Vision Research 141 (2017) 127–135

Contents lists available at ScienceDirect

Vision Research

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

Thresholds for sine-wave corrugations defined by binocular disparity in random dot stereograms: Factor analysis of individual differences reveals two stereoscopic mechanisms tuned for spatial frequency

David H. Peterzell[a,*], Ignacio Serrano-Pedraza[b], Michael Widdall[c], Jenny C.A. Read[c]

[a] *College of Psychology, John F. Kennedy University, Pleasant Hill, CA, USA*
[b] *Facultad de Psicología, Universidad Complutense de Madrid, Madrid, Spain*
[c] *Institute of Neuroscience, Newcastle University, Newcastle upon Tyne, UK*

ARTICLE INFO

Keywords:
Stereo anisotropy
Individual differences
Factor analysis
Vision
Psychophysics
Binocular disparity
Stereopsis
Spatial frequency
Spatial frequency channels
Random dot stereograms
Stereogratings

ABSTRACT

Threshold functions for sinusoidal depth corrugations typically reach their minimum (highest sensitivity) at spatial frequencies of 0.2–0.4 cycles/degree (cpd), with lower thresholds for horizontal than vertical corrugations at low spatial frequencies. To elucidate spatial frequency and orientation tuning of stereoscopic mechanisms, we measured the disparity sensitivity functions, and used factor analytic techniques to estimate the existence of independent underlying stereo channels. The data set ($N = 30$ individuals) was for horizontal and vertical corrugations of spatial frequencies ranging from 0.1 to 1.6 cpd. A principal component analysis of disparity sensitivities (log-arcsec) revealed that two significant factors accounted for 70% of the variability. Following Varimax rotation to approximate "simple structure", one factor clearly loaded onto low spatial frequencies (≤ 0.4 cpd), and a second was tuned to higher spatial frequencies (≥ 0.8 cpd). Each factor had nearly identical tuning (loadings) for horizontal and vertical patterns. The finding of separate factors for low and high spatial frequencies is consistent with previous studies. The failure to find separate factors for horizontal and vertical corrugations is somewhat surprising because the neuronal mechanisms are believed to be different. Following an oblique rotation (Direct Oblimin), the two factors correlated significantly, suggesting some interdependence rather than full independence between the two factors.

1. Introduction

Stereo vision allows us to judge depth from small binocular disparities between the images projected into both eyes. Given that our eyes are offset horizontally in the head, depth perception is based mainly on horizontal disparities. The use of random-dot stereograms (Julesz, 1960, 1971) enables us to present stimuli where the horizontal disparities between eyes is the sole cue to depth. In this way, one can construct the stereoscopic analogue of sinusoidal luminance gratings: corrugations showing sinusoidal depth modulations defined purely by horizontal disparity (Tyler, 1974; Tyler & Raibert, 1975).

1.1. Disparity sensitivity functions (DSFs)

Thresholds for sinusoidal corrugations defined by disparity differ as a function of modulation spatial frequency. The minimum thresholds (highest sensitivity) usually occur at spatial frequencies of 0.2–0.4 cycles/degree (cpd), with sensitivity decreasing markedly above or below the peak. This finding of this representative function was initially established for horizontal corrugations (Tyler, 1974; Rogers & Graham, 1982; Howard & Rogers, 2012).

In later studies adding vertical corrugations, DSFs have been shown to have a similar representative band-pass shape, but have shown a puzzling anisotropy. Corrugations showing sinusoidal modulations of horizontal disparities at low spatial frequencies are much easier to detect when they are horizontally oriented than when they are vertically oriented (Bradshaw & Rogers, 1999; Bradshaw, Hibbard, Parton, Rose, & Langley, 2006; Serrano-Pedraza & Read, 2010; Serrano-Pedraza, Brash, & Read, 2013; Serrano-Pedraza et al., 2016). The same anisotropy also applies to slanted surfaces rotated around the horizontal axis and rotated around the vertical axis (Mitchison & McKee, 1990; Gillam & Ryan, 1992; Cagenello & Rogers, 1993; Hibbard, Bradshaw, Langley, & Rogers, 2002). Recently, Serrano-Pedraza et al. (2016) have shown that the strength of the anisotropy increases with age during development, suggesting a role of visual experience in this anisotropy.

---

[*] Corresponding author at: College of Psychology, John F. Kennedy University, 100 Ellinwood Way, Pleasant Hill, CA, USA.
E-mail addresses: davidpeterzell@mac.com, dpeterzell@jfku.edu (D.H. Peterzell).

https://doi.org/10.1016/j.visres.2017.11.002
Received 15 October 2016; Received in revised form 2 November 2017; Accepted 13 November 2017

1.2. Mechanisms underlying DSFs tuned to spatial frequency and orientation

Marr and Poggio (1979) initially presented a model including stereoscopic disparity channels. They suggested that (1) there is a range of channels with different levels of resolution establishing correspondence between stereo images (see also Tyler, 1973; Richards & Kaye, 1974). They further suggested (2) that the low-frequency channels precede and provide the foundation for higher frequency channels. Although the details are disputed, the basic idea of different channels has wide-ranging support (Mallot, Gillner, & Arndt, 2004; Farell, Li, & McKee, 2004; Menz & Freeman, 2002).

The representative shape of the DSF has often been interpreted as evidence of multiple mechanisms, "channels", or "high-level mechanisms" underlying the DSF. The low spatial frequency decrease in sensitivity has been explained in terms of multiple spatially-tuned disparity mechanisms, which interact through receptive fields' lateral inhibition (Tyler & Julesz, 1978; Schumer & Ganz, 1979; Tyler, 1990). Additionally, to many, the stereo anisotropy implies that distinct neuronal mechanisms are involved in detecting slant about the horizontal and vertical axes. These are regarded as distinct channels.

Classic studies have demonstrated that for horizontally-oriented corrugations, selective tuning also exists for different spatial frequencies (Cobo-Lewis & Yeh, 1994; Tyler, 1983, 1975; Schumer & Ganz, 1979; Tyler & Julesz, 1978). Considerable evidence exists to suggest that there are two or three of such channels, with a bandwidth of around 3 octaves (Serrano-Pedraza & Read, 2010; Tyler, 1990). Until recently no one had examined the mechanisms underlying perception of vertical disparity corrugations, and indeed Serrano-Pedraza and Read (2010) had suggested from circumstantial evidence that there might only be a single channel tuned to vertical. However, more recent evidence has made it clear that both vertical and horizontal stereo corrugations are detected by multiple disparity channels. Serrano-Pedraza et al. (2013), using a critical-band masking paradigm with random dot patterns, concluded there are at least two channels for vertical corrugations, while Witz and Hess (2013), using a detection/discrimination paradigm with spatially band-pass noise, concluded that there are at least three.

1.3. Using individual differences to examine underlying mechanisms

In the present study, we use an alternative method to estimate the minimum number and the nature of mechanisms underlying DSFs. Here, we measure the disparity thresholds of 30 individuals for horizontally- and vertically-oriented depth corrugations of different spatial frequencies depicted in random-dot stereograms (Experiment 1). We compare these with similar thresholds for horizontal and vertical step-edges, which contain many different spatial frequencies (Experiment 2). To estimate the minimum number of and the nature of the mechanisms underlying DSFs, we analyze individual differences in our data using factor analytic techniques.

The essential general assumptions are: (1) individual differences in visual data are determined in part by individual differences in the mechanisms underlying those data, and (2) one can often use correlational and factor-analytic methods to infer the minimum number and nature of the mechanisms underlying those data (Peterzell, 1993; Peterzell & Teller, 2000; Wilmer, 2008; de-Wit & Wagemans, 2016; Peterzell, 2016). The methods for estimating spatiotemporal mechanisms from individual differences have been described elsewhere in a series of studies on contrast sensitivity (Peterzell, 2016; Peterzell, Werner, & Kaplan, 1991, 1993, 1995; Peterzell & Teller, 1996, 2000; Peterzell, Dougherty, & Mayer, 1997; Peterzell & Kelly, 1997; Peterzell, Chang, & Teller, 2000; Peterzell, Schefrin, Tragear, & Werner, 2000).

Several previous investigators have examined individual differences in data to elucidate stereoscopic and other binocular mechanisms (Barendregt, Dumoulin, & Rokers, 2016; Bosten et al., 2015; Chen, Maloney, & Clifford, 2014; Chopin, Levi, Knill, & Bavelier, 2016; Harker, 1982; Hibbard et al., 2002; Harris, Chopin, Zeiner, & Hibbard, 2012; Hildreth & Royden, 2011; Ling, Nefs, Brinkman, Qu, & Heynderickx, 2013; Meredith, 1965; Nefs, O'Hare, & Harris, 2010; Richards, 1970, 1971, 1977; Richards & Lieberman, 1985; van Ee, 2003; Tidbury, Black, & O'Connor, 2015; Wilmer, 2008; Wilmer & Backus, 2007, 2008; Wismeijer, Erkelens, van Ee, & Wexler, 2010). Hibbard et al. (2002) for instance, used individual differences in the stereoscopic anisotropy to provide evidence that sensitivity to surface tilt and slant is in part limited by the sensitivity to luminance-defined orientation and spatial frequency. Others have correlated individual differences in stereopsis and binocular function with individual variability in accommodation and vergence, strabismus, dyslexia, artistic talent, flying and driving performance (Wilmer & Berens, 1920; Henson & Williams, 1980; Rutstein & Eskridge, 1984; Buzzelli, 1991; Livingstone & Conway, 2004; Livingstone, Lafer-Sousa, & Conway, 2011; Wright, Gooch, & Hadley, 2013; Winterbottom et al., 2014). But before this study, none examined the factors underlying disparity sensitivity functions for spatial frequency and orientation.

2. Methods

2.1. Human participants

Both experiments were performed in the Institute of Neuroscience of Newcastle University and were approved by the Ethics Committee of the Newcastle University Faculty of Medical Sciences (approval number 00625). Work was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Informed consent was obtained for experimentation with human subjects. All participants reported having normal or corrected to normal visual acuity. We tested 30 subjects aged between 18 and 26 years, 13 male and 17 females. One 20.4 year-old female did not participate in the second experiment.

2.2. Apparatus

Experiments were carried out in a dark room. Stimuli were presented on a 23-inch LG 3D monitor (D2342P) of the passive pattern-retarder type, with left and right images row-interleaved and separated by circular polarization. The spatial resolution of the monitor was 1920 × 1080 pixels (51 cm × 28.5 cm) and the refresh rate was 60 Hz. Observers sat at a viewing distance of 100 cm, so that a pixel subtended 54 s of arc. Participants used a forehead- and chin-rest and wore appropriate passive 3D glasses. They recorded their responses by pressing the left or right button of a standard computer mouse. All experiments were programmed in Matlab (R2012b) (www.mathworks.com) with the Psychophysics Toolbox extensions (www.psychtoolbox.org) (Pelli, 1997; Brainard, 1997; Kleiner, Brainard, & Pelli, 2007) and run on a DELL workstation with a NVIDIA Quadro K600 graphics card.

2.3. Stimuli and procedure (general)

The stimuli were static random dot stereograms consisting of white two-dimensional Gaussian dots each with a standard deviation of 1 arc min, with a density of 30 dots/deg^2 and without overlapping, presented on a black background. The disparity structure of the stimuli is described for each experiment below. The 3D was rendered with the monitor in standard 2D mode, using the line-interleaved stereo mode of Psychtoolbox's Psychimaging function. That is, our software generated left and right stimuli each 1920 pixels wide by 540 high, and interleaved them row by row to produce a single 1920 × 1080 image to send to the monitor. We did not use the monitor's own 3D function.

Thresholds, defined as a performance of 82% correct on two-interval forced choice tasks (2IFC), were estimated by an adaptive Bayesian staircase procedure, as described in Serrano-Pedraza et al.