Three-dimensional shape perception from chromatic orientation flows

QASIM ZAIDI[1] AND ANDREA LI[2]

[1]SUNY College of Optometry, Department of Vision Sciences, New York, New York
[2]Queens College, CUNY, Department of Psychology, Flushing, New York

Abstract

The role of chromatic information in 3-D shape perception is controversial. We resolve this controversy by showing that chromatic orientation flows are sufficient for accurate perception of 3-D shape. Chromatic flows required less cone contrast to convey shape than did achromatic flows, thus ruling out luminance artifacts as a problem. Luminance artifacts were also ruled out by a protanope's inability to see 3-D shape from chromatic flows. Since chromatic orientation flows can only be extracted from retinal images by neurons that are responsive to color modulations and selective for orientation, the psychophysical results also resolve the controversy over the existence of such neurons. In addition, we show that identification of 3-D shapes from chromatic flows can be masked by luminance modulations, indicating that it is subserved by orientation-tuned neurons sensitive to both chromatic and luminance modulations.

Keywords: Three-dimensional shape, Color, Orientation flows

Introduction

One of the remarkable abilities of the visual system is the accurate perception of 3-dimensional (3-D) shapes in 2-dimensional (2-D) monocular images. The role of chromatic information in this process is controversial. Psychophysical studies disagree on whether purely chromatic information is sufficient for 3-D shape perception (Cavanagh, 1991; Troscianko et al., 1991). To provide a definitive answer, we designed this study on the basis of recent results showing that qualitatively accurate perception of 3-D shape depends critically on the visibility of orientation flows (Breton et al., 1992; Li & Zaidi, 2000, 2001; Ben-Shahar & Zucker, 2001; Zaidi & Li, 2002; Fleming et al., 2004). The extraction of chromatic orientation flows from retinal images requires neurons that respond to chromatic modulations and are also selective for local orientation. There is disagreement in the neurophysiological literature about the existence of such neurons (Kiper, 2003). Cells in striate cortex that prefer color modulations show little or no specificity for orientation (Livingstone & Hubel, 1984, 1988; Ts'o & Gilbert, 1988), but cells that respond to both chromatic and luminance modulations do show orientation selectivity (Leventhal et al., 1995; Gegenfurtner et al., 1997; Kiper et al., 1997; Johnson et al., 2001; Shapley & Hawken, 2002; Friedman et al., 2003). Recent psychophysical work also suggests the existence of mechanisms simultaneously selective for chromaticity and orientation

(Flanagan et al., 1990; Pearson & Kingdom, 2002; Clifford et al., 2003). We resolve both psychophysical and neural controversies by showing that 3-D shape can be perceived accurately from purely chromatic orientation flows, but that the inferred neural mechanisms are responsive to both chromatic and luminance modulations.

Consider the case of a sinusoidal corrugation carved from a solid with the surface pattern repeated in depth. Fig. 1a shows a solid of this type with the line to be carved drawn on the top. Figs. 1b and 1d show two different front-surface patterns on uncarved solids. For the two patterns, Figs. 1c and 1e show perspective projections of front surfaces of solids carved into sinusoidal corrugations. Distortions of the surface patterns are visible in both perspective images. It is easy to perceive a 3-D corrugation with a central concavity in Fig. 1c, even though the image is presented on a flat surface, without stereo, motion, occlusion, or shading cues. However, it is not possible to see the 3-D shape in Fig. 1e. What is the reason that the surface pattern in Fig. 1b can support 3-D shape perception, and that in Fig. 1d cannot?

This question can be answered by analyzing the pre-carved surface patterns. The surface patterns are composed of oriented sinusoidal gratings of identical contrast added in random phase with respect to each other (Fig. 2a). The pattern in Fig. 1b contains six orientations equal to 0, 45, 67.5, 90, −67.5, and −45 deg (0 deg is the horizontal component). The pattern in Fig. 1d contains eight orientations: all of the above and ±22.5 deg. Below each grating, Fig. 2b shows the perspective projection of a corrugated surface with that component. We have shown previously that the

Address correspondence and reprint requests to: Qasim Zaidi, SUNY College of Optometry, Department of Vision Sciences, 33 W 42nd Street, New York, NY 10036, USA. E-mail: qz@sunyopt.edu
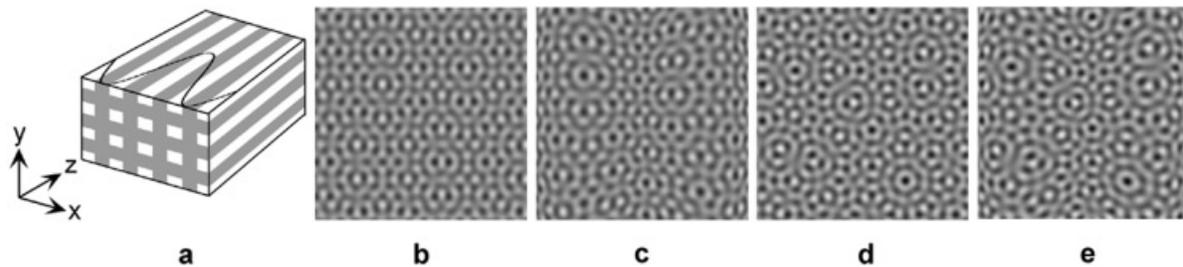
Fig. 1. (a) A solid, simulated by layering identical planar patterns along the depth axis, was carved sinusoidally in depth along the curved line shown on the top face. (b) Six-component surface pattern; sinusoidal components at the orientations shown in Fig. 2(a), except ±22.5 deg. (c) Perspective image of sinusoidal corrugation with six-component surface pattern. (d) Eight-component surface pattern containing all components shown in Fig. 2(a). (e) Perspective image of sinusoidal corrugation with eight-component surface pattern.

critical clues for perceiving the correct curvatures and slants of carved, folded, and stretched corrugations are provided by the orientation flows of the 0-deg component (Li & Zaidi, 2000, 2001, 2003, 2004; Zaidi & Li, 2002). These flows distinctly identify 3-D shape features: they bow inward towards the center of the image at concavities, outward at convexities, converge to the right at rightward slants, and converge to the left at leftward slants. The 0-deg component is parallel to the axis of maximum intrinsic curvature of the corrugated surface. For the upright corrugation, this is also the axis of maximum curvature modulation with respect to the observer.

Orientation modulations of the 0-deg component are physically present in both Figs. 1c and 1e. They are not visible in Fig. 1e because they are masked by orientation modulations of the ±22.5-deg components. (Notice similarities in the orientations in Fig. 2b). As a result, the 3-D corrugation is not perceived correctly in Fig. 1e by most observers. The masking effect can be demonstrated in three ways. The critical flows become visible and the corrugation is perceived correctly, first, if the ±22.5-deg components are removed from the pattern (Li & Zaidi, 2004) (Fig. 1c); second, if the contrast of the 0-deg component is increased relative to the other components (Fig. 3a); and third, if the 0-deg component is a colored grating (Fig. 3b).

We exploit established knowledge about early physiological stages of the visual system to understand the role of color in 3-D perception. For analyzing the colors of the environment, the 3-D space of cone responses (L, M, S) is transformed at the ganglion cell level into $\pm(L+M)$, $\pm(L-M)$ and $\pm(L+M-S)$ signals (Der-

rington et al., 1984; Sun et al., 2006). Variations in light intensity have proportional effects on the responses of the three cone types, so the difference signals ($L-M$ and $L+M-S$) remove common intensity variations, are correlated with spectral variations in the physical input, and are perceived as pure chromatic variations. The $L+M$ signal is correlated with physical intensity variations and is defined as luminance (Wagner & Boynton, 1972; Anstis & Cavanagh, 1983; Lee et al., 1988). In this study, we compare the efficacy of red–green patterns that provide $L-M$ modulations at constant $L+M$ and $L+M-S$ levels, to the efficacy of achromatic modulations that vary all three cone signals proportionately, thus providing $L+M$ (luminance) modulations at constant $L-M$ and $L+M-S$ levels. To show that chromatic orientation flows can convey 3-D shape requires that all luminance artifacts be ruled out.

Whether an orientation- or direction-selective neuron also responds to chromatic modulation can be determined by titrating the luminance ratio of the end-points of high contrast color modulation (Gegenfurtner et al., 1994). To create an extended isoluminant image for testing whether percepts of 3-D shape can result from chromatic information, however, is quite different from nulling luminance contrast for a neuron with a small receptive field. A number of procedures (Wagner & Boynton, 1972; Anstis & Cavanagh, 1983) can be used to make putatively isoluminant red–green versions of Fig. 1c, and such images do convey 3-D shape, although the image tends to fade because the total cone contrast in each component is about 1%. However, there is no psychophysical procedure whose minimum settings will set the
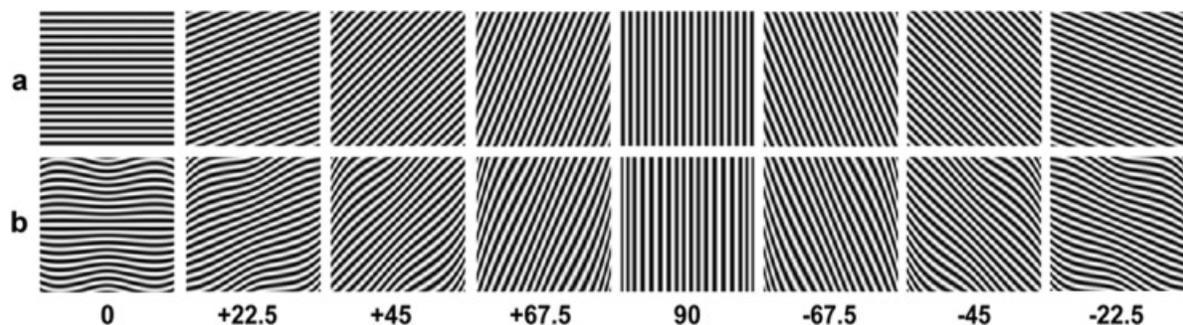
Fig. 2. (a) Orientations of component gratings used to make surface patterns. (b) Perspective images of sinusoidal corrugations with component gratings as surface patterns.
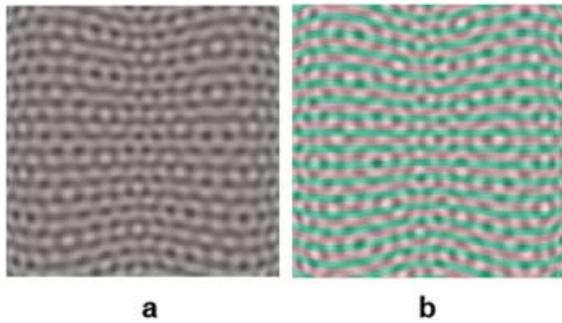
Fig. 3. Perspective images of eight-component sinusoidal corrugation from Fig. 1(c) with (a), the contrast of the 0-deg component increased by a factor of 3, and (b) a chromatic 0-deg component.

whole image to isoluminance. Macular pigment density and cone ratios differ between the fovea and periphery (Zaidi et al., 1989) leading to variations in isoluminance across the retina. In addition, chromatic aberrations due to the optics of the eyes add luminance modulations to what would otherwise be pure chromatic modulations (Flitcroft, 1989). Consequently, luminance artifacts cannot be ruled out as the source of the shape cues in the isoluminant image.

A method that unequivocally rules out luminance artifacts is to find tasks for which chromatic information is superior to luminance information of equal or greater cone contrast. We show that, for six-component patterns like Fig. 1b, less total cone contrast is needed to judge a 3-D shape accurately if the 0-deg luminance component is replaced by a chromatic component. We rule out luminance artifacts in two additional ways. First, the advantage of chromatic information is enhanced when ±22.5-deg luminance masking components are added to the pattern. Second, a protanope is unable to see 3-D shape from chromatic flows.

Materials and methods

Volumetric solids were simulated by layering identical planar patterns along the depth axis. Each solid was then carved sinusoidally in depth as a function of horizontal position with respect to the observer. Perspective images of 1.5 cycles of the sinusoidal corrugation were presented on a CRT monitor at a distance of 1 m. When viewed monocularly, the retinal image coincided with that of a real 3-D sinusoidally curved surface with peak-to-trough amplitude of 14 cm and wavelength of 10 cm. To restrict the shape cues solely to texture variations in the image, all surfaces were presented in fronto-parallel view, without occluding contours. The room lights were turned off during the experiments, and observers' heads were fixed in a chinrest at the proper eye-height and viewing distance for the perspective projections.

Stimuli were generated using Matlab, and presented on a SONY GDM-F500 flat screen monitor with a 800×600 pixel screen running at a refresh rate of 100 frames/s *via* a Cambridge Research Systems Visual Stimulus Generator (CRS VSG 2/3) controlled through a 400 MHz Pentium II PC. Through the use of 12-bit DACs, after gamma correction, the VSG was able to generate 2861 linear levels per gun. Each image was 267×267 pixels.

The patterns used to form the layered solid were the sum of eight (Fig. 1d) or six (Fig. 1b) sinusoidal gratings. Each image subtended 9×9 deg of visual angle. Each sinusoidal grating had

a frequency of 2 cycles/deg. To be able to set the contrast and chromaticity of the 0-deg component independently from the other texture components, the image of the 0-deg component was generated separately from an image containing the other components. In the isolated frame, the cone coordinates of the mid-white were $(L, M, S)/(L+M) = (0.636, 0.364, 0.02)$. The coordinates of the extreme reddish and greenish values were (0.670, 0.330, 0.02) and (0.602, 0.398, 0.02), respectively. The two images were interleaved in alternating frames at 100 Hz, at which rate the flicker was invisible and the images appeared completely merged. As a result of the interleaving, the contrast in the merged image was half the contrast in each frame. For simplicity of exposition, the contrast units used in this paper are for the isolated frames. The mean luminance of all stimuli was 44 cd/m^2.

Surfaces were presented in one of four phases of the sinusoidal corrugation at the central vertical mid-line of the image: a concavity, a convexity, a right slant, or a left slant. Using a 3-button response box, observers indicated whether the shape of the central 2- $\times$ 9-deg region of each image, delineated by a dark grey rectangular outline, appeared concave, convex, slanted to the right, slanted to the left, or flat and fronto-parallel. Viewing was monocular, with unlimited time and no feedback. Data for parts of Experiment 1, 2, and 3 were collected in a single session. Each session contained images for each of the 4 shapes $\times$ 2 texture patterns (achromatic eight-component pattern, achromatic six-component pattern) $\times$ 5 contrast levels of the 0-deg component, and 4 shapes of the eight-component pattern with a full-contrast chromatic horizontal component, for a total of 44 conditions. Each condition was presented ten times, randomly interleaved, for a total of 440 trials. The rest of the data was collected in a single session a few weeks later. This session contained images for each of 4 shapes $\times$ 2 patterns (eight-component pattern with chromatic 0-deg component, six-component pattern with chromatic 0-deg component) $\times$ 5 contrast levels of the 0-deg component, for a total of 40 conditions. Each condition was presented ten times, randomly interleaved, for a total of 400 trials. Each session was preceded by 1 min of adaptation to the whole screen set at the mid-grey. An initial practice session containing three trials of each condition helped observers get acquainted with the task. Percent correct of responses were plotted against the contrast of the 0-deg component, and fitted with Weibull functions. Thresholds were determined conservatively at 62.5% correct (since there were only four different shapes presented).

Six observers participated in these experiments. Observers were not informed about the purposes of the study until all the data had been collected. All had normal or corrected-to-normal acuity. Five of the observers had normal color vision. Only three of the color-normal observers were available for the second set of measurements. Observer AC was a protanope, as determined by the Ishihara color plates, and the Nagel anomaloscope. On the anomoloscope, he was able to match the most saturated red and green by modulating the luminance of a yellow field (luminance settings = 2.3 and 30 for red and green, respectively).

In studying 3-D shape from monocular cues, there is always a danger that observers learn to associate certain 2-D cues with certain responses, instead of actually perceiving 3-D shapes. We designed the study to avoid these problems. Many of the observers had no prior psychophysical experience, and observers were given no feedback about the correctness of their responses. Therefore, in the 5AFC task, they had no way of associating particular 3-D flows with the correct 3-D responses, unless they actually perceived the 3-D shapes.

Results

Experiment 1: 3-D shape from chromatic orientation flows

We measured the threshold contrast of the 0-deg component required for accurate perception of 3-D shapes. The 0-deg component was either an achromatic grating, or an isoluminant reddish–greenish grating. The other components were achromatic gratings. In a method of constant stimuli, the contrast of the 0-deg component was varied in five steps while the contrast of each of the other components was kept constant at 14.28% (the maximum contrast possible for a seven-component image). There is no perfect method to compare the strength of chromatic to luminance inputs. However, if inputs are equated at the photoreceptor level, any differences in the efficiency of a chromatic versus a luminance system are due to postreceptoral factors (Sachtler & Zaidi, 1992; Scharff & Geisler, 1992). Therefore, for comparing isoluminant red–green to achromatic gratings, a metric is provided by adding the absolute values of contrast for each cone type after weighting them by the proportion of each type of cone in the retina: $P_L * |\text{L contrast}| + P_M * |\text{M contrast}|$, where the ratio of proportions $P_L : P_M = 2:1$. For the six-component pattern, the black bars in Fig. 4a show the ratio of luminance to chromatic contrasts of the 0-deg component required for correct shape identification. The light bars do the same for the eight-component patterns. The three observers required between 2.23 to 5.27 times more total cone contrast to perceive 3-D shape accurately with the luminance component than with the chromatic component. The chromatic grating is equal to an L-cone grating added to an opposite phase M-cone grating. The maximum possible luminance artifact would occur if optical or neural processes had the effect of shifting the phases of the L and M cone modulations to form an L+M grating. In that case, the contrast of the luminance artifact would be identical to the total cone contrast of the chromatic grating. The actual luminance artifact due to optical, pre-retinal or neural processes will be a small fraction of the worst-case value. The results show that even in the worst possible case, the luminance artifact at the shape threshold for the red–green flows would be 2.23 to 5.27 times less than the luminance contrast required to perceive shapes with the same accuracy from the achromatic flows. Luminance artifacts thus cannot be responsible for 3-D shape perception from chromatic flows.

Masking of achromatic and chromatic orientation flows

In visual cortex, two classes of neurons respond to pure chromatic modulation: neurons preferentially tuned to equiluminant chromatic modulation are either cosine-tuned or narrower and do not respond to pure luminance modulation, whereas neurons that respond best to some mixture of chromatic and luminance stimulation, also respond to pure luminance and pure chromatic modulations (Johnson et al., 2004). Accurate 3-D shape identification from chromatic flows, in itself, does not distinguish which of the two classes of neurons subserves this performance. For this distinction, we make the reasonable assumption that the response of the first class of neurons should be immune to superposition of luminance modulations, but the response of the second class, even to chromatic modulation, should be susceptible to masking by luminance modulations at neighboring orientations.

The magnitude of masking was calculated as the threshold contrast of the 0-deg component required to make correct shape judgments for the pattern consisting of all eight components
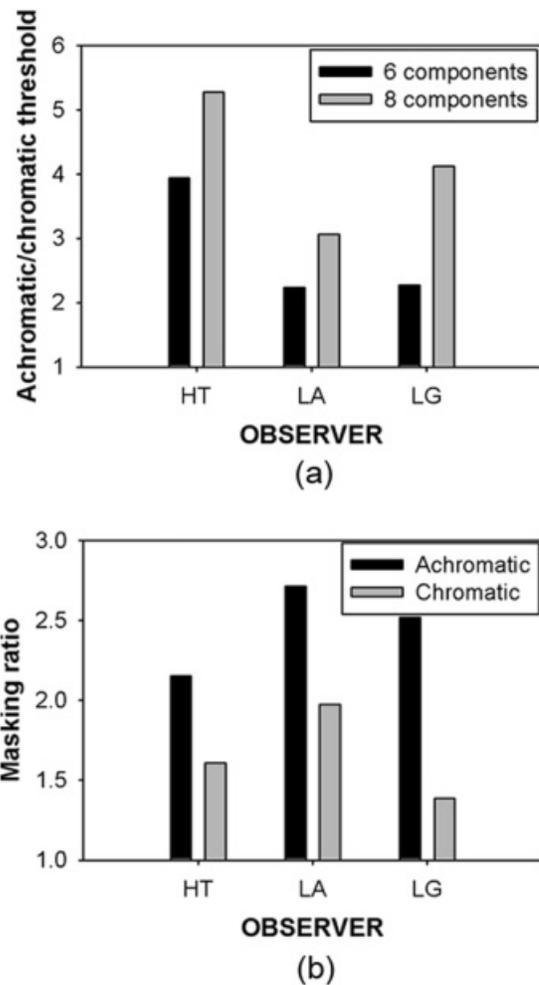
Fig. 4. (a) Ratios of luminance to chromatic cone contrast thresholds required to make correct shape identifications for the six-component corrugation (black) and the eight-component corrugation (grey). (b) Masking ratios for achromatic stimuli and for stimuli in which flows are defined by color, for three observers. In both conditions, 1.0 indicates no masking by achromatic components at neighboring orientations.

(Fig. 1d), divided by the threshold contrast for the pattern without the ±22.5-deg components (Fig. 1b). A masking ratio greater than one would indicate masking of the 0-deg orientation modulations by the ±22.5-deg orientation luminance modulations. Fig. 4b shows the masking results for the achromatic and chromatic orientation flows. For all three observers, the ratios were greater than one. The ratios for chromatic orientation flows, 1.4 to 2.0, were significantly greater than 1.0, but decidedly lower than the ratios for luminance flows, 2.1 to 2.7. These results have two implications. First, if shape identification from chromatic flows was due solely to luminance artifacts, the achromatic and chromatic masking ratios would have been similar (This requires only that the masking effect have the same multiplier on all luminance signals, whether they be artifacts created by the visual system or physically present in the stimulus). Contrary to this expectation, after masking by luminance modulations at neighboring orientations, chromatic flows were even more efficient at conveying 3-D shapes than were luminance flows. Second, if the neurons that extract chromatic flows for shape perception were insensitive to luminance

modulations, their responses would not be affected by luminance modulations at neighboring orientations. The fact that chromatic orientation flows are masked by luminance modulations, indicates that neurons that extract chromatic flows are sensitive to both chromatic and luminance modulations. Under the simplifying assumption that the responses of most neurons are proportional to the projections of a stimulus vector (in chromatic + luminance space) on their preferred vectors, chromatic and luminance flows would be extracted by overlapping but different populations of neurons. The masking luminance modulations will be closer in color angle to the populations of neurons extracting luminance flows than to the neurons extracting chromatic flows. Consequently, the result that the masking ratio for chromatic flows has a lower magnitude than the ratio for luminance flows suggests that the relative masking effects of a stimulus on the responses of neurons tuned to different directions, are proportional to the angles between the vector representing the masking stimulus and the vectors representing the preferred stimuli of each of the neurons. A similar scheme could also account for the results of cross-masking of chromatic and luminance gratings of the same spatial frequency and orientation (Bradley et al., 1988).

Experiment 2: Inability of a protanope to perceive 3-D shape from chromatic orientation flows

To further rule out any luminance artifact, we tested whether a protanope could use red–green chromatic modulations to perceive 3-D shape. This experiment used only the eight-component pattern. As a baseline, in Fig. 5a, the threshold contrasts of the 0-deg achromatic component required for correct shape identification, measured in Experiment 1, are shown for all six observers. The other seven components, as described before, were luminance gratings at contrasts of 14.28%. The threshold cone contrasts range from 13% to 28%, so that only two out of six observers could identify shape correctly in forced-choice trials if the contrast of the 0-deg luminance component was equal to that of the other components, that is, 14.28%.

For the new test, on the critical trials, the 0-deg luminance component was replaced with an isoluminant red–green grating fixed at the highest cone-contrasts possible for our monitor. The maximum luminance artifact possible from this grating was 6.6% contrast. This value was roughly half the shape-identification luminance threshold for the most sensitive observer. The chromatic-flow stimuli were randomly interleaved with the luminance threshold measurements. As in Experiment 1, there were four shape features and five possible responses. If an observer could perceive a 3-D surface, but could not identify specific features correctly, performance would be at 25% correct. If an observer could not see the surfaces as three-dimensional, and reported them as fronto-parallel flat, performance would be at 0% correct. Fig. 5b shows performance in the shape identification task for the six observers. The five observers on the left had normal color vision, and performed well above the dotted line, which indicates the 25% correct level. Across observers, percent corrects were similar for different shapes. Observer AC is a protanope, missing the L-cone pigment. He reported that all the stimuli appeared achromatic. He could identify shapes correctly if the 0-deg component had luminance contrast above 24%. However, if the 0-deg component had only chromatic contrast, he identified nearly all surface shapes as flat. This result indicates that pure luminance tuned neurons are not sufficient to extract chromatic flows from our red–green

(a)

(b)
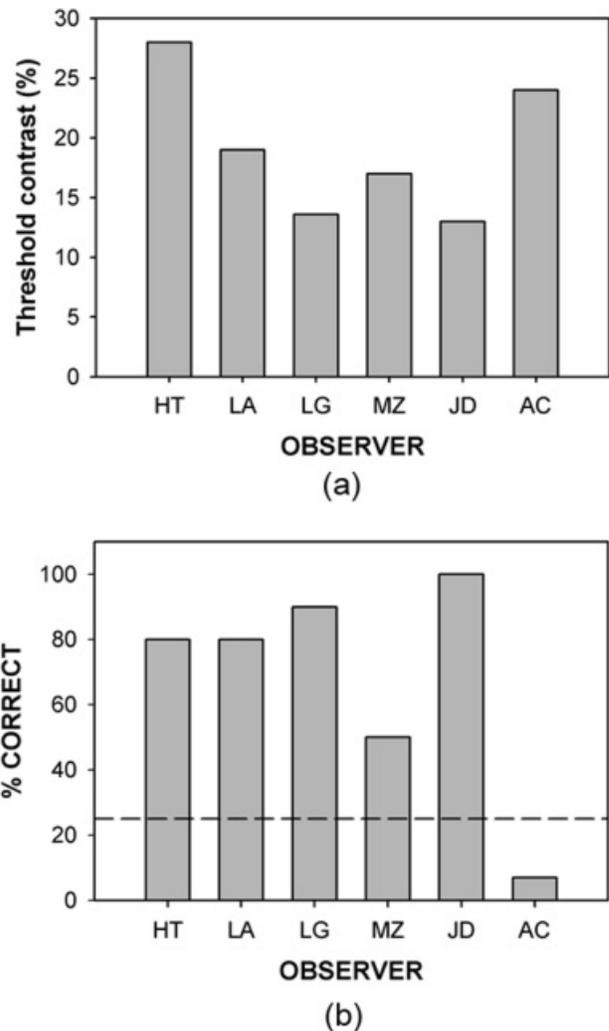
Fig. 5. (a) Contrast thresholds of achromatic horizontal component required for correct shape identifications in eight-component pattern for six observers. (b) Performance in shape identification task for stimuli in which orientation flows are defined by maximum contrast chromatic modulation. All observers except AC, who is a protanope and for whom the orientation flows were invisible, were able to correctly identify 3-D shapes above chance (dashed line). Observer AC reported nearly all surfaces as flat.

stimuli, and strengthens the conclusion that color-normal observers identified the correct shapes by using chromatic neural signals.

Discussion

The main conclusion of this study is that purely chromatic orientation flows can provide sufficient information for observers to accurately perceive 3-D shape from texture variations. The second conclusion of this study is that the neurons that extract chromatic orientation flows are sensitive to both chromaticity and luminance. The individual outputs of such neurons cannot distinguish between a chromatic stimulus and an achromatic stimulus. Consequently, the results of this study raise four interesting questions to which present knowledge provides only partial, but suggestive, answers:

(1) What are the functions served by neurons most sensitive to combinations of chromaticity and luminance as opposed to those

served by neurons that prefer pure chromatic or luminance stimulation? By transforming cone outputs into L+M, L−M, and L+M−S signals, ganglion cells act as minimally correlated linear projectors providing efficient coverage of 3-D cone space and good information transmission (Buchsbaum & Goldstein, 1979; Zaidi, 1997). In addition, separation of signals does have some utility, even if there are hardly any purely chromatic or luminance variations in the real world. In evolution there was a real advantage in finding reddish/yellowish fruit against green foliage (Regan et al., 1998; Parraga et al., 2002) without being distracted by the luminance and yellowish–bluish spatial noise caused by leaves varying in angles to the sun [sunlight is brighter and yellower than skylight (Taylor & Kerr, 1941)]. Moreover, boundaries between materials have both chromatic and luminance components, but shadows and shading create mainly luminance variations in the retinal image, though there will also be a yellowish–bluish component (Taylor & Kerr, 1941). It has been claimed that the visual system uses chromatic variations as cues to variations in surface reflectance, so that only those luminance variations that are uncorrelated with chromatic variations are used for shape-from-shading or shadow identification (Kingdom, 2003; Kingdom et al., 2004).

Striate cortex has orders of magnitude more neurons than there are ganglion cells, so questions of efficiency have less to do with encoding and transmission, and more to do with decoding and representation. To understand why the preferred directions of neurons are distributed widely over 3-D color space (Lennie et al., 1990; Johnson et al., 2004), we consider an analogy with orientation processing. Any possible orientation in the retinal image could be computed from the vector sum of the outputs of just two classes of linear orientation projectors with orthogonal receptive fields. However, the vector sum gives just one orientation, whereas there are a number of visual percepts that require the extraction of multiple orientations per retinotopic location. This can only be done by processing the ouputs of neurons with preferred responses distributed over the complete range of orientations. For example, observers can see multiple orientations at the same retinotopic location. Such stimuli evoke multiple peaks in the population response of a hyper-column of cells tuned to the complete range of orientations (Hubel & Wiesel, 1962, 1968; Ohki et al., 2005). At the least, the percept requires neural circuits that detect distinct peaks, and it is even possible that the overall shape of the population response has to be used (Treue et al., 2000). Similarly, textures can be segmented on the basis of orientation information, for example, the six-component pattern in Fig. 1b from the eight-component pattern in Fig. 1d. The first pattern will evoke six peaks in the population response of each hyper-column, whereas the second will evoke eight. Since these patterns differ only in orientation composition, they can only be differentiated by comparing the two population responses. Further, multiple surface curvatures can be seen simultaneously at the same retinal region when they are defined by distinct orientation flows (Li & Zaidi, 2004). Two higher-level cells responding simultaneously to the presence of distinct orientation flows thus also need distinct orientation signals from the same retinotopic locations.

There are also several reasons why it is useful to have cortical cells tuned to the complete range of directions in 3-D cone space, despite the fact that no new local color information can be added to that provided by lateral geniculate nucleus (LGN) cells. First, analyses of natural scenes reveal a preponderance of combined chromatic and luminance edges (Caywood et al., 2004). In general, cells tuned to pure chromatic or pure luminance stimulation will

not give vigorous responses to oriented edges, so orientation flow extraction will be deficient unless there are cells tuned preferentially to intermediate color directions. Second, the stimuli that evoke vigorous neural responses increase in complexity as signals proceed through areas of the visual cortex (Kobatake & Tanaka, 1994; Op de Beeck et al., 2001; Brincat & Connor, 2004). As a consequence, cortical processing progressively reduces the number of neurons whose outputs are sufficient to decode what is being seen. For this purpose, the efficiency of schemes for tiling input space (e.g. orientation or color space) by receptive fields covering discrete volumes of the space depends partially on what statistics of the population response are used as the output. For example, the density of receptive fields should be proportional to the input probability distribution function in every unit volume of input space if the peak of the population response is the effective statistic (Linsker, 1989). In chromatic + luminance space, this scheme would require that a majority of neurons be preferentially tuned to mixtures of chromaticity and luminance variations.

(2) Why is the visual system more sensitive to chromatic orientation flows than to luminance flows? The relative efficacy of chromatic and luminance signals depends on the spatial and temporal frequencies contained in the stimulus (Thorell et al., 1984). In the experiments in this paper, the gratings forming the orientation flows had a spatial frequency of 2.0 cycles/deg with no limit on viewing time. In the past we have found that 500-ms presentations were sufficiently long for observers to make accurate shape decisions. For a temporal frequency of 1.0 Hz, visual cortical areas V1 and V2 show much stronger fMRI (BOLD) responses to chromatic contrast reversing checkerboards than to luminance checkerboards (Engel et al., 1997). If the BOLD response represents inputs to a cortical area (Logothetis, 2003), then the fMRI results suggest that the population responses of both LGN and V1 are strongest for chromatic stimulation and weakest for luminance at the predominantly low spatial frequencies contained in the checkerboards. Since LGN is dominated by Parvo-cells (Wassle & Boycott, 1991), which are most responsive to chromatic modulation at low temporal frequencies, chromatic signals could stimulate larger population responses in cortical areas that receive input predominantly from the Parvo-cells; for example, areas V2 and V3 in the ventral stream (Gegenfurtner et al., 1997; Kiper et al., 1997).

(3) If chromatic orientation flows are extracted by cells tuned to combinations of luminance and chromatic modulations, should chromatic flows look colored, achromatic, or some mixture of colored and achromatic? In the experimental displays, chromatic flows appeared to be similar to isolated equiluminant red–green patterns (the reproduced colors in Fig. 3 are unlikely to be close to isoluminant). This suggests two possible explanations. First, the color of the flow could be computed as the peak or vector average of the population distribution of the orientation tuned neurons that extract the flow. Even though the class of orientation tuned cells does not include neurons that are preferentially tuned to pure chromatic variations (Johnson et al., 2004), the vector average of the population response to chromatic flows will be in the chromatic direction. A second possibility is that pattern matchers for 3-D curvature signal the shape, but not the color of the shape. Instead, the perceived color of every point in the image is calculated separately at every retinotopic location from a population response that includes nonoriented, spatially low-pass cells tuned to pure chromatic variations. Given that V1 neurons tuned to the chromaticity of the stimulus will give the largest response to chromatic modulation, a colored appearance results from simple population statistics like the peak of the population distribution. It

is worth making explicit, that the second possibility does not imply that any stage of 3-D shape processing is achromatic (Delorme et al., 2000). Even if the computations of 3-D curvature and color appearance are dissociated at some stage past the striate cortex, neurons responsive to 3-D shape will respond to chromatic signals even at the shortest latencies, consistent with the recordings of inferotemporal cortex neurons that respond to complex shapes (Edwards et al., 2003). Judicious sets of texture patterns could be used in single-cell electrophysiology to decide whether the colors of flows are subserved just by cells preferring pure chromatic modulations, or whether inputs from cells preferring combinations of luminance and chromaticity are also important.

(4) How critical are orientation flows to the perception of 3-D shape from texture variations? Images can be parsed into orientation variations and spatial-frequency variations. In the absence of orientation modulations, observers do perceive 3-D shapes from frequency modulations. The problem is accuracy of percepts. For a fixed texture pattern, changes in frequencies in the projected image can be due either to relative slants or to relative distances of segments of the 3-D surface. For example, for surfaces where intrinsic depth changes are small compared to distance from the observer, if the texture on the surface is statistically homogeneous, frequency modulations in the retinal image are mostly a function of surface slant (Li & Zaidi, 2004). On the other hand, if the same surface is carved as in Fig. 1a, the spatial frequency on the surface is a decreasing function of slant, and frequency modulations in the image are due almost entirely to relative distance (Li & Zaidi, 2004). We have previously shown that observers make systematic mistakes in judging 3-D shapes defined by frequency modulations. These mistakes indicate that the visual system acts on the potentially erroneous assumption that frequency modulations are a function of relative distance, but not of relative slant (Li & Zaidi, 2000, 2001; Zaidi & Li, 2002).

Orientation flows have been shown to be critical not just in the perception of 3-D shape from texture (Knill, 2001), but also in the perception of 3-D shape from shading (Breton et al., 1992; Ben-Shahar & Zucker, 2001), and reflections (Fleming et al., 2004). Another major clue to 3-D shape that requires orientation extraction, is the curvature of occluding contours (Tse, 2002). The mathematics of 3-D shape allow an observer to infer with certainty that convexities in a 2-D projection are due to local ovoid shapes and concavities to saddle shapes (Koenderink, 1984). Consequently, it will be worth testing these other cues with chromatic flows to see whether chromatic signals and orientation-tuned, chromatically sensitive neurons play a general role in 3-D shape perception.

Acknowledgments

References

ANSTIS, S. & CAVANAGH, P. (1983). A minimum motion technique for judging equiluminance. In *Colour Vision: Psychophysics and Physiology*, eds. MOLLON, J.D. & SHARPE, L.T., pp. 155–166. London: Academic Press.

BEN-SHAHAR, O. & ZUCKER, S.W. (2001). On the perceptual organization of texture and shading flows: From a geometrical model to coherence computation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR01)*. Hawaii: Kauai.

BRADLEY, A., SWITKES, E. & DE VALOIS, K. (1988). Orientation and spatial frequency selectivity of adaptation to color and luminance gratings. *Vision Research* **28**, 841–856.

BRETON, P., IVERSON, L., LANGER, M. & ZUCKER, S.W. (1992). Shading flows and Scenel bundles: A new approach to shape from shading. In *Second European Conference on Computer Vision (ECCV92)*, ed. SANDINI, S., pp. 135–150. New York: Springer-Verlag.

BRINCAT, S.L. & CONNOR, C.E. (2004). Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nature Neuroscience* **7**, 880–886.

BUCHSBAUM, G. & GOLDSTEIN, J.L. (1979). Optimum probabilistic processing in colour perception. I. Colour discrimination. *Proceedings of the Royal Society B* (London) **205**, 229–247.

CAVANAGH, P. (1991). Vision at equiluminance. In *Vision and Visual Dysfunction Volume V: Limits of Vision*, eds. KULIKOWSKI, J.J. & WALSH, V., pp. 234–250. Boca Raton, Florida: CRC Press.

CAYWOOD, M.S., WILLMORE, B. & TOLHURST, D.J. (2004). Independent components of color natural scenes resemble V1 neurons in their spatial and color tuning. *Journal of Neurophysiology* **91**, 2859–2873.

CLIFFORD, C.W., SPEHAR, B., SOLOMON, S.G., MARTIN, P.R. & ZAIDI, Q. (2003). Interactions between color and luminance in the perception of orientation. *Journal of Vision* **3**, 106–115.

DELORME, A., RICHARD, G. & FABRE-THORPE, M. (2000). Ultra-rapid categorisation of natural scenes does not rely on colour cues: a study in monkeys and humans. *Vision Research* **40**, 2187–2200.

DERRINGTON, A.M., KRAUSKOPF, J. & LENNIE, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *Journal of Physiology* **357**, 241–265.

EDWARDS, R., XIAO, D., KEYSERS, C., FOLDIAK, P. & PERRETT, D. (2003). Color sensitivity of cells responsive to complex stimuli in the temporal cortex. *Journal of Neurophysiology* **90**, 1245–1256.

ENGEL, S., ZHANG, X. & WANDELL, B. (1997). Colour tuning in human visual cortex measured with functional magnetic resonance imaging. *Nature* **388**, 68–71.

FLANAGAN, P., CAVANAGH, P. & FAVREAU, O.E. (1990). Independent orientation-selective mechanisms for the cardinal directions of colour space. *Vision Research* **30**, 769–778.

FLEMING, R.W., TORRALBA, A. & ADELSON, E.H. (2004). Specular reflections and the perception of shape. *Journal of Vision* **4**, 798–820.

FLITCROFT, D.I. (1989). The interactions between chromatic aberration, defocus and stimulus chromaticity: Implications for visual physiology and colorimetry. *Vision Research* **29**, 349–360.

FRIEDMAN, H.S., ZHOU, H. & VON DER HEYDT, R. (2003). The coding of uniform colour figures in monkey visual cortex. *Journal of Physiology* **548**, 593–613.

GEGENFURTNER, K.R., KIPER, D.C., BEUSMANS, J.M., CARANDINI, M., ZAIDI, Q. & MOVSHON, J.A. (1994). Chromatic properties of neurons in macaque MT. *Visual Neuroscience* **11**, 455–466.

GEGENFURTNER, K.R., KIPER, D.C. & LEVITT, J.B. (1997). Functional properties of neurons in macaque area V3. *Journal of Neurophysiology* **77**, 1906–1923.

HUBEL, D.H. & WIESEL, T.N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology* **160**, 106–154.

HUBEL, D.H. & WIESEL, T.N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology* **195**, 215–243.

JOHNSON, E.N., HAWKEN, M.J. & SHAPLEY, R. (2001). The spatial transformation of color in the primary visual cortex of the macaque monkey. *Nature Neuroscience* **4**, 409–416.

JOHNSON, E.N., HAWKEN, M.J. & SHAPLEY, R. (2004). Cone inputs in macaque primary visual cortex. *Journal of Neurophysiology* **91**, 2501–2514.

KINGDOM, F.A. (2003). Color brings relief to human vision. *Nature Neuroscience* **6**, 641–644.

KINGDOM, F.A., BEAUCE, C. & HUNTER, L. (2004). Colour vision brings clarity to shadows. *Perception* **33**, 907–914.

KIPER, D.C. (2003). Colour and form in the early stages of cortical processing. *Journal of Physiology* **548**, 335.

KIPER, D.C., FENSTEMAKER, S.B. & GEGENFURTNER, K.R. (1997). Chromatic properties of neurons in macaque area V2. *Visual Neuroscience* **14**, 1061–1072.

KNILL, D.C. (2001). Contour into texture: Information content of surface contours and texture flow. *Journal of the Optical Society of America A* **18**, 12–35.

Kobatake, E. & Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology* **71**, 856–867.

Koenderink, J.J. (1984). What does the occluding contour tell us about solid shape? *Perception* **13**, 321–330.

Lee, B.B., Martin, P.R. & Valberg, A. (1988). The physiological basis of heterochromatic flicker photometry demonstrated in the ganglion cells of the macaque retina. *Journal of Physiology* **404**, 323–347.

Lennie, P., Krauskopf, J. & Sclar, G. (1990). Chromatic mechanisms in striate cortex of macaque. *Journal of Neuroscience* **10**, 649–669.

Leventhal, A.G., Thompson, K.G., Liu, D., Zhou, Y. & Ault, S.J. (1995). Concomitant sensitivity to orientation, direction, and color of cells in layers 2, 3, and 4 of monkey striate cortex. *Journal of Neuroscience* **15**, 1808–1818.

Li, A. & Zaidi, Q. (2000). Perception of three-dimensional shape from texture is based on patterns of oriented energy. *Vision Research* **40**, 217–242.

Li, A. & Zaidi, Q. (2001). Veridicality of three-dimensional shape perception predicted from amplitude spectra of natural textures. *Journal of the Optical Society of America A* **18**, 2430–2447.

Li, A. & Zaidi, Q. (2003). Observer strategies in perception of 3-D shape from isotropic textures: Developable surfaces. *Vision Research* **43**, 2741–2758.

Li, A. & Zaidi, Q. (2004). Three-dimensional shape from non-homogeneous textures: Carved and stretched surfaces. *Journal of Vision* **4**, 860–878.

Linsker, R. (1989). How to generate ordered maps by maximizing the mutual information between input and output signals. *Neural Computation* **1**, 1402–1411.

Livingstone, M. & Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science* **240**, 740–749.

Livingstone, M.S. & Hubel, D.H. (1984). Anatomy and physiology of a color system in the primate visual cortex. *Journal of Neuroscience* **4**, 309–356.

Logothetis, N.K. (2003). The underpinnings of the BOLD functional magnetic resonance imaging signal. *Journal of Neuroscience* **23**, 3963–3971.

Ohki, K., Chung, S., Ch'ng, Y.H., Kara, P. & Reid, R.C. (2005). Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. *Nature* **433**, 597–603.

Op de Beeck, H., Wagemans, J. & Vogels, R. (2001). Inferotemporal neurons represent low-dimensional configurations of parameterized shapes. *Nature Neuroscience* **4**, 1244–1252.

Parraga, C.A., Troscianko, T. & Tolhurst, D.J. (2002). Spatiochromatic properties of natural images and human vision. *Current Biology* **12**, 483–487.

Pearson, P.M. & Kingdom, F.A. (2002). Texture-orientation mechanisms pool colour and luminance contrast. *Vision Research* **42**, 1547–1558.

Regan, B.C., Julliot, C., Simmen, B., Vienot, F., Charles-Dominique, P. & Mollon, J.D. (1998). Frugivory and colour vision in Alouatta seniculus, a trichromatic platyrrhine monkey. *Vision Research* **38**, 3321–3327.

Sachtler, W. & Zaidi, Q. (1992). Chromatic and luminance signals in visual memory. *Journal of the Optical Society of America A* **9**, 877–894.

Scharff, L.V. & Geisler, W.S. (1992). Stereopsis at isoluminance in the absence of chromatic aberrations. *Journal of the Optical Society of America A* **9**, 868–876.

Shapley, R. & Hawken, M. (2002). Neural mechanisms for color perception in the primary visual cortex. *Current Opinions in Neurobiology* **12**, 426–432.

Sun, H., Smithson, H., Lee, B. & Zaidi, Q. (2006). Specificity of cone inputs to macaque retinal ganglion cells. *Journal of Neurophysiology* **95**, 837–849.

Taylor, A.H. & Kerr, G.P. (1941). The distribution of energy in the visible spectrum of daylight. *Journal of the Optical Society of America* **31**, 3–8.

Thorell, L.G., De Valois, R.L. & Albrecht, D.G. (1984). Spatial mapping of monkey V1 cells with pure color and luminance stimuli. *Vision Research* **24**, 751–769.

Treue, S., Hol, K. & Rauber, H.J. (2000). Seeing multiple directions of motion-physiology and psychophysics. *Nature Neuroscience* **3**, 270–276.

Troscianko, T., Montagnon, R., Le Clerc, J., Malbert, E. & Chanteau, P.L. (1991). The role of colour as a monocular depth cue. *Vision Research* **31**, 1923–1929.

Ts'o, D.Y. & Gilbert, C.D. (1988). The organization of chromatic and spatial interactions in the primate striate cortex. *Journal of Neuroscience* **8**, 1712–1727.

Tse, P.U. (2002). A contour propagation approach to surface filling-in and volume formation. *Psychological Reviews* **109**, 91–115.

Wagner, G. & Boynton, R.M. (1972). Comparison of four methods of heterochromatic photometry. *Journal of the Optical Society of America* **62**, 1508–1515.

Wassle, H. & Boycott, B.B. (1991). Functional architecture of the mammalian retina. *Physiological Reviews* **71**, 447–480.

Zaidi, Q. (1997). Decorrelation of L- and M-cone signals. *Journal of the Optical Society of America A* **14**, 3430–3431.

Zaidi, Q. & Li, A. (2002). Limitations on shape information provided by texture cues. *Vision Research* **42**, 815–835.

Zaidi, Q., Pokorny, J. & Smith, V. (1989). Sources of individual differences in anomaloscope equations for tritan defects. *Clinical Vision Sciences* **4**, 89–94.