Linear binocular combination of responses to contrast modulation: Contrast-weighted summation in first- and second-order vision

Jiawei Zhou

McGill Vision Research, Department of Ophthalmology, McGill University, Montreal, QC, Canada

Mark A. Georgeson

School of Life & Health Sciences, Aston University, Birmingham, UK

Robert F. Hess

McGill Vision Research, Department of Ophthalmology, McGill University, Montreal, QC, Canada

Binocular combination for first-order (luminance-defined) stimuli has been widely studied, but we know rather little about this binocular process for spatial modulations of contrast (second-order stimuli). We used phase-matching and amplitude-matching tasks to assess binocular combination of second-order phase and modulation depth simultaneously. With fixed modulation in one eye, we found that binocularly perceived phase was shifted, and perceived amplitude increased almost linearly as modulation depth in the other eye increased. At larger disparities, the phase shift was larger and the amplitude change was smaller. The degree of interocular correlation of the carriers had no influence. These results can be explained by an initial extraction of the contrast envelopes before binocular combination (consistent with the lack of dependence on carrier correlation) followed by a weighted linear summation of second-order modulations in which the weights (gains) for each eye are driven by the first-order carrier contrasts as previously found for first-order binocular combination. Perceived modulation depth fell markedly with increasing phase disparity unlike previous findings that perceived first-order contrast was almost independent of phase disparity. We present a simple revision to a widely used interocular gain-control theory that unifies first- and second-order binocular summation with a single principle—*contrast-weighted summation*—and we further elaborate the model for first-order combination. Conclusion: Second-order combination is controlled by first-order contrast.

Introduction

The human visual system is able to combine different visual inputs in the two eyes and to fuse them, creating the appearance of a single cyclopean image. To understand binocular vision, then, it is crucial to understand the nature of this process of binocular combination. This has been addressed by systematically varying the difference between the monocular inputs. Several types of image difference have been studied, including local phase (Ding & Sperling, 2006; Huang, Zhou, Lu, Feng, & Zhou, 2009; Zhou, Jia, Huang, & Hess, 2013), luminance contrast (D. H. Baker, Wallis, Georgeson, & Meese, 2012a; Ding, Klein, & Levi, 2013a, 2013b; Huang, Zhou, Lu, & Zhou, 2011; Huang, Zhou, Zhou, & Lu, 2010; Legge, 1984a, 1984b), stereo disparity (Hou, Huang, Liang, Zhou, & Lu, 2013; Legge & Gu, 1989; Reynaud, Zhou, & Hess, 2013), global motion (Hess, Hutchinson, Ledgeway, & Mansouri, 2007; Mansouri, Thompson, & Hess, 2008), and global orientation (Zhou, Huang, & Hess, 2013). Different models have been constructed for binocular combination involving these different kinds of visual information. One common feature of these models is the need for an interocular contrast-gain control weighting stage (Ding et al., 2013a, 2013b; Ding & Sperling, 2006; Hou et al., 2013; Huang et al., 2009; Huang et al., 2011; Huang et al., 2010; Meese, Georgeson, & Baker, 2005, 2006; Meese & Hess, 2004; Meese & Summers, 2009).

The visual stimuli used in most of the above studies were based on variations of luminance across space, so-

Citation: Zhou, J., Georgeson, M. A., & Hess, R. F. (2014). Linear binocular combination of responses to contrast modulation: Contrast-weighted summation in first- and second-order vision. *Journal of Vision*, *14*(13):24, 1–19, http://www.journalofvision.org/content/14/13/24, doi:10.1167/14.13.24.

doi: 10.1167/14.13.24 Received June 30, 2014; published November 25, 2014 ISSN 1534-7362

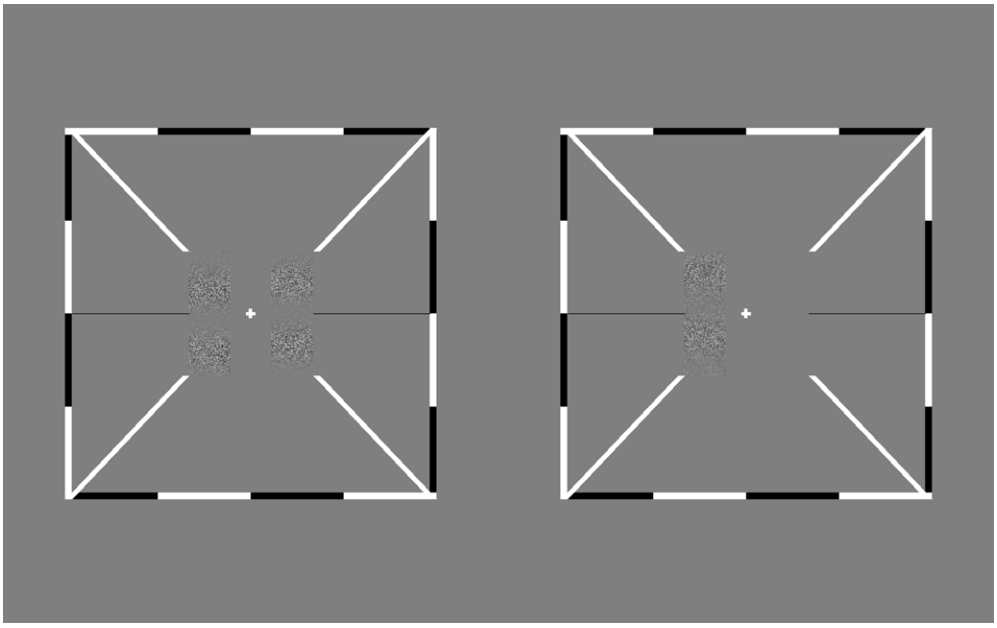

Figure 1. Stimulus display. The left and right panels were presented to the left and right eyes using goggles. The two gratings on the left side of fixation in the two eyes' views had equal and opposite phase shift relative to the center of the frame and were perceptually combined to form a binocular cyclopean image. A monocular "probe" grating was presented to the right side of fixation in the left or the right eye in different trials. Observers adjusted the phase and modulation depth of the probe grating to match the binocular cyclopean image.

called first-order stimuli, as distinct from visual information defined by spatial variation in contrast (Dakin & Mareschal, 2000; Schofield & Georgeson, 2003), texture (Cavanagh & Mather, 1989; Dosher & Lu, 2006; Regan, 2000; Sutter & Graham, 1995; Werkhoven, Sperling, & Chubb, 1993), or orientation (Larsson, Landy, & Heeger, 2006). Images that have structure defined by such higher-order variations are often referred to as second-order stimuli (Cavanagh & Mather, 1989), and it has been suggested that first- and second-order stimuli are processed separately within the visual cortex (C. L. Baker, 1999; Chubb & Sperling, 1988, 1989; Dosher & Lu, 2006; Schofield & Georgeson, 1999; Wilson, 1999).

Many studies have focused on binocular combination of first-order information, but only a few have examined the combination of second-order stimuli. Recently, Zhou, Liu, Zhou, and Hess (2014) studied the binocular phase combination of second-order gratings by varying the relative amounts of contrast modulation (CM) seen by the left and right eyes. Their results were close to the prediction made by simple linear summation (or averaging) of the two envelopes. At first sight, this appears to be different from first-order binocular combination because there a nonlinear mechanism (contrast-gain control) best described performance. However, as we shall see, the linear summation of CM signals can also be interpreted within a contrast-gain control framework similar to that describing first-order binocular combination.

The study by Zhou et al. (2014) also found that second-order phase combination was the same when the carriers in the two eyes were correlated, anticorrelated, or uncorrelated. This is consistent with two previous reports. In one, Wilcox and Hess (1996) found that stereo acuity for CM was the same for correlated and uncorrelated carriers, and in another, M. Georgeson and Schofield (2011) found that binocular summation of CM at threshold was consistent with linear summation and was similar for carriers in the two eyes that were correlated, anticorrelated, or uncorrelated. All three studies found that second-order binocular combination was not affected by the degree of interocular correlation of the carriers, and this implies that binocular combination of second-order CM signals occurs after monocular extraction of the second-order modulations.

Motivation

In this paper, we examine binocular combination of CM by looking at perceived spatial phase and perceived modulation depth (i.e., amplitude) of the CM. This was done with a dichoptic phase- and amplitude-matching task (Figure 1) as used for first-order binocular phase and contrast combination by

Huang et al. (2010). They showed that binocular contrast combination was independent of the interocular phase difference (up to 90°) and suggested that binocular phase combination and binocular contrast combination involved a common nonlinear contrast-gain control stage but different pathways thereafter. With that in mind, we were interested in knowing whether binocular combination of CM amplitudes

- Is phase-invariant like first-order contrast combination
- Involves linear summation similar to second-order binocular phase combination
- Is independent of the interocular correlation of the carriers, similar to second-order binocular phase combination

We found that the perceived CM amplitude varied markedly with both the interocular ratio of modulation depths and the interocular phase difference but not with the carrier correlation.

Road map for this paper

We describe new psychophysical experiments on the binocular combination of second-order (CM) gratings. We find that both second-order phase and amplitude can be explained by a weighted linear summation of second-order modulations, with weights that depend only on first-order carrier contrast. This model is described in Appendix 1.

We discuss the idea that a common principle (contrast-weighted summation) is at work in both first-order and second-order combination.

We offer an elaboration of this model to account for a wide range of previous experimental data *for first-order combination only*. The details of this analysis and model fitting are in Appendix 2.

We return to the theme of commonality and a shared outcome in first- and second-order vision: "ocularity invariance."

Materials and methods

Observers

Three adults (ages: 28, 31, and 33 years old) with normal or corrected-to-normal vision participated in the experiments. Except for the first author, all subjects were naive as to the purpose of the experiment. Written informed consent was obtained from each of them. The study was approved by the Institutional Review Boards of McGill University.

Apparatus

All stimuli were generated and controlled by an Apple Mac computer running Matlab (MathWorks, Natick, MA) with the PsychToolBox 3.0.9 extension (Brainard, 1997; Pelli, 1997). The stimuli were dichoptically presented with Z800 pro goggles (eMagin Corp., Washington, DC), which had a simulated viewing distance of 3.6 m, a spatial resolution of 800 × 600 (corresponding to 30.26° × 22.69° in visual field), a refresh rate of 60 Hz, and a mean luminance of 160 cd/m^2 in each eye. These organic light-emitting diode microdisplays are linear in luminance response (Black, Thompson, Maehara, & Hess, 2011) and exhibit pixel independence in image presentation (Cooper, Jiang, Vildavski, Farrell, & Norcia, 2013); thus we would not expect any nonlinear distortions of our stimuli due to the display equipment (Klein, Hu, & Carney, 1996).

Design

A dichoptic phase- and amplitude-matching paradigm (Huang et al., 2011; Huang et al., 2010) was used to quantify the perceived phase and modulation depth of the binocular percept. In the test stimulus, two horizontal sine-wave CM gratings with equal and opposite phase shifts of $\theta/2$ ($\theta = 0°, 45°, 90°$) were presented dichoptically to the left side of the central fixation point in the two eyes. This pair of CM test gratings was viewed through the goggles and combined perceptually to create a single cyclopean grating. A monocular horizontal sine-wave CM grating (the "probe") was presented to the right side of fixation in the left or the right eye (Figure 1). Observers adjusted the phase and modulation depth of the monocular probe grating to match the binocular percept of the test stimulus.

All gratings (test and probe) had the same mean carrier contrast of 0.2. The modulation depth of the test gratings was fixed at 0.8 in the nondominant eye and varied across trials from 0 to 0.8 in the dominant eye. The initial phase and modulation depth of the probe grating were randomized in each trial and then adjusted by subject to match the binocular percept of the test. The probe was presented either to the left or the right eye, and the other eye saw uniform mean luminance (no carrier noise); the results were averaged over these two conditions to cancel potential eye bias.

Two configurations were used in measuring the perceived phase of the binocular cyclopean image to cancel any positional bias: the phase shift was $+\theta/2$ in the nondominant eye and $-\theta/2$ in the dominant eye or vice versa. Perceived phase at each interocular modulation ratio (δ) was quantified as half the difference

between the matched phases in these two configurations.

In total, there were 72 conditions (three interocular phase differences × six interocular modulation ratios × two probe eye conditions × two phase configurations). Each condition was tested four times. Two functions, i.e., perceived phase versus interocular modulation ratio (PvR function) and perceived modulation depth versus interocular modulation ratio (MvR function), were then derived for each interocular phase difference θ.

All three observers took part in Experiment 1, in which the two test gratings had correlated carriers. In Experiment 2, the first author and one of the naive observers ran additional experiments in which the carriers were either anticorrelated or uncorrelated across the eyes.

Stimuli

Stimuli were horizontally oriented sine-wave CM gratings. The gratings extended for two spatial cycles in the vertical direction (Figure 1), which subtended 6.81° of visual angle (i.e., 0.29 c/°), and 2.27° of visual angle in the horizontal direction. The center of the test gratings was 2.27° to the left of the central fixation in both eyes, and the probe grating was 2.27° to the right in either the left or the right eye. A high-contrast frame (width 0.38°; length 20.43°) with four white diagonal lines (width 0.38°; length 9.63°) was presented surrounding the gratings in each eye to help observers maintain binocular convergence and fusion. A 1-pixel-wide black reference line, aligned to the center of the low-contrast bar of the probe grating, was presented horizontally at the two sides of the gratings to assist the phase matching.

The test gratings in the two eyes were defined by

$$Lum_{nonDE}(x,y) = L_0\left\{1 - C_g.g_1(x,y) \times \left[1 - M_0 cos\left(2\pi fy \pm \frac{\theta}{2}\right)\right]\right\} \quad (1)$$

and

$$Lum_{DE}(x,y) = L_0\left\{1 - C_g.g_2(x,y) \times \left[1 - \delta.M_0 cos\left(2\pi fy \mp \frac{\theta}{2}\right)\right]\right\} \quad (2)$$

where L_0 is the background luminance, and $g_1(x, y)$ and $g_2(x, y)$ are the 2-D binary, white-noise carriers in the two eyes with correlated carriers $g_1 = g_2$, with anticorrelated carriers $g_1 = -g_2$, and with uncorrelated carriers $g_1 \neq g_2$. $C_g = 0.2$ is the contrast of carrier, $f = 0.29$ c/° is the spatial frequency of the sine-wave envelope, $M_0 = 0.8$ is the fixed modulation depth in the nondominant eye, and δ (the *modulation ratio*) is the interocular ratio of modulation depths ($\delta = 0$, 0.2, 0.4, 0.6, 0.8, 1.0). The two dichoptic gratings in the test had an equal and opposite phase shift of $\theta/2$ (relative to the center of the screen), where $\theta = 0°$, 45°, 90°.

We emphasize that for the one-eyed test condition ($\delta = 0$), the test images presented to the left side of fixation were CM gratings with modulation depth of 0.8 in the nondominant eye and 0 in the dominant eye (i.e., carrier noise without modulation). This configuration was different from that of the stimuli presented to the right of fixation, i.e., a monocular probe grating in one eye and uniform mean luminance (no carrier noise) in the other eye.

Procedure

An alignment task was provided at the beginning of each trial to ensure that the two eyes' images were correctly fused. In the alignment task, a fixation display was presented in the center of the larger high-contrast frame together with four white diagonal lines. This display consisted of binocular fixation crosses (3.78° × 3.78° squared) and four monocular dots (0.38° diameter): two in the first and third quadrants for the left eye and two in the second and fourth quadrants for the right eye. Observers were instructed to move the image seen by their nondominant eye using up, down, left, and right arrow keys to align the images from the two eyes. After achieving stable fusion, observers were asked to press the space bar. The corresponding coordinates between the two eyes were then used in subsequent measurements. Observers adjusted the phase and modulation depth of the probe grating to match the corresponding features of the test grating. They were free to select which dimension (phase or modulation depth) to adjust first and to go back and forth between them. Subjects were asked to attend to the difference of contrast of the sine-wave gratings in matching modulation depth. When satisfied with the match in both dimensions, observers pressed the space bar again to report the result. The next trial started 1 s later. During one trial, the gratings, frames, and reference lines were presented continuously in both eyes until the phase and amplitude adjustments were completed. A typical trial lasted about 10 s. Voluntary breaks were allowed. Practice trials were provided prior to data collection.

Model predictions

We considered a variety of ways in which responses to CM in each eye might be combined: simple

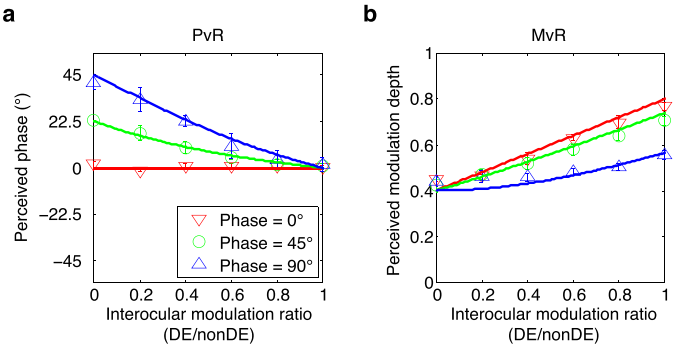

Figure 2. Experiment 1. Binocular combination of second-order stimuli with correlated carriers. (a) Perceived phase versus interocular modulation ratio (PvR) and (b) perceived modulation depth versus interocular modulation ratio (MvR) averaged over three observers at three interocular phase differences (0°, 45°, and 90°). Solid curves are predictions of the contrast-weighted linear summation model (see Appendix 1). Error bars represent standard errors.

summing, simple averaging, weighted averaging, and so on. To make predictions about matching behavior, any such rule must be applied to both the binocular test stimulus and the monocular probe in order to derive response values that match each other. We found that one simple model was consistent with all our results: weighted linear summation of the modulation waveforms in each eye, in which the weights assigned to each eye depend on the relative carrier contrasts shown to each eye. Second-order combination was controlled by first-order contrast.

On this model (see Appendix 1), the binocularly perceived phase θ' for an interocular phase difference θ and modulation ratio $\boldsymbol{\delta}$ was predicted to be

$$\theta' = tan^{-1}\left[\frac{1-\delta}{1+\delta}tan\left(\frac{\theta}{2}\right)\right] \quad (3)$$

and the matching probe modulation depth M' was

$$M' = 0.5M_0\sqrt{\delta^2 + 1 + 2\delta cos(\theta)} \quad (4)$$

Results

Experiment 1: Binocular combination of second-order stimuli with correlated carriers

The way that perceived phase varies with interocular modulation ratio (PvR curves) is shown in Figure 2a. Different colored symbols represent results for the three interocular phase differences (0°, 45°, 90°). When the test modulation was in only one eye (modulation ratio = 0), perceived phase appeared offset and matched the phase seen by that eye, but as the amplitude in the other eye increased toward equality (ratio = 1), the perceived phase shifted toward a balanced, central position (0° on the y-axis). These data replicate and extend those of Zhou et al. (2014). Colored lines represent the very close fit given by the contrast-weighted linear summation model with no free parameters. Appendix 1 gives details of this model, but it is easily summarized. The test carrier contrast is the same in both eyes; this makes the left and right eye weights equal (0.5 each), so weighted summation averages the two spatial contrast envelopes. But the probe contrast is present only in one eye; this yields ocular weights of 1 and 0 (winner take all, WTA), giving complete dominance of the probed eye with no binocular combination. These different ocular weightings lead directly to the predicted curves in Figure 2.

Matched modulation depths increased with interocular modulation ratio (MvR curves) as shown in Figure 2b. Matched modulation depth was close to 0.4 when the test modulation (0.8) was monocular (modulation ratio = 0). This is consistent with averaging of test modulations across the eyes (average of 0.8 in one eye and 0 in the other). Perceived amplitude increased as more modulation was added to the other eye (modulation ratio increased toward 1), but this increase was shallow when the phase disparity was large (90°) rather than small (0° or 45°). A repeated-measures within-subject ANOVA showed that matched modulation depth depended significantly on both modulation ratio, $F(5, 10) = 124.82$, $p < 0.001$, and phase difference, $F(2, 4) = 151.14$, $p < 0.001$, with a significant interaction, $F(10, 20) = 28.29$, $p < 0.001$. Such phase dependency of second-order combination is different from previous reports on first-order combination with the same paradigm (Huang et al., 2011; Huang et al., 2010), in which the perceived contrast of first-order gratings did not depend on the interocular phase difference. However, all our data—for both phase and amplitude matching—are in good agreement with contrast-weighted linear summation (solid curves in Figure 2a, b).

Experiment 2: Binocular combination of second-order stimuli with anticorrelated and uncorrelated carriers

The above experiment implies linear binocular combination of second-order signals when the dichoptic test gratings have correlated carriers. In a previous study of CM phase judgments, binocular combination was found to be similar when the carriers in the two eyes were correlated, anticorrelated, or uncorrelated (Zhou et al., 2014), and this suggests that second-order binocular combination occurs after monocular extrac-

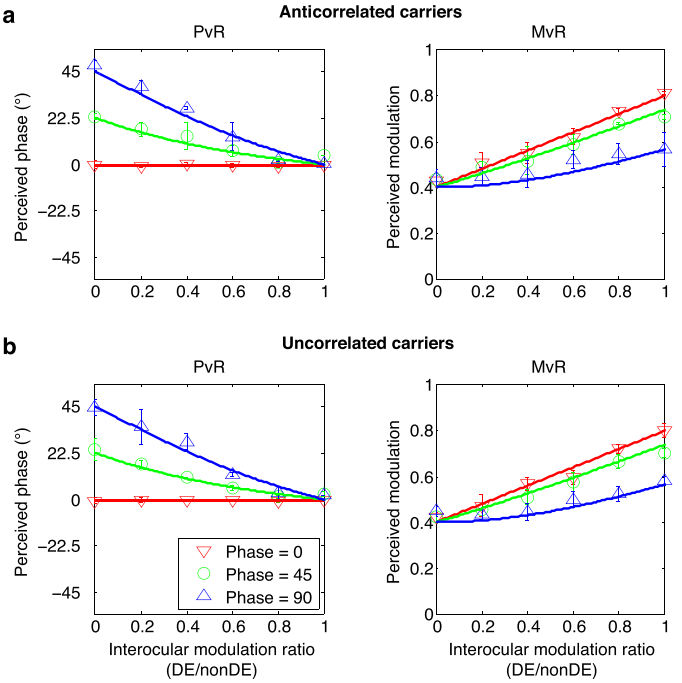

Figure 3. Experiment 2. Binocular combination of second-order stimuli with anticorrelated and uncorrelated carriers. PvRs (left) and MvRs (right) obtained when the dichoptic CM test stimulus had (a) anticorrelated carriers and (b) uncorrelated carriers. Mean of two observers. Other conventions are as in Figure 2.

tion of the second-order modulations. To test whether this is true for perceived amplitude of modulation, we ran a second experiment similar in all respects to Experiment 1 but testing uncorrelated and anticorrelated carriers. Two observers took part. Their averaged PvRs (phase) and MvRs (amplitude) for binocular CM combination with anticorrelated and uncorrelated carriers are shown in Figure 3a and b, respectively.

We ran a three-factor repeated-measures ANOVA on data for the two observers who did both experiments. Perceived phase of the cyclopean image depended significantly on the modulation ratio, $F(5, 5) = 52.06$, $p < 0.001$, and this effect was not different for the three dichoptic carrier types—no significant interaction between modulation ratio and carrier type, $F(10, 10) = 0.31$, $p = 0.96$. The PvRs for all three dichoptic carriers (Figures 2a, 3a, b) are in good agreement with the prediction of linear summation. Similarly, the MvRs for anticorrelated and uncorrelated carriers were almost identical to those for correlated carriers and again closely matched the prediction of contrast-weighted linear summation. This was supported by the ANOVA, which showed that perceived amplitude of the cyclopean image depended significantly on the interocular phase difference, $F(2, 2) = 58.44$, $p = 0.017$, and this was true for all three dichoptic carriers because the interaction of interocular phase difference and carrier type was not significant, $F(4, 4) = 0.30$, $p = 0.86$.

Discussion

In this study, we used a dichoptic phase and amplitude-matching paradigm (Huang et al., 2010) to measure the binocularly perceived phase and modulation depth of the cyclopean image formed from two monocular CM gratings, which differed in spatial phase and modulation depth. We found that the binocularly perceived modulation depended on the interocular modulation ratio and the interocular phase difference (vertical disparity). The results for both phase and amplitude can be explained simply and accurately by a weighted linear summation of second-order modulations, in which the weights depend on first-order carrier contrast but not on modulation depth.

Monocular envelope extraction

Experiments 1 and 2 together showed that binocular combination of CM signals was not affected by the extent to which the left and right eyes' carriers were correlated. This was previously found for second-order binocular phase combination (Zhou et al., 2014) and for second-order stereo processing in human (Wilcox & Hess, 1996) and cat (Tanaka & Ohzawa, 2006) and in the detection of second-order contrast modulation at threshold (M. Georgeson & Schofield, 2011). Our findings extend those earlier studies and imply that insensitivity to carrier correlation might be a general property of second-order binocular combination, suggesting an architecture in which second-order signals about contrast modulation are combined binocularly only after monocular extraction of the second-order envelopes.

Linear binocular summation (weighted averaging) for CM

Binocular combination of CM phase and modulation depth were assessed simultaneously. Our results on CM phase combination replicate those of Zhou et al. (2014), showing that binocularly perceived phase varied with interocular modulation ratio in a nearly linear fashion, consistent with linear summation of the two envelope waveforms. Here we found that matched amplitude of modulation also fitted the prediction of contrast-weighted linear summation of second-order spatial waveforms. These results suggest that binocular combination of CM phase and amplitude are processed within a common pathway.

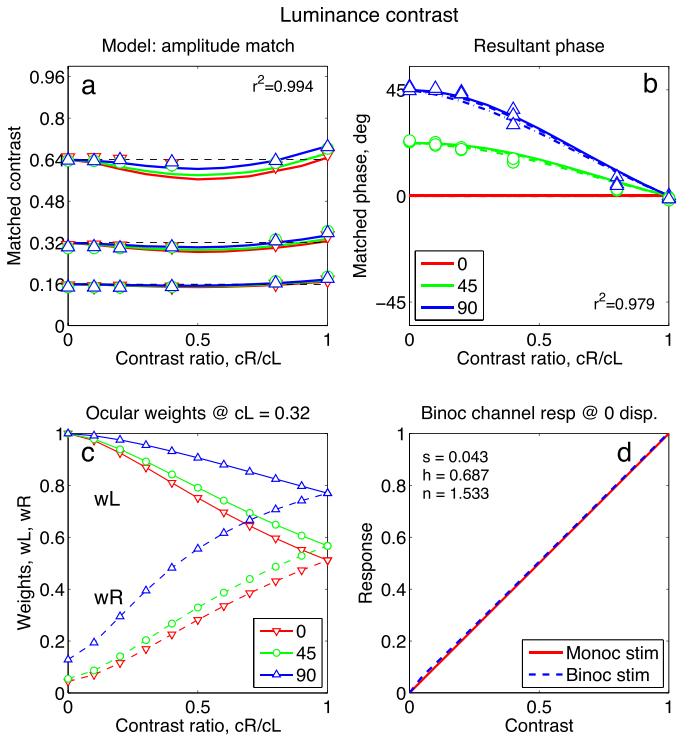

Figure 4. Contrast-weighted summation model accounts for first-order binocular combination. The extended model (see Discussion) was fitted to first-order data to derive parameters $s = 0.043$, $h = 0.687$, $n = 1.533$. (a) Predicted contrast matching was nearly invariant with relative contrast (cR/cL) and with phase disparity (0°, 45°, or 90°) as observed experimentally (symbols; data from Huang et al., 2010). From top to bottom, the three data sets are for fixed base contrast cL = 0.64, 0.32, 0.16. RMS error of the fit to contrast-matching data was small, 0.39 dB. (b) Predicted spatial phase also matched experimental data accurately; RMS error of the fit was 2.4°. Predicted and observed phase matches were unaffected by the base contrast level and so overlap almost completely in this plot. (c) Ocular weights (wL, wR) in the model were driven by relative contrast but were higher with phase disparity (blue) than without (red). This higher gain with disparity tended to offset the reduction of binocularly summed amplitude that would result from increasing disparity and so yielded almost phase-invariant perception of contrast in the disparity range from 0° to 90° (see a). (d) Ocularity invariance: Model responses to monocular contrast were linear and almost equal to the response to binocular contrast. (Note: At very low contrasts, below about 0.05, the binocular contrast response is greater than monocular because of binocular summation, but that is not easily seen on these linear axes.) See also Figure A2.

Interocular contrast-gain control sets the weights for first- and second-order binocular combination

Contrast-gain control is an important aspect of binocular processing that has been analyzed in many studies with first-order stimuli (Ding et al., 2013a, 2013b; Ding & Sperling, 2006; Hou et al., 2013; Huang et al., 2009; Huang et al., 2011; Huang et al., 2010; Meese et al., 2005, 2006; Meese & Hess, 2004; Meese & Summers, 2009). Our results show that second-order binocular combination can also be understood in terms of interocular contrast-gain control. This suggests that a similar contrast-gain control process is used for weighting the monocular visual inputs in both first- and second-order binocular combination.

Using the same paradigm as here, previous studies with first-order gratings found that perceived binocular contrast was phase-invariant up to 90° of disparity (Huang et al., 2011; Huang et al., 2010), at least for contrasts higher than about 4% (D. H. Baker et al., 2012a). This is quite different from the present second-order results in which perceived modulation depth decreased markedly with increasing phase difference for second-order stimuli over the same 0°–90° disparity range (Figure 2b). This difference in phase dependence of first- and second-order binocular combination might be taken to imply that different kinds of binocular processes are involved, but we show here that a similar interocular contrast-gain control mechanism can account for both sets of findings in relation to both phase and amplitude.

Extending the contrast-weighted summation model to previous findings on first-order combination

The simple contrast-weighted summation model described for CM in Appendix 1 is a special case of a more general model for weighting the two eyes' inputs in first-order binocular summation (Ding & Sperling, 2006, 2007). But, as Ding et al. (2013b) have pointed out, without modification that model cannot explain the repeated finding that contrast perception is largely invariant with disparity (Figure 4a). The difficulty lies in accounting for both contrast- and phase-matching data at the same time. In an effort to solve this, Huang et al. (2010) proposed separate pathways for contrast and phase, and Ding et al. (2013b) introduced a hypothetical sensorimotor vergence process ("motor/sensory fusion") that estimated the disparity and then effectively reduced or eliminated the stimulus disparity before combination. Both these proposals considerably increase the complexity of the summation model, and so we have searched for simpler alternatives. In general, if the left- and right-eye weights depend only on contrast, then the combined binocular amplitude must fall with increasing disparity (Equation A7). Thus, it seems plausible to suppose, within the gain-control framework (Ding & Sperling, 2006, 2007), that interocular suppression is arranged to compensate for

this drop in amplitude to achieve contrast constancy over disparity.

We propose that for first-order gratings interocular suppression, which sets the weights W_L, W_R for binocular combination, is influenced directly by *first-order* phase disparity (θ). Thus generalizing from Equations A1 and A2 in Appendix 1,

$$W_L = \frac{s^n + C_L^n}{s^n + C_L^n + g(\theta).C_R^n} \quad (5)$$

$$W_R = \frac{s^n + C_R^n}{s^n + C_R^n + g(\theta).C_L^n} \quad (6)$$

where $g(\theta) = max(1 - h + h.cos\theta, 0)$, and h is a constant ($0 \leq h \leq 1$) that controls the degree of disparity dependence. Note from this definition that, to prevent negative suppression, $g \geq 0$. Increasing h increases the extent to which interocular suppression depends on phase disparity: The strength factor $g(\theta)$ for interocular suppression falls from 1 to $1 - h$ as θ increases from 0 to 90°. The exponent n corresponds to the Ding and Sperling (2006) model's γ term (see Appendix 1), and shapes the way the weights vary with overall contrast level.

In general, Equations 5 and 6 show that contrast in the right eye suppresses the left eye's gain and vice versa, but when $h > 0$, the strength factor g for interocular suppression decreases when disparity is present. As a result, for first-order gratings, the weight for each eye *increases* with phase disparity (blue curves in Figure 4c) because suppression from the other eye decreases. This serves to eliminate (Figure 4a) the disparity dependence of contrast matching that must otherwise be expected.

Weighted response amplitudes $\boldsymbol{a_L}$, $\boldsymbol{a_R}$ for each eye are defined by

$$a_L = W_L r_L, \quad a_R = W_R r_R \quad (7)$$

where $\boldsymbol{r_L}$, $\boldsymbol{r_R}$ are the unweighted monocular response amplitudes, and the binocular response amplitude $\boldsymbol{r_B}$ is given from the vector sum of the two weighted responses $\boldsymbol{a_L}$, $\boldsymbol{a_R}$, taking phase difference into account (Equation A7). For simplicity and parsimony, we assume here that $\boldsymbol{r_L} = \boldsymbol{C_L}$, $\boldsymbol{r_R} = \boldsymbol{C_R}$, but in general, the monocular contrast response is likely to be more complex. If the system uses the combined response $\boldsymbol{r_B}$ to evaluate contrast, two gratings will appear to match in contrast when their $\boldsymbol{r_B}$ values are the same.

Using the simplex algorithm (*fminsearch* in *Matlab*) to minimize the value of chi-square summed over all the group-mean data (contrast and phase matches; cf. Ding et al., 2013b), we found that best-fitting values $h = 0.69$, $s = 0.04$, and $n = 1.53$ gave a very good quantitative account of the first-order results of Huang et al. (2010) for both contrast matching (Figure 4a) and phase matching (Figure 4b). When we switched off phase disparity dependence in the gain control (set $h = 0$) or refitted the model with $h = 0$, the predicted contrast matches decreased markedly at 90° disparity (not shown) and were a poor fit to the data, but phase matches were almost unchanged. Thus contrast matches were sensitive to the influence of first-order phase disparity on gain control (Equations 5 and 6), but phase matches were not.

In short, a simple modification to the Ding and Sperling (2006) gain-control model (the introduction of a disparity term; Equations 5 and 6) can explain why first-order contrast matching is almost constant across disparity while preserving the correct phase-matching behavior. We think this is a plausible and perhaps simpler alternative to the introduction of more elaborate processes, such as motor-sensory fusion (Ding et al., 2013b).

First-order combination at large disparities: Rivalry suppression via the MAX operator

The model of first-order binocular summation described so far works well for disparities up to at least 90° (Figure 4). But in common with several other models, it suffers complete cancellation between left and right inputs when disparity approaches 180°, and human vision does not. Perceived contrast is reduced only a little or not at all (0–4 dB) at these large disparities (see Figure 5a; data from D. H. Baker et al., 2012a), and it seems likely that binocular rivalry suppression in some way takes over from binocular summation (Ding et al., 2013b; Ding & Sperling, 2006; M. A. Georgeson & Wallis, 2014). We show here that one simple extra assumption—representing a form of rivalry suppression—extends the first-order model to account well for perceived contrast across all disparities.

Using preset parameters (from Figure 4), we compared the first-order model predictions with contrast-matching data of D. H. Baker et al. (2012a) as a function of interocular phase disparity over the full range from 0° to ±180° at several levels of contrast (2%–32%). The pattern of their mean results (Figure 5a) was fairly well reproduced by the model up to disparities of about ±120° (Figure 5b), but as expected, the predicted contrast fell profoundly at larger disparities. Ding et al.'s (2013b) Ding-Sperling-Klein-Levi (DSKL) model behaved in a similar way (their figure 17A). Removing phase-dependent interocular suppression (setting $h = 0$ in our model; hence $g = 1$), only made matters worse: The predictions collapsed onto linear summation, making contrast matching vary sharply with the cosine of disparity (Equation A7) at all contrast levels (Figure 5c), quite unlike the data. Something else is needed.

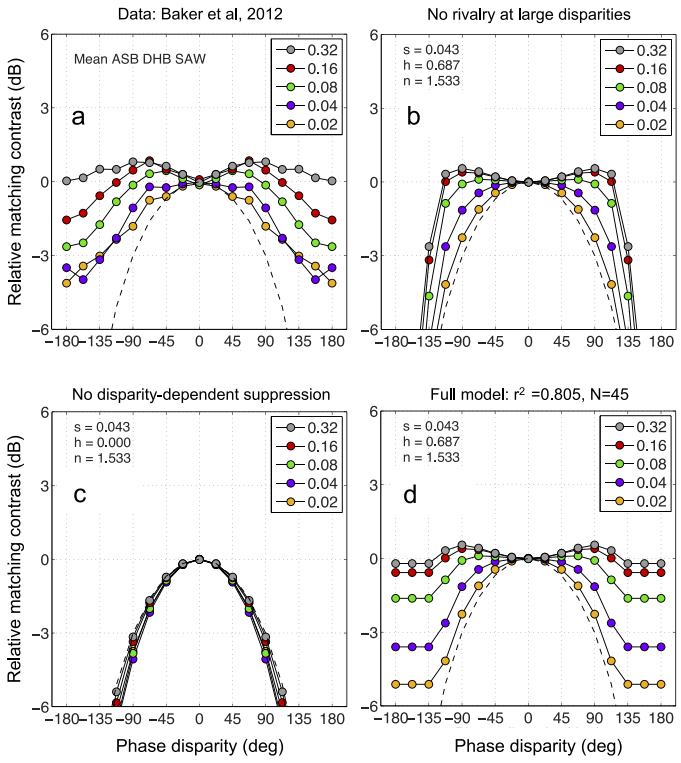

Figure 5. Monocular contrast response determines contrast-matching at large phase disparities. (a) Data on contrast matching for horizontal 1 c/° gratings as a function of phase disparity from D. H. Baker et al. (2012a). Test gratings had one of five contrasts with the same contrast in both eyes. A similar comparison grating with zero disparity was matched in contrast to each test grating. Y-axis shows relative matching contrast in dB [defined as $20 \times \log_{10}(C_{match}/C_{test})$]. Points at 0 dB are a veridical match; those below 0 dB represent lower perceived test contrasts. (b) Our model of binocular summation (including phase-dependent interocular suppression) gave a poor match to the data at large disparities ($r^2 < 0$). Model parameters were the same as Figure 4. (c) Removing phase-dependent suppression (setting $h = 0$) made the model fit much worse ($r^2 < 0$). (d) Full model included phase-dependent interocular suppression and a simple form of WTA rivalry suppression (Equation 8). The similarity to the experimental data is good ($r^2 = 0.805$).

We propose that a nonlinear competitive interaction at the binocular combination stage represents a simple constraint: *If binocular summation will reduce response amplitude, then don't do it*. This idea is easily implemented by supposing that contrast judgments are based on a modified response $\boldsymbol{r_B}'$:

$$r'_B = max(r_L, r_B, r_R) \quad (8)$$

In other words, if the combined response $\boldsymbol{r_B}$ is too low, it is replaced by the larger of the two unweighted monocular responses. This substitution, using the *MAX* operator, can be seen as a form of WTA rivalry, and it occurs mostly at large disparities at which $\boldsymbol{r_B}$ falls dramatically. With no free parameters, it strikingly improved contrast-matching predictions for large disparities as shown in Figure 5d (goodness of fit, $r^2 = 0.805$; RMS error = 0.62 dB), compared with the poor fit of Figure 5b and c. Thus, two different forms of suppression are likely to play a role in binocular combination: (a) the graded, disparity-sensitive interocular suppression that sets the balance (the weights, Equations 5 and 6) for left- and right-eye inputs to the summation process and (b) a more profound, perhaps later stage, WTA suppression that effectively compares the two monocular inputs with the binocularly combined response and picks the largest of the three while vetoing the other two. Selection of the monocular response at large disparities serves to prevent the binocular cancellation of antiphase signals from feeding through to perception. (See Appendix 2 for further analysis and discussion.)

No Fechner paradox in first-order contrast perception

Despite the generally good fit seen in Figure 4, our model (like several others) predicts the occurrence of a *Fechner paradox* that is not seen in the first-order data. The (predicted) paradox is that when one eye's contrast is fixed and the other eye's contrast is increased from zero, then the combined binocular amplitude goes down before going up. The paradoxical reduction is predicted when interocular suppression outweighs the increase in signal strength given by adding contrast to the second eye. Although the Fechner paradox is reliably found in dichoptic brightness matching against a dark background (Anstis & Ho, 1998; D. H. Baker, Wallis, Georgeson, & Meese, 2012b; Engel, 1970; Levelt, 1965), it seems clear from several recent detailed studies that it does not normally occur for luminance contrast against a mid-gray background—neither for increments, decrements, step edges, nor gratings (D. H. Baker et al., 2012b; Ding et al., 2013b). Instead, perceived (matched) contrast is usually close to the higher of the two eyes' contrasts: WTA (although in amblyopic vision, a substantial, asymmetric Fechner paradox has been observed; Ding et al., 2013a). The deviation between model and data (Figure 4a) is not large, but it is systematic and deserves attention.

Ding et al. (2013b) considered the possibility that this WTA behavior arises directly from a WTA competition between the two eyes' inputs. But they rejected this idea because perceived phase does not switch abruptly from one eye's phase to the other when

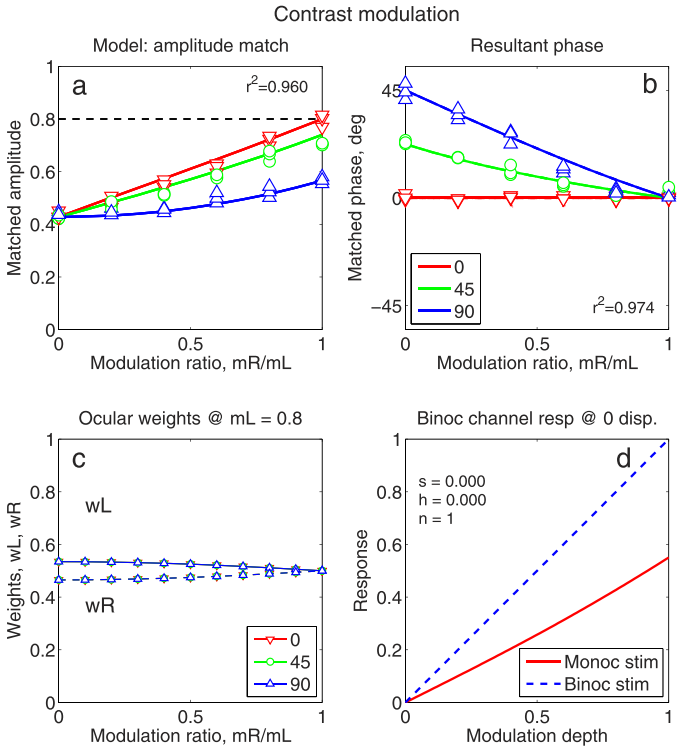

Figure 6. Contrast-weighted summation of modulation depths accounts for second-order binocular combination. Data from Experiments 1 and 2 plotted together, compared with model predictions for CM tasks in the same format as Figure 4. Here $s = 0$, $h = 0$, $n = 1$, and RMS contrast was used in Equations 5 and 6 (or A1 and A2) to derive the ocular weights; see last section of Appendix 1 for rationale and evidence supporting this choice. (a) Predicted matching of CM amplitude increased with modulation ratio (mR/mL) but decreased with increasing phase disparity (0°, 45°, or 90°) in close agreement with the data (RMS error = 0.34 dB). (b) Predicted spatial phase also matched experimental data accurately (RMS error = 2.2°). Repeated symbols in (a) and (b) represent the three levels of carrier correlation and overlap almost completely in these plots. (c) The model's ocular weights (wL, wR) for the dichoptic test images were again driven by relative (carrier) contrast, but differed a little from 0.5 each because images with higher modulation have slightly higher RMS contrast and therefore higher weights than those with lower modulation. (d) With the carrier present in both eyes, model response to monocular modulation was almost linear and half the response to binocular modulation.

one eye's contrast exceeds the other. Instead, they proposed that the apparent WTA behavior in contrast perception arises from the balance between suppressive gain control and a proposed gain enhancement mechanism. But we argue, in the previous section, that it may again be fruitful to consider the contribution made by monocular mechanisms in dichoptic experiments. We show (in Appendix 2) that with the addition of *noise*, a WTA process (the *MAX* operator, Equation 8) can give a good account not only of perceived contrast (Figures 5d, A2a, and A4), but also of perceived phase for dichoptic first-order gratings (Figure A2b). Without noise, the predicted phase switched far too abruptly with increasing contrast ratio, just as Ding et al. (2013b) had supposed (not shown). This analysis strengthens the case for a fairly direct role of monocular signals in binocular vision. In short: *When the summed binocular response is lower than both its monocular inputs, use the larger monocular response instead.*

On the other hand, binocular summation alone gave an almost perfect account of the second-order results (Figure 6) with no free parameters, and so it seems unlikely that monocular mechanisms make a direct contribution to perception of CM aside from their input to the binocular sum.

Summary: Different behaviors but a similar underlying process

In summary, we found that for CM gratings, the perceived phase and modulation depth varied with both phase disparity and interocular ratio of modulation depths (Figure 6a, b). The pattern of results was quite different from the way variations in disparity and contrast ratio influenced perception of first-order luminance contrast and phase (Figure 4a, b) in exactly analogous experiments. This striking difference is now understandable, and a similar principle of contrast-weighted linear summation correctly predicts results for first- and second-order signal combination. Variations in relative contrast alter the weights assigned to each eye, but variations in modulation depth do not (or do so only slightly if RMS contrast is what drives the gain control; Figure 6b). Hence the left and right first-order weights fall and rise respectively as the right-eye contrast increases (Figure 4c), but the second-order weights remain constant or nearly constant for each eye (Figure 6c) as modulation ratio varies. This difference in the adjustment of weights arises because the *same* rule (contrast weighting) is being applied to different circumstances; relative contrast varies, or it doesn't.

The same formula (Equations 5 and 6) can be used to compute the first-order and second-order weights. The weights depend on first-order phase disparity (if present), but they should *not* depend on CM phase disparity because the weights are driven by (first-order) carrier contrast, and this is unaffected by the absolute or relative phase of modulation. For CM, the required lack of phase dependence is achieved by

setting $h = 0$ in the weights for CM stimuli. This is a logical requirement that represents the change in type of stimulus rather than a change in the model itself. We also find from model fitting that for CM the constant s has to be at or close to 0. These two values ($h = 0$, $s = 0$) lead to the simple linear averaging behavior for dichoptic CM seen in Figures 2, 3, and 6.

Ocularity invariance

Contrast-weighted summation ensures constancy or "ocularity invariance" for first-order contrast perception: Responses to monocular and binocular contrasts are nearly the same (Figures 4d and A2d). One might think that this invariance fails for CM because the response to binocular modulation is twice that for monocular modulation (Figure 6d). But that is true only when the same carrier is present in both eyes: The two equal carriers force the weights to be about 0.5 each, thus averaging the two eyes' modulations and halving the response to monocular modulation. In natural viewing, however, closing one eye removes all contrast from that eye, the weight for the other eye goes to 1, and its response to modulation will therefore be the same as for two eyes with weights of 0.5 each. Hence in natural viewing (one eye closed vs. both eyes open), ocularity invariance holds for both first-order and second-order modulation in agreement with everyday experience.

Keywords: binocular combination, binocular disparity, contrast modulation, first-order, second-order, contrast-gain control, interocular suppression, rivalry suppression

Acknowledgments

This work was supported by CIHR grant (#53346) and NSERC grant (#46528-11) to RFH, a RI-MUHC postdoctoral award to JZ, and BBSRC grant (BB/H00159/X1) to MAG. JZ, MAG, and RFH designed the research; JZ performed the experiments; JZ, MAG, and RFH analyzed data and models and wrote the paper. The authors declare no conflict of interest.

Commercial relationships: none.
Corresponding author: Robert F. Hess.
Email: robert.hess@mcgill.ca.
Address: McGill Vision Research, Department of Ophthalmology, McGill University, Montreal, QC, Canada.

References

Anstis, S., & Ho, A. (1998). Nonlinear combination of luminance excursions during flicker, simultaneous contrast, afterimages and binocular fusion. *Vision Research, 38*(4), 523–539.

Baker, C. L., Jr. (1999). Central neural mechanisms for detecting second-order motion. *Current Opinion in Neurobiology, 9*(4), 461–466.

Baker, D. H., Meese, T. S., & Georgeson, M. A. (2007). Binocular interaction: Contrast matching and contrast discrimination are predicted by the same model. *Spatial Vision, 20*(5), 397–413.

Baker, D. H., Wallis, S. A., Georgeson, M. A., & Meese, T. S. (2012a). The effect of interocular phase difference on perceived contrast. *PLoS One, 7*(4), e34696.

Baker, D. H., Wallis, S. A., Georgeson, M. A., & Meese, T. S. (2012b). Nonlinearities in the binocular combination of luminance and contrast. *Vision Research, 56,* 1–9.

Black, J. M., Thompson, B., Maehara, G., & Hess, R. F. (2011). A compact clinical instrument for quantifying suppression. *Optometry and Vision Science, 88*(2), 334–343.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433–436.

Cavanagh, P., & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision, 4,* 103–129.

Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A, 5*(11), 1986–2007.

Chubb, C., & Sperling, G. (1989). Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proceedings of the National Academy of Sciences, USA, 86*(8), 2985–2989.

Cooper, E. A., Jiang, H., Vildavski, V., Farrell, J. E., & Norcia, A. M. (2013). Assessment of OLED displays for vision research. *Journal of Vision, 13*(12):16, 1–13, http://www.journalofvision.org/content/13/12/16, doi:10.1167/13.12.16. [PubMed] [Article]

Dakin, S. C., & Mareschal, I. (2000). Sensitivity to contrast modulation depends on carrier spatial frequency and orientation. *Vision Research, 40*(3), 311–329.

Ding, J., Klein, S. A., & Levi, D. M. (2013a). Binocular combination in abnormal binocular vision. *Journal of Vision, 13*(2):14, 1–31, http://www.

journalofvision.org/content/13/2/14, doi:10.1167/13.2.14. [PubMed] [Article]

Ding, J., Klein, S. A., & Levi, D. M. (2013b). Binocular combination of phase and contrast explained by a gain-control and gain-enhancement model. *Journal of Vision, 13*(2):13, 1–37, http://www.journalofvision.org/content/13/2/13, doi:10.1167/13.2.13. [PubMed] [Article]

Ding, J., & Sperling, G. (2006). A gain-control theory of binocular combination. *Proceedings of the National Academy of Sciences, USA, 103*(4), 1141–1146.

Ding, J., & Sperling, G. (2007). Binocular combination: Measurements and a model. In L. Harris & M. Jenkin (Eds.), *Computational vision in neural and machine systems* (pp. 257–305). Cambridge, UK: Cambridge University Press.

Dosher, B. A., & Lu, Z.-L. (2006). Level and mechanisms of perceptual learning: Learning first-order luminance and second-order texture objects. *Vision Research, 46*(12), 1996–2007.

Engel, G. (1970). Tests of a model of binocular brightness. *Canadian Journal of Psychology/Revue Canadienne de Psychologie, 24*(5), 335.

Georgeson, M., & Schofield, A. (2011). Binocular functional architecture for detection of luminance- and contrast-modulated gratings. *Journal of Vision, 11*(11):305, http://www.journalofvision.org/content/11/11/305, doi:10.1167/11.11.305. [Abstract]

Georgeson, M. A., & Wallis, S. A. (2014). Binocular fusion, suppression and diplopia for blurred edges. *Ophthalmic and Physiological Optics, 34*(2), 163–185.

Hess, R. F., Hutchinson, C., Ledgeway, T., & Mansouri, B. (2007). Binocular influences on global motion processing in the human visual system. *Vision Research, 47*(12), 1682–1692.

Hou, F., Huang, C. B., Liang, J., Zhou, Y., & Lu, Z. L. (2013). Contrast gain-control in stereo depth and cyclopean contrast perception. *Journal of Vision, 13*(8):3, 1–19, http://www.journalofvision.org/content/13/8/3, doi:10.1167/13.8.3. [PubMed] [Article]

Huang, C. B., Zhou, J., Lu, Z. L., Feng, L., & Zhou, Y. (2009). Binocular combination in anisometropic amblyopia. *Journal of Vision, 9*(3):17, 1–16, http://www.journalofvision.org/content/9/3/17, doi:10.1167/9.3.17. [PubMed] [Article]

Huang, C. B., Zhou, J., Lu, Z. L., & Zhou, Y. (2011). Deficient binocular combination reveals mechanisms of anisometropic amblyopia: Signal attenuation and interocular inhibition. *Journal of Vision, 11*(6):4, 1–17, http://www.journalofvision.org/content/11/6/4, doi:10.1167/11.6.4. [PubMed] [Article]

Huang, C. B., Zhou, J., Zhou, Y., & Lu, Z. L. (2010). Contrast and phase combination in binocular vision. *PLoS One, 5*(12), e15075.

Klein, S. A., Hu, Q. J., & Carney, T. (1996). The adjacent pixel nonlinearity: Problems and solutions. *Vision Research, 36*(19), 3167–3181.

Larsson, J., Landy, M. S., & Heeger, D. J. (2006). Orientation-selective adaptation to first- and second-order patterns in human visual cortex. *Journal of Neurophysiology, 95*(2), 862–881.

Legge, G. E. (1984a). Binocular contrast summation: I. Detection and discrimination. *Vision Research, 24*(4), 373–383.

Legge, G. E. (1984b). Binocular contrast summation: II. Quadratic summation. *Vision Research, 24*(4), 385–394.

Legge, G. E., & Gu, Y. C. (1989). Stereopsis and contrast. *Vision Research, 29*(8), 989–1004.

Levelt, W. (1965). Binocular brightness averaging and contour information. *British Journal of Psychology, 56*(1), 1–13.

Mansouri, B., Thompson, B., & Hess, R. F. (2008). Measurement of suprathreshold binocular interactions in amblyopia. *Vision Research, 48*(28), 2775–2784.

Meese, T. S., Georgeson, M. A., & Baker, D. H. (2005). Interocular masking and summation indicate two stages of divisive contrast gain control. *Perception, 34,* 42–43.

Meese, T. S., Georgeson, M. A., & Baker, D. H. (2006). Binocular contrast vision at and above threshold. *Journal of Vision, 6*(11):7, 1224–1243, http://www.journalofvision.org/content/6/11/7, doi:10.1167/6.11.7. [PubMed] [Article]

Meese, T. S., & Hess, R. F. (2004). Low spatial frequencies are suppressively masked across spatial scale, orientation, field position, and eye of origin. *Journal of Vision, 4*(10):2, 843–859, http://www.journalofvision.org/content/4/10/2, doi:10.1167/4.10.2. [PubMed] [Article]

Meese, T. S., & Summers, R. J. (2009). Neuronal convergence in early contrast vision: Binocular summation is followed by response nonlinearity and area summation. *Journal of Vision, 9*(4):7, 1–16, http://www.journalofvision.org/content/9/4/7, doi:10.1167/9.4.7. [PubMed] [Article]

Pelli, D. G. (1997). The VideoToolbox software for

visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442.

Regan, D. (2000). *Human perception of objects: Early visual processing of spatial form defined by luminance, colour, texture, motion, and binocular disparity*. Sunderland, MA: Sinauer Associates.

Reynaud, A., Zhou, J., & Hess, R. F. (2013). Stereopsis and mean luminance. *Journal of Vision, 13*(11):1, 1–11, http://www.journalofvision.org/content/13/11/1, doi:10.1167/13.11.1. [PubMed] [Article]

Schofield, A. J., & Georgeson, M. A. (1999). Sensitivity to modulations of luminance and contrast in visual white noise: Separate mechanisms with similar behaviour. *Vision Research, 39*(16), 2697–2716.

Schofield, A. J., & Georgeson, M. A. (2003). Sensitivity to contrast modulation: The spatial frequency dependence of second-order vision. *Vision Research, 43*(3), 243–259.

Sutter, A., & Graham, N. (1995). Investigating simple and complex mechanisms in texture segregation using the speed-accuracy tradeoff method. *Vision Research, 35*(20), 2825–2843.

Tanaka, H., & Ohzawa, I. (2006). Neural basis for stereopsis from second-order contrast cues. *The Journal of Neuroscience, 26*(16), 4370–4382.

Werkhoven, P., Sperling, G., & Chubb, C. (1993). The dimensionality of texture-defined motion: A single channel theory. *Vision Research, 33*(4), 463–485.

Wilcox, L. M., & Hess, R. F. (1996). Is the site of non-linear filtering in stereopsis before or after binocular combination? *Vision Research, 36*(3), 391–399.

Wilson, H. R. (1999). Non-Fourier cortical processes in texture, form, and motion perception. In P. Ulinski, E. Jones, & A. Peters (Eds.), *Cerebral cortex, vol. 13, models of cortical circuits* (pp. 445–477). New York: Kluwer Academic/Plenum.

Zhou, J., Huang, P.-C., & Hess, R. F. (2013). Interocular suppression in amblyopia for global orientation processing. *Journal of Vision, 13*(5):19, 1–14, http://www.journalofvision.org/content/13.5.19, doi:10.1167/13.5.19. [PubMed] [Article]

Zhou, J., Jia, W., Huang, C.-B., & Hess, R. F. (2013). The effect of unilateral mean luminance on binocular combination in normal and amblyopic vision. *Scientific Reports, 3*(2012), 1–7.

Zhou, J., Liu, R., Zhou, Y., & Hess, R. F. (2014). Binocular combination of second-order stimuli. *PLoS One, 9*(1), e84632.

Appendix 1

A simple model for CM: Contrast-weighted linear summation of modulations across the eyes

Ocular weights

Let the modulation depths for the left and right eyes be M_L, M_R with carrier contrasts C_L, C_R. We define the carrier-dependent weights W_L, W_R assigned to the left and right eyes as

$$W_L = \frac{s + C_L}{s + C_L + C_R} \quad \text{(A1)}$$

$$W_R = \frac{s + C_R}{s + C_L + C_R} \quad \text{(A2)}$$

where s is typically a small-valued constant that prevents division by 0 and ensures that the weight for one eye is 1 when the other eye's contrast is 0. Here, because C_L, C_R are never both 0 at the same time, we can simplify and let $s = 0$. With this simplification, the model has no free parameters and so makes direct predictions about the observer's matching of CM phase and amplitude (shown in Figures 2 and 3).

For the dichoptic CM test gratings, because both carriers are present, $W_L = W_R = 0.5$, implying simple averaging. But for the monocular probe (assumed here to be in the left eye), $C_R = 0$, so $W_L = 1$, and $W_R = 0$, implying no binocular combination of any kind.

Envelope recovery: Monocular FRF processes

Most models of second-order vision have proposed some version of the standard FRF (filter–rectify–filter) concept to explain how the modulation signal (contrast envelope) is recovered from the modulated image. We take such a process as given, and so our starting point is the envelope signal that is recovered. We further assume that envelope responses are recovered by separate FRF processes for the left and right eyes before binocular combination. (Note: another possibility is that the two eyes first combine local contrast values, and then the envelope signal is recovered. We reject this idea because it would predict quite different results when the two eyes have correlated, anticorrelated, and uncorrelated carriers; such differences were not seen in our study.) Then, assuming that monocular responses to the contrast envelope are linear and directly proportional to modulation depth (a simple but strong assumption), the weighted monocular response profiles (R_L, R_R) to modulation at spatial frequency f, with phases ϕ_L, ϕ_R, as a function of position (y) are given by

$$R_L(y) = W_L M_L \cos(2\pi f y + \phi_L) \quad \text{(A3)}$$

$$R_R(y) = W_R M_R \cos(2\pi f y + \phi_R) \quad \text{(A4)}$$

Binocular summation

Weighted monocular responses are summed to give the binocular response $R_B(y) = R_L(y) + R_R(y)$. Thus Equations A3 and A4 sum linearly; each can be treated as a vector with amplitudes $W_L M_L$, $W_R M_R$ and phases $\phi_L = \theta/2$, $\phi_R = -\theta/2$. The resultant phase θ' for interocular phase difference θ ($=\phi_L - \phi_R$) and modulation ratio $\boldsymbol{\delta}$ ($=M_R/M_L$) is

$$\theta' = tan^{-1}\left[\frac{W_L - \delta.W_R}{W_L + \delta.W_R} tan\left(\frac{\theta}{2}\right)\right] \quad \text{(A5)}$$

For our dichoptic CM grating, the weights are equal (see above), and so this reduces to

$$\theta' = tan^{-1}\left[\frac{1 - \delta}{1 + \delta} tan\left(\frac{\theta}{2}\right)\right] \quad \text{(A6)}$$

The probe amplitude (M') will appear to match a given dichoptic test when both deliver the same binocular response amplitude. In general, if two vectors have amplitudes a_L, a_R and angular separation $\boldsymbol{\theta}$, the amplitude a_B of their vector sum is

$$a_B = \sqrt{a_L^2 + 2a_L a_R cos\theta + a_R^2} \quad \text{(A7)}$$

For our CM test stimulus with base modulation M_0 in the left eye, $a_L = W_L M_0$, $a_R = \delta.W_R M_0$, and from Equations A1 and A2, both weights W_L, W_R are 0.5. Inserting these into Equation A7, we get

$$a_B(test) = 0.5 M_0 \sqrt{\delta^2 + 1 + 2\delta cos\theta} \quad \text{(A8)}$$

For the monocular probe (M'), the weights are 1 and 0, and so the outcome is simply

$$a_B(probe) = M' \quad \text{(A9)}$$

Thus when a match is achieved, this model predicts

$$M' = 0.5 M_0 \sqrt{\delta^2 + 1 + 2\delta cos(\theta)} \quad \text{(A10)}$$

These are the amplitude-matching curves plotted in Figures 2b and 3. We emphasize, however, that this model is *not* the same as simply assuming that the monocular envelopes are averaged. The difference lies in the probe response. With averaging, the response to the probe would be $M'/2$ (average of M' and 0) instead of M'. All predicted amplitude matches would double and completely fail to fit the data. The key assumption in our model is that the weights are driven by luminance contrast, not by modulation depth, and with this contrast weighting, the same combination rule gives different behavior for the binocular test and the monocular probe. For the test grating, when the carrier contrast is present and equal in both eyes, we get binocular averaging, but for the probe, when one eye has no carrier, we get WTA (weights of 1 and 0). The surprising consequence is that a monocular ($\delta = 0$) test amplitude of 0.8 is matched by a monocular probe amplitude of 0.4. These two rules (averaging and WTA) emerge from the same weighting scheme, which can be seen as a mechanism for achieving "ocularity invariance"—perception of form and contrast remain the same with one eye and two eyes (D. H. Baker, Meese, & Georgeson, 2007; Ding & Sperling, 2006).

Relation to the Ding & Sperling model

Our calculation of weights for the two eyes (Equations A1 and A2) is equivalent to a special case of Ding and Sperling's (2006) contrast-gain control theory for first-order combination. They assumed that each eye (a) exerts gain control on the other eye's signal in proportion to the contrast energy of its own input and (b) additionally exerts gain control on the other eye's gain control. They derived the weights for the left and right eyes as

$$W_L = \frac{1 + \rho C_L^\gamma}{1 + \rho C_L^\gamma + \rho C_R^\gamma} \quad \text{(A11)}$$

and

$$W_R = \frac{1 + \rho C_R^\gamma}{1 + \rho C_L^\gamma + \rho C_R^\gamma} \quad \text{(A12)}$$

where ρ is the gain-control efficiency of the signal sine-wave grating, and γ is the exponent of the nonlinear transducer. Equivalence with Equations A1 and A2 is exact when $s = 1/\rho$ and $\gamma = 1$. Note that the meaning of C_L (or C_R) remains unchanged; it is the luminance contrast of the image, determined by the carrier for second-order and the grating itself for first-order.

Suppose the weights for second-order combination are driven by first-order contrast in the same way as for first-order combination. For our dichoptic second-order test gratings, because both carriers were present and $C_L = C_R$, the Ding and Sperling Equations A11 and A12 would always give equal weights, and $W_L = W_R = 0.5$, if $\rho \gg 1$, again implying simple averaging. Note that when $C_L = C_R$ this equality of weights holds true for any value of the nonlinear exponent (γ). And similarly for a monocular (left eye) probe $C_R = 0$, so $W_L = 1$ and $W_R = 0$, again implying no binocular combination for the probe. This also is true for any γ and implies

that the present experiments on CM are insensitive to the value of γ. If the ratio C_L:C_R is systematically varied, then the ocular weights vary—favoring the eye that has the higher contrast—and the value of γ can be estimated from phase-matching data. This has been done for first-order signals but not yet for second-order. For first-order summation, the geometric mean estimate of γ was close to 1 (1.16 ± 0.18, pooled over two previous studies with a total of seven observers; Ding & Sperling, 2007; Huang et al., 2010). Our formulation (Equations A1 and A2) is equivalent to assuming $\gamma = 1$ for CM.

Is the contrast weighting for CM driven by mean contrast or RMS contrast?

All the analysis and discussion in Appendix 1 so far has tacitly presumed that the carrier contrast values C_L, C_R needed to compute the weights (Equations A1 and A2) are the mean pixel contrasts or Michelson contrasts (which are the same for binary noise). This is simple, and it makes the contrast values and the weights independent of modulation depth, and so the weights are equal at all modulation ratios and equal to 0.5 when $s = 0$. But is mean contrast the most relevant measure? The sharp-eyed reader may notice that in Figures 2b and 3 the observed amplitude matches at low modulation ratios (0, 0.2) lie slightly but consistently *above* the model's asymptote of 0.4 that is predicted by equal weighting. It turns out that this discrepancy disappears, and the model fit improves, if RMS contrast is used in place of mean contrast. For binary noise (as used here) Schofield and Georgeson (1999) showed (their equation 8) that $C_{RMS} = c\sqrt{1 + m^2/2}$, where C is the carrier pixel contrast and m the contrast modulation depth. Thus RMS contrast increases with modulation. In our experiments, at a modulation ratio of zero, the modulated image in one eye has a higher RMS contrast than its unmodulated partner. The result of using RMS contrast to compute the weights is shown in Figure 6. Model predictions, still with no free parameters, were compared with all the data points from Experiments 1 and 2. The fit was excellent: For the amplitude matches $r^2 = 0.960$ (RMS error = 0.335 dB), compared with $r^2 = 0.943$ (RMS error = 0.461 dB) when mean contrast was used for the weights (as in Figures 2 and 3). The improvement clearly lay at the low modulation ratios, and Figure 6c reveals why: The weight is a little higher for the eye with more modulation. This pushes the binocular sum a bit higher than it would be if the weights were equal, but the difference in weights shrinks as the other eye's modulation increases toward equality. We conclude that, although the improvement in model fit is small, it looks convincing and has a clear rationale: Ocular weights for CM are driven by RMS contrast not mean or Michelson contrast. (Note: This analysis has no implications for first-order combination because, for luminance gratings, Michelson contrast and RMS contrast are directly proportional to each other and therefore functionally interchangeable.)

Appendix 2

A role for monocular responses in binocular perception of first-order contrast and phase

We first show how the use monocular and binocular signals, selected in a WTA fashion via the MAX operator (Equation 8), leads to good predictions for *first-order* contrast-matching across all phase disparities seen in Figure 5d. Then we show that adding independent noise to each of the three signals before the MAX operator can yield good predictions for perception of binocular phase while, at the same time, improving the predictions for contrast-matching by reducing the Fechner paradox.

WTA 1: The MAX operator

Recall our proposal (Equation 8) that perceived contrast is determined by whichever of the three response amplitudes (two monocular, one binocular) is the largest:

$$r'_B = max(r_L, r_B, r_R) \quad (8)$$

Note also that the binocular response amplitude r_B varies with disparity for two reasons: (a) the vector sum varies with phase difference (Equation A7), and (b) the component amplitudes in the vector sum also vary with phase difference because the ocular weights vary with disparity (when $h > 0$; Equations 5 and 6). Red curves in Figure A1 illustrate how the resultant amplitude r_B varies with phase disparity, falling steeply to zero as disparity increases from ±120° to ±180°. On the other hand, the monocular responses do not depend on disparity and plot as circles (blue) in these polar diagrams.

The response of the max operator is the outer envelope of the red and blue curves (plotted behind in gray). The predicted contrast matches (gray symbols) lie very close to this outer envelope. (They differ slightly at low contrasts because the matching of test and comparison gratings is also taken into account. The predicted match is the contrast of the comparison grating—with 0 disparity—that produces the same

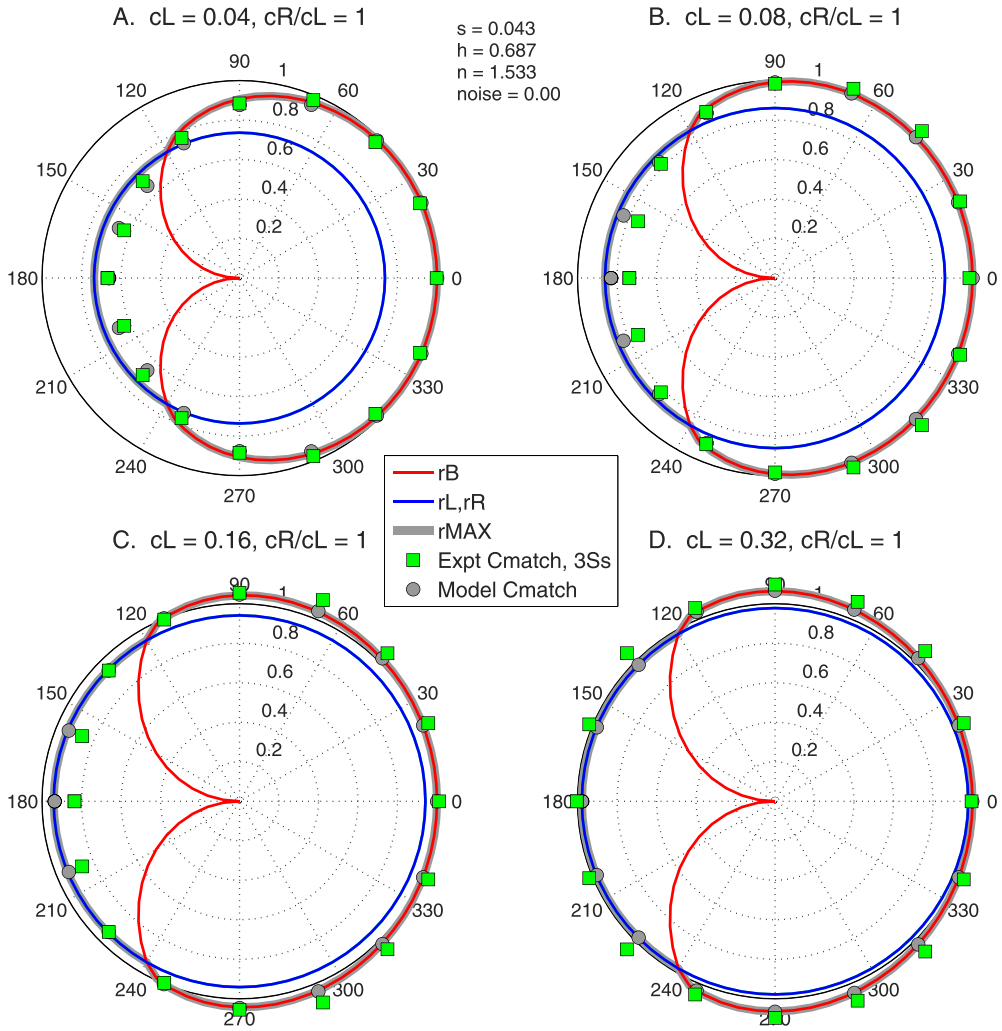

Figure A1. Polar plots illustrate how the first-order model with MAX operator (Equation 8) explains contrast matching at all phase disparities and contrast levels (Figure 5). Polar angle represents phase disparity (0°–360°); radial distance represents either relative response level (colored curves) or relative contrast match (symbols), normalized to the point at 0 disparity. Red curves show that the model's binocular response (r_B) is greater than monocular responses (r_L, r_R, blue) for disparities up to ±120° but falls well below them at larger disparities. Green squares: experimental contrast matches from Figure 5a; gray circles: corresponding model matches from Figure 5d. Note how the experimental data closely follow the outer envelope of the response curves (rMAX, gray curve). Model parameters were as Figure 5 with no noise in the *MAX* operator. Predicted matches (gray symbols) agree well with experimental data.

response r_B' as the test grating does.) Importantly, the experimental contrast matches (green squares; mean of 3 Ss; D. H. Baker et al., 2012a) lie close to the predicted ones (gray symbols) at all disparities and all contrast levels. This figure makes it clear how, on this model, contrast-matching up to ±120° disparity is explained by the weighted binocular sum, r_B (red curve), and at larger disparities, it is the monocular response (blue curve) that matters. More subtly, the relationship between monocular and binocular response curves varies with contrast level because the binocular weights vary with contrast (Equations 5 and 6). The monocular response is relatively weaker at low contrast, and this accounts for the interaction between disparity and contrast level seen most clearly in Figure 5a and d, a stronger effect of phase disparity at low contrasts.

WTA 2: The noisy MAX operator

Ding et al. (2013b) noted that the difficult challenge for models of binocular combination is to account for contrast matching and phase matching at the same time. Some models that did well on one task fared badly on the other or vice versa. Indeed, when we reapplied the above model, with MAX operator, to the analysis of Figure 4, the contrast-matching predictions improved ($r^2 = 0.996$), and the Fechner paradox was reduced, but the phase matches were poor ($r^2 = 0.64$,

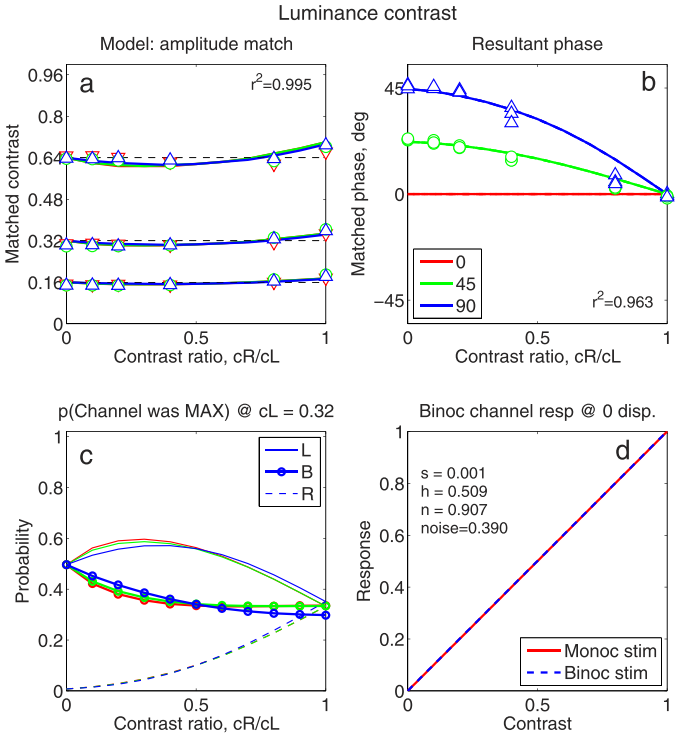

Figure A2. Similar to Figure 4, but illustrating how, with a noisy MAX operator (Equation 8), the first-order model can (a) explain the near-absence of Fechner paradox in dichoptic contrast matching and (b) give a good account of binocular phase matching. Color in (a, b, c) denotes phase disparity as shown in (b). (a) Data (symbols) as Figure 4a; curves show best fitting model (see Appendix 2) with four free parameters listed in (d). (c) Contributions made to perception made by left or right eye contrast (L, R; thin lines) and by the binocularly combined response (B; thick lines). "Contribution" is defined as the probability that a given random variable (r_L, r_B, or r_R) gave the MAX response in a given condition.

RMS error = 10.1° of phase; not shown). As expected, they were not graded with contrast ratio (unlike Figure 4b), but were constant, determined by the higher contrast input until the contrast ratio was close to 1.0. However, if (as seems likely) the three inputs to the MAX operator are noisy over time or from trial to trial, then all three inputs should contribute somewhat to any given condition, and perceived phase might show a smoother transition as contrast ratio increased from 0 to 1 (like the data in Figure 4b). This idea was tested with the same first-order model as before (*WTA 1*) except that the effects of adding independent Gaussian noise to the signals before the MAX operator were also included. Rather than running a Monte Carlo simulation (too slow), we wrote a *Matlab* function that computed the statistics of the MAX operator (mean, *SD*, and probability distribution of the output) given the means and standard deviations of the N input signals ($N = 3$). This useful function is available as Supplementary Material to this paper.

The means of the three inputs were $\boldsymbol{r_L}$, $\boldsymbol{r_B}$, $\boldsymbol{r_R}$ with a common standard deviation defined as $\sigma = noise.\max(\boldsymbol{r_L}, \boldsymbol{r_B}, \boldsymbol{r_R})$, where *noise* is a single new parameter of the model. Thus three small inputs were less noisy than three large inputs, but one large input would make all three inputs more noisy. Contrast matches were, as before, obtained when the mean output for a test grating matched the mean output for the comparison grating. Phase matches were given by the probability-weighted vector sum of the three signal phases (ϕ_L, ϕ_B, ϕ_R). Thus,

$$\phi_{match} = \arg\left\{ \sum_{j=L,B,R} p_j \exp(i\phi_j) \right\} \quad \text{(A13)}$$

where p_j is the probability that in a given test condition the jth signal (j = L, B, R) yielded the *max* response. The probability p_j is also returned by the *Matlab* function described above, and we refer to it as the *contribution* made by the jth signal to perception of a given test condition. Obviously, the three contributions always sum to 1.

This model was fitted simultaneously to the contrast and phase matching data of Figure 4 (Huang et al., 2010) in the same way as before, but now there were four free parameters: *s*, *h*, *n*, and *noise*. The fit to the contrast data (Figure A2a) was excellent ($r^2 = 0.995$; RMS error = 0.31 dB). But now the fit to the phase data (Figure A2b) was also very good ($r^2 = 0.963$; RMS error = 3.2° of phase). Phase-matching behavior of the model was sensitive to the level of noise. For example,

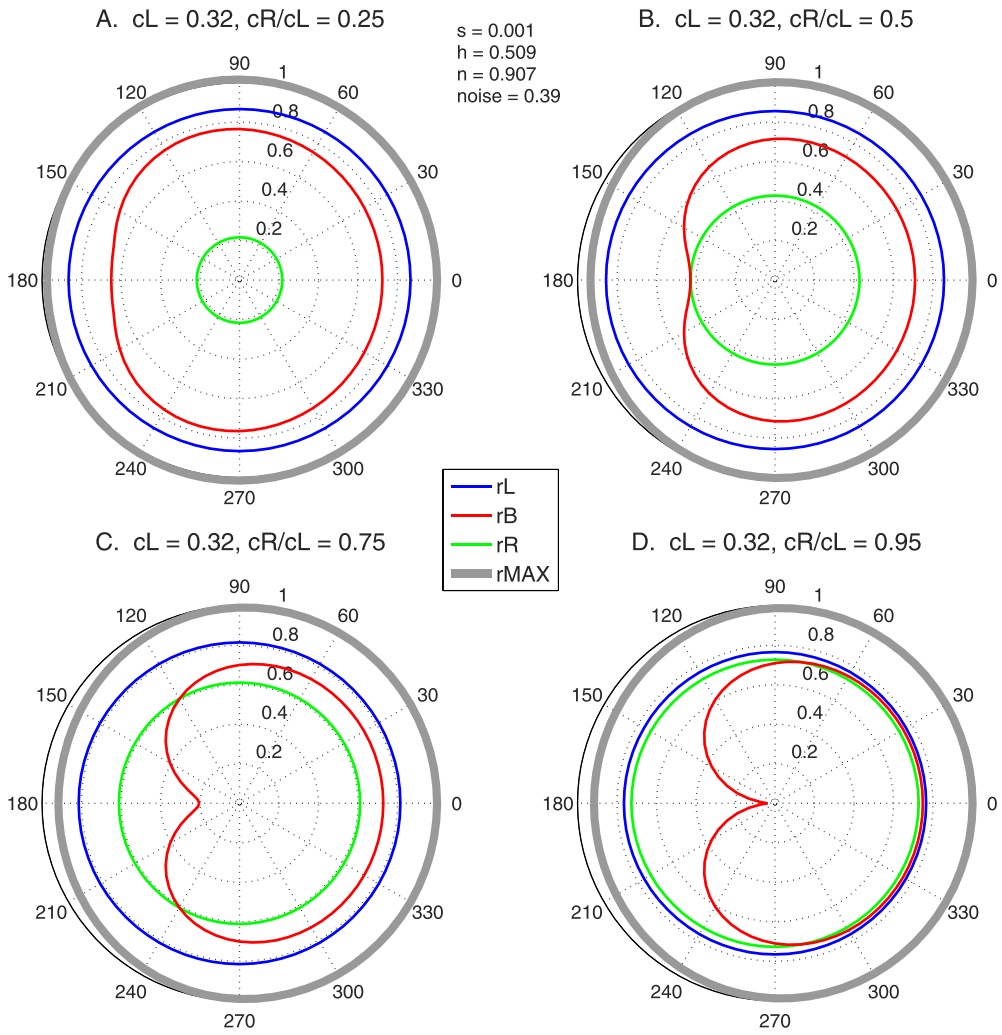

Figure A3. Behavior of the MAX operator with noisy inputs. The four model parameters (inset, top) are from fits to first-order data (Figure A2). Polar plots again illustrate the mean monocular (blue, green) and binocular (red) responses, now expressed relative to the mean MAX operator response obtained at 0 disparity. Gray curve shows how the mean MAX response varied with disparity. Note how, with noise, the mean MAX response is necessarily greater than all the input means. The *contribution* made by each input can be gauged from its relative mean value (red, blue, or green). The Fechner paradox is implicit here ($r_B < r_L$) but largely hidden from the final output because r_B's contribution is correspondingly lower.

doubling or halving the *noise* parameter degraded the quality of phase predictions much more than the contrast matches.

Importantly, the noisy MAX model allows us to estimate the contributions made by monocular and binocular responses (L, B, R) to these dichoptic tasks (Figure A2c). For this plot, contrast C_L is fixed (0.32) while C_R rises from 0 to 0.32. Not surprisingly the R contribution rises with C_R. But a surprising implication is that the binocular (B) contribution is greatest (0.5) for a monocular input and falls to about one third when the input is fully binocular (i.e., left and right eye contrasts are equal), irrespective of disparity. Using the same parameters, Figure A3 gives more detail on these contributions across all phase disparities.

Compared with Figure 4a, the fit to contrast-matching data is improved (Figure A2a) because the predicted Fechner paradox is much reduced. The paradox depends on r_B falling, but when this happens, the MAX operator automatically draws a greater contribution from r_L (Figures A2c and A3), and the paradox is averted. Thus, not only at large disparities but also at intermediate contrast ratios, monocular responses substitute for reduced binocular ones and render perceived contrast more nearly constant than it would otherwise be.

Finally, we asked whether the addition of noise might degrade predictions about contrast matching (Figure 5d) that were already successful without noise. Figure A4 (right) shows that this was not so. The noise actually improved the similarity of model and data.

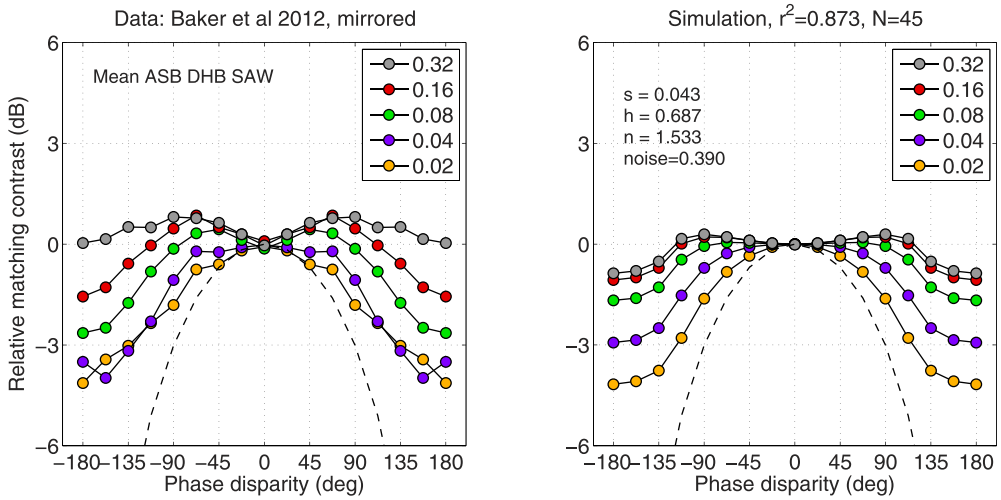

Figure A4. First-order contrast-matching versus phase disparity. Similar to Figure 5a and d but now with noise in the *MAX* operator. Model parameters *s*, *h*, *n* were as in Figure 5; noise level was also fixed from Figure A2. Noise did not degrade but rather improved the similarity between model (right) and data (left).

It may seem counterintuitive that binocular combination contributes no more than one third of the input to binocular contrast perception. One's intuition might be that binocular summation makes responses stronger, and so they should dominate over monocular responses. But that intuition is incorrect. The consensus from the present work and much recent research (cited earlier) is that, except at low contrast, binocular combination takes the form of weighted averaging, not summing, and on this view, binocular responses are rarely much larger than monocular ones (Figures A1 and A3). The special value of binocular vision presumably lies elsewhere in extending the visual field, improving the reliability of signals, and enabling the encoding of stereo disparity and depth. But for nonstereo tasks, we cannot easily dismiss the noisy MAX model because it accounts parsimoniously for such a wide range of contrast- and phase-matching data (Figures A2 and A4).