

Available online at www.sciencedirect.com



Vision Research 46 (2006) 2798-2803



www.elsevier.com/locate/visres

Second-order spatial frequency and orientation channels in human vision

Dave Ellemberg^{a,b,*}, Harriet A. Allen^c, Robert F. Hess^d

^a Université de Montréal, Département de Kinésiologie, Canada

^b Centre de Recherche en Neuropsychologie Expérimentale et Cognition (CERNEC), Canada

^c School of Psychology, University of Birmingham, UK

^d McGill Vision Research Unit, McGill University, Canada

Received 9 December 2005; received in revised form 23 January 2006

Abstract

We compared the number of spatial frequency and orientation mechanisms underlying first- versus second-order processing by measuring discrimination at detection threshold for first- and second-order Gabors to determine the smallest difference in spatial frequency and orientation that permits accurate discrimination at threshold. For second-order gratings, the number of channels is the same as for first-order gratings for spatial frequencies up to about 2 cpd; however, there are fewer second-order channels at higher spatial frequencies. In contrast, the number of labeled channels for orientation is the same for first- and second-order gratings. In conclusion, our findings provide evidence for distinct spatial frequency and orientation labeled detectors in second-order visual processing. We also show that, relative to first-order, there are fewer second-order channels processing higher spatial frequencies. This is consistent with a filter-rectify-filter scheme for second-order in which the second stage of filtering is at lower spatial frequencies. © 2006 Elsevier Ltd. All rights reserved.

Keywords: First-order; Second-order; Spatial frequency; Orientation; Channels; Visual discrimination; Detection threshold

1. Introduction

It is now well accepted that the early stages of visual processing consist of a series of filter-like mechanisms that are selective to specific image properties such as spatial frequency and orientation (reviewed in Graham, 1989). In fact, accumulating evidence from psychophysical studies in humans as well as from single-cell recording studies in monkeys suggests that these mechanisms have relatively limited spatial frequency (± 1 octave) and orientation bandwidths (± 12 –15°) (reviewed in DeValois & DeValois, 1988). In an effort to determine the number of channels making up early spatial frequency processing, Watson and Robson (1981) used a discrimination task at detection threshold. This task consists in finding the smallest differ-

* Corresponding author. Fax: +1 514 3432181.

E-mail address: dave.ellemberg@umontreal.ca (D. Ellemberg).

ence in spatial frequency or orientation that enables discrimination at detection threshold. This is based on the argument that at detection threshold, stimulus energy for the given spatial frequency is only sufficient to activate the most sensitive channel. Therefore, if two stimuli can be perfectly discriminated at detection threshold, this suggests that they are processed by different channels. In their study, Watson and Robson (1981) found that there are seven spatial frequency channels between 0.25 and 30 cycles per degree (cpd). Hess and Norbdy (1986) extended Watson and Robson's findings and identified an additional channel that covers spatial frequencies below 0.2 cpd. Other studies have applied the same approach and determined the number of temporal frequency channels (Hess & Plant, 1985; Thompson, 1983).

The investigation of spatial frequency channels has mainly been done for mechanisms that are sensitive for first-order image properties, such as luminance. However, the human visual system is also able to detect second-order

^{0042-6989/\$ -} see front matter @ 2006 Elsevier Ltd. All rights reserved. doi:10.1016/j.visres.2006.01.028

image attributes other than luminance, such as texture and contrast, in which there is no difference in mean luminance (Cavanagh & Mather, 1989; Chubb & Sperling, 1988; Dumoulin, Baker, Hess, & Evans, 2003; Ellemberg, Allen, & Hess, 2004; McGraw, Levi, & Whitaker, 1999; Prins & Kingdom, 2003). Accumulating evidence indicates that our ability to detect stimuli defined by second-order features is done via separate mechanisms (Mareschal & Baker, 1998; Zhu & Baker, 1993) designed to extract modulations in contrast or texture rather than as a consequence of a processing non-linearity in early visual processing (Burton, 1973; Langley, Fleet, & Hibbard, 1996; Taub, Victor, & Conte, 1997).

Computational modeling suggests that the detection of second-order image properties requires not only a first stage of linear filtering, as does the detection of first-order image properties, but also an additional stage of filtering that is preceded by rectification (Chubb & Sperling, 1988; Wilson, Ferrara, & Yo, 1992). Recently, Prins and Kingdom (2003) provided evidence of such additional processing for second-order perception in humans. They found that the sensitivity to the second-order component of a texture (i.e., the envelope created by the orientation and frequency modulation) composed of densely packed Gabor elements (i.e., the carrier) is decreased by previous adaptation to a first-order grating that matched the characteristics of the first-order carrier in the stimulus. These data support a 'filter-rectify-filter' model, in which early linear filtering via mechanisms that are sensitive to the high spatial frequencies of the carrier are followed by a non-linear processing stage (e.g., full-wave, half-wave rectification, or squaring), and subsequently by a second stage filtering by mechanisms that are sensitive to the lower spatial frequencies of the envelope. Single-cell recording studies in cats report that the second stage of filtering is tuned to spatial frequencies that can be lower than the first stage of filtering by a factor of 10-20 (Mareschal & Baker, 1999).

Although the perception of second-order information has been widely investigated over the past decade our understanding of the neural architecture of the second-order processing system is at an early stage compared with that of its first-order counterpart. For example, while it is currently presumed that the human visual system possesses second-order mechanisms tuned to spatial frequency and orientation (Sutter, Sperling, & Chubb, 1995), there is at present limited support for this proposal. Furthermore, nothing is known of the number or bandwidths of such sub-mechanisms. By means of a discrimination at threshold paradigm, like that used by Watson and Robson (1981), the goal of the present study was to compare the number of first- and second-order channels used by the human visual system to extract spatial frequency and orientation information. We hypothesize that to exploit the degree of correlation that exists between the first- and second-order information contained in natural scenes (Johnson & Baker, 2004), it would be advantageous for the neural architecture supporting first- and second-order processing to be comparable over as much of the visual range as possible. However, we do expect to see this breakdown in the higher spatial frequency range because of the intrinsic lower spatial frequency limitation of second-compared with first-order detectors.

2. Methods

2.1. Observers

One of the authors and two observers, who were unaware of the issues examined, participated in this study. Two had normal acuity and the other had corrected to normal acuity.

2.2. Stimuli and apparatus

The stimuli consisted of localized two-dimensional Gabors and equally sized patches of two-dimensional binary noise. A Gabor is a sinusoidal modulation of luminance that is windowed by a two-dimensional Gaussian envelope. The first-order stimuli were created by adding the sinusoidal component of a Gabor to two-dimensional binary noise. This stimulus is represented by the following equation:

$$G(x, y) = L_{\text{mean}} + L_{\text{mean}} (G_{\text{sine}}(2\pi f x) + RC) \exp(-x^2/\sigma_x^2) \exp(-y^2/\sigma_y^2), \quad (1)$$

where L_{mean} is the mean luminance of the pattern, *f* is the spatial frequency of the sinusoidal modulation, *G* is the contrast of the grating, *R* is the random carrier (having contrast *C*), and (σ_x) and (σ_y) are, respectively, the horizontal and vertical space constants (the distance from the centre of the normal distribution where amplitude decreases to 1/e). The noise carrier had a contrast of 50% and each noise element was 1.9 by 1.9 arc min.

The second-order stimuli were created by multiplying the sinusoidal component of a Gabor by two-dimensional binary noise. This produced Gabors with an internal sinusoidal structure that varied in contrast and had a mean luminance that was constant across the pattern. The geometry of the second-order stimulus is represented by the following equation:

$$G(x,y) = L_{\text{mean}}(1 + (R(M_{\text{sine}}(2\pi f x) + 1)\exp(-x^2/\sigma_x^2)\exp(-y^2/\sigma_y^2)C), \quad (2)$$

where M is the modulation depth of the sinusoidal component and all other parameters are the same as indicated above. We created our second-order stimuli so that any change in modulation depth varied both the high and low contrast parts of the pattern.

The stimuli and presentation routine were programmed in MatlabTM using the Psychophysics Toolbox routine (Brainard, 1997; Pelli, 1997). The experiments were run on a Macintosh G4 computer and the images were displayed on a monitor that had a frame rate of 75 Hz, a resolution of 1152×870 pixels, and a mean luminance of 42 cd/m^2 . The relationship between voltage and screen luminance was measured with a photometer. The Gabors were produced with a subset of achromatic luminance values that were ordered linearly.

2.3. Procedure

The display was viewed binocularly in dim lighting. At the beginning of each trial, observers were instructed to fixate a cross at the centre of a uniformly illuminated screen. The method used to measure discrimination at threshold was a *two-by-two alternative forced-choice* procedure utilizing the method of constant stimuli (Watson & Robson, 1981). Randomly, for each trial, one interval contained an unmodulated noise pattern whilst the other contained a Gabor pattern. In all, there were four separate conditions: first- and second-order spatial frequencies (0.1–22 cpd), and first- and second-order orientation at threshold across all orientations (0–180 degrees).

2.3.1. Experimental conditions

2.3.1.1. Spatial frequency discrimination at threshold. The observer sat at 57 cm from the screen for baseline spatial frequencies from 0.1 to 12 cpd and at 128 cm from the screen for the baseline spatial frequencies of 12 and 22 cpd. To verify the effect of viewing distance, 12 cpd was tested at both viewing distances. The size of the stimuli was two periods of the smallest spatial frequency tested. The stimuli were truncated at $\pm 2\sigma$. The orientation of the sine-wave component was vertical. For each bloc of trials, a pair of spatial frequencies (the baseline and comparative spatial frequencies) was presented at five contrast levels, spanning the observer's threshold. Each of the two spatial frequencies making up a pair was presented 33 times at each contrast level in each of three separate runs, for a total of 990 trials per datum point (e.g., filled or open symbols on Figs. 2 and 3). For a given block of trials, spatial frequency (higher versus lower) was varied randomly across trails. The observer's task was to indicate, using one of two keys on the keyboard, which of the two intervals (first or second) contained the stimulus (detection) as well as indicate using two other keys, which spatial frequency was presented (discrimination). Prior to each block of trials, the observer was familiarized with high contrast versions of the stimuli. Each stimulus was presented for 350 ms, separated by a 500 ms interval during which the screen returned to mean luminance. Each interval was accompanied by a tone.

2.3.1.2. Orientation discrimination at threshold. All testing was done at 57 cm from the screen and spatial frequency remained constant at 1.4 cpd. The observer's task was to indicate, using one of two keys on the keyboard, which of the two intervals (first or second) contained the stimulus (detection) as well as indicate using two other keys, which orientation was presented (discrimination). Prior to each block of trials, the observer was familiarized with high contrast versions of the stimuli. All other details were the same as for the spatial frequency condition.

2.4. Data analysis

To determine whether the two stimuli presented in each pair (spatial frequency or orientation) were perfectly discriminated at threshold (i.e., indicating different underlying mechanisms), the psychometric functions for detection and discrimination were fitted by Weibull functions using the maximum likelihood procedure and the two ensuing functions were compared statistically (for details see Prins & Kingdom, 2003). Specifically, if the likelihood for "same parameters" is not significantly different from the likelihood for "different parameters," then the detection and discrimination data can be fitted as well by a single Weibull and the data are considered to be perfectly discriminated. We defined lambda (λ), an arbitrary variable, to be

$$\lambda = -2\ln(L_{\rm S}/L_{\rm D}),\tag{3}$$

where $L_{\rm S}$ is the likelihood for the "same parameters" (detection and discrimination Weibulls have equal parameters) and $L_{\rm D}$ is the likelihood for the "different parameters" (detection and discrimination Weibulls have different parameters). An advantage with this model is that it includes a bias parameter for discrimination. Further, it should be noted that combining the discrimination data also controls for part of this bias.

3. Results

Fig. 1 presents a typical example of the psychometric functions for detection and discrimination. The left panel shows a spatial frequency pairing that is not perfectly discriminated and the right panel shows a spatial frequency pairing that is perfectly discriminated. Fig. 2 presents the spatial frequency discrimination at threshold data for the first- and second-order stimuli for the two observers. The horizontal axis displays spatial frequency and the vertical axis shows the threshold difference (discrimination threshold minus detection threshold). Therefore, a value of zero indicates equal detection and discrimination thresholds. The arrows at the top correspond to the baseline spatial frequency to which the following points on the line are compared. Open symbols represent spatial frequencies that are not perfectly discriminated from the baseline, whilst filled ones show those that are (based on the calculation of lambda presented in Section 2.4). Circles present data that were collected at a viewing distance of 57 cm and the squares data that were collected at a viewing distance of 128 cm. For both participants, there is almost complete overlap between the spatial frequency pairs that were tested at both distances, indicating that changing viewing distance for the higher spatial frequencies did not influence the results.

The results are similar for both participants, and for both the first- and second-order conditions. For the firstorder stimuli, each observer produced seven pairs of spatial frequencies that were perfectly discriminated, consistent with there being eight labeled channels. In contrast, for the second-order stimuli, there are fewer pairs of spatial frequencies that are perfectly discriminated. The number of perfectly discriminated pairs is the same as for first-order



Fig. 1. In both panels, the dark triangles represent detection of 0.10 cpd and the white triangles represent discrimination of 0.10 cpd. The dark squares represent detection of 0.16 cpd (left) or 0.19 cpd (right) and the white squares represent discrimination of 0.16 cpd (left) or 0.19 cpd (right). The spatial frequency pairing on the left panel is not perfectly discriminated whilst the pairing on the right is. These data are from subject PG.



Fig. 2. Spatial frequency discrimination at threshold. The baseline spatial frequencies are indicated by the arrows at the top and the symbols present the comparative spatial frequencies tested. Open symbols show spatial frequency pairs that are not perfectly discriminated and filled symbols present those that are. Circles present data collected at a viewing distance of 57 cm and squares at a viewing distance of 128 cm. The left panel shows data for the first-order stimuli and the right panel shows data for second-order stimuli. The upper figures present data from DE and the lower figures present data from participant PG.

gratings for spatial frequencies up to about 2 cpd, and the difference in spatial frequency between these pairs is also similar. However, beyond 2 cpd there is only one pair of perfectly discriminated spatial frequencies for the second-order condition, compared to four pairs for the first-order condition. Further, for the second-order condition the highest spatial frequency tested was 19 cpd and it was not discriminated from 5 cpd.

Fig. 3 presents the data for orientation discrimination at threshold for the first- and second-order stimuli for the two observers. The horizontal axis displays orientation and the vertical axis shows the threshold difference (discrimination threshold minus detection threshold). All other details are the same as for Fig. 2. Again, the results are similar for both participants, and that both for the first- and second-order conditions. However, in contrast with the spatial frequency data, we find no difference between the results for the firstversus second-order conditions. For example, HS produced five pairs of perfectly discriminated orientations, both for the first- and second-order conditions, and the difference in orientation between these pairs is also similar.

4. Discussion

The results from the present study provide the first strong support for separate spatial frequency and orientation tuned mechanisms underlying the detection of second-order information. The lowest spatial frequencytuned mechanism is around 0.1 cpd and the highest around 6 cpd. We find evidence for five separate spatial frequency tuned channels. In terms of orientation, five discriminable steps can be made at threshold across the entire range, suggesting the existence of six separate orientation tuned channels, at least for the spatial frequency tested (i.e., 1.4 cpd).

We find a number of important differences and similarities between the first- and second-order channels. At the lower spatial frequencies, between the limits of 0.1–2 cpd, the number and bandwidth of the channels appear to be similar for first- and second-order processing. However, at higher spatial frequencies, there are fewer spatial frequency tuned channels underlying second-order compared to firstorder pattern vision. Specifically, between 2 and 6 cpd there are two second-order spatial frequency tuned channels,



Fig. 3. Orientation discrimination at threshold. The baseline orientations are indicated by the arrows at the top and the symbols present the comparative orientations tested. Open symbols show orientation pairs that are not perfectly discriminated and filled symbols present those that are. The left panel shows data for the first-order stimuli and the right panel shows data for second-order stimuli. The upper figures present data from DE and the lower figures present data from participant HS.

whilst between 2 and 20 cpd there are five first-order channels. The reduced number of spatial frequency tuned mechanisms covering higher frequencies and their wider bandwidths implies a decrease in efficiency of second-order processing with respect to first-order. This is consistent with evidence of reduced modulation sensitivity and reduced resolution at higher spatial frequencies for second-order compared to first-order pattern vision (Manahilov, Calvert, & Simpson, 2003; Schofield & Georgeson, 1999). Our pattern of results is also consistent with the 'filter-rectify-filter' processing scheme proposed to account for second-order processing (e.g., Cavanagh & Mather, 1989; Chubb & Sperling, 1988; Prins & Kingdom, 2003) and single-cell studies in cats that find that second-order filtering is tuned to spatial frequencies that are lower than the first stage of filtering by a factor of 10–20 (Mareschal & Baker, 1999).

In contrast, we do not find any difference in the number of orientation selective channels between the first- and second-order conditions. This finding might also suggest that the first- and second-order mechanisms also have similar bandwidths. However, it is important to note that although the current paradigm is a powerful tool to determine the number of separate labeled channels underlying spatial frequency and orientation filtering, it does not give a modelfree measure of bandwidth. In fact, using a paradigm designed to assess lateral inhibitory interactions among first- and second-order filters, we previously documented slightly narrower orientation tuning for first- compared to second-order processing (Ellemberg et al., 2004). Further, although other studies also suggest several similarities between the mechanisms underlying first- and second-order orientation processing (Allen, Hess, Mansouri, & Dakin, 2003; Mareschal & Baker, 1999; Smith, Clifford, & Wenderoth, 2001), recent imaging data in humans show that orientation-selective first-order adaptation originates in V1, whereas second-order stimulus orientation is extracted by additional extra-striate processing beyond V1 (Larsson, Landy, & Heeger, 2006).

The present findings of comparable number and bandwidth of orientation channels and spatial channels at lower spatial frequencies for first- and second-order mechanisms in human vision are not unexpected. In the early stages of cortical processing where cells exhibit selectivity for both spatial frequency and orientation, it has been shown that single-cells process both first- and second-order information via separate intracellular circuits (Mareschal & Baker, 1998; Zhu & Baker, 1993). A sizeable proportion of both simple and complex cells exhibit such a "double-duty" in which the spatial frequency and orientation tuning to the first- and second-order characteristics of the stimulus are matched (particularly for position of optima but also to a lesser extent, bandwidth) on a cell by cell basis (Mareschal & Baker, 1999; Song & Baker, 2004).

From a computational point of view, a recent analysis of the correlations between first- and second-order structure in natural scenes (Johnson & Baker, 2004) shows that there is a strong correlation between first- and second-order information contained in similar regions of a natural scene. In fact, these correlations occur for information within similar spatial frequency and orientation bands and are thought to be the basis of a mechanism to distinguish between real edges that are characterized by correlated changes in luminance (first-order) and contrast/texture (second-order) from shadows that do not contain such a correlation between first- and second-order image structure (Martin, Fowlkes, & Malik, 2004). Having comparable sets of spatial frequency and orientationally tuned first- and second-order mechanisms would seem to an important prerequisite for such an analysis. Furthermore, such a nice correspondence between the spatial frequency and orientation architecture of the early filters for these two types of visual information processing potentially facilitates future comparison of how their outputs are combined for suprathreshold discriminations.

Acknowledgments

This research was supported by CIHR Grant (MOP10818) and NSERC Grant (OGP0046528) to Robert F. Hess. Dave Ellemberg was supported by a NSERC research fellowship. We are also grateful to Nicolaas Prins for modifying and allowing use to use his statistical software.

References

- Allen, H. A., Hess, R. F., Mansouri, B., & Dakin, S. C. (2003). Integration of first- and second-order orientation. *Journal of the Optical Society of America A*, 20, 974–986.
- Brainard, D. H. (1997). The psychophysics toolbox. Spatial Vision, 10, 433–436.
- Burton, G. J. (1973). Evidence for non-linear response processes in human visual system from the measurements on the threshold of spatial beat frequencies. *Vision Research*, 13, 1211–1225.
- Cavanagh, P., & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision*, *4*, 103–129.
- Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. *Journal of* the Optical Society of America A, 5, 1986–2007.
- DeValois, R. L., & DeValois, K. K. (1988). *Spatial vision*. New York: Oxford University Press.
- Dumoulin, S. O., Baker, C. L., Jr., Hess, R. F., & Evans, A. C. (2003). Cortical specialization for processing first- and second-order motion. *Cerebral Cortex*, 13, 1375–1385.
- Ellemberg, D., Allen, H. A., & Hess, R. F. (2004). The investigation of local network interactions for first- and second-order information. *Vision Research*, 44, 1787–1797.

- Hess, R. F., & Norbdy, K. (1986). Spatial and temporal properties of human rod vision in the achromat. In: R. F. Hess, L. T. Sharpe, K. Nordby (Eds.), *Night vision: Basic, clinical and applied aspects* (pp. 390–413).
- Hess, R. F., & Plant, G. T. (1985). Temporal frequency discrimination in human vision: Evidence for an additional mechanism in the low spatial and high temporal frequency region. *Vision Research*, 25, 1493–1500.
- Graham, N. V. S. (1989). *Visual pattern analysers*. New york: Oxford University Press.
- Johnson, A. P., & Baker, C. L. Jr., (2004). First- and second-order information in natural images: A filter-based approach to image statistics. *Journal of the Optical Society of America A*, 21, 913–925.
- Langley, K., Fleet, D. J., & Hibbard, P. B. (1996). Linear filtering precedes nonlinear processing in early vision. *Current Biology*, 6, 891–896.
- Larsson, J., Landy, M. S., & Heeger, D. J. (2006). Orientation-selective adaptation to first- and second-order patterns in human visual cortex. *Journal of Neurophysiology*, 95, 862–881.
- Manahilov, J., Calvert, J., & Simpson, W. A. (2003). Temporal properties of the visual responses to luminance and contrast modulated noise. *Vision Research*, 43, 1855–1867.
- Mareschal, I., & Baker, C. L. Jr., (1998). Temporal and spatial response to second-order stimuli in cat A18. *Journal of Neurophysiology*, 80, 2811–2823.
- Mareschal, I., & Baker, C. L. Jr., (1999). Cortical processing of secondorder motion. *Visual Neuroscience*, 16, 527–540.
- Martin, D. R., Fowlkes, C. C., & Malik, J. (2004). Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(1).
- McGraw, P. V., Levi, D. M., & Whitaker, D. (1999). Spatial characteristics of the second-order visual pathway revealed by positional adaptation. *Nature Neuroscience*, 2, 479–484.
- Pelli, D. G. (1997). The Videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442.
- Prins, N., & Kingdom, F. A. A. (2003). Detection and discrimination of texture modulations defined by orientation, spatial frequency, and contrast. *Journal of Optical Society of America A*, 20, 401–410.
- Schofield, A. J., & Georgeson, M. A. (1999). Sensitivity to modulations of luminance and contrast in visual white noise: Separate mechanisms with similar behaviour. *Vision Research*, 39, 2697–2716.
- Smith, S., Clifford, C. W. G., & Wenderoth, P. (2001). Interaction between first- and second-order orientation channels revealed by the tilt illusion: Psychophysics and computational modelling. *Vision Research*, 41, 1057–1071.
- Sutter, A., Sperling, G., & Chubb, C. (1995). Measuring the spatial frequency selectivity of second-order texture mechanisms. *Vision Research*, 35, 915–924.
- Song, Y., & Baker, C. L. Jr., (2004). A common mechanism underlying neuronal processing of contrast envelopes and illusory contours [abstract]. *Journal of Vision*, 4, 66a.
- Taub, E., Victor, J. D., & Conte, M. M. (1997). Non-linear preprocessing in short-range motion. *Vision Research*, 37, 1459–1477.
- Thompson, P. G. (1983). Discrimination of moving gratings at and above detection threshold. *Vision Research*, 23, 1533–1538.
- Watson, A. B., & Robson, J. G. (1981). Discrimination at threshold: Labeled detectors in human vision. *Vision Research*, 21, 1115–1122.
- Wilson, H. R., Ferrara, V. P., & Yo, C. (1992). A Psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience*, 9, 79–97.
- Zhu, Y.-X., & Baker, C. L. Jr., (1993). A processing stream in mammalian visual cortex neurons for non-Fourier responses. *Science*, 261, 98–101.