Vision Research 86 (2013) 115–127

Contents lists available at SciVerse ScienceDirect

Vision Research

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

Reverse correlation reveals how observers sample visual information when estimating three-dimensional shape

Peter Scarfe [a,*], Paul B. Hibbard [b,c]

[a] Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB1 5QJ, United Kingdom
[b] School of Psychology and Neuroscience, University of St. Andrews, St Mary's College, South Street, St. Andrews, Fife KY16 9JP, United Kingdom
[c] Department of Psychology, University of Essex, Wivenhoe Park, Colchester C04 3SQ, United Kingdom

ARTICLE INFO

Article history:
Received 12 July 2012
Received in revised form 14 March 2013
Available online 9 May 2013

Keywords:
Stereopsis
3D shape

ABSTRACT

Human observers exhibit large systematic distance-dependent biases when estimating the three-dimensional (3D) shape of objects defined by binocular image disparities. This has led some to question the utility of disparity as a cue to 3D shape and whether accurate estimation of 3D shape is at all possible. Others have argued that accurate perception is possible, but only with large continuous perspective transformations of an object. Using a stimulus that is known to elicit large distance-dependent perceptual bias (random dot stereograms of elliptical cylinders) we show that contrary to these findings the simple adoption of a more naturalistic viewing angle completely eliminates this bias. Using behavioural psychophysics, coupled with a novel surface-based reverse correlation methodology, we show that it is binocular edge and contour information that allows for accurate and precise perception and that observers actively exploit and sample this information when it is available.

1. Introduction

1.1. Estimating three-dimensional object shape

The human visual system gains valuable information about the 3D structure of our environment from the fact that our eyes view the world from two slightly different positions (Julesz, 1971; Wheatstone, 1838). This difference means that in the shared field of vision a given point on an object typically projects to slightly different positions on the retina of each eye, producing horizontal and vertical retinal image disparities (Howard & Rogers, 2002). Horizontal disparities have been considered particularly important for depth perception because, with accurate information about object distance, they can be used to geometrically specify the full 3D structure of the environment (Hershenson, 1999; Johnston, 1991). Distance information is needed because the same pattern of horizontal disparity is consistent with infinitely many objects depending on the distance. Once distance information is known this depth ambiguity can be resolved.

Distance can be estimated from convergence (Brenner & van Damme, 1998), vertical disparity (Rogers & Bradshaw, 1993), or other cues (Hershenson, 1999), and could be used to scale horizontal disparities to specify shape accurately and unambiguously. Despite this possibility, observers typically show large systematic distance-dependent biases when estimating 3D properties of the environment from binocular disparities. This causes the same object to have a dramatically different perceived 3D shape when placed at different distances from the observer (Johnston, 1991) and objects need to morph in shape when moving in depth to be perceived as physically constant (Scarfe & Hibbard, 2006). This has led some to question the utility of disparity as a cue to 3D shape (Pizlo, 2008; Pizlo, Li, & Steinman, 2008; Todd & Norman, 2003) and whether shape can be accurately recovered from disparity and other visual cues at all (Todd & Norman, 2003). Others have concluded that whilst the perception of 3D shape is typically biased, it can be accurately estimated, but only when observers are provided with large, continuous, perspective transformations of an object (Bingham & Lind, 2008).

This latter point raises a more general issue, namely that studies of the estimation of 3D shape from disparity often present static random dot stereograms at eye height, orientated face-onto a static observer. An example of such a stimulus, in this case a random dot stereogram of a cylinder, is shown in Fig. 1a. The aim of this type of viewing situation is to experimentally control the information available to the observer so as to constrain the ways that they could estimate 3D shape. Whilst this is in many ways a sensible experimental approach, it results in a highly unnatural viewpoint that is not at all characteristic of our natural interactions with objects in real life. It is therefore possible that the biases in perceived shape demonstrated in previous studies, using both computer generated and real world stimuli (Johnston, 1991; Watt et al., 2005), are due to the restricted information that is available to

---

* Corresponding author.
E-mail address: ps611@cam.ac.uk (P. Scarfe).

0042-6989/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.visres.2013.04.016

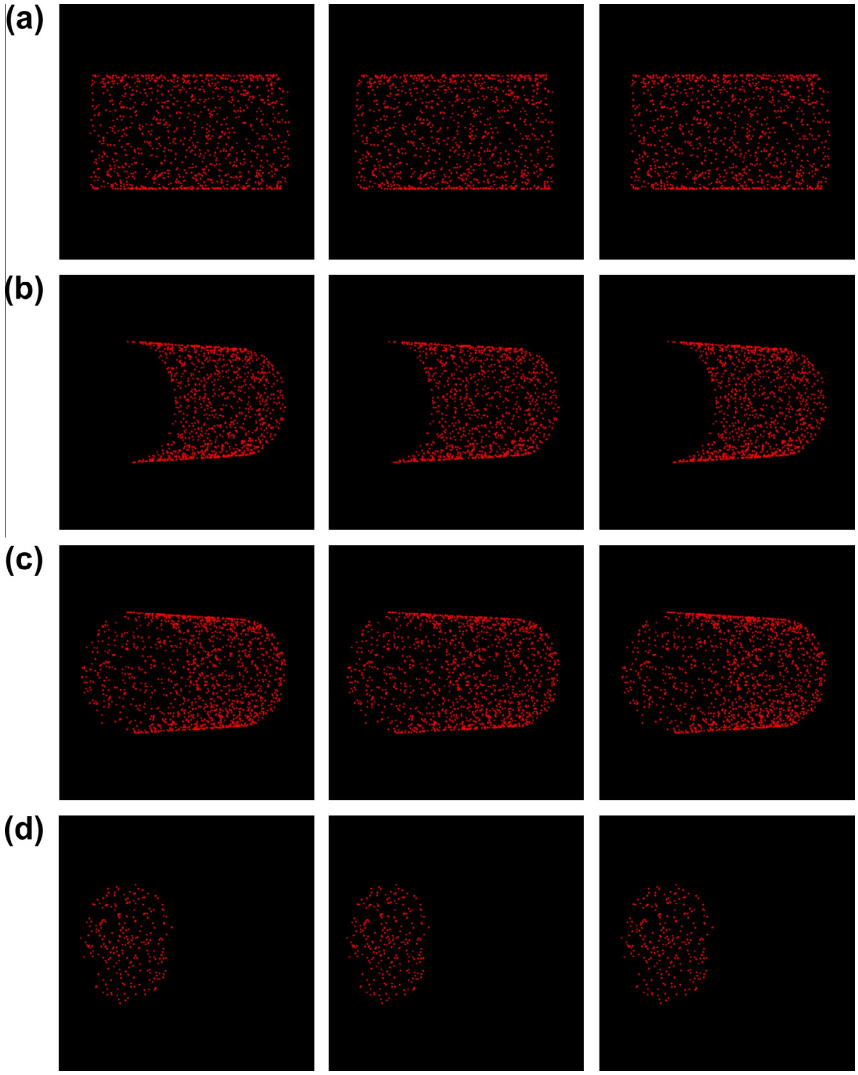

Fig. 1. Stereograms depicting the four stimulus conditions used in Experiment 1. (a) ACC condition (b) Non-Lidded condition (c) Lidded condition (d) Lid only condition. Left and middle columns for divergent fusion, middle and right columns for crossed fusion.

the observer, not some inherent inability to estimate 3D shape accurately (Todd & Norman, 2003).

The aims of the present study were therefore twofold. Firstly, we wished to determine whether veridical perception of 3D shape is possible with random dot stereograms, a type of stimulus that is routinely used to demonstrate the inaccuracies of 3D shape perception. Secondly, if accurate perception of shape is possible, we wished to determine what type of information allows for this. To achieve this we compared performance with the traditional face-on stimulus (Fig. 1a), to that obtained when cylinders were rotated around their vertical axis, either with (Fig. 1b), or without (Fig. 1c), a rendering of their "lid". Rotating the cylinder is a fairly trivial change, but despite this fact, additional information becomes available that previous research suggests could be very informative as regards 3D shape.

1.2. Additional cues available with cylinder rotation

With cylinder rotation, observers gain a view of the abrupt disparity discontinuities present along the end contour of the cylinder. Disparity discontinuities arise where abrupt changes in the spatial gradient of disparity occur, such as at the corners or edges of a surface, or where two different surfaces abut one another (Gillam, Blackburn, & Brooks, 2007). It is thought that the visual system might exhibit greater sensitivity to disparity discontinuities or higher derivatives of disparity gradients (Howard & Rogers, 2002; Stevens & Brookes, 1988). Consistent with this, the presence of disparity discontinuities has been shown to greatly increase both the speed of stereoscopic fusion and the accuracy of the ensuing 3D percept (in this case, stimulus slant) (Bradshaw, Hibbard, & Gillam, 2002; Gillam, Chambers, & Russo, 1988; Gillam, Flagg, & Finlay, 1984). As such it has been argued that edges and discontinuities are important primitives for stereopsis (Gillam, Chambers, & Russo, 1988). Indeed, cells in early cortical areas such as V2 of the macaque monkey have been show to selectively respond to stereoscopic contours, edges and corners (von der Heydt, Zhou, & Friedman, 2000), and this selectivity is thought to provide valuable information to upstream cortical areas that are responsive to more complicated aspects of 3D structure (Janssen, Vogels, & Orban, 1999; Orban, Janssen, & Vogels, 2006).

In addition to information from disparity, a rotated view of a cylinder makes available, or alters, other cues useful for estimating 3D structure. The rotated cylinder's body provides enhanced perspective cues that could be used to estimate its orientation (Hershenson, 1999; Howard & Rogers, 2002; Saunders & Backus, 2006; Saunders & Knill, 2001), and therefore possibly shape as