Vision Research 49 (2009) 84–95

Contents lists available at ScienceDirect

Vision Research

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

Depth propagation and surface construction in 3-D vision

Mark A. Georgeson [a,*], Tim A. Yates [a], Andrew J. Schofield [b]

[a] *School of Life & Health Sciences, Aston University, Birmingham B4 7ET, UK*
[b] *School of Psychology, Birmingham University, Birmingham B15 2TT, UK*

ARTICLE INFO

Article history:
Received 29 February 2008
Received in revised form 26 September 2008

Keywords:
Depth perception
Depth propagation
Surface construction
Lateral interaction
Stereopsis

ABSTRACT

In stereo vision, regions with ambiguous or unspecified disparity can acquire perceived depth from unambiguous regions. This has been called stereo capture, depth interpolation or surface completion. We studied some striking induced depth effects suggesting that depth interpolation and surface completion are distinct stages of visual processing. An inducing texture (2-D Gaussian noise) had sinusoidal modulation of disparity, creating a smooth horizontal corrugation. The central region of this surface was replaced by various test patterns whose perceived corrugation was measured. When the test image was horizontal 1-D noise, shown to one eye or to both eyes without disparity, it appeared corrugated in much the same way as the disparity-modulated (DM) flanking regions. But when the test image was 2-D noise, or vertical 1-D noise, little or no depth was induced. This suggests that horizontal orientation was a key factor. For a horizontal sine-wave luminance grating, strong depth was induced, but for a square-wave grating, depth was induced only when its edges were aligned with the peaks and troughs of the DM flanking surface. These and related results suggest that disparity (or local depth) propagates along horizontal 1-D features, and then a 3-D surface is constructed from the depth samples acquired. The shape of the constructed surface can be different from the inducer, and so surface construction appears to operate on the results of a more local depth propagation process.

1. Introduction

In the perception of 3-D surfaces, image elements or regions with ambiguous or unspecified depth can acquire perceived depth from unambiguous regions. In various contexts, such co-operative perceptual interactions have been called stereo capture (Kham & Blake, 2000; Ramachandran, 1986; Wu, Zhou, Qi, & Wang, 1998) where disparity-matching is ambiguous; disparity interpolation (Mitchison & McKee, 1987b; Yang & Blake, 1995) where features are sparse or monocular; and surface integration (Yin, Kellman, & Shipley, 2000), surface interpolation (Wilcox, 1999; Wilcox & Duke, 2005) or surface completion (Rubin, 2001; Yin, Kellman, & Shipley, 1997) where parts of a surface are missing or occluded. The general theme that emerges from such studies is that the visual system constructs a representation of visible surfaces (Marr, 1982) often from sparsely sampled and/or ambiguous data, and in doing so must interpolate depth values and/or surface shape in regions where explicit information is absent.

Like the filling-in of brightness, colour and texture (Pessoa & De Weerd, 2003), the interpolation of depth values may involve propagation of information across visual space. We describe here some striking new perceptual effects, where depth corrugation (disparity modulation) in flanking regions can induce vivid perceived depth in a central test region (Fig. 1B). We used a monocular gauge figure to measure the perceived shape of a variety of test images, and to quantify the magnitude of the induction effect. The nature and specificity of these induction effects lead us to propose that depth propagation and surface construction are different, perhaps successive, stages of visual processing.

2. Methods

2.1. Image generation

Images of a random, Gaussian white noise texture (512 × 512 pixels) were created in *Matlab* 5.2 on a Macintosh G4 computer, and displayed on a Clinton fast-phosphor monitor via a CRS Bits++ interface in true 14-bit greyscale mode. *PsychToolbox* (Brainard, 1997) and in-house software were used to calibrate the display system and run the experiments. Stereo viewing was achieved by using frame-interleaving FE1 goggles (CRS Ltd.) to present separate images to the two eyes. The high frame rate (150 Hz; 75 Hz per eye) ensured that the alternating display appeared as a steady 3-D image with no visible flicker. Luminance measurements and calculations showed that physical crosstalk between the two eyes' views was very low and would be undetectable. We calculated that for sine-wave gratings up to 80% contrast less than 1% of that contrast would effectively 'leak' through to the other eye.

* Corresponding author.
E-mail address: m.a.georgeson@aston.ac.uk (M.A. Georgeson).

0042-6989/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2008.09.030

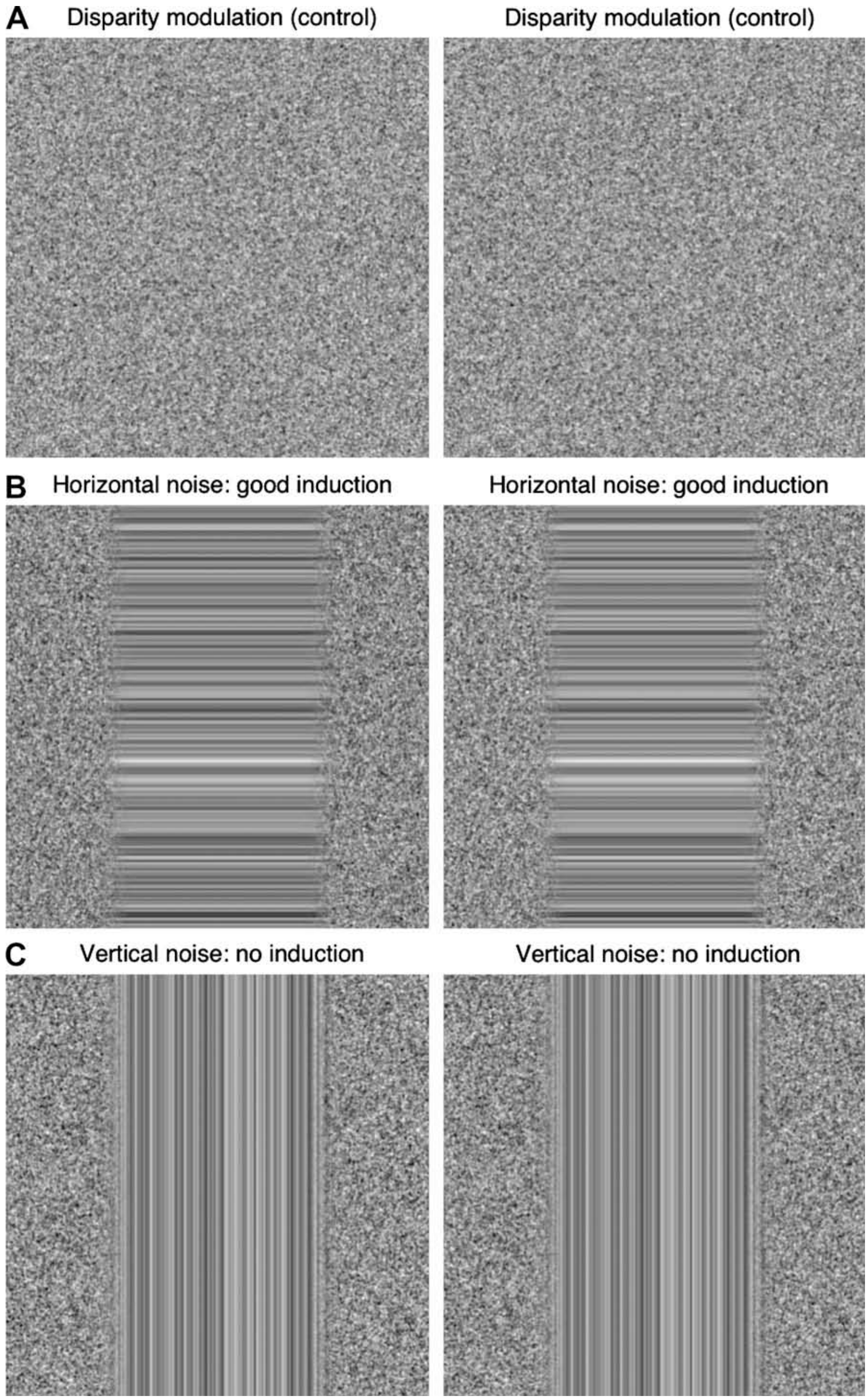

Fig. 1. (A) The slightly blurred Gaussian noise texture used in the experiments. Fusion of the left and right images should reveal a horizontal corrugated surface, produced by sinusoidal modulation of disparity. (B) Configuration of the central test region, flanked by the corrugated (disparity-modulated) inducing surface. Test texture is horizontal 1-D noise, and is the same in both eyes. Depth induction from the flanks was strong. (C) As B, but the test texture is vertical. Depth induction was absent. For more example images, see Supplementary Figures S1-S4.

To create the appearance of horizontal corrugations (Fig. 1A), with sinusoidal modulations of disparity, each row of the texture had to be shifted to the left or right by a small (often sub-pixel) distance that was a sinusoidal function of the row's vertical position (y) in the image. These small shifts were achieved by blurring each row separately with a Gaussian kernel whose space constant (σ) was 1 pixel (1.17 min arc), and whose peak was displaced by $\pm\delta/2$ for the left and right eyes respectively, where δ is the desired disparity of the row. Thus the convolution (blur) kernel r was defined by:

$$r(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x \pm \delta/2)^2}{2\sigma^2}\right) \quad (1)$$

where x is sampled in 1-pixel steps. To ensure that the blurring was isotropic (circularly symmetric), a similar blurring was then imposed on each column of pixels, but with no spatial offset.

The adequacy of sub-pixel resolution is influenced both by the effects of quantization (number of grey levels) in the graphics display system, and by the effects of gamma-correction. To evaluate the rendering of disparity one would, ideally, record the CRT