Evidence from vergence eye movements that disparities defined by luminance and contrast are sensed by independent mechanisms

Holger A. Rambold

Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, MD, USA

Boris M. Sheliga

Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, MD, USA

Frederick A. Miles

Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, MD, USA

We recorded the initial disparity vergence responses (DVRs) elicited by 1-D sinusoidal gratings differing in phase at the two eyes by 1/4 wavelength and defined by luminance modulation (LM) or contrast modulation (CM) of dynamic binary noise. Both LM and CM stimuli elicited DVRs, but those to CM had longer latency (on average by ~20 ms). DVRs showed sigmoidal dependence on depth of modulation, with higher thresholds for CM than for LM. With both LM and CM stimuli, fixing the modulation at one eye well above threshold rendered the DVR hypersensitive to low-level modulation at the other eye (dichoptic facilitation). Disparities defined by LM at one eye and CM at the other generated weak DVRs in the "wrong" direction, consistent with mediation entirely by distortion products associated with the CM stimulus. These (reversed) DVRs could be nulled by adding LM to the CM stimulus (in phase), and the greater the depth of the CM, the greater the added LM required for nulling, exactly as predicted by a simple compressive non-linearity. We conclude that disparities defined by LM and by CM are sensed by independent cortical mechanisms, at least for the purposes of generating short-latency vergence eye movements to disparity steps.

Keywords: luminance, contrast, human vision, first-order disparity, second-order disparity, distortion products, disparity vergence


Introduction

When large random-dot patterns are viewed dichoptically and subjected to small binocular misalignments (disparities), corrective vergence eye movements are elicited at ultra-short latencies in humans and monkeys (Busettini, Fitzgibbon, & Miles, 2001; Busettini, Miles, & Krauzlis, 1996; Masson, Busettini, & Miles, 1997; Masson, Yang, & Miles, 2002; Rambold & Miles, 2008; Takemura, Inoue, & Kawano, 2002; Takemura, Inoue, Kawano, Quaia, & Miles, 2001; Takemura, Kawano, Quaia, & Miles, 2002; Takemura, Murata, Kawano, & Miles, 2007; Yang, FitzGibbon, & Miles, 2003). Thus, crossed disparities elicit convergence, uncrossed disparities elicit divergence, left-hyper disparities elicit left sursumvergence and right-hyper disparities elicit right sursumvergence, as expected of negative-feedback control mechanisms that use binocular disparity to eliminate vergence errors. When the disparity stimuli consist of 1-D sinusoidal luminance gratings that are identical at the two eyes except for a phase difference of 1/4 wavelength, the initial disparity vergence responses (DVRs) always operate to reduce the imposed 1/4 wavelength disparity, i.e., give greatest weight to the nearest-neighbor binocular matches (Sheliga, FitzGibbon, & Miles, 2006, 2007). Vergence responses can also be elicited at ultra-short latencies by binocular disparities applied to dense anticorrelated random-dot patterns—in which the dots seen by the two eyes have opposite contrast (Masson et al., 1997; Takemura et al., 2001)—even when those patterns are perceived as rivalrous and do not support depth perception.
(Cogan, Kontsevich, Lomakin, Halpern, & Blake, 1995; Cogan, Lomakin, & Rossi, 1993; Cumming, Shapiro, & Parker, 1998; Masson et al., 1997), consistent with the idea that these eye movements derive their visual input from an early stage of cortical processing prior to the level at which depth percepts are elaborated (Masson et al., 1997). The initial vergence responses to these anticorrelated stimuli are in the reverse direction of those to normal correlated stimuli (Masson et al., 1997; Takemura et al., 2001), one of many properties shared with disparity-selective neurons in the primate striate cortex that are well captured by the so-called disparity-energy model (Cumming & DeAngelis, 2001; Cumming & Parker, 1997; Fleet, Wagner, & Heeger, 1996; Ohzawa, DeAngelis, & Freeman, 1990; Parker & Cumming, 2001; Prince, Cumming, & Parker, 2002; Prince, Pointon, Cumming, & Parker, 2002; Qian, 1994; Read & Cumming, 2003; Read, Parker, & Cumming, 2002). The stereo matching in striate cortex relies on the local interocular correlations between the filtered signals from the two eyes, and by analogy with low-level motion detectors, the underlying disparity detectors can be thought of as first-order, Fourier or energy-based.

There is substantial evidence for stereopsis in the presence of disparity stimuli that are not defined by luminance—and so lack first-order disparity energy—and the underlying disparity-sensing mechanisms are referred to as second-order, non-Fourier or feature-based (Edwards, Pope, & Schor, 1999a, 2000; Ellemberg, Allen, & Hess, 2004; Frisby & Mayhew, 1978; Halpern, 1991; Hess, Baker, & Wilcox, 1999; Hess & Wilcox, 1994, 2001; Kovacs & Feher, 1997; Langley, Fleet, & Hibbard, 1998, 1999; McKee, Verghese, & Farell, 2004; Pope, Edwards, & Schor, 1999b; Schor, Edwards, & Sato, 2001; Wilcox & Hess, 1995, 1996, 1997, 1998; Ziegler & Hess, 1999). Clinical studies have described patients who show stereo perception with non-Fourier stimuli but fail standard stereo tests using luminance (McCull, Ziegler, & Hess, 2000). Further, a recent study in area V2 of cat cortex described a subset of neurons that respond to disparities defined by either first-order luminance or second-order contrast (Tanaka & Ohzawa, 2006). It has long been known that human subjects can initiate vergence eye movements to binocular images whose detailed form is quite different at the two eyes (Jones & Kerr, 1972; Mitchell, 1970; Westheimer & Mitchell, 1969). More recently, Schor and colleagues used Gabor patches in the competition paradigm of Jones and Kerr (1972) and reported vergence responses to patches that had orthogonal carriers or opposite luminance polarity, especially with larger disparities, i.e., the disparity of the Gaussian envelope alone could suffice to initiate vergence (Pope, Edwards, & Schor, 1999a; Sato, Edwards, & Schor, 2001). Experiments on humans (Sheliga et al., 2006) and monkeys (Miura, Sugita, Matsuura, Inaba, Kawano, & Miles, 2008) examined the stereo matching underlying the initial vergence responses to binocular stimuli whose second-order features and principal first-order component had disparities of opposite sign. These workers reported that the very earliest horizontal and vertical vergence responses were dominated by the first-order (luminance-defined) influences and any contribution from the second-order features was, at best, very weak. On the other hand, Stevenson (2002a, 2002b) measured the closed-loop gain of the vergence eye movements elicited by sinusoidal temporal variations in disparity defined solely by contrast-modulated dynamic noise—a pure second-order stimulus—and reported robust horizontal responses but no vertical responses.

We here report the results from a series of experiments in which we recorded the initial hDVRs and vDVRs elicited by disparity stimuli in the form of 1-D sinusoidal gratings with a binocular phase difference of 1/4 wavelength, created by luminance modulation (LM) and/or contrast modulation (CM) of a dynamic noise carrier that was uncorrelated at the two eyes (see Figure 1 for examples of such stimuli). In Experiment 1 we show that gratings defined by CM at both eyes (“CM + CM stimulus”; see Figure 1E), whether horizontal or vertical, elicit robust DVRs that work to reduce the 1/4 wavelength disparity, i.e., give greatest weight to the nearest-neighbor matches. However, there is substantial evidence for compressive non-linearities early in the visual pathway, even in the retina (e.g., MacLeod & He, 1993; MacLeod, Williams, & Makous, 1992), and these result in first-order distortion products that can render even pure second-order stimuli visible to first-order (energy-based) detectors; for discussion see Lu and Sperling (2001) and Scott-Samuel and Georgeson (1999). In Experiment 1 we also show that disparity stimuli defined by CM at both eyes still elicit robust DVRs, although at slightly longer latency, when such early first-order distortion products have been nulled by adding in-phase LM at both eyes, consistent with mediation by cortical mechanisms selectively sensitive to disparities defined by second-order CM. In Experiment 2, we show that hybrid first- and second-order disparity stimuli—in which one eye saw a grating defined by LM and the other saw a grating defined by CM with a binocular phase difference of 1/4 wavelength (“LM + CM stimulus”; see Figure 1F)—generated only weak DVRs and these were always in the “wrong” direction, i.e., opposite to the imposed disparity, and had the ultra-short latencies associated with first-order DVRs, all consistent with mediation by first-order distortion products associated with the CM stimulus. These (reversed) DVRs could be eliminated entirely by adding a small amount of LM to the CM stimulus in phase, equivalent in magnitude to the first-order distortion products uncovered in Experiment 1. Interestingly, the DVRs here showed a sigmoidal dependence on the added LM—with no evidence of a threshold as DVRs passed through zero and reversed sign—and the greater the depth of the CM, the greater the added LM required for nulling, again consistent with mediation entirely by first-order distortion products. Indeed, we were able to simulate these data using...
a simple model of the early compressive non-linearity. Various controls indicated that the failure of the hybrid LM + CM stimulus to elicit DVRs after nulling any first-order distortion products was not due to differences in the amplitude, spatial phase or timing of the inputs from the two eyes, leading to the conclusion that the cortical detectors sensing disparities defined by first-order LM and second-order CM are independent. In the remaining experiments, in any given trial the two eyes always saw the same type of stimulus, i.e., both saw either LM (“LM + LM stimulus”; see Figure 1D) or CM (the CM + CM stimulus), although the distortion products in the latter were always nulled with an appropriate (small) amount of in-phase LM at both eyes. This allowed us to characterize the properties of the pure first- and second-order DVRs. In Experiment 3, we show that the initial first-order DVRs had appreciably lower modulation thresholds for luminance-defined disparities than did the initial second-order DVRs for contrast-defined disparities. However, if the modulation at one eye was fixed well above these threshold levels, then the minimum modulation at the other eye required to generate a vergence response was

Figure 1. The disparity stimuli. (A–C) One-dimensional sinusoidal gratings created from a dynamic binary noise pattern (A), by luminance modulation (B), and contrast modulation (C); the upper plots show sample cross sections of the luminance and the lower panels show the patterns as seen by the observer. (D–F) Patterns as seen by the observer when disparity stimuli in the form of 1-D sinusoidal gratings with a binocular phase difference of 1/4 wavelength were created by luminance modulation at the two eyes (D), contrast modulation at the two eyes (E), and luminance modulation at one eye and contrast modulation at the other eye (F). LE, left eye. RE, right eye.
Experiment 1: Disparities defined by CM elicit robust vergence eye movements at short latency even after all distortion products have been nulled by adding LM

In this experiment, noise patterns were subject to CM, creating 1-D vertical or horizontal sinusoidal gratings that differed in phase at the two eyes by 1/4 wavelength (binocular disparity). These contrast-defined second-order disparity stimuli generated robust vergence eye movements that were corrective insofar as they worked to reduce the 1/4 wavelength disparity, i.e., gave greatest weight to the nearest-neighbor matches. However, a major concern was to determine the extent to which these DVRs might have resulted from first-order (luminance-defined) distortion products due to compressive nonlinearities that are known to exist early in the visual system. To address this question we used the nulling procedure of Stevenson (2002a) who had employed it to determine the contribution of first-order distortion products to the closed-loop disparity-vergence responses associated with contrast-defined disparities that modulated sinusoidally over time. Stevenson had based his approach on methodology that had been developed earlier to determine the contribution of first-order distortion products to the perception of second-order motion (Lu & Sperling, 2001; Scott-Samuel & Georgeson, 1999). The rationale behind the procedure is that any first-order distortion products can be offset by adding sinusoidal LM in phase with the sinusoidal CM, raising the luminance of the high-contrast regions and lowering the luminance of the low-contrast regions. The expectation in our experiment was that the amplitude of the initial DVRs would show a U-shaped dependence on the added LM, with the minimum corresponding to complete nulling of the first-order distortion products. Any residual DVR at the null point would be attributed to cortical mechanisms sensitive to the second-order contrast-defined disparity. On the other hand, if there were no residual DVRs at the null point, i.e., the minimum amplitude was zero, the implication would be that the disparity-vergence mechanism under study here does not receive inputs from cortical neurons capable of extracting binocular disparities defined solely by CM. We report that the added (in-phase) LM required to null the distortion products was generally small (<5%) and the minimum DVRs—attributed to cortical mechanisms sensitive to the second-order disparities defined by CM—were very robust with both horizontal and vertical stimuli, although of slightly longer latency than the DVRs to first-order disparities defined by LM.

Methods

Some of the techniques were very similar to those used previously (Rambold & Miles, 2008; Sheliga et al., 2006, 2007) and these will be described only in brief here.

Subjects

Three subjects participated in this study: two were authors (BMS and FAM) and the third (SB) was uninformed about the purpose of the experiments. All subjects had normal or corrected-to-normal vision. Ophthalmological and neurological findings were unremarkable and stereoacuity was 40 seconds of arc on the Titmus graded circles test.

Visual stimuli

Dichoptic stimuli were presented in a Wheatstone mirror stereoscope. In an otherwise darkened room, each eye saw a computer monitor (Sony or ViewSonic 21” CRT) through a 45° mirror, creating a binocular image straight ahead at a distance of 521 mm from the eyes’ corneal vertices, which was also the optical distance to the images on the monitor screens. Each monitor was driven by an independent PC (Dell Precision 380) and the outputs of each computer’s video card (PC NVIDIA Quadro FX 5600) were frame-locked via NVIDIA Quadro G-Sync cards. This arrangement allowed the presentation of independent images simultaneously to each eye. The monitor screens were each 400 mm wide and 300 mm high (41.8° × 32.0°), with 1600 by 1200 pixels (i.e., 36.6 pixels° directly ahead of each eye), and the two were synchronously refreshed at a rate of 100 Hz. Each monitor was driven via an attenuator (Pelli, 1997), allowing presentation of black and white images with 11-bit grayscale resolution. Two look-up tables (one for each monitor), each with 2048 equally spaced luminance levels from 0 to 41.6 cd/m², were created from luminance measures (Minolta LS_100 luminance meter) of different gray levels.

The disparity stimuli were 1-D vertical or horizontal sinusoidal gratings (0.25 cycles/p°) that differed in phase at the two eyes by 1/4 wavelength and were created by modulating the contrast of a high-frequency carrier (CM + CM stimulus). This carrier consisted of dynamic binary noise (DBN) that was generated independently for each eye and changed with every vertical refresh. The DBN filled the two screens (mean luminance, 20.8 cd/m²) and each noise element, which was composed of 4 × 4 pixels (6.6 × 6.6 min arc), was randomly assigned one of two luminance values, $L_{\text{max}}$ or $L_{\text{min}}$, yielding a luminance...
(Michelson) Contrast, \( C \), of 32\% given by the following Equation:

\[
\frac{L_{\text{max}} - L_{\text{min}}}{L_{\text{max}} + L_{\text{min}}} \times 100\%.
\]  

(1)

Gratings defined by CM were produced by multiplying the luminance profile of the DBN by 1-D sinusoids with 1-pixel resolution. The Depth of the Contrast Modulation, DCM, was given by the following Equation:

\[
\frac{C_{\text{max}} - C_{\text{min}}}{C_{\text{max}} + C_{\text{min}}} \times 100\%,
\]  

(2)

where \( C_{\text{max}} \) and \( C_{\text{min}} \) were the maximum and minimum local (Michelson) contrasts in the image. The DCM was always 80\% with a mean, space-averaged, luminance contrast equal to that of the unmodulated noise carrier (32\%).

We determined the contribution of any first-order distortion products to the initial DVRs elicited with gratings defined by CM at both eyes by adding various amounts of LM at both eyes, as in the studies of Stevenson (2002a) on closed-loop DVRs and of Scott-Samuel and Georgeson (1999) on motion perception. This was achieved by summing 1-D sinusoids (offset and scaled to preserve constancy of the DCM, see below) in phase with the luminance profile of the contrast-defined grating, so that the maximum luminance of the added LM was in phase with the maximum contrast of the CM. The depth of the added luminance modulation, DLM, was given by Equation 1 (where \( L_{\text{max}} \) and \( L_{\text{min}} \) were the mean local luminances—averaged over the light and dark areas—at the peak and at the trough, respectively, of the modulating sinusoid) and could be 0\%, 1\%, 2\%, 3\%, 4\%, 5\%, 8\%, 16\%, or 24\%. In other trials, the added LM was 180° out of phase, i.e., its maximum luminance was in phase with the minimum contrast of the CM, and the DLM in these cases could be 2\%, 4\% or 8\%. However, insofar as this added LM changed the local mean luminance (\( L_{\text{mean}} \)), it would also change the local luminance contrast. This effect is most readily appreciated by restating Equation 1, which defines the Luminance (Michelson) Contrast (\( C \)), as follows:

\[
\frac{A}{L_{\text{mean}}} \times 100\%.
\]  

(3)

where \( L_{\text{mean}} \) is given by \( (L_{\text{max}} + L_{\text{min}}) / 2 \) and \( A \), the Amplitude of the LM, is given by \( (L_{\text{max}} - L_{\text{min}}) / 2 \). 

Equation 3 indicates that increases (decreases) in the local \( L_{\text{mean}} \) result in proportional decreases (increases) in the local contrast, \( C \). This means that when the LM is added in phase with the CM, \( C_{\text{max}} \) will decrease and \( C_{\text{min}} \) will increase, the net result being a reduction in the DCM. On the other hand, adding LM 180° out of phase with the CM has the reverse effect, increasing the DCM. Thus, to ensure that the DCM was not altered when x\% LM was added in phase (180° out of phase) it was first necessary to increase (decrease) the DCM by x\%. This kind of adjustment was performed on each and every one of the LM stimuli used in all of the experiments described in this paper and no further reference will be made to this.

In our convention, the disparity of the binocular stimulus is defined by the position of the stimulus seen by the left eye minus the position of the stimulus seen by the right eye, and because rightward (upward) positions are positive in our convention, crossed (left-hyper) disparities have positive values whereas uncrossed (right-hyper) disparities have negative values.

Technical note: Even after completing the linearization of the monitor displays (using uniform gray images), a very slight luminance modulation could still be detected in the contrast-defined gratings by averaging the output of a photometer (cf. Scott-Samuel & Georgeson, 1999). To compensate for this it was generally necessary to add a very small amount of sinusoidal LM in phase with the sinusoidal CM, i.e., maximum luminance in phase with maximum contrast, consistent with a slight compressive non-linearity in the monitors (cf. Smith & Ledgeway, 1997). When the CM was 80\% the LM required for compensation averaged 0.35\%. When the CM was 40\% the LM required for compensation averaged 0.20\%. These data were used to estimate the (small) amount of LM required to compensate for the effects of the slight non-linearities in the monitors on all of the CM stimuli used in all of the experiments described in this paper and no further reference will be made to this.

Eye movement recordings

The horizontal and vertical positions of both eyes were recorded with the electromagnetic induction technique (CNC Engineering, Seattle, USA) using scleral search coils embedded in silastin rings (Skalar, Delft, The Netherlands), and the signals from each eye were sampled at 1 kHz (Collewijn, Van Der Mark, & Jansen, 1975; Robinson, 1963). At the beginning of each recording session, a calibration procedure was performed for the coil in each eye using monocularly viewed fixation targets.

Procedures

At the start of each trial, a central fixation cross (width 10°, height 5°, thickness 0.05°) appeared on both monitors. When the two eye-position monitors had signified locations within 2° of the centers of the crosses for 800–1100 ms, the crosses were replaced with the disparity stimuli (randomly chosen from a lookup table) for 200 ms (20 frames), after which the screen changed to
uniform gray, marking the end of the trial. After an inter-trial interval of 500 ms the fixation crosses reappeared signaling the start of another trial. The subjects were asked to fixate the centers of the crosses and to refrain from blinking or making saccades except during the inter-trial interval. Eye velocities in excess of 18°/s were assumed to be due to saccadic intrusions and trials with such contaminants were aborted and repeated later. A complete block of trials had 48 stimulus conditions: 4 directions of CM (crossed, uncrossed, left-hyper, right-hyper), 12 levels of added LM (-8%, -4%, -2%, 0%, 1%, 2%, 3%, 4%, 5%, 8%, 16%, and 24%).

The experimental paradigms were controlled by three PCs, which communicated via Ethernet using a TCP/UDP/IP protocol. One of the PCs was running the Real-time EXPerimentation software (REX) developed by Hays, Richmond, and Optican (1982), and provided the overall control of the experiment as well as acquiring, displaying, and storing the eye-movement data. The other two PCs were running a Matlab subroutine (MathWorks Inc., Natick, USA) utilizing the Psychophysics Toolbox extension (Brainard, 1997) and generated the binocular visual stimuli upon receiving a start signal from the REX machine.

**Data analysis**

The horizontal and vertical eye-position measures obtained for the right and left eyes during the calibration procedure were each fitted with second-order polynomials whose parameters were then used to linearize the corresponding eye-position data recorded during the experiment proper. The linearized eye-position measures were smoothed with an acausal sixth-order Butterworth filter (3 dB at 30 Hz) and mean temporal profiles were computed for each stimulus condition. Trials with small saccadic intrusions that had failed to reach the eye-velocity criterion of 18°/s used during the experiment were deleted. We used the convention that rightward and upward deflections of the eyes (and stimuli) were positive. The horizontal (vertical) vergence angle was computed by subtracting the horizontal (vertical) position of the right eye from the horizontal (vertical) position of the left eye. This meant that convergence and left-sursumvergence had positive signs. To improve the signal-to-noise, the mean vergence response profile to each uncrossed (right-hyper) disparity stimulus was subtracted from the mean vergence response profile to the corresponding crossed (left-hyper) disparity stimulus: “mean pooled-difference response measures.” As convergence and left-sursumvergence were positive in our sign convention, the horizontal (vertical) pooled-difference measures were positive when in the forward/compensatory/corrective direction defined by our 1/4 wavelength disparity stimuli (requiring that greatest weight be given to the nearest-neighbor matches). Response latency was estimated by determining the time after stimulus onset when the mean pooled-difference vergence velocity first exceeded 0.1°/s. In line with previous studies (Sheliga et al., 2006, 2007), the minimum latency of vergence onset with the luminance-defined stimuli varied a little from one subject to another and the initial vergence responses to a given stimulus were quantified by measuring the changes in the mean pooled-difference position measures over the initial open-loop period, i.e., over the period up to twice the minimum response latency. However, to permit within-subject comparisons across different paradigms and experiments, for a given subject the duration of this measurement window was always the same throughout the entire study (65 ms for subject, SB; 70 ms for BMS; 75 ms for FAM), and for all of the data obtained from a given subject with a given stimulus set, this window always commenced at the same time after the stimulus onset (“stimulus-locked measures”), the actual time being determined by the shortest response latency in the particular data set. Additional response measurements were made in which the start of the measurement window coincided with the measured onset of the individual mean vergence responses based on the 0.1°/s threshold criterion (“response-locked measures”). The criterion for statistical significance was $p \leq 0.05$ throughout the study.

**Results**

**Vertical disparity vergence responses (vDVRs)**

Figure 2 shows the vDVRs recorded from subject FAM when 1/4 wavelength disparities were applied to horizontal sine-wave stimuli defined by 80% CM with various amounts of added LM. The vertical vergence velocity recorded over time (in the form of mean pooled-difference profiles) is shown in Figures 2A (LM added in phase) and 2B (LM added 180° out of phase) and indicates that all stimuli elicited robust vDVRs that were always evident as upward deflections, which means that the initial vDVRs always operated to reduce the imposed 1/4 wavelength disparity, i.e., to reduce the disparity of the nearest-neighbor binocular matches (cf. Sheliga et al., 2006, 2007). This was the case whether the stimuli were defined solely by CM (dotted traces in Figures 2A and 2B, labeled “0%,” indicating that there was no added LM) or defined by both CM and LM, regardless of whether the LM was added in phase (continuous traces in Figure 2A) or 180° out of phase (continuous traces in Figure 2B). The vergence velocity response profile over time elicited by the CM stimulus alone (“0%” trace) is transient and bell-shaped, with a latency that is slightly longer than typically reported in the literature for the vDVRs to luminance-defined stimuli (Bussettini et al., 2001; Sheliga et al., 2006, 2007). The addition of in-phase LM >2% brought out responses at shorter latency and, within limits, the larger these additions the larger the amplitude of these initial
responses and the earlier their onset (Figure 2A). Out-of-phase LM additions also brought out responses at shorter latency and these peaked at earlier times with larger additions (Figure 2B). On the other hand, in-phase LM additions ≤2% tended to increase onset latency and decrease response amplitude, although these changes were quite small and require quantitative measures to fully appreciate them.

The quantitative effects of adding LM to the contrast-defined stimuli are illustrated in Figures 2C and 2D, which show the mean latency and the mean pooled-difference (stimulus-locked) response measures, respectively, plotted...
against the added LM (with in-phase additions having a positive sign and out-of-phase additions a negative sign by convention). These latency and amplitude plots have slightly skewed normal distributions that are well captured by the following Equation ($r^2$ values = 0.968 and 0.948, respectively), which consists of a Gaussian function and its integral, the latter determining the skewness:

$$\text{Off set} + A \cdot \exp \left( -\frac{(X - X_0)^2}{2 \cdot \sigma^2} \right) \cdot \left( 1 + \text{erf} \left( \frac{X - X_0}{\sqrt{2} \cdot \sigma} \right) \right),$$

(4)

where $X_0$ is the added LM (in %) at which the Gaussian function has its peak value, $A$ (in degrees or ms), $\sigma$ is the standard deviation of the Gaussian function, and $\alpha$ is a variable that determines the skewness. The continuous curves in Figures 2C and 2D are the least-squares best fits obtained with Equation 4, whose parameters are listed in Table S1 (latency data) and Table S2 (amplitude data: main entries, which are “stimulus-locked measures”) in the Supplementary Material. Table S1 includes the following parameters which were computed from the fitted functions: (1) $\text{Lat}_{\text{max}}$, the maximum latency (in ms); (2) $\text{LM}_{\text{peak}}$, the added LM (in %) at $\text{Lat}_{\text{max}}$, which was 1% for the data in Figure 2; (3) $\text{LM}_{\text{hh}}$, the half-height width (in %), i.e., the width mid-way between the peak and the asymptote; (4) $d\text{Lat}_{\text{max}}$, the maximum change in latency (in ms) given by the difference between the peak and the asymptote, which was 31.2 ms for the data in Figure 2. Table S2 includes the following parameters which were computed from the fitted functions: (1) $R_{\text{min}}$, the minimum response; (2) $\text{LM}_{\text{peak}}$, the added LM (in %) at $R_{\text{min}}$, which was 1.7% for the data in Figure 2; and (3) $\text{LM}_{\text{hh}}$, the half-height width (in %), i.e., the width mid-way between the peak and the asymptote. The changes in latency occurred over a much narrower range of LM than the changes in response magnitude and this is reflected in the smaller values of $\text{LM}_{\text{hh}}$: 5.3% vs. 12.8%, respectively, for the data in Figure 2. The best fits obtained with Equation 4 also show slight negative skewness in the latency data ($\alpha = -1.52$) and positive skewness in the magnitude data ($\alpha = 0.3$).

Tables S1 and S2 also list the data obtained from the other two subjects, which were generally similar and well fitted by Equation 4. In particular, $\text{LM}_{\text{peak}}$ ranged from 1.0% to 2.4% (mean = 1.8 ± 0.7%) for the latency plots and from 1.7% to 3.4% (mean = 2.5 ± 0.9%) for the magnitude plots, and the minimum $\text{vDVR}$s ($R_{\text{min}}$) were always robust. The maximum changes in latency, based on $d\text{Lat}_{\text{max}}$, averaged 25.6 ± 7.9 ms. Also, $\text{LM}_{\text{hh}}$ was consistently smaller for the latency data than for the amplitude data: 6.5 ± 1.0% vs. 12.4 ± 1.6%, respectively.

Response-locked (pooled-difference) amplitude measures also showed a U-shaped dependence on the added LM that was well fitted by Equation 4 and the parameters of these fits are included in Table S2 in parentheses (“response-locked measures”). These data indicate that the U-shaped dependence seen with the stimulus-locked measures was not entirely due to the changes in latency.

**Horizontal disparity vergence responses (hDVRs)**

Figure 3, which shows the hDVRs recorded from subject FAM when 1/4 wavelength disparities were applied to vertical sine-wave stimuli defined by 80% CM with various amounts of added LM, is organized like Figure 2. The horizontal vergence velocity profiles (mean pooled-differences) recorded over time are shown in Figures 3A (LM added in phase) and 3B (LM added 180° out of phase), and the associated quantitative estimates of latency and response amplitude (“stimulus-locked measures”) are plotted in Figures 3C and 3D, respectively. It is immediately apparent that the hDVR velocity profiles in Figures 3A and 3B are more prolonged than the transient $\text{vDVR}$ velocity profiles in Figures 2A and 2B and this was a consistent finding in all three subjects. Nonetheless, the dependencies of both the latency and the magnitude of the hDVRs (stimulus-locked measures) on the added LM were qualitatively the same as for the vDVRs and were always well fitted by Equation 4 in all three subjects; see the continuous curves in Figures 3C and 3D as well as the listings of the best-fit and computed parameters in Tables S1 and S2 in the Supplementary Material. Concerning the parameters of these fits: (1) the minimum hDVRs ($R_{\text{min}}$) were always robust; (2) the values of $\text{LM}_{\text{peak}}$ were always slightly higher than for the vDVR data, ranging from 2.0% to 5.2% (mean = 3.4 ± 1.6%) for the latency plots and from 2.4% to 6.0% (mean = 3.9 ± 1.9%) for the magnitude plots; and (3) the values of $d\text{Lat}_{\text{max}}$ were always slightly lower than for the vDVR data, averaging 18.8 ± 3.1 ms.

Response-locked (pooled-difference) amplitude measures again showed a U-shaped dependence on the added LM that was well fitted by Equation 4 and the parameters of these fits are included in Table S2 in parentheses.

**Discussion of Experiment 1**

The initial DVR amplitude measures showed U-shaped dependence on the added LM, the stimulus-locked measures reaching a minimum with in-phase LM additions that averaged 1.8% ($\text{vDVR}$) and 3.9% ($\text{hDVR}$), consistent with modest contributions from first-order distortion products due to compressive non-linearities in the early visual pathway (cf. Stevenson, 2002a). In this scheme, the response minimum is achieved when the added LM exactly nulls the first-order distortion products and any residual DVRs at this point are assumed to reflect the operation of cortical mechanisms that can extract disparities defined solely by second-order CM. It is
apparent that these residual responses, which we will refer to as “second-order DVRs,” were robust with both horizontal and vertical disparities defined by CM at both eyes, although their latencies were slightly greater than those to disparities defined by first-order LM, on average by 18.8 ms (hDVR) and 25.6 ms (vDVR), based on the values of dLat_{max}, which is our best estimate of the difference in the minimum latency of the first- and second-order DVRs. Note that the absolute magnitude of the initial DVRs was rather small, necessitating high-resolution recording techniques and the averaging of many responses to obtain robust data.

The data plotted in Figures 2C and 2D indicate that the temporal profile in Figure 2A that had the smallest measured amplitude, i.e., the profile that most nearly approximated the pure second-order vDVR, also had the longest measured latency and was obtained with 2% LM added in phase. Subtracting this putative second-order vDVR profile from each of the others in Figures 2A and 2B yields profiles that represent the responses to the pure luminance-defined additions in disparity, i.e., the first-order vDVRs. Figures 4A and 4B each reproduce this second-order vDVR profile as a thick gray trace (labeled “2%”) together with the residual first-order vDVR profiles as thin black traces, those obtained with in-phase LM additions being shown in A and those obtained with out-of-phase LM additions being shown in B. These traces indicate that the first- and second-order vDVRs are both transient in the velocity domain and clearly differ in latency. In fact, the first-order vDVRs to the larger LM...
Figure 4. The DVRs to disparity stimuli defined by contrast modulation (CM + CM stimulus) with added LM: separation into first- and second-order components (subject FAM). Mean pooled-difference vergence velocity traces recorded over time. (A) Gray trace labeled “2%” is reproduced from Figure 2A, where it is the response whose distortion products have been most completely nulled (with 2% LM added in phase) and therefore represents our best estimate of the “pure” second-order vDVR to the CM stimulus; black traces were obtained by subtracting this gray trace from all other traces in Figure 2A and represent our best estimate of the first-order vDVRs to the in-phase LM additions; labels indicate added LM and correspond to those in Figure 2A. (B) Gray trace labeled “2%” is the same as in panel A; black traces were obtained by subtracting this gray trace from all traces in Figure 2B and represent our best estimate of the first-order vDVRs to the out-of-phase LM additions; labels indicate added LM and correspond to those in Figure 2B. (C) Gray trace labeled “3%” is reproduced from Figure 3A, where it is the response whose distortion products have been most completely nulled (with 3% LM added in phase) and therefore represents our best estimate of the “pure” second-order hDVR to the CM stimulus; black traces were obtained by subtracting this gray trace from all other traces in Figure 3A and represent our best estimate of the first-order hDVRs to the in-phase LM additions; labels indicate added LM and correspond to those in Figure 3A. (D) Gray trace labeled “3%” is the same as in panel C; black traces were obtained by subtracting this gray trace from all traces in Figure 3B and represent our best estimate of the first-order hDVRs to the out-of-phase LM additions; labels indicate added LM and correspond to those in Figure 3B. Traces are each means of 203–212 responses (vDVRs) and 212–220 responses (hDVRs).
additions have reached their peaks before the second-order DVRs even get underway. However, given that latency often varies with response amplitude it is important to note that this difference in the latency of the first- and second-order DVRs is evident even when the two kinds of responses are of comparable amplitude; for example, in Figure 4, compare the latency of the gray traces with that of the black ones when the LM additions were 16% in phase (Figure 4A) and 8% out of phase (Figure 4B).

Figures 3C and 3D indicate that the temporal profile in Figure 3A that most nearly approximates the second-order hDVR is that obtained with 3% LM added in phase, and subtracting this profile from the others in that figure yields the residual first-order hDVR profiles. These first- and second-order hDVR profiles are shown in Figures 4C and 4D in black and gray line, respectively, and like the corresponding vDVR profiles, these clearly differ in onset latency. However, the hDVR velocity profiles are generally much more sustained than the vDVR profiles, often outlasting our 200-ms recording window.

Our finding that the first-order DVRs have a slightly shorter latency than the second-order DVRs (on average by 22.2 ms, based on the values of dLat_{max}) is reminiscent of the finding of Masson and Castet (2002) that the latency of the ocular following responses (OFRs) elicited by the Fourier component of unikinetic plaid motion is generally ~20 ms less than that of the OFRs elicited by the non-Fourier component. In addition, Lindner and Ilg (2000) have shown that pursuit eye movements initiated by some kinds of second-order motion are delayed ~30 ms compared to those initiated by first-order targets, and Hawken and Gegenfurtner (2001) reported even longer delays with slower targets. Such latency differences are consistent with the proposal that the perception of second-order motion requires more processing than does the perception of first-order motion, as in the filter–rectify–filter model of second-order processing (Baker, 1999; Chubb & Sperling, 1988; Nishida, Ledeweg, & Edwards, 1997; Scott-Samuel & Georgeson, 1999; Wilson, Ferrera, & Yo, 1992). Interestingly, filter–rectify–filter models have also been invoked to explain the processing of second-order disparity (McKee et al., 2004; Tanaka & Ohzawa, 2006; Wilcox & Hess, 1996). However, it is important to remember that non-Fourier motion defined by contrast-modulated noise does not elicit either early OFRs (Hayashi, Miura, Tabata, & Kawano, 2008) or optokinetic responses (Harris & Smith, 1992).

Our finding that contrast-defined disparities elicit robust vDVRs and hDVRs even after all distortion products have been nulled is clearly at odds with Stevenson’s (2002a, 2002b) findings with low-frequency temporal modulation (0.25 Hz) of the disparity of 1-D sinusoidal gratings defined by CM (0.5 cycles/°). Stevenson reported that adding in-phase LM could eliminate all vDVRs (although not all hDVRs). This raises the question as to why Stevenson’s contrast-defined stimuli failed to elicit second-order vDVRs. Actually, the vDVRs in Stevenson’s study were always appreciably weaker than the hDVRs, even with luminance-defined stimuli, whereas we found the opposite tendency. This directional anisotropy in our vergence data can be seen by comparing the data in Figures 2 and 3, as well as the data in Figures 4A–4D, and will become even more evident later when we characterize the dependencies of vDVRs and hDVRs on the depth of modulation and the spatial frequency in detail in Experiments 3 and 4. Stevenson’s use of low-frequency (0.25 Hz) temporal modulations might have been an important factor here, especially given our finding that vDVRs were much more transient than hDVRs. In addition, Stevenson recorded steady-state, closed-loop vergence responses whereas we recorded initial, open-loop vergence responses. This difference might have been critical because the disparity stimuli in Stevenson’s study appeared inside a circular window with a salient black surround to which the disparity was not applied. This surround would therefore work against any change in the closed-loop DVRs and might have been especially effective in damping any vDVRs, which often give more weight to peripheral inputs than do hDVRs (Howard, Fang, Allison, & Zacher, 2000).

Recent studies have argued that many of the fundamental visual properties of the earliest open-loop DVRs—whose latencies are commensurate with the first-order responses in the present study—directly reflect disparity processing in the striate cortex (Sheliga et al., 2007), i.e., bottom-up processing, although there is also substantial evidence from monkeys that these responses are mediated by the middle temporal (MT) and medial superior temporal (MST) regions of the cortex (Ponce, Lomber, & Born, 2008; Takemura et al., 2001, 2007). Comparable evidence is not available regarding the cortical mediation of either second-order DVRs or steady-state, closed-loop DVRs, and it is entirely possible that they are mediated by different (higher?) cortical regions such as the prearcuate area 8A (Gamlin, 2002; Gamlin & Yoon, 2000).

**Experiment 2: Disparities defined by CM at one eye and by LM at the other eye do not elicit vergence eye movements when distortion products have been nulled**

In Experiment 2, we attempted to elicit DVRs by applying 1/4 wavelength disparities to 1-D sinusoids defined by CM at one eye and LM at the other eye: the LM + CM stimulus. Edwards et al. (2000) have reported that such hybrid disparity stimuli give rise to percepts of depth and concluded that first- and second-order signals are pooled prior to the extraction of disparity and perceived
depth. Hybrid first- and second-order motion stimuli have been used by various authors (Ledgeway & Smith, 1994; Mather & West, 1993; Scott-Samuel & Georgeson, 1999). Mather and West (1993) used two-frame movies in which the images were random dots defined by contrast in one frame and by luminance in the other. Such movies failed to evoke perceived motion whereas two-frame movies in which the dots in both frames had the same defining characteristic—either luminance or contrast—resulted in robust percepts of motion. Ledgeway and Smith (1994) and Scott-Samuel and Georgeson (1999) used moving 1-D gratings that advanced in 1/4 wavelength steps but alternate frames in the sequence were defined by CM and LM. With this arrangement, successive frames that had the same type of modulation—CM or LM—were 180° out of phase and hence, on their own, would result in stationary flicker: only motion detectors that sense both the LM and the CM would respond here. When subjects did perceive motion here it was in the opposite direction to the 1/4 wavelength steps, consistent with an early compressive non-linearity that induces an out-of-phase LM component into the CM signal, i.e., a first-order distortion product. Scott-Samuel and Georgeson measured the amplitude of this distortion product by adding in-phase LM to the CM frames to exactly null the perceived motion. In Experiment 2A, we now report that 1/4 wavelength disparities applied to LM + CM stimuli elicit vergence responses but in the opposite direction to the applied disparity. Further, these “reversed” responses have the ultra-short latencies typical of first-order DVRs—rather than the slightly longer latencies expected with second-order DVRs—and can be eliminated by adding in-phase LM to the CM in an amount that approximates the magnitude of the distortion products recorded in Experiment 1. The clear implication is that nulling the distortion products eliminates all DVRs to the LM + CM disparity stimuli. In Experiment 2B, we reported that the added LM required to achieve this complete nulling increases linearly with the depth of the CM, and these data can be simulated by a simple compressive non-linearity that has previously been used by Scott-Samuel and Georgeson to account for first-order distortion products. Control experiments indicate that the failure of the LM + CM stimuli to generate DVRs (other than those due to distortion products) is unlikely to have resulted from binocular matching problems emanating from differences in the amplitude (Experiment 2C), timing (Experiment 2D), or phase (Experiment 2E) of the inputs from the two eyes.

**Methods**

With a few exceptions, which will be noted, the subjects, eye-movement-recording techniques, procedures, and data analysis were as in Experiment 1.

**Visual stimuli**

The disparity stimuli were 1-D vertical or horizontal sinusoidal gratings (0.25 cycles/°) that differed in phase at the two eyes by 1/4 wavelength and were created by modulating the contrast or the luminance of DBN patterns as in Experiment 1. There were five paradigms, designated Experiments 2A–2E.

In Experiment 2A, the gratings were defined by LM at one eye and CM at the other eye (LM + CM stimuli), and the 1/4 wavelength phase difference describing the binocular disparity specified the separation of the peak luminance of the CM pattern and the peak contrast of the CM pattern; see Figure 1F. Because hDVRs were often much weaker than vDVRs—even with optimal LM stimuli seen by both eyes (LM + LM stimuli)—we mostly used horizontal gratings and recorded the associated vDVRs. In any given trial with these horizontal gratings, the LM could have one of four values (DLM: 8%, 16%, 32%, 48%) whereas the CM was always the same (DCM: 80%). Given the findings in Experiment 1, we expected the CM pattern to have first-order distortion products that were 180° out of phase with the CM, i.e., the peak of the luminance modulation due to the distortion products would be in phase with the trough of the contrast modulation. This suggested that the first-order distortion products associated with the eye viewing the CM would combine with the first-order LM at the other eye to produce a 1/4 wavelength disparity that was 180° out of phase with the applied 1/4 wavelength disparity, and hence would tend to elicit a “reversed” DVR at the ultra-short latencies characteristic of first-order DVRs (cf. Scott-Samuel & Georgeson, 1999). We sought to null the influence of any such distortion products by adding small amounts of in-phase LM to the CM and, in any given trial, this added LM could have any one of 7 or 8 depths of modulation (ranging from 0% to between 3% and 4%, so as to range above and below the expected first-order distortion products based on Experiment 1). We also used vertical gratings in Experiment 2A to record the hDVRs when one eye saw 16% LM while the other eye saw 80% CM and again added small amounts of in-phase LM to the CM to bracket the expected first-order distortion products.

It will be seen that the data obtained in Experiment 2A strongly suggested that the DVRs elicited by the LM + CM stimulus were entirely due to first-order distortion products associated with the CM component. Such distortion products would be expected to vary with the depth of the CM but this parameter was always the same in Experiment 2A (DCM: 80%). In Experiment 2B, therefore, we again used LM + CM stimuli like those in Experiment 2A except that now the LM component was always the same (DLM: 16%) and, in any given trial, the CM could have one of four depths of modulation (DCM: 80%, 63.5%, 50.4%, 40%). Again, we sought to null the influence of the distortion products by adding small...
amounts of LM to the CM components in phase. For this, four values of added LM were used with each CM stimulus, again carefully selected to encompass the expected distortion products (based on linear interpolation of the data obtained from each subject in Experiments 1 and 2A).

A concern in Experiments 2A and 2B was that the “binocular matching” achieved with the LM + CM stimuli might have been compromised by an interocular difference in the stimulus magnitude, especially given that the depth of the modulation at the two eyes—given by DLM and DCM—was defined differently. In Experiment 2C, we therefore sought to determine if the detectors that sense disparities defined by luminance or contrast are tolerant of interocular contrast. For this, we used LM + LM stimuli in which the DLM at the two eyes could be either the same (8% + 8% or 16% + 16%) or different (8% + 16%) as well as CM + CM stimuli in which the DCM at the two eyes could be either the same (40% + 40% or 80% + 80%) or different (40% + 80%).

Experiment 1 had indicated that the first-order DVRs elicited by luminance-defined stimuli tended to have shorter latencies than the second-order DVRs of comparable magnitude elicited by contrast-defined stimuli, raising the possibility that the “binocular matching” achieved with the hybrid LM + CM stimuli in Experiments 2A and 2B might have been compromised by differences in the timing of the visual signals from the two eyes. For example, if the difference in latency originated in monocular sensory pathways and the neural response to the LM stimulus was very transient—decaying away completely before the delayed neural response to the CM stimulus reached the disparity sensing mechanism—then simultaneous binocular activation by such hybrid stimuli would be precluded. We addressed this issue in Experiment 2D using LM + CM stimuli in which the grating defined by LM (DLM: 12%) could appear at the same time as, or up to 6 frames after, the appearance of the grating defined by CM (DCM: 80%), each frame introducing a 10-ms interocular delay. Experiment 2D also included LM + LM stimuli (DLM: 12% + 12%) and CM + CM stimuli (DCM: 80% + 80%) in which the images seen by the two eyes could appear either simultaneously or with interocular delays of up to 6 frames. The rationale in these last experiments was to monitor any decay over time in the monocular neural signals generated by the LM and CM stimuli.

Another possible concern with the LM + CM stimuli was that the neural representation of the two monocular images might undergo differential shifts in spatial phase that could modify their effective binocular disparity. A differential phase shift of exactly 1/4 wavelength, for example, would completely negate the applied 1/4 wavelength of disparity, although such a fortuitous outcome might seem unlikely. We addressed this issue in Experiment 2E using LM + CM stimuli (DLM: 16%; DCM: 80%) with binocular disparities defined by phase differences of 0, 1/8, 1/4, 3/8, and 1/2 wavelength. We also used LM + CM stimuli (DLM: 16% + 16%) and CM + CM stimuli (DCM: 80% + 80%) in which the images seen by the two eyes could differ in spatial phase by 1/8 or 1/4 wavelength in order to gain further insight into the sensitivity of the initial DVRs to the magnitude of the applied disparity. A related problem is raised if the neural signals encoding the stimuli undergo full-wave rectification such as that invoked by Pope et al. (1999b) to explain the depth seen with disparity stimuli of opposite polarity. Edwards et al. (2000) have argued that such full-wave rectification would double the effective spatial frequency of first-order stimuli but not of second-order stimuli. If such frequency doubling were to occur with our LM stimuli, then applying 1/4 wavelength disparities to the LM + CM stimulus would align the peaks and troughs of the neural representations of the two stimuli and there would be no net disparity between them. To address this possibility, Experiment 2E included LM + CM stimuli (DLM: 24%; DCM: 80%) in which the spatial frequency of the CM stimulus (0.25 cycles/in) was twice that of the LM stimulus (0.125 cycles/in) and the applied disparity was 1/4 of the wavelength of the CM stimulus (1°).

Technical note regarding the stimuli in Experiments 2C–2E: Any grating containing CM was supplemented with in-phase LM to null any first-order distortion products, the amount of added LM being based on the findings in a pilot experiment. No further reference will be made to this.

Results

**Experiment 2A: LM + CM stimuli elicit low-level DVRs at ultra-short latency that are in the opposite direction to the applied disparity, consistent with first-order distortion products**

Quarter wavelength disparities applied to 1-D horizontal sinusoids defined by CM at one eye and LM at the other eye always elicited small but robust DVRs in the opposite direction to the applied 1/4 wavelength disparity. This is evident from Figure 5A, which shows the vDVR profiles (mean pooled-difference traces) recorded over time from subject BMS when the “pure” LM at one eye had one of four values (DLM: 8%, 16%, 32%, 48%) while the “pure” CM at the other eye was fixed (DCM: 80%). In the figure, the DLM is indicated by the numbers to the right of the traces, which are aligned with the peak responses. In all cases, the initial vDVRs consist of transient downward deflections, indicating “reversed” responses, and these were strongest when the “pure” LM was at its lowest level (8%). Significantly, these vDVRs have the ultra-short latency characteristic of this subject’s first-order vDVRs in Experiment 1 (79 ms).

The “reversed” direction and ultra-short latency of these vDVRs to LM + CM stimuli are consistent with an early compressive non-linearity that induces an out-of-phase
The LM component—a first-order distortion product—into the CM signal, whose nearest-neighbor (1/4 wavelength) match with the "pure" LM signal at the other eye is in the "reverse" direction. The addition of in-phase LM to the CM shifted the vDVRs toward the normal "forward" direction, and this is apparent in Figure 5B, which shows sample traces for the same subject for the case in which the other eye always saw "pure" LM of 8%. In the figure, the magnitude of the added LM is indicated by the numbers to the right side aligned with the peaks of the associated traces; upward deflections denote responses that reduce the imposed 1/4 wavelength disparity (note DVRs here are all "reversed"); horizontal gray line beneath the traces indicates the response measurement window for the 8% trace; abscissa shows time since onset of stimulus; each trace is mean of 140–145 responses. (B) Mean pooled-difference vergence velocity traces recorded over time when the CM was 80%, the LM at the other eye was 8%, and various levels of LM (indicated by numbers at right side aligned with peaks of the associated traces) were added to the CM in phase; other arrangements as in A; each trace is mean of 139–145 responses. (C) Dependence of pooled-difference (stimulus-locked) response measures on the LM added (in phase) to the CM (whose DCM was 80%) while the LM at the other eye was fixed at 8%, 16%, 32%, and 48% (see key); positive values denote responses that reduce the imposed 1/4 wavelength disparity; data points are each means of 139–145 measures (SDs ranged from 0.007° to 0.012°); continuous curves are the least-squares best fits obtained with Equation 5 (parameters given in Table S3 in the Supplementary Material).

Figure 5. The vDVRs to LM + CM stimuli: dependence on the LM added to the CM in phase (subject BMS). (A) Mean pooled-difference vergence velocity traces recorded over time when the CM was 80% and the LM at the other eye was 8%, 16%, 32%, and 48% (indicated by numbers at right side aligned with the peaks of the associated traces); upward deflections denote responses that reduce the imposed 1/4 wavelength disparity (note DVRs here are all "reversed"); horizontal gray line beneath the traces indicates the response measurement window for the 8% trace; abscissa shows time since onset of stimulus; each trace is mean of 140–145 responses. (B) Mean pooled-difference vergence velocity traces recorded over time when the CM was 80%, the LM at the other eye was 8%, and various levels of LM (indicated by numbers at right side aligned with peaks of the associated traces) were added to the CM in phase; other arrangements as in A; each trace is mean of 139–145 responses. (C) Dependence of pooled-difference (stimulus-locked) response measures on the LM added (in phase) to the CM (whose DCM was 80%) while the LM at the other eye was fixed at 8%, 16%, 32%, and 48% (see key); positive values denote responses that reduce the imposed 1/4 wavelength disparity; data points are each means of 139–145 measures (SDs ranged from 0.007° to 0.012°); continuous curves are the least-squares best fits obtained with Equation 5 (parameters given in Table S3 in the Supplementary Material).

LM component—a first-order distortion product—into the CM signal, whose nearest-neighbor (1/4 wavelength) match with the “pure” LM signal at the other eye is in the “reverse” direction. The addition of in-phase LM to the CM shifted the vDVRs toward the normal “forward” direction, and this is apparent in Figure 5B, which shows sample traces for the same subject for the case in which the other eye always saw “pure” LM of 8%. In the figure, the magnitude of the added LM is indicated by the numbers to the right side aligned with the peaks of the associated traces, which are again aligned with the peak responses. The mean pooled-difference (stimulus-locked) response measures for these data are plotted as a function of the added LM in Figure 5C (in half-filled black symbols). Figure 5C also shows the response measures obtained from the same subject (BMS) when the “pure” LM at the other eye was fixed at 16%, 32% and 48% and, in each case, the dependence on the added LM is sigmoidal. All four data sets in Figure 5C were well fitted by the following Equation (\(r^2 = 0.998-0.999\)), which is a variant of the Naka–Rushton equation (Naka & Rushton, 1966):

\[
A + R_{\text{max}} \frac{m^n}{m^n + m_{50}^n},
\]

where \(A\) is an offset to permit a non-zero minimum (negative) asymptote, \(R_{\text{max}}\) is the maximum attainable (positive) response, \(m\) is the added LM (in %), \(m_{50}\) is the added LM (in %) when the response is mid-way between its asymptotic values, and \(n\) is an exponent. The four curves in Figure 5C are the best fits obtained with Equation 5 and all intersect at a common point just slightly below the zero response level where the slopes tend to be inversely related to the “pure” LM at the other eye. Likewise, the responses when there was no added LM were inversely related to the “pure” LM at the other eye, and this is also reflected in the values of the offset, \(A\); see Table S3 in the Supplementary Material, which lists
the best-fit parameters obtained with Equation 5. The added LM at which each of these fitted functions intersects the zero-response level (designated, \( \text{DLM}_{\text{DVR} = 0} \)) tends to increase very slightly with the depth of the “pure” LM, ranging from 1.70% to 1.84%, and these computed measures are included in Table S3. These measures indicate that the added LM required to null the “reversed” responses in Experiment 2A was always within 0.6% of the first-order distortion products estimated for this subject in Experiment 1 (2.3%, which is included in Table S3 in square brackets for easy comparison). A surprising feature of the data in Figure 5C is the lack of a clear threshold as the vDVRs approach and pass through zero. This phenomenon will be examined further in Experiment 3.

The vDVR data from the other two subjects showed the same general characteristics. Thus, when no LM was added, responses to the LM + CM stimuli were in the opposite direction to the applied 1/4 wavelength disparity and of ultra-short latency, consistent with first-order distortion products. Further, these reversed responses could be nulled by adding LM in phase to the CM and the dependence of the responses on the added LM was well fitted by Equation 5. The best-fit parameters are included in Table S3, together with the computed values of \( \text{DLM}_{\text{DVR} = 0} \), which ranged from 1.10% to 1.29% (FAM) and from 2.73% to 3.24% (SB), values that are similar to the distortion products estimated for these subjects in Experiment 1 (1.7% and 3.4%, respectively). The clear implication is that the vDVRs to the LM + CM stimuli resulted almost entirely from first-order distortion products.

The hDVR data obtained with vertical LM + CM gratings stimuli were often problematical: pilot recordings had indicated that responses were often smaller and the noise levels appreciably greater. For this reason, we used only one LM + CM stimulus (16% LM and 80% CM) and obtained data from only two subjects. However, the basic findings were generally similar to those for the vDVRs: responses to these LM + CM stimuli were reversed, had ultra-short latencies, could be nulled by adding LM to the CM in phase, and showed a dependence on the added LM that was reasonably well fitted by Equation 5; see Table S4 in the Supplementary Material, which shows the best-fit parameters for both subjects, along with the values of \( \text{DLM}_{\text{DVR} = 0} \) and the estimated distortion products (from Experiment 1) in square brackets. The values of \( \text{DLM}_{\text{DVR} = 0} \) were only slightly less than the distortion products estimated in Experiment 1—by <0.8% (FAM) and <1.2% (BMS)—consistent with the idea that the hDVRs to the LM + CM stimuli also result from first-order distortion products.

**Experiment 2B: Evidence that the distortion products associated with the LM + CM stimuli are a linear function of the depth of the CM**

If the DVRs elicited by the hybrid LM + CM disparity stimuli in Experiment 2A were due mostly to distortion products secondary to compressive non-linearities early in the visual pathways then their magnitude should depend on the peak luminance of the CM. In Experiment 2B, we sought to determine if this was the case by repeating Experiment 2A with four depths of CM (and hence, four peak luminances). Thus, one eye always saw 16% LM while the other eye saw CM with one of four DCM values ranging from 40% to 80%. As in the previous experiment, the LM + CM stimuli elicited DVRs in the direction opposite to the applied 1/4 wavelength disparity, consistent with first-order distortion products. Various amounts of in-phase LM were added to each of the four CM stimuli to null these putative distortion products and thereby determine their magnitude in equivalent LM (i.e., the values of \( \text{DLM}_{\text{DVR} = 0} \)). Figure 6A shows the dependence of the vDVRs (mean pooled-difference, stimulus-locked measures) on the added LM for each of the four levels of CM (see the key) for subject BMS. The LM additions varied over only a small range and, for a given DCM, the vDVRs showed a linear dependence on the added LM (mean \( r^2 \) for these data = 0.989 ± 0.006). When the LM addition was zero (vertical dotted line in Figure 6A), the reversed vDVRs increased with increases in the DCM, as expected of distortion products. The coefficients for the best-fit linear regressions in Figure 6A are listed in Table S5 in the Supplementary Material, together with the added LM at which the fitted regressions intersected the zero-response level, i.e., the added LM required for nulling \( \text{DLM}_{\text{DVR} = 0} \). It is evident from the plots in Figure 6A and from the associated regression coefficients in Table S5 that increasing the DCM mainly increased the vertical offset (in the “backward” direction) and had little impact on the slope of the dependence on the added LM. The values of \( \text{DLM}_{\text{DVR} = 0} \) for the data in Figure 6A are plotted as a function of the DCM in Figure 6B—see the open squares—and clearly increase linearly (\( r^2 = 0.984 \)); see Table S6 in the Supplementary Material for the parameters of the best-fit linear regressions of \( \text{DLM}_{\text{DVR} = 0} \) on the DCM, which is plotted as a black line in Figure 6B. The thick gray lines in Figures 6B and 6C are best fits for a model of the compressive non-linearity and its associated distortion products and these will be described in the Discussion of Experiment 2.

Similar vDVR data were obtained from the other two subjects; see the open circles (FAM) and open diamonds (SB) plotted in Figure 6B and the associated best-fit linear regression coefficients listed in Tables S5 and S6. Interestingly, the offsets of the three regression lines in Figure 6B were very close to zero (especially BMS and SB) and their slopes varied with their magnitudes. Thus, in Figure 6B, the greater a subject’s responses, the greater that subject’s sensitivity to the DCM.

Useful hDVR data were obtained only from two subjects (FAM, BMS) and, for both, \( \text{DLM}_{\text{DVR} = 0} \) tended to increase with the magnitude of the CM, but this dependence was convincingly linear only for subject BMS (see Figure 6C). The coefficients for the best-fit
linear regressions of $DLM_{DVR=0}$ on $DCM$ for these $hDVR$ data are listed in Table S6.

**Experiment 2C: Sensitivity of the initial DVRs to differences in the depth of the stimulus modulation at the two eyes ("interocular contrast")**

In all of our previous experiments with the LM + LM and CM + CM stimuli, the depth of stimulus modulation at the two eyes was always identical, and we now recorded the DVRs to such stimuli when the depth of modulation at the two eyes was different. Figure 7A shows the $vDVR$ temporal profiles (mean pooled-difference traces) obtained from subject FAM with LM + LM stimuli and indicates that when the depth of modulation was different at the two eyes ($16% + 8%$) the responses were intermediate between those obtained with matching binocular stimuli that had the lower of the two modulations ($8% + 8%$) and those obtained with matching binocular stimuli that had the higher of the two modulations ($16% + 16%$). Any effects on latency were very minor. Figure 7B shows the mean $vDVR$ temporal profiles obtained from the same subject with CM + CM stimuli and indicates that when the depth of modulation was different at the two eyes ($80% + 80%$), the responses were again intermediate between those obtained with matching stimuli that had the lower modulation ($40% + 40%$) and the higher modulation ($80% + 80%$) in both magnitude and latency. Thus, the disparity-sensing mechanisms mediating the $vDVR$s to LM + LM and CM + CM stimuli tolerate large interocular differences in the depth of modulation ("interocular contrast"). Indeed, selectively increasing the depth of modulation at one eye increases the $vDVR$ even when it decreases the binocular matching.

Qualitatively similar $hDVR$ data were obtained from this same subject and can be seen in Figures 7C and 7D. The other two subjects (BMS and SB) yielded DVR data that were similar to those in Figure 7 in all essentials.

**Experiment 2D: Sensitivity of the initial DVRs to interocular differences in the timing of the stimuli at the two eyes**

The impact of interocular delays on the $vDVR$s is illustrated for one subject (FAM) in Figure 8. The temporal profiles (mean pooled-difference vergence velocity traces) obtained with LM + LM stimuli ($12% + 12%$) whose onsets at the two eyes were separated in time by $0–60$ ms are shown in Figure 8A. All traces are aligned on the appearance of the second (delayed) stimuli, and it is evident that interocular delays attenuated the initial $vDVR$s by an amount that increased with the delay. (In Figure 8A, the delays are indicated by the numbers to the right of the traces, and these are aligned with the associated...
Figure 7. The DVRs to LM + LM and CM + CM stimuli: effects of interocular contrast (Experiment 2C). Mean pooled-difference vergence velocity traces over time (subject FAM). (A) vDVRs with LM stimuli. (B) vDVRs with CM stimuli. (C) hDVRs with LM stimuli. (D) hDVRs with CM stimuli. Labels indicate the type and depth of modulation of the gratings seen by the two eyes, e.g., the trace labeled “LM8 + LM8” indicates that both eyes saw 8% LM and the trace labeled “CM80 + CM40” indicates that one eye saw 80% CM and the other eye saw 40% CM. Traces in continuous line indicate that the two eyes saw the same depth of modulation and traces in dotted line indicate that the two eyes saw a different depth of modulation. Each trace is mean of 136–140 responses. Upward deflections denote responses that reduce the imposed 1/4 wavelength disparity; abscissas show time since onset of stimuli.
Figure 8. **Experiment 2D**: dependence of vDVRs on interocular delay (subject FAM). (A) LM + LM stimuli; mean pooled-difference vergence velocity traces over time with interocular delays of 0–60 ms (indicated by numbers at right side aligned with the peaks of the associated traces); abscissa shows the time (in ms) measured from the appearance of the second (delayed) stimulus; upward deflections denote responses that reduce the imposed 1/4 wavelength disparity; horizontal gray line indicates the response measurement window; each trace is mean of 184–202 responses. (B) CM + CM stimuli; mean pooled-difference vergence velocity traces over time with interocular delays of 0–60 ms; conventions as in panel A; each trace is mean of 185–199 responses. (C) LM + CM stimuli; mean pooled-difference vergence velocity traces over time when the LM stimulus appeared 0–60 ms after the CM stimulus; the two horizontal gray lines beneath the traces indicate the response measurement windows, which were the same as for the LM + LM data and the CM + CM data (see labels); abscissa shows the time (in ms) measured from the appearance of the (delayed) LM stimulus; the trace for the 20-ms delay is shown in dotted line; other conventions as in panel A; each trace is mean of 186–197 responses. (D) Dependence of the pooled-difference (stimulus-locked) response measures (normalized) on interocular delay; open circles, LM + LM data; filled circles, CM + CM data; cross-in-circles linked by dotted lines, LM + CM data based on the time window used for the LM + LM data; half-filled circles linked by dotted lines, LM + CM data based on the time window used for the CM + CM data; arrow indicates the interocular delay (23 ms) corresponding to the difference in the latency of the responses to LM + LM in panel A and to CM + CM in panel B with 0-ms interocular delays; data points are each means of 184–202 measures (SDs ranged from 0.012° to 0.016°); smooth curves are the least-squares best fits obtained with Equation 6 (parameters given in Table S7 in the Supplementary Material).
initial peaks.) The impact on response latency was relatively minor. The usual stimulus-locked quantitative response measures for these data effectively map the decay of the responses to the first stimulus and show a sigmoidal dependence on the interocular delay, attenuation reaching ~45% with the longest delay interval (60 ms); see the open circles plotted in Figure 8D (and note that the plotted measures are normalized). Similar findings were obtained with the CM + CM stimuli (80% + 80%) although the attenuation was more severe, reaching ~70% with the longest delay interval; see the traces in Figure 8B and the filled circles plotted in Figure 8D. Note that the stimulus-locked response measures for a given stimulus set (LM + LM or CM + CM) used a measurement window that was time-locked to the onset of the second (delayed) stimulus and commenced always at a time determined by the shortest measured latency in the set. Thus, for the data in Figures 8A and 8B, the measurement windows extended over the periods 75–150 ms (LM + LM data in A) and 98–173 ms (CM + CM data in B), measured with respect to the onset of the second stimulus; see the gray horizontal bars beneath the traces. The normalized response measures in Figure 8D for the data obtained with the LM + LM and CM + CM stimuli were fitted with the following Equation, which is another variant of the Naka–Rushton equation (Naka & Rushton, 1966):

\[ A - R_{\text{max}} \frac{t^n}{t^n + t_{50}^n}, \]

where \( A \) is an offset (here fixed at 1), \( R_{\text{max}} \) is the asymptotic level with long delays (in normalized units), \( t \) is the interocular delay (in ms), \( t_{50} \) is the delay when response attenuation is 50% maximal (in ms), and \( n \) is an exponent. The continuous curves in Figure 8D are the best fits obtained with Equation 6 and their parameters are listed in Table S7 in the Supplementary Material.

These data indicate that the neural responses to the LM and the CM stimuli decay over time. Experiment 1 had indicated that the DVRs elicited by LM + LM stimuli have shorter latencies than the DVRs of comparable magnitude elicited by CM + CM. If this difference in latency originates in monocular sensory pathways, then there is the possibility with the LM + CM stimuli that the neural responses to the LM component reach the (hypothetical) disparity detectors and then decay away completely before the responses to the CM component even arrive, precluding any DVR. It is therefore critical to measure the extent of the decay in the DVR elicited by the LM + LM stimulus when the interocular delay approximates this difference in the latency. The measured difference in latency of the responses in Figures 8A and 8B when the interocular delays were zero is 23 ms. The curve in Figure 8D fitted to the data obtained with the LM + LM stimuli indicates that the normalized DVR is still 94.9% with an interocular delay of 23 ms (marked by a vertical arrow in Figure 8D). Thus, even if the difference in latency originates entirely in the monocular pathways—the worst case—the neural response to the LM component of the LM + CM stimulus will have decayed only 5.1% by the time the neural response to the CM component reaches the disparity detector. Experiment 2C indicated that such changes in the amplitude of the signal emanating from one eye are well tolerated by both first- and second-order disparity sensing mechanisms.

That a latency difference is not responsible for the failure of the LM + CM stimuli to generate DVRs is further reinforced by the data in Figure 8C, which are the temporal profiles (mean pooled-difference traces) obtained from the same subject with LM + CM stimuli when the LM stimulus appeared 0, 10, 20, 30, 40, 50 or 60 ms after the CM stimulus. None of the delays brought out a clear vergence response although there was a hint of a very weak one with the time delay that most nearly approximated the difference in the latencies of the vDVRs elicited by LM + LM and CM + CM stimuli (with zero interocular delay), i.e., the 20-ms delay trace (shown in dotted line in Figure 8C). Two response measures are plotted for these LM + CM data in Figure 8D based on the time windows used for the LM + LM data (cross-in-circles linked by dotted lines) and for the CM + CM data (half-filled circles linked by dotted lines). These two measures clearly gave very similar data and failed to uncover any responses that were significantly different from zero in 10 out of 14 cases (Student’s \( t \)-test). However, the two measures for the data obtained with the 20-ms delay were both statistically significant—although their normalized values were only 8% and 10% (LM + LM and CM + CM measures, respectively).

We also obtained vDVR data like those in Figure 8 from two other subjects (BMS, SB), whose responses to LM + LM and CM + CM stimuli showed a similar dependence on interocular delay; see the parameters of the best fits obtained with Equation 6 listed in Table S7. The attenuation of the vDVRs to the LM + LM stimuli when the interocular delay approximated the difference in the latency of the vDVRs to LM + LM and CM + CM stimuli (BMS: 24 ms; SB: 12 ms) was only 8.3% (BMS) and 0% (SB). In addition, with these two other subjects, selectively delaying the LM component of the LM + CM stimulus by up to 60 ms failed to uncover responses that were significantly different from zero in 18 out of 28 cases (Student’s \( t \)-test). Further, none of the response measures obtained with the time delays that most nearly approximated the differences in the latencies of the vDVRs elicited by LM + LM and CM + CM stimuli (with zero interocular delay)—20 ms (BMS), 10 ms (SB)—were significant (Student’s \( t \)-test). Even in the 10 out of 28 cases that were statistically significant, the mean normalized responses were only 9.8% ± 4.5%.

We also obtained hDVR data with vertical gratings from two subjects (FAM, BMS) and these showed a dependency on interocular delay similar to that seen in the
vDVR data. However, the attenuation with interocular delays was greater and faster than for the vDVRs, the hDVRs reaching asymptotes with delays of only 30–50 ms. These hDVR data were again well fitted by Equation 6, and the best-fit parameters are listed in Table S7. The attenuation of the hDVRs to the LM + LM stimuli when the interocular delay approximated the difference in the latency of the hDVRs to LM + LM and CM + CM stimuli (BMS: 22 ms; FAM: 13 ms) was 55.3% (BMS) and 7.8% (FAM). Selectively delaying the LM component of the LM + CM stimulus by up to 60 ms failed to bring out any significant hDVRs to these stimuli in either subject. (Student’s t-test).

Experiment 2E: Sensitivity of the initial DVRs to changes in the magnitude of the applied binocular disparity (interocular spatial phase)

Varying the magnitude of the disparity applied to the LM + CM stimulus from 0 to 1/2 wavelength in 1/8 wavelength steps failed to bring out any clear vergence responses; see the temporal profiles (mean pooled-difference velocity traces) in Figure 9B, which are the vDVRs obtained from subject FAM. Stimulus-locked response measures failed to uncover any responses that were significantly different from zero (Student’s t-test). That the somewhat coarse sampling of disparities is unlikely to have missed any responses is indicated by the temporal profiles (mean pooled-difference velocity traces) in Figure 9A, which show the vDVRs obtained from the same subject when disparities of 1/4 wavelength (black) and 1/8 wavelength (gray) were applied to LM + LM stimuli (continuous lines) and CM + CM stimuli (dotted lines). Although the responses to 1/8 wavelength were smaller than those to 1/4 wavelength, they are nonetheless still robust—the reduction here was 29% (LM + LM) and 57% (CM + CM). The hDVR data obtained from this same subject with vertical gratings were similar in all essentials, i.e., all of the LM + CM stimuli failed to elicit any significant responses (Student’s t-test) and the DVRs to 1/8 wavelength disparities were smaller than those to 1/4 wavelength disparities by 26% (LM + LM) and 19% (CM + CM).

A second subject (BMS) produced similar data, although the responses to the LM + CM stimuli were contaminated by low-level first-order distortion products, which were evident at ultra-short latency with the 1/8, 1/4, and 3/8 wavelength disparity stimuli: normalized responses were 11.6% and 11.0% (of those to the LM + LM and CM + CM stimuli, respectively, with 1/4 wavelength interocular phase shift and comparable depths of modulation). This subject’s responses to LM + LM and CM + CM stimuli were smaller when disparities were 1/8 wavelength than when they were 1/4 wavelength by 25% and 47% (vDVR) and by 12% and 27% (hDVR), respectively.

A third subject (SB) also produced very similar vDVR data, although the responses to the LM + CM stimuli were again occasionally contaminated by low-level first-order distortion products, which were evident at ultra-short latency with the 1/8 and 3/8 wavelength disparity stimuli: normalized responses were 7.7% and 3.1% (of those to the LM + LM and CM + CM stimuli, respectively, with 1/4 wavelength interocular phase shift and comparable depths of modulation). This subject’s responses to LM + LM and CM + CM stimuli were smaller when disparities were 1/8 wavelength than when they were 1/4 wavelength by 28% and 36%.

Figure 9. Experiment 2E: dependence of vDVRs on interocular spatial phase (subject FAM). Mean pooled-difference vergence velocity traces over time. (A) Data obtained with LM + LM stimuli (LM: 16%; CM: 80%) using disparities of 0, 1/8, 1/4, 3/8, and 1/2 wavelength. Upward deflections denote responses that reduce the imposed 1/4 wavelength disparity; abscissa shows time since onset of stimuli. Each trace is mean of 132–139 responses.
Discussion of Experiment 2

That the DVRs elicited by LM + CM stimuli in Experiment 2A were in the reverse direction of the applied disparity and had ultra-short latencies is consistent with compressive non-linearities in the early visual pathways that result in first-order distortion products in the signals emanating from the eye viewing the CM. Further, these reversed responses could be nullled by adding LM to the CM in phase, and in those cases in which the CM was of the same magnitude as in Experiment 1 (80%), the added LM required for nulling approximated the magnitude of the distortion products estimated in Experiment 1. Thus, any genuine DVRs elicited by the LM + CM stimuli were at best extremely weak, at least in the time period under consideration here, and must if not all of the recorded responses resulted from first-order distortion products.

Experiment 2B revealed that the amount of LM that had to be added to the CM to null the reversed responses to the LM + CM stimuli was a linear function of the DCM. We now sought to determine if this result could also be explained entirely by distortion products secondary to early compressive non-linearities. For this, we subjected our CM stimuli (with a range of LM additions) to a non-linear filter that Scott-Samuel and Georgeson (1999) had used to model the early compressive non-linearity and is given by Equation 5 (Experiment 2A) with the following parametric arrangements: the offset (A) was fixed at zero, and the exponent (n) was fixed at 1 so that the degree of compression was determined entirely by the semi-saturation constant (m50). The maximum response (Rmax) is not important here and, like Scott-Samuel and Georgeson, we set its value to 2 + m50 so that the input and output ranged from zero to 2. For each of the four CM stimuli (DCM: 40%, 50.4%, 63.5%, 80%), we added LM in various amounts up to 20%. For each stimulus combination of CM and added LM, the two sine waves describing the upper and lower envelopes of the luminance profiles were each transformed with Equation 5 and their outputs were then summed. We then determined the LM additions that gave zero summed responses for a range of values of m50. From these, we selected the single value of m50 for which the LM additions needed to achieve these zero sums with all four CM stimuli were closest to the LM additions that were needed to achieve nulling in Experiment 2B, i.e., the values of m50 that gave the best fits to each of the data plots in Figures 6B and 6C.

This procedure indicated that when the early compressive non-linearities were simulated with Equation 5, the distortion products associated with our CM stimuli were always a linear function of the DCM with zero vertical offset and a slope inversely related to the value of the semi-saturation constant, m50; see the thick gray lines plotted in Figures 6B and 6C and the best-fit values of m50 listed under “Model” in Table S6. Remarkably, these estimates of the distortion products generally provided good fits to our DVR data although there was only one free parameter (m50). Note that if the “noisy” hDVR data of FAM are excluded, the mean r2 value for the remaining data obtained with the hybrid LM + CM stimulus. These findings provide strong support for the view that distortion products secondary to compressive non-linearities early in the visual pathway are responsible for almost all of the DVRs elicited by the LM + CM stimulus. This failure of the LM + CM stimulus to generate DVRs other than distortion products is consistent with the hypothesis that the cortical mechanisms sensing disparities defined by luminance and by contrast are independent.

However, there were also some technical factors related to the binocular matching of the images seen by the two eyes that might have been responsible for this negative result with the LM + CM stimulus. These factors were addressed in three control experiments, which were concerned with the potentially disruptive effects of differences in the amplitude (Experiment 2C), timing (Experiment 2D), and phase (Experiment 2E) of the inputs from the two eyes. Experiment 2C indicated that the disparity-sensing mechanisms mediating the vergence responses to LM + LM and CM + CM stimuli tolerated large interocular differences in the depth of modulation (“interocular contrast”): A vergence response was assured so long as the two eyes each saw a stimulus that was effective when paired with one of the same type and depth of modulation. The two components of our hybrid LM + CM stimuli always met this requirement, i.e., each was very effective in eliciting DVRs when incorporated into a matching
stereo pair (LM + LM and CM + CM). This indicates that interocular differences in the amplitude and/or efficacy of the stimuli at the two eyes are unlikely to have been responsible for the failure of our hybrid LM + CM stimuli to activate the disparity-sensing mechanisms that mediate the responses when both eyes see the same type of modulation (LM + LM and CM + CM stimuli).

Experiment 1 showed that first-order DVRs have shorter latencies than second-order DVRs of comparable amplitude, and if this latency difference were to originate in the monocular pathways—the worst-case scenario—then the two signals generated by the LM + CM stimulus would arrive at the (hypothetical) disparity detectors at different times. This could be a serious problem if the response to the LM component was very transient and had sufficient time to decay away completely before the response to the CM component arrived. Experiment 2D showed that introducing interocular delays resulted in an attenuation of the DVR to LM + LM stimuli that developed over time—presumably reflecting a gradual decay in the sensory responses to the first stimulus—but this decay was sufficiently gradual that the neural signals generated by the LM component of the LM + CM stimulus would still be very robust when the delayed neural response to the CM component arrived. In fact, we estimated that the neural response to the LM component would still retain 44.7% to 100% of its initial maximal value (mean = 84.7%) when the response to the CM component arrived, and Experiment 2C had indicated that this would suffice to generate a robust DVR. Thus, interocular delays per se were unlikely to have been responsible for the failure of our hybrid LM + CM stimuli. In fact, selectively delaying the LM component of the LM + CM stimulus brought out relatively few statistically significant DVRs (overall, 14/70 cases) and these averaged <10% of those to the LM + LM and CM + CM stimuli (with comparable depths of modulation).

If the neural responses to luminance- and contrast-defined stimuli undergo differential spatial filtering this could affect their phase relationship and hence alter the effective disparity of the LM + CM stimulus. It will be seen in Experiment 4 that the first- and second-order DVRs have similar spatial frequency tuning functions, implying similar spatial filtering and suggesting that large differences in spatial phase are unlikely. Regardless, Experiment 2E indicated that changes in disparity of 1/8 wavelength were well tolerated by both first- and second-order DVRs, and varying the binocular phase of the LM + CM stimuli from 0 to 1/2 wavelength in 1/8 wavelength steps failed to bring out any significant vergence responses. We conclude that spatial phase shifts are unlikely to have been a major factor in the failure of the LM + CM stimuli.

It is common to invoke a major non-linearity such as half-wave or full-wave rectification to explain how the brain extracts disparities from second-order stimuli (Edwards et al., 2000; Hess et al., 1999; Hess & Wilcox, 1994; Langley et al., 1999; Pope et al., 1999b; Wilcox & Hess, 1995, 1996). Edwards et al. (2000) have argued that full-wave rectification would double the effective spatial frequency of first-order stimuli but not of second-order stimuli, rendering 1/4 wavelength disparity stimuli ineffective with the LM + CM stimulus. However, Experiment 2E included a paradigm in which the spatial frequency of the LM component of the LM + CM stimulus was half that of the CM component and this too failed to elicit significant vergence eye movements.

Taken together, these various findings indicate that the extraction of binocular disparities defined by first-order luminance and second-order contrast is accomplished by two independent cortical mechanisms, at least for the purpose of generating vergence eye movements at short latency. In addition, our luminance- and contrast-defined stimuli activate those cortical mechanisms selectively, eliciting what we have termed first- and second-order DVRs. That the responses to our contrast-defined stimuli lack any responses at the ultra-short latency typical of first-order responses reinforces the view that they are pure second-order DVRs. However, the vDVR temporal profiles with luminance-defined stimuli often showed two distinct peaks, the later one coinciding roughly with the (single) peak in the temporal profiles recorded with contrast-defined stimuli; see Figures 7A, 8A, 9A (continuous traces), and 10A. It is tempting to attribute this delayed second peak to second-order processing, but we think this unlikely because it failed to show up when we used the hybrid LM + CM stimuli, i.e., the responses to our luminance-defined stimuli are pure first-order DVRs.

The idea of separate first- and second-order DVRs is consistent with the proposal that stereopsis has separate mechanisms for sensing depth based on first- and second-order information (Hess & Wilcox, 1994, 2008; Langley et al., 1998, 1999; Wilcox & Hess, 1997, 1998). Studies of other visual functions also report that stimuli defined by luminance are processed independently of stimuli defined by contrast. For example, Schofield and Georgeson (1999) reported that a background defined by LM could facilitate the detection of LM but not of CM, and one defined by CM could facilitate the detection of CM but not of LM. There is also considerable evidence that motion defined by luminance is processed separately from motion defined by contrast (Derrington, Badcock, & Henning, 1993; Ledgeway & Smith, 1994; Lu & Sperling, 1995; Mather & West, 1993; Scott-Samuel & Georgeson, 1999). However, Edwards et al. (2000) have reported that LM + CM stimuli give rise to percepts of depth and they concluded that first- and second-order monocular signals are pooled prior to the extraction of depth. The stimuli in the study of Edwards et al. were very similar to those in the present one in terms of overall size, temporal frequency, spatial frequency, depth of modulation, duration, and the magnitude of the binocular disparities. One concern, therefore, is that the subjects in the experiments of Edwards et al. might have been responding to distortion products, as in the present study. Indeed, the authors state “that ‘correct’
performance does not necessarily mean that the observer
saw the depth in the direction that corresponded to the
nearest-neighbor match” (Edwards and Schor, 1999,
p. 2647), perhaps indicating that at least some of the
depth percepts were in the “reversed” direction, like the
distortion products in the present study. Nonetheless,
Edwards et al. (2000) argue that the depth percepts in
their study were not mediated by distortion products, in
part because “unambiguous motion was not perceived”
when the same first- and second-order stimuli were used
in alternation in the motion paradigm of Ledgeway and
Smith (1994) and Scott-Samuel and Georgeson (1999).
Whatever the resolution of this matter, we conclude that
the DVRs in the present study do not have access to the
pooled first- and second-order monocular signals invoked
by Edwards et al. There are other data which indicate that

Figure 10. **Experiment 3A**: dependence of vDVRs on the depth of modulation when the type and depth are identical at the two eyes. (A) Mean pooled-difference vergence velocity traces over time (subject BMS) obtained with LM + LM stimuli as the DLM ranged from 1% to 48% (indicated by numbers at right side aligned with the peaks of the associated traces); upward deflections denote responses that reduce the imposed 1/4 wavelength disparity; horizontal gray line beneath the traces indicates the response measurement window; abscissa shows time since onset of stimuli; each trace is mean of 93–95 responses. (B) Mean pooled-difference vergence velocity traces over time (subject BMS) obtained with CM + CM stimuli as the DCM ranged from 8% to 80%; conventions as in panel A; each trace is mean of 93–95 responses. (C) Dependence of latency of the vDVRs on the depth of modulation (subject BMS); open squares, LM + LM data; filled squares, CM + CM data; smooth curves are the least-squares best fits obtained with Equation 6. (D) Dependence of the pooled-difference (stimulus-locked) response measures on the depth of modulation (three subjects; see key); open symbols, LM + LM data; filled symbols, CM + CM data; data points are means of 93–95 (BMS; SDs ranged from 0.005° to 0.009°), 195–200 (FAM; SDs ranged from 0.009° to 0.015°), and 39–45 (SB; SDs ranged from 0.005° to 0.010°) measures; smooth curves are the least-squares best fits obtained with Equation 5 (parameters given in Table S8 in the Supplementary Material).
vergence eye movements and stereopsis do not always utilize the same disparity signals. For example, vergence eye movements can be elicited at short latencies when binocular disparities are applied to dense anticorrelated random-dot patterns (Masson et al., 1997; Takemura et al., 2001), even when those patterns are perceived as rivalrous and do not support depth perception (Cogan et al., 1993, 1995; Cumming et al., 1998; Masson et al., 1997). These latter findings were seen as consistent with the idea that the initial DVR derives its visual input from an early stage of cortical processing prior to the level at which depth percepts are elaborated (Masson et al., 1997).

**Experiment 3: The dependence of the initial first- and second-order DVRs on the depth of modulation**

In the previous experiment, we designed special luminance- and contrast-defined stimuli—appropriately supplemented with CM and LM, respectively—with the object of selectively activating the neural mechanisms that sense first- and second-order disparities. These special stimuli provided evidence that first-order luminance and second-order contrast are processed by separate neural mechanisms, at least for the purpose of generating short-latency vergence eye movements. This meant that, with appropriately designed stimuli, these vergence eye movements could be used to selectively explore the visual properties of those first- and second-order mechanisms. In the present experiments (Experiment 3), in any given trial, the two eyes always saw the same type of modulation, i.e., both eyes saw either luminance- or contrast-defined stimuli, and our major concern was with the dependence of the associated first- and second-order DVRs on the depth of the stimulus modulation. We report that the DVRs to both types of stimuli showed sigmoidal dependence on the depth of the modulation, although the first-order DVRs showed lower thresholds and also began to saturate at lower modulation levels than the second-order DVRs. (Of course, a direct quantitative comparison of their sensitivities to the depth of modulation is difficult because the adequate stimuli for first- and second-order DVRs are defined differently.) In any given trial, the images seen by the two eyes in this experiment always had the same depth of modulation. In a second experiment, we fixed the modulation at one eye at a moderately high level and documented the dependence of the DVRs on the depth of modulation at the other eye. We report that high-level modulation at one eye, whether luminance- or contrast-defined, renders the associated first- and second-order DVRs hypersensitive to low-level modulation (of the same type) at the other eye.

**Methods**

The subjects, eye-movement-recording techniques, procedures, and data analysis were as in Experiments 1 and 2.

**Visual stimuli**

The disparity stimuli were 1-D vertical or horizontal sinusoidal gratings that differed in phase at the two eyes by 1/4 wavelength and were created by modulating the luminance or the contrast of DBN patterns as in Experiment 1. The first-order distortion products associated with the CM were nulled by adding in-phase LM in accordance with the findings in Experiments 1, 2A, and 2B: the absence of any DVRs with the ultra-short latencies characteristic of first-order distortion products indicated that this nulling was successful.

**Experiment 3A: Dependence on the depth of modulation when identical at the two eyes**

The parameters of the luminance-defined gratings were as follows: Spatial frequencies were selected to be close to optimal for the vDVRs and hDVRs of each subject (based on data in Experiment 4): 0.29 and 0.15 cycles/° (BMS), 0.29 and 0.16 cycles/° (FAM), and 0.33 and 0.11 cycles/° (SB), respectively; in any given trial, the DLM could be 0%, 1%, 2%, 4%, 8%, 16%, 32%, or 48%.

The parameters of the contrast-defined gratings were as follows: Spatial frequencies were selected to be close to optimal for the vDVRs and hDVRs of each subject (based on data in Experiment 4): 0.24 and 0.17 cycles/° (BMS), 0.28 and 0.22 cycles/° (FAM), and 0.31 and 0.16 cycles/° (SB), respectively; in any given trial, the DCM could be 8%, 16%, 32%, 48%, 64%, or 80%.

**Experiment 3B: Dependence on the depth of modulation at one eye when the other eye sees a fixed, high-level modulation**

The parameters of the luminance-defined gratings were as follows: Spatial frequencies were always 0.25 cycles/°; the DLM at one eye was always 16% and at the other eye varied from trial to trial: 0.25%, 0.5%, 0.75%, 1.0%, 1.25%, 1.5%, and 2.0%; because these levels of modulation were substantially less than those previously used in Experiment 3A, additional controls were included in which this same selection of modulations was used (0.25%, 0.5%, 1.0%, 1.5%, and 2.0%), plus a few additions (2.5%, 3.0%, and 3.5%) while the two eyes always saw the same modulation in any given trial.

The parameters of the contrast-defined gratings were as follows: Spatial frequencies were optimized for each subject and were as for the contrast-defined gratings in Experiment 3A; the DCM at one eye was always 80% and at the other eye varied from trial to trial: 1%, 2%, 4%, 8%, 16%, 32%, 48%, 64%; because these levels of modulation were again less than those previously used...
in Experiment 3A, additional controls were included in which this same selection of modulations was used (1%, 2%, ..., 64%) while the two eyes always saw the same modulation in any given trial.

Results

**Experiment 3A: Dependence on the depth of modulation when identical at the two eyes**

The initial DVRs showed clear increases in amplitude as the depth of modulation increased (at both eyes) and this is evident from the traces in Figures 10A (LM data) and 10B (CM data), which each show the mean pooled-difference vDVR profiles over time obtained from subject BMS with particular depths of modulation (indicated in % by the numbers at the right side aligned with the peaks of the associated traces). The growth in the amplitude of the responses with increases in the depth of modulation was accompanied by decreases in the latency of onset, and this time shift was much more pronounced for the responses to the CM stimuli. The minimum latency of the vDVRs to LM stimuli is obviously less than that to the CM stimuli—note that the abscissas in Figures 10A and 10B are identical—and this difference is clear from the measured values plotted against the depth of modulation in Figure 10C, with the LM data shown in open symbols and the CM data in filled symbols. Larger responses always tended to have shorter latencies, especially with the CM stimuli, and the complete absence of any ultra-short-latency responses to the CM stimuli indicates that first-order distortion products were successfully nulled by the added in-phase LM. It is also evident that although LM stimuli of low contrast produced responses of long latency—comparable with the latency of the responses to CM stimuli—the amplitudes of these LM responses were much smaller than those of the CM responses of similar latency: If the CM responses had been due to distortion products, their amplitudes would have been much lower than they are, like those to the low-contrast LM stimuli. The usual stimulus-locked quantitative response measures are plotted in Figure 10D and show sigmoidal dependence on the depth of modulation, but BMS’s threshold for a response to the LM stimuli (open squares) is almost an order of magnitude less than that to the CM stimuli (filled squares). The continuous curves in Figure 10D are the least-squares best fits obtained with Equation 5 (with offset A fixed at zero), and the values of the associated parameters are listed in Table S8 in the Supplementary Material. The values of the free parameters, \( m_{50} \) and \( n \), are substantially larger for the CM data than for the LM data and this was also true for the other two subjects (FAM, SB), whose data are included in Figure 10D and Table S8. When the latency measures were used to obtain response-locked measures of the vDVRs, dependence on contrast showed the same trends as with the stimulus-locked measures, and the data for each subject were again well fit by Equation 5, although these fits were based on fewer data points and were never quite as good as for the stimulus-locked data. The best-fit values of the free parameters using the response-locked measures are listed in round parentheses in Table S8.

The hDVR data of all three subjects showed the same tendencies as their vDVR data and this is evident from the parameters of the best fits obtained with Equation 5 (with offset \( A \) fixed at zero), which are listed separately in Table S8. Thus, for each of the three subjects, the values of \( m_{50} \) and \( n \) are again substantially larger for the CM data than for the LM data.

**Experiment 3B: Dependence on the depth of modulation at one eye when the other eye sees a fixed, high-level modulation**

The initial vDVRs to luminance-defined vertical disparities when the modulation at one eye was always 16% while the modulation at the other eye on any given trial ranged from 0.25% up to 2% (designated “LM16 + LM” stimuli) are shown for subject BMS in Figures 11A and 11C. It is apparent from the mean pooled-difference vDVR profiles over time shown in Figure 11A that, even with such low-level stimuli, there was a very clear growth of the responses as modulation increased. Indeed, the usual stimulus-locked quantitative response measures plotted in Figure 11C as a function of the DLM show a very rapid exponential growth with no obvious sign of an initial threshold (see the open squares). These data deviate dramatically from those obtained when the modulation was applied equally to the two eyes (designated “LM + LM” stimuli). The latter data are shown as mean pooled-difference vDVR profiles over time in Figure 11B and as filled squares in Figure 11C. It is clear from these figures that when the modulation was equal at the two eyes the dependence on DLM was sigmoidal, incrementing only after exceeding a threshold, and all responses fell far short of those obtained when one eye saw a high modulation. Clearly, applying a high modulation to one eye rendered the other eye hypersensitive to modulation, a phenomenon we refer to as, **dichoptic facilitation**. In fact, when one eye saw a high modulation, robust responses could be obtained by applying modulations at the other eye that, if applied equally to the two eyes, would have been subthreshold. The plots in Figure 11C were both well fitted by Equation 5 (with offset \( A \) fixed at zero) and the best-fit parameters are listed in Table S9 in the Supplementary Material. Note, in particular, that the parameter \( n \) was almost three times higher for the fits to the LM + LM data than for the fits to the LM16 + LM data, and this was also true for the vDVR data of the other two subjects, which are also listed in Table S9 and show the same general features. This dichoptic facilitation was also seen in the hDVR data obtained with the LM16 + LM horizontal disparity stimuli, which is apparent from the best-fit...
parameters obtained with Equation 5 (with offset $A$ fixed at zero) that are also listed in Table S9. However, hDVR data were obtained from only two subjects (BMS, FAM).

Sample vDVR data obtained from subject BMS with contrast-defined vertical disparities are shown in Figure 12, which is organized like Figure 11. The data obtained when the DCM at one eye was always 80% while the DCM at the other eye on any given trial ranged from 1% up to 64% (designated “CM80 + CM” stimuli) are shown in Figures 12A and 12C (open squares), while those obtained when both eyes saw the same DCM (designated “CM + CM” stimuli) are shown in Figures 12B and 12C (filled squares). The general form of the response plots in Figure 12C shows a striking resemblance to those in Figure 11 and the data obtained with the CM80 + CM stimuli show clear dichoptic facilitation. The plots in Figure 12C were also well fitted by Equation 5 (with offset $A$ fixed at zero) and their best-fit parameters are listed in Table S10 in the Supplementary Material, together with those for the other two subjects, which showed the same general features, including dichoptic facilitation. The hDVR data obtained with the CM80 + CM horizontal disparity stimuli showed these same features and the best-fit parameters obtained with Equation 5 (with offset $A$ fixed at zero) are also listed in Table S10, although we obtained such data for only two subjects (BMS, FAM).

Discussion of Experiment 3

When the depth of modulation at the two eyes was the same, the initial DVRs showed sigmoidal dependence on the depth of modulation but the thresholds for activation were almost an order of magnitude lower for the luminance-defined stimuli than for the contrast-defined stimuli. That the thresholds to contrast-defined stimuli were higher than to luminance-defined stimuli was not due to the longer latencies of the former and our use of stimulus-locked measurement windows: for each data set, the measurement window commenced with the onset of the response that had the shortest measured latency in the set. The contrast dependence data obtained with the luminance-defined stimuli strongly resemble previously published data obtained with disparity stimuli created by direct sinusoidal modulation of luminance (Sheliga et al., 2006). In fact, this last study included data for two of the subjects in the present study (FAM and BMS) and these were also well fit by Equation 5 so we included their

Figure 11. Experiment 3B: dependence of vDVRs on the depth of LM at one eye when the other eye saw a fixed, high-level LM (subject BMS). (A) Mean pooled-difference vergence velocity traces over time when one eye always saw 16% LM while the other eye saw LM ranging from 0.25% to 2% (LM16 + LM stimuli); upward deflections denote responses that reduce the imposed 1/4 wavelength disparity; horizontal gray line indicates the response measurement window; abscissa shows time since onset of stimuli; each trace is mean of 134–140 responses. (B) Mean pooled-difference vergence velocity traces over time when LM was applied equally to the two eyes and ranged from 0.25% to 3.5% (LM + LM stimuli); other conventions as in panel A; each trace is mean of 136–140 responses. (C) Dependence of the pooled-difference (stimulus-locked) response measures on the depth of modulation; open squares, LM16 + LM data; filled squares, LM + LM data; SDs ranged from 0.005° to 0.008°; smooth curves are the least-squares best fits obtained with Equation 5 (parameters given in Table S9 in the Supplementary Material).
best-fit parameters in Table S8 (in square brackets) to permit easy comparison.

A high-amplitude luminance- or contrast-defined stimulus at one eye rendered the DVR hypersensitive to lower-amplitude stimuli of the same modality at the other eye, a phenomenon that we termed “dichoptic facilitation.” Indeed, modulations could be subthreshold when applied equally to the two eyes but generate robust DVRs when paired with a suprathreshold stimulus at the other eye. Thus, selectively increasing the modulation at one eye increased the associated vergence responses even when it reduced the similarity of the stereo images, a phenomenon that we also reported in Experiment 2C. A related observation was reported by Edwards, Pope, and Schor (1998), who recorded the vergence eye movements of human subjects when confronted with three Gabor patches defined by luminance modulation, a central one at fixation and two others, one on either side, which when matched with the central patch had competing crossed and uncrossed disparities. These workers reported that selectively raising the contrast of one of the two competing patches increased the likelihood that it would be the target of the vergence eye movements, and similar observations have been made on monkeys (Dearworth, Davison, Li, & Gamlin, 2005). Clearly, selective increases in contrast can bias ambiguous stereo matches. As pointed out by Petrov (2004), “interocular contrast is ubiquitous in the natural environment” because “only truly matte surfaces ... scatter incoming light equally in all directions,” so that tolerance of differences in contrast at the two eyes is useful in the real world. In line with this, Petrov showed that the perception of ambiguous stereo matches within global patterns was also resolved in favor of the higher (dissimilar) contrast rather than the lower (similar) contrast. However, a number of studies have reported that stereoacuity is degraded when the paired images are of unequal contrast. Thus, increasing the contrast equally at the two eyes improves stereoacuity but increasing the contrast at only one eye degrades it, especially with low spatial frequencies like those used in the present study, a phenomenon often referred to as contrast paradox (Cormack, Stevenson, & Landers, 1997; Halpern & Blake, 1988; Legge & Gu, 1989; Schor, Edwards, & Pope, 1998; Schor & Heckmann, 1989; Stevenson & Cormack, 2000). It is apparent that the impact of interocular contrast depends critically on the nature of the binocular task.

Figure 12. Experiment 3B: dependence of vDVRs on the depth of CM at one eye when the other eye saw a fixed, high-level CM (subject BMS). (A) Mean pooled-difference vergence velocity traces over time when one eye always saw CM of 80% while the other eye saw CM ranging from 1% to 48% (CM80 + CM stimuli); upward deflections denote responses that reduce the imposed 1/4 wavelength disparity; horizontal gray line indicates the response measurement window; abscissa shows time since onset of stimuli; each trace is mean of 198–205 responses. (B) Mean pooled-difference vergence velocity traces over time when CM was applied equally to the two eyes and ranged from 1% to 48% (CM + CM stimuli); other conventions as in panel A; each trace is mean of 199–205 responses. (C) Dependence of the pooled-difference (stimulus-locked) response measures on the depth of modulation; open squares, CM80 + CM data; filled squares, CM + CM data; SDs ranged from 0.006° to 0.010°; smooth curves are the least-squares best fits obtained with Equation 5 (parameters given in Table S10 in the Supplementary Material).
Phenomena analogous to dichoptic facilitation have been described with monocular stimuli. For example, in a motion sequence involving 1/4 wavelength shifts between frames that have high and low contrast, the strength of the motion signal is proportional to the product of the two contrasts (Lu & Sperling, 2001; van Santen & Sperling, 1984). One consequence of this so-called product rule is that even low-contrast frames that would be subthreshold if they made up all the frames in the sequence can be effective when interleaved with high-contrast frames. However, this effect is seen only when the higher-contrast frames have contrast <8.5%, which is well below the 16% LM that we used to elicit dichoptic facilitation in our study. Another point of similarity with dichoptic facilitation is that the product rule applies to both first- and second-order motion stimuli (Werkhoven, Sperling, & Chubb, 1994). More recently, a similar effect has been demonstrated for orientation discrimination, whereby adding a low-contrast textured stimulus (that is below its own detection threshold) to an ambiguously oriented high-contrast texture renders the orientation salient, a phenomenon referred to as contrast amplification (Appelbaum, Lu, & Sperling, 2007).

**Experiment 4: The dependence of the initial first- and second-order DVRs on spatial frequency**

In this experiment, the two eyes always saw the same type and level of modulation, and our concern was with the dependence of the first- and second-order DVRs on the spatial frequency of the modulation. We report that the initial DVRs elicited by luminance- and contrast-defined stimuli both showed Gaussian dependence on log spatial frequency, even peaking at similar frequencies, although the vDVR data invariably peaked at higher frequencies than the hDVR data.

**Methods**

The subjects, eye-movement-recording techniques, procedures, and data analysis were as in Experiment 1.

**Visual stimuli**

The disparity stimuli were 1-D vertical or horizontal sinusoidal gratings that differed in phase at the two eyes by 1/4 wavelength and were created by modulating the luminance or the contrast of DBN patterns as in Experiment 1. In any given trial, the two eyes always saw the same modulation and this could be either 16% LM or 80% CM. The first-order distortion products associated with the contrast modulations were nulled by adding in-phase LM in accordance with the findings in Experiments 1, 2A, and 2B. Spatial frequencies in any given trial could be 0.025, 0.05, 0.1, 0.2, 0.4, 0.8, 1.6, or 3.2 cycles/°, although because of their differing response ranges, the highest value was omitted from the horizontal disparities and the lowest value was omitted from the vertical disparities.

**Results**

The initial DVRs showed clear dependence on spatial frequency and this is evident from the traces in Figures 13A (LM data) and 13B (CM data), which each show the mean pooled-difference vDVR profiles over time obtained from subject FAM with particular spatial frequencies (each indicated in cycles/° by a number at the right side aligned with the peak of the associated trace). With both luminance- and contrast-defined stimuli, response amplitudes tend to peak with spatial frequencies of 0.2–0.4 cycles/°. Note that responses to spatial frequencies above the optimum are shown in dotted line in Figures 13A and 13B (and indicated by the numbers in square brackets to the right of the traces). Interestingly, the response profiles with luminance-defined stimuli below the optimum frequency show two distinct peaks, the later one roughly coinciding in time with the (single) peak in the response to the contrast-defined stimulus, but this later peak is absent with luminance-defined stimuli above the optimum frequency. As usual, the minimum latency of the vDVRs to LM stimuli is clearly less than that to the CM stimuli—by ~40 ms in Figure 13—and this difference is apparent from the measured latency values plotted against the spatial frequency (on a log abscissa) in Figure 13C, with the LM data shown in open symbols and the CM data in filled symbols. Larger responses tend to have shorter latencies, especially with the LM stimuli, and the absence of any ultra-short-latency responses to the CM stimuli indicates that the early first-order distortion products were again successfully nulled by the added in-phase LM. The usual stimulus-locked quantitative response measures are plotted as a function of the spatial frequency (on a logarithmic scale) in Figure 13D and the vDVR data obtained with LM and CM stimuli both display a band-pass dependence on log spatial frequency that is well captured by Gaussian functions (continuous curves in Figure 13D), cf. the spatial frequency dependence of the initial vergence responses when 1/4 wavelength disparities are applied to 1-D sinusoidal gratings composed of pure luminance modulation (Sheliga et al., 2006). In fact, all four data sets (vDVR and hDVR data obtained with LM and CM stimuli) obtained from all three subjects showed this Gaussian dependence on log spatial frequency, and the three parameters of the best-fit Gaussian functions—peak amplitude (Apeak), spatial frequency of the peak (f0), and standard deviation (σ)—are listed in Table S1 in the
Supplementary Material. It is apparent from these listings that, for the vDVR data, the values of $f_o$ were slightly higher with the LM stimuli than with the CM stimuli (mean values = $0.300 \pm 0.022$ vs. $0.275 \pm 0.036$ cycles/°, respectively) whereas, for the hDVR data, the opposite was true (mean values = $0.138 \pm 0.025$ vs. $0.183 \pm 0.033$ cycles/°, respectively). It is also evident from this that the values of $f_o$ were somewhat higher for the vDVR data than for the hDVR data with both LM and CM stimuli. However, the values of $\sigma$ were consistently slightly higher with LM stimuli than with CM stimuli for both the vDVR data (mean values = $0.45 \pm 0.03$ vs. $0.37 \pm 0.06$ cycles/°, respectively).
Discussion

The DVRs to both luminance- and contrast-defined stimuli showed Gaussian dependencies on log spatial frequency, and although the responses to these two kinds of modulation showed consistent differences in their peaks ($f_o$) and standard deviations ($\sigma$), these differences were, in general, quite small. Experiment 2 had indicated that the two kinds of modulation are processed independently so that these new results must reflect similarities in the spatial characteristics of the underlying sensing mechanisms. It is of some interest, therefore, that Schofield and Georgeson (1999) have reported that detection thresholds for luminance- and contrast-modulated noise stimuli show very similar dependencies on spatial frequency—although rather more low-pass than in the present study—and yet, as pointed out in the Discussion of Experiment 2, these workers also obtained clear evidence that the two kinds of stimuli are processed quite separately. In a previous study (Sheliga et al., 2006), we recorded the initial DVRs elicited with gratings defined by sinusoidal luminance modulation and they too showed Gaussian dependence on log spatial frequency. Those data compare well with the present data obtained with the LM stimuli, the only difference being that the preferred spatial frequencies ($f_o$) were slightly lower in the present study. Two of our present subjects (FAM and BMS) were in both studies and their data from that previous study are included in Table S11 (in square brackets).

Interestingly, in Figure 13 the response profiles obtained with luminance-defined stimuli below the optimum frequency show two distinct peaks, the second later one roughly coinciding with the (single) peak in the response to the contrast-defined stimulus. However, this delayed second peak was absent with luminance-defined stimuli when defined by contrast and the opposite was true with the hybrid LM + CM stimuli. We think it significant that in Figure 13 the 0.05 cycles/° stimulus elicited a delayed second peak when defined by luminance but no response when defined by contrast and the opposite was true with the 0.8 cycles/° stimulus. Clearly, the delayed second peak did not conform to the spatial frequency characteristics of the second-order DVRs, reinforcing the idea that the responses to the luminance-defined stimuli are pure first-order DVRs.

Acknowledgments

This research was supported by the Intramural Program of the National Eye Institute at the NIH. HAR was also supported by the Alexander von Humboldt Foundation (Germany). Commercial relationships: none. Corresponding author: Boris M. Sheliga. Email: bms@lsr.nei.nih.gov. Address: Laboratory of Sensorimotor Research, National Institutes of Health, Building 49 Room 2A50, 49 Convent Drive, Bethesda, MD 20892-4435, USA.

Footnotes

1Subject SB contributed vDVR data only.
2Experiment 3 will show that a moderately high contrast LM stimulus at one eye—as in the present experiment—renders the DVR extremely sensitive to even very low-level LM at the other eye: dichoptic facilitation. This makes it very difficult to exactly null all of the first-order distortion products.
3Subject SB contributed vDVR data only.
4Note that the abscissas in Figures 10C and 10D have logarithmic scales.
5However, quantitative comparisons of their dependencies on the depth of modulation are problematic because these stimuli are defined so differently.

References


disparity in macaque V1. *Journal of Neurophysiology, 87*, 209–221. [PubMed] [Article]


