Vision Research 94 (2014) 41–50

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

The role of the foreshortening cue in the perception of 3D object slant

Iliya V. Ivanov [1], Daniel J. Kramer, Kathy T. Mullen *

McGill Vision Research, Department of Ophthalmology, McGill University, Montreal, Quebec H3A 1A1, Canada

ARTICLE INFO

Article history:
Received 8 November 2012
Received in revised form 21 October 2013
Available online 9 November 2013

Keywords:
Psychophysics
Color vision
Slant
Isoluminance
3D object perception
Radial frequency

ABSTRACT

Slant is the degree to which a surface recedes or slopes away from the observer about the horizontal axis. The perception of surface slant may be derived from static monocular cues, including linear perspective and foreshortening, applied to single shapes or to multi-element textures. It is still unclear the extent to which color vision can use these cues to determine slant in the absence of achromatic contrast. Although previous demonstrations have shown that some pictures and images may lose their depth when presented at isoluminance, this has not been tested systematically using stimuli within the spatio-temporal passband of color vision. Here we test whether the foreshortening cue from surface compression (change in the ratio of width to length) can induce slant perception for single shapes for both color and luminance vision. We use radial frequency patterns with narrowband spatio-temporal properties. In the first experiment, both a manual task (lever rotation) and a visual task (line rotation) are used as metrics to measure the perception of slant for achromatic, red–green isoluminant and S-cone isolating stimuli. In the second experiment, we measure slant discrimination thresholds as a function of depicted slant in a 2AFC paradigm and find similar thresholds for chromatic and achromatic stimuli. We conclude that both color and luminance vision can use the foreshortening of a single surface to perceive slant, with performances similar to those obtained using other strong cues for slant, such as texture. This has implications for the role of color in monocular 3D vision, and the cortical organization used in 3D object perception.

1. Introduction

A sensitivity to color contrast is known to be an integral part of the ventral stream of the human visual cortex, at least up to visual area V4 and the ventral-occipital (VO) region (Brewer et al., 2005; Bushnell et al., 2011; Hadjikhani et al., 1998; Liu & Wandell, 2005; Mullen et al., 2007; Wade et al., 2002; Zeki et al., 1991). Although color vision was initially suggested to play little part in form perception (Livingstone & Hubel, 1987, 1988), subsequently, it has emerged that color contrast can be used effectively in 2-dimensional form processing providing the stimulus components are presented at sufficiently low spatial frequencies to fall within the spatial passband of color vision (Hamburger, Hansen, & Gegenfurtner, 2007; McIlhagga & Mullen, 1996; Mullen, 1985; Mullen & Beaudot, 2002; Mullen, Beaudot, & Ivanov, 2011; Mullen, Beaudot, & McIlhagga, 2000; Reisbeck & Gegenfurtner, 1998; Wuerger & Morgan, 1999).

The visual system can reconstruct 3-dimensional (3D) object form quite effortlessly from the two-dimensional (2D) achromatic retinal representation by the use of different static monocular depth cues, including perspective cues, shape from shading, and texture gradients. Emerging evidence suggests that different cues relating to object surfaces (e.g. texture, surface color) may be analysed in different areas of LOC in human visual cortex from those processing object shape and form (Cant, Arnott, & Goodale, 2009; Cavina-Pratesi et al., 2010; Tsutsui, Taira, & Sakata, 2005). One of the outstanding questions is the role that color contrast can play in recovering the 3D form of objects based on static 2D monocular cues. For perspective cues of objects, based on demonstrations, Livingstone and Hubel (1987, 1988) suggested that isoluminant chromatic stimuli do not induce the sensation of three-dimensional form: the image of a bicycle slanting away from the viewer appeared difficult to recognize at isoluminance, as did a line drawings of a collection of jumbled, overlapping 3D shapes. On the other hand, Cavanagh (1991) argued that isoluminant line drawings of very simple 3D shapes (e.g. a cube) are recognizable at isoluminance. In terms of 'shape from shading', 3D representations by definition disappear at isoluminance and so it is not surprising that there is poor object and face recognition when shaded stimuli are rendered chromatic and isoluminant (Gregory, 1977). In terms of multi-element texture gradients, Livingstone and Hubel (1988)

---

* Corresponding author. Address: McGill Vision Research, Department of Ophthalmology, H4.14, 687 Pine Ave West, Montreal, Quebec H3A 1A1, Canada. Fax: +1 514 843 1691.

E-mail address: kathy.mullen@mcgill.ca (K.T. Mullen).

[1] Present address: Low Vision Clinic and Research Lab, Center for Ophthalmology, University Eye Hospital, University of Tuebingen, Tuebingen, Germany.

0042-6989/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.visres.2013.10.019

reported a loss of perceived slant in line-drawn texture representations of planar surfaces at isoluminance, although this was contradicted by study using regular arrays of hard-edged rectangles to represent a receding surface (Troscianko et al., 1991). Using more complex stimuli involving the modulation of textured patterns (orientation flows of multi-element patterns) at lower spatial frequencies, however, Zaidi and Li (2006) reported that 3D modulations could be determined at isoluminance.

Here we return to the role of perspective cues in color vision in determining 3D form for single elements. We test whether color contrast can contribute to the perception of 3D form (slant) based on the projective geometry of a single cue, that of stimulus foreshortening. We investigate the perception of slant using the foreshortening cue for isoluminant chromatic stimuli, presented within the spatial passband of color vision, compared to the equivalent achromatic stimuli. Slant represents the slope of an objects' surface about the horizontal axis as it recedes away from the observer and is an important aspect of 3D vision that can potentially be determined by linear perspective cues as well as foreshortening cues. The foreshortening cue comes from the compression of a single surface (change in the ratio of width to length) in the absence of linear perspective cues (Blake, Bulthoff, & Sheinberg, 1993; Buckley, Frisby, & Blake, 1996; Cutting & Millard, 1984; Stevens, 1981).

It is unclear whether the local foreshortening cue can independently induce slant perception in single shapes, whether achromatic or chromatic. Within the context of texture perception, it has been debated whether inter-element comparisons of size and density across the surface (a global cue) or the foreshortening of individual elements (a local cue) is the predominant cause for the sensation of slant (Cutting & Millard, 1984; Knill, 1998a, 1998b; Todd, Christensen, & Guckes, 2010). Todd, Christensen, and Guckes (2010) suggest that foreshortening of elements within the context of a complete texture fail to produce slant perception, but foreshortening has yet to be investigated as a local cue on its own. We focus on whether the projective geometry of a single, non-textured cue can induce the perception of a slanted surface. This is a simple local cue that uses a single form to represent slant in the absence of the global cues present in arrays of texture elements.

Our aims are twofold: first to determine whether we can use the foreshortening cue in isolation to make 3D slant judgments of single shape stimuli. Second, by comparing our slant perception discrimination for achromatic and isoluminant chromatic stimuli, we wish to determine whether there is a role for color contrast in 3D perception from foreshortening. We apply the foreshortening cue to shapes based on radial frequency (RF) patterns, which use curved contours to represent concentric shapes. By manipulation of the local curvature and the tangent orientation along the contour in the projecting plane, we represented an object foreshortening corresponding to a particular slant. We chose RF patterns because they are well established in the literature for studying global shape perception, requiring the integration of multiple, curved, contour elements into a complete shape (Wilkinson, Wilson, & Habak, 1998; Wilson & Wilkinson, 1997). They also have the added convenience of being narrow band in the spatial frequency domain, allowing the perception of both achromatic and chromatic stimuli to be based on the same spatial frequency range. The use of narrowband stimuli also allows chromatic and achromatic stimuli to be matched in visibility, based on the contrast detection of the same spatial frequency range. If our visual system could use local compressions arising from foreshortening to support global 3D shape perception, it would suggest a single cue is adequate to induce slant perception. We show that both achromatic and isoluminant, chromatic RF patterns induce very similar slant perception and slant discrimination, demonstrating that both color and luminance vision can use the foreshortening of a surface to perceive slant, with performances similar to those obtained using other strong cues for slant, such as texture. This implies that the human visual system can use foreshortening cues with both color and achromatic contrast to determine slant and 3D form, and supports the hypothesis that the areas of the IT cortex responsible for 3D object perception based on static, monocular cues receive inputs from both color and achromatic contrast.

2. Methods

2.1. Stimuli

Stimuli were achromatic or chromatic radial frequency patterns projected orthographically with different depicted slants, as illustrated in Fig. 1. The chromatic patterns either isolated the L/M cone opponent pathway (red–green, isoluminant) or isolated the S-cones. The radial frequency patterns were radially modulated D4s, the fourth derivative of a Gaussian (Wilkinson, Wilson, & Habak, 1998; Wilson & Wilkinson, 1997), band-limited in the spatial frequency domain, and defined by the following equations:

$$RF(r) = L_m[1 + c(1 - 4r^2 + 4r^4/3)e^{-r^2}] \tag{1}$$

$$r(x, y) = \frac{\sqrt{x^2 + y^2} - R(x, y)}{\sigma} \tag{2}$$

$$R(x, y) = R_m\{1 + A\sin[f_r \arctan(y/x) + \theta]\} \tag{3}$$

$$\sigma = \frac{\sqrt{2}}{\pi\omega_p} \tag{4}$$

where σ is the space constant of the $RF(r)$ in degrees, ω_p is the D4 peak spatial frequency (the contour, fixed at 1 cpd), $R(x, y)$ is the sinusoidal radial modulation of the D4 (defining its shape), R_m is the mean radius (2.0 deg before any applied slant). R_m is randomly varied between trials by a factor of 0.5–1 in conditions indicated in the text. f_r is the radial frequency (fixed at 4), A is the amplitude of the radial modulation fixed at 0.15, and θ is the phase of the modulation, which was randomly selected in each trial of the experiments to prevent adaptation. L_m is the mean luminance (for achromatic patterns) or mean chromaticity (for chromatic patterns) and c is cone contrast, as defined below.

We created the depicted optical slant at the centre of the stimulus by modulating the orthographic projection of the stimulus onto the screen such that:

$$\cos(\phi) = \omega/\lambda \tag{5}$$

where ϕ is the depicted physical slant, ω and λ are the width and the height of the stimulus respectively.

Stimuli were represented in the three-dimensional cone-contrast space (Cole, Hine, & McIlhagga, 1993; Sankeralli & Mullen, 1996). A linear transform was calculated to convert between the red, green and blue phosphor contrasts of the monitor and the three cone contrasts (L_C, M_C and S_C). Stimulus contrast is defined as the root mean square of the vector length in cone-contrast units (C_C):

$$C_C = \sqrt{(L_C)^2 + (M_C)^2 + (S_C)^2} \tag{6}$$

Isoluminance was estimated by a minimum motion task in the cone-contrast space (Cavanagh, Tyler, & Favreau, 1984), in which the perceived minimum motion of a Gabor stimulus (3.6 deg^2) was measured using a method of adjustment. Isoluminance was calculated as the arithmetic mean of at least 20 settings. We