color vision for form perception is that it may capture more directly the spatial distribution of objects independently of spatial variations in lighting. A particularly compelling example of this is the problem of detecting a fruit among foliage. The complex 3D structure of the variegated leaves clutters the image with luminance edges that effectively camouflage the luminance edges in the fruit. Yet the more uniform colors characterizing the fruit and foliage often allow the fruit to be readily detectable on the basis of chromaticity.

If the purpose of color is to discriminate and identify particular classes of objects, then the dimensions we have evolved to experience as color ought to be tuned to specific properties of the natural color environment. There is some evidence in favor of this view. Based upon samples of the spectral reflectances of objects in the visual environments of several primate species, and calculation of different possible photopigment combinations, it has been concluded that the spectral position of the L- and M-cone photopigments is optimized for detecting fruit17,18 or edible leaves19 against the background of foliage. Individuals who are lacking one class of photoreceptor, therefore, might be impaired in their discrimination of many natural
stimuli. This can be appreciated from simulations by Viénot et al.\textsuperscript{20} of the reduced color gamut of color-blind individuals, dichromats, who are missing one of the three classes of cone photoreceptors. Although these simulations do not permit us to know what dichromats perceive, one can see what happens to various forms with a reduced color gamut. The dichromat can match a natural scene that is missing much salient information about objects that can be discriminated by a trichromat. By this matching criterion, there can be no doubt that some individuals experience colors (and by implication, forms) differently from others. These studies reveal the importance of color for the perception of objects.

Most natural images do not contain all possible spatial and chromatic variations, and it is likely that visual mechanisms evolved to take advantage of these constraints. For example, analyses of the spatial statistics of natural scenes demonstrate a characteristic distribution of spatial frequencies in which the amplitudes are greatest for low spatial frequencies and fall off with increasing frequency at a rate of 1/frequency. This applies to both the achromatic\textsuperscript{21} and chromatic content of natural images.\textsuperscript{22-25} The domination of image content by low spatial frequencies implies that much of the information can be sampled by chromatic mechanisms, even though they have lower resolution. Studies of the grayscale statistics of natural scenes have provided powerful insights into the design principles underlying cortical receptive fields.\textsuperscript{26} Similar analyses of the spatial variations in color should prove equally illuminating.

If properties that are characteristic of natural images drove the evolutionary adaptations of the visual system, what of properties that instead vary idiosyncratically across different images? In this case the visual system adapts very rapidly to match coding for the prevailing scene. Color provides dramatic examples of these adjustments. For example, the color of illuminants varies widely, yet the visual system adapts to discount these variations so that the colors of objects remain relatively stable, a point we return to in the following section. Measurements of natural images also reveal a second way in which color varies across scenes. For example, Webster and Mollon\textsuperscript{27} examined color distributions for a large number of individual outdoor scenes. They specified these images in terms of cone contrasts along three postreceptoral chromatic axes thought to characterize second stage chromatic mechanisms \{M+L, L-M or S-(L+M)\}. Chromatic contrasts of most scenes tended to be biased along bluish to yellowish-green axes, but the principal color direction defining these axes varied widely, from nearly pure blue-yellow (for arid, panoramic scenes) to variations along a yellow-green axis of pure S-cone stimulation (for scenes dominated by foliage). Such color variations are too large to be encoded efficiently by mechanisms with fixed characteristics, for they often introduced strong correlations in the responses across different color mechanisms. Webster and Mollon showed, however, that the visual system adapts to the chromatic contrasts in these scenes in a manner that tends to decorrelate channel responses. This adjustment, called contrast adaptation, serves to enhance chromatic discriminations within a particular scene for objects that deviate from the prevailing gamut of colors. In this way, color mechanisms provide a basis for discrimination and identification of objects, and in particular, may serve to highlight the salience of novel objects. These adaptive processes also adjust to the spatial properties of images. For example, contrast adaptation is selective for spatial scale. Because images typically have more energy at lower spatial frequencies, we might expect the visual system to adapt to this imbalance, by reducing sensitivity to lower frequencies. In fact, something like this is observed when CSF's are measured while observers are exposed (and thus adapted) to sequences of natural scenes. Notably, the changes in sensitivity may be large and selective enough to suggest that the color CSF becomes nearly bandpass when measured under "natural" viewing conditions (i.e., while adapted to natural scenes; Webster\textsuperscript{7}).

4. COLOR INTEGRATES SPACE AND SEPARATES OBJECTS

Color can significantly enhance search and identification of information on visual displays. Human factors studies have demonstrated that in many situations, color is more effective than shape or size in helping to locate information quickly.\textsuperscript{28} The attention-getting nature of color facilitates search while at the same time provides a good basis for grouping or organizing information on a display which may help the user segregate multiple types of information and reduce clutter.

Long-range cortical interactions may be necessary to integrate color signals across space, for without them we might well see the forms of objects as skeletons (based on receptive fields responsive only to border contrast) rather than filled-in surfaces. The filling-in of color extends beyond the boundaries of the physical stimulus as can be observed over short ranges by assimilation and long ranges by the water color effect. The latter effect is observed readily when a thin colored line flanks a border on a white piece of paper. The previously white area inside the border will now be uniformly filled with a desaturated color similar to the chromatic flank. Pinna and Spillmann\textsuperscript{29} compared the effectiveness of the water color effect in determining figure-ground relations with Gestalt factors (proximity, similarity, good continuation, closure, symmetry) that are known to facilitate the perception of form. They found that when regions perceived as ground were filled in with illusory water color, the figure-ground organization reversed such that the area filled with illusory color was perceived as figure. These results provide compelling evidence of the importance of color perception in determining perception of form.
Better known than long-range assimilation, such as the water color effect, is long-range color contrast. When surfaces are delineated by chromatic or luminance edges, the opponent color is induced in surrounding or adjacent areas. Examples of this effect were described by Goethe as colored shadows, and illustrated by Monet in his haystack series. In one haystack painting, the morning light falls upon the snow, and the yellow haystacks are surrounded with blue colored shadows. In another canvas, a greenish shadow is induced by the reddish color of a haystack in the late afternoon sun. We now regard this simultaneous contrast effect as due to the reciprocal neural-opponent responses across the visual field, and to higher-level processes that parse the perceived color and lightness according to our interpretations of shading and transparency. Monet exaggerates chromatic contrast in shadows on canvas to draw attention to what normally can be perceived with no direct physical counterpart in the stimulus. It may seem from these phenomena of assimilation and contrast that the visual system is easily fooled and subject to illusion, but the mechanisms producing these effects are what make normal color vision possible. As Hering pointed out: “The most important consequences of reciprocal interactions are not at all those expressed in contrast phenomena, that is, in the alleged false seeing of ‘real’ colors of objects. On the contrary, it is precisely the so-called correct seeing of these colors that depends in its very essence on such reciprocal interactions” (pp. 123-124). In full agreement, Monet would later say: “For me, a landscape does not exist in its own right, since its appearance changes at every moment; but its surroundings bring it to life – the air and the light which vary continually.... For me, it is only the surrounding atmosphere which gives objects their true value” (p. 36). Thus, while some phenomena of assimilation and contrast do illustrate imperfect color constancy, they seldom lead to confusion about an object’s identity based on color. More generally, they support correct identification of object color by accentuating differences between object and shadow, with assimilation enhancing the uniformity of a single surface and contrast enhancing differences between figure and ground. The long-range mechanisms mediating these effects are what keep us from being fooled most of the time about the appearance of objects when the spectral content of the illumination changes – they make color constancy possible. This is the experience that the colors of most objects appear to be about the same in a wide variety of lighting conditions, even though they may reflect very different spectral distributions to the eye. Color constancy would not be possible if the visual system were not able to adjust its chromatic sensitivity as the illumination varies across space and time.

5. CONCLUSIONS

Color vision is dependent on neural pathways that concurrently carry information about form. Even achromatopsic individuals who do not report phenomenal color vision are able in some cases to use chromatic information for perception of form. The utility of chromatic vision in form perception is made apparent in scenes with variegated chromatic and spatial structure. Identification of stimuli on the basis of luminance cues is often impossible, but chromatic vision renders certain objects visible. The specific characteristics of color and spatial coding can often be predicted from characteristics of the natural visual environment. However, because natural scenes vary, a fixed visual sensitivity could not in general be optimally tuned for individual environments. The visual system is capable of adapting to the global spatial-chromatic conditions and thereby shifts its tuning to optimize detection and discrimination. Finally, color vision provides a basis for form perception by filling-in and contrast, with the former promoting uniformity within a surface and the latter facilitating the separation of figure from ground.

ACKNOWLEDGEMENTS

Supported by NIH grants AG04058 and EY10834.

REFERENCES