ELSEVIER

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Vision Research 46 (2006) 1080–1090

Vision Research

www.elsevier.com/locate/visres

Neural correlates of shape-from-shading

Chuan Hou *, Mark W. Pettet, Vladimir Y. Vildavski, Anthony M. Norcia

The Smith-Kettlewell Eye Research Institute, 2318 Fillmore Street, San Francisco, CA 94115, USA

Received 14 June 2005; received in revised form 14 October 2005

Abstract

Visual evoked potentials were recorded during presentation of a single stimulus that generated bi-stable perceptual alternation between two different three-dimensional percepts. One interpretation (asymmetric) changed depth structure from flat to corrugated in depth and the other (symmetric) had the appearance of a flat surface translating laterally behind a set of apertures. Responses during perception of the asymmetric three-dimensional structure contained larger negative components than did responses during perception of the symmetric three-dimensional structure. Control experiments suggest that the interpretation of depth structure is selected after junction information caused by the interplay between shading and object shape is extracted.

Keywords: Shading; 3D-shape; Junctions; Bi-stable perception; Visual evoked potentials

1. Introduction

Shading, combined with boundary contour information provides a powerful monocular cue for three-dimensional (3D) shape. The ability to perceive shape-from-shading has been studied extensively using psychophysical methods (Aks & Enns, 1992; Braun, 1993; Enns & Rensink, 1990; Kleffner & Ramachandran, 1992; Ramachandran, 1988a, 1988b; Sun & Perona, 1997, 1996), but the underlying neural mechanisms are poorly understood. An early functional imaging study in humans used shading gradients to portray a series of convexities and concavities that were either lit from above or from the side (Humphrey et al., 1997). The observers reported stronger and more stable percepts of 3D when the lighting was from above, as previously described (Ramachandran, 1988b). A comparison of the activation in the two lighting conditions indicated differential responses in areas V1, V2, and V3. These results are consistent with a physiological study of V1 and V2 neurons of macaque monkeys (Lee, Yang, Romero, & Mumford, 2002) which demonstrated that similar stimuli evoked a contextual pop-out response (120–320 ms after stimulus onset), although this response was influenced by higher-order stimulus attributes and task experience. Human evoked response data (Mamassian, Jentzsch, Bacon, & Schweinberger, 2003) suggest that the direction of lighting is disambiguated as early as 100 ms, which is consistent with activation of early visual areas. A more recent fMRI study, using attention to either shading or color as the task has, on the other hand, emphasized the intra-parietal sulcus as being important in the perception of shape-from-shading (Taira, Nose, Inoue, & Tsutsui, 2001).

Although these results provide important clues about the neural mechanisms that process shading cues and extract shape information, there are still some open questions. The physiological and imaging studies have used lighting-direction or the shape of the shading gradient to vary the 3D percept. Here we varied the consistency of shading and boundary information to study the time course of shading/border interactions. Stimuli whose border and shading cues were consistent with a 3D interpretation elicited enhanced visual evoked potentials (VEP) responses and associated spectral components that were not present in control conditions that had inconsistent border and

---

[*] Corresponding author. Tel.: +1 415 345 2073; fax: +1 415 345 8455.
E-mail address: chuanhou@ski.org (C. Hou).

0042-6989/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2005.10.017

shading cues, but were matched for motion transients. Of greater interest, in the course of the experiments, we found a new bi-stable stimulus for which the same images yielded two different 3D interpretations, one alternating between flat and 3D and the other alternating between two translationally symmetric states with the same depth ordering. These stimuli have allowed us to separate response activity associated with the different depth interpretations from that due to low-level junction cues present in the stimuli.

2. Methods and materials

2.1. Observers

A total of 16 visually normal observers, aged 19 to 49 years participated. Eleven subjects participated in each experiment. Five of the subjects took part in both experiments. Each observer had or was corrected to 20/20 or better acuity in each eye, had normal stereopsis on the Titmus stereo test and was fully refracted for the viewing distance. The research followed the tenets of the World Medical Association Declaration of Helsinki and informed consent was obtained from the subjects after explanation of the nature and possible consequences of the study. The research was approved by the Institutional Human Experimentation Committee.

2.2. Stimuli

Stimuli were generated by a MacIntosh G4 computer equipped with an nVidia GeForce 2 graphics card and were presented on a 19-in., multi-synch monochrome CRT monitor (MRHB2000, Richardson Electronics) set to a graphics mode of 800 × 600 pixels, 72 Hz vertical refresh, at 256 gray levels (video bandwidth, 150 MHz). We varied the spatial relationship between shading and object border cues, either making the two consistent with a 3D interpretation or a predominantly two-dimensional (2D) interpretation. We used several variants of Ramachandran's windowed grating (Ramachandran, 1988a), as illustrated in Fig. 1 (see Supplementary Materials for animations). The shading information comprised a square-wave grating (denoted the *carrier*). The carrier pattern was a 0.32 cpd square-wave grating with 95% contrast and mean luminance of 123 Cd/m^2. The carrier was partially hidden by an occluding gray frame of 123 Cd/m^2 with three horizontal apertures, each 3° tall, with 3° between each aperture. The carrier was windowed by this mid-gray occluder whose aperture (denoted the *envelope*) running along the direction of the carrier followed a serrated path, with the period of the serrations matching the period of the grating. When the peaks of the envelope serrations aligned with the black-white transitions of the carrier, the stimulus looked like a folded surface (denoted the 3D image). This surface could appear as a set of triangular corrugations resembling a series of adjacent rooftops or a fan-fold card standing upright. When the carrier and envelope were misaligned, the stimulus appeared as a relatively flat, jagged, black-and-white-striped ribbon.

Note that none of the stimuli were actually 3D—they contained monocular cues that were either consistent with or largely inconsistent with a 3D percept. Prior to beginning the recordings, we reviewed each of the stimuli with

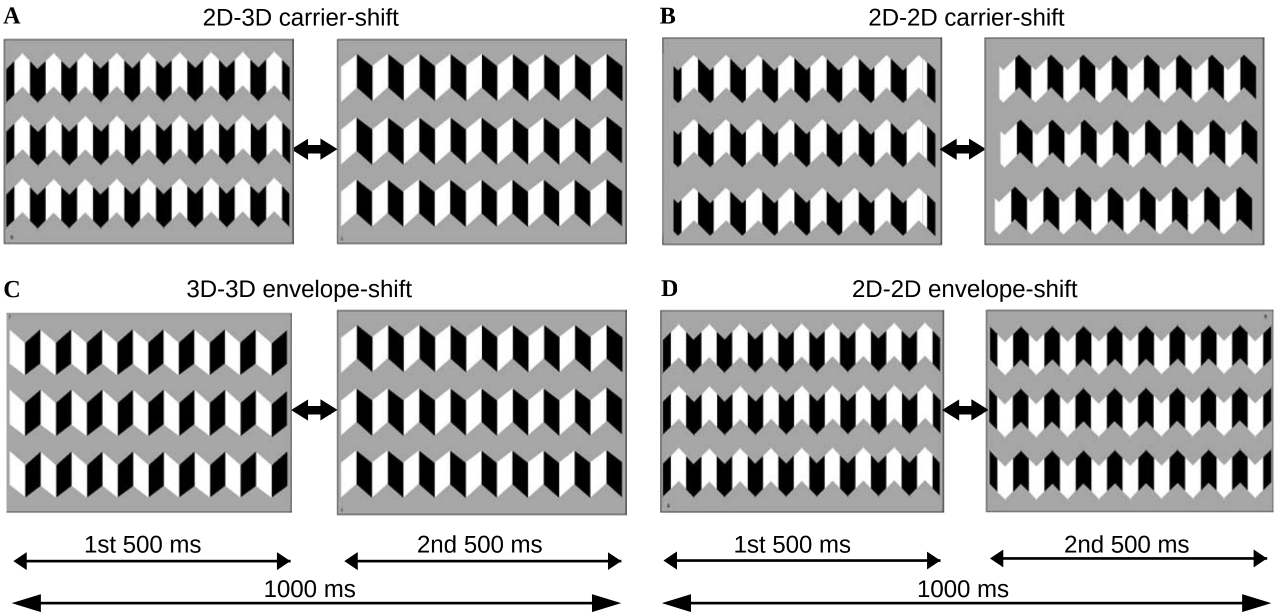

Fig. 1. Stimuli and experimental paradigm. (A) Test condition: 2D–3D carrier-shift, the alternation was between misaligned (~2D image) and aligned states (3D image); only the carrier modulated by 90°. (B) Control condition 1: 2D–2D carrier-shift, the same grating motion (carrier shift of 90°) was made symmetrically between misaligned states (~2D images) again without an envelope shift. (C) Control condition 2: 3D–3D envelope-shift, the carrier remained constant and the envelope shifted by 180°, yielding an alternation between two 3D images. (D) Control condition 3: 2D–2D envelope-shift, the carrier remained static and the envelope was shifted through 180° symmetrically about the zero crossing of the carrier grating. The images were updated every 500 ms and repeated every second. See Supplementary Materials for animations of the stimuli (set QuickTime to "loop" for continuous viewing).