Available online at www.sciencedirect.com
ScienceDirect
Vision Research 48 (2008) 1327–1334
Vision Research
www.elsevier.com/locate/visres

The transient nature of 2nd-order stereopsis

Robert F. Hess [a,*], Laurie M. Wilcox [b]

[a] McGill Vision Research, Department of Ophthalmology, McGill University, 687 Pine Avenue W (H4-14), Montreal, Que., Canada H3A 1A1
[b] Centre for Vision Research, Department of Psychology, York University, Toronto, Ont., Canada M3J 1P3

Received 12 June 2007; received in revised form 6 February 2008

Abstract

There are currently two competing dichotomies used to describe how local stereoscopic information is processed by the human visual system. The first is in terms of the type of the spatial filtering operations used to extract relevant image features prior to stereoscopic analysis (i.e. 1st- vs 2nd-order stereo; [Hess, R. F., & Wilcox, L. M. (1994). Linear and non-linear filtering in stereopsis. *Vision Research, 34*, 2431–2438]). The second is in terms of the temporal properties of the mechanisms used to process stereoscopic information (i.e. sustained vs transient stereo; [Schor, C. M., Edwards, M., & Pope, D. R. (1998). Spatial-frequency and contrast tuning of the transient-stereopsis system. *Vision Research*, *38*(20), 3057–3068]). Here we compare the dynamics of 1st- and 2nd-order stereopsis using several types of stimuli and find a clear dissociation in which 1st-order stimuli exhibit sustained properties while 2nd-order patterns show more transient properties. Our results and analyses unify and simplify two complimentary bodies of work.

Keywords: Stereopsis; Dynamics; 2nd-order; Transient; Sustained; Dmin

1. Introduction

There is strong evidence that human stereo-processing can operate in one of two modes, one in which the disparity of luminance-defined image features is extracted and another in which the disparity of contrast-image features is extracted (Kovács & Fehér, 1997; Langley, Fleet, & Hibbard, 1999; Lin & Wilson, 1995; McKee, Verghese, & Farell, 2004; McKee, Verghese, & Farell, 2005; Sato, 1983; Wilcox & Hess, 1995, 1997, 1998). This has been known for some time, the work of (Mitchell, 1966) and (Ramachandran, Rao, & Vidyasagar, 1973) both suggested that there was more to stereo than the processing of luminance-defined disparity. It has since been proposed that there are two separate stereoscopic processing systems. One is specialized for luminance-defined or 1st-order stimuli, depends on the spatial frequency content of image features, and is optimally sensitive to small disparities relative to the size of the object. Another is specialized for the processing of 2nd-order image structure, is relatively insensitive to spatial frequency and depends critically on the size of image features, particularly at large disparities (Wilcox & Hess, 1995, 1997, 1998).

More recently another mechanistic dichotomy has immerged based solely on response dynamics (Edwards, Pope, & Schor, 1999; Pope, Edwards, & Schor, 1999b; Schor, Edwards, & Pope, 1998; Schor, Edwards, & Sato, 2001). This distinction evolved from a similar dichotomy in the vergence literature (Edwards, Pope, & Schor, 1998; Pope, Edwards, & Schor, 1999a). It has been assumed, based on the properties of the sustained vergence system, that the sustained stereo system extracts depth for durations of up to 1 s and may be particularly sensitive to small disparities. This system is thought to be polarity sensitive and to exhibit narrowband tuning for spatial frequency and orientation of image features (Mitchell & O'Hagan, 1972; Schor & Wood, 1983). Schor and colleagues have argued that the transient stereo system on the other hand is polarity-insensitive (Pope et al., 1999b) and exhibits

* Corresponding author. Fax: +1 514 843 1691.
E-mail address: Robert.hess@mcgill.ca (R.F. Hess).

0042-6989/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2008.02.008

broadband tuning for orientation (Edwards et al., 1999) and spatial frequency (Schor et al., 1998) and is sensitive to a range of disparities (Schor et al., 2001).

On the face of it, there is more than a passing similarity between the sub-system processing properties that these dichotomies purport to represent. In principle, either could map on to a much earlier dichotomy (i.e. quantitative vs qualitative) based primarily on the size of the disparity (Ogle & Weil, 1958). For example, the properties of the so-called 1st- and 2nd-order stereo systems appear to correspond to the so-called sustained and transient stereo systems, respectively. To confirm this one would need to show that the 1st-order processing system exhibits *only* sustained dynamics and the 2nd-order system exhibits *only* transient dynamics. Other possibilities exist. For example, either the 1st-order or 2nd-order system (or both) could exhibit sustained as well as transient components. We would argue that if the primary distinction is in terms of the dynamic rather than the image features operated on then this would be the expected outcome. If the dichotomy is primarily based on what image features are processed (i.e. luminance vs contrast) then dynamics should be included as one of the many distinguishing features of these two systems (i.e. 1st-order = sustained vs 2nd-order = transient). To resolve this issue, here we compare the dynamics of stereoscopic detection of a set of stimuli designed to stimulate 1st- or 2nd-order mechanisms. Such comparisons are not available from the existing literature due to the wide range of stimuli and configurations that have been used, but also because in their investigations of 1st- and 2nd-order stereopsis Hess and Wilcox did not vary exposure duration, but held it constant at a brief duration to avoid eye movement artifacts.

Since the objective of this work is to make a careful comparison of the temporal properties of 1st- and 2nd-order stereopsis, it is important that the stimuli be chosen to discriminate between the two types of processing, but otherwise be as similar as possible. In a previous study, we undertook a comprehensive assessment of 1st-order stereo dynamics as a function of stimulus spatial frequency by covaring envelope size, spacing and stimulus bandwidth. We based our current stimulus parameter and configuration choices on the results of that previous study (Hess & Wilcox, 2006).

We used three different varieties of 2nd-order stimuli and their spatially equivalent 1st-order counterparts. In one stimulus set, we used bandpass 1-D spatial noise stimuli (bandwidth 0.6 octaves) whose stereo-pairs were either correlated (1st-order) or uncorrelated (2nd-order). Another stimulus set comprised Gaussian-windowed 1-D broadband spatial noise stimuli whose stereo-pairs were either correlated (1st-order) or uncorrelated (2nd-order). The final stimulus set consisted of vertical or horizontal Gabor stimuli (bandwidth approximate 0.9 octaves) that were either in-phase (1st-order) or out-of-phase (2nd-order), respectively. The latter two stimulus sets (Gaussian-windowed noise and the vertical and horizontal Gabors) were tested at three different spatial scales. All results were fitted with a model so that the degree to which the dynamics are sustained vs transient could be derived. The modeling results were then compared with the large body of data from our previous study of 1st-order stereopsis (Hess & Wilcox, 2006).

2. Methods

2.1. Apparatus

Stimuli were presented as grey-level variations on a single fast phosphor Clinton monitor. A full screen display of 1024 × 768 pixels was used. At a viewing distance of 1.15 m this subtended 17° by 14° of visual angle. The mean luminance was 69 cd/m^2 and the screen remained at mean luminance except when stimuli were presented. The monitor was controlled by a Cambridge Research Systems VSG2/3 graphics card which implements a resistor network to sum DAC outputs and allows a pseudo 12 bit grey-level representation after gamma correction. The frame rate was 120 Hz. Stereo-pairs were displayed on alternate frames and seen by each eye using LCD goggles.

2.2. Observers

Two observers were tested. Each of the subjects had normal or corrected to normal vision with normal stereo vision (using the Randot Stereo-test and by their performance in previous stereoacuity experiments).

2.3. Noise stimuli

Two different vertical 1-D spatial noise stimuli were used. The 1st-order noise stimulus (see Fig. 1A) was constructed by convolving a spatial Gabor (i.e. narrowband noise with a peak spatial frequency of 5.76 c/d, a sigma of 0.17° and bandwidth 0.6 octaves) by 1-D white noise (termed Gabor-filtered noise). This stimulus was then windowed with a 2-D Gaussian envelope with a standard deviation of 34.2 min. Correlated and uncorrelated stereo-pairs were generated, each at a range of relative disparities. The 2nd-order noise stimulus (Fig. 1B) consisted of the spatial Gaussian windowing of vertical 1-D white noise (termed Gaussian-windowed noise) at three different spatial scales (i.e. broadband noise with Gaussian sigmas of 8.28, 24.8 and 49.6 min).

2.4. Gabor stimuli

Gabor stimuli (approximate bandwidth 0.9 octaves) were oriented horizontally or vertically (see below) and presented at three different spatial scales (sigmas of 8.28, 24.8 and 49.6 min and peak spatial frequencies of 10.9, 3.6 and 1.8 c/d). Vertical in-phase Gabors were used to assess 1st-order stereopsis (Fig. 1C) and horizontal out-of-phase Gabors were used to assess 2nd-order stereopsis (Fig. 1D).

2.5. Contrast modulated noise stimuli

We initially tried to compare the dynamics of stereo-processing using a contrast modulated noise stimulus (2nd-order) with that of a spatially equivalent, luminance modulated noise stimulus (same spatial components but added). However, for reasons we do not understand (and only for some subjects), the addition of noise to a 1st-order stimulus changes its dynamics (compared with no noise), a finding also previously documented for stimulus detectability (Manahilov, Calvert, & Simpson, 2003) and which invalidates the luminance modulated stimulus as the ideal control for its contrast modulated counterpart. Owing to a lack of a valid 1st-order control, we did not continue with this stimulus.